Inverse and Metric Problems in Random Media

Habilitationsschrift

vorgelegt von

Heinrich Matzinger

Fakultät für Mathematik

der Universität Bielefeld

Juli 2004

Acknowledgements

My first word of thanks goes to Professor Friedrich Götze, my habilitation supervisor. Working in his research group at the University of Bielefeld has been a rewarding experience. I am especially indebted for his suggestion to study the Longest Common Subsequence problem.
I am very grateful to my Ph.D. advisor Harry Kesten who encouraged me to investigate the Scenery Reconstruction problem beyond my Ph.D. thesis. The questions he asked defined to a large extent the direction of research in the area of Scenery Reconstruction.
I am grateful to Servet Martinez for inviting me several times to the Centro de Modelamiento Matematico in Chile. Two of the articles in my habilitation have been written there.
Finally I thank my friends and coauthors:

A.N. Avramidis, University of Montreal, Canada
F. Bonetto, Georgia Tech, U.S.A.
C. Durringer, Univerity of Toulouse, France
A. Hart, Centro de Modelamiento Matematico, Chile
R. Hauser, Oxford University, England
J. Lember, Tartu University, Estonia
M. Löwe, University Muenster, Germany
S. Martinez, Centro de Modelamiento Matematico, Chile.
F. Merkl, University Leiden, the Netherlands
S. Rolles, UCLA, U.S.A.

Working with these colleagues has been very enjoyable.

Contents

Introduction

This habilitation is a collection of twelve papers. Papers [1]-[10] are devoted to the Scenery Reconstruction problem. Papers [11] and [12] investigate the asymptotic properties of the Longest Common Subsequence (LCS) of two random sequences.

paper 1 Franz Merkl, Heinrich Matzinger and Matthias Löwe. *Reconstructing a Multicolor Random Scenery seen along a Random Walk Path with Bounded Jumps.* Electronic Journal of Probability, 9(15):436–507, 2004.

paper 2 Heinrich Matzinger and Silke Rolles. *Reconstructing a random scenery observed with random errors along a random walk path.* Probab. Theory Related Fields, 125(4):539–577, 2003.

paper 3 Heinrich Matzinger. *Reconstructing random scenery seen along a simple random walk path.* accepted in Ann. Appl. Probab., 2004.

paper 4 Matthias Löwe and Heinrich Matzinger. *Scenery Reconstruction in Two Dimensions with Many Colors.* Ann. Appl. Probab., 12(4):1322–1347, 2002.

paper 5 Matthias Löwe and Heinrich Matzinger. *Reconstruction of Sceneries with Correlated Colors.* Stochastic Process. Appl., 105(2):175–210, 2003.

paper 6 Jüri Lember and Heinrich Matzinger. *Information recovery from a randomly mixed up message-text.* Submitted (2004).

paper 7 Jüri Lember and Heinrich Matzinger. *Reconstructing a piece of 2-color scenery.* Preprint (2002).

paper 8 Heinrich Matzinger and Silke Rolles. *Retrieving random media.* Submitted (2003).

paper 9 Heinrich Matzinger and Silke Rolles. *Finding blocks and other patterns in a random coloring of $\mathbb{Z}$.* Submitted (2003).

paper 10 Andrew Hart, Servet Martinez and Heinrich Matzinger. *Markers for error-corrupted observations.* Submitted (2004).

paper 11 Raphael Hauser, Servet Martinez and Heinrich Matzinger. *Large deviation based upper bounds for the LCS-Problem.* Submitted (2003).

paper 12 Jüri Lember, Heinrich Matzinger and Clement Durringer. *Deviation from the mean in sequence comparaison when one sequence is periodic.* Preprint (2004).

The problem of scenery reconstruction and the mathematics behind the LCS-statistics used to check the alignment of genetic sequences, belong to the same class of problems. In both cases, one is concerned with identifying randomly altered information.
To define the scenery reconstruction problem, consider a recurrent random walk $\{S(t)\}_{t\in\mathbb{N}}$ on $\mathbb{Z}$ and a coloring of the integers $\xi : \mathbb{Z} \to \{0, 1, 2, \ldots, C-1\}$. The coloring ξ is called a (C-color) scenery. We are allowed to observe the scenery ξ along the path of S. This means that at time t, we see the color $\chi(t) := \xi(S(t))$. The color record

$$\chi := (\chi(0), \chi(1), \chi(2), \ldots).$$

is known. The scenery reconstruction problem is to recover ξ, given only the observations χ. It varies greatly in difficulty depending on the number of colors in ξ and the distribution of S. I showed in my PhD-thesis [21], that for a simple random walk with holding it is possible to reconstruct a.s. almost any scenery. He takes the scenery i.i.d.. However, Kesten noticed [17], that the method used there completely fails in many other situations. For these other cases, it was necessary to develop a completely new methodology. In this habilitation thesis, these new approaches are presented and worked out in the papers [paper 1]-[paper 10]. The articles on Scenery Reconstruction in this habilitation have been written after my PhD-thesis.

An important method used in the context of sequence alignment problems is the longest common subsequence method. Let I^n denote the integer interval $I^n := [1, n]$. Let $X = (X_i)_{i\in I^n}$ and $Y = (Y_j)_{j\in I^n}$ denote two finite sequences. Denote by L_n the length of the longest common subsequence of X and of Y. (A common subsequence is a sequence, which is a subsequence of X and of Y.) If the longest common subsequence is abnormally long, one decides that X and Y are related to each other. The mathematics of the LCS-problem is badly understood. Let the sequences X and Y be i.i.d. sequences, independent of each other, and with known distribution. $E[L_n]/n$ is known to converge to a (non-random) number γ, as $n \to \infty$. However, in general, γ and the asymptotics of $L_n - E[L_n]$ are unknown. In [paper 11] we present a large deviation based upper bound for γ. The second LCS-paper, investigates the asymptotics of $L_n - E[L_n]$.

0.1 List of Coauthors

Here is a list of the coauthors (and friends) who worked with me on the papers in my habilitation:

- A.N. Avramidis, University of Montreal
- F. Bonetto, Georgia Tech, U.S.A.
- C. Durringer, Univerity of Toulouse, France
- A. Hart, Centro de Modelamiento Matematico, Chile
- R. Hauser, Oxford University, England
- J. Lember, Tartu University, Estonia

- M. Löwe, University Muenster, Germany
- F. Merkl, University Leiden, the Netherlands
- S. Rolles, UCLA, U.S.A.
- M. Servet, Centro de Modelamiento Matematico, Chile

0.2 Scenery Reconstruction

Work on the scenery reconstruction problem started by Kesten's question, whether one can recognize a single defect in a random scenery. Kesten was motived by the T, T^{-1}-problem as well as a conjecture by Keane and den Hollander, and Benjamini about distinguishing sceneries.
The T, T^{-1}-problem is a problem from ergodic theory. The origin of this problem is a famous conjecture by Kolmogorov. He demonstrated that every Bernoulli shift T has a trivial tail-field (let us call the class of all transformations having a trivial tail-field $\mathcal{K}$) and conjectured that also the converse is true. This was proved to be wrong by Ornstein, who presented an example of a transformation which is $\mathcal{K}$ but not Bernoulli. Evidently his transformation was constructed for the particular purpose to resolve Kolmogorov's conjecture. In 1971 Ornstein, Adler, and Weiss came up with a very natural example which is $\mathcal{K}$ but appeared not to be Bernoulli. This was the T, T^{-1}-transformation, and the T, T^{-1}-problem was to show that it was not Bernoulli. In a famous paper Kalikow [14] showed that the T, T^{-1}-transformation is not even loosely Bernoulli and therefore solved the T, T^{-1}-problem. A generalization of this result was recently proved by den Hollander and Steif [7].
The T, T^{-1}-transformation gives rise to a random process of pairs. The first coordinate of these pairs can be regarded as the position of a realization of simple random walk on the integers at time i. The second coordinate tells which color the walker would read at time i, if the integers were colored by an i.i.d. process with black and white in advance.

This is the original setup of the scenery distinction and scenery reconstruction problems. They are related to the T, T^{-1}-problem, but actually we also consider them interesting in their own rights. Moreover, it was pointed out to us that techniques related to those we use for the scenery reconstruction might also be useful in the context of reconstruction of DNA sequences.

Specification of the scenery reconstruction problem. Two sceneries $\xi : \mathbb{Z} \to \{0, 1, \ldots, C-1\}$ and $\tilde{\xi} :\to \{0, 1, \ldots, C-1\}$ are said to be equivalent if one of them is obtained from the other by a translation or reflection. In this case, we write $\xi \approx \tilde{\xi}$. The scenery reconstruction problem can also be formulated as follows:
Does one path realization of the process $\{\chi(t)\}_{t\geq 0}$ uniquely determine ξ? The answer in those general terms is "no". However, under appropriate restrictions, the answer will become "yes". Let us explain these restrictions: First, if ξ and $\tilde{\xi}$ are equivalent, we can in general not distinguish whether the observations come from ξ or from $\tilde{\xi}$. Thus, we can only reconstruct ξ up to equivalence modulo $\approx$. Second, the reconstruction works in the best case only almost surely. Eventually, Lindenstrauss in [12] exhibits sceneries which

can not be reconstructed. However, a lot of "typical" sceneries can be reconstructed up to equivalence. For this we usually take the scenery ξ to be the outcome of a random process and prove that almost every scenery can be reconstructed up to equivalence. Most scenery reconstruction results assume that ξ and S are independent of each other and distributed according to given laws μ and ν. Usually, scenery reconstruction results are formulated as follows:
Given that ξ and S are independent and follow the laws μ, respectively ν, there exists a measurable function

$$\mathcal{A} : \{0,1\}^{\mathbb{N}} \longrightarrow \{0,1\}^{\mathbb{Z}}$$

such that

$$P(\mathcal{A}(\chi) \approx \xi) = 1.$$

In most cases, we prove the above type of theorem by explicitly describing how to reconstruct ξ from χ. The constructed scenery $\bar{\xi}$ is shown to be a.s. equivalent to ξ.
This problem which at first glance seem inaccessible, can also be understood as an "inverse problem". It asks for the unobservable marginal distribution of a combination of a random walk and a random walk we can observe.

Connection to ergodic theory. As mentionned, scenery reconstruction is part of a research area which investigates the ergodic properties of the color record χ. One of the motivations comes from the T, T^{-1}problem; see Kalikow [14]. The ergodic properties of the observations χ were investigated by Heicklen, Hoffman, Rudolph in [10], Keane and den Hollander in [15], den Hollander in [8], and den Hollander and Steif in [7].

A related topic: distinguishing sceneries. A related important problem is to distinguish sceneries: Benjamini, den Hollander, and Keane independently asked whether all non-equivalent sceneries could be distinguished. Here is an outline of this problem: Let η_1 and η_2 be two given sceneries. Assume that either η_1 or η_2 is observed along a random walk path, but we do not know which one. Can we figure out which of the two sceneries was taken? Kesten and Benjamini proved that one can distinguish almost every pair of sceneries even in two dimensions and with only two colors. Before that, Howard had proved in [11], [12], and [13] that any two periodic one dimensional non-equivalent sceneries are distinguishable, and that one can almost surely distinguish single defects in periodic sceneries. The problem of distinguishing two sceneries which differ only in one point is called "detecting a single defect in a scenery". Kesten in [16] proved that one can a.s. recognize a single defect in a random scenery with at least five colors. He asked whether one can distinguish a single defect even if there are only two colors in the scenery.

Solution of the scenery reconstruction problem. Kesten's question was answered by the following result, proved in my Ph.D. thesis [21]: Typical 2-color sceneries can be reconstructed almost surely up to equivalence. The colors are taken i.i.d. uniformly distributed. I showed that almost every 2-color scenery can be almost surely reconstructed. In [22], I proved that almost every 3-color scenery can be almost surely reconstructed. In [17], Kesten noticed that my proofs in [21], [22] heavily relie on the skip-free property of the random walk as well as the one-dimensionality of the scenery. He asked whether the result might still hold in the other cases.

Scenery reconstruction in two dimensions. Together with Löwe, I proved in [paper 4], that one can still reconstruct sceneries in two dimensions, provided there are sufficiently many colors.

Scenery reconstruction using random walk with bounded jumps. Furthermore, in [paper 1], Löwe, Merkl, and I proved the following result: If the random walk can reach every integer with positive probability and is recurrent with at most bounded jumps, and if moreover there are strictly more colors than possible single steps for the random walk, then one can almost surely reconstruct almost every scenery up to equivalence. The case of two colors, $C_0 = 2$, is more difficult than the case investigated in this paper. The major difficulty in the 2-color case is to reconstruct a finite piece of ξ close to the origin. Lember and I [paper 6,7] showed that in the 2-color case one can reconstruct with high probability some information contained in a piece of ξ near by the origin. We then showed how this implies that one can reconstruct a 2-color random scenery seen along the path of a random walk with bounded jumps.

Scenery reconstruction given disturbed input data. In [paper 2], Rolles and I adapted the method proposed by Löwe, Merkl and me to the case where random errors occur in the observed color record. We showed that the scenery can still be reconstructed provided the probability of the errors is small enough. When the observations are seen with random errors, the reconstruction of sceneries is closely related to some coin tossing problems. These have been investigated by Harris and Keane [9] and Levin, Pemantle and Peres [19]. Our paper on reconstruction with errors was motivated by their work and by a question of Peres: He asked for generalizations of the existing results on random coin tossing for the case of many biased coins. With Andrew Hart [paper 6] we solved part of the reconstruction problem for a two color scenery seen along a random walk with bounded jumps. The part which we solved is the finite amount reconstruction which correspond to the content of [paper 10].

Scenery reconstruction using correlated sceneries. Den Hollander asked if scenery reconstruction is possible if the scenery is not i.i.d.. In a joint paper [paper 5], Löwe and I investigated that problem in the case of three colors: we characterized the distributions of sceneries for which the three color method of Matzinger is still applicable. For two colors and for random walks with jumps this still remains an open question.

Related work: Indistinguishable sceneries. As mentioned above, it is in general not possible to reconstruct ξ; one can at most expect a reconstruction *up to equivalence*. As a matter of fact, even this is impossible: By a theorem of Lindenstrauss [20], there exist non-equivalent sceneries that cannot be distinguished and thus cannot be reconstructed.

Time complexity of scenery reconstruction. For sceneries that can be reconstructed, Benjamini asked if it is possible to reconstruct a finite piece in polynomial time (in the length of the piece.) In [paper 8] and [paper 9], Rolles and I answered this question in the affirmative. Den Hollander asked if it is possible to reconstruct a finite piece of scenery in polynomial time. With Rolles we were able to prove that it works

indeed in $n^{2+\varepsilon}$ time where $\varepsilon > 0$ is an arbitrary small constant and n denotes the length of the piece of ξ we reconstruct.

Some ideas on how the reconstruction techniques work. We present here a simplified example in order to give the reader a glimpse on how scenery reconstruction works. Assume for a moment that the scenery ξ is non-random, and instead of being a two color scenery, would be a four color scenery, i.e. $\xi : \mathbb{Z} \longrightarrow \{0, 1, 2, 3\}$. Let us imagine furthermore, that there are two integers x, y such that $\xi(x) = 2$ and $\xi(y) = 3$, but outside x and y the scenery has everywhere color 0 or 1, (i.e. for all $z \in \mathbb{Z}$ with $z \neq x, y$ we have that $\xi(z) \in \{0, 1\}$.) The simple random walk $\{S(k)\}_{k\geq 0}$ can go with each step one unit to the right or one unit to the left. This implies that the shortest possible time for the random walk $\{S(k)\}_{k\geq 0}$ to go from the point x to the point y is $|x-y|$. When the random walk $\{S(k)\}_{k\geq 0}$ goes in shortest possible time from x to y it goes in a straight way, which means that between the time it is at x and until it reaches y it only moves in one direction. During that time, the random walk $\{S(k)\}_{k\geq 0}$ reveals the portion of ξ lying between x and y. If between time t_1 and t_2 the random walk goes in a straight way from x to y, (that is if $|t_1 - t_2| = |x-y|$ and $S(t_1) = x, S(t_2) = y$), then the word $\chi(t_1), \chi(t_1+1), \ldots, \chi(t_2)$ is a copy of the scenery ξ restricted to the interval $[\min\{x, y\}, \max\{x, y\}]$. In this case, the word $\chi(t_1), \chi(t_1+1), \ldots, \chi(t_2)$ is equal to the word $\xi(x), \xi(x+u), \xi(x+2u), \ldots, \xi(y)$, where $u := (y-x)/|y-x|$. Since the random walk $\{S(k)\}_{k\geq 0}$ is recurrent it a.s. goes at least once, in the shortest possible way from the point x to the point y. Because we are given infinitely many observations we can a.s. figure out what the distance between x and y is: the distance between x and y is the shortest time laps that a "3" will ever appear in the observations χ after a "2". When, on the other hand, a "3" appears in the observations χ in shortest possible time after a "2", then between the time we see that "2" and until we see the next "3", we observe a copy of $\xi(x), \xi(x+u), \xi(x+2u), \ldots, \xi(y)$ in the observations χ. This fact allows us to reconstruct the finite piece $\xi(x), \xi(x+u), \xi(x+2u), \ldots, \xi(y)$ of the scenery. Choose any couple of integers t_1, t_2 with $t_2 > t_1$, minimizing $|t_2 - t_1|$ under the condition that $\chi(t_1) = 2$ and $\chi(t_2) = 3$. A.s. then $\chi(t_1), \chi(t_1+1), \ldots, \chi(t_2)$ is equal to $\xi(x), \xi(x+u), \xi(x+2u), \ldots, \xi(y)$.

A numerical example: Let the scenery ξ be such that: $\xi(-2) = 0$, $\xi(-1) = 2$, $\xi(0) = 0$, $\xi(1) = 1$, $\xi(2) = 1$, $\xi(3) = 3$, $\xi(4) = 0$. Assume furthermore that the scenery ξ has a 2 and a 3 nowhere else then in the points -1 and 3. Imagine that χ the observation given to us would start as follows:

$$\chi = (0, 2, 0, 1, 0, 1, 3, 0, 3, 1, 1, 1, 1, 0, 2, 0, 1, 1, 3, \ldots)$$

By looking at all of χ we would see that the shortest time a 3 occurs after a 2 in the observations is 4. In the first observations given above there is however already a 3 only four time units after a 2. The binary word appearing in that place , between the 2 and the 3 is 011. We deduce from this that between the place of the 2 and the 3 the scenery must look like: 011.

In reality the scenery we want to reconstruct is i.i.d. and does not have a 2 and a 3 occuring in only one place . So, instead of the 2 and the 3 in the example above we will use a special pattern in the observations which will tell us when the random walk is back at the same spot. One possibility (although not yet the one we will eventually use) would be to use binary words of the form: 001100 and 110011. It is easy to verify that the only possibility for the word 001100, resp. 110011 to appear in the observations, is when the same word 001100, resp. 110011 occurs in the scenery and the random walk reads it. So, imagine (to give another pedagogical example of a simplified case) the scenery would be such that in a place x there occurs the word 001100, and in the place y there occurs the word 110011 , but these two words occur in no other place in the scenery. These words can then be used as markers: In order to reconstruct the piece of the scenery ξ comprised between x and y we could proceed as follows: take in the observations the

place where the word 110011 occurs in shortest time after the word 001100. In that place in the observations we see a copy of the piece of the scenery ξ comprised between x and y. The reason why the very last simplified example is not realistic is the following: we take the scenery to be the outcome of a random process itself where the $\xi(k)$'s are i.i.d. variables themselves. Thus any word will occur infinitely often in the scenery ξ. The simple markers described above are not good enough for most scenery reconstruction problems. Instead we use sophisticated markers. Depending on the number of colors of ξ and the distribution of S, the techniques use to build efficient markers differ very much.

For example, take the random walk S to be a simple walk with holding. Let the scenery ξ be a two color scenery. Let $x \in \mathbb{Z}$ be a point such that $\xi_z = 0$ and $\xi_{z+1} = 1$. Then S can generate any pattern by just moving back and forth between the points z and $z+1$. Hence, most patterns can be generated in many locations of the scenery ξ. This implies that there exists no markers based on some simple combinatorial ideas. Rather, the markers have to be subtle statistical "localization tests".

Another important problem for scenery reconstruction is to develop statistical tests to find out when the random walk is close to the origin. If we know when the random walk is in the vicinity of the origin, we can recognize the markers close to the place where we want to reconstruct a piece of ξ.

0.3 Longest Common Subsequences

The investigation of the longest common subsequences (LCS) of two finite words is one of the main problems in the theory of pattern matching. The LCS-problem plays a role for DNA- and Protein-alignments, file-comparison, speech-recognition and so forth.
In everything that follows we assume that:
$X_i, i \in \mathbb{N}$ and $Y_i, i \in \mathbb{N}$ are two independent, i.i.d. sequences of uniform random variables. The state space of these variables is a finite alphabet $A = \{0, 1, \ldots, C-1\}$. (In the simplest case, the variables are just i.i.d. Bernoulli variables with parameter $1/2$.) Let X and Y denote the finite random words:

$$X = X_1 \ldots X_n$$

and

$$Y = Y_1 \ldots Y_n.$$

Let L_n designate the length of a longest common subsequence of X and Y. (A common subsequence of X and Y is a sequence which is a subsequence of X and of Y. Let $j \leq n$. Then, X and Y admit a common subsequence of lenght j iff there exist two increasing integer maps $\pi : [0, j-1] \to [0, n]$ and $\sigma : [0, j-1] \to [0, n]$ such that $X_{\pi(k)} = Y_{\sigma(k)}$ for all $k \leq j-1$). The random variable L_n and several of its variants have been studied intensively by probabilists, computer-scientists and mathematical biologists; for applications of LCS-algorithms in biology see Waterman [24].

The most widely used method for the comparison of genetic data is a generalization of the LCS-method. (For an excellent overview of this subject see Waterman-Vingron [26].) In this generalization a maximal score is sought over the set of all possible alignments of

the two sequences, where gaps are penalized with a fixed parameter $\delta > 0$ and mismatches are penalized by a fixed amount $\mu > 0$: consider for example the two words "brot" and "bat". One possible alignment $\mathbb{A}$ of these words is

b	r	o	t
b	a	$-$	t

The score of this alignment is $1 - \mu - \delta + 1 = S(\mathbb{A})$. The matching pairs of letters "b" and "t" are each valued with a weight of 1. The gap $-$ in "bat" after the "a" costs $-\delta$. Furthermore, the mismatch between "r" and "a" is penalized by adding $-\mu$ to the total score. If $M_{\mu,\delta}(X, Y)$ denotes the maximal score amongst all possible alignments of the two words X and Y, and if $M_n(\mu, \delta)$ is the random variable defined by $M_n(\mu, \delta) = M_{\mu,\delta}(X, Y)$, where X and Y are two i.i.d. random sequences of length n, then the LCS-problem is a special case of the investigation of $M_n(\mu, \delta)$, because $L_n = M_n(\infty, 0)$. Generalizing the arguments from the LCS-problem, one can prove that the limit

$$a(\mu, \delta) = \lim_{n\to\infty} \frac{\mathbb{E}[M_n]}{n}$$

exists. Arratia-Waterman [2] showed that there is a phase transition phenomenon defined by critical values of μ and δ. In one phase M_n is of linear order in n, whereas in the other it is logarithmically small in n. Waterman [25] conjectures that the deviation of M_n from its mean behaves like $\sqrt{n}$.

Let us mention a few further details on the history of these problems and the state of knowledge about them: Waterman-Arriata [2] derive a law of large deviation for L_n for fluctuations on scales larger than $\sqrt{n}$. Using first passage percolation methods, Alexander [1] proves that $\mathbb{E}[L_n]/n$ converges at a rate of order $\sqrt{\log n/n}$.

Large deviation based upper bounds for γ. Using a sub-additivity argument, Chvatal-Sankoff [4] prove that the limit

$$\gamma := \lim_{n\to\infty} \mathbb{E}[L_n]/n$$

exists. The exact value of γ remains however unknown. Chvatal-Sankoff [4] derive upper and lower bounds for γ, and similar upper bounds were found by Baeza-Yates, Gavalda, Navarro and Scheihing [3] using an entropy argument.

These bounds have been improved by Deken [6], and subsequently by Dančik-Paterson [5, 23]. For a very large alphabet, M. Kiwi, M. Loebl and J. Matousek [18] where able to determine the value of γ. On the other hand, in [paper 11], R. Hauser, S. Martinez and I developed a method which allows to further improve the upper bound on γ. Our approach can be seen as a generalization of the method of Dančik-Paterson. It is a large deviation and Monte-Carlo simulation based technique.

Fluctuation when one sequence is periodic. Let $(Z_i)_{i\in\mathbb{N}}$ be a (non-random) periodic sequence. Let $LCS_{Z,n}$ designate the length of the longest common subsequence of the two

finite sequences $(X_i)_{i\in I^n}$ and $(Z_i)_{i\in I^n}$, (where as before I^n designates the integer interval $I^n = [1, n]$.) It is known that the limit

$$\lim_{n\to\infty} \frac{LCS_{Z,n}}{n}$$

converges to a constant γ_Z. This constant is not known. Of course γ_Z depends on the periodic sequence $Z_i, \imath \in \mathbb{N}$. In [paper 12], J. Lember, C. Durringer and I , studied the asymptotic deviation from the mean of the random variable $LCS_{Z,n}$. Our main result is that

$$DEV_{Y,n} := LCS_{Z,n} - E\left[LCS_{Z,n}\right]$$

has an order of magnitude of $\sqrt{n}$ when the period of $Z_i, \imath \in \mathbb{N}$ is small in comparison to n.

References

[1] Kenneth S. Alexander. The rate of convergence of the mean length of the longest common subsequence. *Ann. Appl. Probab.*, 4(4):1074–1082, 1994.

[2] Richard Arratia and Michael S. Waterman. A phase transition for the score in matching random sequences allowing deletions. *Ann. Appl. Probab.*, 4(1):200–225, 1994.

[3] R.A. Baeza-Yates, R. Gavaldà, G. Navarro, and R. Scheihing. Bounding the expected length of longest common subsequences and forests. *Theory Comput. Syst.*, 32(4):435–452, 1999.

[4] Václáv Chvatal and David Sankoff. Longest common subsequences of two random sequences. *J. Appl. Probability*, 12:306–315, 1975.

[5] Vlado Dancik and Mike Paterson. Upper bounds for the expected length of a longest common subsequence of two binary sequences. *Random Structures Algorithms*, 6(4):449–458, 1995.

[6] Joseph G. Deken. Some limit results for longest common subsequences. *Discrete Math.*, 26(1):17–31, 1979.

[7] F. den Hollander and J. E. Steif. Mixing properties of the generalized T, T^{-1}-process. *J. Anal. Math.*, 72:165–202, 1997.

[8] W. Th. F. den Hollander. Mixing properties for random walk in random scenery. *Ann. Probab.*, 16(4):1788–1802, 1988.

[9] M. Harris and M. Keane. Random coin tossing. *Probab. Theory Related Fields*, 109(1):27–37, 1997.

[10] D. Heicklen, C. Hoffman, and D. J. Rudolph. Entropy and dyadic equivalence of random walks on a random scenery. *Adv. Math.*, 156(2):157–179, 2000.

[11] C. D. Howard. Detecting defects in periodic scenery by random walks on $\mathbb{Z}$. *Random Structures Algorithms*, 8(1):59–74, 1996.

[12] C. D. Howard. Orthogonality of measures induced by random walks with scenery. *Combin. Probab. Comput.*, 5(3):247–256, 1996.

[13] C. D. Howard. Distinguishing certain random sceneries on $\mathbb{Z}$ via random walks. *Statist. Probab. Lett.*, 34(2):123–132, 1997.

[14] S. A. Kalikow. T, T^{-1} transformation is not loosely Bernoulli. *Ann. of Math. (2)*, 115(2):393–409, 1982.

[15] M. Keane and W. Th. F. den Hollander. Ergodic properties of color records. *Phys. A*, 138(1-2):183–193, 1986.

[16] H. Kesten. Detecting a single defect in a scenery by observing the scenery along a random walk path. In *Itô's stochastic calculus and probability theory*, pages 171–183. Springer, Tokyo, 1996.

[17] H. Kesten. Distinguishing and reconstructing sceneries from observations along random walk paths. In *Microsurveys in discrete probability (Princeton, NJ, 1997)*, pages 75–83. Amer. Math. Soc., Providence, RI, 1998.

[18] Marcos Kiwi, Martin Loebl, and Jiri Matousek. Expected length of the longset common subsequence for large alphabets. *preprint*, 2003.

[19] D. A. Levin, R. Pemantle, and Y. Peres. A phase transition in random coin tossing. *Ann. Probab.*, 29(4):1637–1669, 2001.

[20] E. Lindenstrauss. Indistinguishable sceneries. *Random Structures Algorithms*, 14(1):71–86, 1999.

[21] H. Matzinger. *Reconstructing a 2-color scenery by observing it along a simple random walk path with holding.* PhD thesis, Cornell University, 1999.

[22] H. Matzinger. Reconstructing a three-color scenery by observing it along a simple random walk path. *Random Structures Algorithms*, 15(2):196–207, 1999.

[23] Mike Paterson and Vlado Dancik. Longest common subsequences. In *Mathematical foundations of computer science 1994 (Kosice, 1994)*, volume 841 of *Lecture Notes in Comput. Sci.*, pages 127–142. Springer, Berlin, 1994.

[24] Michael S. Waterman. General methods of sequence comparison. *Bull. Math. Biol.*, 46(4):473–500, 1984.

[25] Michael S. Waterman. Estimating statistical significance of sequence alignments. *Phil. Trans. R. Soc. Lond. B*, 344:383–390, 1994.

[26] M.S. Waterman and M. Vingron. Sequence comparison significance and poisson approximation. *Statistical Science*, 9(3):367–381, 1994.

Chapter 1

Reconstructing a Multicolor Random Scenery seen along a Random Walk Path with Bounded Jumps

Electronic Journal of Probability, 9(15):436–507, 2004.

By Matthias Löwe, Heinrich Matzinger, and Franz Merkl

Kesten [12] noticed that the scenery reconstruction method proposed by Matzinger [17] relies heavily on the skip-free property of the random walk. He asked if one can still reconstruct an i.i.d. scenery seen along the path of a non-skip-free random walk. In this article, we positively answer this question. We prove that if there are enough colors and if the random walk is recurrent with at most bounded jumps, and if it can reach every integer, then one can almost surely reconstruct almost every scenery up to translations and reflections. Our reconstruction method works if there are more colors in the scenery than possible single steps for the random walk. [1]

1.1 Introduction and Result

A (one dimensional) *scenery* is a coloring ξ of the integers $\mathbb{Z}$ with C_0 colors $\{1, \ldots, C_0\}$. Two sceneries ξ, ξ' are called *equivalent*, $\xi \approx \xi'$, if one of them is obtained from the other by a translation or reflection. Let $(S(t))_{t\geq 0}$ be a recurrent random walk on the integers. Observing the scenery ξ along the path of this random walk, one sees the color $\xi(S(t))$ at time t. The *scenery reconstruction problem* is concerned with trying to retrieve the scenery ξ, given only the sequence of observations $\chi := (\xi(S(t)))_{t\geq 0}$. Quite obviously retrieving a scenery can only work up to equivalence. Work on the scenery reconstruction problem started by Kesten's question, whether one can recognize a single defect in a random scenery. Kesten [11] answered this question in the affirmative in the case of four

[1] *MSC 2000 subject classification:* Primary 60K37, Secondary 60G10, 60J75.
Key words: Scenery reconstruction, jumps, stationary processes, random walk, ergodic theory.

colors. He takes the colors to be i.i.d. uniformly distributed. In his Ph.D. thesis [17], see also [18] and [20], Matzinger proved that typical sceneries can be reconstructed: He takes the sceneries as independent uniformly distributed random variables, too. He showed that almost every scenery can be almost surely reconstructed. In [12], Kesten noticed that this proof in [17] heavily relies on the skip-free property of the random walk. He asked whether the result might still hold in the case of a random walk with jumps. This article gives a positive answer to Kesten's question: If the random walk can reach every integer with positive probability and is recurrent with bounded jumps, and if there are strictly more colors than possible single steps for the random walk, then one can almost surely reconstruct almost every scenery up to equivalence.

More formally: Let $\mathcal{C} = \{1, \dots, C_0\}$ denote the set of colors. Let μ be a probability measure over $\mathbb{Z}$ supported over a finite set $\mathcal{M} := \operatorname{supp} \mu \subseteq \mathbb{Z}$. With respect to a probability measure P, let $S = (S(k))_{k \in \mathbb{N}}$ be a random walk starting in the origin and with independent increments having the distribution μ. We assume that $E[S(1)] = 0$; thus S is recurrent. Furthermore we assume that $\operatorname{supp} \mu$ has the greatest common divisor 1, thus S can reach every $z \in \mathbb{Z}$ with positive probability. Let $\xi = (\xi(j))_{j \in \mathbb{Z}}$ be a family of i.i.d. random variables, independent of S, uniformly distributed over $\mathcal{C}$. We prove:

Theorem 1.1.1. *If $|\mathcal{C}| > |\mathcal{M}|$, then there exists a measurable map $\mathcal{A} : \mathcal{C}^{\mathbb{N}} \to \mathcal{C}^{\mathbb{Z}}$ such that*

$$P[\mathcal{A}(\chi) \approx \xi] = 1. \tag{1.1.1}$$

Research on random sceneries started by work by Keane and den Hollander [10], [4]. They thoroughly investigated ergodic properties of a color record seen along a random walk. These questions were motivated among others by the work of Kalikow [9] and den Hollander, Steif [3], in ergodic theory.

As was shown in [20] the two color scenery reconstruction problem for a scenery which is i.i.d. is equivalent to the following problem: let $(R(k))_{k \in \mathbb{Z}}$ and $(S(k))_{k \geq 0}$ be two independent simple random walks on $\mathbb{Z}$ both starting at the origin and living on the same probability space. Does one path realization of the iterated random walk $(R(S(k)))_{k \geq 0}$ uniquely determine the path of $(R(k))_{k \in \mathbb{Z}}$ a.s. up to shift and reflection around the origin? This is a discrete time analogue to a problem solved by Burdzy [2] concerning the path of iterated Brownian motion.

A preform of the scenery reconstruction problem is the problem of distinguishing two given sceneries. It has been investigated by Benjamini and Kesten in [1] and [11]. Howard in a series of articles [8], [7], [6] also contributed to this area; see below. The scenery distinguishing problem is the following: Given two different sceneries ξ, ξ' and observations $(\tilde{\xi}(S(j)))_{j \geq 0}$, where $\tilde{\xi}$ equals either ξ or ξ', the question is: Can we distinguish whether $\tilde{\xi} = \xi$ or $\tilde{\xi} = \xi'$? Benjamini and Kesten [1] showed that one can almost surely distinguish almost all pairs of sceneries ξ, ξ', if they are drawn independently with i.i.d. entries. Their result even holds in the two dimensional case. This result is not beaten by a reconstruction result: the reconstruction method in two dimensions by Löwe and Matzinger [16] holds only when we have many colors. When ξ and ξ' differ in precisely one point, the distinguishing problem was examined by Kesten [11] and Howard [6]. Kesten proved that almost all pairs of those sceneries (ξ, ξ') can be distinguished in the 5-color case.

He assumes the sceneries to be i.i.d. Howard proved that all periodic sceneries can be distinguished.

As mentioned above, it is in general not possible to reconstruct ξ; one can at most expect a reconstruction *up to equivalence*. As a matter of fact, even this is impossible: By a theorem of Lindenstrauss [14], there exist non-equivalent sceneries that cannot be distinguished. Of course, they also cannot get reconstructed.

For sceneries that can be reconstructed Benjamini asked whether the reconstruction works also in polynomial time. This question was positively answered by Matzinger [19] in the case of a two color scenery and a simple random walk with holding. Löwe and Matzinger [15] proved that reconstruction works in many cases even if the scenery is not i.i.d., but has some correlations. For the setting of our article den Hollander asked if the finite bound on the length of the jumps is necessary for scenery reconstruction.

In a way a result by Lenstra and Matzinger complements the present paper. If the random walk might jump more than distance 1 only with very small probability and if the tail of the distribution of the jumps decays sufficiently fast, Lenstra and Matzinger [13] proved that scenery reconstruction is still possible.

Let us explain how this article is organized. In order to avoid getting lost among the many details of the rather complex proof, this article is ordered in a "top-down" approach: In order to show the global structure of the reconstruction procedure in a compact but formal way, we start with a section called "Skeleton". This section collects the main theorems and main definitions of the reconstruction method, using "lower level" procedures as black boxes. In the "Skeleton" section, we only show how these theorems fit together to yield a proof of the reconstruction theorem 1.1.1; all proofs of the "ingredient" theorems are postponed to later sections. Although this approach is more abstract than a "bottom-up" structure would be, we hope that it allows the reader to more quickly see the global structure.

Overview on some steps for the reconstruction procedure The reconstruction starts with an ergodicity argument: It suffices to consider only sceneries which produce a very untypical initial piece of observations; in particular we may condition on a large but finite initial piece of the observations to be constant. We apply a reconstruction procedure, which works only in this untypical situation, again and again to the observations with larger and larger initial pieces dropped, disregarding all instances that do not produce the prescribed "untypical" outcome. Finally we will see even the prescribed "untypical situation" sufficiently frequent to successfully reconstruct the scenery. The "untypical initial piece" serves to identify locations close to the origin at later times again, at least up to a certain time horizon.

The reconstruction procedure consists in a hierarchy of partial reconstruction procedures; these try to reconstruct larger and larger pieces of the scenery around the origin. The hierarchy of partial reconstruction procedures is defined recursively.

To reconstruct a large piece in the $(m+1)$st hierarchical level, we need some information where the random walker is located while producing its color records. This

information is encoded in stopping times, which stop the random walk with high probability sufficiently close to the origin, at least up to a certain time horizon.

The stopping times for the $(m+1)$st hierarchical level are built using the mth level partial reconstruction procedure: Given a reconstructed piece around the origin from the mth level, one starts the whole mth level partial reconstruction procedure again at a later "candidate time". Whenever the piece of scenery obtained in this way has a sufficiently high overlap with the reconstructed piece around the origin, then one has a high chance that the random walk is close to the origin at the "candidate time".

The global structure of this recursive construction is formally described in the "Skeleton" Sections 1.3 and 1.4, and we prove in Sections 1.7 and 1.8 that the stopping times fulfill their specification.

The heart of the reconstruction procedure, i.e. the construction of the partial reconstruction algorithm given the stopping times, is described in Section 1.5 and proven to be correct in Section 1.6. Roughly speaking, to reconstruct a piece of scenery of size 2^n, we collect a "puzzle" of words of size proportional to n, i.e. logarithmically in the size of the piece to be reconstructed. The puzzle contains (with high probability) all correct subwords of the given size in the "true" piece of scenery to be reconstructed, but also some "garbage" words. We play a kind of puzzle game with these pieces: starting with seed words, we reconstruct larger and larger pieces by adjoining more and more pieces of the puzzle that fit to the growing piece.

Although the actual construction is much more complicated than the idea described now, let us describe an (over)simplified version of how to collect pieces in the puzzle: Suppose we have two "characteristic signals" A and B in the scenery, which occur only once in the scenery. Suppose that the distance between A and B is a multiple of the maximal step size $l_\rightarrow$ of the random walk to the right. Then we can almost surely identify the whole "ladder" word read while stepping from A to B with step size $l_\rightarrow$ as follows: Look at all occurrences of A and B in the color record with minimal distance. The words occurring in the color record between those A and B should (a.s.) be always the same in the whole record, and it is the "ladder" word we are looking for. Of course, by ergodicity there are almost surely no (bounded) signals A and B in the scenery that occur only once; this is why the simple idea described here cannot be applied without considerable refinement.

The "pieces of puzzle" obtained are $l_\rightarrow$-spaced pieces; not pieces with spacing 1. This is why our puzzle game leads to reconstructions of modulo classes of the scenery modulo $l_\rightarrow$ only. In order to successfully reconstruct the whole scenery, we need to arrange these modulo classes correctly, using some "neighborship" relation between pieces of the puzzle. Unfortunately, the correct arrangement of modulo classes is a technically intricate step in the reconstruction procedure.

1.2 Some notation

We collect some globally used nonstandard notations and conventions in this section.

Sets, functions, and integers: For functions f and sets D the notation $f\lceil D$ means the restriction of f to the set D. D need not be contained in the domain of f; thus $f\lceil D$ is defined on $D \cap \text{domain}(f)$. If f and g are functions, the notation $f \subseteq g$ means

that f is a restriction of g; this notation is consistent with the set theoretic definition of functions. By convention, $0 \in \mathbb{N}$. The integer part of a real number r is denoted by $\lfloor r \rfloor := \max\{z \in \mathbb{Z} \mid z \le r\}$; similarly $\lceil r \rceil := \min\{z \in \mathbb{Z} \mid z \ge r\}$.

Integer intervals: Unless explicitly stated otherwise, intervals are taken over the integers, e.g. $[a,b] = \{n \in \mathbb{Z} \ : \ a \le n \le b\}$, $]a,b[= \{n \in \mathbb{Z} \ : \ a < n < b\}$. Given a fixed number C_0, we define the *set of colors* $\mathcal{C} := [1, C_0] = \{1, \ldots, C_0\}$, $|\mathcal{C}| = C_0$.

In the rest of this section I will denote an arbitrary subset of $\mathbb{Z}$ unless otherwise specified.

Sceneries and equivalence: By definition, a *scenery* is an element of $\mathcal{C}^{\mathbb{Z}}$. If $I \subseteq \mathbb{Z}$, then the elements of $\mathcal{C}^I$ are called *pieces of scenery*. The *length* $|\zeta|$ of a piece of scenery $\zeta \in \mathcal{C}^I$ is the cardinality $|I|$ of its index set. $\zeta^{\leftrightarrow} := (\zeta_{-i})_{i \in -I}$ denotes the reflection of a piece of scenery $\zeta \in \mathcal{C}^I$ at the origin. Two pieces of scenery $\zeta \in \mathbb{Z}^I$ and $\zeta' \in \mathbb{Z}^{I'}$ are called *strongly equivalent*, $\zeta \equiv \zeta'$, if ζ is obtained by some translation of ζ', i.e. $I' = I + b$ for some $b \in \mathbb{Z}$, and $\zeta = (\zeta'_{i+b})_{i \in I}$. ζ and ζ' are called *equivalent*, $\zeta \approx \zeta'$, if ζ is obtained by some translation or reflection of ζ', i.e. $I' = aI + b$ for some $a \in \{\pm 1\}$, $b \in \mathbb{Z}$, and $\zeta = (\zeta'_{ai+b})_{i \in I}$. If $T : \mathbb{Z} \to \mathbb{Z}$, $T(z) = az + b$, denotes this translation or reflection, then $T[\zeta] := \zeta'$ denotes the transport of ζ' by T; the same notation is used for the domains: $T[I] = I'$. By definition, $\zeta \preccurlyeq \zeta'$ means that $\zeta \approx \zeta' \lceil J$ for some $J \subseteq I'$. If additionally such a subset $J \subseteq I'$ and its reading direction (i.e. either $\zeta \equiv \zeta' \lceil J$ or $\zeta \equiv (\zeta' \lceil J)^{\leftrightarrow}$) is unique, we write $\zeta \preccurlyeq_1 \zeta'$. Similarly $\zeta \sqsubseteq \zeta'$ (in words: "ζ occurs in ζ'") means that $\zeta \equiv \zeta' \lceil J$ for some $J \subseteq I'$.

Words: The elements of $\mathcal{C}^* := \bigcup_{n \in \mathbb{N}} \mathcal{C}^n = \bigcup_{n \in \mathbb{N}} \mathcal{C}^{\{0,\ldots,n-1\}}$ are called *words* (over $\mathcal{C}$). We identify $\mathcal{C}$ with $\mathcal{C}^1$. The concatenation of two words $w_1 \in \mathcal{C}^n$ and $w_2 \in \mathcal{C}^m$ is denoted by $w_1 w_2 \in \mathcal{C}^{n+m}$.

Probability distributions: The law of a random variable X with respect to a probability measure P is denoted by $\mathcal{L}_P(X)$. The n-fold convolution of a probability distribution μ over $\mathbb{R}$ is denoted by μ^{*n}.

Random sceneries and random walks: As mentioned before, let μ be a probability measure over $\mathbb{Z}$ supported over a finite set $\mathcal{M} = \operatorname{supp} \mu \subseteq \mathbb{Z}$. Let $\Omega_2 \subseteq \mathbb{Z}^{\mathbb{N}}$ denote the set of all paths with starting point $S(0) = 0$ and jump sizes $S(t+1) - S(t) \in \mathcal{M}$, $t \in \mathbb{N}$. Let Q_0 denote the law of a random walk $S = (S(k))_{k \in \mathbb{N}}$ with start in $0 \in \mathbb{Z}$ and with independent increments having the distribution μ. Furthermore, let $\xi = (\xi_j)_{j \in \mathbb{Z}}$ be a family of i.i.d. random variables, independent of S, with uniform distribution $\mathcal{L}(\xi_j) = \nu$ over $\mathcal{C}$. We realize (ξ, S) as canonical projections of $\Omega = \mathcal{C}^{\mathbb{Z}} \times \Omega_2$ endowed with its canonical product σ-algebra and the probability measure $P := \nu^{\mathbb{Z}} \otimes Q_0$. (The restriction of the random walk paths not to have forbidden jumps even on null sets is technically convenient.) We assume that $E[S(1) - S(0)] = 0$ ($k \in \mathbb{N}$); thus S is recurrent. Furthermore we assume that $\operatorname{supp} \mu$ has the greatest common divisor 1, thus S eventually reaches every $z \in \mathbb{Z}$ with probability one. For fixed sceneries $\xi \in \mathcal{C}^{\mathbb{Z}}$, we set $P_\xi := \delta_\xi \otimes Q_0$, where δ_ξ denotes the Dirac measure at ξ. Thus P_ξ is the "canonical" version of the conditioned probability

$P[\,\cdot\,|\xi]$. We use the notations P_ξ and $P[\,\cdot\,|\xi]$ as synonyms; i.e. we will never work with a different version of the conditioned measure $P[\,\cdot\,|\xi]$ than P_ξ.

Filtrations: We define the filtration $\mathcal{F} := (\mathcal{F}_n)_{n\in\mathbb{N}}$, $\mathcal{F}_n := \sigma(\xi, (S(k))_{k=0,\ldots,n})$ over Ω. We further introduce the filtration $\mathcal{G} := (\mathcal{G}_n)_{n\in\mathbb{N}}$ over $\mathcal{C}^{\mathbb{N}}$, where $\mathcal{G}_n$ is the σ-algebra generated by the projection map $\mathcal{C}^{\mathbb{N}} \to \mathcal{C}^{[0,n]}$, $\chi \mapsto \chi\lceil[0,n]$.

Observations of the scenery along the random walk and shifts:
Let $\chi = (\chi_n)_{n\in\mathbb{N}} := (\xi_{S(n)})_{n\in\mathbb{N}}$. We sometimes write simply $\chi = \xi \circ S$; this is to be understood in the sense $\chi(\omega) = \xi(\omega) \circ S(\omega)$ for all $\omega \in \Omega$. Let $\mathcal{H} = (\mathcal{H}_n)_{n\in\mathbb{N}}$, $\mathcal{H}_n := \sigma(\chi_k, 0 \le k \le n)$ denote the filtration obtained by observing the scenery along initial pieces of the random walk. We define the shift operations $\theta : \mathcal{C}^{\mathbb{N}} \to \mathcal{C}^{\mathbb{N}}$, $(\chi_n)_{n\in\mathbb{N}} \mapsto (\chi_{n+1})_{n\in\mathbb{N}}$, and $\Theta : \Omega \to \Omega$, $(\xi, S) \mapsto ((\xi_{n+S(1)})_{n\in\mathbb{Z}}, (S(k+1)-S(1))_{k\in\mathbb{N}})$; thus $\chi\circ\Theta = \theta\circ\chi$. Intuitively, Θ spatially shifts both the scenery and the random walk by the location $S(1)$ of the random walk after one step, and it drops the first time step. One observes $\xi \approx \xi\circ\Theta$.

Admissible paths: A piece of path $\pi = (\pi_i)_{i\in I} \in \mathbb{Z}^I$ over an integer interval I is called *admissible* if $\pi_{i+1}-\pi_i \in \mathcal{M}$ for all $\{i, i+1\} \subseteq I$. For finite $I \neq \varnothing$, $\pi_{\min I}$ and $\pi_{\max I}$ are called starting point and end point of π, respectively. We set $\mathsf{TimeShift}(\pi) := (\pi_{i-1})_{i\in I+1}$. By definition, the length $|\pi|$ of the path π is the cardinality $|I|$. For $x, t > 0$ let $\mathsf{AdPaths}(x, t)$ denote the set of all admissible pieces of path $\pi \in [-x, x]^{[0,t[}$.

Ladder intervals and ladder paths: Let $l_\rightarrow := \max\mathcal{M}$, $l_\leftarrow := |\min\mathcal{M}|$; thus $l_\rightarrow$ and $l_\leftarrow$ are the maximal possible jump sizes of S to the right and to the left, respectively. We abbreviate $l := \max\{l_\rightarrow, l_\leftarrow\}$ and $h := l|\mathcal{M}|$. By definition, d-spaced intervals ($d \in \mathbb{N}$) are sets of the form $I \cap (a+d\mathbb{Z})$ with a bounded interval I and a modulo class $a+d\mathbb{Z} \in \mathbb{Z}/d\mathbb{Z}$. $l_\rightarrow$-spaced intervals are also called right ladder intervals. Similarly, $l_\leftarrow$-spaced intervals are called left ladder intervals. By definition, a right ladder path is a piece of path that steps through the points of some right ladder interval in increasing order. Similarly, a left ladder path is a piece of path that steps through the points of some left ladder interval in decreasing order.

Reading words from pieces of sceneries: For $I = \{i_0, \ldots, i_{n-1}\} \subseteq \mathbb{Z}$ with $i_0 < \ldots < i_{n-1}$ and a piece of scenery $\zeta \in \mathcal{C}^I$, we define $\zeta_\rightarrow := (\zeta_{i_k})_{k=0,\ldots,n-1} \in \mathcal{C}^n$ and $\zeta_\leftarrow := (\zeta_{i_{n-1-k}})_{k=0,\ldots,n-1} \in \mathcal{C}^n$; thus $\zeta_\rightarrow$ and $\zeta_\leftarrow$ are the words obtained by reading ζ from the left to the right and from the right to the left, respectively. The right ladder word of a scenery ξ over a right ladder interval I is defined to be $(\xi\lceil I)_\rightarrow$; similarly one defines left ladder words $(\xi\lceil J)_\leftarrow$ over left ladder intervals J.

1.2.1 Conventions concerning constants

Four fixed "sufficiently large" positive integer parameters c_2, c_1, α, and n_0 globally play a role. The meaning of these parameters is explained below at the location of their occurrence; at this point we only describe their mutual dependence:

- $c_2 \in \mathbb{N}$ is chosen first sufficiently large; say $c_2 \ge c_2^{\min}(|\mathcal{C}|, \mu)$.

- Then $c_1 \in 2\mathbb{N}$ is chosen to be even and sufficiently large; say $c_1 \geq c_1^{\min}(c_2, |\mathcal{C}|, \mu)$.
- Then $\alpha \in \mathbb{N}$ is chosen to be sufficiently large; say $\alpha \geq \alpha^{\min}(c_1, |\mathcal{C}|, \mu)$.
- Finally $n_0 \in 2\mathbb{N}$ is chosen to be even and sufficiently large; say $n_0 \geq n_0^{\min}(c_1, \alpha, |\mathcal{C}|, \mu)$.

We do not specify explicitly here how large the allowed lower bounds $c_2^{\min}$, $c_1^{\min}$, $\alpha^{\min}$ and $n_0^{\min}$ actually need to be; but we emphasize that the constructions below will work if they are sufficiently large.

All other positive constants are denoted by "c_i" with a counting index $i > 2$; they keep their meaning globally during the whole article. Unless explicitly stated otherwise, these constants may depend only on the number of colors $|\mathcal{C}|$ and on the jump distribution μ of the random walk; in particular they may depend on the upper bound l of the jump size, but not on n_0.

1.3 Skeleton of the Reconstruction Procedure

Our first "ingredient" theorem reduces the problem of almost surely reconstructing sceneries to the following simpler one: We only need to find an auxiliary reconstruction procedure $\mathcal{A}_{\mathrm{B}}$ which may fail to give an answer, and it may sometimes even give the wrong answer, if only giving the correct answer is more probable than giving a wrong one. Roughly speaking, we apply the auxiliary reconstruction procedure $\mathcal{A}_{\mathrm{B}}$ repeatedly to the observations with initial pieces dropped, taking the answer of the majority as our result; here ergodicity of the observations plays a key role.

Theorem 1.3.1. *If there exists a measurable map $\mathcal{A}_{\mathrm{B}} : \mathcal{C}^{\mathbb{N}} \to \mathcal{C}^{\mathbb{Z}} \cup \{\texttt{fail}\}$ with*

$$P[\mathcal{A}_{\mathrm{B}}(\chi) \neq \texttt{fail}, \mathcal{A}_{\mathrm{B}}(\chi) \approx \xi] > P[\mathcal{A}_{\mathrm{B}}(\chi) \neq \texttt{fail}, \mathcal{A}_{\mathrm{B}}(\chi) \not\approx \xi], \tag{1.3.1}$$

then there exists a measurable map $\mathcal{A} : \mathcal{C}^{\mathbb{N}} \to \mathcal{C}^{\mathbb{Z}}$ such that

$$P[\mathcal{A}(\chi) \approx \xi] = 1. \tag{1.3.2}$$

The auxiliary reconstruction procedure $\mathcal{A}_{\mathrm{B}}$ gives the output "`fail`" if one does not see a long block of 1's in the initial piece of the observations. Thus failure of $\mathcal{A}_{\mathrm{B}}$ is a very frequent event; however, non-failure still occurs with a positive but small probability, and conditioned on this event the most probable answer will be the correct one. Roughly speaking, when we apply $\mathcal{A}_{\mathrm{B}}$ again and again to the observations with initial pieces dropped, we will finally see sufficiently many long blocks of 1's to make the whole procedure work correctly.

The required long block of 1's in the initial piece should have length n_0^{20} for some sufficiently large but fixed even number $n_0 \in 2\mathbb{N}$. The parameter n_0, which parametrizes the size of this required block, is chosen fixed but large enough (see Subsection 1.2.1).

Definition 1.3.1. *With the abbreviation $J_1 = [-2ln_0^{20}, 2ln_0^{20}]$, we define the following events:*

$$E_{\mathrm{B}}(k) := \{\chi_n = 1 \text{ for all } n \leq k\} \quad \text{for } k \in \mathbb{N}, \tag{1.3.3}$$

$$\mathrm{BigBlock} := \left\{ \begin{array}{l} \text{There is an integer interval } J_0 \subseteq J_1 \text{ with } |J_0| \geq n_0^4 \text{ such} \\ \text{that } \xi\lceil J_0 = (1)_{j \in J_0} \text{ is a constant piece of scenery with} \\ \text{value 1.} \end{array} \right\}. \tag{1.3.4}$$

Let P_{B} denote the image of the conditioned law $P[\,\cdot\,|E_{\mathrm{B}}(n_0^{20})]$ with respect to the shift $\Theta^{n_0^{20}}$. Furthermore, we define the conditioned law

$$\tilde{P} := P_{\mathrm{B}}[\,\cdot\,|\mathrm{BigBlock}]. \tag{1.3.5}$$

The event $E_{\mathrm{B}}(n_0^{20})$ occurs when we see a large block of 1's in an initial piece of the observations, while BigBlock occurs when there is a large block of 1's close to the origin in the (unobservable) real scenery ξ. The next lemma tells us that such a large block in the real scenery is very probable whenever we see a large initial block of 1's in the observations:

Lemma 1.3.1. *There exists $c_3 > 0$ such that $P_{\mathrm{B}}[\mathrm{BigBlock}] \geq 1 - e^{-c_3 n_0^{1.2}}$.*

We describe the intuitive meaning of $\tilde{P}$: After having seen a large initial block of 1's in the observations, we drop this initial piece and take the present point as our new starting point. Since then a large block of 1's close to the origin in the unobservable real scenery ξ is typical, it does not change much when we even condition on this (unobservable) event.

Almost all the proofs using the measure $\tilde{P}$ will not explicitly use its definition (1.3.5), but only the following properties of $\tilde{P}$ (and n_0):

Lemma 1.3.2. *The probability measure $\tilde{P}$ fulfills:*

1. *ξ and S are independent with respect to $\tilde{P}$;*
2. *The common distributions of $(S, \xi\lceil(\mathbb{Z}\setminus J_1))$ with respect to $\tilde{P}$ and with respect to P coincide.*
3. *With respect to $\tilde{P}$, the restriction $\xi\lceil J_1$ is independent of $\xi\lceil(\mathbb{Z}\setminus J_1)$.*
4. *$\tilde{P}[\mathrm{BigBlock}] = 1$.*

The next theorem shows that whenever we have a reconstruction procedure $\mathcal{A}'$ that works sufficiently probably with respect to the modified measure $\tilde{P}$, then there exists the auxiliary reconstruction procedure $\mathcal{A}_{\mathrm{B}}$ that we needed above:

Theorem 1.3.2. *Assume that there exists a measurable map $\mathcal{A}' : \mathcal{C}^{\mathbb{N}} \to \mathcal{C}^{\mathbb{Z}}$ with*

$$\tilde{P}[\mathcal{A}'(\chi) \approx \xi] \geq \frac{2}{3}. \tag{1.3.6}$$

Then there exists a measurable map $\mathcal{A}_{\mathrm{B}} : \mathcal{C}^{\mathbb{N}} \to \mathcal{C}^{\mathbb{Z}} \cup \{\mathtt{fail}\}$ such that

$$P[\mathcal{A}_{\mathrm{B}}(\chi) \neq \mathtt{fail}, \mathcal{A}_{\mathrm{B}}(\chi) \approx \xi] > P[\mathcal{A}_{\mathrm{B}}(\chi) \neq \mathtt{fail}, \mathcal{A}_{\mathrm{B}}(\chi) \not\approx \xi]. \tag{1.3.7}$$

The reconstruction function $\mathcal{A}'$ required by the last theorem is built by putting together a hierarchy of partial reconstruction algorithms $\mathcal{A}^m$, $m \geq 1$. The partial reconstruction algorithms $\mathcal{A}^m$ try to reconstruct longer and longer pieces around the origin; the relevant length scale in the m-th hierarchy is given by 2^{n_m}, where n_m is defined as follows:

Definition 1.3.2. *We define recursively a sequence $(n_m)_{m\in\mathbb{N}}$: n_0 was already chosen above; we set*

$$n_{m+1} := 2^{\lfloor\sqrt{n_m}\rfloor}. \tag{1.3.8}$$

The partial reconstruction algorithms may sometimes, but not too frequently, give the wrong answer:

Theorem 1.3.3. *Assume that there exists a sequence $(\mathcal{A}^m)_{m\geq 1}$ of measurable maps $\mathcal{A}^m : \mathcal{C}^{\mathbb{N}} \to \mathcal{C}^{[-5\cdot 2^{n_m}, 5\cdot 2^{n_m}]}$ such that*

$$\tilde{P}\left[\bigcup_{m=1}^{\infty} (E^m)^c\right] \leq \frac{1}{3}, \tag{1.3.9}$$

where

$$E^m := \{\xi\lceil[-2^{n_m}, 2^{n_m}] \preccurlyeq \mathcal{A}^m(\chi) \preccurlyeq \xi\lceil[-9\cdot 2^{n_m}, 9\cdot 2^{n_m}]\}. \tag{1.3.10}$$

Then there exists a measurable map $\mathcal{A}' : \mathcal{C}^{\mathbb{N}} \to \mathcal{C}^{\mathbb{Z}}$ such that the following holds :

$$\tilde{P}[\mathcal{A}'(\chi) \approx \xi] \geq \frac{2}{3}. \tag{1.3.11}$$

Before describing it formally, let us intuitively explain how the hierarchy of partial reconstruction algorithms $\mathcal{A}^m$ is constructed: The $\mathcal{A}^m$ are built recursively in a "zig-zag" way simultaneously with a hierarchy of stopping times:

These stopping times have the task to estimate times when the random walk S is sufficiently close back to the origin, at least up to a certain time horizon. For this estimation, one may use only an initial piece of the color record χ. To find "higher level" stopping times, we try to reconstruct a piece of scenery both at the present candidate location and at the starting point, using a "lower level" partial reconstruction algorithm. If the two obtained pieces of scenery have a high overlap with each other, then there is a good chance that the candidate location and the starting point are close to each other. This is the "zig" part of the "zig-zag" recursion.

The "zag" part of the recursion uses the stopping times as follows to construct a "higher level" partial reconstruction algorithm $\mathcal{A}^m$: Whenever the stopping times indicate that one might be sufficiently close to the origin, one collects "typical signals" which one expects to be characteristic of the local environment in the scenery. The data obtained in this way are then matched together similarly to playing a puzzle game. This procedure is the heart of the whole reconstruction method.

To get the whole construction started, one needs some initial stopping times which indicate that one might be sufficiently close to the origin. A simple way to get such times is the following: Whenever one observes a sufficiently long block of 1's in the color record, then one has a high chance to be close to the origin. (Remember: We conditioned on seeing a long block of 1's at an initial piece of the color record.) This is the reason why we introduce the modified measure $\tilde{P}$, since with respect to $\tilde{P}$ one can be (almost) sure to have a big block of 1's in the scenery close to the origin. However, the such constructed stopping times are not reliable enough to base the first partial reconstruction algorithm on them. Instead, these stopping times are used as ingredients to construct more reliable stopping times.

We treat the "zig" part and the "zag" part of the recursion separately, starting with the formal specification of the "zig" part: Given an abstract partial reconstruction algorithm f, we build stopping times out of it:

The specification of the stopping times depends on a fixed, sufficiently large parameter $\alpha \in \mathbb{N}$. Informally speaking, α influences how many stopping times in each step should be valuable, and what the time horizon for the m-th partial reconstruction algorithm in the

hierarchy should be. The parameter α is chosen fixed but large enough; recall Subsection 1.2.1.

Definition 1.3.3. *Let $m \geq 1$. Let a function $f : \mathcal{C}^{\mathbb{N}} \to \mathcal{C}^{[-5\cdot 2^{n_m}, 5\cdot 2^{n_m}]}$ be given. Assume that $f(\chi)$ depends only on $\chi\lceil[0, 2\cdot 2^{12\alpha n_m}[$. We define the random set*

$$\mathbb{T}_f(\chi) := \left\{ t \in [0, 2^{12\alpha n_{m+1}} - 2\cdot 2^{12\alpha n_m}[\;\middle|\; \exists w \in \mathcal{C}^{2\cdot 2^{n_m}} \colon\; w \preccurlyeq f(\chi) \text{ and } w \preccurlyeq f(\theta^t(\chi)) \right\}. \tag{1.3.12}$$

We define a sequence $T_f = (T_{f,k})_{k\geq 0}$ of $\mathcal{G}$-adapted stopping times with values in $[0, 2^{12\alpha n_{m+1}}]$: Let $t(0) < \ldots < t(|\mathbb{T}_f(\chi)| - 1)$ be the elements of $\mathbb{T}_f(\chi)$ arranged in increasing order. For $k \in \mathbb{N}$, we set

$$T_{f,k}(\chi) := \begin{cases} t(2\cdot 2^{2n_{m+1}}k) + 2\cdot 2^{12\alpha n_m} & \text{if } 2\cdot 2^{2n_{m+1}}k < |\mathbb{T}_f(\chi)|, \\ 2^{12\alpha n_{m+1}} & \text{otherwise.} \end{cases} \tag{1.3.13}$$

Observe that the stopping times $T_f(\chi)$ depend only on $\chi\lceil[0, 2^{12\alpha n_{m+1}}[$.

In the next definition, we introduce events $E^m_{\mathrm{stop},\tau}$; they specify what the stopping times should fulfill: There should be sufficiently many of them, they should be separated by at least $2\cdot 2^{2n_m}$, and they should stop the random walk sufficiently close to the origin. Furthermore, given any abstract partial reconstruction algorithm f, we define an event $E^m_{\mathrm{reconst},f}$; it measures whether f correctly reconstructs a piece of the scenery around the origin.

Definition 1.3.4. *Let $m \in \mathbb{N}$.*

1. *Given a sequence $\tau = (\tau_k)_{k\in\mathbb{N}}$ of $\mathcal{G}$-adapted stopping times, we define*

$$\begin{aligned} &E^m_{\mathrm{stop},\tau} \\ &:= \bigcap_{k=0}^{2^{\alpha n_m}} \left\{ \tau_k(\chi) < 2^{12\alpha n_m},\ |S(\tau_k(\chi))| \leq 2^{n_m}, \tau_j(\chi) + 2\cdot 2^{2n_m} \leq \tau_k(\chi) \text{ for } j < k \right\}. \end{aligned} \tag{1.3.14}$$

2. *We set for $f : \mathcal{C}^{\mathbb{N}} \to \mathcal{C}^{[-5\cdot 2^{n_m}, 5\cdot 2^{n_m}]}$:*

$$E^m_{\mathrm{reconst},f} \;:=\; \{\xi\lceil[-2^{n_m}, 2^{n_m}] \preccurlyeq f(\chi) \preccurlyeq \xi\lceil[-9\cdot 2^{n_m}, 9\cdot 2^{n_m}]\}. \tag{1.3.15}$$

Roughly speaking, the following theorem states: there are stopping times "to get started" which solve their task with high probability:

Theorem 1.3.4. *There exists a sequence of $\mathcal{G}$-adapted stopping times $T^1 = (T^1_k)_{k\in\mathbb{N}}$ with values in $[0, 2^{12\alpha n_1}]$ and a constant $c_4 > 0$, such that*

$$\tilde{P}\left[(E^1_{\mathrm{stop},T^1})^{\mathrm{c}}\right] \leq e^{-c_4 n_0}. \tag{1.3.16}$$

The next theorem states that the "zig"-part of the construction works correctly with high probability. As a premise, the "zig"-part needs the underlying "lower level" partial reconstruction algorithm f to work correctly when f is applied at the beginning. Furthermore, the "zig"-part needs f to have a sufficiently high probability to work correctly on the given scenery ξ whenever it is applied again. Informally speaking, the reason is: In

the "zig"-part we can only reconstruct, if we know where we are. The idea is to start the whole lower-level reconstruction procedure again whenever we want to find out whether we are close to the origin. As mentioned before, if the result has a large overlap with the piece we have have already reconstructed, we can be rather sure that we are close to the origin.

Theorem 1.3.5. *Under the assumptions of Definition 1.3.3, we have that*

$$P\left[(E^{m+1}_{\mathrm{stop},T_f})^{\mathrm{c}} \cap E^m_{\mathrm{reconst},f} \cap \left\{ P\left[E^m_{\mathrm{reconst},f} \,|\, \xi\right] \geq \frac{1}{2} \right\}\right] \leq e^{-n_{m+1}}. \tag{1.3.17}$$

We remark: in the "zig part" (Theorem 1.3.5) we work with the event $E^{m+1}_{\mathrm{stop},T_f}$, while in the "zag part" (Theorem 1.3.6 below) we work with E^m_{stop,T_f}.

Intuitively, in order to successfully recognize locations close to the origin, we need not only the "lower level" reconstruction to work correctly the first time (i.e. $E^m_{\mathrm{reconst},f}$ needs to hold), but also the scenery must be such that whenever one starts the "lower level" reconstruction again, one has a sufficiently high chance to reconstruct again a correct piece; this is why we need the event "$P[E^m_{\mathrm{reconst},f} \,|\, \xi] \geq 1/2$".

Finally the heart of the reconstruction algorithm consists of the "zag"-part: there are partial reconstruction algorithms Alg^{n_m} which take an initial piece of the color record as input data, and abstract "lower level" stopping times τ as "argument procedures". Intuitively, the following theorem states that the algorithms Alg^{n_m} reconstruct correctly with high probability, provided the "argument procedures" τ fulfill their specification $E^m_{\mathrm{stop},\tau}$.

Theorem 1.3.6. *For every $m \in \mathbb{N}$, there is a map*

$$\mathrm{Alg}^{n_m} : [0, 2^{12\alpha n_m}]^{\mathbb{N}} \times \mathcal{C}^{2\cdot 2^{12\alpha n_m}} \to \mathcal{C}^{[-5\cdot 2^{n_m}, 5\cdot 2^{n_m}]} \tag{1.3.18}$$

such that for every vector $\tau = (\tau_k)_{k\in\mathbb{N}}$ of $\mathcal{G}$-adapted stopping times with values in $[0, 2^{12\alpha n_m}]$ one has

$$P\left[(E^m_{\mathrm{reconst},\mathrm{Alg}^{n_m}(\tau,\cdot)})^{\mathrm{c}} \cap E^m_{\mathrm{stop},\tau}\right] \leq c_5 e^{-c_6 n_m} \tag{1.3.19}$$

for some positive constants c_6 and c_5, where $\mathrm{Alg}^{n_m}(\tau,\cdot) : \chi \mapsto \mathrm{Alg}^{n_m}(\tau(\chi), \chi\lceil[0, 2\cdot 2^{12\alpha n_m}[)$.

To motivate the allowed range for the abstract arguments τ in this theorem, recall that $T_{f,k}(\chi)$ in (1.3.13) take their values in $[0, 2^{12\alpha n_{m+1}}]$.

Note that Theorems 1.3.5 and 1.3.6 use the original probability measure P, while Theorem 1.3.4 uses the modified probability measure $\tilde{P}$.

An algorithm Alg^n is defined in the next Section 1.5, but its correctness, i.e. Theorem 1.3.6, is proven in Section 1.6, below. Theorems 1.3.5 and 1.3.4 are proven below in separate Sections 1.7 and 1.8, respectively. Right now we show how to use these three theorems: Provided these three theorems are true, the hypothesis of Theorem 1.3.3 holds, i.e. there exists a sequence of measurable maps $\mathcal{A}^m : \mathcal{C}^{\mathbb{N}} \to \mathcal{C}^{[-5\cdot 2^{n_m}, 5\cdot 2^{n_m}]}$ such that (1.3.9) is valid. We take the maps Alg^{n_m} and the sequences of stopping times T^1, T_f from Theorems 1.3.4, 1.3.5, and 1.3.6 to define recursively maps $\mathcal{A}^m$. Then we prove: the properties guaranteed by Theorems 1.3.4, 1.3.5, and 1.3.6 imply that the sequence of maps $(\mathcal{A}^m)_{m\geq 1}$ satisfies (1.3.9). We are ready to describe the "zig-zag"-recursion formally:

Definition 1.3.5. *We define $\mathcal{A}^m : \mathcal{C}^{\mathbb{N}} \to \mathcal{C}^{[-5\cdot 2^{n_m}, 5\cdot 2^{n_m}]}$ and sequences $T^m = (T^m_k)_{k\in\mathbb{N}}$ of $\mathcal{G}$-adapted stopping times by simultaneous recursion over $m \geq 1$:*

- *T^1 is chosen using Theorem 1.3.4.*
- *$\mathcal{A}^m(\chi) := \mathrm{Alg}^{n_m}(T^m(\chi), \chi\lceil[0, 2 \cdot 2^{12\alpha n_m}[)$, with Alg^{n_m} taken from Theorem 1.3.6.*
- *$T^{m+1} := T_{\mathcal{A}^m}$, with the notation of Definition 1.3.3.*

Recall Definition (1.3.10) of the events E^m. ¿From now on, we use our specific choice for $\mathcal{A}^m$ from Definition 1.3.5. Using (1.3.15), we rewrite (1.3.10) in the form

$$E^m = E^m_{\mathrm{reconst},\mathcal{A}^m}. \tag{1.3.20}$$

Theorem 1.3.7. *For the sequence $(\mathcal{A}^m)_{m\geq 1}$ as defined in Definition 1.3.5 and $(E^m)_{m\in\mathbb{N}}$ as in (1.3.20), the bound (1.3.9) is valid.*

All theorems of this section together yield the proof of our main theorem:

Proof of Theorem 1.1.1. By Theorem 1.3.7, (1.3.9) holds; then (1.3.11) holds by Theorem 1.3.3; moreover (1.3.7) holds by Theorem 1.3.2; finally Theorem 1.3.1 implies the claim (1.1.1) of Theorem 1.1.1. ■

1.4 Proofs concerning the Skeleton Structure

Lemma 1.4.1. *The shift $\Theta : \Omega \to \Omega$, $(\xi, S) \mapsto (\xi(\cdot + S(1)), S(\cdot + 1) - S(1))$ is measure-preserving and ergodic with respect to P.*

Proof. The shift Θ is measure-preserving: Since the distribution of ξ is invariant under (deterministic) translations, and since $S(1)$ is independent of ξ, we get: $\xi(\cdot + S(1))$ has the same distribution as ξ. Furthermore, $(S(t+1) - S(1))_{t\in\mathbb{N}}$ has the same distribution as S. Since ξ, $S(1)$ and $(S(t+1) - S(1))_{t\in\mathbb{N}}$ are independent, $\xi(\cdot + S(1))$ and $(S(t+1) - S(1))_{t\in\mathbb{N}}$ are independent, too. Consequently $\Theta(\xi, S)$ has the same distribution as (ξ, S).

For the ergodicity part, we condition first on deterministic ξ: Recall from Section 1.2 that for fixed $\xi \in \mathbb{C}^{\mathbb{Z}}$, $P_\xi = \delta_\xi \otimes Q_0$ denotes the "canonical" version of the conditioned measure $P[\,\cdot \mid \xi]$. We claim first that the shift Θ is ergodic (but in general not measure-preserving) with respect to P_ξ. To prove this claim, note that the standard shift

$$\hat{\Theta} : \mathbb{Z}^{\mathbb{N}} \to \mathbb{Z}^{\mathbb{N}}, \quad (s(t))_{t\in\mathbb{N}} \mapsto (s(t+1))_{t\in\mathbb{N}} \tag{1.4.1}$$

is ergodic (but not measure-preserving) with respect to the probability measure P_S induced by S. Consider the measurable map

$$f_\xi : \mathbb{Z}^{\mathbb{N}} \to \Omega, \quad f_\xi(s) := (\xi(\cdot + s(0)), s - s(0)); \tag{1.4.2}$$

then $f_\xi \circ \hat{\Theta} = \Theta \circ f_\xi$, and P_ξ is the image measure of P_S with respect to f_ξ. Thus Θ is ergodic with respect to P_ξ, since $\hat{\Theta}$ is ergodic with respect to P_S.

Let $A \subseteq \Omega$ be measurable and shift-invariant: $\Theta^{-1}[A] = A$. According to the above we have for every $\xi \in \mathcal{C}^{\mathbb{Z}}$ that $P_\xi[A] \in \{0, 1\}$. Consider the set

$$M := \left\{\xi \in \mathcal{C}^{\mathbb{Z}} \mid P_\xi[A] = 1\right\}. \tag{1.4.3}$$

We claim that for all $a \in \mathbb{Z}$ holds $\xi \in M$ if and only if $\xi(\cdot + a) \in M$. To prove this claim, let $\xi \in M$ and choose $N \in \mathbb{N}$ such that $P[S(N) = a] > 0$. Then the image measure of $P_\xi[\,\cdot \mid S(N) = a]$ with respect to Θ^N equals $P_{\xi(\cdot+a)}$. Assume $\xi \in M$. Then $1 = P_\xi[A \mid S(N) = a] = P_\xi[\Theta^{-N}A \mid S(N) = a] = P_{\xi(\cdot+a)}[A]$; this shows that $\xi \in M$ implies $\xi(\cdot + a) \in M$. The same argument, applied to the translated scenery $\xi' = \xi(\cdot + a)$ and $a' = -a$, shows that $\xi(\cdot + a) \in M$ implies $\xi \in M$; hence M is a translation invariant set.

By the ergodicity of the translation operator on sceneries, M has measure $P[\xi \in M] = 0$ or $P[\xi \in M] = 1$. If $P[\xi \in M] = 0$, then $P[P_\xi[A] = 0] = 1$; in this case Fubini's theorem yields $P[A] = 0$. Otherwise $P[P_\xi[A] = 1] = 1$; thus $P[A] = 1$, again by Fubini's theorem. ■

Proof of Theorem 1.3.1. The idea of this proof is to apply the reconstruction function $\mathcal{A}_B$ to all the shifted observations $\theta^k(\chi)$ for each $k \in \mathbb{N}$. Every time one does this, one gets either a scenery or the state `fail` as result.

Given $\mathcal{A}_B : \mathcal{C}^{\mathbb{N}} \to \mathcal{C}^{\mathbb{Z}} \cup \{\texttt{fail}\}$ as in the hypothesis of the theorem, we define measurable functions $\mathcal{A}_B^k : \mathcal{C}^{\mathbb{N}} \to \mathcal{C}^{\mathbb{Z}}$, $k \in \mathbb{N}$, as follows:

- If there exists $j \in [0, k[$ such that $\mathcal{A}_B(\theta^j(\chi)) \neq \texttt{fail}$ and

$$\begin{aligned} & \left|\left\{ j' \in [0,k[\;\middle|\; \mathcal{A}_B(\theta^{j'}(\chi)) \neq \texttt{fail},\ \mathcal{A}_B(\theta^{j'}(\chi)) \approx \mathcal{A}_B(\theta^j(\chi)) \right\}\right| \qquad (1.4.4) \\ > & \left|\left\{ j' \in [0,k[\;\middle|\; \mathcal{A}_B(\theta^{j'}(\chi)) \neq \texttt{fail},\ \mathcal{A}_B(\theta^{j'}(\chi)) \not\approx \mathcal{A}_B(\theta^j(\chi)) \right\}\right|, \end{aligned}$$

  then let j_0 be the smallest j with this property, and define $\mathcal{A}_B^k(\chi) := \mathcal{A}_B(\theta^{j_0}(\chi))$.

- Else define $\mathcal{A}_B^k(\chi)$ to be the constant scenery $(1)_{j\in\mathbb{Z}}$.

Finally define the measurable function $\mathcal{A} : \mathcal{C}^{\mathbb{N}} \to \mathcal{C}^{\mathbb{Z}}$ by

$$\mathcal{A}(\chi) := \begin{cases} \lim_{k\to\infty} \mathcal{A}_B^k(\chi) & \text{if this limit exists pointwise,} \\ (1)_{j\in\mathbb{Z}} & \text{otherwise.} \end{cases} \qquad (1.4.5)$$

We check that the such defined function $\mathcal{A}$ fulfills the claim (1.1.1) of the Theorem 1.3.1: Let us give the general idea: by the hypothesis (1.3.1) and an ergodicity argument, "on the long run" the proportion of sceneries $\mathcal{A}_B(\theta^k(\chi))$ (for $k \in \mathbb{N}$) which are equivalent to ξ is strictly bigger than the proportion of sceneries which are not equivalent to ξ. More formally, define for $k \in \mathbb{Z}$ the Bernoulli variables X^k_{sce} and $X^k_{\text{wrong sce}}$: we set X^k_{sce} equal to 1 iff $\mathcal{A}_B(\theta^k(\chi)) \neq \texttt{fail}$ and $\mathcal{A}_B(\theta^k(\chi)) \approx \xi$. Similarly, $X^k_{\text{wrong sce}}$ is equal to 1 iff $\mathcal{A}_B(\theta^k(\chi)) \neq \texttt{fail}$ and $\mathcal{A}_B(\theta^k(\chi)) \not\approx \xi$. Define

$$Y^k_{\text{sce}} := \frac{1}{k}\sum_{i=0}^{k-1} X^i_{\text{sce}} \quad \text{and} \quad Y^k_{\text{wrong sce}} := \frac{1}{k}\sum_{i=0}^{k-1} X^i_{\text{wrong sce}}. \qquad (1.4.6)$$

Observe that if $Y^k_{\text{sce}} > Y^k_{\text{wrong sce}}$ holds, then $\mathcal{A}_B^k(\chi) \approx \xi$. As a consequence of Lemma 1.4.1, the sequences $(X^k_{\text{sce}})_{k\geq 0}$ and $(X^k_{\text{wrong sce}})_{k\geq 0}$ are stationary and ergodic, since they can be viewed as a measurable function of the sequence $k \mapsto \Theta^k(\xi, S)$. Note that $\xi \approx \xi(\cdot + S(k))$. By the ergodic theorem, we have almost surely:

$$Y^k_{\text{sce}} \overset{k\to\infty}{\longrightarrow} P\left[\mathcal{A}_B(\theta^k(\chi)) \neq \texttt{fail}, \mathcal{A}_B(\chi) \approx \xi\right], \qquad (1.4.7)$$

$$Y^k_{\text{wrong sce}} \overset{k\to\infty}{\longrightarrow} P\left[\mathcal{A}_B(\theta^k(\chi)) \neq \texttt{fail}, \mathcal{A}_B(\chi) \not\approx \xi\right]. \qquad (1.4.8)$$

Thus by the assumption (1.3.1) there exists a.s. a (random) k_0 such that for all $k \geq k_0$ we have $Y^k_{\text{sce}} > Y^k_{\text{wrong sce}}$ and hence $\mathcal{A}^k_B(\chi) = \mathcal{A}^{k_0}_B(\chi) \approx \xi$; recall that we chose the smallest possible j_0 in the definition of $\mathcal{A}^k_B$. Thus a.s. $\mathcal{A}(\chi) \approx \xi$. ■

Definition 1.4.1. *For $k, \kappa \in \mathbb{N}$, let $\Xi_{\text{Block}}(k,\kappa)$ be the event of sceneries*

$$\Xi_{\text{Block}}(k,\kappa) := \left\{ \xi \in \mathcal{C}^{\mathbb{Z}} \,\middle|\, \begin{array}{l} \textit{There is an integer interval } J_0 \subseteq [-lk, lk] \textit{ with } |J_0| \geq \kappa \\ \textit{such that } \xi\lceil J_0 = (1)_{j \in J_0} \textit{ is a constant piece of scenery with} \\ \textit{value 1.} \end{array} \right\}. \tag{1.4.9}$$

In this section, only the case $k = n_0^{20}$, $\kappa = n_0^4$ is relevant. In Section 1.8 below, another case is used as well.

Lemma 1.4.2. *If $\kappa \in 2\mathbb{N}$ is large enough, $k \geq \kappa^2$, $k \in \mathbb{N}$, and if ξ is a scenery with $\xi \notin \Xi_{\text{Block}}(k,\kappa)$, then $P\left[E_{\text{B}}(k) \middle| \xi\right] \leq e^{-c_7 k/\kappa^2}$ with some constant $c_7 > 0$. As a consequence, $P\left[E_{\text{B}}(k) \middle| \xi \notin \Xi_{\text{Block}}(k,\kappa)\right] \leq e^{-c_7 k/\kappa^2}$.*

Proof. Let $\xi \in \mathcal{C}^{\mathbb{Z}} \setminus \Xi_{\text{Block}}(k,\kappa)$. The idea of the proof is to split the time interval $[0,k]$ into pieces of size κ^2. Let us examine at first one of these time intervals of size κ^2: A typical length scale for the distance that the random walks travels in this time interval is κ; in particular the probability that it travels farther than distance κ is bounded away from 0, at least if κ is large enough. If the random walk travels that far, it gets close by a point not colored with "1", assuming that $\xi \notin \Xi_{\text{Block}}(k,\kappa)$. (Note that the random walk does not leave the interval $[-lk, lk]$ up to the time horizon k.) But once the random walk is close enough to a point not colored with "1", the probability to hit this point a few steps later is bounded away from 0, too. Thus in every κ^2-sized interval the random walk has a probability bounded away from 0 to see not only the color 1. There are roughly k/κ^2 such intervals in $[0,k]$; thus the probability to see only 1's up to the time horizon k is exponentially small in k/κ^2.

Formally, we proceed as follows: We define stopping times $(\tau_j)_{j=0,\ldots,\lfloor \kappa^{-2} k \rfloor - 1}$:

$$\tau_j := \inf\left\{ t \in [j\kappa^2, (j+1/2)\kappa^2] \mid \xi\lceil [S(t), S(t) + l_{\rightarrow}] \text{ is not constant } 1 \right\}. \tag{1.4.10}$$

In other words, τ_j is the smallest time in the interval $[j\kappa^2, (j+1/2)\kappa^2]$ when there is a point sufficiently close to the right of the location of the random walker which is not colored with "1". If no such time exists, $\tau_j = \infty$. We claim: For some constant $c_8 > 0$ holds

$$P\left[\tau_j < \infty \mid \xi, S\lceil [0, j\kappa^2]\right] \geq c_8. \tag{1.4.11}$$

This means: Uniformly in $\xi \in \mathcal{C}^{\mathbb{Z}} \setminus \Xi_{\text{Block}}(k,\kappa)$ and in the history of the random walker up to time $j\kappa^2$, the chance that the random walk will get sufficiently close to a point not colored by "1" during the next $\kappa^2/2$ steps is bounded from below by a positive constant.

To prove (1.4.11), we observe by the Markov property, $\Delta_j := S((j+1/2)\kappa^2) - S(j\kappa^2)$ has the distribution $\mu^{*\kappa^2/2}$ with respect to the conditioned law $P'_j := P[\cdot \mid \xi, S\lceil [0, j\kappa^2]]$; recall that κ is even. Since $\mu^{*\kappa^2/2}$ has the standard deviation $c_9\kappa$ for some constant c_9, the Central Limit Theorem implies

$$P\left[\Delta_j \geq \kappa\right] \geq c_{10} \tag{1.4.12}$$

for large enough κ. Here c_{10} denotes a fixed positive constant less than $P[X \geq c_9^{-1}]$, and X is a standard normal random variable. Observe that whenever $\xi \in \mathcal{C}^{\mathbb{Z}} \setminus \Xi_{\mathrm{Block}}(k,\kappa)$ and $\Delta_j \geq \kappa$ hold (i.e. in the interval of interest the random walk S moves at least the distance κ to the right), then τ_j is finite (i.e. the random walk passes close to a point which is not colored with 1). This is true since $\xi \in \mathcal{C}^{\mathbb{Z}} \setminus \Xi_{\mathrm{Block}}(k,\kappa)$ implies that $\xi\lceil[S((j+1/2)\kappa^2), S((j+1/2)\kappa^2)+\kappa]$ cannot be a constant piece 1, and since the random walk does not perform jumps to the right larger than $l_\rightarrow$. Since the jump distribution μ is not supported on a strict sublattice $\gamma\mathbb{Z}$ of $\mathbb{Z}$, there is a fixed $L \in \mathbb{N}$ such that $[0, l_\rightarrow] \subseteq \bigcup_{\ell=0}^{L} \mathrm{supp}(\mu^{*\ell})$. If the event $\{\tau_j < \infty\}$ holds, then ξ is not constant 1 on the interval $[S(\tau_j), S(\tau_j)+l_\rightarrow]$. Let A_j denote the event that $\chi\lceil[j\kappa^2,(j+1)\kappa^2]$ is constant 1. Then we have for some constant $1 > c_{11} > 0$ and $\kappa^2/2 \geq L$:

$$\begin{aligned} P'_j[A_j^{\mathrm{c}}] &\geq P'_j[\tau_j < \infty] P'_j[\exists \ell \in [0,L] :\ \chi(\tau_j+\ell) \neq 1 \mid \tau_j < \infty] \\ &\geq c_8 P'_j[\exists \ell \in [0,L] :\ \chi(\tau_j+\ell) \neq 1 \mid \tau_j < \infty] \geq c_{11}. \end{aligned} \tag{1.4.13}$$

Hence we obtain by the Markov property:

$$P\left[E_{\mathrm{B}}(k) \mid \xi\right] \leq P\left[\left.\bigcap_{j=0}^{\lfloor\kappa^{-2}k\rfloor-1} A_j \;\right|\; \xi\right] = E\left[\left.\prod_{j=0}^{\lfloor\kappa^{-2}k\rfloor-1} P'_j[A_j] \;\right|\; \xi\right] \leq (1-c_{11})^{\lfloor\kappa^{-2}k\rfloor}. \tag{1.4.14}$$

This proves Lemma 1.4.2 ■

Proof of Lemma 1.3.1. Let $\sigma^2 := \mathrm{Var}[S(1)]$ be the variance of the single step distribution μ. Consider the integer interval $I := [-2\sigma n_0^{10}, 2\sigma n_0^{10}] \cap \mathbb{Z}$; then

$$P\left[\xi\lceil I = (1)_{j\in I}\right] = |\mathcal{C}|^{-|I|} \geq |\mathcal{C}|^{-4\sigma n_0^{10}-1}. \tag{1.4.15}$$

The first submartingale inequality states $P[\max_{0\leq j\leq t} X_j > \lambda] \leq E[X_t]/\lambda$, for $\lambda > 0$ and nonnegative submartingales X_t. Recall that S^2 is a submartingale, since $x \mapsto x^2$ is convex. Applying the submartingale inequality yields:

$$P\left[\exists j \in [0, n_0^{20}] :\ S(j) \notin I\right] = P\left[\max_{0\leq j\leq n_0^{20}} S(j)^2 > 4\sigma^2 n_0^{20}\right] \leq (4\sigma^2 n_0^{20})^{-1} E\left[S(n_0^{20})^2\right] = \frac{1}{4}. \tag{1.4.16}$$

If $S(j) \in I$ is valid for all $j \in [0, n_0^{20}]$ and if $\xi\lceil I = (1)_{j\in I}$ holds, then $\chi\lceil[0, n_0^{20}] = (1)_{j\in[0,n_0^{20}]}$. Thus (1.4.15) and (1.4.16) and the independence of S and ξ imply

$$P\left[E_{\mathrm{B}}(n_0^{20})\right] \geq \frac{3}{4}|\mathcal{C}|^{-4\sigma n_0^{10}-1}. \tag{1.4.17}$$

Hence we get for some constant $c_3 > 0$, using Lemma 1.4.2 and the abbreviations $\Xi^{\mathrm{c}}_{\mathrm{Block}} = \mathcal{C}^{\mathbb{Z}} \setminus \Xi_{\mathrm{Block}}(n_0^{20}, n_0^4)$ and $E_{\mathrm{B}} = E_{\mathrm{B}}(n_0^{20})$:

$$P\left[\xi \in \Xi^{\mathrm{c}}_{\mathrm{Block}} \middle|\, E_{\mathrm{B}}\right] \leq \frac{P\left[E_{\mathrm{B}} \middle|\, \xi \in \Xi^{\mathrm{c}}_{\mathrm{Block}}\right]}{P\left[E_{\mathrm{B}}\right]} \leq \frac{4}{3}|\mathcal{C}|^{4\sigma n_0^{10}+1} e^{-c_7 n_0^{12}} \leq e^{-c_3 n_0^{12}}. \tag{1.4.18}$$

The shift operation $\Theta^{n_0^{20}}$ applied to (ξ, S) cannot shift the scenery ξ by more than $l n_0^{20}$ steps, and every shift of the interval $[-l n_0^{20}, l n_0^{20}]$ by not more than $l n_0^{20}$ steps is contained in J_1. Thus the shifted event $\Theta^{-n_0^{20}}\mathrm{BigBlock}$ occurs whenever the event $\xi \in \Xi_{\mathrm{Block}}$ holds; thus (1.4.18) implies $P[\Theta^{-n_0^{20}}\mathrm{BigBlock}^{\mathrm{c}} \mid E_{\mathrm{B}}] \leq e^{-c_3 n_0^{12}}$, which is equivalent to the claim of Lemma 1.3.1. ■

Proof of Lemma 1.3.2. We abbreviate $k := n_0^{20}$.

1. We observe first that $\xi \circ \Theta^k$ and the event $E_{\mathrm{B}}(k)$ are both measurable with respect to the σ-field $\sigma(\xi, (S(j))_{j\le k})$, and $S \circ \Theta^k$ is measurable with respect to $\sigma((S(j) - S(k)))_{j>k})$. Since $\sigma(\xi, (S(j))_{j\le k})$ and $\sigma((S(j) - S(k)))_{j>k})$ are independent with respect to P, this implies that $\xi \circ \Theta^k$ and $S \circ \Theta^k$ are independent with respect to $P[\,\cdot\,|E_{\mathrm{B}}(k)]$. Hence ξ and S are independent with respect to the image measure $P_{\mathrm{B}} = (P[\,\cdot\,|E_{\mathrm{B}}(k)]) \circ (\Theta^k)^{-1}$. Since BigBlock $\in \sigma(\xi)$, this implies part 1.

2. By the independence proven in 1., it suffices to show the two claims $\mathcal{L}_{\tilde{P}}(S) = \mathcal{L}_P(S)$ and $\mathcal{L}_{\tilde{P}}(\xi\lceil(\mathbb{Z} \setminus J_1)) = \mathcal{L}_P(\xi\lceil(\mathbb{Z} \setminus J_1))$:

   - With respect to P, $S \circ \Theta^k$ and S both have i.i.d. μ-distributed increments and the starting point 0; thus their distributions coincide. By the above argument, $S \circ \Theta^k$ and $E_{\mathrm{B}}(k)$ are independent with respect to P. Hence the laws of $S \circ \Theta^k$ with respect to P and with respect to $P[\,\cdot\,|E_{\mathrm{B}}(k)]$ coincide with the law $\mathcal{L}_P(S)$ of S with respect to P. Hence $\mathcal{L}_P(S) = \mathcal{L}_{P[\,\cdot\,|E_{\mathrm{B}}(k)]}(S \circ \Theta^k) = \mathcal{L}_{P_B}(S)$. Since ξ and S are independent with respect to P_B, and since BigBlock $\in \sigma(\xi)$, we obtain the first claim $\mathcal{L}_{\tilde{P}}(S) = \mathcal{L}_P(S)$.
   - We condition on fixed values of $\xi\lceil[-lk, lk]$ and $S\lceil[0, k]$:
     We know that $\xi \circ \Theta^k$ is a translation of ξ by $S(k)$ steps, which is not more than kl; this translation maps $[-lk, lk]$ to a subset of J_1. Thus $(\xi \circ \Theta^k)\lceil(\mathbb{Z} \setminus J_1)$ is obtained by translating a ($S(k)$-dependent) subpiece of $\xi\lceil(\mathbb{Z} \setminus [-lk, lk])$. Thus by our i.i.d. and independence assumptions for ξ and S we get: $(\xi \circ \Theta^k)\lceil(\mathbb{Z} \setminus J_1)$ has the distribution $\mathcal{L}_P(\xi\lceil(\mathbb{Z} \setminus J_1)) = \nu^{\mathbb{Z}\setminus J_1}$ with respect to $P[\,\cdot\,|\xi\lceil[-lk, lk], S\lceil[0, k]]$. Furthermore, $(\xi \circ \Theta^k)\lceil(\mathbb{Z} \setminus J_1)$ and $(\xi \circ \Theta^k)\lceil J_1$ are independent with respect to $P[\,\cdot\,|\xi\lceil[-lk, lk], S\lceil[0, k]]$.
     Since $E_{\mathrm{B}}(k)$ depends only on $\xi\lceil[-lk, lk]$ and $S\lceil[0, k]$, this implies

     $$\mathcal{L}_{P_{\mathrm{B}}}(\xi\lceil(\mathbb{Z} \setminus J_1)) = \mathcal{L}_{P[\,\cdot\,|E_{\mathrm{B}}(k)]}((\xi \circ \Theta^k)\lceil(\mathbb{Z} \setminus J_1)) = \nu^{\mathbb{Z}\setminus J_1}, \tag{1.4.19}$$

     and $\xi\lceil(\mathbb{Z} \setminus J_1)$ is independent of $\xi\lceil J_1$ with respect to P_{B}. Since the event BigBlock depends only on $\xi\lceil J_1$, this independence implies

     $$\mathcal{L}_{\tilde{P}}(\xi\lceil(\mathbb{Z} \setminus J_1)) = \mathcal{L}_{P_{\mathrm{B}}}(\xi\lceil(\mathbb{Z} \setminus J_1)) = \nu^{\mathbb{Z}\setminus J_1} = \mathcal{L}_P(\xi\lceil(\mathbb{Z} \setminus J_1)); \tag{1.4.20}$$

     recall our choice of $\tilde{P}$. This proves our second claim.

3. We have just seen: BigBlock $\in \sigma(\xi\lceil J_1)$, and the random pieces $\xi\lceil(\mathbb{Z} \setminus J_1)$ and $\xi\lceil J_1$ are mutually independent with respect to P_{B}. These two facts imply part 3.

4. This is an immediate consequence of the definition $\tilde{P} = P_{\mathrm{B}}[\,\cdot\,|\text{BigBlock}]$.

■

Proof of Theorem 1.3.2. Assume $\mathcal{A}' : \mathcal{C}^{\mathbb{N}} \to \mathcal{C}^{\mathbb{Z}}$ is a measurable map satisfying (1.3.6):

$$P_{\mathrm{B}}[\,\mathcal{A}'(\chi) \approx \xi\,|\text{BigBlock}] \ge \frac{2}{3}. \tag{1.4.21}$$

So,

$$P_{\mathrm{B}}\left[\{\mathcal{A}'(\chi) \approx \xi\} \cap \text{BigBlock}\right] \ge \frac{2}{3} P_{\mathrm{B}}\left[\text{BigBlock}\right]. \tag{1.4.22}$$

By Lemma 1.3.1 it follows, since n_0 is large enough (see Subsection 1.2.1):

$$P_{\mathrm{B}}[\mathcal{A}'(\chi) \approx \xi] \geq \frac{2}{3}\left(1 - e^{-c_3 n_0^{12}}\right) > \frac{1}{2}. \tag{1.4.23}$$

Now, by definition of P_{B},

$$P_{\mathrm{B}}[\mathcal{A}'(\chi) \approx \xi] = P\left[\mathcal{A}'\left(\chi \circ \Theta^{n_0^{20}}\right) \approx \xi \circ \Theta^{n_0^{20}} \;\middle|\; E_{\mathrm{B}}(n_0^{20})\right]. \tag{1.4.24}$$

Obviously $\xi \circ \Theta^{n_0^{20}} \approx \xi$. Thus

$$P\left[\mathcal{A}'\left(\chi \circ \Theta^{n_0^{20}}\right) \approx \xi \;\middle|\; E_{\mathrm{B}}(n_0^{20})\right] > \frac{1}{2}. \tag{1.4.25}$$

We define $\mathcal{A}_{\mathrm{B}} : \mathcal{C}^{\mathbb{N}} \to \mathcal{C}^{\mathbb{Z}} \cup \{\mathtt{fail}\}$:

$$\mathcal{A}_{\mathrm{B}}(\chi) := \begin{cases} \mathcal{A}'(\chi \circ \Theta^{n_0^{20}}) & \text{if } E_{\mathrm{B}}(n_0^{20}) \text{ holds,} \\ \mathtt{fail} & \text{otherwise;} \end{cases} \tag{1.4.26}$$

this is well defined since $E_{\mathrm{B}}(n_0^{20}) \in \sigma(\chi)$. By (1.4.25), the such defined $\mathcal{A}_{\mathrm{B}}$ satisfies (1.3.7). ■

Lemma 1.4.3. *For all events $E \subseteq \Omega$ we have*

$$\tilde{P}(E) \leq |\mathcal{C}|^{4ln_0^{20}+1} P(E). \tag{1.4.27}$$

Proof of Lemma 1.4.3. Define $\Omega' := \mathcal{C}^{\mathbb{Z}\setminus J_1} \times \Omega_2$ and write $\Omega = \mathcal{C}^{\mathbb{Z}} \times \Omega_2 = \mathcal{C}^{J_1} \times \mathcal{C}^{\mathbb{Z}\setminus J_1} \times \Omega_2 = \mathcal{C}^{J_1} \times \Omega'$. Then by definition of the measure P and by Lemma 1.3.2 we have $P = \nu^{J_1} \otimes P_{\Omega'}$ and $\tilde{P} = \tilde{P}_{J_1} \otimes P_{\Omega'}$ where $P_{\Omega'}$ and $\tilde{P}_{J_1}$, respectively, are the marginal distributions of $\tilde{P}$ on Ω' and $\mathcal{C}^{J_1}$, respectively. Thus we have for all measurable cylinder-sets of the form $E = \{e_1\} \times E_2 \subseteq \Omega$, where $e_1 \in \mathcal{C}^{J_1}$ and $E_2 \subseteq \Omega'$:

$$\tilde{P}[E] = \tilde{P}_{J_1}[\{e_1\}] P_{\Omega'}[E_2] \leq |\mathcal{C}|^{4ln_0^{20}+1} \nu^{J_1}[\{e_1\}] P_{\Omega'}[E_2] = |\mathcal{C}|^{4ln_0^{20}+1} P[E] \tag{1.4.28}$$

where the inequality follows because ν is the uniform distribution on $\mathcal{C}$, $|J_1| = 4ln_0^{20} + 1$, and $\tilde{P}_{J_1}$ is bounded from above by one. Since $\mathcal{C}^{J_1}$ is finite, every measurable subset of Ω can be written as a finite disjoint union of sets of the above form $\{e_1\} \times E_2$ with $e_1 \in \mathcal{C}^{J_1}$ and $E_2 \subseteq \Omega'$. This proves the result. ■

Proof of Theorem 1.3.3. For pieces of scenery ψ, φ, we define the piece of scenery $\Phi(\psi, \varphi)$ as follows: If $\psi \preccurlyeq_1 \varphi$, then $\Phi(\psi, \varphi)$ denotes the unique piece of scenery with $\Phi(\psi, \varphi) \approx \varphi$ such that $\psi \subseteq \Phi(\psi, \varphi)$; otherwise we set $\Phi(\psi, \varphi) := \varphi$. Let $\mathcal{A}^m$ as in the hypothesis of the theorem and $\chi \in \mathcal{C}^{\mathbb{N}}$. With the abbreviation $\xi^m := \mathcal{A}^m(\chi)$, we define recursively

$$\zeta^1 := \xi^1, \tag{1.4.29}$$

$$\zeta^{m+1} := \Phi(\zeta^m, \xi^{m+1}), \tag{1.4.30}$$

$$\mathcal{A}'(\chi) := \begin{cases} \lim_{m\to\infty} \zeta^m & \text{if this limit exists pointwise on } \mathbb{Z}, \\ (1)_{j\in\mathbb{Z}} & \text{else.} \end{cases} \tag{1.4.31}$$

(By convention, a sequence $(\zeta^m)_{m\in\mathbb{N}}$ of pieces of sceneries converges pointwise to a scenery ζ if the following holds: $\liminf_{m\to\infty} \mathrm{domain}(\zeta^m) = \mathbb{Z}$, and for every $z \in \mathbb{Z}$ there is $m_z > 0$ such that for all $m \geq m_z$ one has $\zeta^m(z) = \zeta(z)$.) Being a pointwise limit of measurable maps, the map $\mathcal{A}' : \mathcal{C}^{\mathbb{N}} \to \mathcal{C}^{\mathbb{Z}}$ is measurable. For the purpose of the proof, we abbreviate $\underline{\xi}^m := \xi\lceil[-2^{n_m}, 2^{n_m}]$ and $\overline{\xi}^m := \xi\lceil[-9\cdot 2^{n_m}, 9\cdot 2^{n_m}]$ and we define the events

$$E^m_{1\mathrm{fit}} := \left\{ \underline{\xi}^m \preccurlyeq_1 \overline{\xi}^{m+1} \right\}. \tag{1.4.32}$$

We claim:

1. $\liminf_{m\to\infty} E^m_{1\mathrm{fit}}$ holds $\tilde{P}$-a.s.,
2. If the event $\liminf_{m\to\infty} E^m_{1\mathrm{fit}} \cap \bigcap_{m=1}^{\infty} E^m$ occurs, then $\mathcal{A}'(\chi) \approx \xi$.

These two statements together with the hypothesis (1.3.9) imply the claim (1.3.11) of the theorem.

Proof of claim 1.: By Lemma 1.4.3 we may replace "$\tilde{P}$-a.s." in the claim by "P-a.s.". If $I_1 \neq I_2$ are fixed integer intervals with $|I_1| = |I_2|$, then $P[\xi\lceil I_1 \approx \xi\lceil I_2] \leq 2c_{12}e^{-c_{13}|I|}$ holds for some constants $c_{12}, c_{13} > 0$, even if I_1 and I_2 are not disjoint. (See also the similar Lemma 1.6.18, in particular estimate (1.6.66), below. The factor 2 makes the notation consistent with this lemma; recall the binary choice: $\xi\lceil I_1 \approx \xi\lceil I_2$ means $\xi\lceil I_1 \equiv \xi\lceil I_2$ or $\xi\lceil I_1 \equiv (\xi\lceil I_2)^{\leftrightarrow}$.) We apply this for $I_1 = [-2^{n_m}, 2^{n_m}]$ and all integer intervals $I_2 \subseteq [-9\cdot 2^{n_{m+1}}, 9\cdot 2^{n_{m+1}}]$ with $|I_1| = |I_2| = 2\cdot 2^{n_m} + 1$, $I_1 \neq I_2$; there are at most $18\cdot 2^{n_{m+1}}$ choices of I_2. We obtain $P[(E^m_{1\mathrm{fit}})^{\mathrm{c}}] \leq 18\cdot 2^{n_{m+1}} \cdot 2c_{12}e^{-2c_{13}2^{n_m}}$, which is summable over m; recall $n_{m+1} = o(2^{n_m})$ as $m \to \infty$. Hence $(E^m_{1\mathrm{fit}})^{\mathrm{c}}$ occurs P-a.s. only finitely many times by the Borel-Cantelli lemma; this proves claim 1.

Next we prove the second claim: By the assumption made there, there is a (random) M such that the events $E^m_{1\mathrm{fit}}$ and E^m hold for all $m \geq M$. Let $m \geq M$. In the considerations below, we use several times the following rule: For pieces of sceneries $\alpha, \beta, \gamma, \delta$:

$$\text{If } \alpha \preccurlyeq \beta \preccurlyeq \gamma \preccurlyeq \delta \text{ and } \alpha \preccurlyeq_1 \delta, \text{ then } \beta \preccurlyeq_1 \gamma. \tag{1.4.33}$$

In particular, this applies to

$$\underline{\xi}^m \preccurlyeq_1 \overline{\xi}^{m+1} \quad \text{and} \quad \underline{\xi}^m \preccurlyeq \xi^m \preccurlyeq \overline{\xi}^m \preccurlyeq \underline{\xi}^{m+1} \preccurlyeq \xi^{m+1} \preccurlyeq \overline{\xi}^{m+1}; \tag{1.4.34}$$

we obtain $\xi^m \preccurlyeq_1 \xi^{m+1}$. By the definition of ζ^m and Φ, we know $\zeta^m \approx \xi^m$; hence we obtain $\zeta^m \preccurlyeq_1 \xi^{m+1}$. Using the definition of Φ again, we see $\zeta^m \subseteq \Phi(\zeta^m, \xi^{m+1}) = \zeta^{m+1}$. Using (1.4.33), (1.4.34), $\zeta^m \approx \xi^m$, $\zeta^{m+1} \approx \xi^{m+1}$ again, we get

$$\zeta^m \preccurlyeq_1 \overline{\xi}^m \preccurlyeq_1 \zeta^{m+1} \preccurlyeq_1 \overline{\xi}^{m+1} \text{ and } \zeta^m \preccurlyeq_1 \overline{\xi}^{m+1}. \tag{1.4.35}$$

Let $h^m : \mathbb{Z} \to \mathbb{Z}$, $m \geq M$, denote the unique translation or reflection that maps ζ^m onto a subpiece of $\overline{\xi}^m$. As a consequence of $\zeta^m \subseteq \zeta^{m+1}$, $\overline{\xi}^m \subseteq \overline{\xi}^{m+1}$, and (1.4.35) we see that h^m does not depend on m for $m \geq M$. Hence h^m maps $\zeta := \bigcup_{m\geq M} \zeta^m$ to a subpiece of $\xi = \bigcup_{m\geq M} \overline{\xi}^m$; thus $\zeta \preccurlyeq \xi$. In fact the domain of ζ is $\mathbb{Z}$; to see this we observe that $\mathrm{domain}(\zeta)$ contains all $(h^m)^{-1}[\mathrm{domain}(\overline{\xi}^m)] = (h^m)^{-1}[-9\cdot 2^{n_m}, 9\cdot 2^{n_m}]$, which cover all of $\mathbb{Z}$. To summarize, we have shown that $(\zeta^m)_{m\geq M}$ converges pointwise to a scenery $\zeta \approx \xi$; thus $\mathcal{A}'(\chi) = \zeta \approx \xi$ by the definition of $\mathcal{A}'(\chi)$. This finishes the proof of the second claim and also the proof of Theorem 1.3.3. ■

Definition 1.4.2. *We define events of sceneries*

$$\Xi_{\mathrm{I}} := \left\{ \xi \in \mathcal{C}^{\mathbb{Z}} \;\middle|\; P\left[(E^1_{\mathrm{stop},T^1})^{\mathrm{c}} \;\middle|\; \xi \right] \le e^{-c_4 n_0/2} \right\}, \tag{1.4.36}$$

$$\begin{aligned}\Xi_{\mathrm{II}} &:= \bigcap_{m=1}^{\infty} \left\{ \xi \in \mathcal{C}^{\mathbb{Z}} \;\middle|\; \text{If } P\left[E^m \mid \xi\right] \ge \frac{1}{2}, \text{ then } P\left[(E^{m+1}_{\mathrm{stop},T^{m+1}})^{\mathrm{c}} \cap E^m \;\middle|\; \xi \right] \le e^{-n_{m+1}/2} \right\} \\ &= \bigcap_{m=1}^{\infty} \left\{ \xi \in \mathcal{C}^{\mathbb{Z}} \;\middle|\; P\left[(E^{m+1}_{\mathrm{stop},T^{m+1}})^{\mathrm{c}} \cap E^m \cap \left\{ P\left[E^m \mid \xi\right] \ge \frac{1}{2} \right\} \;\middle|\; \xi \right] \le e^{-n_{m+1}/2} \right\}, \end{aligned} \tag{1.4.37}$$

$$\Xi_{\mathrm{III}} := \bigcap_{m=1}^{\infty} \left\{ \xi \in \mathcal{C}^{\mathbb{Z}} \;\middle|\; P\left[(E^m)^{\mathrm{c}} \cap E^m_{\mathrm{stop},T^m} \mid \xi \right] \le c_5^{1/2} e^{-c_6 n_m/2} \right\}, \tag{1.4.38}$$

$$\Xi := \Xi_{\mathrm{I}} \cap \Xi_{\mathrm{II}} \cap \Xi_{\mathrm{III}}, \tag{1.4.39}$$

where c_5 and c_6 are taken from Theorem 1.3.6 and c_4 is taken from Theorem 1.3.4.

Note the similarity between these events and the bounds in (1.3.16), (1.3.17) and (1.3.19). The following lemma provides a link between bounds with and without conditioning on the scenery ξ:

Lemma 1.4.4. *Let A be an event, $r \ge 0$, and Q be a probability measure on Ω such that $Q[A] \le r^2$. Then*

$$Q\left[Q[A|\xi] > r \right] \le r. \tag{1.4.40}$$

Proof of Lemma 1.4.4. This follows directly from

$$r^2 \ge Q[A] \ge \int_{\{Q[A|\xi]>r\}} Q[A|\xi]\, dQ \ge r Q\left[Q[A|\xi] > r \right]. \tag{1.4.41}$$

■

Lemma 1.4.5. *For some constant $c_{14} > 0$ it holds:*

$$\tilde{P}\left[\xi \notin \Xi \right] \le e^{-c_{14} n_0}. \tag{1.4.42}$$

Proof of Lemma 1.4.5. Using the bound (1.3.16), Lemma 1.4.4 for $Q = \tilde{P}$, the fact $\tilde{P}[\cdot \mid \xi] = P[\cdot \mid \xi]$, and the definition (1.4.36) of Ξ_{I}, we obtain for a sufficiently small constant $c_{14} > 0$

$$\tilde{P}\left[\xi \notin \Xi_{\mathrm{I}} \right] \le e^{-c_4 n_0/2} \le \frac{e^{-c_{14} n_0}}{3}; \tag{1.4.43}$$

recall that n_0 was chosen large enough, see Subsection 1.2.1. As a consequence of the bounds (1.3.17) and (1.3.19) we know

$$P\left[(E^{m+1}_{\mathrm{stop},T^{m+1}})^{\mathrm{c}} \cap E^m \cap \left\{ P\left[E^m \mid \xi\right] \ge \frac{1}{2} \right\} \right] \le e^{-n_{m+1}}, \tag{1.4.44}$$

$$P\left[(E^m)^{\mathrm{c}} \cap E^m_{\mathrm{stop},T^m} \right] \le c_5 e^{-c_6 n_m}. \tag{1.4.45}$$

We obtain by the bound (1.4.44), Lemmas 1.4.3 and 1.4.4 with $Q = P$, and (1.4.37):

$$\tilde{P}\left[\xi \notin \Xi_{\mathrm{II}} \right] \le |\mathcal{C}|^{4ln_0^{20}+1} P\left[\xi \notin \Xi_{\mathrm{II}} \right] \le |\mathcal{C}|^{4ln_0^{20}+1} \sum_{m=1}^{\infty} e^{-n_{m+1}/2} \le \frac{e^{-c_{14} n_0}}{3}. \tag{1.4.46}$$

Here we used again that n_0 is large, and that $(n_m)_{m\in\mathbb{N}}$ grows fast; see Definition 1.3.2. The same argument yields, this time using (1.4.45) and (1.4.38):

$$\tilde{P}\left[\xi \notin \Xi_{\mathrm{III}}\right] \;\leq\; |\mathcal{C}|^{4ln_0^{20}+1} P\left[\xi \notin \Xi_{\mathrm{III}}\right] \leq |\mathcal{C}|^{4ln_0^{20}+1} \sum_{m=1}^{\infty} c_5^{1/2} e^{-c_6 n_m/2} \leq \frac{e^{-c_{14}n_0}}{3} \tag{1.4.47}$$

The combination of (1.4.43), (1.4.46), (1.4.47), and (1.4.39) proves Lemma 1.4.5. ■

Lemma 1.4.6. *For all $\xi \in \Xi$ and all $m \in \mathbb{N}$ the following holds for some constants $c_{15} > 0$, $c_{16} > 0$:*

$$P[E^m \mid \xi] \geq 1 - \sum_{k=0}^{m} c_{16} e^{-c_{15} n_k} \geq \frac{1}{2} \tag{1.4.48}$$

and

$$P[E^m \setminus E^{m+1} \mid \xi] \leq c_{16} e^{-c_{15} n_{m+1}}. \tag{1.4.49}$$

Proof of Lemma 1.4.6. Let $\xi \in \Xi$. We prove (1.4.48) and (1.4.49) simultaneously by induction over m: For $m = 1$ we obtain, since $\xi \in \Xi_{\mathrm{I}}$ and $\xi \in \Xi_{\mathrm{III}}$; see (1.4.36) and (1.4.38):

$$\begin{aligned} P[E^1 \mid \xi] \;&\geq\; P[E^1_{\mathrm{stop},T^1} \mid \xi] - P[(E^1)^{\mathrm{c}} \cap E^1_{\mathrm{stop},T^1} \mid \xi] \\ &\geq\; 1 - e^{-c_4 n_0/2} - c_5^{1/2} e^{-c_6 n_1/2} \geq 1 - \sum_{m=0}^{1} c_{16} e^{-c_{15} n_m} \geq \frac{1}{2}; \end{aligned} \tag{1.4.50}$$

for some constants c_{16}, c_{15}; recall that $n_1 \geq n_0$ and n_0 is large enough by Subsection 1.2.1. Thus (1.4.48) holds for $m = 1$. Let $m \geq 1$. Using $\xi \in \Xi_{\mathrm{II}}$, (1.4.37), and our induction hypothesis (1.4.48), we see $P[(E^{m+1}_{\mathrm{stop},T^{m+1}})^{\mathrm{c}} \cap E^m \mid \xi] \leq e^{-n_{m+1}/2}$. Hence we obtain (1.4.49), using $\xi \in \Xi_{\mathrm{III}}$ and (1.4.38):

$$\begin{aligned} P[E^m \setminus E^{m+1} \mid \xi] \;&\leq\; P\left[(E^{m+1})^{\mathrm{c}} \cap E^{m+1}_{\mathrm{stop},T^{m+1}} \,\middle|\, \xi\right] + P\left[(E^{m+1}_{\mathrm{stop},T^{m+1}})^{\mathrm{c}} \cap E^m \,\middle|\, \xi\right] \\ &\leq\; c_5^{1/2} e^{-c_6 n_{m+1}/2} + e^{-n_{m+1}/2} \leq c_{16} e^{-c_{15} n_{m+1}}. \end{aligned} \tag{1.4.51}$$

Consequently we get, using our induction hypothesis (1.4.48) again:

$$P[E^{m+1} \mid \xi] \;\geq\; P[E^m \mid \xi] - P[E^m \setminus E^{m+1} \mid \xi] \geq 1 - \sum_{k=0}^{m+1} c_{16} e^{-c_{15} n_k} \geq \frac{1}{2}; \tag{1.4.52}$$

this completes our induction step. ■

Lemma 1.4.7. *For some constant $c_{17} > 0$ and for all $\xi \in \Xi$,*

$$\tilde{P}\left[\bigcup_{m=1}^{\infty} (E^m)^{\mathrm{c}} \,\middle|\, \xi\right] \leq e^{-c_{17} n_0}. \tag{1.4.53}$$

Proof of Lemma 1.4.7. By Lemma 1.4.6 we have for $\xi \in \Xi$:

$$P\left[\bigcup_{m=1}^{k} (E^m)^{\mathrm{c}} \,\middle|\, \xi\right] \leq P\left[(E^1)^{\mathrm{c}} \mid \xi\right] + \sum_{m=1}^{k} P[E^m \setminus E^{m+1} \mid \xi] \leq \sum_{m=0}^{k} c_{16} e^{-c_{15} n_m} \leq e^{-c_{17} n_0}, \tag{1.4.54}$$

where $c_{17} < c_{15}$ is a small positive constant; recall that n_0 is large. In the limit as $k \to \infty$, this yields the result (1.4.53). ■

Proof of Theorem 1.3.7. Using Lemma 1.4.5 we have

$$\begin{aligned} \tilde{P}\left[\bigcup_{m=1}^{\infty}(E^m)^{\mathrm{c}}\right] &\le \tilde{P}\left[\xi \notin \Xi\right] + \tilde{P}\left[\{\xi \in \Xi\} \cap \bigcup_{m=1}^{\infty}(E^m)^{\mathrm{c}}\right] \qquad (1.4.55)\\ &\le e^{-c_{14}n_0} + \int_{\{\xi\in\Xi\}} \tilde{P}\left[\bigcup_{m=1}^{\infty}(E^m)^{\mathrm{c}} \,\middle|\, \xi\right] \mathrm{d}\tilde{P} \\ &\le e^{-c_{14}n_0} + \sup_{\xi\in\Xi} \tilde{P}\left[\bigcup_{m=1}^{\infty}(E^m)^{\mathrm{c}} \,\middle|\, \xi\right]. \end{aligned}$$

We bound the argument of the last supremum, using Lemma 1.4.7:

$$\tilde{P}\left[\bigcup_{m=1}^{\infty}(E^m)^{\mathrm{c}} \,\middle|\, \xi\right] = P\left[\bigcup_{m=1}^{\infty}(E^m)^{\mathrm{c}} \,\middle|\, \xi\right] \le e^{-c_{17}n_0}. \qquad (1.4.56)$$

The combination of (1.4.55) and (1.4.56) yields, since n_0 is large (by Subsection 1.2.1):

$$\tilde{P}\left[\bigcup_{m=1}^{\infty}(E^m)^{\mathrm{c}}\right] \le e^{-c_{14}n_0} + e^{-c_{17}n_0} \le \frac{1}{3}. \qquad (1.4.57)$$

■

1.5 Heart of the Reconstruction Procedure: Definition of the Algorithm Alg^n

This section contains the heart of the reconstruction procedure: for every $n \in \mathbb{N}$, we define an algorithm Alg^n; it is designed to reconstruct long pieces of scenery with high probability. In Section 1.6 below we show that it fulfills the formal specification given in Theorem 1.3.6.

Informally speaking, the observation χ allows us to collect many pieces of "puzzle words". These puzzle words are chosen to have size $c_1 n$ with a fixed parameter c_1; recall subsection 1.2.1. To obtain them, we collect triples of words (w_1, w_2, w_3) which occur in sequence in the observations χ soon after a stopping time $\tau(k)$; an initial piece of χ is represented below by a formal argument η. We put those words w_2 into our puzzle which are already uniquely determined by w_1 and w_3. This means that w_1 and w_3 should be be very "characteristic signals"; if w_1 and w_3 could be read at very different locations in the scenery close to a stopping time, then it is unprobable that they will enclose always the same word w_2. Frequently, w_2 turns out to be a ladder word: Whenever one reads a w_2 in the context $w_1 w_2 w_3$ along a non-ladder path sufficiently close to the origin, one reads with high probability a different word w_2' in the context $w_1 w_2' w_3$, too, along a different path with the same starting point and the same end point; but then w_2 is not collected as a puzzle word.

Here is the formal construction: We take input data $\tau \in [0, 2^{12\alpha n}]^{\mathbb{N}}$ and $\eta \in \mathcal{C}^{2\cdot 2^{12\alpha n}}$. A side remark: although for formal reasons there are infinitely many $\tau(k)$ given in the input data, the construction below actually uses only the first $2^{\alpha n}$ of them.

Definition 1.5.1. *We define for $m \geq 0$ the random sets:*

$$\mathsf{PrePuzzle}^n(\tau,\eta) := \tag{1.5.1}$$
$$\left\{(w_1,w_2,w_3) \in (\mathcal{C}^{c_1 n})^3 \mid \exists k \in [0, 2^{\alpha n}[\colon\ w_1 w_2 w_3 \sqsubseteq \eta\lceil[\tau(k), \tau(k) + 2^{2n}]\right\},$$

$$\mathsf{Puzzle}^n_{\mathrm{I}}(\tau,\eta) := \tag{1.5.2}$$
$$\{(w_1,w_2,w_3) \in \mathsf{PrePuzzle}^n(\tau,\eta) \ \mid \forall (w_1,w_2',w_3) \in \mathsf{PrePuzzle}^n(\tau,\eta)\colon w_2' = w_2\},$$

$$\mathsf{Puzzle}^n_{\mathrm{II}}(\tau,\eta) := \tag{1.5.3}$$
$$\{w_2 \ \in \mathcal{C}^{c_1 n} \mid \exists w_1, w_3 \in \mathcal{C}^{c_1 n}\colon (w_1,w_2,w_3) \in \mathsf{Puzzle}^n_{\mathrm{I}}(\tau,\eta)\}.$$

Let us explain the idea behind the following constructions: Although many of the words w_2 in "$\mathsf{Puzzle}_{\mathrm{II}}$" turn out to be ladder words of a central piece in the true scenery ξ, some of them are not: There are "garbage words" in the puzzle. We play a "puzzle-game" with the words in "$\mathsf{Puzzle}_{\mathrm{II}}$": We try to fit larger and larger pieces together. In order to distinguish "real" pieces from "garbage" pieces, we need some "seed words" which are guaranteed (with high probability) not to be garbage words; every piece that fits to a piece containing a seed word has a high chance not to be garbage, too. This is what the set $\mathsf{Seed}_{\mathrm{II}}$ defined below is good for. We identify "seed" words as "puzzle" words that occur in the observations almost immediately after a stopping time $\tau(k)$, when we expect the random walk to be close to the origin.

Recall the abbreviation $h = l|\mathcal{M}|$. Formally, we proceed as follows:

Definition 1.5.2.

$$\mathsf{Seed}^n_{\mathrm{I}}(\tau,\eta) := \tag{1.5.4}$$
$$\left\{(w_1,w_2,w_3) \in \mathsf{Puzzle}^n_{\mathrm{I}}(\tau,\eta) \ \middle|\ \begin{array}{l}\exists k \in [0,2^{\alpha n}[\ \exists j \in [0, 7c_1 nl]: \\ w_1 w_2 w_3 \equiv \eta\lceil(\tau(k) + j + [0, 3c_1 n[)\end{array}\right\},$$

$$\mathsf{Seed}^n_{\mathrm{II}}(\tau,\eta) := \{w_2 \in \mathcal{C}^{c_1 n} \mid (w_1,w_2,w_3) \in \mathsf{Seed}^n_{\mathrm{I}}(\tau,\eta)\}, \tag{1.5.5}$$

$$\mathsf{Seed}^n_{\mathrm{III}}(\tau,\eta) := \tag{1.5.6}$$
$$\left\{u \in \mathsf{Seed}^n_{\mathrm{II}}(\tau,\eta) \ \middle|\ \begin{array}{l}\exists v \in \mathsf{Seed}^n_{\mathrm{II}}(\tau,\eta): \\ (u\lceil([0, c_2 nl_{\leftarrow}] \cap l_{\leftarrow}\mathbb{Z}))_{\rightarrow} = (v\lceil([0, c_2 nl_{\rightarrow}] \cap l_{\rightarrow}\mathbb{Z}))_{\leftarrow}\end{array}\right\},$$

$$\mathsf{Neighbors}^n(\tau,\eta) := \tag{1.5.7}$$
$$\left\{(w_1,w_2) \in (\mathcal{C}^{c_1 n})^2 \mid \exists k \in [0,2^{\alpha n}[, w \in \mathcal{C}^{h-1}:\ w_1 w w_2 \sqsubseteq \eta\lceil[\tau(k), \tau(k) + 2^{2n}]\right\}.$$

Let us explain what "$\mathsf{Seed}_{\mathrm{III}}$" is intended for: We need to identify the orientation of the pieces (whether they are to be read "forward" or "backward"). This task consists of two problems: The identification of the relative orientation of two pieces with respect to each other, and the identification of the absolute orientation with respect to the "true" scenery ξ. Of course, we have no chance to identify the absolute orientation if the random walk is symmetric; we even bother about identifying the absolute orientation only in the very unsymmetric case $l_{\rightarrow} \neq l_{\leftarrow}$. The set $\mathsf{Seed}_{\mathrm{III}}$ helps us to identify the absolute orientation in this case: Suppose we read every $l_{\rightarrow}$-th letter in a word from the left to the right, and every $l_{\leftarrow}$-th letter in the same word from the right to the left; then every $l_{\rightarrow} l_{\leftarrow}$-th letter appears in both words, when at least one letter is read both times. This turns out

to be characteristic enough to identify the reading directions "left" and "right" in the case $l_\rightarrow \neq l_\leftarrow$. The fixed parameter c_2 influences the length of the sample pieces in this procedure.

The relation "Neighbors" serves as an estimation for the geometric neighborship relation between ladder words: ladder words that occur closely together in the observation χ are expected to occur on geometrically neighboring intervals in the "true" scenery ξ. The next definition defines a "true" geometric neighborship relation $\triangleright_n$. We try to reconstruct the corresponding "true" neighborship relation for ladder words in a piece of ξ using only the "estimated" neighborship relation "Neighbors".

Recall that μ^{*k} denotes the k-fold convolution of μ; in particular

$$\operatorname{supp} \mu^{*k} := \left\{ \sum_{i=1}^{k} s_i \;\middle|\; \forall i:\ s_i \in \operatorname{supp} \mu \right\}. \tag{1.5.8}$$

Definition 1.5.3. *Let I, J be right ladder intervals. By definition, $I \triangleright_n J$ means $|I| = |J| = c_1 n$ and $\min J - \max I \in \operatorname{supp} \mu^{*h}$. Similarly for I', J' being left ladder intervals, $I' \triangleleft_n J'$ means $|I'| = |J'| = c_1 n$ and $\max J' - \min I' \in \operatorname{supp} \mu^{*h}$.*

The next definition is the heart of our method: We describe how to obtain reconstructed pieces of sceneries. All pieces of scenery $w \in \mathcal{C}^{[-5\cdot 2^n, 5\cdot 2^n]}$ are tested as candidates in a sequence of "Filters": Reconstructed ladder words should be in "$\mathsf{Puzzle}_{\mathrm{II}}$", the "estimated" and the "reconstructed" neighborship relation should be consistent with each other, the reconstructed pieces should contain "$\mathsf{Seed}_{\mathrm{III}}$" words, and no piece of the puzzle should be used twice.

Only candidate pieces that pass all Filters are considered as a solution of the partial reconstruction problem.

Definition 1.5.4. *Let $\mathsf{Filter}_i^n(\tau, \eta)$, $i = 1, \ldots, 5$, denote the set of all $w \in \mathcal{C}^{[-5\cdot 2^n, 5\cdot 2^n]}$ which fulfill the following condition $1., \ldots, 5.$, respectively:*

1. *For every right ladder interval $I \subseteq [-5 \cdot 2^n, 5 \cdot 2^n]$, $|I| = c_1 n$, one has $(w\lceil I)_\rightarrow \in \mathsf{Puzzle}_{\mathrm{II}}^n(\tau)$.*

2. *For all right ladder intervals $I, J \subseteq [-5 \cdot 2^n, 5 \cdot 2^n]$:*
   if $I \triangleright_n J$, then $((w\lceil I)_\rightarrow, (w\lceil J)_\rightarrow) \in \mathsf{Neighbors}^n(\tau, \eta)$.

3. *For all right ladder intervals $I, J \subseteq [-5 \cdot 2^n, 5 \cdot 2^n]$, $|I| = |J| = c_1 n$:*
   if $((w\lceil I)_\rightarrow, (w\lceil J)_\rightarrow) \in \mathsf{Neighbors}^n(\tau, \eta)$, then there is $q \in \mathbb{N}$ such that $I \triangleright_n J + q l_\rightarrow$.

4. *For every right modulo class $Z \in \mathbb{Z}/l_\rightarrow\mathbb{Z}$ there exists a right ladder interval $I \subseteq Z \cap [-2 \cdot 2^n, 2 \cdot 2^n]$ such that $(w\lceil I)_\rightarrow \in \mathsf{Seed}_{\mathrm{III}}^n(\tau, \eta)$.*

5. *For all right ladder intervals $I, J \subseteq [-5 \cdot 2^n, 5 \cdot 2^n]$, $|I| = |J| = c_1 n$:*
   if $(w\lceil I)_\rightarrow = (w\lceil J)_\rightarrow$, then $I = J$.

We set

$$\mathsf{SolutionPieces}^n(\tau, \eta) := \bigcap_{i=1}^{5} \mathsf{Filter}_i^n(\tau, \eta). \tag{1.5.9}$$

The output of the algorithm Alg^n could be any of these pieces $w \in \mathsf{SolutionPieces}^n(\tau,\eta)$; we choose one of them, if it exists.

Definition 1.5.5. *We define* $\mathrm{Alg}^n(\tau,\eta)$ *as follows:*

- *If* $\mathsf{SolutionPieces}^n(\tau,\eta)$ *is nonempty, then we define* $\mathrm{Alg}^n(\tau,\eta)$ *to be its lexicographically smallest element.*
- *Otherwise* $\mathrm{Alg}^n(\tau,\eta)$ *is defined to be the constant scenery* $(1)_{j\in[-5\cdot 2^n,5\cdot 2^n]}$.

We could have equally well taken any element of $\mathsf{SolutionPieces}^n(\tau,\eta)$ in Definition 1.5.5; we choose the lexicographically smallest one just for definiteness.

1.6 Playing Puzzle: Correctness of the Algorithm Alg^n

In this section we prove Theorem 1.3.6 by showing that the Algorithm Alg^n defined in Definition 1.5.5 fulfills the specification described by this theorem: Let $n = n_m$, $m \in \mathbb{N}$. A remark concerning notation: Events defined in this section are labeled with an upper index n, not m, since the "hierarchy level" m plays no role here, in contrast to the "Skeleton" section. Only events that also occur in the "Skeleton" section keep their old index m. Hopefully, this should not cause any confusion.

Let $\tau = (\tau_k)_{k\in\mathbb{N}}$ denote a fixed vector of $\mathcal{G}$-adapted stopping times with values in $[0, 2^{12\alpha n}]$. We abbreviate $\mathsf{Input} := (\tau(\chi), \chi\lceil[0, 2\cdot 2^{12\alpha n}[)$.

Definition 1.6.1. *We define the following events:*

$$E^n_{\text{xi does it}} := \{\xi\lceil[-5\cdot 2^n, 5\cdot 2^n] \in \mathsf{SolutionPieces}^n(\mathsf{Input})\}\,, \tag{1.6.1}$$

$$E^n_{\text{all pieces ok}} := \left\{ \begin{array}{l} \forall w \in \mathsf{SolutionPieces}^n(\mathsf{Input}) : \\ \xi\lceil[-2^n, 2^n] \preccurlyeq w \preccurlyeq \xi\lceil[-9\cdot 2^n, 9\cdot 2^n] \end{array} \right\}. \tag{1.6.2}$$

Lemma 1.6.1.

$$E^n_{\text{xi does it}} \cap E^n_{\text{all pieces ok}} \subseteq E^m_{\text{reconst},\mathrm{Alg}^n(\tau,\cdot)} \tag{1.6.3}$$

Proof of Lemma 1.6.1. When the event $E^n_{\text{xi does it}}$ holds, then the set $\mathsf{SolutionPieces}^n(\mathsf{Input})$ is not empty. Thus $\mathrm{Alg}^n(\mathsf{Input})$ is the lexicographically smallest element of $\mathsf{SolutionPieces}^n(\mathsf{Input})$. When the event $E^n_{\text{all pieces ok}}$ also holds, then $\xi\lceil[-2^n, 2^n] \preccurlyeq \mathrm{Alg}^n(\mathsf{Input}) \preccurlyeq \xi\lceil[-9\cdot 2^n, 9\cdot 2^n]$. ■

Here is the main theorem of this section; it states that the events $E^n_{\text{xi does it}}$ and $E^n_{\text{all pieces ok}}$ occur very probably whenever the stopping times τ fulfill their task specified by $E^m_{\text{stop},\tau}$:

Theorem 1.6.1. *For some constant* $c_6 > 0$, $c_5 > 0$*:*

$$P\left[E^m_{\text{stop},\tau} \setminus (E^n_{\text{xi does it}} \cap E^n_{\text{all pieces ok}})\right] \le c_5 e^{-c_6 n}. \tag{1.6.4}$$

This theorem is proven the following three subsections. We split the proof into a purely combinatoric part and a probabilistic part. The combinatoric part (subsection 1.6.1 below for $E^n_{\text{xi does it}}$ and subsection 1.6.2 below for $E^n_{\text{all pieces ok}}$) shows that whenever some more "basic" events (named $B^n_{\dots}$ below, where "..." stands for a varying label) and $E^m_{\text{stop},\tau}$ occur, then the events $E^n_{\text{xi does it}}$ and $E^n_{\text{all pieces ok}}$ occur, too. In the probabilistic part (subsection 1.6.3 below) we show that these basic events $B^n_{\dots}$ are highly probable, at least when $E^m_{\text{stop},\tau}$ occurs.

The **Proof of Theorem 1.3.6** is an immediate consequence of Lemma 1.6.1 and Theorem 1.6.1.

1.6.1 Combinatorics concerning $E^n_{\text{xi does it}}$

In this subsection, we show that a piece of ξ centered at the origin passes all the tests specified by the Filter_i, provided some basic events $B^n_{\dots}$ (specified below) hold.

Definition 1.6.2. *For $n \in \mathbb{N}$ we define the following events:*

$$B^n_{\text{sig rl}} := \left\{ \begin{array}{l} \text{For every right ladder path } \pi \in [-2\cdot l2^{2n}, 2\cdot l2^{2n}]^{[0,c_1 n/2[} \text{ and} \\ \text{for every admissible piece of path } \pi' \in \mathsf{AdPath}(2\cdot l2^{2n}, c_1 n/2): \\ \text{If } \xi\circ\pi = \xi\circ\pi', \text{ then } \pi(c_1 n/2-1) \geq \pi'(c_1 n/2-1). \end{array} \right\}, \tag{1.6.5}$$

$$B^n_{\text{sig rr}} := \left\{ \begin{array}{l} \text{For every right ladder path } \pi \in [-2\cdot l2^{2n}, 2\cdot l2^{2n}]^{[0,c_1 n/2[} \text{ and} \\ \text{for every admissible piece of path } \pi' \in \mathsf{AdPath}(2\cdot l2^{2n}, c_1 n/2): \\ \text{If } \xi\circ\pi = \xi\circ\pi', \text{ then } \pi(0) \leq \pi'(0). \end{array} \right\}. \tag{1.6.6}$$

Let $B^n_{\text{sig ll}}$ and $B^n_{\text{sig lr}}$ be defined just as $B^n_{\text{sig rl}}$ and $B^n_{\text{sig rr}}$ with "right ladder path" replaced by "left ladder path" and with "$\leq$" and "$\geq$" exchanged in (1.6.5) and (1.6.6). We set

$$B^n_{\text{signals}} := B^n_{\text{sig rl}} \cap B^n_{\text{sig rr}} \cap B^n_{\text{sig ll}} \cap B^n_{\text{sig lr}}, \tag{1.6.7}$$

$$E^n_{\text{signals II}} := \left\{ \begin{array}{l} \text{For every ladder path } \pi \in [-2\cdot l2^{2n}, 2\cdot l2^{2n}]^{[0,c_1 n[} \text{ and for every} \\ \text{admissible piece of path } \pi' \in \mathsf{AdPath}(2\cdot l2^{2n}, c_1 n): \\ \text{If } \xi\circ\pi = \xi\circ\pi', \text{ then } \pi(c_1 n/2) = \pi'(c_1 n/2). \end{array} \right\}. \tag{1.6.8}$$

Lemma 1.6.2. $B^n_{\text{signals}} \subseteq E^n_{\text{signals II}}$.

Proof of Lemma 1.6.2. Assume that the event B^n_{signals} occurs. Let $\pi \in [-2\cdot l2^{2n}, 2\cdot l2^{2n}]^{[0,c_1 n[}$ be a right ladder path and $\pi' \in \mathsf{AdPath}(2\cdot l2^{2n}, c_1 n)$. Assume that $\xi\circ\pi = \xi\circ\pi'$ holds. Looking at the first half of π and π' only (with the first points $(0,\pi(0))$, $(0,\pi'(0))$ dropped), we see $\pi(c_1 n/2) \geq \pi'(c_1 n/2)$, since $B^n_{\text{sig rl}}$ holds. Similarly, looking at the second half of π and π' only, we infer $\pi(c_1 n/2) \leq \pi'(c_1 n/2)$, since $B^n_{\text{sig rr}}$ holds. Therefore $\pi(c_1 n/2)$ and $\pi'(c_1 n/2)$ coincide. The case of left ladder paths is treated similarly. This shows that $E^n_{\text{signals II}}$ holds. ■

Definition 1.6.3. *By definition, the event $B^n_{\text{all paths}}$ occurs if and only if the following holds: every admissible piece of path $R \in [-12\cdot 2^n, 12\cdot 2^n]^{[0,3c_1 n[}$ occurs in the random walk S with start at most 2^{2n} time steps after some stopping time $\tau(k)$, $k < 2^{\alpha n}$. More formally:*

$$B^n_{\text{all paths}} := \left\{ \begin{array}{l} \forall R \in \mathsf{AdPaths}(12\cdot 2^n, 3c_1 n)\, \exists k \in [0, 2^{\alpha n}[\, \exists j \in [0, 2^{2n}] : \\ \mathsf{TimeShift}^{\tau(k)+j}(R) \subseteq S \end{array} \right\}. \tag{1.6.9}$$

The following auxiliary lemma helps us to show below that the true scenery ξ passes the test Filter_1. Roughly speaking, it tells us that sufficiently many ladder words occur in the puzzle. This is important, since playing our puzzle game would lead to a failure when pieces were missing.

Lemma 1.6.3. *Assume that the event $B^n_{\text{all paths}} \cap B^n_{\text{signals}} \cap E^m_{\text{stop},\tau}$ holds. Let $I \subseteq [-6 \cdot 2^n, 6 \cdot 2^n]$ be a right (or left) ladder interval with $|I| = 3c_1 n$, and let $w_1, w_2, w_3 \in \mathcal{C}^{c_1 n}$ with $(\xi\lceil I)_{\rightarrow} = w_1 w_2 w_3$ (or $(\xi\lceil I)_{\leftarrow} = w_1 w_2 w_3$ in the case of a left ladder interval). Then $(w_1, w_2, w_3) \in \mathsf{Puzzle}^n_{\mathrm{I}}(\mathsf{Input})$.*

Proof of Lemma 1.6.3. Assume that I is a right ladder interval; the case of left ladder intervals can be treated in the same way by exchanging "left" and "right". Let $I = I_1 \cup I_2 \cup I_3$, where I_1, I_2, and I_3 denote the left, middle, and right third of I, respectively; thus $(\xi\lceil I_i)_{\rightarrow} = w_i$, $i = 1, 2, 3$. Since the event $B^n_{\text{all paths}}$ holds, the straight path which steps through the elements of I from the left to the right in $3c_1 n$ steps is realized at least once by the random walk $(S(t))_{t \geq 0}$ within time 2^{2n} of a stopping time $\tau(k)$, $k < 2^{\alpha n}$. Observing ξ along such a straight path generates the word $w_1 w_2 w_3$. Thus

$$(w_1, w_2, w_3) \in \mathsf{PrePuzzle}^n(\mathsf{Input}). \tag{1.6.10}$$

Let w_2' be such that $(w_1, w_2', w_3) \in \mathsf{PrePuzzle}^n(\mathsf{Input})$. In order to prove the claim $(w_1, w_2, w_3) \in \mathsf{Puzzle}^n_{\mathrm{I}}(\mathsf{Input})$ it remains to show: $w_2 = w_2'$. When the event $E^m_{\text{stop},\tau}$ holds, the stopping times of $\tau(k)$, $k < 2^{\alpha n}$, all stop the random walk $(S(t))_{t \geq 0}$ somewhere in the interval $[-2^n, 2^n]$. Within time 2^{2n} the random walk moves at most a distance $l 2^{2n}$. Because of $w_1 w_2' w_3 \in \mathsf{PrePuzzle}^n(\mathsf{Input})$, the word $w_1 w_2' w_3$ occurs somewhere in the observations at most 2^{2n} time steps after a stopping time $\tau(k)$, $k < 2^{\alpha n}$. Within time 2^{2n} after a stopping time, the random walk cannot be further away from the origin than $l 2^{2n} + 2^n \leq 2 \cdot l 2^{2n}$, since the event $E^m_{\text{stop},\tau}$ holds. Thus there exists an admissible piece of path $R' : [0, 3c_1 n[\rightarrow [-2 \cdot l 2^{2n}, 2 \cdot l 2^{2n}]$ such that $\xi \circ R' = w_1 w_2' w_3$. Let $R : [0, 3c_1 n[\rightarrow I \subseteq [-2 \cdot l 2^{2n}, 2 \cdot l 2^{2n}]$ denote the right ladder path which passes through I from the left to the right. We know $\xi \circ R'\lceil[0, c_1 n[= \xi \circ R\lceil[0, c_1 n[= w_1$ and $(\xi \circ R'\lceil[2c_1 n, 3c_1 n[)_{\rightarrow} = (\xi \circ R\lceil[2c_1 n, 3c_1 n[)_{\rightarrow} = w_3$. Furthermore, the event $E^n_{\text{signals II}} \supseteq B^n_{\text{signals}}$ holds; see Lemma 1.6.2. Abbreviating $x := c_1 n/2$ and $y := 5c_1 n/2$, this implies $R'(x) = R(x)$ and $R'(y) = R(y)$. But $R\lceil[x, y]$ is a right ladder path; thus $R'\lceil[x, y]$ must be the same right ladder path, since only right ladder paths can travel equally fast to the right as R does. Hence $w_2 = (\xi \circ R\lceil[c_1 n, 2c_1 n[)_{\rightarrow} = (\xi \circ R'\lceil[c_1 n, 2c_1 n[)_{\rightarrow} = w_2'$. This finishes the proof of Lemma 1.6.3. ■

Corollary 1.6.1. *If the event $B^n_{\text{all paths}} \cap B^n_{\text{signals}} \cap E^m_{\text{stop},\tau}$ holds, then $\xi\lceil[-5 \cdot 2^n, 5 \cdot 2^n] \in \mathsf{Filter}^n_1(\mathsf{Input})$.*

Proof of Corollary 1.6.1. Assume that $B^n_{\text{all paths}} \cap B^n_{\text{signals}} \cap E^m_{\text{stop},\tau}$ holds, and let $I_2 \subseteq [-5 \cdot 2^n, 5 \cdot 2^n]$, $|I_2| = c_1 n$, be a right ladder interval. Set $I_1 := I_2 - c_1 n l_{\rightarrow}$ and $I_3 := I_2 + c_1 n l_{\rightarrow}$; these are right ladder intervals adjacent to the left and to the right of I_2, respectively. Thus $I := I_1 \cup I_2 \cup I_3$ is a right ladder interval, $|I| = 3c_1 n$. Since $n \geq n_0$ and n_0 is large enough, we obtain $I \subseteq [-6 \cdot 2^n, 6 \cdot 2^n]$. We set $w_i := (\xi\lceil I_i)_{\rightarrow}$, $i = 1, 2, 3$. We have $(w_1, w_2, w_3) \in \mathsf{Puzzle}^n_{\mathrm{I}}(\mathsf{Input})$ by Lemma 1.6.3; thus $w_2 \in \mathsf{Puzzle}^n_{\mathrm{II}}(\mathsf{Input})$. This finishes the proof of Corollary 1.6.1. ■

The following definitions are analogous to the definition of Filter^n_2 and Filter^n_3, with the "reconstructed candidate" w replaced by the true scenery ξ, and with the domain

$[-5\cdot 2^n, 5\cdot 2^n]$ replaced by the larger domain $[-9\cdot 2^n, 9\cdot 2^n]$. We insert the corresponding statements for left ladder intervals, too; this turns out to be useful only in the next subsection.

Definition 1.6.4.

$$E^n_{\text{neighbor I}} := \tag{1.6.11}$$
$$\left\{\begin{array}{l}\text{For all right ladder intervals } I,J \subseteq [-9\cdot 2^n, 9\cdot 2^n]\text{: if } I \triangleright_n J, \text{ then}\\ ((\xi\lceil I)_\rightarrow, (\xi\lceil J)_\rightarrow) \in \mathsf{Neighbors}^n(\tau,\eta).\\ \text{For all left ladder intervals } I,J \subseteq [-9\cdot 2^n, 9\cdot 2^n]\text{: if } I \triangleleft_n J, \text{ then}\\ ((\xi\lceil I)_\leftarrow, (\xi\lceil J)_\leftarrow) \in \mathsf{Neighbors}^n(\tau,\eta).\end{array}\right\},$$

$$E^n_{\text{neighbor II}} := \tag{1.6.12}$$
$$\left\{\begin{array}{l}\text{For all right ladder intervals } I,J \subseteq [-9\cdot 2^n, 9\cdot 2^n],\ |I| = |J| = c_1 n\text{:}\\ \text{if } ((\xi\lceil I)_\rightarrow, (\xi\lceil J)_\rightarrow) \in \mathsf{Neighbors}^n(\tau,\eta), \text{ then there is } q\in\mathbb{N} \text{ such that}\\ I \triangleright_n J + q l_\rightarrow.\\ \text{For all left ladder intervals } I,J \subseteq [-9\cdot 2^n, 9\cdot 2^n],\ |I| = |J| = c_1 n\text{:}\\ \text{if } ((\xi\lceil I)_\leftarrow, (\xi\lceil J)_\leftarrow) \in \mathsf{Neighbors}^n(\tau,\eta), \text{ then there is } q\in\mathbb{N} \text{ such that}\\ I \triangleleft_n J - q l_\leftarrow.\end{array}\right\}.$$

Lemma 1.6.4. *If the event $B^n_{\text{all paths}}$ holds, then the event $E^n_{\text{neighbor I}}$ holds too, and consequently $\xi\lceil[-5\cdot 2^n, 5\cdot 2^n] \in \mathsf{Filter}^n_2(\mathsf{Input})$.*

Proof of Lemma 1.6.4. Assume that the event $B^n_{\text{all paths}}$ holds. We treat only the case of right ladder intervals; the case of left ladder intervals can be treated in the same way by exchanging right with left, $\rightarrow$ with $\leftarrow$, and $\triangleright_n$ with $\triangleleft_n$.

Let $I, J \subseteq [-9\cdot 2^n, 9\cdot 2^n]$ be right ladder intervals such that $I \triangleright_n J$. We need to prove $((\xi\lceil I)_\rightarrow, (\xi\lceil J)_\rightarrow) \in \mathsf{Neighbors}^n(\mathsf{Input})$. Let $i_l := \min I$, $i_r := \max I$, $j_l := \min J$, and $j_r := \max J$. Since $I \triangleright_n J$, there exists an admissible piece of path consisting of $h+1 = l|\mathcal{M}|+1$ points starting in i_r and ending in j_l. Since $I \triangleright_n J$ we have $|I|, |J| = c_1 n$. Thus there exists an admissible piece of path $R : [0, 2c_1 n + h - 1[\rightarrow [i_l, j_r]$ starting at i_l and ending in j_r; furthermore we can require that $R\lceil[0, c_1 n[$ and $R\lceil(c_1 n + h - 1 + [0, c_1 n[)$ are right ladder paths. Set $w_1 = (\xi\lceil I)_\rightarrow$ and $w_2 = (\xi\lceil J)_\rightarrow$; then $\xi \circ R = w_1 w w_2$ where $w \in \mathcal{C}^{h-1}$. Since $n \geq n_0$ holds and n_0 is large enough, we have $h \leq c_1 n$. Thus the piece of path R has length shorter than or equal to $3c_1 n$. The range $\mathrm{rng}(R)$ of R fulfills $\mathrm{rng}(R) \subseteq [-10\cdot 2^n, 10\cdot 2^n]$, and since $B^n_{\text{all paths}}$ holds, the random walk $(S(t))_{t\geq 0}$ "follows the path" R at least once within time 2^{2n} after a stopping time of τ. In other words, there exists $k \in [0, 2^{\alpha n}[$ and $j \in [0, 2^{2n} - 2c_1 n - h + 1]$ such that for all $i \in [0, 2c_1 n + h - 1[$ we have $S(\tau(k) + j + i) = R(i)$. Thus we get $\xi \circ S\lceil(\tau(k) + j + [0, 2c_1 n + h - 1[) \equiv w_1 w w_2$. This implies that $(w_1, w_2) \in \mathsf{Neighbors}^n(\mathsf{Input})$ and thus $((\xi\lceil I)_\rightarrow, (\xi\lceil J)_\rightarrow) \in \mathsf{Neighbors}^n(\mathsf{Input})$. ■

The following elementary number theoretic lemma serves to replace admissible pieces of path with more than h steps by admissible pieces of path with h steps, up to a sequence of maximal steps in one direction:

Lemma 1.6.5. *Let $s = (s_j)_{j=1,\ldots,K} \in \mathcal{M}^K$, $K \in \mathbb{N}$. Then there is $(r_j)_{j=1,\ldots,h} \in \mathcal{M}^h$ with*

$$\sum_{j=1}^{h} r_j + (K-h) l_\rightarrow - \sum_{j=1}^{K} s_j \in l_\rightarrow \mathbb{N}. \tag{1.6.13}$$

Similarly, there is $(r'_j)_{j=1,\dots,h} \in \mathcal{M}^h$ *with*

$$\sum_{j=1}^{h} r'_j - (K-h)l_{\leftarrow} - \sum_{j=1}^{K} s_j \in -l_{\leftarrow}\mathbb{N}. \tag{1.6.14}$$

Proof. In order to treat (1.6.13) and (1.6.14) simultaneously, let $l_{\leftrightarrow}$ denote either $l_{\rightarrow}$ or $-l_{\leftarrow}$. For $a \in \mathcal{M}$ let n_a denote the number of $j = 1, \dots, K$ such that $s_j = a$. Let $n'_a \in [0, |l_{\leftrightarrow}|[\cap(n_a + l_{\leftrightarrow}\mathbb{Z})$ denote the remainder of n_a modulo $l_{\leftrightarrow}$. Then $\sum_{a\in\mathcal{M}} n'_a \le h$. Choose any list $(r_j)_{j=1,\dots,h} \in \mathcal{M}^h$ having n'_a entries a for every $a \in \mathcal{M}\setminus\{l_{\leftrightarrow}\}$ and $h - \sum_{a\in\mathcal{M}\setminus\{l_{\leftrightarrow}\}} n'_a$ entries $l_{\leftrightarrow}$. Set

$$q := \frac{1}{l_{\leftrightarrow}} \sum_{a\in\mathcal{M}} (n_a - n'_a)(l_{\leftrightarrow} - a) \in \mathbb{N}; \tag{1.6.15}$$

note $(l_{\leftrightarrow} - a)/l_{\leftrightarrow} \ge 0$ and $n_a - n'_a \in |l_{\leftrightarrow}|\mathbb{N}$. Then

$$\begin{aligned} \sum_{j=1}^{K} (l_{\leftrightarrow} - s_j) &= \sum_{a\in\mathcal{M}} n_a(l_{\leftrightarrow} - a) \qquad (1.6.16)\\ &= ql_{\leftrightarrow} + \sum_{a\in\mathcal{M}} n'_a(l_{\leftrightarrow} - a) = ql_{\leftrightarrow} + \sum_{j=1}^{h} (l_{\leftrightarrow} - r_j), \end{aligned}$$

which implies the claim (1.6.13) or (1.6.14), respectively. ■

Lemma 1.6.6. *If the event* $E^n_{\text{signals II}} \cap E^m_{\text{stop},\tau}$ *holds, then the event* $E^n_{\text{neighbor II}}$ *holds, too, and consequently* $\xi\lceil[-5\cdot 2^n, 5\cdot 2^n] \in \mathsf{Filter}^n_3(\mathsf{Input})$.

Proof. Assume that the events $E^n_{\text{signals II}}$ and $E^m_{\text{stop},\tau}$ hold. We treat here the case of right ladder intervals:

Let $I, J \subseteq [-9\cdot 2^n, 9\cdot 2^n]$ be right ladder intervals with $|I| = |J| = c_1 n$, and assume $((\xi\lceil I)_{\rightarrow}, (\xi\lceil J)_{\rightarrow}) \in \mathsf{Neighbors}^n(\mathsf{Input})$. We need to show $I \triangleright_n J + ql_{\rightarrow}$ for some $q \in \mathbb{N}$.

Using Definition 1.5.7 of $\mathsf{Neighbors}^n$ and the abbreviations $w_1 := (\xi\lceil I)_{\rightarrow}$ and $w_2 := (\xi\lceil J)_{\rightarrow}$, we see: There is an admissible piece of path $R : [0, 2c_1 n + h - 1[\to \mathbb{Z}$ with the following properties:

- R is realized by the random walk S in during some time interval $D \subseteq \tau(k) + [0, 2^{2n}]$, $|D| = 2c_1 n + h - 1$, for some $k \in [0, 2^{\alpha n}[$. This means: R equals $S\lceil D$ when time-shifted back to the origin.
- Observing the scenery ξ along R produces $w_1 w w_2$ for some $w \in \mathcal{C}^{h-1}$; i.e.: $\xi \circ R = w_1 w w_2$.

We know $|\tau(k)| \le 2^n$ since the event $E^m_{\text{stop},\tau}$ holds; thus R takes all its values in $[-(2^n + l2^{2n}), 2^n + l2^{2n}] \subseteq [-2\cdot l2^{2n}, 2\cdot l2^{2n}]$, since the random walk cannot travel faster than distance l per step. We examine the first $c_1 n$ steps of R: $(\xi \circ R\lceil[0, c_1 n[)_{\rightarrow} = w_1 = (\xi\lceil I)_{\rightarrow}$ implies $R(c_1 n/2) = \min I + c_1 n l_{\rightarrow}/2$, since the event $E^n_{\text{signals II}}$ holds; note that $x := \min I + c_1 n l_{\rightarrow}/2$ is the point in the middle of a right ladder path walking through I. The same argument applies to the last $c_1 n$ steps of R: $(\xi \circ R\lceil(c_1 n + h - 1 + [0, c_1 n[))_{\rightarrow} = w_2 = (\xi\lceil J)_{\rightarrow}$ implies $R(3c_1 n/2 + h - 1) = \min J + c_1 n l_{\rightarrow}/2 =: y$; y is the point in the middle of J. The path R travels from x to y in $K := c_1 n + h - 1 \ge h$ steps,

using some step sizes $(s_j)_{j=1,\dots,K} \in \mathcal{M}^K$. As a consequence of (1.6.13) in Lemma 1.6.5, there is $(r_j)_{j=1,\dots,h} \in \mathcal{M}^h$ with $\sum_{j=1}^h r_j + (K-h)l_\rightarrow - \sum_{j=1}^K s_j = ql_\rightarrow$ for some $q \in \mathbb{N}$. Since $\max I - x = (c_1 n/2 - 1)l_\rightarrow$ and $y - \min J = c_1 n l_\rightarrow/2$, we obtain $\min J - \max I = y - x - (c_1 n - 1)l_\rightarrow = \sum_{j=1}^K s_j - (c_1 n - 1)l_\rightarrow = \sum_{j=1}^h r_j - ql_\rightarrow$. This means $I \triangleright_n (J + ql_\rightarrow)$, as we wanted to show.

Summarizing, this implies $\xi\lceil[-5\cdot 2^n, 5\cdot 2^n] \in \mathsf{Filter}_3^n(\mathsf{Input})$ and the first statement in the definition of $E^n_{\text{neighbor II}}$, which treats right ladder intervals.

The proof for left ladder intervals can be treated analogously. Altogether, we see that the event $E^n_{\text{neighbor II}}$ is valid. ■

Definition 1.6.5. *We define the event*

$$B^n_{\text{seed I}} := \left\{ \begin{array}{l} \text{For every modulo class } Z \in \mathbb{Z}/l_\rightarrow\mathbb{Z} \text{ there exists } k \in [0, 2^{\alpha n}[\\ \text{such that } S(\tau(k)+h) \in Z,\ S\lceil(\tau(k)+h+[0,3c_1 n l_\leftarrow]) \text{ is a right} \\ \text{ladder path, and } S\lceil(\tau(k)+h+3c_1 n l_\leftarrow + [0,3c_1 n l_\rightarrow]) \text{ is a left} \\ \text{ladder path.} \end{array} \right\}. \quad (1.6.17)$$

Lemma 1.6.7. *If the events $B^n_{\text{all paths}}$, B^n_{signals}, $B^n_{\text{seed I}}$ and $E^m_{\text{stop},\tau}$ hold, then $\xi\lceil[-5\cdot 2^n, 5\cdot 2^n] \in \mathsf{Filter}_4^n(\mathsf{Input})$.*

Proof of Lemma 1.6.7. Assume that the event $B^n_{\text{all paths}} \cap B^n_{\text{signals}} \cap B^n_{\text{seed I}} \cap E^m_{\text{stop},\tau}$ holds. Let $Z \in \mathbb{Z}/l_\rightarrow\mathbb{Z}$. Since $B^n_{\text{seed I}}$ holds, there exists a $k \in [0, 2^{\alpha n}[$ such that $S(\tau(k)+h) \in Z$, $R_1 := S\lceil(\tau(k)+h+[0,3c_1 n l_\leftarrow])$ is a right ladder path, and $R_2 := S\lceil(\tau(k)+h+3c_1 n l_\leftarrow + [0,3c_1 n l_\rightarrow])$ is a left ladder path. Since $E^m_{\text{stop},\tau}$ holds, we know $S(\tau(k)) \in [-2^n, 2^n]$. Thus the random walk S cannot leave the interval $[-2\cdot 2^n, 2\cdot 2^n]$ during the time interval $\tau(k) + [h + 3c_1 n l_\leftarrow + 3c_1 n l_\rightarrow]$, since $(h + 3c_1 n l_\leftarrow + 3c_1 n l_\rightarrow)l < 2^n$, and the random walk cannot travel faster than l per step. Thus R_1 and R_2 take all their values in $[-2\cdot 2^n, 2\cdot 2^n]$. Note that the right ladder path R_1 and the left ladder path walk R_2 traverse precisely the same interval, R_1 using step size $l_\rightarrow$ to the right, and R_2 with step size $-l_\leftarrow$ back. The same is true when we restrict R_1 and R_2 to the smaller time intervals $[t_1, t_1'] := \tau(k)+h+c_1 n l_\leftarrow + [0, c_2 n l_\leftarrow]$ and $[t_2, t_2'] := \tau(k)+h+3c_1 n l_\leftarrow + 2c_1 n l_\rightarrow + [-c_2 n l_\rightarrow, 0]$, respectively: We have $S(t_1) = S(t_2') =: a$, and $S(t_1') = S(t_2) =: b$, and $S\lceil[t_1, t_1']$ is a right ladder path: it traverses $[a,b]$ from the left to the right, while on $S\lceil[t_2, t_2']$ it is a left ladder path; it traverses $[a,b]$ in opposite direction. In particular, reading only every $l_\leftarrow$th letter in $\chi\lceil[t_1, t_1']$ and only every $l_\rightarrow$th letter in $\chi\lceil[t_2, t_2']$ yield the same word, only in reversed direction:

$$(\chi\lceil([t_1, t_1'] \cap (t_1 + l_\leftarrow\mathbb{Z}))_\rightarrow = (\xi\lceil([a,b] \cap (a + l_\rightarrow l_\leftarrow \mathbb{Z})))_\rightarrow = (\chi\lceil[t_2, t_2'] \cap (t_1 + l_\leftarrow\mathbb{Z}))_\leftarrow. \quad (1.6.18)$$

We consider the words $u_1u_2u_3 := \chi\lceil(t_1 - c_1 n + [0, 3c_1 n[)$ and $v_1v_2v_3 := \chi\lceil(t_2 - c_1 n + [0, 3c_1 n[)$ with $u_i, v_i \in \mathcal{C}^{c_1 n}$; note that $t_1 - c_1 n + [0, 3c_1 n[\subseteq \mathrm{domain}(R_1)$ and $t_2 - c_1 n + [0, 3c_1 n[\subseteq \mathrm{domain}(R_2)$. We get $(u_1, u_2, u_3), (v_1, v_2, v_3) \in \mathsf{Puzzle}^n_{\text{I}}(\mathsf{Input})$ by Lemma 1.6.3. Hence we obtain $(w_1, w_2, w_3) \in \mathsf{Seed}^n_{\text{I}}(\mathsf{Input})$ by Definition (1.5.4), since the words $u_1u_2u_3$ and $v_1v_2v_3$ occur in the observations sufficiently close to a stopping time $\tau(k)$; more specifically: $t_1 - c_1 n, t_2 - c_1 n \in \tau(k) + [0, 7c_1 n l]$. Consequently $u_2, v_2 \in \mathsf{Seed}^n_{\text{II}}(\mathsf{Input})$ by Definition (1.5.5). Finally we observe

$$(u_2\lceil([0, c_2 n l_\leftarrow] \cap l_\leftarrow\mathbb{Z}))_\rightarrow = (\xi\lceil([a,b] \cap (a + l_\rightarrow l_\leftarrow\mathbb{Z})))_\rightarrow = (v_2\lceil([0, c_2 n l_\rightarrow] \cap l_\rightarrow\mathbb{Z}))_\leftarrow \quad (1.6.19)$$

by (1.6.18). Thus we have shown $u_2 \in \mathsf{Seed}^n_{\mathrm{III}}(\mathsf{Input})$, see (1.5.6). Since $u_2 = \xi \circ S\lceil(t_1 + [0, c_1 n[)$, and since $S\lceil(t_1 + [0, c_1 n[)$ is a right ladder path with values in $Z \cap [-2 \cdot 2^n, 2 \cdot 2^n]$, this implies $\xi\lceil[-5 \cdot 2^n, 5 \cdot 2^n] \in \mathsf{Filter}^n_4(\mathsf{Input})$. ■

Definition 1.6.6. *For $n \in \mathbb{N}$, we define the following event:*

$$B^n_{\text{unique fit}} := \left\{ \begin{array}{l} \text{For every } i, j \in \{1, \dots, l^2\}, \text{ every } i\text{-spaced interval } I \subseteq [-11 \cdot \\ 2^n, 11 \cdot 2^n], \text{ and every } j\text{-spaced interval } J \subseteq [-11 \cdot 2^n, 11 \cdot 2^n] \\ \text{with } |I| = |J| \geq c_2 n \text{ holds } (\xi\lceil I)_{\leftarrow} \neq (\xi\lceil J)_{\rightarrow}, \text{ and if } I \neq J, \\ \text{then } (\xi\lceil I)_{\rightarrow} \neq (\xi\lceil J)_{\rightarrow}. \end{array} \right\}, \quad (1.6.20)$$

Lemma 1.6.8. *If the event $B^n_{\text{unique fit}}$ holds, then $\xi\lceil[-5 \cdot 2^n, 5 \cdot 2^n] \in \mathsf{Filter}^n_5(\mathsf{Input})$.*

Proof. Using $c_2 \leq c_1$ (see subsection 1.2.1), this follows immediately from Definition 1.6.6 of the event $B^n_{\text{unique fit}}$, and of Definition 1.5.4 of Filter^n_5. ■

Theorem 1.6.2. $B^n_{\text{all paths}} \cap B^n_{\text{signals}} \cap B^n_{\text{seed I}} \cap B^n_{\text{unique fit}} \cap E^m_{\text{stop},\tau} \subseteq E^n_{\text{xi does it}}$

Proof. We collect the statements of Lemmas/Corollary 1.6.2, 1.6.1, 1.6.4, 1.6.6, 1.6.7, and 1.6.8 in the following list:

$$\begin{array}{lcl} B^n_{\text{signals}} & \subseteq & E^n_{\text{signals II}}, \\ B^n_{\text{all paths}} \cap B^n_{\text{signals}} \cap E^m_{\text{stop},\tau} & \subseteq & \{\xi\lceil[-5 \cdot 2^n, 5 \cdot 2^n] \in \mathsf{Filter}^n_1(\mathsf{Input})\}, \\ B^n_{\text{all paths}} & \subseteq & \{\xi\lceil[-5 \cdot 2^n, 5 \cdot 2^n] \in \mathsf{Filter}^n_2(\mathsf{Input})\}, \\ E^n_{\text{signals II}} \cap E^m_{\text{stop},\tau} & \subseteq & \{\xi\lceil[-5 \cdot 2^n, 5 \cdot 2^n] \in \mathsf{Filter}^n_3(\mathsf{Input})\}, \\ B^n_{\text{all paths}} \cap B^n_{\text{signals}} \cap B^n_{\text{seed I}} \cap E^m_{\text{stop},\tau} & \subseteq & \{\xi\lceil[-5 \cdot 2^n, 5 \cdot 2^n] \in \mathsf{Filter}^n_4(\mathsf{Input})\}, \\ B^n_{\text{unique fit}} & \subseteq & \{\xi\lceil[-5 \cdot 2^n, 5 \cdot 2^n] \in \mathsf{Filter}^n_5(\mathsf{Input})\}. \end{array}$$

The theorem is an immediate consequence these statements, using (1.6.1) and (1.5.9). ■

1.6.2 Combinatorics concerning $E^n_{\text{all pieces ok}}$

In this subsection, we show that a piece w that passes all the Filter_i occurs in the true scenery ξ near the origin, provided some "basic" events $B^n_{\dots}$ hold.

Definition 1.6.7. *We define the events*

$$B^n_{\text{recogn straight}} := \quad (1.6.21)$$
$$\left\{ \begin{array}{l} \text{For every } R \in \mathsf{AdPaths}(11 \cdot 2^n, c_1 n) \text{ with } R(c_1 n - 1) - R(0) \notin \{(c_1 n - 1) l_{\rightarrow}, -(c_1 n - \\ 1) l_{\leftarrow}\} \text{ there is } \bar{R} \in \mathsf{AdPaths}(12 \cdot 2^n, c_1 n) \text{ such that } R(0) = \bar{R}(0),\ R(c_1 n - 1) = \\ \bar{R}(c_1 n - 1), \text{ and } \xi \circ R \neq \xi \circ \bar{R}. \end{array} \right\},$$

$$E^n_{\text{only ladder}} := \left\{ \begin{array}{l} \text{For all } (w_1, w_2, w_3) \in \mathsf{Puzzle}^n_{\mathrm{I}}(\mathsf{Input}) \text{ and every admissible} \\ \text{piece of path } R : [0, 3c_1 n[\rightarrow [-11 \cdot 2^n, 11 \cdot 2^n] \text{ with } \xi \circ R = \\ w_1 w_2 w_3 \text{ holds: } w_2 \text{ is a ladder word of } \xi\lceil[-11 \cdot 2^n, 11 \cdot 2^n]. \end{array} \right\}. \quad (1.6.22)$$

Lemma 1.6.9. *We have*

$$B^n_{\text{all paths}} \cap B^n_{\text{recogn straight}} \subseteq E^n_{\text{only ladder}}. \quad (1.6.23)$$

Proof of Lemma 1.6.9. Assume that the event $B^n_{\text{all paths}} \cap B^n_{\text{recogn straight}}$ holds. Let $w_1w_2w_3 \in \mathsf{Puzzle}^n_{\mathrm{I}}(\mathsf{Input})$, and let $R : [0, 3c_1n[\to [-11 \cdot 2^n, 11 \cdot 2^n]$ be an admissible piece of path with $\xi \circ R = w_1w_2w_3$. We prove by contradiction that the event $E^n_{\text{only ladder}}$ holds: Assume w_2 is not a ladder word of $\xi\lceil[-11 \cdot 2^n, 11 \cdot 2^n]$. Since $B^n_{\text{recogn straight}}$ holds, there exists an admissible piece of path $\bar{R} : [c_1n, 2c_1n[\to [-11 \cdot 2^n, 11 \cdot 2^n]$ such that $R(c_1n) = \bar{R}(c_1n)$ and $R(2c_1n - 1) = \bar{R}(2c_1n - 1)$, but $w_2 \neq (\xi \circ \bar{R})_{\to} =: w_2'$. Let $\check{R} : [0, 3c_1n[\to [-11 \cdot 2^n, 11 \cdot 2^n]$ be the admissible piece of path which on $[c_1n, 2c_1n[$ is equal to $\bar{R}$ and otherwise is equal to R. We have $\xi \circ \check{R} = w_1w_2'w_3$. Since $B^n_{\text{all paths}}$ holds, too, this implies that the random walk S follows the path of $\check{R}$ within time 2^{2n} from a stopping time of $\tau(k)$, $k < 2^{\alpha n}$. The same is valid for R, maybe with a different stopping time $\tau(k')$. In other words: $w_1w_2'w_3 \in \mathsf{PrePuzzle}^n(\mathsf{Input})$ and $w_1w_2w_3 \in \mathsf{PrePuzzle}^n(\mathsf{Input})$. This implies the contradiction $w_1w_2w_3 \notin \mathsf{Puzzle}^n_{\mathrm{I}}(\mathsf{Input})$; thus we have proved Lemma 1.6.9. ■

Definition 1.6.8. *We define the events*

$$B^n_{\text{outside out}} := \tag{1.6.24}$$
$$\left\{\begin{array}{l}\text{For every admissible piece of path} \\ R \in ([-2 \cdot l2^{2n}, 2 \cdot l2^{2n}] \setminus [-10 \cdot 2^n, 10 \cdot 2^n])^{[0, c_1n/2[}\text{: } \xi \circ R \text{ is not strongly} \\ \text{equivalent to any ladder word of length } c_1n/2 \text{ of } \xi\lceil[-9 \cdot 2^n, 9 \cdot 2^n].\end{array}\right\},$$

$$E^n_{\text{mod class}} := \tag{1.6.25}$$
$$\left\{\begin{array}{l}\text{For all } w \in \mathsf{Filter}^n_1(\mathsf{Input}) \text{ and for all right ladder intervals } I \subseteq [-2 \cdot 2^n, 2 \cdot 2^n], \\ |I| = c_1n\text{:} \\ \text{If there is a right ladder interval } J_r \subseteq [-2 \cdot 2^n, 2 \cdot 2^n] \text{ with } w\lceil I \equiv \xi\lceil J_r, \text{ then} \\ \xi\lceil([-2^n, 2^n] \cap (J_r + l_{\to}\mathbb{Z})) \sqsubseteq w\lceil(I + l_{\to}\mathbb{Z}) \sqsubseteq \xi\lceil([-9 \cdot 2^n, 9 \cdot 2^n] \cap (J_r + l_{\to}\mathbb{Z})), \text{ and} \\ \text{if } l_{\to} = l_{\leftarrow} \text{ and if there is a (left) ladder interval } J_l \subseteq [-2 \cdot 2^n, 2 \cdot 2^n] \text{ with } (w\lceil I)^{\leftrightarrow} \equiv \\ \xi\lceil J_l, \text{ then } \xi\lceil([-2^n, 2^n] \cap (J_l + l\mathbb{Z})) \sqsubseteq (w\lceil(I + l\mathbb{Z}))^{\leftrightarrow} \sqsubseteq \xi\lceil([-9 \cdot 2^n, 9 \cdot 2^n] \cap (J_l + l\mathbb{Z})).\end{array}\right\}.$$

Informally speaking, the meaning of the event $E^n_{\text{mod class}}$ is the following: If a "reconstructed" piece of scenery w contains a correct "seed piece" $w\lceil I$ over a sufficiently long ladder word, then the whole modulo class generated by I is reconstructed correctly. The reconstruction may generate the wrong orientation, but this is only allowed if left ladder intervals and right ladder intervals coincide, and if already the "seed piece" $w\lceil I$ is reversed compared with the true scenery ξ.

The next lemma formalizes the intuitive idea of "playing a puzzle game": We start with a seed word as reconstructed piece; then we append successively pieces of our puzzle that match to an ending of the growing reconstructed piece. This procedure continues until the reconstructed piece is large enough.

Lemma 1.6.10. *We have*

$$B^n_{\text{outside out}} \cap B^n_{\text{unique fit}} \cap E^n_{\text{only ladder}} \cap E^m_{\text{stop},\tau} \subseteq E^n_{\text{mod class}} \tag{1.6.26}$$

Proof of Lemma 1.6.10. Assume that the events on the left hand side of (1.6.26) hold. We claim that then $E^n_{\text{mod class}}$ holds, too. To prove this claim, let $w \in \mathsf{Filter}^n_1(\mathsf{Input})$, and let $I \subseteq [-2 \cdot 2^n, 2 \cdot 2^n]$, $|I| = c_1n$ be a right ladder interval. Assume that $J \subseteq [-2 \cdot 2^n, 2 \cdot 2^n]$ is a ladder interval. We assume one of the following two cases:

A) J is a right ladder interval, and $w\lceil I \equiv \xi\lceil J$;

B) $l_\rightarrow = l_\leftarrow$ and $(w\lceil I)^\leftrightarrow \equiv \xi\lceil J$.

We treat both cases simultaneously as far as possible; in order to unify the notation, let $\cdot^\sim$ denote the reversion operation $\cdot^\leftrightarrow$ in case B and the identity operation in case A. We set $Z := J + l_\rightarrow\mathbb{Z} \in \mathbb{Z}/l_\rightarrow\mathbb{Z}$; then it remains to show:

$$\xi\lceil([-2^n, 2^n] \cap Z) \sqsubseteq (w\lceil(I + l_\rightarrow\mathbb{Z}))^\sim \sqsubseteq \xi\lceil([-9 \cdot 2^n, 9 \cdot 2^n] \cap Z). \tag{1.6.27}$$

To prove the right hand side of (1.6.27), we prove by induction over all right ladder intervals I' with $I \subseteq I' \subseteq [-5 \cdot 2^n, 5 \cdot 2^n]$:

$$(w\lceil I')^\sim \sqsubseteq \xi\lceil([-9 \cdot 2^n, 9 \cdot 2^n] \cap Z). \tag{1.6.28}$$

Once we have proven this, the right hand side of (1.6.27) follows from the special case $I' = [-5 \cdot 2^n, 5 \cdot 2^n] \cap (I + l_\rightarrow\mathbb{Z})$.

The induction starts with $I = I'$: in this case (1.6.28) holds since our assumption A) or B), respectively, implies $(w\lceil I)^\sim \sqsubseteq \xi\lceil([-9 \cdot 2^n, 9 \cdot 2^n] \cap Z)$. For the induction step, assume that (1.6.28) holds for some I'. We enlarge I' by a single new point: let $I'' = I' \cup \{i\} \subseteq [-5 \cdot 2^n, 5 \cdot 2^n] \cap (I + l_\rightarrow\mathbb{Z})$ be a right ladder interval, $i \notin I'$. Let $I_i \subseteq I''$ be a right ladder interval with $|I_i| = c_1 n$ and $i \in I_i$. Using $w \in \mathsf{Filter}_1^n(\mathsf{Input})$ we see $w_2 := (w\lceil I_i)_\rightarrow \in \mathsf{Puzzle}_{\mathrm{II}}^n(\mathsf{Input})$. Hence there are $w_1, w_3 \in \mathcal{C}^{c_1 n}$ such that $(w_1, w_2, w_3) \in \mathsf{Puzzle}_{\mathrm{I}}^n(\mathsf{Input}) \subseteq \mathsf{PrePuzzle}^n(\mathsf{Input})$. Thus $w_1 w_2 w_3$ occurs in the observation χ at most 2^{2n} time steps after a stopping time $\tau(k)$, $k < 2^{\alpha n}$; say $w_1 w_2 w_3$ is read there in χ while the random walk follows an admissible piece of path $R : [0, 3c_1 n[\to \mathbb{Z}$; (we shifted the time domain of R back to the origin). Since the event $E^m_{\mathrm{stop},\tau}$ holds, we have $|S(\tau(k))| \le 2^n$. Within time 2^{2n} the random walk cannot travel farther than distance $l2^{2n}$; thus R has all its values in $[-(2^n + l2^{2n}), 2^n + l2^{2n}] \subseteq [-2 \cdot l2^{2n}, 2 \cdot l2^{2n}]$. Consider the ladder interval $I_i' := I_i \setminus \{i\} = I_i \cap I'$, $|I_i'| = c_1 n - 1 \ge c_1 n/2$: the induction hypothesis (1.6.28) implies $(w\lceil I_i')^\sim \sqsubseteq \xi\lceil([-9 \cdot 2^n, 9 \cdot 2^n] \cap Z)$; say $(w\lceil I_i')^\sim \equiv \xi\lceil D'$ for some right ladder interval $D' \subseteq [-9 \cdot 2^n, 9 \cdot 2^n] \cap Z$. Furthermore, $w_2' := (w\lceil I_i')_\rightarrow$ is a subword of $w_2 = (w\lceil I_i)_\rightarrow$ and thus also a subword of $\xi \circ R$. Hence we see, using that the event $B^n_{\text{outside out}}$ holds: R cannot take all of its values outside $[-10 \cdot 2^n, 10 \cdot 2^n]$; thus it has all its values in $[-10 \cdot 2^n - 3c_1 nl, 10 \cdot 2^n + 3c_1 nl] \subseteq [-11 \cdot 2^n, 11 \cdot 2^n]$. Since the event $E^n_{\text{only ladder}}$ holds, $w_2 = (w\lceil I_i)_\rightarrow$ is a ladder word of $\xi\lceil[-11 \cdot 2^n, 11 \cdot 2^n]$; say $w_2 = (\xi\lceil D)_\rightarrow$ for some right ladder interval $D \subseteq [-11 \cdot 2^n, 11 \cdot 2^n]$ (we call this "case A_1"), or $w_2 = (\xi\lceil D)_\leftarrow$ for some left ladder interval $D \subseteq [-11 \cdot 2^n, 11 \cdot 2^n]$ (call this "case B_1"). Thus w_2' occurs as a (possibly reversed) ladder word

- as a subword of $(\xi\lceil D)_\rightarrow$ in case A_1, or as a subword of $(\xi\lceil D)_\leftarrow$ in case B_1;
- as $w_2' = (\xi\lceil D')_\rightarrow$ in case A, or as $w_2' = (\xi\lceil D')_\leftarrow$ in case B.

Since the event $B^n_{\text{unique fit}}$ holds, this implies $D' \subseteq D$, and furthermore the reading directions have to coincide: If case A holds, then case A_1 occurs, and if case B holds, then case B_1 occurs. Let $T : \mathbb{Z} \to \mathbb{Z}$ denote the translation (case A) or reflection (case B) that transports $w\lceil I_i$ to $\xi\lceil D$. Then T transports $w\lceil I_i'$ to $\xi\lceil D'$, and thus – using once more that $B^n_{\text{unique fit}}$ holds – T is also the map that transports $w\lceil I'$ to a subpiece of

$\xi\lceil([-9\cdot 2^n, 9\cdot 2^n]\cap Z)$ according to the induction hypothesis (1.6.28). Hence T transports $w\lceil(I_i\cup I') = w\lceil I''$ to an equivalent subpiece of $\xi\lceil[-11\cdot 2^n, 11\cdot 2^n]$. To see that $T[w\lceil I'']$ is already a subpiece of $\xi\lceil([-9\cdot 2^n, 9\cdot 2^n]\cap Z)$, we proceed as follows: T maps the nonempty seed interval $I\subseteq[-2\cdot 2^n, 2\cdot 2^n]$ to $J\subseteq[-2\cdot 2^n, 2\cdot 2^n]\cap Z$; thus it has the form $T(z) = \pm z + a$ with $|a|\le 4\cdot 2^n$. Consequently T maps the domain $[-5\cdot 2^n, 5\cdot 2^n]$ of w to a subset of $[-9\cdot 2^n, 9\cdot 2^n]$. This shows $(w\lceil I'')^\sim \sqsubseteq \xi\lceil([-9\cdot 2^n, 9\cdot 2^n]\cap Z)$, which finishes our induction step and also the proof of the right hand side of the claim (1.6.27).

To prove the left hand side of (1.6.27), we observe that T^{-1} maps $[-2^n, 2^n]$ to a subset of $[-5\cdot 2^n, 5\cdot 2^n]$. Since T maps I to J, it maps the modulo class $I + l_\to\mathbb{Z}$ to $Z = J + l_\to\mathbb{Z}$; thus T^{-1} maps $[-2^n, 2^n]\cap Z$ to a subset of $(I+l_\to\mathbb{Z})\cap[-5\cdot 2^n, 5\cdot 2^n] = (I+l_\to\mathbb{Z})\cap\mathrm{domain}(w)$. Since T^{-1} maps a subpiece of $\xi\lceil([-9\cdot 2^n, 9\cdot 2^n]\cap Z)$ to $w\lceil(I + l_\to\mathbb{Z})$, this implies the left hand side of the claim (1.6.27). This finishes the proof of Lemma 1.6.10. ■

Definition 1.6.9. *We define the event*

$$E^n_{\text{seed II}} := \left\{\begin{array}{l}\text{Every } u\in \mathsf{Seed}^n_{\mathrm{II}}(\mathsf{Input}) \text{ is a left or right ladder word of } \xi\lceil[-2\cdot\\ 2^n, 2\cdot 2^n]. \text{ If } l_\to\neq l_\leftarrow, \text{ then every } u\in \mathsf{Seed}^n_{\mathrm{III}}(\mathsf{Input}) \text{ is a right}\\ \text{ladder word of } \xi\lceil[-2\cdot 2^n, 2\cdot 2^n].\end{array}\right\}. \quad (1.6.29)$$

Lemma 1.6.11. *We have*

$$B^n_{\text{unique fit}}\cap B^n_{\text{signals}}\cap B^n_{\text{all paths}}\cap B^n_{\text{recogn straight}}\cap E^m_{\text{stop},\tau}\subseteq E^n_{\text{seed II}}. \quad (1.6.30)$$

Proof of Lemma 1.6.11. Assume that the events on the left hand side of (1.6.30) hold. In order to show that the $E^n_{\text{seed II}}$ holds, let $w_2\in \mathsf{Seed}^n_{\mathrm{II}}(\mathsf{Input})$. We need to show that w_2 is a ladder word of $\xi\lceil[-2\cdot 2^n, 2\cdot 2^n]$. Using (1.5.5), we take $w_1, w_3\in\mathcal{C}^{c_1 n}$ with $(w_1, w_2, w_3)\in \mathsf{Seed}^n_{\mathrm{I}}(\mathsf{Input})$; thus $w_1w_2w_3\equiv\eta\lceil(\tau(k)+j+[0,3c_1n[)$ for some $k<2^{\alpha n}$ and $j\in[0,7c_1nl]$. Since $E^m_{\text{stop},\tau}$ holds, we have $|S(\tau(k))|\le 2^n$. Using $2^n + 7c_1nl^2 + 3c_1ln\le 2\cdot 2^n - c_1nl$, we see that the random walk S is located inside the interval $[-2\cdot 2^n + c_1nl, 2\cdot 2^n - c_1nl]$ during the time interval $\tau(k) + 7c_1nl + [0, 3c_1n[$. The word $w_1w_2w_3$ is read along an admissible piece of path, say $R\in\mathsf{AdPath}(2\cdot 2^n - c_1nl, 3c_1n)$ with $\xi\circ R = w_1w_2w_3$; (the time interval is shifted back to the origin). The event $E^n_{\text{only ladder}}$ holds by Lemma 1.6.9, and we have $(w_1, w_2, w_3)\in\mathsf{Puzzle}^n_{\mathrm{I}}(\mathsf{Input})$; hence w_2 is a ladder word of $\xi\lceil[-11\cdot 2^n, 11\cdot 2^n]$; say $w_2 = \xi\circ\pi$ for a ladder path $\pi : [0, c_1n[\to[-11\cdot 2^n, 11\cdot 2^n]$. Let $\pi' = R\lceil[c_1n, 2c_1n[$ be the middle piece of R, along which one observes $(\xi\circ\pi')_\to = w_2 = \xi\circ\pi$. Since the event $E^n_{\text{signals II}}$ holds by Lemma 1.6.2, we get $\pi'((3/2)c_1n) = \pi(c_1n/2)$; thus π takes least one value in $[-(2\cdot 2^n - c_1nl), 2\cdot 2^n - c_1nl]$; therefore all the values of π are in $[-2\cdot 2^n, 2\cdot 2^n]$. Thus w_2 is a ladder word of $\xi\lceil[-2\cdot 2^n, 2\cdot 2^n]$.

For the rest of the proof we assume $l_\to\neq l_\leftarrow$ and let $u\in\mathsf{Seed}^n_{\mathrm{III}}(\mathsf{Input})$. It remains to show: u is a right ladder word of $\xi\lceil[-2\cdot 2^n, 2\cdot 2^n]$. Using Definition (1.5.6) of $\mathsf{Seed}^n_{\mathrm{III}}$, we choose $v\in\mathsf{Seed}^n_{\mathrm{II}}(\mathsf{Input})$ with $(u\lceil(l_\leftarrow\mathbb{Z}\cap[0, c_2nl_\leftarrow]))_\to = (v\lceil(l_\to\mathbb{Z}\cap[0, c_2nl_\to]))_\leftarrow$. ¿From the first part of the proof we get: u and v are ladder words of $\xi\lceil[-2\cdot 2^n, 2\cdot 2^n]$, since $u, v\in\mathsf{Seed}^n_{\mathrm{II}}(\mathsf{Input})$. We distinguish three cases:

1. u is a right ladder word;

2. u and v are left ladder words;

3. u is a left ladder word and v is a right ladder word.

We need to show that case 1. holds; thus we prove that the cases 2. and 3. lead to a contradiction:

In **case 2.**, let $u = (\xi\lceil I)_\leftarrow$ and $v = (\xi\lceil J)_\leftarrow$ for some left ladder intervals $I, J \subseteq [-2\cdot 2^n, 2\cdot 2^n]$, $|I| = |J| = c_1 n$. We get $(u\lceil(l_\leftarrow\mathbb{Z}\cap[0, c_2 n l_\leftarrow]))_\rightarrow = (\xi\lceil I')_\leftarrow$ for some $l_\leftarrow^2$-spaced interval $I' \subseteq I$, $|I'| = c_2 n + 1$. Similarly, $(v\lceil(l_\rightarrow\mathbb{Z}\cap[0, c_2 n l_\rightarrow]))_\leftarrow = (\xi\lceil J')_\rightarrow$ for some $l_\leftarrow l_\rightarrow$-spaced interval $J' \subseteq J$, $|J'| = c_2 n + 1$. Thus $(\xi\lceil I')_\leftarrow = (\xi\lceil J')_\rightarrow$, which is incompatible with the event $B^n_{\text{unique fit}}$.

In **case 3.**, let $u = (\xi\lceil I)_\leftarrow$ for some left ladder interval $I \subseteq [-2\cdot 2^n, 2\cdot 2^n]$ and $v = (\xi\lceil J)_\rightarrow$ for some right ladder interval $J \subseteq [-2\cdot 2^n, 2\cdot 2^n]$, $|I| = |J| = c_1 n$. We get again $(u\lceil(l_\leftarrow\mathbb{Z}\cap[0, c_2 n l_\leftarrow]))_\rightarrow = (\xi\lceil I')_\leftarrow$ for some $l_\leftarrow^2$-spaced interval $I' \subseteq I$, $|I'| = c_2 n + 1$. This time we have $(v\lceil(l_\rightarrow\mathbb{Z}\cap[0, c_2 n l_\rightarrow]))_\leftarrow = (\xi\lceil J')_\leftarrow$ for some $l_\rightarrow^2$-spaced interval $J' \subseteq J$, $|J'| = c_2 n + 1$. Since $l_\leftarrow^2 \neq l_\rightarrow^2$, we have $I' \neq J'$. We obtain $(\xi\lceil I')_\leftarrow = (\xi\lceil J')_\leftarrow$, which is incompatible with the event $B^n_{\text{unique fit}}$;

Thus cases 2. and 3. cannot occur. Summarizing, we have proven that the event $E^n_{\text{seed II}}$ holds. ■

Definition 1.6.10. *If $l_\rightarrow = l_\leftarrow$, we define the event*

$$E^n_{\text{dist}} := \tag{1.6.31}$$
$$\left\{\begin{array}{l}\text{For all ladder intervals } I, J \subseteq [-9\cdot 2^n, 9\cdot 2^n],\ |I| = |J| = c_1 n\text{: if at least one}\\ \text{of } ((\xi\lceil I)_\rightarrow, (\xi\lceil J)_\rightarrow),\ ((\xi\lceil I)_\rightarrow, (\xi\lceil J)_\leftarrow),\ ((\xi\lceil I)_\leftarrow, (\xi\lceil J)_\rightarrow),\text{ or } ((\xi\lceil I)_\leftarrow, (\xi\lceil J)_\leftarrow)\\ \text{is in } \mathsf{Neighbors}^n(\mathsf{Input}),\text{ then distance}(I, J) \leq 3\cdot l c_1 n.\end{array}\right\}$$

In the case $l_\rightarrow \neq l_\leftarrow$, we set E^n_{dist} to be the sure event.

Lemma 1.6.12. $B^n_{\text{signals}} \cap E^m_{\text{stop},\tau} \subseteq E^n_{\text{dist}}$

Proof of Lemma 1.6.12. Assume that the event $B^n_{\text{signals}} \cap E^m_{\text{stop},\tau}$ holds, and that $l_\rightarrow = l_\leftarrow = l$. Let $I, J \subseteq [-9\cdot 2^n, 9\cdot 2^n]$, $|I| = |J| = c_1 n$ be right ladder intervals, and assume that there is a (w_1, w_2) among $((\xi\lceil I)_\rightarrow, (\xi\lceil J)_\rightarrow)$, $((\xi\lceil I)_\rightarrow, (\xi\lceil J)_\leftarrow)$, $((\xi\lceil I)_\leftarrow, (\xi\lceil J)_\rightarrow)$, or $((\xi\lceil I)_\leftarrow, (\xi\lceil J)_\leftarrow)$ with $(w_1, w_2) \in \mathsf{Neighbors}^n(\mathsf{Input})$. By definition (1.5.7), some word $w_1 w w_2$ with $w \in \mathcal{C}^{h-1}$ occurs in the observations χ at most 2^{2n} time steps after a stopping time $\tau(k)$, $k < 2^{\alpha n}$. Since $E^m_{\text{stop},\tau}$ holds, the random walk remains in the interval $[-2\cdot l 2^{2n}, 2\cdot l 2^{2n}]$ during that time interval; say the random walk follows an admissible piece of path $R : [0, 2c_1 n + h - 1[\rightarrow [-2\cdot l 2^{2n}, 2\cdot l 2^{2n}]$ while producing the observations $\xi \circ R = w_1 w w_2$; (we shifted the time domain back to the origin). R consists of the three pieces $\pi_1' = R\lceil[0, c_1 n[$, $\pi' = R\lceil(c_1 n + [0, h-1[)$, and $\pi_2' = R\lceil(c_1 n + h - 1 + [0, c_1 n[)$ with $\xi\circ\pi_1' = w_1$, $(\xi\circ\pi')_\rightarrow = w$, and $(\xi\circ\pi_2')_\rightarrow = w_2$. Let $x_1 := c_1 n/2$ and $x_2 := (3/2)c_1 n + h - 1$ be the points in the middle of the domain of π_1' and π_2', respectively. Then

$$|\pi_1'(x_1) - \pi_2'(x_2)| \leq (c_1 n + h - 1) l, \tag{1.6.32}$$

since the path R cannot travel faster than l per step. The event $E^n_{\text{signals II}}$ holds by Lemma 1.6.2. Let $\pi_1 : [0, c_1 n[\rightarrow I$ and $\pi_2 : c_1 n + h - 1 + [0, c_1 n[\rightarrow J$ be ladder paths with range I and J, respectively; we choose these paths to be left or right ladder paths according to whether the reading direction is "$\leftarrow$" or "$\rightarrow$". Hence, using $\xi\circ\pi_1 = w_1 = \xi\circ\pi_1'$ and $(\xi\circ\pi_2)_\rightarrow = w_2 = (\xi\circ\pi_2')_\rightarrow$, we obtain $\pi_1'(x_1) = \pi_1(x_1)$ and $\pi_2'(x_2) = \pi_2(x_2)$. Consequently (1.6.32) implies

$$\text{distance}(I, J) \leq |\pi_1(x_1) - \pi_2(x_2)| \leq 3\cdot l c_1 n. \tag{1.6.33}$$

Summarizing, we have shown that the event E^n_{dist} holds. ■

The following event $E^n_{\mathrm{mod}\,\gamma\,\mathrm{ok}}$ compares modulo classes (modulo some γ) in "reconstructed" pieces w with modulo classes in the "true" scenery ξ. Roughly speaking, it states that all modulo classes are reconstructed correctly, and either all of them are reconstructed in the correct orientation ("case A"), or all of them are reversed ("case B"). Even more, reversion is only allowed for symmetric maximal jumps of the random walk. Our goal is to show that this event holds for $\gamma = 1$ (at least if the basic events $B_{\dots}$ hold), but as intermediate steps, other values of γ are relevant, too.

Definition 1.6.11. *For all divisors $\gamma \geq 1$ of $l_\rightarrow$, we define the event*

$$E^n_{\mathrm{mod}\,\gamma\,\mathrm{ok}} := \tag{1.6.34}$$
$$\left\{ \begin{array}{l} \text{For all } w \in \mathsf{SolutionPieces}^n(\mathsf{Input}) \text{ there is a bijection } \iota_\gamma : \mathbb{Z}/\gamma\mathbb{Z} \to \mathbb{Z}/\gamma\mathbb{Z} \\ \text{such that (at least) one of the following two cases holds:} \\ \quad \textbf{A)}\ \forall Z \in \mathbb{Z}/\gamma\mathbb{Z}\colon\ \xi\lceil([-2^n, 2^n] \cap \iota_\gamma(Z)) \sqsubseteq w\lceil Z \sqsubseteq \xi\lceil([-9\cdot 2^n, 9\cdot 2^n] \cap \iota_\gamma(Z)) \\ \quad \textbf{B)}\ l_\rightarrow = l_\leftarrow \text{ and} \\ \qquad \forall Z \in \mathbb{Z}/\gamma\mathbb{Z}\colon\ \xi\lceil([-2^n, 2^n] \cap \iota_\gamma(Z)) \sqsubseteq (w\lceil Z)^\leftrightarrow \sqsubseteq \xi\lceil([-9\cdot 2^n, 9\cdot 2^n] \cap \\ \qquad \iota_\gamma(Z)) \end{array} \right\}.$$

Lemma 1.6.13. *For $\gamma = l_\rightarrow$, we have $E^n_{\mathrm{seed\,II}} \cap E^n_{\mathrm{mod\,class}} \cap E^n_{\mathrm{dist}} \cap B^n_{\mathrm{unique\,fit}} \subseteq E^n_{\mathrm{mod}\,l_\rightarrow\,\mathrm{ok}}$.*

Proof of Lemma 1.6.13. Assume that the event $E^n_{\mathrm{seed\,II}} \cap E^n_{\mathrm{mod\,class}} \cap E^n_{\mathrm{dist}} \cap B^n_{\mathrm{unique\,fit}}$ holds. Let $w \in \mathsf{SolutionPieces}^n(\mathsf{Input})$. Let $Z \in \mathbb{Z}/l_\rightarrow\mathbb{Z}$. In order to define $\iota(Z) = \iota_{l_\rightarrow}(Z)$, we proceed as follows: Since $w \in \mathsf{Filter}_4(\mathsf{Input})$, there exists a right ladder interval $I \subseteq Z \cap [-2\cdot 2^n, 2\cdot 2^n]$ such that $(w\lceil I)_\rightarrow \in \mathsf{Seed}^n_{\mathrm{III}}(\tau, \eta)$. We choose such an I. Then $(w\lceil I)_\rightarrow$ is a left or right ladder word of $\xi\lceil[-2\cdot 2^n, 2\cdot 2^n]$, since the event $E^n_{\mathrm{seed\,II}}$ holds. More specifically: for some right ladder interval $J \subseteq [-2\cdot 2^n, 2\cdot 2^n]$, at least one of the following two cases holds true:

$$\begin{array}{ll} \text{Case A}(Z)\colon & w\lceil I \equiv \xi\lceil J, \\ \text{Case B}(Z)\colon & l_\rightarrow = l_\leftarrow \text{ and } (w\lceil I)^\leftrightarrow \equiv \xi\lceil J. \end{array}$$

We define $\iota(Z) := J + l_\rightarrow\mathbb{Z} \in \mathbb{Z}/l_\rightarrow\mathbb{Z}$. Since the event $E^n_{\mathrm{mod\,class}}$ holds, we get

$$\begin{array}{ll} \text{for Case A}(Z)\colon & \xi\lceil([-2^n, 2^n] \cap \iota(Z)) \sqsubseteq w\lceil Z \sqsubseteq \xi\lceil([-9\cdot 2^n, 9\cdot 2^n] \cap \iota(Z)), \\ \text{for Case B}(Z)\colon & \xi\lceil([-2^n, 2^n] \cap \iota(Z)) \sqsubseteq (w\lceil Z)^\leftrightarrow \sqsubseteq \xi\lceil([-9\cdot 2^n, 9\cdot 2^n] \cap \iota(Z)). \end{array}$$

We claim that one of the following two cases occurs:

$$\begin{array}{ll} \text{Case A:} & \text{For all modulo classes } Z \in \mathbb{Z}/l_\rightarrow\mathbb{Z} \text{ holds Case A}(Z); \\ \text{Case B:} & \text{For all modulo classes } Z \in \mathbb{Z}/l_\rightarrow\mathbb{Z} \text{ holds Case B}(Z). \end{array}$$

This is obvious for $l_\rightarrow \neq l_\leftarrow$, since then Case B$(Z)$ cannot occur. To prove the claim for $l_\rightarrow = l_\leftarrow$, we proceed as follows: For $Z \in \mathbb{Z}/l\mathbb{Z}$, let $T_Z : \mathbb{Z} \to \mathbb{Z}$ denote a translation (Case A(Z)) or reflection (Case B(Z)) which transports $w\lceil Z$ to a subpiece of $\xi\lceil[-9\cdot 2^n, 9\cdot 2^n] \cap \iota(Z)$. Let $Z, W \in \mathbb{Z}/l\mathbb{Z}$. We choose two right ladder intervals $I_1 \subseteq Z \cap [4\cdot 2^n, 5\cdot 2^n]$, $I_2 \subseteq W \cap [4\cdot 2^n, 5\cdot 2^n]$, $|I_1| = |I_2| = c_1 n$, with $I_1 \rhd_n I_2$; such intervals exist, since $\operatorname{supp} \mu^{*h}$ meets every modulo class (modulo l) and since $n \geq n_0$ is large enough. We abbreviate $I'_1 := T_Z[I_1]$ and $I'_2 := T_W[I_2]$. Since $w \in \mathsf{Filter}_2(\mathsf{Input})$ one has $((w\lceil I_1)_\rightarrow, (w\lceil I_2)_\rightarrow) \in \mathsf{Neighbors}^n(\mathsf{Input})$. Let X_Z denote the symbol "$\rightarrow$" in the Case A(Z) and "$\leftarrow$" in the Case

B(Z). Then $((w\lceil I_1)_\rightarrow, (w\lceil I_2)_\rightarrow) = ((\xi\lceil I_1')_{X_Z}, (\xi\lceil I_2')_{X_W})$. Since the event E^n_{dist} holds, this implies distance$(I_1', I_2') \le 3 \cdot l c_1 n$. However, T_Z maps $[-5\cdot 2^n, 5\cdot 2^n]$ to $[-9\cdot 2^n - l, 9\cdot 2^n + l]$; (the extra summand l arises since T_Z was specified only by its action on a modulo class). Thus it maps $I_1, I_2 \subseteq [4\cdot 2^n, 5\cdot 2^n]$ to a subset of $[4\cdot 2^n - l, 9\cdot 2^n + l]$ in the Case A(Z), and to a subset of $[-9\cdot 2^n - l, -4\cdot 2^n + l]$ in the Case B(Z). The same statement holds with Z replaced by W. The intervals $[4\cdot 2^n - l, 9\cdot 2^n + l]$ and $[-9\cdot 2^n - l, -4\cdot 2^n + l]$ are farther apart than $3 \cdot l c_1 n \ge$ distance(I_1', I_2'); thus either both T_Z and T_W must be translations, or both must be reflections. Summarizing, we have shown so far that Case A holds or Case B holds.

It only remains to show that $\iota : \mathbb{Z}/l_\rightarrow\mathbb{Z} \to \mathbb{Z}/l_\rightarrow\mathbb{Z}$ is bijective. Since $\mathbb{Z}/l_\rightarrow\mathbb{Z}$ is finite, it suffices to show that ι is injective: Let $Z, W \in \mathbb{Z}/l_\rightarrow\mathbb{Z}$ with $\iota(Z) = \iota(W)$. Using the above maps T_Z, T_W again, we know

$$T_Z[Z \cap \text{domain}(w)] \quad = T_Z[Z \cap [-5\cdot 2^n, 5\cdot 2^n]] \subseteq \quad \iota(Z) \cap [-9\cdot 2^n, 9\cdot 2^n], \tag{1.6.35}$$

$$T_W[W \cap \text{domain}(w)] \quad = T_W[W \cap [-5\cdot 2^n, 5\cdot 2^n]] \subseteq \quad \iota(W) \cap [-9\cdot 2^n, 9\cdot 2^n]. \tag{1.6.36}$$

The sets on the right hand of (1.6.35) and (1.6.36) coincide; thus $T_Z[Z \cap [-5\cdot 2^n, 5\cdot 2^n]]$ and $T_W[W \cap [-5\cdot 2^n, 5\cdot 2^n]]$ overlap at least in $K \cap \iota(Z)$ for some interval K of length 2^n. We choose any right ladder interval $D \subseteq K \cap \iota(Z)$ with $|D| = c_1 n$ and set $D_1 := T_Z^{-1}[D]$ and $D_2 := T_W^{-1}[D]$. Then

$$\begin{aligned} &\text{Case A:} \quad (w\lceil D_1)_\rightarrow = (\xi\lceil D)_\rightarrow = (w\lceil D_2)_\rightarrow, \\ &\text{Case B:} \quad (w\lceil D_1)_\rightarrow = (\xi\lceil D)_\leftarrow = (w\lceil D_2)_\rightarrow; \end{aligned}$$

thus $w \in \mathsf{Filter}_5(\mathsf{Input})$ implies $D_1 = D_2$; hence $Z = D_1 + l_\rightarrow\mathbb{Z} = D_2 + l_\rightarrow\mathbb{Z} = W$. This shows that ι is indeed injective. ■

The next lemma contains a "step down" procedure in order to arrange correctly larger and larger modulo classes in a reconstructed piece of scenery w. Here is a rough idea for the rather complex construction:

Suppose we have already correctly reconstructed large pieces of the scenery ξ restricted to modulo classes (mod γ, say) up to a translation (and possibly a global reflection for all classes). Our task is to identify the relative translation between different modulo classes.

We start with a "reference" ladder word; it occurs over both, a ladder interval I in the reconstructed "candidate" scenery w, and a ladder interval J in the "true" scenery ξ (possibly reflected). Then we look for the rightmost "neighboring" ladder words that occur *not* in the same modulo class as the reference word, both in the candidate scenery and in the true scenery; we use here the "estimated" neighborship relation "$\mathsf{Neighbors}$". Taking the rightmost "neighboring" words as our new starting point, we repeat this construction until we are sure after γ steps to re-enter the modulo class that we started with; say we arrive at ladder intervals I_γ and J_γ, respectively. In this way we obtain two "chains" (I_i) and (J_i) of neighboring ladder intervals; (J_i) belongs to the the "true" scenery, and (I_i) belongs to the "reconstructed candidate" w.

Using the Definition of the tests "$\mathsf{Filter}_{2/3}$", and of the events $E_{\text{neighbor I/II}}$, we know that the "estimated" and the "geometrical" neighborship relations coincide at least when taking only rightmost neighbors as above; this holds for both, the "reconstructed" piece w and for the "true" scenery ξ. The distance between I_γ and I equals the distance between J_γ and J, since this distance is not affected by a relative translation between different modulo classes; recall that I_γ and I belong to the same class modulo γ, and so do J_γ and

J. Having identified the starting point and the end point of our two chains of intervals, there also no ambiguity left for the relative position of the intervals in between in the chain; but then we have successfully reconstructed the larger modulo class spanned by the whole chain (I_i).

This construction is repeated recursively until we have correctly reconstructed the whole piece of scenery.

We describe the procedure formally:

Lemma 1.6.14. *Assume that the events $B^n_{\text{unique fit}}$, $E^n_{\text{neighbor I}}$, and $E^n_{\text{neighbor II}}$ hold true. Let $\gamma > 1$ be a divisor of $l_\rightarrow$ and assume that the event $E^n_{\text{mod}\,\gamma\,\text{ok}}$ is valid. Then there is a divisor γ' of $l_\rightarrow$ with $1 \le \gamma' < \gamma$ such that the event $E^n_{\text{mod}\,\gamma'\,\text{ok}}$ is valid, too.*

Proof. Let γ be as in the hypothesis of the lemma. Every modulo class $Z \in \mathbb{Z}/\gamma\mathbb{Z}$ is a union of modulo classes $Z' \in \mathbb{Z}/l_\rightarrow\mathbb{Z}$. Furthermore, every such modulo class $Z' \in \mathbb{Z}/l_\rightarrow\mathbb{Z}$ has a nonempty intersection with $\operatorname{supp}\mu^{*h}$. (One can see this as follows: Since 1 is the greatest common divisor of the elements of $\operatorname{supp}\mu$, every integer can be written in the form $\beta l_\rightarrow + \sum_{j=1}^K s_j$ with $\beta \in \mathbb{Z}$, $K \in \mathbb{N}$, and $s_j \in \operatorname{supp}\mu$ for $1 \le j \le K$. By Lemma 1.6.5 it suffices to take $K = h$; thus we get $\mathbb{Z} = \operatorname{supp}\mu^{*h} + l_\rightarrow\mathbb{Z}$, which is equivalent to the above claim.)

Since we assume $\gamma > 1$, the set difference $\mathbb{Z} \setminus \gamma\mathbb{Z}$ contains at least one $Z \in \mathbb{Z}/\gamma\mathbb{Z}$ as a subset; thus $\mathbb{Z} \setminus \gamma\mathbb{Z}$ has at least one element in common with $\operatorname{supp}\mu^{*h}$. Let $M_\rightarrow := \max[(\mathbb{Z} \setminus \gamma\mathbb{Z}) \cap \operatorname{supp}\mu^{*h}]$ and $M_\leftarrow := -\min[(\mathbb{Z} \setminus \gamma\mathbb{Z}) \cap \operatorname{supp}\mu^{*h}]$. Define γ' to be the greatest common divisor of γ and $M_\rightarrow$; thus $\gamma' < \gamma$ since $M_\rightarrow \notin \gamma\mathbb{Z}$.

Let $w \in \mathsf{SolutionPieces}^n(\mathsf{Input})$. According to Definition (1.6.34) of $E^n_{\text{mod}\,\gamma\,\text{ok}}$ we have to distinguish two cases A and B; however, we treat both cases simultaneously as far as possible. We set

$$\zeta := \begin{cases} \xi\lceil[-9\cdot 2^n, 9\cdot 2^n] & \text{in case A of (1.6.34),} \\ (\xi\lceil[-9\cdot 2^n, 9\cdot 2^n])^{\leftrightarrow} & \text{in case B of (1.6.34).} \end{cases} \tag{1.6.37}$$

For $Z \in \mathbb{Z}/\gamma Z$ we set $\tilde{\iota}_\gamma(Z) := \pm\iota_\gamma(\pm Z)$ with "+" in case A and "−" in case B; here the bijection $\iota_\gamma : \mathbb{Z}/\gamma\mathbb{Z} \to \mathbb{Z}/\gamma\mathbb{Z}$ is taken from Definition (1.6.34) of the event $E^n_{\text{mod}\,\gamma\,\text{ok}}$. The introduction of $\tilde{\iota}_\gamma$ takes care of the inversion of modulo classes in ζ in case B. Since the event $E^n_{\text{mod}\,\gamma\,\text{ok}}$ is valid, we have for all $Z \in \mathbb{Z}/\gamma\mathbb{Z}$:

$$\zeta\lceil(\tilde{\iota}_\gamma(Z) \cap [-2^n, 2^n]) \sqsubseteq w\lceil Z \sqsubseteq \zeta\lceil\tilde{\iota}_\gamma(Z). \tag{1.6.38}$$

For $Z \in \mathbb{Z}/\gamma\mathbb{Z}$, let $T_Z : \mathbb{Z} \to \mathbb{Z}$ denote the translation which transports $w\lceil Z$ to some $T_Z[w\lceil Z] \subseteq \zeta\lceil\tilde{\iota}_\gamma(Z)$; in particular $T_Z[Z] = \tilde{\iota}_\gamma(Z)$. T_Z is uniquely determined, since the event $B^n_{\text{unique fit}}$ holds. Of course, T_Z also depends on γ, but we suppress this in the notation, since γ is considered fixed for the moment. For $W \in \mathbb{Z}/\gamma\mathbb{Z}$, we set $\tilde{T}_W := (T_{\tilde{\iota}_\gamma^{-1}(W)})^{-1}$; thus $\tilde{T}_W[W] = \tilde{\iota}_\gamma^{-1}(W)$. For later use, we note

$$\zeta\lceil(\tilde{\iota}_\gamma(Z) \cap [-2^n, 2^n]) \subseteq T_Z[w\lceil Z] \subseteq \zeta\lceil\tilde{\iota}_\gamma(Z). \tag{1.6.39}$$

We define

$$\zeta' := \bigcup_{Z \in \mathbb{Z}/\gamma\mathbb{Z}} T_Z[w\lceil Z] \subseteq \zeta. \tag{1.6.40}$$

Note that $[-2^n, 2^n] \subseteq \mathrm{domain}(\zeta')$. For (nonempty) ladder intervals I and J, we abbreviate $T_I := T_{I+\gamma\mathbb{Z}}$ and $\tilde{T}_J := \tilde{T}_{J+\gamma\mathbb{Z}}$.

Let the following data be given: $u \in \{w, \zeta'\}$, a right ladder interval I contained in the domain of u with $|I| = c_1 n$, and $k \in [0, \gamma]$. We define $\mathsf{Seq}(I, u, k)$ to denote the set of all $(I_0, \ldots, I_k)$ with the following properties:

1. $I_0 = I$;
2. $I_0, \ldots, I_k$ are right ladder intervals contained in the domain of u with $|I_j| = c_1 n$, $0 \le j \le k$.
3. For all $j \in [0, k[$: $I_j + \gamma\mathbb{Z} \neq I_{j+1} + \gamma\mathbb{Z}$.
4. For all $j \in [0, k[$: $((u\lceil I_j)_\rightarrow, (u\lceil I_{j+1})_\rightarrow) \in \mathsf{Neighbors}^n(\mathsf{Input})$.

Of course $\mathsf{Seq}(I, u, k)$ also depends on γ, Input, and n, but these parameters are considered fixed for the moment.

Let $\mathsf{MaxSeq}(I, u, k)$ denote the set of all $(I_j)_{j=0,\ldots,k} \in \mathsf{Seq}(I, u, k)$ for which $\min I_k - \min I_0$ is maximal.

Given a modulo class $Z \in \mathbb{Z}/\gamma\mathbb{Z}$, we take a fixed right ladder interval $J \subseteq \tilde{\iota}_\gamma^{-1}(Z) \cap [0, (c_1 n + 1) l_\rightarrow] \subseteq \tilde{\iota}_\gamma^{-1}(Z) \cap \mathrm{domain}(\zeta')$, $|J| = c_1 n$. Furthermore, we set $I := \tilde{T}_J J \subseteq Z \cap \mathrm{domain}(w)$.

J serves as a "reference" interval in the "true" (only possibly reflected) piece of scenery ζ', while I serves as a "reference" interval in the "reconstructed" piece of scenery w.

We prove by induction over k:

- $\mathsf{MaxSeq}(I, w, k)$ contains a unique element $(I_j)_{j=0,\ldots,k}$, namely
$$I_j = M_\rightarrow j + (c_1 n - 1) l_\rightarrow j + I. \tag{1.6.41}$$
- $\mathsf{MaxSeq}(J, \zeta', k)$ contains a unique element $(J_j)_{j=0,\ldots,k}$, too, namely
$$J_j = Mj + (c_1 n - 1) l_\rightarrow j + J, \tag{1.6.42}$$
where $M = M_\rightarrow$ in case A and $M = M_\leftarrow$ in case B.

This is obvious for $k = 0$. Here is the induction step $k - 1 \mapsto k$:
If $(I_j)_{j=0,\ldots,k}$, $(J_j)_{j=0,\ldots,k}$ are given by (1.6.41) and (1.6.42), then
$$(I_j)_{j=0,\ldots,k} \in \mathsf{Seq}(I, w, k) \quad \text{and} \quad (J_j)_{j=0,\ldots,k} \in \mathsf{Seq}(J, \zeta', k). \tag{1.6.43}$$
To see this, we check the conditions 1.–4. in the definition of Seq:

1. This is obvious.
2. The only nontrivial claims are $J_j \subseteq \mathrm{domain}(\zeta')$ and $I_j \subseteq \mathrm{domain}(w)$, $0 \le j \le k$. To prove the first claim, we observe $|\min J - \min J_j| \le (M + c_1 n l) k \le 2 c_1 n l \gamma \le 2 c_1 n l^2$; thus we obtain for all $i \in J_j$: $|i| \le 2 c_1 n l^2 + (c_1 n + 1) l_\rightarrow \le 2^n$; hence $J_j \subseteq [-2^n, 2^n] \subseteq \mathrm{domain}(\zeta')$. To prove the second claim, we observe that $J + [-2^n/2, 2^n/2] \subseteq [-2^n, 2^n] = \mathrm{domain}(\zeta')$ (recall $n \ge n_0$, and n_0 is large enough). We apply the translation $\tilde{T}_J$ to $J + ([-2^n/2, 2^n/2] \cap \gamma\mathbb{Z})$ to obtain $I + ([-2^n/2, 2^n/2] \cap \gamma\mathbb{Z}) = \tilde{T}_J[J + ([-2^n/2, 2^n/2] \cap \gamma\mathbb{Z})] \subseteq \mathrm{domain}(w) = [-5 \cdot 2^n, 5 \cdot 2^n]$. This implies $I + [-2^n/2 + \gamma, 2^n/2 - \gamma] \subseteq [-5 \cdot 2^n, 5 \cdot 2^n]$, since $[-5 \cdot 2^n, 5 \cdot 2^n]$ is an interval; consequently $I_j \subseteq I + [-2^n/2 + \gamma, 2^n/2 - \gamma] \subseteq \mathrm{domain}(w)$, which proves the second claim.

3. This is a consequence of $\min I_{j+1} - \max I_j = M_\rightarrow \notin \gamma\mathbb{Z}$ and $\min J_{j+1} - \max J_j = M \notin \gamma\mathbb{Z}$.

4. Because of $\min I_{j+1} - \max I_j = M_\rightarrow \in \operatorname{supp}\mu^{*h}$ we get $I_j \rhd_n I_{j+1}$; thus the fact $w \in \mathsf{Filter}_2^n(\mathsf{Input})$ implies $((w\lceil I_j)_\rightarrow, (w\lceil I_{j+1})_\rightarrow) \in \mathsf{Neighbors}^n(\mathsf{Input})$; see Definition 1.5.4. Similarly $\min J_{j+1} - \max J_j = M_\rightarrow \in \operatorname{supp}\mu^{*h}$ in case A and $\min J_{j+1} - \max J_j = M_\leftarrow \in -\operatorname{supp}\mu^{*h}$ in case B. Hence we get $J_j \rhd_n J_{j+1}$ in case A and $-J_j \lhd_n -J_{j+1}$ in case B; this implies $((\zeta'\lceil J_j)_\rightarrow, (\zeta'\lceil J_{j+1})_\rightarrow) \in \mathsf{Neighbors}^n(\mathsf{Input})$ in both cases, since the event $E^n_{\text{neighbor I}}$ holds; see Definition 1.6.4.

Thus the conditions 1.–4. are indeed valid.

To check the defining property of MaxSeq, consider another sequence

$$(I'_j)_{j=0,\dots,k} \in \mathsf{Seq}(I,w,k) \quad \text{and} \quad (J'_j)_{j=0,\dots,k} \in \mathsf{Seq}(J,\zeta',k). \tag{1.6.44}$$

Using our induction hypotheses

$$\mathsf{MaxSeq}(I,w,k-1) = \{(I_j)_{j=0,\dots,k-1}\}, \tag{1.6.45}$$
$$\mathsf{MaxSeq}(J,\zeta',k-1) = \{(J_j)_{j=0,\dots,k-1}\} \tag{1.6.46}$$

and

$$(I'_j)_{j=0,\dots,k-1} \in \mathsf{Seq}(I,w,k-1), \qquad (J'_j)_{j=0,\dots,k-1} \in \mathsf{Seq}(J,\zeta',k-1), \tag{1.6.47}$$

we know

$$\min I_{k-1} - \min I_0 \geq \min I'_{k-1} - \min I'_0, \tag{1.6.48}$$
$$\min J_{k-1} - \min J_0 \geq \min J'_{k-1} - \min J'_0, \tag{1.6.49}$$

with equality only if $(I'_j)_{j=0,\dots,k-1} = (I_j)_{j=0,\dots,k-1}$ or $(J'_j)_{j=0,\dots,k-1} = (J_j)_{j=0,\dots,k-1}$.

We treat first case of the I's: Using $((w\lceil I'_{k-1})_\rightarrow, (w\lceil I'_k)_\rightarrow) \in \mathsf{Neighbors}^n(\mathsf{Input})$ and $w \in \mathsf{Filter}_3^n(\mathsf{Input})$ we get $I'_{k-1} \rhd_n I'_k + al_\rightarrow$ for some $a \in \mathbb{N}$; thus

$$\min I'_k - \max I'_{k-1} \leq \min I'_k + al_\rightarrow - \max I'_{k-1} \leq M_\rightarrow \tag{1.6.50}$$

by the maximality of $M_\rightarrow$ and $I_k + \gamma\mathbb{Z} \neq I_{k-1} + \gamma\mathbb{Z}$; (see condition 3. in the definition of Seq, and recall $l_\rightarrow \in \gamma\mathbb{Z}$). Hence

$$\begin{aligned} \min I'_k - \min I'_0 &= (\min I'_k - \max I'_{k-1}) + (c_1 n - 1)l_\rightarrow + (\min I'_{k-1} - \min I'_0) \\ &\leq M_\rightarrow + (c_1 n - 1)l_\rightarrow + (\min I_{k-1} - \min I_0) = \min I_k - \min I_0. \end{aligned} \tag{1.6.51}$$

This proves

$$(I_j)_{j=0,\dots,k} \in \mathsf{MaxSeq}(I,w,k). \tag{1.6.52}$$

Furthermore, using our induction hypothesis, equality in (1.6.51) can hold only if $(I'_j)_{j=0,\dots,k-1} \in \mathsf{MaxSeq}(I,w,k-1)$ and $\min I'_k - \max I'_{k-1} = M_\rightarrow$, which is equivalent to $(I'_j)_{j=0,\dots,k} = (I_j)_{j=0,\dots,k}$.

We treat $(J_j)_{j=0,\dots,k}$ similarly: Since the event $E^n_{\text{neighbor II}}$ holds, $((\zeta'\lceil J'_{k-1})_\rightarrow, (\zeta'\lceil J'_k)_\rightarrow) \in \mathsf{Neighbors}^n(\mathsf{Input})$ implies

$$J'_{k-1} \rhd_n J'_k + al_\rightarrow \qquad \text{in case A}, \tag{1.6.53}$$
$$-J'_{k-1} \lhd_n -J'_k - al_\leftarrow \qquad \text{in case B} \tag{1.6.54}$$

for some $a \in \mathbb{N}$; see Definition (1.6.12). This implies in both cases A and B, analogously to (1.6.50):

$$\min J'_k - \max J'_{k-1} \le \min J'_k + a l_\rightarrow - \max J'_{k-1} \le M \tag{1.6.55}$$

by the maximality of M; recall that $M = M_\rightarrow$ in case A and $M = M_\leftarrow$ in case B, and that $l_\rightarrow = l_\leftarrow \in \gamma\mathbb{Z}$ holds in case B; furthermore recall that J'_k and J'_{k-1} belong to different classes modulo γ. We repeat arguments similar to (1.6.51):

$$\begin{aligned} &\min J'_k - \min J'_0 \\ &= (\min J'_k - \max J'_{k-1}) + (c_1 n - 1) l_\rightarrow + (\min J'_{k-1} - \min J'_0) \\ &\le M + (c_1 n - 1) l_\rightarrow + (\min J_{k-1} - \min J_0) = \min J_k - \min J_0, \end{aligned} \tag{1.6.56}$$

with equality only if $(J'_k)_{j=0,\ldots,k} \in \mathsf{MaxSeq}(J, \zeta', k-1)$ and $\min J'_k - \max J'_{k-1} = M$. This proves in analogy to (1.6.52):

$$\mathsf{MaxSeq}(J, \zeta', k) = \{(J_j)_{j=0,\ldots,k}\}. \tag{1.6.57}$$

Since $\tilde{\iota}_\gamma$ is bijective, the facts $T_{I_j} I_j \subseteq \mathrm{domain}(\zeta')$, $(\zeta' \lceil T_{I_j} I_j)_\rightarrow = (w \lceil I_j)_\rightarrow$, and $(I_j)_j \in \mathsf{Seq}(I, w, k)$ imply

$$(T_{I_j} I_j)_j \in \mathsf{Seq}(J, \zeta', k). \tag{1.6.58}$$

Similarly, $\tilde{T}_{J_j} J_j \subseteq \mathrm{domain}(w)$, $(w \lceil \tilde{T}_{J_j} J_j)_\rightarrow = (\zeta' \lceil J_j)_\rightarrow$, and $(J_j)_j \in \mathsf{Seq}(J, \zeta', k)$ imply

$$(\tilde{T}_{J_j} J_j)_j \in \mathsf{Seq}(I, w, k). \tag{1.6.59}$$

Now we set $k = \gamma$. Observe that $I_\gamma + \gamma\mathbb{Z} = I_0 + \gamma\mathbb{Z}$ and $J_\gamma + \gamma\mathbb{Z} = J_0 + \gamma\mathbb{Z}$; hence $T_{I_0} = T_{I_\gamma}$ and $\tilde{T}_{J_0} = \tilde{T}_{J_\gamma}$. Thus, using (1.6.52), (1.6.57), (1.6.58), (1.6.59), and the defining property of MaxSeq, we obtain

$$\begin{aligned} \min I_\gamma - \min I_0 &= \min T_{I_\gamma} I_\gamma - \min T_{I_0} I_0 \\ &\le \min J_\gamma - \min J_0 \\ &= \min \tilde{T}_{J_\gamma} J_\gamma - \min \tilde{T}_{J_0} J_0 \\ &\le \min I_\gamma - \min I_0. \end{aligned} \tag{1.6.60}$$

Since the first and last term in (1.6.60) are identical, equality holds everywhere in (1.6.60). Hence, using (1.6.57), (1.6.58), and the defining property of MaxSeq again, we see

$$(T_{I_j} I_j)_j \in \mathsf{MaxSeq}(J, \zeta', \gamma) \tag{1.6.61}$$

and thus $(T_{I_j} I_j)_j = (J_j)_j$, since $\mathsf{MaxSeq}(J, \zeta', \gamma)$ is a singleton. Furthermore the facts (1.6.41), (1.6.42), $\gamma \neq 0$, and $T_{I_0} = T_{I_\gamma}$ imply $M_\rightarrow = M$, since

$$0 = (\min I_\gamma - \min I_0) - (\min J_\gamma - \min J_0) = M_\rightarrow \gamma - M\gamma. \tag{1.6.62}$$

A side remark: consequently case B cannot occur whenever $M_\rightarrow \neq M_\leftarrow$. Using (1.6.41) and (1.6.42) again, we see that all translations T_{I_j}, $j = 0, \ldots, \gamma$, coincide: $T_{I_j} = T_I$. We observe

$$(I_0 \cup \ldots \cup I_\gamma) + \gamma\mathbb{Z} = I + \{j M_\rightarrow \mid j = 0, \ldots, \gamma\} + \gamma\mathbb{Z} = I + \gamma'\mathbb{Z}; \tag{1.6.63}$$

recall that γ' was defined to be the greatest common divisor of $M_\to$ and γ. Thus we have shown: the translations T_Z, $Z \in \mathbb{Z}/\gamma\mathbb{Z}$, depend only on the rougher modulo class $Z' = Z + \gamma'\mathbb{Z} \in \mathbb{Z}/\gamma'\mathbb{Z}$; hence $T_{Z+\gamma'\mathbb{Z}} := T_Z$ and $\iota_{\gamma'} : \mathbb{Z}/\gamma'\mathbb{Z} \to \mathbb{Z}/\gamma'\mathbb{Z}$, $\iota_{\gamma'}(Z') := \bigcup_{Z\subseteq Z', Z\in\mathbb{Z}/\gamma\mathbb{Z}} \iota_\gamma(Z)$ are well-defined. Since $\iota_\gamma : \mathbb{Z}/\gamma\mathbb{Z} \to \mathbb{Z}/\gamma\mathbb{Z}$ is a bijection, $\iota_{\gamma'}$ is a bijection, too. In analogy to $\tilde\iota_\gamma$, we introduce $\tilde\iota_{\gamma'}(Z') := \pm\iota_{\gamma'}(\pm Z')$, ("+" in case A, "−" in case B). As a consequence of (1.6.39), we obtain for all $Z' \in \mathbb{Z}/\gamma'\mathbb{Z}$:

$$\zeta\lceil(\tilde\iota_{\gamma'}(Z') \cap [-2^n, 2^n]) \subseteq T_{Z'}[w\lceil Z'] \subseteq \zeta\lceil\tilde\iota_{\gamma'}(Z'). \tag{1.6.64}$$

Hence the event $E^n_{\mathrm{mod}\,\gamma'\,\mathrm{ok}}$ is valid. This finishes the proof of Lemma 1.6.14. ■

Lemma 1.6.15. $E^n_{\mathrm{mod}\,l_\to\,\mathrm{ok}} \cap B^n_{\mathrm{unique\,fit}} \cap E^n_{\mathrm{neighbor\,I}} \cap E^n_{\mathrm{neighbor\,II}} \subseteq E^n_{\mathrm{mod}\,1\,\mathrm{ok}}$.

Proof. Assume that $E^n_{\mathrm{mod}\,l_\to\,\mathrm{ok}} \cap B^n_{\mathrm{unique\,fit}} \cap E^n_{\mathrm{neighbor\,I}} \cap E^n_{\mathrm{neighbor\,II}}$ holds. Let Γ denote the (random) set of all divisors $\gamma \geq 1$ of $l_\to$ for which the event $E^n_{\mathrm{mod}\,\gamma\,\mathrm{ok}}$ is valid. $\Gamma \neq \varnothing$, since $E^n_{\mathrm{mod}\,l_\to\,\mathrm{ok}}$ holds. The smallest element of Γ cannot be bigger than 1 by Lemma 1.6.14; thus it must be equal to 1. This means that $E^n_{\mathrm{mod}\,1\,\mathrm{ok}}$ holds. ■

Lemma 1.6.16. *For $\gamma = 1$, we have $E^n_{\mathrm{mod}\,1\,\mathrm{ok}} \subseteq E^n_{\mathrm{all\,pieces\,ok}}$.*

Proof. This is obvious, since there is only the trivial "modulo class" $Z = \iota_1(Z) = \mathbb{Z}$ remaining for $\gamma = 1$: In case A, one has $\xi\lceil[-2^n, 2^n] \sqsubseteq w \sqsubseteq \xi\lceil[-9\cdot 2^n, 9\cdot 2^n]$, and in case B, one has $\xi\lceil[-2^n, 2^n] \sqsubseteq w^\leftrightarrow \sqsubseteq \xi\lceil[-9\cdot 2^n, 9\cdot 2^n]$. ■

Theorem 1.6.3. $B^n_{\mathrm{seed\,I}} \cap B^n_{\mathrm{unique\,fit}} \cap B^n_{\mathrm{all\,paths}} \cap B^n_{\mathrm{outside\,out}} \cap B^n_{\mathrm{recogn\,straight}} \cap B^n_{\mathrm{signals}} \cap E^m_{\mathrm{stop},\tau} \subseteq E^n_{\mathrm{all\,pieces\,ok}}$

Proof. We collect the results of Lemmas 1.6.2, 1.6.4, 1.6.6, 1.6.9, 1.6.10, 1.6.11, 1.6.12, 1.6.13, 1.6.15, and 1.6.16 in the following list:

$$\begin{array}{lcl}
B^n_{\mathrm{signals}} & \subseteq & E^n_{\mathrm{signals\,II}}, \\
B^n_{\mathrm{all\,paths}} & \subseteq & E^n_{\mathrm{neighbor\,I}}, \\
E^n_{\mathrm{signals\,II}} \cap E^m_{\mathrm{stop},\tau} & \subseteq & E^n_{\mathrm{neighbor\,II}}, \\
B^n_{\mathrm{all\,paths}} \cap B^n_{\mathrm{recogn\,straight}} & \subseteq & E^n_{\mathrm{only\,ladder}}, \\
B^n_{\mathrm{outside\,out}} \cap B^n_{\mathrm{unique\,fit}} \cap E^n_{\mathrm{only\,ladder}} \cap E^m_{\mathrm{stop},\tau} & \subseteq & E^n_{\mathrm{mod\,class}}, \\
B^n_{\mathrm{unique\,fit}} \cap B^n_{\mathrm{signals}} \cap B^n_{\mathrm{all\,paths}} \cap B^n_{\mathrm{recogn\,straight}} \cap E^m_{\mathrm{stop},\tau} & \subseteq & E^n_{\mathrm{seed\,II}}, \\
B^n_{\mathrm{signals}} \cap E^m_{\mathrm{stop},\tau} & \subseteq & E^n_{\mathrm{dist}}, \\
E^n_{\mathrm{seed\,II}} \cap E^n_{\mathrm{mod\,class}} \cap E^n_{\mathrm{dist}} \cap B^n_{\mathrm{unique\,fit}} & \subseteq & E^n_{\mathrm{mod}\,l_\to\,\mathrm{ok}}, \\
E^n_{\mathrm{mod}\,l_\to\,\mathrm{ok}} \cap B^n_{\mathrm{unique\,fit}} \cap E^n_{\mathrm{neighbor\,I}} \cap E^n_{\mathrm{neighbor\,II}} & \subseteq & E^n_{\mathrm{mod\,1\,ok}}, \\
E^n_{\mathrm{mod\,1\,ok}} & \subseteq & E^n_{\mathrm{all\,pieces\,ok}}.
\end{array}$$

The claim of the theorem is a simple combination of these inclusions. ■

1.6.3 Probabilistic estimates for basic events

In this subsection we show that the "basic events" $B_{\ldots}$ occur very probably. Together with the result of the previous subsections this shows that the partial reconstruction algorithms Alg^n yield with high probability a correctly reconstructed piece of scenery.

We start with an elementary auxiliary lemma:

Lemma 1.6.17. *Let $f : I_0 \to J$ be a finite injection without fixed points. Then there is $I' \subseteq I_0$ with $|I'| \geq |I_0|/3$ and $f[I'] \cap I' = \varnothing$.*

Proof. We construct recursively finite sequences (I_k) and (I'_k), for $k - 1 < |I_0|/3$, of subsets of I_0. The "loop invariants" of the recursion are: $f[I_k] \cap I'_k = \varnothing$, $f[I'_k] \cap I_k = \varnothing$, $f[I'_k] \cap I'_k = \varnothing$, $I_k \cap I'_k = \varnothing$, $|I'_k| = k$, and $|I_k| \geq |I_0| - 3k$.

The recursion starts with the given I_0 and with $I'_0 = \varnothing$. In the $(k+1)$st step, $k < |I_0|/3$, we choose any point $x \in I_k$, and define $I'_{k+1} := I'_k \cup \{x\}$. If $f^{-1}(x)$ exists, then we set $I_{k+1} := I_k \setminus \{x, f(x), f^{-1}(x)\}$; else we set $I_{k+1} := I_k \setminus \{x, f(x)\}$.

Note that the validity of the above "loop invariants" is indeed preserved by the recursion; the fact $f(x) \neq x$ is used for the third loop invariant.
Finally we set $I' := I'_k$ for $k := \min\{j \in \mathbb{N} \mid 3j \geq |I_0|\}$; then $I' \subseteq I_0$ is well-defined and fulfills the claims in Lemma 1.6.17. ■

Lemma 1.6.18. *There exists constants $c_{18}, c_{19} > 0$ not depending on n such that:*

$$P\left[(B^n_{\text{unique fit}})^{\text{c}}\right] \leq c_{18} e^{-c_{19} n}. \tag{1.6.65}$$

Proof. Let $i, j \in \{1, \ldots, l^2\}$, and let $I \subseteq [-11 \cdot 2^n, 11 \cdot 2^n]$ be a i-spaced interval, and $J \subseteq [-11 \cdot 2^n, 11 \cdot 2^n]$ be a j-spaced interval with $|I| = |J| \geq 1$. Let $f : I \to J$ be a monotonically increasing or decreasing bijection, but not the identity map; thus the case $I = J$ can only occur if f is decreasing.

We claim: For some constants $c_{12} > 0$ and $c_{13} > 2 \log 2 / c_2$ (not depending on i, j, I, or J) we have

$$P[\xi \circ f = \xi\lceil I] \leq c_{12} e^{-c_{13}|I|}. \tag{1.6.66}$$

Note that $\xi \circ f = \xi\lceil I$ is equivalent to $(\xi\lceil J)_{\to} = (\xi\lceil I)_{\to}$ if f is increasing, and it is equivalent to $(\xi\lceil J)_{\to} = (\xi\lceil I)_{\leftarrow}$ if f is decreasing.

Before proving (1.6.66), let us show how it implies (1.6.65): There are at most l^2 choices for (i, j), and given (i, j), there are at most $(22 \cdot 2^n + 1)^2 \leq 500 \cdot 2^{2n}$ choices for (I, J) with $|I| = |J| = c_2 n$; finally there is one binary choice: f is increasing or decreasing. If $\xi \circ f \neq \xi\lceil I$ holds for all of these choices (with the trivial exception $I = J$ and $f = \mathrm{id}$), then the event $B^n_{\text{unique fit}}$ is valid; note that it suffices to consider $|I| = |J| = c_2 n$ instead of $|I| = |J| \geq c_2 n$, since it suffices to consider subintervals of I, J consisting only of $c_2 n$ points. Hence (1.6.66) implies (1.6.65):

$$P\left[(B^n_{\text{unique fit}})^{\text{c}}\right] \leq l^2 \cdot 500 \cdot 2^{2n} \cdot 2 \cdot c_{12} e^{-c_{13} c_2 n} = c_{18} e^{-c_{19} n}, \tag{1.6.67}$$

where $c_{18} := 1000 l^2 c_{12}$ and $c_{19} := c_{13} c_2 - 2 \log 2 > 0$.

We prove (1.6.66) next: unless f is the identity map, it can have at most a single fixed point, since it is the restriction of some affine-linear map to the ladder interval I. Remove this fixed point from I, if it exists; call I_0 the set of all remaining points. By Lemma 1.6.17, there is $I' \subseteq I_0$ with $|I'| \geq |I_0|/3 \geq (|I| - 1)/3$ and $f[I'] \cap I' = \varnothing$. Hence $\xi\lceil f[I']$ and $\xi\lceil I'$ are independent random pieces of scenery; thus

$$P\left[\xi \circ f = \xi\lceil I\right] \leq P\left[\xi \circ f\lceil I' = \xi\lceil I'\right] = |\mathcal{C}|^{-|I'|} \leq |\mathcal{C}|^{-(|I|-1)/3}; \tag{1.6.68}$$

thus (1.6.66) follows with $c_{13} := (\log |\mathcal{C}|)/3$ and $c_{12} := |\mathcal{C}|^{1/3}$. Note that $c_{13} c_2 - 2 \log 2 > 0$ since c_2 was required to be large enough; recall subsection 1.2.1. ■

Lemma 1.6.19. *There exist constants $c_{20}, c_{21} > 0$ not depending on n such that:*

$$P\left[(B^n_{\text{all paths}})^{\mathrm{c}} \cap E^m_{\text{stop},\tau}\right] \le c_{21} e^{-c_{20} n}. \tag{1.6.69}$$

Proof of Lemma 1.6.19. Let $k < 2^{\alpha n}$ and $R \in \mathsf{AdPath}(12 \cdot 2^n, 3c_1 n)$. We set

$$B^{n,k}_R := \left\{\exists j \in [0, 2^{2n}] : \mathsf{TimeShift}^{\tau(k)+j}(R) \subseteq S\right\}, \tag{1.6.70}$$

$$E^m_{\text{stop},\tau,k} := \left\{ \begin{array}{l} \tau_k(\chi) < 2^{12\alpha n_m},\ |S(\tau_k(\chi))| \le 2^{n_m}, \\ \tau_j(\chi) + 2 \cdot 2^{2n_m} \le \tau_k(\chi) \text{ for } j < k \end{array} \right\}, \tag{1.6.71}$$

$$A^{n,k}_R := E^m_{\text{stop},\tau,k} \setminus B^{n,k}_R. \tag{1.6.72}$$

Note that $B^n_{\text{all paths}} = \bigcap_{R \in \mathsf{AdPath}(12\cdot 2^n, 3c_1 n)} \bigcup_{k=0}^{2^{\alpha n}-1} B^{n,k}_R$ and $E^m_{\text{stop},\tau} \subseteq E^m_{\text{stop},\tau,k}$ for $k \le 2^{\alpha n}$, and thus

$$E^m_{\text{stop},\tau} \setminus B^n_{\text{all paths}} \subseteq \bigcup_{R \in \mathsf{AdPath}(12\cdot 2^n, 3c_1 n)} \bigcap_{k=0}^{2^{\alpha n}-1} A^{n,k}_R. \tag{1.6.73}$$

In the following, R runs over the set $\mathsf{AdPath}(12 \cdot 2^n, 3c_1 n)$:

$$P\left[(B^n_{\text{all paths}})^{\mathrm{c}} \cap E^m_{\text{stop},\tau}\right] \le |\mathsf{AdPath}(12 \cdot 2^n, 3c_1 n)| \max_R P\left[\bigcap_{k=0}^{2^{\alpha n}-1} A^{n,k}_R\right], \tag{1.6.74}$$

$$|\mathsf{AdPath}(12 \cdot 2^n, 3c_1 n)| \le 25 \cdot 2^n |\mathcal{M}|^{3c_1 n}, \tag{1.6.75}$$

$$P\left[\bigcap_{k=0}^{2^{\alpha n}-1} A^{n,k}_R\right] = \prod_{k=0}^{2^{\alpha n}-1} P\left[A^{n,k}_R \,\middle|\, \bigcap_{j<k} A^{n,j}_R\right], \tag{1.6.76}$$

$$P\left[A^{n,k}_R \,\middle|\, \bigcap_{j<k} A^{n,j}_R\right] \le P\left[(B^{n,k}_R)^{\mathrm{c}} \,\middle|\, E^m_{\text{stop},\tau,k} \cap \bigcap_{j<k} A^{n,j}_R\right]; \tag{1.6.77}$$

the last statement follows from the elementary fact $P[A \cap B|C] \le P[A|B \cap C]$. Since $2^{2n} + 3c_1 n \le 2 \cdot 2^{2n}$, we have $C^{n,k}_R := E^m_{\text{stop},\tau,k} \cap \bigcap_{j<k} A^{n,j}_R \in \mathcal{F}_{\tau_k}$, i.e. one can decide whether the event $C^{n,k}_R$ holds by observing ξ and $S(0), \ldots, S(\tau_k)$. Furthermore, if $C^{n,k}_R$ holds, then $|S(\tau_k(\chi))| \le 2^n$, and as a consequence of the local Central Limit Theorem [5], Theorem 5.2 (page 132) we get: there is a constant $c_{22} > 0$ such that for all x, y with $|x| \le 12 \cdot 2^n$ and $|y| \le 2^n$: $P[y + S(j) = x \text{ for some } j \in [0, 2^{2n}]] \ge c_{22} 2^{-n}$; note that $y + S$ is a random walk starting in the point y. Note that we do not need the random walk to be aperiodic; it suffices that it can reach every integer, i.e. that the greatest common divisor of the elements of $|\mathcal{M}|$ is 1. Thus by the strong Markov property:

$$\inf_{|x| \le 12\cdot 2^n} P\left[S(\tau(k) + j) = x \text{ for some } j \in [0, 2^{2n}] \;\middle|\; C^{n,k}_R\right] \ge c_{22} 2^{-n}. \tag{1.6.78}$$

Once it is in the starting point x, the probability that S follows an admissible path $R \in \mathsf{AdPath}(12 \cdot 2^n, 3c_1 n)$ for the next $3c_1 n - 1$ steps is bounded from below by $\mu^{3c_1 n}_{\min}$. Here $\mu_{\min} := \min\{\mu(\{x\}) \,|\, x \in \mathcal{M}\}$ is the smallest positive probability for a jump. Therefore, using the strong Markov property again:

$$P\left[B^{n,k}_R \;\middle|\; C^{n,k}_R\right] \ge c_{22} 2^{-n} \mu^{3c_1 n}_{\min}. \tag{1.6.79}$$

We combine (1.6.74)–(1.6.77) and (1.6.79) to obtain

$$\begin{aligned} P\left[E^m_{\mathrm{stop},\tau} \setminus B^n_{\mathrm{all\,paths}}\right] &\leq 25 \cdot 2^n |\mathcal{M}|^{3c_1 n} (1 - c_{22} 2^{-n} \mu_{\min}^{3c_1 n})^{2^{\alpha n}} \\ &\leq 25 \cdot 2^n |\mathcal{M}|^{3c_1 n} \exp\left\{-c_{22} 2^{-n} \mu_{\min}^{3c_1 n} 2^{\alpha n}\right\} \\ &\leq 25 \exp\left\{n(\log 2 + 3c_1 \log |\mathcal{M}|) - c_{22} e^{n(\alpha \log 2 + 3c_1 \log \mu_{\min} - \log 2)}\right\}. \end{aligned} \tag{1.6.80}$$

Now $\alpha > 1 - 3c_1 \log_2 \mu_{\min}$ by our choice of α in subsection 1.2.1; thus the right hand side of the last inequality converges to 0 superexponentially fast as $n \to \infty$. Note that we may choose an upper bound $c_{21} e^{-c_{20} n}$ for the right hand side in 1.6.80, where neither c_{21} nor c_{20} depend on α or c_1. This is true since $n \geq n_0$, and n_0 was chosen large enough, depending on c_1 and α; recall subsection 1.2.1. This proves the lemma. ■

Lemma 1.6.20. *There exists a constant $c_{23} > 0$ not depending on n such that:*

$$P\left[(B^n_{\mathrm{outside\,out}})^{\mathrm{c}}\right] \leq 160 e^{-c_{23} n}. \tag{1.6.81}$$

Proof of Lemma 1.6.20. The set $[-2 \cdot l 2^{2n}, 2 \cdot l 2^{2n}] \setminus [-10 \cdot 2^n, 10 \cdot 2^n]$ contains less than $4 \cdot l \cdot 2^{2n}$ points, and for every fixed starting point the number of admissible paths with $c_1 n/2$ points is equal to $|\mathcal{M}|^{c_1 n/2 - 1}$. Hence there are less than $4 \cdot l 2^{2n} |\mathcal{M}|^{c_1 n/2}$ paths $R \in \mathsf{AdPaths}(2 \cdot l 2^{2n}, c_1 n/2)$ with $R(i) \notin [-10 \cdot 2^n, 10 \cdot 2^n]$ for all $i = 0, \ldots, c_1 n/2$. On the other hand, there are less than $40 \cdot 2^n$ ladder words of length $c_1 n/2$ in $[-9 \cdot 2^n, 9 \cdot 2^n]$. The colors $\xi \circ R$ that a path in $R \in \mathsf{AdPaths}(2^{2n}, c_1 n/2)$ with $R(i) \notin [-10 \cdot 2^n, 10 \cdot 2^n]$ for all $i = 0, \ldots, c_1 n/2 - 1$ reads are independent of the colors inside $[-9 \cdot 2^n, 9 \cdot 2^n]$. Thus the probability that a given path $R \in \mathsf{AdPaths}(2^{2n}, c_1 n/2)$ with $R(i) \notin [-10 \cdot 2^n, 10 \cdot 2^n]$ for all $i = 0, \ldots, c_1 n/2 - 1$ reads the same colors as a fixed ladder word in $[-9 \cdot 2^n, 9 \cdot 2^n]$ is $|\mathcal{C}|^{-c_1 n/2}$. Thus

$$P\left[(B^n_{\mathrm{outside\,out}})^{\mathrm{c}}\right] \leq 160 l 2^{3n} |\mathcal{M}|^{c_1 n/2} |\mathcal{C}|^{-c_1 n/2}. \tag{1.6.82}$$

Since $|\mathcal{M}| < |\mathcal{C}|$, the last expression becomes exponentially decreasing in n since $c_1 > 6/\log \frac{|\mathcal{C}|}{|\mathcal{M}|}$ since c_1 was chosen large enough; see subsection 1.2.1. This proves the lemma. ■

We prepare the treatment of the event $B^n_{\mathrm{recognstraight}}$ by the following combinatoric lemma:

Lemma 1.6.21. *Let $c_{24} := 1/(2|\mathcal{M}|(l_\rightarrow + l_\leftarrow))$. There are two intervals $I_1, I_2 \subseteq [0, c_1 n[$ with $|I_1| = |I_2| \geq c_{24} c_1 n - 1$ such that the following statement is valid: For all $R \in \mathsf{AdPaths}(11 \cdot 2^n, c_1 n)$ with $R(c_1 n - 1) - R(0) \notin \{(c_1 n - 1) l_\rightarrow, -(c_1 n - 1) l_\leftarrow\}$, there is $I \in \{I_1, I_2\}$ and an admissible path $\bar{R} \in \mathsf{AdPaths}(12 \cdot 2^n, c_1 n)$ with the following properties:*

- $R(0) = \bar{R}(0), \quad R(c_1 n - 1) = \bar{R}(c_1 n - 1)$.
- *At least one of the following holds:*
  1. *for all $(i, j) \in I \times I$ with $j < i$: $\bar{R}(i) \notin \{R(j), \bar{R}(j)\}$;*
  2. *for all $(i, j) \in I \times I$ with $i < j$: $\bar{R}(i) \notin \{R(j), \bar{R}(j)\}$.*

Proof. We define $k := \lfloor c_{24} c_1 n \rfloor$, $I' := [1, 2k] \subseteq [0, c_1 n[$, $I_1 := [1, k]$, and $I_2 := [k + 1, 2k]$. We observe $|I_1|, |I_2| \geq c_{24} c_1 n - 1$ and $I_1, I_2 \subseteq [0, c_1 n[$.

Let $R \in \mathsf{AdPaths}(11 \cdot 2^n, c_1 n)$ be not a ladder path. We show first: There are $R', R'' \in \mathsf{AdPaths}(12 \cdot 2^n, c_1 n)$ such that $R'(0) = R''(0) = R(0)$, $R'(c_1 n - 1) = R''(c_1 n - 1) =$

$R(c_1 n - 1)$, $R'\lceil I'$ and $R''\lceil I'$ are ladder paths, and $R''\lceil I' = r + R'\lceil I'$ for some $r \neq 0$, i.e. $R''\lceil I'$ is obtained from $R'\lceil I'$ by a spatial translation.

To prove this claim, let $d = (d_i)_{i=1\ldots,c_1 n-1} \in \mathcal{M}^{c_1 n-1}$, $d_i := R(i) - R(i-1)$, be the jump sizes in R. Every other $\tilde{d} \in \mathcal{M}^{c_1 n-1}$ with $\sum_{i=1}^{c_1 n-1} \tilde{d}_i = \sum_{i=1}^{c_1 n-1} d_i$ gives rise to an admissible path $\tilde{R} \in \mathsf{AdPaths}(12 \cdot 2^n, c_1 n)$, too, with $\tilde{R}(0) = R(0)$, $\tilde{R}(c_1 n - 1) = R(c_1 n - 1)$, and with jump sizes $\tilde{d}_i = \tilde{R}(i) - \tilde{R}(i-1)$; namely $\tilde{R}(i) := R(0) + \sum_{j=1}^{i} \tilde{d}_j$. Since $\tilde{R}$ has its starting point and end point in $[-11 \cdot 2^n, 11 \cdot 2^n]$ and since $c_1 nl < 2^n$, the path $\tilde{R}$ can indeed not leave the range $[-12 \cdot 2^n, 12 \cdot 2^n]$.

There are at most $|\mathcal{M}|$ possible values for d_i, but there are $c_1 n$ possible indices i; thus at least one value $a \in \mathcal{M}$ occurs in the d_i at least $c_1 n/|\mathcal{M}|$ times. We choose $2k(a + l_{\leftarrow}) \geq 0$ indices i with $d_i = a$ and replace them by $l_{\rightarrow}$, and we choose $2k(l_{\rightarrow} - a) \geq 0$ different indices i with $d_i = a$ and replace them by $-l_{\leftarrow}$; note that $2k(a + l_{\leftarrow}) + 2k(l_{\rightarrow} - a) = 2k(l_{\rightarrow} + l_{\leftarrow}) \leq c_1 n/|\mathcal{M}|$. We end up with a new vector $\tilde{d} \in \mathcal{M}^{c_1 n-1}$ with $\sum_{i=1}^{c_1 n-1} \tilde{d}_i = \sum_{i=1}^{c_1 n-1} d_i$, since $2k(l_{\leftarrow} + l_{\rightarrow})a = 2k(a + l_{\leftarrow})l_{\rightarrow} + 2k(l_{\rightarrow} - a)(-l_{\leftarrow})$. $\tilde{d}$ contains at least $2k$ entries with value $l_{\rightarrow}$, or it contains at least $2k$ entries with value $-l_{\leftarrow}$, since already the described replacement procedure has produced sufficiently many such entries. However, not all entries of $\tilde{d}$ can equal $l_{\rightarrow}$; similarly not all its entries can equal $-l_{\leftarrow}$, since R is not a ladder path. We permute the entries of $\tilde{d}_i$ in two different ways; the resulting vectors are called d' and d'': First to obtain d', permute the entries in $\tilde{d}$ such that the first $2k$ permuted entries d'_i, $i = 1, \ldots, 2k$ either all equal $l_{\rightarrow}$ or all equal $-l_{\leftarrow}$; the order of the remaining entries is irrelevant. Second to obtain d'', transpose the first entry d'_1 with a different entry $d'_i \neq d'_1$. Let R' and R'' be admissible pieces of paths with $R'(0) = R''(0) = R(0)$ and step sizes $d'_i = R'(i) - R'(i-1)$ and $d''_i = R''(i) - R''(i-1)$, respectively. Recall $I' = [1, 2k]$; then $R'\lceil I'$ and $R''\lceil I'$ are ladder paths, and $R''\lceil I'$ is obtained from translating $R'\lceil I'$ by $r := d''_1 - d'_1 \neq 0$. Thus our first claim holds.

$R'\lceil I'$ is a right ladder path or a left ladder path. Without loss of generality, we assume that it is a right ladder path; the case of left ladder paths can be treated similarly by reversing directions in the arguments below. Furthermore, we assume without loss of generality $r > 0$; otherwise we exchange R' with R''.

We are ready to prove the claim of the lemma; recall that k is a point in the middle of I'. There are two cases:

- If $R(k) > R'(k)$, then we take $I := I_1$ and $\bar{R} := R'$. Since $R'\lceil I$ is a right ladder path, it moves with maximal speed $l_{\rightarrow}$ to the right. R cannot move faster than that to the right; thus $R(j) > R'(i)$ and $R'(j) > R'(i)$ for all $i, j \in I$ with $i < j$.

- If $R(k) \leq R'(k)$, then $R(k) < r + R'(k) = R''(k)$; this time we take $I := I_2$ and $\bar{R} := R''$. The same argument as above yields $R(j) < R''(i)$ and $R''(j) < R''(i)$ for all $i, j \in I$ with $j < i$.

This proves Lemma 1.6.21. ■

Lemma 1.6.22. *There exist positive constants c_{25} and c_{26} not depending on n such that:*

$$P\left[(B^n_{\text{recogn straight}})^{\mathrm{c}}\right] \leq c_{25} e^{-c_{26} n}. \tag{1.6.83}$$

Proof of Lemma 1.6.22. Given $R \in \mathsf{AdPaths}(11 \cdot 2^n, c_1 n)$ with $R(c_1 n) - R(0) \notin \{(c_1 n - 1)l_{\rightarrow}, -(c_1 n - 1)l_{\leftarrow}\}$, we take $I = I(R) \subseteq [0, c_1 n[$ and $\bar{R} \in \mathsf{AdPaths}(12 \cdot 2^n, c_1 n)$ as

in Lemma 1.6.21. Without loss of generality assume that condition 1. in Lemma 1.6.21 is satisfied. We prove for all $I' \subseteq I$ by induction on $|I'|$:

$$P[(\xi \circ R)\lceil I' = (\xi \circ \bar{R})\lceil I'] = |\mathcal{C}|^{-|I'|}. \tag{1.6.84}$$

This is obvious for $I' = \varnothing$. For other I', let $I'' := I' \setminus \max I'$. Then $\xi(\bar{R}(\max I'))$ is independent of $(\xi \circ R\lceil I'', \xi \circ \bar{R}\lceil I'')$, since they are generated by disjoint parts of the scenery. Thus

$$\begin{aligned} &P[(\xi \circ R)\lceil I' = (\xi \circ \bar{R})\lceil I'] \\ &= P[\xi(R(\max I')) = \xi(\bar{R}(\max I'))] \ \cdot \ P[(\xi \circ R)\lceil I'' = (\xi \circ \bar{R})\lceil I''] \\ &= |\mathcal{C}|^{-1}|\mathcal{C}|^{-|I''|} = |\mathcal{C}|^{-|I'|}. \end{aligned} \tag{1.6.85}$$

By taking $I' = I$, we conclude $P[(\xi \circ R)\lceil I = (\xi \circ \bar{R})\lceil I] \le |\mathcal{C}|^{-|I|}$.

Unfortunately, it does not suffice to multiply the last bound with the bound $23 \cdot 2^n|\mathcal{M}|^{c_1 n} \ge |\mathsf{AdPaths}(11 \cdot 2^n, c_1 n)|$: the product may sometimes be bigger than 1.

To overcome this difficulty, we partition $\mathsf{AdPaths}(12 \cdot 2^n, c_1 n) \ni R$ into equivalence classes $[R]$: we put two paths into the same class if and only if they are mapped to the same value by the map $R \mapsto (R(0), R(c_1 n - 1), I(R), R\lceil I(R))$; here $I(R) \in \{I_1, I_2\}$ is taken from Lemma 1.6.21. We bound the number of equivalence classes from above: For our purposes, a simple but rough bound suffices: There are at most $25 \cdot 2^n$ choices for each of $R(0)$, $R(c_1 n - 1)$, and $R(\min I(R))$, and there is a binary choice $I(R) \in \{I_1, I_2\}$; finally given $R(\min I(R))$, there are not more than $|\mathcal{M}|^k$ choices for $R\lceil I(R)$, where again $k = \lfloor c_{24} c_1 n \rfloor = |I(R)|$. Altogether the number of equivalence classes is bounded by $c_{27} 2^{3n} |\mathcal{M}|^k$, where $c_{27} := 2 \cdot 25^3$. We may choose a map $\mathsf{AdPaths}(11 \cdot 2^n, c_1 n) \to \mathsf{AdPaths}(12 \cdot 2^n, c_1 n)$, $R \mapsto \bar{R}$ such that $\bar{R}$ depends only of the equivalence class $[R]$ and fulfills the claim in Lemma 1.6.21. We get

$$\begin{aligned} &P[(B^n_{\text{recogn straight}})^{\mathrm{c}}] \le P[\exists R \in \mathsf{AdPaths}(11 \cdot 2^n, c_1 n) : \ \xi \circ R\lceil I(R) = \xi \circ \bar{R}\lceil I(R)] \\ &\le \sum_{[R]} P[\xi \circ R\lceil I(R) = \xi \circ \bar{R}\lceil I(R)] \\ &\le c_{27} 2^{3n} (|\mathcal{M}|/|\mathcal{C}|)^k \le c_{27} (|\mathcal{C}|/|\mathcal{M}|) \exp\{(3 \log 2 - c_{24} c_1 \log(|\mathcal{M}|/|\mathcal{C}|))n\}. \end{aligned} \tag{1.6.86}$$

We emphasize: the sum in the last but one expression runs over equivalence classes $[R]$, not over paths R; the event $\{\xi \circ R\lceil I(R) = \xi \circ \bar{R}\lceil I(R)\}$ does not depend on the choice of $R \in [R]$. We have $c_{24} c_1 \log(|\mathcal{M}|/|\mathcal{C}|) - 3 \log 2 \ge 1$; recall from subsection 1.2.1 that c_1 is large enough. The estimate (1.6.86) proves the lemma with $c_{26} = 1$, $c_{25} = c_{27}|\mathcal{C}|/|\mathcal{M}|$. ■

Lemma 1.6.23. *There exist constants $c_{28} > 0$, $c_{29} > 0$ such that:*

$$P\left[(B^n_{\text{signals}})^{\mathrm{c}}\right] \le c_{29} e^{-c_{28} n}. \tag{1.6.87}$$

Proof of Lemma 1.6.23. We show that

$$P[B^n_{\text{sig rr}}] \ge 1 - c_{30} e^{-c_{28} n} \tag{1.6.88}$$

for some constants $c_{30} > 0$ and $c_{28} > 0$. The proof for $B^n_{\text{sig rl}}$, $B^n_{\text{sig lr}}$, and $B^n_{\text{sig ll}}$ can be done analogously. Take a right ladder path $\pi \in [-2 \cdot l2^{2n}, 2 \cdot l2^{2n}]^{[0, c_1 n/2[}$ and an admissible

piece of path $\pi' \in \mathsf{AdPath}(2 \cdot l2^{2n}, c_1 n/2)$ with $\pi(0) > \pi'(0)$. We show by induction over $j \in [0, c_1 n/2[$ with the abbreviation $I = [0, j+1[$ and $I' = [0, j[$:

$$P\left[\xi \circ \pi \lceil I' = \xi \circ \pi' \lceil I'\right] = |\mathcal{C}|^{-j}. \tag{1.6.89}$$

Indeed, (1.6.89) is trivial for $j = 0$. For the step $j \mapsto j+1$, we observe that $\pi(j)$ is right of all $\pi(i)$ and $\pi'(i)$, $i < j$, since π is a right ladder path and $\pi(0) > \pi'(0)$. Thus $\xi \circ \pi(j)$ is independent of the family $(\xi \circ \pi \lceil I', \xi \circ \pi' \lceil I')$. Therefore, using our induction hypothesis,

$$\begin{aligned} &P[\xi \circ \pi \lceil I = \xi \circ \pi' \lceil I] \\ &= \quad P[\xi \circ \pi \lceil I' = \xi \circ \pi' \lceil I'] \cdot P[\xi \circ \pi(j) = \xi \circ \pi'(j)] = |\mathcal{C}|^{-j-1}, \end{aligned} \tag{1.6.90}$$

For $j = c_1 n/2$ we obtain that

$$P[\xi \circ \pi \lceil [0, c_1 n/2[= \xi \circ \pi' \lceil [0, c_1 n/2[] \le |\mathcal{C}|^{-c_1 n/2}. \tag{1.6.91}$$

There are no more than $4 \cdot l2^{2n} + 1 \le 5 \cdot l2^{2n}$ such π and not more than $5 \cdot l2^{2n} |\mathcal{M}|^{c_1 n/2}$ such π'. Therefore

$$P[(B^n_{\mathrm{sig\,r}})^{\mathrm{c}}] \le (5 \cdot l2^{2n})^2 |\mathcal{M}|^{c_1 n/2} |\mathcal{C}|^{-c_1 n/2} \tag{1.6.92}$$

holds; consequently (1.6.88) is valid with $c_{29} = 25 \cdot l^2$ and $c_{28} := 1 \le c_1 \log(|\mathcal{C}|/|\mathcal{M}|)/2 - 4 \log 2$. The last inequality holds, since $|\mathcal{C}| > |\mathcal{M}|$ and c_1 was chosen large enough; see subsection 1.2.1. ■

Lemma 1.6.24. *There exist constants $c_{31} > 0$ and $c_{32} > 0$ such that:*

$$P\left[(B^n_{\mathrm{seed\,I}})^{\mathrm{c}} \cap E^m_{\mathrm{stop},\tau}\right] \le c_{32} e^{-c_{31} n}. \tag{1.6.93}$$

Proof. We proceed similarly to the proof of Lemma 1.6.19. In the following, Z runs over all classes $Z \in \mathbb{Z}/l_{\rightarrow}\mathbb{Z}$. We set for all Z (compare with Definition (1.6.17) of $B^n_{\mathrm{seed\,I}}$):

$$B^{n,k}_Z := \left\{ \begin{array}{l} S(\tau(k)+h) \in Z,\ S\lceil(\tau(k)+h+[0, 3c_1 n l_{\leftarrow}]) \text{ is a right ladder} \\ \text{path, and } S\lceil(\tau(k)+h+3c_1 n l_{\leftarrow}+[0, 3c_1 n l_{\rightarrow}]) \text{ is a left ladder} \\ \text{path.} \end{array} \right\}, \tag{1.6.94}$$

$$A^{n,k}_Z := E^m_{\mathrm{stop},\tau,k} \setminus B^{n,k}_Z. \tag{1.6.95}$$

where $E^m_{\mathrm{stop},\tau,k}$ is given by (1.6.71). Note that $B^n_{\mathrm{seed\,I}} = \bigcap_Z \bigcup_{k=0}^{2^{\alpha n}-1} B^{n,k}_Z$ and still $E^m_{\mathrm{stop},\tau} \subseteq E^m_{\mathrm{stop},\tau,k}$ for $k < 2^{\alpha n}$; thus

$$E^m_{\mathrm{stop},\tau} \setminus B^n_{\mathrm{seed\,I}} \subseteq \bigcup_Z \bigcap_{k=0}^{2^{\alpha n}-1} A^{n,k}_Z. \tag{1.6.96}$$

We obtain

$$\begin{aligned} P\left[E^m_{\mathrm{stop},\tau} \setminus B^n_{\mathrm{seed\,I}}\right] &\le l_{\rightarrow} \max_Z P\left[\bigcap_{k=0}^{2^{\alpha n}-1} A^{n,k}_Z\right] \\ &= l_{\rightarrow} \max_Z \prod_{k=0}^{2^{\alpha n}-1} P\left[A^{n,k}_Z \,\middle|\, \bigcap_{j<k} A^{n,j}_Z\right], \end{aligned} \tag{1.6.97}$$

$$P\left[A^{n,k}_Z \,\middle|\, \bigcap_{j<k} A^{n,j}_Z\right] \le P\left[(B^{n,k}_Z)^{\mathrm{c}} \,\middle|\, E^m_{\mathrm{stop},\tau,k} \cap \bigcap_{j<k} A^{n,j}_Z\right]. \tag{1.6.98}$$

Since $h + 3c_1 nl_\leftarrow + 3c_1 nl_\rightarrow < 2 \cdot 2^{2n}$, we have $C_Z^{n,k} := E^m_{\text{stop},\tau,k} \cap \bigcap_{j<k} A_Z^{n,j} \in \mathcal{F}_{\tau(k)}$. Using Lemma 1.6.5, we know $l_\rightarrow \mathbb{Z} + \operatorname{supp} \mu^h = \mathbb{Z}$; hence $c_{33} := \inf_{x\in\mathbb{Z}} P[x + S(h) \in Z] > 0$; (note that the random walk $x + S$ starts in the point x). Moreover, given that $S(h) \in Z$, the probability to follow a right ladder path in Z in the subsequent $3c_1 nl_\leftarrow$ steps is $\mu(\{l_\rightarrow\})^{c_1 nl_\leftarrow}$, and the probability to follow then a left ladder path in the next $3c_1 nl_\rightarrow$ steps is $\mu(\{-l_\leftarrow\})^{c_1 nl_\rightarrow}$.

Thus by the strong Markov property:

$$P\left[B_Z^{n,k} \mid C_Z^{n,k}\right] \geq c_{33}\mu(\{l_\rightarrow\})^{3c_1 nl_\leftarrow}\mu(\{-l_\leftarrow\})^{3c_1 nl_\rightarrow} = c_{33}e^{-c_{34}c_1 n}, \tag{1.6.99}$$

where $c_{34} := -3l_\leftarrow \log\mu(\{l_\rightarrow\}) - 3l_\rightarrow \log\mu(\{-l_\leftarrow\})$. We combine (1.6.97), (1.6.98) and (1.6.99) to obtain

$$P\left[E^m_{\text{stop},\tau} \setminus B^n_{\text{seed I}}\right] \leq l_\rightarrow(1 - c_{33}e^{-c_{34}c_1 n})^{2^{\alpha n}} \leq l_\rightarrow \exp\left\{-c_{33}e^{-c_{34}c_1 n}2^{\alpha n}\right\}. \tag{1.6.100}$$

We have $c_{34}c_1 < \alpha \log 2$, since α was chosen large enough; see subsection 1.2.1. Thus the right hand side of the last inequality converges to 0 superexponentially fast as $n \to \infty$. This proves the lemma, since $n \geq n_0$ and n_0 was chosen large enough. ■

Finally we reap the results of this section:

Proof of Theorem 1.6.1. By Theorems 1.6.2 and 1.6.3 we have

$$E^n_{\text{xi does it}} \cap E^n_{\text{all pieces ok}} \supseteq \tag{1.6.101}$$

$$B^n_{\text{seed I}} \cap B^n_{\text{unique fit}} \cap B^n_{\text{all paths}} \cap B^n_{\text{outside out}} \cap B^n_{\text{recogn straight}} \cap B^n_{\text{signals}} \cap E^m_{\text{stop},\tau}, \tag{1.6.102}$$

hence

$$\begin{aligned} E^m_{\text{stop},\tau} \setminus (E^n_{\text{xi does it}} \cap E^n_{\text{all pieces ok}}) \subseteq\ & (B^n_{\text{unique fit}})^{\text{c}} \cup ((B^n_{\text{all paths}})^{\text{c}} \cap E^m_{\text{stop},\tau}) \\ \cup\ & (B^n_{\text{outside out}})^{\text{c}} \cup (B^n_{\text{recogn straight}})^{\text{c}} \cup (B^n_{\text{signals}})^{\text{c}} \cup ((B^n_{\text{seed I}})^{\text{c}} \cap E^m_{\text{stop},\tau}). \end{aligned} \tag{1.6.103}$$

Thus Theorem 1.6.1 follows from the main Lemmas 1.6.18, 1.6.19, 1.6.20, 1.6.22, 1.6.23, and 1.6.24 of this subsection. ■

1.7 How to find back: Correctness of the stopping times T_f

In this section, we prove Theorem 1.3.5.

Definition 1.7.1. *Let $T = (T_k)_{k\in\mathbb{N}}$ be a sequence of $\mathcal{G}$-adapted stopping times. We define the events*

$$E^m_{\text{no error},T} := \left\{\forall k \geq 0 : \text{if } T_k(\chi) < 2^{12\alpha n_m}, \text{ then } |S(T_k(\chi))| \leq 2^{n_m}\right\}, \tag{1.7.1}$$

$$E^m_{\text{enough back}} := \left\{\begin{matrix}\text{Up to time } 2^{12\alpha n_m}/8,\ S \text{ visits 0 at least} \\ 2^{3\alpha n_m} \text{ times}\end{matrix}\right\}. \tag{1.7.2}$$

We abbreviate

$$\Xi^m_{\text{reconst},f} := \left\{\xi \in \mathcal{C}^{\mathbb{Z}} \;\middle|\; P\left[E^m_{\text{reconst},f} \mid \xi\right] \geq \frac{1}{2}\right\}; \tag{1.7.3}$$

recall Definition (1.3.15) of the event $E^m_{\text{reconst},f}$.

Lemma 1.7.1. *For some constant c_{35} and all $m \geq 0$:*

$$1 - P\left[E^m_{\text{enough back}}\right] \leq c_{35} 2^{-\alpha n_m}. \tag{1.7.4}$$

Proof of Lemma 1.7.1. Let $(X_i)_{i\geq 1}$ denote the time difference between the $(i+1)$st and the i-th visit of S at the origin. By recurrence, $(X_i)_{i\geq 1}$ is a.s. well defined, and by the strong Markov property it is i.i.d. with respect to P. Since S starts in the origin, X_1 is the first return time to the origin, and $\sum_{i=1}^{j} X_i$ is (a.s.) the time of the j-th visit at the origin. For the sake of this proof, we abbreviate: $x = 2^{12\alpha n_m}/8$ and $y = 2^{3\alpha n_m}$. Using

$$\left(E^m_{\text{enough back}}\right)^{\text{c}} = \left\{\sum_{i=1}^{y} X_i \geq x\right\} \subseteq \left\{\left(\sum_{i=1}^{y} X_i^{1/3}\right)^3 \geq x\right\} \tag{1.7.5}$$

and the Chebyshev-Markov inequality, we obtain the claim (1.7.4):

$$\begin{aligned} 1 - P\left[E^m_{\text{enough back}}\right] &\leq P\left[\sum_{i=1}^{y} X_i^{1/3} \geq x^{1/3}\right] \leq x^{-1/3} E\left[\sum_{i=1}^{y} X_i^{1/3}\right] \\ &= x^{-1/3} y E\left[X_1^{1/3}\right] = 2E\left[X_1^{1/3}\right] 2^{-\alpha n_m}. \end{aligned} \tag{1.7.6}$$

The fact $E\left[X_1^{1/3}\right] < \infty$ is an immediate consequence of a lemma proved on page 382 of [21]. In our context, this lemma states that there exists a constant $c_{58} > 0$ such that $P[S(k) \neq 0; k = 1, 2, \ldots, n] \leq c_{58} n^{-1/2}$ for all $n > 0$. ■

Definition 1.7.2. *Let $v(k)$, $k \geq 0$, denote the $(k+1)$st visit of S to the origin. We introduce a random set $\mathbb{T}'_f(\xi, \chi)$ and an event $E^{m+1}_{\text{when back recog}}$:*

$$\mathbb{T}'_f(\xi,\chi) := \left\{t \in \mathbb{N} \,\middle|\, \xi\lceil[-2^{n_m}, 2^{n_m}] \preccurlyeq f(\theta^t(\chi)) \preccurlyeq \xi\lceil[-9\cdot 2^{n_m}, 9\cdot 2^{n_m}]\right\}, \tag{1.7.7}$$

$$E^{m+1}_{\text{when back recog},f} := \left\{ \begin{array}{l} \text{For more than } 1/4 \text{ of the points } k \in [0, 2^{2\alpha n_{m+1}}[\text{ holds} \\ v(k2^{\alpha n_{m+1}}) \in \mathbb{T}'_f(\xi,\chi) \end{array} \right\}. \tag{1.7.8}$$

Lemma 1.7.2. *If the event $E^m_{\text{reconst},f}$ holds, then $\mathbb{T}_f(\chi) \supseteq \mathbb{T}'_f(\xi,\chi) \cap [0, 2^{12\alpha n_{m+1}} - 2\cdot 2^{12\alpha n_m}[$.*

Proof. We know $\xi\lceil[-2^{n_m}, 2^{n_m}] \preccurlyeq f(\chi)$ by $E^m_{\text{reconst},f}$. Let $t \in \mathbb{T}'_f(\xi,\chi)$, $t < 2^{12\alpha n_{m+1}} - 2\cdot 2^{12\alpha n_m}$. Then we also have $\xi\lceil[-2^{n_m}, 2^{n_m}] \preccurlyeq f(\theta^t\chi)$. Hence $t \in \mathbb{T}_f(\chi)$; to this end recall Definition (1.3.12) of the random set $\mathbb{T}_f(\chi)$. This implies the lemma. ■

Lemma 1.7.3. *Assume that the events $E^{m+1}_{\text{no error},T_f} \cap E^{m+1}_{\text{enough back}} \cap E^{m+1}_{\text{when back recog},f}$ and $\mathbb{T}_f(\chi) \supseteq \mathbb{T}'_f(\xi,\chi) \cap [0, 2^{12\alpha n_{m+1}} - 2\cdot 2^{12\alpha n_m}[$ hold. Then $E^{m+1}_{\text{stop},T_f}$ holds, too.*

Proof. Using $E^{m+1}_{\text{enough back}}$, we know

$$v(k2^{\alpha n_{m+1}}) \in [0, 2^{12\alpha n_{m+1}}/8] \subseteq [0, 2^{12\alpha n_{m+1}} - 2\cdot 2^{12\alpha n_m}[\tag{1.7.9}$$

for all $k \in [0, 2^{2\alpha n_{m+1}}[$. Since the event $E^{m+1}_{\text{when back recog},f}$ holds, we obtain $|\mathbb{T}_f(\chi)| \geq |\mathbb{T}'_f(\xi,\chi) \cap [0, 2^{12\alpha n_{m+1}} - 2\cdot 2^{12\alpha n_m}[| \geq 2^{2\alpha n_{m+1}}/4$. By Definition (1.3.13) of the stopping times T_f, this yields $T_{f,k}(\chi) < 2^{12\alpha n_{m+1}}$ for all $k < (2^{2\alpha n_{m+1}}/4)/(2\cdot 2^{2n_{m+1}}) = 2^{2(\alpha-1)n_{m+1}}/8$. The event $E^{m+1}_{\text{no error},T_f}$ holds, and $2^{2(\alpha-1)n_{m+1}}/8 \geq 2^{\alpha n_{m+1}}$; recall that α and $n_{m+1} \geq n_0$ are

large (see Section 1.2.1). Hence we obtain $|S(T_{f,k}(\chi))| \le 2^{n_{m+1}}$ for all $k \in [0, 2^{\alpha n_{m+1}}[$; recall (1.7.1). Using Definition (1.3.13) again, we see that $T_{f,j}(\chi) + 2 \cdot 2^{2n_{m+1}} \le T_{f,k}(\chi)$ is automatically fulfilled for $j < k$ whenever $T_{f,k}(\chi) < 2^{12\alpha n_{m+1}}$, which is the case at least for $k \in [0, 2^{\alpha n_{m+1}}[$. Summarizing, we have proven that the event $E^{m+1}_{\mathrm{stop},T_f}$ holds; recall its definition (1.3.14). ■

Lemma 1.7.4. $P\left[\left(E^{m+1}_{\text{when back recog},f}\right)^{\mathrm{c}} \cap \left\{\xi \in \Xi^m_{\mathrm{reconst},f}\right\}\right] \le 0.9^{2^{2\alpha n_{m+1}}}$.

Proof. We define Bernoulli random variables Y_k, $k \ge 0$, by $Y_k := 1$ if $v(k2^{\alpha n_{m+1}}) \in \mathbb{T}'_f(\xi, \chi)$, and $Y_k := 0$ else. Note that $v\left((k+1)2^{\alpha n_{m+1}}\right) - v(k2^{\alpha n_{m+1}}) \ge 2^{\alpha n_{m+1}} > 2 \cdot 2^{12\alpha n_m}$. Also note that $E^{m+1}_{\text{when back recog},f} = \left\{2^{-2\alpha n_{m+1}} \sum_{k=0}^{2^{2\alpha n_{m+1}}-1} Y_k \ge 1/4\right\}$. Because of the strong Markov property of the random walk $(S(k))_{k\ge 0}$ we have that conditioned under ξ the variables $(Y_k)_{k\ge 0}$ are i.i.d.; recall that $f(\chi)$ depends at most on $\chi\lceil[0, 2 \cdot 2^{12\alpha n_m}[$. If furthermore $\xi \in \Xi^m_{\mathrm{reconst},f}$ holds, then $E[Y_k \mid \xi] \ge 1/2$. Hence we obtain for these ξ, using the exponential Chebyshev inequality for the binomial variable $\sum_{k=0}^{2^{2\alpha n_{m+1}}-1} Y_k$:

$$P\left[\left(E^{m+1}_{\text{when back recog},f}\right)^{\mathrm{c}} \mid \xi\right] = P\left[2^{-2\alpha n_{m+1}} \sum_{k=0}^{2^{2\alpha n_{m+1}}-1} Y_k \le \frac{1}{4} \,\middle|\, \xi\right] \tag{1.7.10}$$

$$\le E\left[e^{1/4-Y_1} \mid \xi\right]^{2^{2\alpha n_{m+1}}} \le \left(\frac{e^{1/4}+e^{-3/4}}{2}\right)^{2^{2\alpha n_{m+1}}} \le 0.9^{2^{2\alpha n_{m+1}}}.$$

This yields the claim of the lemma:

$$P\left[\left(E^{m+1}_{\text{when back recog},f}\right)^{\mathrm{c}} \cap \left\{\xi \in \Xi^m_{\mathrm{reconst},f}\right\}\right] \tag{1.7.11}$$
$$\le P\left[\left(E^{m+1}_{\text{when back recog},f}\right)^{\mathrm{c}} \mid \xi \in \Xi^m_{\mathrm{reconst},f}\right] \le 0.9^{2^{2\alpha n_{m+1}}}.$$

■

Lemma 1.7.5. $P\left[(E^{m+1}_{\text{no error},T_f})^{\mathrm{c}} \cap E^m_{\mathrm{reconst},f}\right] \le \frac{1}{3}e^{-n_{m+1}}$.

Proof. Let υ_i denote the $(i+1)$st time when the random walk S visits a point of $\mathbb{Z}\backslash[-2^{n_{m+1}} + 2l2^{12\alpha n_m}, 2^{n_{m+1}} - 2l2^{12\alpha n_m}]$. We set

$$E^{m+1}_{\mathrm{wrong},i} := \left\{\exists w \in \mathcal{C}^{2\cdot 2^{n_m}} : w \preccurlyeq \xi\lceil[-9 \cdot 2^{n_m}, 9 \cdot 2^{n_m}] \text{ and } w \preccurlyeq f(\theta^{\upsilon_i}(\chi))\right\}. \tag{1.7.12}$$

If the event $E^{m+1}_{\mathrm{wrong},i}$ occurs, then our procedure might fail to estimate correctly the location of the random walk: we might be misled to think that at time $\upsilon_i + 2 \cdot 2^{12\alpha n_m}$ we are close to the origin while we are not.

We claim that the following holds:

$$(E^{m+1}_{\text{no error},T_f})^{\mathrm{c}} \cap E^m_{\mathrm{reconst},f} \subseteq \bigcup_{i=0}^{2^{12\alpha n_{m+1}}-1} E^{m+1}_{\mathrm{wrong},i}. \tag{1.7.13}$$

Indeed: If $(E^{m+1}_{\text{no error},T_f})^{\mathrm{c}}$ holds, then $|S(T_{f,k}(\chi))| > 2^{n_{m+1}}$ for some k with $T_{f,k} < 2^{12\alpha n_{m+1}}$ (see (1.7.1)); thus $|S(T_{f,k}(\chi) - 2 \cdot 2^{12\alpha n_m})| > 2^{n_{m+1}} - 2l2^{12\alpha n_m}$, since S cannot travel faster than speed l. This means $T_{f,k}(\chi) - 2 \cdot 2^{12\alpha n_m} = \upsilon_i$ for some $i < 2^{12\alpha n_{m+1}}$. Using

Definition 1.3.3 of $T_{f,k}(\chi)$, this implies $\upsilon_i \in \mathbb{T}_f(\chi)$; hence there is $w \in \mathcal{C}^{2\cdot 2^{n_m}}$ such that $w \preccurlyeq f(\chi)$ and $w \preccurlyeq f(\theta^{\upsilon_i}(\chi))$. Assuming that the event $E^m_{\text{reconst},f}$ holds, too, this implies $w \preccurlyeq f(\chi) \preccurlyeq \xi\lceil[-9\cdot 2^{n_m}, 9\cdot 2^{n_m}]$; see (1.3.15). This yields that $E^{m+1}_{\text{wrong},i}$ holds; recall (1.7.12). Summarizing, we have shown that (1.7.13) holds.

For all i, $f(\theta^{\upsilon_i}(\chi))$ depends only on $\chi\lceil[\upsilon_i, \upsilon_i + 2\cdot 2^{12\alpha n_m}[$, and S does not visit $[-9\cdot 2^{n_m}, 9\cdot 2^{n_m}]$ in this time interval $[\upsilon_i, \upsilon_i + 2\cdot 2^{12\alpha n_m}[$, since the distance between $[-9\cdot 2^{n_m}, 9\cdot 2^{n_m}]$ and $\mathbb{Z}\backslash[-2^{n_{m+1}} + 2l2^{12\alpha n_m}, 2^{n_{m+1}} - 2l2^{12\alpha n_m}]$ is larger than $2l2^{12\alpha n_m}$, and since the random walk cannot travel faster than l steps per time unit.

Thus by the strong Markov property and by independence of S and ξ, we get: $\chi\lceil[\upsilon_i, \upsilon_i + 2^{12\alpha n_m}]$ is independent of $\xi\lceil[-9\cdot 2^{n_m}, 9\cdot 2^{n_m}]$; therefore $f(\theta^{\upsilon_i}(\chi))$ is independent of $\xi\lceil[-9\cdot 2^{n_m}, 9\cdot 2^{n_m}]$, too.

The probability that a random word of length $2\cdot 2^{n_m}$ which has i.i.d. letters with uniform distribution in $\mathcal{C}$ is equal to a word which is independent of it is equal to $|\mathcal{C}|^{-2\cdot 2^{n_m}}$. There are at most $37\cdot 2^{n_m}$ words of a fixed length in $\xi\lceil[-9\cdot 2^{n_m}, 9\cdot 2^{n_m}]$ and also in $f(\theta^{\upsilon_k}(\chi))$, counting all reversed words, too. Thus there are at most $37^2 2^{2n_m}$ pairs of such words. It follows that

$$P\left[E^{m+1}_{\text{wrong},i}\right] \le 37^2 2^{2n_m} |\mathcal{C}|^{-2\cdot 2^{n_m}}. \tag{1.7.14}$$

Hence we get the claim of the lemma, using (1.7.13):

$$\begin{aligned} P\left[(E^{m+1}_{\text{no error},T_f})^{\text{c}} \cap E^m_{\text{reconst},f}\right] &\le \sum_{i=0}^{2^{12\alpha n_{m+1}}-1} P\left[E^{m+1}_{\text{wrong},i}\right] \\ &\le 2^{12\alpha n_{m+1}} \cdot 37^2 2^{2n_m} |\mathcal{C}|^{-2\cdot 2^{n_m}} \le \frac{1}{3} e^{-n_{m+1}}. \end{aligned} \tag{1.7.15}$$

For the last inequality, recall that $n_m \ge n_0$ is large enough, and note that $|\mathcal{C}|^{-2\cdot 2^{n_m}}$ is the leading term of the last but one expression; also recall that $n_{m+1} = 2^{\lfloor\sqrt{n_m}\rfloor}$ is of a much smaller order than 2^{n_m}. ■

Proof of Theorem 1.3.5. By Lemmas 1.7.2 and 1.7.3, we know $E^{m+1}_{\text{no error},T_f} \cap E^{m+1}_{\text{enough back}} \cap E^{m+1}_{\text{when back recog},f} \cap E^m_{\text{reconst},f} \subseteq E^{m+1}_{\text{stop},T_f}$. Using some Boolean algebra, this implies

$$\begin{aligned} &(E^{m+1}_{\text{stop},T_f})^{\text{c}} \cap E^m_{\text{reconst},f} \cap \left\{\xi \in \Xi^m_{\text{reconst},f}\right\} \\ &\subseteq \left(E^{m+1}_{\text{enough back}}\right)^{\text{c}} \cup \left((E^{m+1}_{\text{no error},T_f})^{\text{c}} \cap E^m_{\text{reconst},f}\right) \cup \left(\left(E^{m+1}_{\text{when back recog},f}\right)^{\text{c}} \cap \left\{\xi \in \Xi^m_{\text{reconst},f}\right\}\right). \end{aligned} \tag{1.7.16}$$

Consequently, using Definition (1.7.3) of $\Xi^m_{\text{reconst},f}$ and Lemmas 1.7.1, 1.7.4, and 1.7.5:

$$\begin{aligned} &P\left[(E^{m+1}_{\text{stop},T_f})^{\text{c}} \cap E^m_{\text{reconst},f} \cap \left\{P\left[E^m_{\text{reconst},f} \mid \xi\right] \ge \frac{1}{2}\right\}\right] \\ &\le P\left[\left(E^{m+1}_{\text{enough back}}\right)^{\text{c}}\right] + P\left[(E^{m+1}_{\text{no error},T_f})^{\text{c}} \cap E^m_{\text{reconst},f}\right] \\ &\quad + P\left[\left(E^{m+1}_{\text{when back recog},f}\right)^{\text{c}} \cap \left\{\xi \in \Xi^m_{\text{reconst},f}\right\}\right] \\ &\le c_{35} 2^{-\alpha n_{m+1}} + \frac{1}{3} e^{-n_{m+1}} + 0.9^{2^{2\alpha n_{m+1}}} \le e^{-n_{m+1}}; \end{aligned} \tag{1.7.17}$$

recall that α and $n_{m+1} \ge n_0$ are large (see Section 1.2.1). This proves Theorem 1.3.5. ■

1.8 Getting started: The first stopping times

In this section, we prove Theorem 1.3.4.

1.8.1 The stopping times T^0

We start with the definition of a sequence $T^0 = (T^0_k)_{k\geq 0}$ of $\mathcal{G}$-adapted stopping times with values in $[0, 2^{12\alpha n_0}]$. Roughly speaking, these times search for long blocks of 1's in the observation χ. Here is intuitive idea behind this construction: Since we conditioned on a large block of 1's to occur in the true scenery ξ close to the origin, observing a long block of 1's at a later time indicates with high probability that the random walk has returned close to the origin. This is true only up to a certain time horizon, since long blocks of 1's in the true scenery will occur far from the origin, as well.

Definition 1.8.1. *Let the random set $\mathbb{T}^0(\chi)$ be defined as follows:*

$$\mathbb{T}^0(\chi) := \left\{ t \in [0, 2^{12\alpha n_0} - n_0^7[\;\middle|\; \chi[\,t, t+n_0^7] \text{ is constant } 1 \right\}. \tag{1.8.1}$$

We arrange the elements of $\mathbb{T}^0(\chi)$ in increasing order: $t_0(0) < \ldots < t_0(|\mathbb{T}^0(\chi)| - 1)$. We set

$$T^0_k(\chi) := \begin{cases} t_0(2 \cdot 2^{2n_0} k) + n_0^7 & \text{if } 2 \cdot 2^{2n_0} k < |\mathbb{T}^0(\chi)|, \\ 2^{12\alpha n_0} & \text{otherwise.} \end{cases} \tag{1.8.2}$$

Recall Definition (1.3.14):

$$E^0_{\text{stop},T^0} = \bigcap_{k=0}^{2^{\alpha n_0}} \left\{ T^0_k(\chi) < 2^{12\alpha n_0},\ |S(T^0_k(\chi))| \leq 2^{n_0}, T^0_j(\chi) + 2 \cdot 2^{2n_0} \leq T^0_k(\chi) \text{ for } j < k \right\}. \tag{1.8.3}$$

Theorem 1.8.1. *For some positive constants c_{36} and c_{37} holds $\tilde{P}\left[E^0_{\text{stop},T^0}\right] \geq 1 - c_{36} e^{-c_{37} n_0}$.*

We prepare the proof of Theorem 1.8.1 by some Definitions and Lemmas. We use again the abbreviation $J_1 = [-2ln_0^{20}, 2ln_0^{20}]$.

Definition 1.8.2 (Analogue of Definition (1.7.2)). *Let $I \subseteq J_1$ be an integer interval. We define*

$$\tilde{E}^0_{\text{enough back},I} := \{\text{Up to time } 2^{12\alpha n_0}/4,\ S \text{ visits the interval } I \text{ at least } 2^{3\alpha n_0} \text{ times } \}. \tag{1.8.4}$$

Lemma 1.8.1. *For some constants $c_{38} > 0$ and $c_{39} > 0$, the following holds: If $I \subseteq J_1$, $|I| \geq l$, is an integer interval, then*

$$\tilde{P}[\tilde{E}^0_{\text{enough back},I}] \geq 1 - c_{38} e^{-c_{39} n_0}. \tag{1.8.5}$$

Proof. Let $T_I := \inf\{t \mid S(t) \in I\}$ be the entrance time of S into I. We show first: For some positive constants c_{40} and c_{41} (depending at most on the distribution μ of $S(1)$) we have:

$$\tilde{P}\left[T_I \geq 2^{12\alpha n_0}/8\right] \leq c_{40} e^{-c_{41} n_0}. \tag{1.8.6}$$

If $0 \in I$, this is trivial, since S starts in 0. Otherwise I contains only positive numbers, or it contains only negative numbers; without loss of generality we assume the first possibility. Let $z = \min I \in]0, 2ln_0^{20}]$. Consider the interval $J :=]-2^{n_0}, z[\subseteq]-2^{n_0}, 2^{n_0}[$, and consider the exit time $H := \inf\{t \mid S(t) \notin J\}$ of J. Note that H is a.s. finite.

On the one hand, we know

$$\tilde{P}[H \geq t] \leq c_{42} e^{-c_{43} 2^{-2n_0} t} \tag{1.8.7}$$

for some constants $c_{42}, c_{43} > 0$ depending at most on the variance of $S(1)$, since in every time interval of size 2^{2n_0} the random walk has a positive probability to exit J, bounded away from 0. In particular, for $t = 2^{12\alpha n_0}/8$ the probability in (1.8.7) is superexponentially small in n_0.

On the other hand, since S is a martingale and since S has jumps sizes bounded by l, we get $\tilde{P}[S(H) > 0] \geq 1 - (z + l)2^{-n_0}$. Furthermore, using again that S has jumps sizes bounded by l, we know the following: If $S(H) > 0$, then $S(H) \in I$ and $T_I = H$, since z is the leftmost point in I and $|I| \geq l$; the random walk cannot cross I without touching it.

Altogether, we have the following upper bound for the left hand side in (1.8.6):

$$\tilde{P}[H \geq 2^{12\alpha n_0}/8 \text{ or } S(H) < 0] \leq c_{40} e^{-c_{41} n_0}. \tag{1.8.8}$$

for some positive constants c_{40} and c_{41}.

Provided the random walk visits a point $x \in I$, the probability to visit this point again at least $2^{3\alpha n_m}$ times in the subsequent $2^{12\alpha n_m}/8$ time steps is at least $1 - c_{28} 2^{-\alpha n_0}$. This follows from Lemma 1.7.1, using the strong Markov property of the random walk; recall that the law of S with respect to P and with respect to $\tilde{P}$ coincide. Combining this with (1.8.6) yields claim (1.8.5) of Lemma 1.8.1. ■

We remark: Lemma 1.8.1 holds not only for deterministic intervals I, but also for random ones, provided that I and S are independent. We use this below for the following specific choice of I, which depends on the scenery ξ, but not on S:

$\tilde{P}$-a.s. there is a (random) integer interval $J_0 \subseteq J_1 = [-2ln_0^{20}, 2ln_0^{20}]$ with $|J_0| \geq n_0^4$ such that $\xi\lceil J_0$ is constant 1; recall Definition 1.3.1. Just for definiteness we take the rightmost such J_0. Let

$$I = I(\xi) := \left\{ z \in J_0 \;\middle|\; \operatorname{dist}(z, \mathbb{Z} \setminus J_0) > n_0^4/4 \right\}; \tag{1.8.9}$$

then I is $\tilde{P}$-a.s. well defined, and it is an integer interval containing $|I| \geq n_0^4/2 \geq l$ points.

Definition 1.8.3 (Modification of Definition 1.7.2). *Let $w(k)$, $k \geq 0$, denote the $(k+1)$st visit to the (random) set $I(\xi)$ by the random walk S. We introduce a random set $\mathbb{T}^{0\prime}$ and an event $\tilde{E}^0_{\text{when back recog}}$:*

$$\mathbb{T}^{0\prime} := \left\{ t \in \mathbb{N} \,\middle|\, S(t) \in I(\xi) \text{ and } |S(j) - S(t)| \leq n_0^4/4 \text{ for } 0 \leq j - t \leq n_0^7 \right\}, \tag{1.8.10}$$

$$\tilde{E}^0_{\text{when back recog}} := \{\text{For more than } 1/4 \text{ of the points } k \in [0, 2^{2\alpha n_0}[\text{ holds } w(k2^{\alpha n_0}) \in \mathbb{T}^{0\prime} \}. \tag{1.8.11}$$

Lemma 1.8.2 (Modification of Lemma 1.7.4). $\tilde{P}\left[\left(\tilde{E}^0_{\text{when back recog}}\right)^c\right] \leq 0.9^{2^{2\alpha n_0}}$.

Proof. We observe as in (1.4.16) by the submartingale inequality:

$$P[|S(j)| \le n_0^4/4 \text{ for } 0 \le j \le n_0^7] \ge 1 - 4^2 n_0^{-8} E\left[S(n_0^7)^2\right] = 1 - \frac{16\,\mathrm{Var}[S(1)]}{n_0} \ge \frac{1}{2}, \quad (1.8.12)$$

since n_0 is large enough; see Subsection 1.2.1. Let Y_k denote the indicator function of the event $\{w(k2^{\alpha n_0}) \in \mathbb{T}^{0\prime}\}$; the Y_k are $\tilde{P}$-a.s. well defined. As a consequence of the strong Markov property, the Y_k, $k \in [0, 2^{2\alpha n_0}[$, are i.i.d. Bernoulli random variables; note that the stopping times $w(k2^{\alpha n_0})$, $k \in \mathbb{N}$, have at least the spacing $2^{\alpha n_0} > n_0^7$. Furthermore $\tilde{P}[Y_k = 1] \ge 1/2$, since this probability equals the left hand side in (1.8.12). The claim of the Lemma now follows by the same large deviation argument as in (1.7.10). ■

Lemma 1.8.3 (Analogue of Lemma 1.7.2). *The inclusion* $\mathbb{T}^0 \supseteq \mathbb{T}^{0\prime} \cap [0, 2^{12\alpha n_0} - n_0^7[$ *holds* $\tilde{P}$*-almost surely.*

Proof. Assume that the event BigBlock holds; this occurs $\tilde{P}$-almost surely. Then $I(\xi)$ is well defined. Let $t \in \mathbb{T}^{0\prime}$, $t < 2^{12\alpha n_0} - n_0^7$. Then $S(t) \in I(\xi)$, and during the subsequent n_0^7 steps, the random walk S cannot leave the interval J_0, since it does not travel farther than $n_0^4/4$ (recall definition (1.8.10)), and since $\mathbb{Z} \setminus J_0$ is more distant than this from $I(\xi)$ (recall the definition of $I(\xi)$). Since $\xi\lceil J_0$ is constant 1 by definition (1.3.4) of the event BigBlock, this implies that $S\lceil[t, t + n_0^7]$ is constant 1; i.e. $t \in \mathbb{T}^0$. ■

Recall Definition (1.7.1):

$$E^0_{\text{no error},T^0} = \{\forall k \in \mathbb{N} : \text{ If } T_k^0(\chi) < 2^{12\alpha n_0}, \text{ then } |S(T_k^0(\chi))| \le 2^{n_0}\}. \quad (1.8.13)$$

Lemma 1.8.4 (Modification of Lemma 1.7.3). *Assume that the events* $E^0_{\text{no error},T^0} \cap \tilde{E}^0_{\text{enough back},I(\xi)} \cap \tilde{E}^0_{\text{when back recog}}$ *and* $\mathbb{T}^0(\chi) \supseteq \mathbb{T}^{0\prime} \cap [0, 2^{12\alpha n_0} - n_0^7[$ *hold. Then* E^0_{stop,T^0} *holds, too.*

Proof. The proof is almost the same as for Lemma 1.7.3. Note that the small differences between the definitions of $\tilde{E}^0_{\text{enough back},I(\xi)}$ and $E^{m+1}_{\text{enough back}}$ are not essential for the validity of the proof. ■

Definition 1.8.4. *We define the event of sceneries*

$$\Xi^0_{\text{no blocks}} := \left\{ \xi \in \mathcal{C}^{\mathbb{Z}} \,\middle|\, \begin{array}{l}\text{For every (integer) interval } J \subseteq [-2l2^{12\alpha n_0}, 2l2^{12\alpha n_0}] \setminus J_1 \\ \text{with } |J| = n_0^2 \text{ holds: } \xi\lceil J \text{ is not constant 1.}\end{array} \right\} \quad (1.8.14)$$

Lemma 1.8.5. *For some positive constants* c_{44}, c_{45} *holds* $\tilde{P}[\xi \in \Xi^0_{\text{no blocks}}] \ge 1 - c_{44}e^{-c_{45}n_0}$.

Proof. For every fixed interval $J \subseteq [-2 \cdot l2^{12\alpha n_0}, 2 \cdot l2^{12\alpha n_0}] \setminus J_1$ with $|J| = n_0^2$ we have

$$\tilde{P}[\xi\lceil J \text{ is constant 1}] = |\mathcal{C}|^{-n_0^2}, \quad (1.8.15)$$

which is superexponentially small in n_0. Furthermore, there are less than $4l2^{12\alpha n_0}$ such intervals. Thus $\tilde{P}[\xi \notin \Xi^0_{\text{no blocks}}] \le 4l2^{12\alpha n_0}|\mathcal{C}|^{-n_0^2}$, which is still superexponentially small in n_0. This implies the lemma. Note that we may choose c_{44}, c_{45} independent α for n_0 large enough, even though $4l2^{12\alpha n_0}|\mathcal{C}|^{-n_0^2}$ does depend on α (see Subsection 1.2.1. ■

Lemma 1.8.6. *For some constants* $c_{46}, c_{47} > 0$ *holds*

$$\tilde{P}[(E^0_{\text{no error},T^0})^{\mathrm{c}}] \le c_{46}e^{-c_{47}n_0}. \quad (1.8.16)$$

Proof. Let $\mathcal{X}$ be defined by

$$\mathcal{X} := \{x \in \mathbb{Z} | x + [-ln_0^7, ln_0^7] \subseteq [-2l2^{12\alpha n_0}, 2l2^{12\alpha n_0}] \setminus J_1\}. \tag{1.8.17}$$

As a consequence of Lemma 1.4.2 (with the parameters $k = n_0^7$ and $\kappa = n_0^2$) we know for every $\xi \in \mathcal{C}^{\mathbb{N}}$ such that $\xi\lceil[-ln_0^7, ln_0^7]$ contains no block of 1's of length n_0^2:

$$P_\xi[\xi \circ S\lceil[0, n_0^7] \text{ is constant } 1] \leq e^{-c_7 n_0^3}. \tag{1.8.18}$$

Let $t \in \mathbb{N}$ and let $\xi \in \Xi^0_{\text{no blocks}}$. Using the Markov property of the random walk, (1.8.18) implies the following:

$$P_\xi[\xi \circ S\lceil(t + [0, n_0^7]) \text{ is constant } 1 \mid S(t) \in \mathcal{X}] \leq e^{-c_7 n_0^3}. \tag{1.8.19}$$

If $t < 2^{12\alpha n_0}$ and $|S(t)| > 2^{n_0}$ holds, then we know $S(t) \in \mathcal{X}$; note that $J_1 = [-2ln_0^{20}, 2ln_0^{20}]$ has a distance larger than ln_0^7 from $\mathbb{Z} \setminus [-2^{n_0}, 2^{n_0}]$, and recall that S cannot travel faster than with speed l, and that n_0 is large by Subsection 1.2.1.

Thus (1.8.19) implies

$$\begin{aligned} &P_\xi[(E^0_{\text{no error},T^0})^c] \\ &\leq P_\xi[\text{There is } t < 2^{12\alpha n_0} \text{ such that } |S(t)| > 2^{n_0} \text{ and } \xi \circ S\lceil(t + [0, n_0^7]) \text{ is constant } 1] \\ &\leq 2^{12\alpha n_0} e^{-c_7 n_0^3} \leq e^{-n_0}; \end{aligned} \tag{1.8.20}$$

for the last inequality recall that n_0 was chosen large enough, depending on α (see Subsection 1.2.1). Combining this with Lemma 1.8.5 yields for some positive constants c_{46}, c_{47}:

$$\tilde{P}[(E^0_{\text{no error},T^0})^c] \leq \tilde{P}[\xi \notin \Xi^0_{\text{no blocks}}] + \int_{\{\xi \in \Xi^0_{\text{no blocks}}\}} P_\xi[(E^0_{\text{no error},T^0})^c]\, d\tilde{P} \leq c_{46} e^{-c_{47} n_0}. \tag{1.8.21}$$

■

Proof of Theorem 1.8.1. ¿From Lemmas 1.8.3 and 1.8.4 we know that

$$\tilde{P}[(E^0_{\text{stop},T^0})^c] \leq \tilde{P}[(E^0_{\text{no error},T^0})^c] + \tilde{P}[(\tilde{E}^0_{\text{enough back},I(\xi)})^c] + \tilde{P}[(\tilde{E}^0_{\text{when back recog}})^c]. \tag{1.8.22}$$

Hence the claim of Theorem 1.8.1 is a consequence of Lemmas 1.8.1, 1.8.2, and 1.8.6. ■

1.8.2 The stopping times T^1

Unfortunately, the constructed stopping times T^0 are not good enough as arguments for the first reconstruction Algorithm "Alg^n": We cannot construct more than roughly $\exp(\text{const}\, n_0^4)$ reliable stopping times based on the way we build the T^0 stopping times; in fact we use much less than this number. Otherwise we cannot guarantee that they really stop the random walk with high probability close to the origin. However, the number $\exp(\text{const}\, n_0^4)$ is much too small to collect a sufficiently large puzzle for reconstructing at least the modified piece $\xi\lceil J_1$ in the scenery using our reconstruction algorithm; to illustrate this fact, we remark that we have only roughly an upper bound $d\tilde{P}/dP \leq \exp(\text{const}\, n_0^{20})$; see Lemma 1.4.3. A modification of the parameters does not

solve this problem; we need an essentially improved series of stopping times T^1 to get the reconstruction algorithm started.

Our construction of T^1 is partially parallel to the construction of the partial reconstruction algorithm Alg^n, but it is also partially parallel to the construction of the stopping times T_f and T^0: Roughly speaking, we collect a set of typical signals ("a puzzle") at the very beginning and another one at a candidate time. Instead of matching the pieces together, we just compare the two puzzles: If the puzzles have a sufficiently high overlap, then they were generated with high probability at roughly the same location.

Fortunately, many constructions of the previous sections can be used again, up to small modifications: There are extra complications due to the presence of a modified domain J_1. We keep the presentation as close as possible to the previous sections to show the parallelism. Here is the formal definition of the "new" puzzles and of T^1:

Definition 1.8.5. *We set, using the abbreviation* $\mathsf{Input} := (T^0(\chi), \chi\lceil[0, 2 \cdot 2^{12\alpha n_0}[)$ *again:*

$$\mathsf{Puzzle}^{n_0}_{\mathrm{III}}(\chi) := \tag{1.8.23}$$
$$\left\{(w_1, w_2, w_3) \in \mathsf{Puzzle}^{n_0}_{\mathrm{I}}(\mathsf{Input}) \,\middle|\, \exists k \in [0, 2^{\alpha n_0}[\colon\ w_1 w_2 w_3 \sqsubseteq \chi\lceil(T^0_k(\chi) + [0, 2^{n_0}/l])\right\},$$

$$\mathsf{Puzzle}^{n_0}_{\mathrm{IV}}(\chi) := \{w_2 \ \in \mathcal{C}^{c_1 n_0} \mid \exists w_1, w_3 \in \mathcal{C}^{c_1 n_0} \colon\ (w_1, w_2, w_3) \in \mathsf{Puzzle}^{n_0}_{\mathrm{III}}(\chi)\}, \tag{1.8.24}$$

$$\mathbb{T}^1(\chi) := \left\{ t \in [0, 2^{12\alpha n_1} - 2 \cdot 2^{12\alpha n_0}[\;\middle|\; \begin{array}{l} |\mathsf{Puzzle}^{n_0}_{\mathrm{IV}}(\chi) \cap \mathsf{Puzzle}^{n_0}_{\mathrm{IV}}(\theta^t\chi)| \geq 2^{n_0/3} \\ \text{and } |\mathsf{Puzzle}^{n_0}_{\mathrm{IV}}(\theta^t\chi)| \leq 50 \cdot 2^{n_0} \end{array} \right\}. \tag{1.8.25}$$

Finally we define another sequence $T^1 = (T^1_k)_{k\geq 0}$ *of* $\mathcal{G}$*-adapted stopping times with values in* $[0, 2^{12\alpha n_1}]$*: Let* $t_1(0) < \ldots < t_1(|\mathbb{T}^1(\chi)| - 1)$ *be the elements of* $\mathbb{T}^1(\chi)$ *arranged in increasing order. For* $k \in \mathbb{N}$*, we set*

$$T^1_k(\chi) := \begin{cases} t_1(2 \cdot 2^{2n_1}k) + 2 \cdot 2^{12\alpha n_0} & \text{if } 2 \cdot 2^{2n_1}k < |\mathbb{T}^1(\chi)|, \\ 2^{12\alpha n_1} & \text{otherwise.} \end{cases} \tag{1.8.26}$$

Note that $T^0(\chi)$ only depends on $\chi\lceil[0, 2^{12\alpha n_0}[$; thus $\mathsf{Puzzle}^{n_0}_{\mathrm{IV}}(\chi)$ only depends on $\chi\lceil[0, 2 \cdot 2^{12\alpha n_0}[$, since $2^{n_0}/l \leq 2^{12\alpha n_0}$.

Definition 1.8.6. *Using the abbreviation* $J_1 = [-2ln_0^{20}, 2ln_0^{20}]$ *from Definition 1.3.1 again, we define the following random sets:*

$$\mathsf{CorPaths} := \left\{ R \in \mathbb{Z}^{[0, c_1 n_0[} \,\middle|\, \begin{array}{l} R \text{ is an admissible piece of path, for every admissible piece of} \\ \text{path } R' : [0, c_1 n_0[\to \mathbb{Z} \text{ with } R'(0) = R(0) \text{ and } R'(c_1 n_0 - 1) = \\ R(c_1 n_0 - 1) \text{ holds } \xi \circ R' = \xi \circ R, \text{ and there is such a path } R' \\ \text{which takes at least one value in } J_1. \end{array} \right\}, \tag{1.8.27}$$

$$\mathsf{Corrupted} := \{\xi \circ R \in \mathcal{C}^{c_1 n_0} \,|R \in \mathsf{CorPaths}\}, \tag{1.8.28}$$

$$\mathsf{Center}_{\mathrm{I}} := \{w \in \mathcal{C}^{c_1 n_0} \mid w \text{ is a (left or right) ladder word of } \xi\lceil([-11 \cdot 2^{n_0}, 11 \cdot 2^{n_0}] \setminus J_1)\}, \tag{1.8.29}$$

$$\mathsf{Center}_{\mathrm{II}} := \mathsf{Center}_{\mathrm{I}} \cup \mathsf{Corrupted}, \tag{1.8.30}$$

$$\mathsf{Center}_{\mathrm{III}} := \left\{ w \in \mathcal{C}^{c_1 n_0} \mid \begin{array}{l} w \text{ is a right ladder word of} \\ \xi\lceil([-2 \cdot 2^{n_0/2}, 2 \cdot 2^{n_0/2}] \setminus [-2^{n_0/2}, 2^{n_0/2}]) \end{array} \right\}. \tag{1.8.31}$$

Some of the definitions and lemmas below are only small modifications of previous definitions and lemmas, respectively. We underline the new pieces to show the differences.

Definition 1.8.7 (Modification of Definition 1.6.6). *We define:*

$$\tilde{B}^{n_0}_{\text{unique fit}} := \left\{ \begin{array}{l} \text{For every } i,j \in \{1,\dots,l^2\}, \text{ every } i\text{-spaced interval} \\ \underline{I \subseteq [-11\cdot 2^{n_0}, 11\cdot 2^{n_0}] \setminus J_1}, \text{ and every } j\text{-spaced interval} \\ \underline{J \subseteq [-11\cdot 2^{n_0}, 11\cdot 2^{n_0}] \setminus J_1} \text{ with } |I| = |J| \geq c_2 n_0 \text{ holds} \\ (\xi\lceil I)_{\leftarrow} \neq (\xi\lceil J)_{\rightarrow}, \text{ and if } I \neq J, \text{ then } (\xi\lceil I)_{\rightarrow} \neq (\xi\lceil J)_{\rightarrow}. \end{array} \right\}, \tag{1.8.32}$$

Lemma 1.8.7 (Modification of Lemma 1.6.18). *There exists constants $c_{18}, c_{19} > 0$ such that the following holds:*

$$\tilde{P}\left[(\tilde{B}^{n_0}_{\text{unique fit}})^{\text{c}}\right] \leq c_{18} e^{-c_{19} n_0}. \tag{1.8.33}$$

Proof. The proof of Lemma 1.6.18 remains literally true when we replace P by $\tilde{P}$, but additionally restrict I and J to be disjoint from J_1, since the distributions of $\xi\lceil(\mathbb{Z}\setminus J_1)$ with respect to $\tilde{P}$ and with respect to P coincide; see Lemma 1.3.2. ■

Lemma 1.8.8. $|\mathsf{Center}_{\text{I}}| \leq 46\cdot 2^{n_0}$, $|\mathsf{Corrupted}| \leq n_0^{41}$, *and thus* $|\mathsf{Center}_{\text{II}}| \leq 50\cdot 2^{n_0}$. *If the event* $\tilde{B}^{n_0}_{\text{unique fit}}$ *holds, then* $|\mathsf{Center}_{\text{III}}| \geq 2^{n_0/3}$.

Proof. The first statement is obvious, since there are at most $23\cdot 2^{n_0}$ choices for the leftmost point of a ladder interval in $[-11\cdot 2^{n_0}, 11\cdot 2^{n_0}]$, and there is the binary choice "left" or "right".

We show $|\mathsf{Corrupted}| \leq n_0^{41}$ next: The number of pairs $(R(0), R(c_1 n_0 - 1)) \in \mathbb{Z}^2$ with $R \in \mathsf{CorPaths}$ is bounded by $(|J_1| + c_1 n_0 l)^2 \leq n_0^{41}$; recall that n_0 was chosen to be large (see Subsection 1.2.1). Furthermore, every such pair gives rise to at most a single element of $\mathsf{Corrupted}$, since different paths $R, R' \in \mathsf{CorPaths}$ with the same starting point and the same end point generate the same word $\xi\circ R = \xi \circ R'$ by Definition (1.8.27). This shows $|\mathsf{Corrupted}| \leq n_0^{41} \leq 4\cdot 2^{n_0}$, since n_0 is large enough by Subsection 1.2.1. Using the definition of $\mathsf{Center}_{\text{II}}$, we obtain $|\mathsf{Center}_{\text{II}}| \leq 50\cdot 2^{n_0}$.

Finally we show $|\mathsf{Center}_{\text{III}}| \geq 2^{n_0/3}$. We observe that $[-2\cdot 2^{n_0/2}, 2\cdot 2^{n_0/2}] \setminus [-2^{n_0/2}, 2^{n_0/2}]$ is disjoint from J_1. Assuming that $\tilde{B}^{n_0}_{\text{unique fit}}$ holds, this implies that all right ladder intervals $I_1, I_2 \subseteq [-2\cdot 2^{n_0/2}, 2\cdot 2^{n_0/2}] \setminus [-2^{n_0/2}, 2^{n_0/2}]$, with $I_1 \neq I_2$ $|I_1| = |I_2| = c_1 n_0 \geq c_2 n_0$ generate pairwise different ladder words $(\xi\lceil I_1)_{\rightarrow} \neq (\xi\lceil I_2)_{\rightarrow}$. Since there are at least $2^{n_0/2} - c_1 n_0 \geq 2^{n_0/3}$ such ladder intervals (n_0 is large enough; see Subsection 1.2.1), there are as least as many ladder words $w \in \mathsf{Center}_{\text{III}}$. ■

Lemma 1.8.9. *For every $x \in \mathbb{Z}$ with $|x| > 2\cdot 2^{n_0} + 2l 2^{12\alpha n_0}$ and for every $t \in [0, 2^{12\alpha n_1}[$ holds:*

$$\begin{aligned} &\tilde{P}\left[S(t) = x,\ |\mathsf{Puzzle}^{n_0}_{\text{IV}}(\theta^t\chi)| \leq 50\cdot 2^{n_0}, \text{ and } |\mathsf{Center}_{\text{II}} \cap \mathsf{Puzzle}^{n_0}_{\text{IV}}(\theta^t\chi)| \geq 2^{n_0/3}\ \right] \\ &\leq \exp\{-2^{n_0/4}\}. \end{aligned} \tag{1.8.34}$$

Proof. We set

$$\mathsf{Outside}_{x,t} := \left\{ \begin{array}{ll} \mathsf{Puzzle}^{n_0}_{\text{IV}}(\theta^t\chi) & \text{if } S(t) = x \text{ and } |\mathsf{Puzzle}^{n_0}_{\text{IV}}(\theta^t\chi)| \leq 50\cdot 2^{n_0}, \\ \varnothing & \text{else.} \end{array} \right. \tag{1.8.35}$$

The random set $\mathsf{Puzzle}_{\mathrm{IV}}^{n_0}(\theta^t\chi)$ only depends on $\chi\lceil[t, t+2\cdot 2^{12\alpha n_0}[$, and the random walk cannot travel a longer distance than $2l2^{12\alpha n_0}$ during the time interval $[t, t+2\cdot 2^{12\alpha n_0}[$. Given $S(t)=x$ and $|x| > 2\cdot 2^{n_0} + 2l2^{12\alpha n_0}$, the random walk S cannot enter the interval $[-2\cdot 2^{n_0}, 2\cdot 2^{n_0}]$ during the time interval $[t, t+2\cdot 2^{12\alpha n_0}[$; thus $\mathsf{Outside}_{x,t}$ depends only on S and $\xi\lceil(\mathbb{Z}\setminus[-2\cdot 2^{n_0}, 2\cdot 2^{n_0}])$. Hence, using Lemma 1.3.2 and $J_1 \subseteq [-2\cdot 2^{n_0}, 2\cdot 2^{n_0}]$, the random piece of scenery $\xi\lceil[-2\cdot 2^{n_0}, 2\cdot 2^{n_0}]$ and the random set $\mathsf{Outside}_{x,t}$ are independent with respect to $\tilde{P}$. Let $\mathcal{I}_{\mathrm{r}}$ denote the set of all right ladder intervals $I \subseteq ([-2\cdot 2^{n_0}, 2\cdot 2^{n_0}] \setminus J_1)$ with $|I| = c_1 n_0$. We define $\mathcal{I}_{\mathrm{l}}$ similarly with "right ladder intervals" replaced by "left ladder intervals". We partition $\mathcal{I}_{\mathrm{r}}$ into $c_1 n_0 l_\rightarrow$ subsets, $\mathcal{I}'_{\mathrm{r}}(1), \ldots, \mathcal{I}'_{\mathrm{r}}(c_1 n_0 l_\rightarrow)$:

$$\mathcal{I}'_{\mathrm{r}}(k) := \{I \in \mathcal{I}_{\mathrm{r}} \mid \min I \in k + c_1 n_0 l_\rightarrow \mathbb{Z}\} \tag{1.8.36}$$

Let $k \in [1, c_1 n_0 l_\rightarrow]$ be fixed. Note that the cardinality $N := |\mathcal{I}'_{\mathrm{r}}(k)|$ fulfills the bounds

$$\frac{2^{n_0}}{c_1 n_0 l_\rightarrow} \le \frac{4\cdot 2^{n_0}}{c_1 n_0 l_\rightarrow} - |J_1| - 2 \le N \le \frac{4\cdot 2^{n_0}}{c_1 n_0 l_\rightarrow}. \tag{1.8.37}$$

Furthermore, the elements of $\mathcal{I}'_{\mathrm{r}}(k)$ are pairwise disjoint; thus the family $(\xi\lceil I)_{I\in\mathcal{I}'_{\mathrm{r}}(k)}$ is i.i.d. and independent of $\mathsf{Outside}_{x,t}$ (with respect to $\tilde{P}$). For $I \in \mathcal{I}_{\mathrm{r}}$, we set $X_I^{\mathrm{r}} := 1$ for $(\xi\lceil I)_\rightarrow \in \mathsf{Outside}_{x,t}$, and $X_I^{\mathrm{r}} := 0$ otherwise. Similarly for $J \in \mathcal{I}_{\mathrm{l}}$, let X_J^{l} denote the indicator function of the event $\{(\xi\lceil J)_\leftarrow \in \mathsf{Outside}_{x,t}\}$. Then, conditioned on a given value of $\mathsf{Outside}_{x,t}$, the Bernoulli random variables X_I^{r}, $I \in \mathcal{I}'_{\mathrm{r}}(k)$, are i.i.d. with respect to $\tilde{P}[\,\cdot\,|\mathsf{Outside}_{x,t}]$. Furthermore we have, using $|\mathsf{Outside}_{x,t}| \le 50\cdot 2^{n_0}$:

$$\tilde{P}[X_I^{\mathrm{r}} = 1 \mid \mathsf{Outside}_{x,t}] \le |\mathsf{Outside}_{x,t}||\mathcal{C}|^{-|I|} \le 50 e^{(\log 2 - c_1 \log|\mathcal{C}|)n_0} =: p. \tag{1.8.38}$$

We set $Y_k^{\mathrm{r}} := \sum_{I\in\mathcal{I}'_{\mathrm{r}}(k)} X_I^{\mathrm{r}}$. Consequently this random variable is stochastically dominated by a Binomial(N,p)–distributed random variable; note that Y_k^{r} is binomially distributed with respect to the conditioned measure $\tilde{P}[\,\cdot\,|\mathsf{Outside}_{x,t}]$. A rough but simple large deviation estimate suffices for our purposes: Using the exponential Chebyshev inequality, we have for $a > 0$ and $\sigma := \log(a/p) > 0$:

$$\begin{aligned}\tilde{P}[Y_k^{\mathrm{r}} \ge Na] &\le E[e^{\sigma Y_k^{\mathrm{r}} - Na}] \le (pe^{\sigma(1-a)} + (1-p)e^{-\sigma a})^N = ((1+a-p)p^a a^{-a})^N \\ &\le (e^a p^a a^{-a})^N = \exp\{Na(1-\log(a/p))\}\end{aligned} \tag{1.8.39}$$

In particular, we obtain for the choice $a = N^{-1}2^{n_0/3}/(4c_1 n_0 l_\rightarrow) \ge 2^{-2n_0/3}/16$ (where we have used (1.8.37)), using (1.8.38): $\sigma = \log(a/p) \ge (c_1 \log|\mathcal{C}| - \frac{5}{3}\log 2)n_0 - \log 800 \ge c_1(\log|\mathcal{C}|)n_0/2 + 1$; the last inequality holds by our choice of c_1 and n_0 (see Subsection 1.2.1). Hence we obtain:

$$\begin{aligned}\tilde{P}\left[\sum_{I\in\mathcal{I}_{\mathrm{r}}} X_I^{\mathrm{r}} \ge \frac{2^{n_0/3}}{4}\right] &\le \sum_{k=1}^{c_1 n_0 l_\rightarrow} \tilde{P}\left[Y_k^{\mathrm{r}} \ge Na\right] \le c_1 n_0 l_\rightarrow \exp\{Na(1-\log(a/p))\} \\ &\le c_1 n_0 l_\rightarrow \exp\left\{-\frac{\log|\mathcal{C}|}{8l_\rightarrow} 2^{n_0/3}\right\} \le \frac{1}{2}\exp\{-2^{n_0/4}\}.\end{aligned} \tag{1.8.40}$$

The same argument works for left ladder intervals, too:

$$\tilde{P}\left[\sum_{J\in\mathcal{I}_{\mathrm{l}}} X_J^{\mathrm{l}} \ge \frac{2^{n_0/3}}{4}\right] \le \frac{1}{2}\exp\{-2^{n_0/4}\}. \tag{1.8.41}$$

Combining (1.8.41), (1.8.40), and $|\mathsf{Corrupted}| \le n_0^{41} \le 2^{n_0/3}/2$ (see Lemma 1.8.8), we obtain

$$\tilde{P}\left[|\mathsf{Center}_{\mathrm{I}} \cap \mathsf{Outside}_{x,t}| \ge 2^{n_0/3}\right] \le \tilde{P}\left[|\mathsf{Center}_{\mathrm{II}} \cap \mathsf{Outside}_{x,t}| \ge \frac{2^{n_0/3}}{2}\right] \tag{1.8.42}$$
$$\le \tilde{P}\left[\sum_{I\in\mathcal{I}_{\mathrm{r}}} X_I^{\mathrm{r}} \ge \frac{2^{n_0/3}}{4}\right] + \tilde{P}\left[\sum_{J\in\mathcal{I}_{\mathrm{l}}} X_J^{\mathrm{l}} \ge \frac{2^{n_0/3}}{4}\right] \le \exp\{-2^{n_0/4}\}.$$

The claim (1.8.34) is an immediate consequence of this bound. ■

Due to the presence of the "modified" part $\xi\lceil J_1$, we define the following modification of $B^{n_0}_{\text{recogn straight}}$, which is a little weaker than the original version:

Definition 1.8.8 (Modification of Definition 1.6.7).

$$\tilde{B}^{n_0}_{\text{recogn straight}} := \tag{1.8.43}$$
$$\left\{\begin{array}{l}\text{For every } R \in \mathsf{AdPaths}(11\cdot 2^{n_0}, c_1 n_0) \text{ with } R(c_1n_0-1) - R(0) \notin \{(c_1n_0 - \\ 1)l_{\rightarrow}, (c_1n_0-1)l_{\leftarrow}\} \text{ there is } \bar{R} \in \mathsf{AdPaths}(12\cdot 2^{n_0}, c_1n_0) \text{ such that } R(0) = \bar{R}(0), \\ R(c_1n_0-1) = \bar{R}(c_1n_0-1), \text{ and } (\underline{\bar{R} \text{ takes at least one value in } J_1}, \text{ or } \xi\circ R \neq \\ \xi\circ\bar{R}).\end{array}\right\},$$

$$\tilde{E}^{n_0}_{\text{only ladder}} := \left\{\begin{array}{l}\text{For all } (w_1,w_2,w_3) \in \mathsf{Puzzle}^{n_0}_{\mathrm{I}}(\mathsf{Input}) \text{ and every admissible} \\ \text{piece of path } R : [0, 3c_1n_0[\to [-11\cdot 2^{n_0}, 11\cdot 2^{n_0}] \text{ with } \xi\circ R = \\ w_1w_2w_3 \text{ holds: } w_2 \text{ is a ladder word of } \xi\lceil[-11\cdot 2^{n_0}, 11\cdot 2^{n_0}], \\ \underline{\text{or } w_2 \in \mathsf{Corrupted}}.\end{array}\right\}. \tag{1.8.44}$$

Lemma 1.8.10 (Modification of Lemma 1.6.22). *There exist positive constants c_{25} and c_{26} not depending on n_0 such that:*

$$\tilde{P}\left[(\tilde{B}^{n_0}_{\text{recogn straight}})^{\mathrm{c}}\right] \le c_{25}e^{-c_{26}n_0}. \tag{1.8.45}$$

Proof. The proof of Lemma 1.6.22 requires only a small modification: Given $R \in \mathsf{AdPaths}(11\cdot 2^{n_0}, c_1n_0)$, there are two cases: Either some $\bar{R} \in \mathsf{AdPaths}(12\cdot 2^{n_0}, c_1n_0)$ with $\bar{R}(0) = R(0)$ and $\bar{R}(c_1n_0-1) = R(c_1n_0-1)$ touches J_1 ("case 1"), or no such $\bar{R}$ touches J_1 ("case 2").

- In the first case, the underlined new condition in definition (1.8.43) of $\tilde{B}^{n_0}_{\text{recogn straight}}$ is certainly satisfied.
- In the second case, we may proceed further just as in the proof of Lemma 1.6.22, with P replaced by $\tilde{P}$: formula (1.6.84) remains true in this case, since neither $R\lceil I(R)$ nor $\bar{R}\lceil I(R)$ touches J_1. Recall that $\xi\lceil(\mathbb{Z}\setminus J_1)$ has the same distribution with respect to P as with $\tilde{P}$.

The rest of the proof of Lemma 1.6.22 still remains true when we replace P by $\tilde{P}$ but remove all paths R from $\mathsf{AdPaths}(11\cdot 2^{n_0}, c_1n_0)$ and $\mathsf{AdPaths}(12\cdot 2^{n_0}, c_1n_0)$ that belong to the first case. ■

Lemma 1.8.11 (Modification of Lemma 1.6.9). *We have*

$$B^{n_0}_{\text{all paths}} \cap \tilde{B}^{n_0}_{\text{recogn straight}} \subseteq \tilde{E}^{n_0}_{\text{only ladder}}. \tag{1.8.46}$$

Proof. We describe the modifications required in the proof of Lemma 1.6.9: Assume the event $B^{n_0}_{\text{all paths}} \cap \tilde{B}^{n_0}_{\text{recogn straight}}$ holds, and let $w_1w_2w_3 \in \mathsf{Puzzle}^{n_0}_{\mathrm{I}}(\mathsf{Input})$, and $R \in \mathsf{AdPaths}(11 \cdot 2^{n_0}, 3c_1n_0)$, $\xi \circ R = w_1w_2w_3$ as in the proof of Lemma 1.6.9. Again, we prove by contradiction that $\tilde{E}^{n_0}_{\text{only ladder}}$ holds: Assume that w_2 is not a ladder word of $\xi\lceil[-11 \cdot 2^{n_0}, 11 \cdot 2^{n_0}]$ and $w_2 \notin \mathsf{Corrupted}$. We distinguish two cases: Either the middle piece $R\lceil[c_1n_0, 2c_1n_0[$ of R belongs to $\mathsf{CorPaths}$ when being time-shifted back to the origin ("case 1"), or it does not ("case 2").

- In case 1, $w_2 = (\xi \circ R\lceil[c_1n_0, 2c_1n_0[)_\to \in \mathsf{Corrupted}$ by Definition (1.8.28), which contradicts our assumption.
- In case 2, using Definition (1.8.27), there is an admissible piece of path $R' : [c_1n_0, 2c_1n_0[\to \mathbb{Z}$ with $R'(c_1n_0) = R(c_1n_0)$ and $R'(2c_1n_0 - 1) = R(2c_1n_0 - 1)$ such that $w_2' := (\xi \circ R')_\to \neq (\xi \circ R\lceil[c_1n_0, 2c_1n_0[)_\to$ ("case 2.1"), or all admissible paths $R' : [c_1n_0, 2c_1n_0[\to \mathbb{Z}$ with $R'(c_1n_0) = R(c_1n_0)$ and $R'(2c_1n_0 - 1) = R(2c_1n_0 - 1)$ do not touch J_1 and fulfill $\xi \circ R' = \xi \circ R\lceil[c_1n_0, 2c_1n_0[$ ("case 2.2").
  - In case 2.1 we proceed just as in the proof of Lemma 1.6.9; this yields the contradiction $w_1w_2w_3 \notin \mathsf{Puzzle}^{n_0}_{\mathrm{I}}(\mathsf{Input})$.
  - In case 2.2, we use that $R\lceil[c_1n_0, 2c_1n_0[$ is not a ladder path, since w_2 is not a ladder word of $\xi\lceil[-11 \cdot 2^{n_0}, 11 \cdot 2^{n_0}]$. Using Definition (1.8.43), this case is contradictory, too, since $\tilde{B}^{n_0}_{\text{recogn straight}}$ holds.

Thus all cases lead to a contradiction; this proves the Lemma. ■

Lemma 1.8.12. *If $\tilde{E}^{n_0}_{\text{only ladder}} \cap E^0_{\text{stop},T^0}$ holds, then $\mathsf{Puzzle}^{n_0}_{\mathrm{IV}}(\chi) \subseteq \mathsf{Center}_{\mathrm{II}}$.*

Proof. Assume that $\tilde{E}^{n_0}_{\text{only ladder}} \cap E^0_{\text{stop},T^0}$ holds, and let $w_2 \in \mathsf{Puzzle}^{n_0}_{\mathrm{IV}}(\chi)$. Take $w_1, w_3 \in \mathcal{C}^{c_1n_0}$ with $(w_1, w_2, w_3) \in \mathsf{Puzzle}^{n_0}_{\mathrm{III}}(\chi)$ by (1.8.24). Then by (1.8.23), $(w_1, w_2, w_3) \in \mathsf{Puzzle}^{n_0}_{\mathrm{I}}(\mathsf{Input})$, and $w_1w_2w_3$ occurs in the observations χ at most $2^{n_0}/l$ time steps after some stopping time $T^0_k(\chi)$, $0 \le k < 2^{\alpha n_0}$. Since E^0_{stop,T^0} holds, we have $|S(T^0_k)| \le 2^{n_0}$; thus $w_1w_2w_3$ is read in χ while the random walk follows some admissible piece of path R with values in $[-2 \cdot 2^{n_0}, 2 \cdot 2^{n_0}] \subseteq [-11 \cdot 2^{n_0}, 11 \cdot 2^{n_0}]$. Since $\tilde{E}^{n_0}_{\text{only ladder}}$ holds, this implies: w_2 is a ladder word of $\xi\lceil[-11 \cdot 2^{n_0}, 11 \cdot 2^{n_0}]$, or $w_2 \in \mathsf{Corrupted}$. In the next argument, we use the following fact: If π is a ladder path and $\bar{\pi}$ is an admissible piece of path with the same length, starting point, and end point as π, then $\bar{\pi} = \pi$. Using this fact and the Definitions (1.8.27) and (1.8.28), we see: if w_2 is a ladder word of $\xi\lceil[-11 \cdot 2^{n_0}, 11 \cdot 2^{n_0}]$, but not of $\xi\lceil([-11 \cdot 2^{n_0}, 11 \cdot 2^{n_0}] \setminus J_1)$, then $w_2 \in \mathsf{Corrupted}$, too. Thus we obtain $w_2 \in \mathsf{Center}_{\mathrm{II}}$. This proves the lemma. ■

Lemma 1.8.13 (Modification of Lemma 1.6.19). *There exist constants $c_{20}, c_{21} > 0$ such that:*

$$\tilde{P}\left[(B^{n_0}_{\text{all paths}})^{\mathrm{c}} \cap E^0_{\text{stop},T^0}\right] \le c_{21}e^{-c_{20}n_0}. \tag{1.8.47}$$

Proof. Again, the proof of Lemma 1.6.19 remains literally true when we replace P by $\tilde{P}$ and τ by T^0; in particular note that the event $B^{n_0}_{\text{all paths}}$ depends only on the random walk S, but not on the scenery ξ, and the strong Markov property for S still holds with respect to $\tilde{P}$. ■

Recall Definition (1.7.1):

$$E^1_{\text{no error},T^1} = \{\forall k \in \mathbb{N} : \text{ If } T^1_k(\chi) < 2^{12\alpha n_1}, \text{ then } |S(T^1_k(\chi))| \le 2^{n_1}\}.$$

Lemma 1.8.14. *For some constants $c_{48}, c_{49} > 0$ holds* $\tilde{P}\left[E^1_{\text{no error},T^1}\right] \geq 1 - c_{48}e^{-c_{49}n_0}$.

Proof. Using Definition 1.8.5 of T^1 and Lemmas 1.8.11, 1.8.12, and 1.8.9, we obtain (see also the explanations below):

$$\begin{aligned}
&\tilde{P}\left[E^1_{\text{no error},T^1}\right] \qquad (1.8.48)\\
&\geq \tilde{P}\left[\begin{array}{l}\text{For all } t \in [0, 2^{12\alpha n_1}[\text{ holds: if } |\mathsf{Puzzle}^{n_0}_{\text{IV}}(\chi) \cap \mathsf{Puzzle}^{n_0}_{\text{IV}}(\theta^t\chi)| \geq 2^{n_0/3} \text{ and}\\ |\mathsf{Puzzle}^{n_0}_{\text{IV}}(\theta^t\chi)| \leq 50 \cdot 2^{n_0}, \text{ then } |S(t + 2^{12\alpha n_0})| \leq 2^{n_1}\end{array}\right]\\
&\geq \tilde{P}[\mathsf{Puzzle}^{n_0}_{\text{IV}}(\chi) \subseteq \mathsf{Center}_{\text{II}}] - \sum_{t=0}^{2^{12\alpha n_1}-1} \tilde{P}\left[\begin{array}{l}|\mathsf{Center}_{\text{II}} \cap \mathsf{Puzzle}^{n_0}_{\text{IV}}(\theta^t\chi)| \geq 2^{n_0/3},\\ |\mathsf{Puzzle}^{n_0}_{\text{IV}}(\theta^t\chi)| \leq 50 \cdot 2^{n_0}, \text{ and}\\ 2\cdot 2^{n_0} + 2l2^{12\alpha n_0} < |S(t)| \leq l2^{12\alpha n_1}\end{array}\right]\\
&\geq \tilde{P}[B^{n_0}_{\text{all paths}} \cap \tilde{B}^{n_0}_{\text{recogn straight}} \cap E^0_{\text{stop},T^0}] - 2^{12\alpha n_1} \cdot 2l2^{12\alpha n_1} \cdot \exp\{-2^{n_0/4}\}\\
&\geq \tilde{P}[E^0_{\text{stop},T^0}] - \tilde{P}[(B^{n_0}_{\text{all paths}})^{\text{c}} \cap E^0_{\text{stop},T^0}] - \tilde{P}[(\tilde{B}^{n_0}_{\text{recogn straight}})^{\text{c}}]\\
&\quad - 2l\exp\{24(\log 2)\alpha n_1 - 2^{n_0/4}\}\\
&\geq 1 - c_{48}e^{-c_{49}n_0}
\end{aligned}$$

for some constants $c_{48}, c_{49} > 0$. For the second inequality in (1.8.48), note that the random walk S cannot travel farther than $l2^{12\alpha n_1}$ within time $2^{12\alpha n_1}$; thus $|S(t + 2^{12\alpha n_0})| \leq 2^{n_1}$ or $2 \cdot 2^{n_0} + 2l2^{12\alpha n_0} < |S(t)| \leq l2^{12\alpha n_1}$ holds for all $t \in [0, 2^{12\alpha n_1}[$. In the last step of (1.8.48) we used Theorem 1.8.1, Lemmas 1.8.10 and 1.8.13, and the fact $n_1 = 2^{\lfloor\sqrt{n_0}\rfloor}$ (recall Definition 1.3.2); especially $\exp\{24(\log 2)\alpha n_1 - 2^{n_0/4}\}$ is superexponentially small in n_0. The constants c_{48} and c_{49} need not depend on α, since n_0 was chosen large and α-dependent (see Subsection 1.2.1). ■

Definition 1.8.9. *We define the event*

$$B^{n_0}_{\text{all paths II}} := \left\{ \begin{array}{l}\forall R \in \mathsf{AdPaths}(3 \cdot 2^{n_0/2}, 3c_1n_0)\ \exists k \in [0, 2^{\alpha n_0}[\ \exists j \in [0, 2^{n_0}/l] :\\ \mathsf{TimeShift}^{T^0_k(\chi)+j}(R) \subseteq S\end{array} \right\}. \qquad (1.8.49)$$

Lemma 1.8.15 (Yet another modification of Lemma 1.6.19). *There exist constants $c_{50}, c_{51} > 0$ not depending on n_0 such that:*

$$\tilde{P}\left[(B^{n_0}_{\text{all paths II}})^{\text{c}} \cap E^0_{\text{stop},T^0}\right] \leq c_{50}e^{-c_{51}n_0}. \qquad (1.8.50)$$

Proof. The proof is almost the same as the proof of Lemma 1.6.19; we only explain the differences. This time, some of the parameters in the proof of Lemma 1.6.19 must be changed: We replace (1.6.70) by $B^{n,k}_R := \left\{\exists j \in [0, 2^{n_0}/l] :\ \mathsf{TimeShift}^{\tau(k)+j}(R)\right\}$ and $\mathsf{AdPath}(12 \cdot 2^n, 3c_1n)$ by $\mathsf{AdPaths}(3 \cdot 2^{n_0/2}, 3c_1n_0)$; then the estimate (1.6.75) is replaced by $|\mathsf{AdPaths}(3 \cdot 2^{n_0/2}, 3c_1n_0)| \leq 7 \cdot 2^{n_0/2}|\mathcal{M}|^{3c_1n_0}$. Then (1.6.78) changes to

$$\inf_{|x|\leq 3\cdot 2^{n_0/2}} \tilde{P}\left[S(\tau(k) + j) = x \text{ for some } j \in [0, 2^{n_0}/l] \;\middle|\; C^{n_0,k}_R\right] \geq c_{52}2^{-n_0/2} \qquad (1.8.51)$$

for some constant $c_{52} > 0$. Hence the right hand side of (1.6.79) gets replaced by the bound $c_{52}2^{-n_0/2}\mu^{3c_1n_0}_{\min}$. We end up with the following modified version of (1.6.80):

$$\begin{aligned}
&\tilde{P}\left[E^0_{\text{stop},T^0} \setminus B^{n_0}_{\text{all paths II}}\right] \qquad (1.8.52)\\
&\leq 7\exp\left\{n_0\left(\frac{\log 2}{2} + 3c_1\log|\mathcal{M}|\right) - c_{52}e^{n_0(\alpha\log 2 + 3c_1\log\mu_{\min} - (\log 2)/2)}\right\},
\end{aligned}$$

which still converges superexponentially fast to 0 as $n_0 \to \infty$. This proves the Lemma. ■

Definition 1.8.10 (Modification of Definition 1.6.2).

$$\tilde{B}^{n_0}_{\text{signals}} := \left\{ \begin{array}{l} \text{For every right ladder path } \pi \in ([-2l2^{2n_0}, 2l2^{2n_0}] \setminus J_1)^{[0,c_1n_0/2[} \\ \text{and for every admissible piece of path } \pi' \in \\ \mathsf{AdPath}(2l2^{2n_0}, c_1n_0/2): \\ \text{If } \xi \circ \pi = \xi \circ \pi', \text{ then } \pi(0) \le \pi'(0) \text{ and } \pi(c_1n_0/2-1) \ge \\ \pi'(c_1n_0/2-1). \end{array} \right\}, \tag{1.8.53}$$

$$\tilde{E}^{n_0}_{\text{signals II}} := \left\{ \begin{array}{l} \text{For every ladder path } \pi \in ([-2l2^{2n_0}, 2l2^{2n_0}] \setminus J_1)^{[0,c_1n_0[} \text{ and} \\ \text{for every admissible piece of path } \pi' \in \mathsf{AdPath}(2l2^{2n_0}, c_1n_0): \\ \text{If } \xi \circ \pi = \xi \circ \pi', \text{ then } \pi(c_1n_0/2) = \pi'(c_1n_0/2). \end{array} \right\}. \tag{1.8.54}$$

Note that π' in the last two definitions may well have some of its values in J_1.

Lemma 1.8.16 (Modification of Lemma 1.6.23). *There exist constants $c_{28} > 0$, $c_{29} > 0$ not depending on n_0 such that:*

$$\tilde{P}\left[(\tilde{B}^{n_0}_{\text{signals}})^{\text{c}}\right] \le c_{29}e^{-c_{28}n_0}. \tag{1.8.55}$$

Proof. The proof of Lemma 1.6.23 remains literally true when we consider paths

$$\pi \in ([-2l2^{2n_0}, 2l2^{2n_0}] \setminus J_1)^{[0,c_1n_0[} \tag{1.8.56}$$

only, but replace P by $\tilde{P}$. Note that in the induction step in proof of Lemma 1.6.23 $\xi \circ \pi(j)$ is independent of the family $(\xi \circ \pi\lceil I', \xi \circ \pi'\lceil I')$ with respect to $\tilde{P}$, too, even if π' touches the "corrupted" domain J_1; see (1.6.90) and a few lines before this formula. This is true because π does not touch J_1, and $\xi\lceil J_1$ is independent of $\xi\lceil(\mathbb{Z} \setminus J_1)$ by Lemma 1.3.2. Thus formula (1.6.90) remains true when P is replaced by $\tilde{P}$. ■

Lemma 1.8.17 (Modification of Lemma 1.6.2). $\tilde{B}^{n_0}_{\text{signals}} \subseteq \tilde{E}^{n_0}_{\text{signals II}}$.

Proof. When we consider paths π only which do not to touch J_1, the proof of Lemma 1.6.2 remains literally true in this modified case, too. ■

Lemma 1.8.18 (Modification of Lemma 1.6.3).
Assume that the event $B^{n_0}_{\text{all paths}} \cap \tilde{B}^{n_0}_{\text{signals}} \cap E^0_{\text{stop},T^0}$ holds. Let $I \subseteq [-6 \cdot 2^{n_0}, 6 \cdot 2^{n_0}] \setminus J_1$ be a right ladder interval with $|I| = 3c_1n_0$, and let $w_1, w_2, w_3 \in \mathcal{C}^{c_1n_0}$ with $(\xi\lceil I)_\rightarrow = w_1w_2w_3$. Then $(w_1, w_2, w_3) \in \mathsf{Puzzle}^{n_0}_{\text{I}}(\mathsf{Input})$.

Proof. The proof of Lemma 1.6.3 remains literally true when we consider only intervals I that are disjoint from J_1, $n = n_0$, and replace $B^{n_0}_{\text{signals}}$ and $E^{n_0}_{\text{signals II}}$ by their modified versions; note that the ladder path R in the proof of Lemma 1.6.3 does not touch J_1, but we need not assume this for R'. ■

Lemma 1.8.19. *If the event $B^{n_0}_{\text{all paths}} \cap B^{n_0}_{\text{all paths II}} \cap \tilde{B}^{n_0}_{\text{signals}} \cap E^0_{\text{stop},T^0}$ holds, then $\mathsf{Center}_{\text{III}} \subseteq \mathsf{Puzzle}^{n_0}_{\text{IV}}(\chi)$.*

Proof. Assume that $B^{n_0}_{\text{all paths}} \cap B^{n_0}_{\text{all paths II}} \cap \tilde{B}^{n_0}_{\text{signals}} \cap E^0_{\text{stop},T^0}$ holds, and let $w_2 \in \mathsf{Center}_{\text{III}}$. Then $w_2 = (\xi\lceil I)_\rightarrow$ for some right ladder interval $I \subseteq [-2 \cdot 2^{n_0/2}, 2 \cdot 2^{n_0/2}] \setminus [-2^{n_0/2}, 2^{n_0/2}]$, $|I| = c_1n_0$. We take the larger right ladder interval $I' \supseteq I$, $|I'| = 3c_1n_0$, with c_1n_0 extra points to the left of I and another c_1n_0 extra points to the right of I; then

$I' \subseteq [-3\cdot 2^{n_0/2}, 3\cdot 2^{n_0/2}] \setminus J_1 \subseteq [-6\cdot 2^{n_0}, 6\cdot 2^{n_0}] \setminus J_1$; note that $\operatorname{dist}(J_1, \mathbb{Z}\setminus[-2^{n_0/2}, 2^{n_0/2}]) > c_1 n_0 l$ and $\operatorname{dist}([-2\cdot 2^{n_0/2}, 2\cdot 2^{n_0/2}], \mathbb{Z}\setminus[-3\cdot 2^{n_0/2}, 3\cdot 2^{n_0/2}]) > c_1 n_0 l$; recall that n_0 is chosen large enough (Subsection 1.2.1). Then $(\xi\lceil I')_\to = w_1 w_2 w_3$ for some $w_1, w_3 \in \mathcal{C}^{c_1 n_0}$, and Lemma 1.8.18 implies $(w_1, w_2, w_3) \in \mathsf{Puzzle}_{\mathrm{I}}^{n_0}(\mathsf{Input})$. Let R denote the (unique) right ladder path $R : [0, 3c_1[\to I'$. Since $B^{n_0}_{\text{all paths II}}$ holds, the random walk S follows R (time-shifted) at most $2^{n_0}/l$ time steps after some stopping time $T^0_k(\chi)$, $k \in [0, 2^{\alpha n_0}[$. Then $\xi \circ R = w_1 w_2 w_3$; thus $(w_1, w_2, w_3) \in \mathsf{Puzzle}^{n_0}_{\mathrm{III}}(\chi)$ by Definition (1.8.23); hence $w_2 \in \mathsf{Puzzle}^{n_0}_{\mathrm{IV}}(\chi)$ by Definition (1.8.24). This proves the lemma. ■

Definition 1.8.11. *We set*

$$E^1_{\text{center}} := \left\{ |\mathsf{Center}_{\mathrm{III}} \cap \mathsf{Puzzle}^{n_0}_{\mathrm{IV}}(\chi)| \geq 2^{n_0/3} \text{ and } |\mathsf{Puzzle}^{n_0}_{\mathrm{IV}}(\chi)| \leq 50\cdot 2^{n_0} \right\}, \tag{1.8.57}$$

$$\Xi^1_{\text{center}} := \left\{ \xi \in \mathcal{C}^{\mathbb{Z}} \,\middle|\, P[E^1_{\text{center}} \mid \xi] \geq \frac{1}{2} \right\}. \tag{1.8.58}$$

The sets E^1_{center} and Ξ^1_{center} play an analogous role for the stopping times T^1 as $E^m_{\text{reconst},f}$ and $\Xi^m_{\text{reconst},f}$ play for the "higher level" stopping times in Section 1.7.

Lemma 1.8.20. *For some positive constants c_{53} and c_{54} holds $\tilde{P}[\xi \in \Xi^1_{\text{center}}] \geq 1 - c_{53}e^{-c_{54}n_0}$.*

Proof. If the events $B^{n_0}_{\text{all paths}}$, $B^{n_0}_{\text{all paths II}}$, $\tilde{B}^{n_0}_{\text{signals}}$, E^0_{stop,T^0}, $\tilde{E}^{n_0}_{\text{only ladder}}$, and $\tilde{B}^{n_0}_{\text{unique fit}}$ hold, then we have $|\mathsf{Center}_{\mathrm{III}} \cap \mathsf{Puzzle}^{n_0}_{\mathrm{IV}}(\chi)| = |\mathsf{Center}_{\mathrm{III}}| \geq 2^{n_0/3}$ and $|\mathsf{Puzzle}^{n_0}_{\mathrm{IV}}(\chi)| \leq |\mathsf{Center}_{\mathrm{II}}| \leq 50\cdot 2^{n_0}$ by Lemmas 1.8.12, 1.8.19, and 1.8.8. By Lemma 1.8.11, we can replace $\tilde{E}^{n_0}_{\text{only ladder}}$ in the above list of events by $\tilde{B}^{n_0}_{\text{recogn straight}}$. Thus we have

$$\tilde{P}\left[E^1_{\text{center}}\right] \geq \tilde{P}\left[B^{n_0}_{\text{all paths}} \cap B^{n_0}_{\text{all paths II}} \cap \tilde{B}^{n_0}_{\text{signals}} \cap E^0_{\text{stop},T^0} \cap \tilde{B}^{n_0}_{\text{recogn straight}} \cap \tilde{B}^{n_0}_{\text{unique fit}}\right] \tag{1.8.59}$$

$$\geq 1 - c_{55}e^{-c_{54}n_0}$$

for some positive constants c_{55} and c_{54} by Theorem 1.8.1 and Lemmas 1.8.13, 1.8.15, 1.8.10, 1.8.16, and 1.8.7.

Hence we obtain the following:

$$\frac{1}{2}\tilde{P}[\xi \notin \Xi^1_{\text{center}}] = \frac{1}{2}\tilde{P}\left[P[(E^1_{\text{center}})^{\mathrm{c}} \mid \xi] > \frac{1}{2}\right] \leq \tilde{P}[(E^1_{\text{center}})^{\mathrm{c}}] \leq c_{55}e^{-c_{54}n_0}; \tag{1.8.60}$$

recall that $\tilde{P}[\cdot|\xi]$ and $P[\cdot|\xi]$ coincide. ■

Definition 1.8.12 (Yet another modification of Definition 1.7.2).
Let $v(k)$ denote again the $(k+1)$st visit of S to the origin. We define

$$\mathbb{T}^{1\prime}(\xi, \chi) := \left\{ t \in \mathbb{N} \,\middle|\, \begin{array}{l} |\mathsf{Center}_{\mathrm{III}} \cap \mathsf{Puzzle}^{n_0}_{\mathrm{IV}}(\theta^t\chi)| \geq 2^{n_0/3} \\ \text{and } |\mathsf{Puzzle}^{n_0}_{\mathrm{IV}}(\theta^t\chi)| \leq 50\cdot 2^{n_0} \end{array} \right\}, \tag{1.8.61}$$

$$E^1_{\text{when back recog}} := \left\{ \begin{array}{l} \text{For more than } 1/4 \text{ of the points } k \in [0, 2^{2\alpha n_1}[\\ \text{holds } v(k2^{\alpha n_1}) \in \mathbb{T}^{1\prime}(\xi, \chi) \end{array} \right\}. \tag{1.8.62}$$

Lemma 1.8.21 (Yet another analogue to Lemma 1.7.2). *If the event $B^{n_0}_{\text{all paths}} \cap B^{n_0}_{\text{all paths II}} \cap \tilde{B}^{n_0}_{\text{signals}} \cap E^0_{\text{stop},T^0}$ holds, then $\mathbb{T}^1(\chi) \supseteq \mathbb{T}^{1\prime}(\xi, \chi) \cap [0, 2^{12\alpha n_1} - 2\cdot 2^{12\alpha n_0}[$.*

Proof. Assuming that the event $B^{n_0}_{\text{all paths}} \cap B^{n_0}_{\text{all paths II}} \cap \tilde{B}^{n_0}_{\text{signals}} \cap E^0_{\text{stop}}$ holds, we know $\mathsf{Center}_{\text{III}} \subseteq \mathsf{Puzzle}^{n_0}_{\text{IV}}(\chi)$ by Lemma 1.8.19; thus $|\mathsf{Puzzle}^{n_0}_{\text{IV}}(\chi) \cap \mathsf{Puzzle}^{n_0}_{\text{IV}}(\theta^t\chi)| \geq |\mathsf{Center}_{\text{III}} \cap \mathsf{Puzzle}^{n_0}_{\text{IV}}(\theta^t\chi)|$ for all t. This implies the claim $\mathbb{T}^1(\chi) \supseteq \mathbb{T}^{1\prime}(\xi,\chi) \cap [0, 2^{12\alpha n_1} - 2 \cdot 2^{12\alpha n_0}[$ of the lemma; recall Definition (1.8.25) of $\mathbb{T}^1(\chi)$. ■

Lemma 1.8.22 (Yet another modification of Lemma 1.7.3). *Assume that the events* $E^1_{\text{no error},T^1} \cap E^1_{\text{enough back}} \cap E^1_{\text{when back recog}}$ *and* $\mathbb{T}^1(\chi) \supseteq \mathbb{T}^{1\prime}(\xi,\chi) \cap [0, 2^{12\alpha n_1} - 2 \cdot 2^{12\alpha n_0}[$ *hold. Then* E^1_{stop,T^1} *holds, too.*

Proof. Replacing $\mathbb{T}^{(\prime)}_f$ by $\mathbb{T}^{1(\prime)}$ and T_f by T^1, the proof of Lemma 1.7.3 remains literally true. ■

Lemma 1.8.23 (Yet another modification of Lemma 1.7.4). *We have the bound*

$$\tilde{P}\left[\left(E^1_{\text{when back recog}}\right)^{\text{c}} \cap \left\{\xi \in \Xi^1_{\text{center}}\right\}\right] \leq 0.9^{2^{2\alpha n_1}}. \tag{1.8.63}$$

Proof. The proof of Lemma 1.7.4 remains literally true when one replaces $E^{m+1}_{\text{when back recog},f}$ by $E^1_{\text{when back recog}}$, $\Xi^m_{\text{reconst},f}$ by Ξ^1_{center}, $m+1$ by 1, and P by $\tilde{P}$; recall $P[\cdot|\xi] = \tilde{P}[\cdot|\xi]$. ■

Proof of Theorem 1.3.4. By Lemmas 1.8.21 and 1.8.22 we know

$$\begin{aligned} &E^1_{\text{no error},T^1} \cap E^1_{\text{enough back}} \cap E^1_{\text{when back recog}} \cap B^{n_0}_{\text{all paths}} \cap B^{n_0}_{\text{all paths II}} \cap \tilde{B}^{n_0}_{\text{signals}} \cap E^0_{\text{stop},T^0} \\ &\subseteq E^1_{\text{stop},T^1}. \end{aligned} \tag{1.8.64}$$

Since $E^1_{\text{enough back}}$ depends only on S but not on ξ, we have

$$\tilde{P}[(E^1_{\text{enough back}})^{\text{c}}] = P[(E^1_{\text{enough back}})^{\text{c}}]. \tag{1.8.65}$$

Thus, using Lemmas 1.8.14, 1.7.1, 1.8.23, 1.8.20, 1.8.13, 1.8.15, 1.8.16, and Theorem 1.8.1, we know

$$\begin{aligned} &\tilde{P}[(E^1_{\text{stop},T^1})^{\text{c}}] \\ &\leq \tilde{P}[(E^1_{\text{no error},T^1})^{\text{c}}] + \tilde{P}[(E^1_{\text{enough back}})^{\text{c}}] + \tilde{P}\left[\left(E^1_{\text{when back recog}}\right)^{\text{c}} \cap \left\{\xi \in \Xi^1_{\text{center}}\right\}\right] \\ &\quad + \tilde{P}\left[\xi \notin \Xi^1_{\text{center}}\right] + \tilde{P}[(B^{n_0}_{\text{all paths}})^{\text{c}} \cap E^0_{\text{stop},T^0}] + \tilde{P}[(B^{n_0}_{\text{all paths II}})^{\text{c}} \cap E^0_{\text{stop},T^0}] \\ &\quad + \tilde{P}[(\tilde{B}^{n_0}_{\text{signals}})^{\text{c}}] + \tilde{P}[(E^0_{\text{stop}})^{\text{c}}] \\ &\leq c_{48}e^{-c_{49}n_0} + c_{35}2^{-\alpha n_1} + 0.9^{2^{2\alpha n_1}} + c_{53}e^{-c_{54}n_0} + c_{21}e^{-c_{20}n_0} \\ &\quad + c_{50}e^{-c_{51}n_0} + c_{29}e^{-c_{28}n_0} + c_{36}e^{-c_{37}n_0} \\ &\leq e^{-c_4 n_0}, \end{aligned} \tag{1.8.66}$$

since n_0 is chosen large enough (see Subsection 1.2.1) ■

References

[1] Itai Benjamini and Harry Kesten. Distinguishing sceneries by observing the scenery along a random walk path. *J. Anal. Math.*, 69:97–135, 1996.

[2] Krzysztof Burdzy. Some path properties of iterated Brownian motion. In *Seminar on Stochastic Processes, 1992 (Seattle, WA, 1992)*, pages 67–87. Birkhäuser Boston, Boston, MA, 1993.

[3] Frank den Hollander and Jeffrey E. Steif. Mixing properties of the generalized T, T^{-1}-process. *J. Anal. Math.*, 72:165–202, 1997.

[4] W. Th. F. den Hollander. Mixing properties for random walk in random scenery. *Ann. Probab.*, 16(4):1788–1802, 1988.

[5] Richard Durrett. *Probability: theory and examples.* Duxbury Press, Belmont, CA, second edition, 1996.

[6] C. Douglas Howard. Detecting defects in periodic scenery by random walks on $\mathbb{Z}$. *Random Structures Algorithms*, 8(1):59–74, 1996.

[7] C. Douglas Howard. Orthogonality of measures induced by random walks with scenery. *Combin. Probab. Comput.*, 5(3):247–256, 1996.

[8] C. Douglas Howard. Distinguishing certain random sceneries on $\mathbb{Z}$ via random walks. *Statist. Probab. Lett.*, 34(2):123–132, 1997.

[9] Steven Arthur Kalikow. T, T^{-1} transformation is not loosely Bernoulli. *Ann. of Math. (2)*, 115(2):393–409, 1982.

[10] M. Keane and W. Th. F. den Hollander. Ergodic properties of color records. *Phys. A*, 138(1-2):183–193, 1986.

[11] Harry Kesten. Detecting a single defect in a scenery by observing the scenery along a random walk path. In *Itô's stochastic calculus and probability theory*, pages 171–183. Springer, Tokyo, 1996.

[12] Harry Kesten. Distinguishing and reconstructing sceneries from observations along random walk paths. In *Microsurveys in discrete probability (Princeton, NJ, 1997)*, pages 75–83. Amer. Math. Soc., Providence, RI, 1998.

[13] A. Lenstra and H. Matzinger. Reconstructing a 4-color scenery by observing it along a recurrent random walk path with unbounded jumps. In preparation, Eurandom, 2001.

[14] Elon Lindenstrauss. Indistinguishable sceneries. *Random Structures Algorithms*, 14(1):71–86, 1999.

[15] M. Löwe and H. Matzinger. Reconstruction of sceneries with correlated colors. Eurandom Report 99-032, Eurandom, 1999. Submitted to *Stochastic Processes and their Applications*.

[16] M. Löwe and H. Matzinger. Scenery reconstruction in two dimensions with many colors. Eurandom Report 99-018, Eurandom, 1999. Submitted to *The Annals of Applied Probability*.

[17] H. Matzinger. *Reconstructing a 2-color scenery by observing it along a simple random walk path with holding.* PhD-thesis, Cornell University, 1999.

[18] H. Matzinger. Reconstructing a 2-color scenery by observing it along a simple random walk path. Eurandom Report 2000-003, Eurandom, 2000. Submitted to *The Annals of Applied Probability*.

[19] H. Matzinger. Reconstructing a 2-color scenery in polynomial time by observing it along a simple random walk path with holding. Eurandom Report 2000-002, Eurandom, 2000. Submitted to *Probability Theory and Related Fields*.

[20] Heinrich Matzinger. Reconstructing a three-color scenery by observing it along a simple random walk path. *Random Structures Algorithms*, 15(2):196–207, 1999.

[21] Frank Spitzer. *Principles of random walks*. Springer-Verlag, New York, second edition, 1976. Graduate Texts in Mathematics, Vol. 34.

Chapter 2

Reconstructing a random scenery observed with random errors along a random walk path

Probab. Theory Related Fields, 125(4):539–577, 2003.

By Heinrich Matzinger, and Silke Rolles,

We show that an i.i.d. uniformly colored scenery on $\mathbb{Z}$ observed along a random walk path with bounded jumps can still be reconstructed if there are some errors in the observations. We assume the random walk is recurrent and can reach every point with positive probability. At time k, the random walker observes the color at her present location with probability $1-\delta$ and an error Y_k with probability δ. The errors Y_k, $k \geq 0$, are assumed to be stationary and ergodic and independent of scenery and random walk. If the number of colors is strictly larger than the number of possible jumps for the random walk and δ is sufficiently small, then almost all sceneries can be almost surely reconstructed up to translations and reflections. [1]

2.1 Introduction and result

We call a coloring of the integers $\mathbb{Z}$ with colors from the set $\mathcal{C} := \{1, 2, \ldots, C\}$ a *scenery*. Let $(S_k; k \in \mathbb{N}_0)$ be a recurrent random walk on $\mathbb{Z}$. At time k the random walker observes the color $\xi(S_k)$ at her current location. Given the color record $\chi := (\xi(S_k); k \in \mathbb{N}_0)$, can we almost surely reconstruct the scenery ξ without knowing the random walk path? This problem is called *scenery reconstruction problem*. In general, one can only hope to reconstruct the scenery up to *equivalence*, where we call two sceneries ξ and ξ' *equivalent* and write $\xi \approx \xi'$ if ξ is obtained from ξ' by a translation and/or reflection.

Early work on the scenery reconstruction problem was done by Kesten in [14]. He proved that a single defect in a 4-color random scenery can be detected if the scenery

[1] *MSC 2000 subject classification:* Primary 60K37, Secondary 60G10, 60J75.
Key words: Scenery reconstruction, jumps, stationary processes, random walk, ergodic theory.

is i.i.d. uniformly colored. Reconstruction of typical 2-color sceneries was proved by Matzinger in his Ph.D. thesis [23] (see also [25] and [24]): Almost all i.i.d. uniformly colored sceneries observed along a simple random walk path (with holding) can be almost surely reconstructed. In [15], Kesten noticed that the proof in [23] heavily relies on the skip-freeness of the random walk. In [22], Löwe, Matzinger, and Merkl showed that scenery reconstruction is possible for random walks with bounded jumps if there are sufficiently many colors.

In this article, we prove that scenery reconstruction still works if the observations are seen with certain random errors. We make the same assumptions on scenery and random walk as in [22]: The random walk can reach every integer with positive probability and is recurrent with bounded jumps, and there are strictly more colors than possible single steps for the random walk. To keep the exposition as easy as possible, we assume in addition that for the random walk maximal jump length to the left and maximal jump length to the right are equal; we believe that the results of this paper remain true without this assumption. At time k the random walker observes color $\xi(S_k)$ with probability $1-\delta$, whereas she observes an error Y_k with probability δ. If the errors are independent of scenery and random walk, the occurences of errors are i.i.d. Bernoulli with parameter δ and Y_k, $k \geq 0$, is stationary and ergodic, then for all δ sufficiently small, almost all sceneries can be almost surely reconstructed up to translations and reflections.

More precisely, we consider the following setup: Let $\delta \in]0,1[$. Let μ be a probability measure over $\mathbb{Z}$ with finite support $\mathcal{M}$. With respect to a probability measure P_δ, let $S = (S_k; k \in \mathbb{N}_0)$ be a random walk starting at the origin with independent μ-distributed increments. We assume that $E[S_1] = 0$ and $\mathcal{M}$ has greatest common divisor 1; hence S is recurrent and can reach every $z \in \mathbb{Z}$ with positive probability. Let $\xi = (\xi_k; k \in \mathbb{Z})$ be a family of i.i.d. random variables, uniformly distributed over $\mathcal{C}$. Let $X := (X_k; k \in \mathbb{N}_0)$ be a sequence of i.i.d. random variables taking values in $\{0,1\}$, Bernoulli distributed with parameter δ, and let $Y := (Y_k; k \in \mathbb{N}_0)$ be as sequence of random variables taking values in $\mathcal{C}$ which is stationary and ergodic under P_δ. We assume that (ξ, S, X, Y) are independent. The scenery observed with errors along the random walk path is the process $\tilde{\chi} := (\tilde{\chi}_k; k \in \mathbb{N}_0)$ defined by $\tilde{\chi}_k := \chi_k = \xi(S_k)$ if $X_k = 0$ and $\tilde{\chi}_k := Y_k$ if $X_k = 1$. Our main theorem reads as follows:

Theorem 2.1.1. *If $|\mathcal{C}| > |\mathcal{M}|$, then there exists $\delta_1 > 0$ and a map $\mathcal{A} : \mathcal{C}^{\mathbb{N}_0} \longrightarrow \mathcal{C}^{\mathbb{Z}}$ which is measurable with respect to the canonical sigma algebras, such that $P_\delta\left(\mathcal{A}(\tilde{\chi}) \approx \xi\right) = 1$ for all $\delta \in]0, \delta_1[$.*

If $\delta = 0$, there are no errors in the observations. In this case, the assertion of Theorem 8.1.2 was proved by Löwe, Matzinger, and Merkl in [22].

Closely related coin tossing problems have been investigated by Harris and Keane [7], Levin, Pemantle, and Peres [18], and Levin and Peres [17]. The present paper has to a large extend been motivated by their work and a question of Peres who asked for generalizations of the existing random coin tossing results for the case of many biased coins.

Let $\chi' := (\chi'_k; k \in \mathbb{N}_0)$ be a coin tossing record, obtained in one of the following ways: a) a (two-sided) fair coin is tossed i.i.d., or b) at renewal times of a renewal process a coin with bias θ is tossed and at all other times a fair coin. Can we almost surely determine from χ' whether we are in case a) or b)?

Let u_n denote the probability of a renewal at time n. Harris and Keane in [7] showed that if $\sum_{n=1}^{\infty} u_n^2 = \infty$ then we can almost surely determine how χ' was produced, whereas this is not possible if $\sum_{n=1}^{\infty} u_n^2 < \infty$ and θ is small enough. Levin, Pemantle, and Peres in [18] showed that to distinguish between a) and b) not only the square-summability of (u_n) but also θ is relevant. They proved that for some renewal sequence (u_n) there is a phase transition: There exists a critical parameter θ_c such that for $|\theta| > \theta_c$ we can almost surely distinguish between a) and b), whereas for $|\theta| < \theta_c$ this is not possible.

The problem we address in this paper can be seen as a generalization of the following coin tossing problem: We have C different coins $\gamma_1, \gamma_2, \ldots, \gamma_C$ each one with C different faces $1, 2, \ldots, C$. Coin γ_i has distribution μ_i which gives probability $1 - \delta + \delta/C$ to face i and probability δ/C to each remaining face. For all $z \in \mathbb{Z}$ we choose i.i.d. uniformly among $\gamma_1, \gamma_2, \ldots, \gamma_C$ a coin $\zeta(z)$. Let $(S_k; k \in \mathbb{N}_0)$ be a random walk on $\mathbb{Z}$ fulfilling the conditions described above, independent of ζ. We generate a coin tossing record $\chi' := (\chi'_k; k \in \mathbb{N}_0)$ by tossing the coin $\zeta(S_k)$ at location S_k at time k. Then χ' has the same distribution as $\tilde{\chi}$ defined above, if we choose Y_k i.i.d. uniformly distributed over $\mathcal{C}$. Theorem 8.1.2 implies that we can almost surely determine ζ up to equivalence from the coin tossing record χ', as long as δ is small enough.

Research on random sceneries started by work by Keane and den Hollander ([13] and [5]) who studied ergodic properties of a color record seen along a random walk. Their questions were motivated among others by the work of Kalikow [12] in ergodic theory. More recently, den Hollander, Steif [4], and Heicklen, Hoffman, Rudolph [8] contributed to this area.

A preform of the scenery reconstruction problem is the scenery distinguishing problem (for a description of the problem see [15]) which started with the question whether any two non-equivalent sceneries can be distinguished. This question was asked by Benjamini and independently by den Hollander and Keane. The problem has been investigated by Benjamini and Kesten in [2] and [14]. Howard in [11], [10], [9] also contributed to this area. Recently, Lindenstrauss [19] showed the existence of uncountably many sceneries which cannot be reconstructed.

Löwe and Matzinger [21] proved that two-dimensional sceneries can be reconstructed if there are enough colors. In the case of a 2-color scenery and simple random walk with holding, the authors ([27], see also [26]) showed that the reconstruction can be done in polynomial time. By a result of Löwe and Matzinger [20], reconstruction is possible in many cases even if the scenery is not i.i.d., but has some correlations. In [16], Lenstra and Matzinger showed that scenery reconstruction is still possible if the random walk might jump more than distance 1 with very small probability and the tail of the jump distribution decays sufficiently fast.

The exposition is organized as follows. In Section 9.2, we introduce some notation and we formally describe our setup. Section 2.3 describes the structure of the proof of Theorem 8.1.2: By an ergodicity argument, it suffices to find a partial reconstruction algorithm $\mathcal{A}'$ which reconstructs correctly with probability $> 1/2$. To construct $\mathcal{A}'$, we build partial reconstruction algorithms $\mathcal{A}^m$, $m \geq 1$, which reconstruct bigger and bigger pieces of scenery around the origin. Section 2.4 contains the proofs of the theorems from Section 2.3. The core of the reconstruction is an algorithm Alg^n which reconstructs a finite piece of scenery around the origin given as input finitely many observations, stopping times, and a small piece of scenery which has been reconstructed earlier. Section 2.5 contains the definition of Alg^n. In Section 2.6, we show that Alg^n fulfills its task with high probability.

2.2 Notation and setup

In this section, we collect frequently used notation.

Sets and functions: The cardinality of a set D is denoted by $|D|$. We write $f|D$ for the restriction of a function f to a set D. For a sequence $\mathcal{S} = (s_i; i \in I)$ we write $|\mathcal{S}| := |I|$ for the number of components of $\mathcal{S}$. If s_i is an entry of $\mathcal{S}$, we write $s_i \in \mathcal{S}$; sometimes we write $s(i)$ instead of s_i. For events B_k, $k \geq 1$, we write $\liminf_{k\to\infty} B_k := \cup_{n=1}^{\infty} \cap_{k=n}^{\infty} B_k$ for the event that all but finitely many B_k's occur.

Integers and integer intervals: $\mathbb{N}$ denotes the set of natural numbers; by definition, $0 \notin \mathbb{N}$. We set $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. If $x \in \mathbb{R}$, we denote by $\lfloor x \rfloor$ the largest integer $\leq x$. Unless explicitly stated otherwise, intervals are taken over the integers, e.g. $[a, b] = \{n \in \mathbb{Z} : a \leq n \leq b\}$, $[a, b[= \{n \in \mathbb{Z} : a \leq n < b\}$.

Sceneries: We fix $C \geq 2$, and denote by $\mathcal{C} := \{1, ..., C\}$ the set of colors. A *scenery* is an element of $\mathcal{C}^{\mathbb{Z}}$. *A piece of scenery* is an element of $\mathcal{C}^I$ for a subset I of $\mathbb{Z}$; here I need not be an integer interval. The cardinality of the set I is called the length of the piece of scenery. We denote by $(1)_I$ the piece of scenery in $\mathcal{C}^I$ which is identically equal to 1. For $I = \{i_1, i_2, \ldots, i_k\} \subseteq \mathbb{Z}$ with $i_1 < i_2 < \ldots < i_k$ and a piece of scenery $\xi \in \mathcal{C}^I$ we define $\xi_{\rightarrow}$ to be the piece of scenery ξ read from left to right and $\xi_{\leftarrow}$ to be ξ read from right to left: $\xi_{\rightarrow} := (\xi(i_j); j \in [1, k])$ and $\xi_{\leftarrow} := (\xi(i_{k-j+1}); j \in [1, k])$.

Equivalence of sceneries: Let $\psi \in \mathcal{C}^I$ and $\psi' \in \mathcal{C}^{I'}$ be two pieces of sceneries. We say that ψ and ψ' are *equivalent* and write $\psi \approx \psi'$ iff I and I' have the same length and there exists $a \in \mathbb{Z}$ and $b \in \{-1, 1\}$ such that for all $k \in I$ we have that $a + bk \in I'$ and $\psi_k = \psi'_{a+bk}$. We call ψ and ψ' *strongly equivalent* and write $\psi \equiv \psi'$ if $I' = a + I$ for some $a \in \mathbb{Z}$ and $\psi_k = \psi'_{a+k}$ for all $k \in I$. We say ψ *occurs in* ψ' and write $\psi \sqsubseteq \psi'$ if $\psi \equiv \psi'|J$ for some $J \subseteq I'$. We write $\psi \preceq \psi'$ if $\psi \approx \psi'|J$ for some $J \subseteq I'$. If the subset J is unique, we write $\psi \preceq_1 \psi'$.

Random walks, random sceneries, and random errors: Let μ be a probability measure on $\mathbb{Z}$ with finite support $\mathcal{M}$. We assume that $|\mathcal{M}| < |\mathcal{C}|$, i.e. the number of colors is strictly larger than the number of possible jumps of the random walk. We assume $\max \mathcal{M} = |\min \mathcal{M}|$, and we write $L := \max \mathcal{M}$ for the maximal jump length of the random walk. Let $\Omega_2 \subseteq \mathbb{Z}^{\mathbb{N}_0}$ denote the set of all paths with jump sizes $S_{k+1} - S_k \in \mathcal{M}$ for all $k \in \mathbb{N}_0$. We denote by Q_x the distribution on $(\Omega_2)^{\mathbb{N}_0}$ of a random walk $(S_k; k \in \mathbb{N}_0)$ starting at x with i.i.d. increments distributed according to μ. We assume that $\sum_{k\in\mathcal{M}} k\mu(k) = 0$ and $\mathcal{M}$ has greatest common divisor 1, consequently the random walk is recurrent and can reach every integer with positive probability.

The scenery $\xi := (\xi_k; k \in \mathbb{Z})$ is i.i.d. with ξ_k uniformly distributed on $\mathcal{C}$. Let $X := (X_k; k \in \mathbb{N}_0)$ be a sequence of i.i.d. Bernoulli random variables with values in $\{0, 1\}$. If $X_k = 0$, then at time k the random walk observes color $\xi(S_k)$, whereas if $X_k = 1$ an error occurs in the observations at time k: the random walker observes Y_k, where $Y := (Y_k; k \in \mathbb{N}_0)$ is a sequence of random variables taking values in $\mathcal{C}$. We assume that (ξ, S, X, Y) are independent and realized as canonical projections on $\Omega := \left(\mathcal{C}^{\mathbb{Z}}, \Omega_2, \{0, 1\}^{\mathbb{N}_0}, \mathcal{C}^{\mathbb{N}_0}\right)$ with the product σ-algebra generated by the canonical projections and probability measures $P_{\delta,x} := \nu^{\otimes\mathbb{Z}} \otimes Q_x \otimes B_\delta^{\otimes\mathbb{N}_0} \otimes \lambda$, $\delta \in [0, 1]$, $x \in \mathbb{Z}$; here ν denotes the uniform distribution on $\mathcal{C}$, B_δ the Bernoulli distribution with parameter δ on $\{0, 1\}$ and λ a probability measure on $\mathcal{C}^{\mathbb{N}_0}$ such that the left-shift is measure-preserving and ergodic with respect to λ. We abbreviate $P_\delta := P_{\delta,0}$ and $P := P_0$. .

We call $\chi := (\chi_k := \xi(S_k); k \in \mathbb{N}_0)$ the *scenery observed along the random walk path*; sometimes we write $\xi \circ S$ instead of χ. We define $\tilde{\chi} := (\tilde{\chi}_k; k \in \mathbb{N}_0)$, the *scenery observed with errors along the random walk path*, by

$$\tilde{\chi}_k := \begin{cases} \chi_k & \text{if } X_k = 0, \\ Y_k & \text{if } X_k = 1. \end{cases}$$

For a fixed scenery $\xi \in \mathcal{C}^{\mathbb{Z}}$ we set $P_\delta^\xi := \delta_\xi \otimes Q_0 \otimes B_\delta^{\otimes \mathbb{N}_0} \otimes \lambda$, where δ_ξ denotes the Dirac measure at ξ. Thus P_δ^ξ is the canonical version of the conditional probability $P_\delta(\cdot|\xi)$. We use P_δ^ξ and $P_\delta(\cdot|\xi)$ as synonyms; i.e. we never work with a different version of the conditional probability $P_\delta(\cdot|\xi)$.

Admissible paths: Let $I = [i_1, i_2]$ be an integer interval. We call a path $R \in \mathbb{Z}^I$ *admissible* if $R_{i+1} - R_i \in \mathcal{M}$ for all $i \in [i_1, i_2 - 1]$. We call $R(i_1)$ the starting point, $R(i_2)$ the endpoint, and $|I|$ the length of R.

Words: We call the elements of $\mathcal{C}^* := \cup_{n \in \mathbb{N}_0} \mathcal{C}^n$ *words*. If $w \in \mathcal{C}^n$, we say that w *has length* n and write $|w| = n$.

Ladder intervals, ladder paths, and ladder words: A *ladder interval* is a set of the form $I \cap (a + L\mathbb{Z})$ with a bounded interval I and a modulo class $a + L\mathbb{Z} \in \mathbb{Z}/L\mathbb{Z}$. Let I be a ladder interval. We call a path R of length $|I|$ which traverses I from left to right or from right to left a *ladder path* or a *straight crossing of* I. The *ladder words* of a scenery ξ over I are $(\xi|I)_\rightarrow$ and $(\xi|I)_\leftarrow$.

Filtration and shift: We define a filtration over Ω: $\mathcal{G} := (\mathcal{G}_n; n \in \mathbb{N}_0)$ with $\mathcal{G}_n := \sigma(\tilde{\chi}_k; k \in [0, n])$ is the natural filtration of the observations with errors. We define the shift $\theta : \mathcal{C}^{\mathbb{N}_0} \to \mathcal{C}^{\mathbb{N}_0}$, $\eta \mapsto \eta(\cdot + 1)$.

2.2.1 Conventions about constants

All constants keep their meaning throughout the whole article. Unless otherwise stated, they depend only on C and μ. Constants α, γ, ε, $\bar{\varepsilon}$, c_1, c_2, and n_1 play a special role in the constructions below; we state here how they are chosen. All other constants are denoted by c_i, $i \geq 3$, δ_i, ε_i, $i \geq 1$.

- We choose $\gamma > 0$.
- We choose $c_2 \in]1, \frac{C}{C-1}[$ and $\bar{\varepsilon} \in]0, \bar{\varepsilon}^{\max}[$ with
$$\bar{\varepsilon}^{\max} := \min\{1/30, \varepsilon_1/90, [\ln C - \ln c_2 - \ln(C-1)]/(90 \ln C)\},$$
where ε_1 is as in Lemma 2.6.7.
- We choose $c_1 \in \mathbb{N}$ to be a multiple of 36 with $c_1 \geq 27/[\ln C - \ln c_2 - \ln(C-1) - 90\bar{\varepsilon} \ln C]$.
- We set $\varepsilon := c_1 \bar{\varepsilon}$.
- We choose $\alpha > \max\{\gamma, 1 + \gamma - [3c_1 \ln \mu_{\min}]/\ln 2\}$, where we abbreviate $\mu_{\min} := \min\{\mu(i) : i \in \mathcal{M}\}$.
- Finally we choose $n_1 \in \mathbb{N}$, $n_1 \geq \min\{25, c_3\}$, large enough that $2^n \geq c_1 L 2^{\lfloor \sqrt{n} \rfloor}$ for all $n \geq n_1$ and $\varepsilon_2(n_1) + (2\varepsilon_3(n_1))^{1/2} + \sum_{m=2}^\infty c_4 e^{-c_5 n_m} < 1/2$ holds, where c_3 is defined in Theorem 2.3.5, $\varepsilon_2(n_1)$ in Lemma 2.4.3, $\varepsilon_3(n_1)$ in Theorem 2.3.3, and c_4 and c_5 in Lemma 2.4.4.

2.3 The structure of the reconstruction

In order to prove Theorem 8.1.2, we reduce the problem of reconstructing the scenery successively to simpler problems. Theorems 2.3.1 and 2.3.2 below show that it suffices to find algorithms which do only partial reconstructions. Proofs are postponed to later sections: Sections 2.5 and 2.6 are dedicated to the proof of Theorem 2.3.5, all other statements of this section are proved in Section 2.4. Our first theorem states that it suffices to find a reconstruction algorithm $\mathcal{A}'$ which reconstructs correctly with probability $> 1/2$:

Theorem 2.3.1. *If there exist $\delta_1 > 0$ and a measurable map $\mathcal{A}' : \mathcal{C}^{\mathbb{N}_0} \longrightarrow \mathcal{C}^{\mathbb{Z}}$ such that $P_\delta\left(\mathcal{A}'(\tilde{\chi}) \approx \xi\right) > 1/2$ for all $\delta \in]0, \delta_1[$, then there exists a measurable map $\mathcal{A} : \mathcal{C}^{\mathbb{N}_0} \longrightarrow \mathcal{C}^{\mathbb{Z}}$ such that $P_\delta\left(\mathcal{A}(\tilde{\chi}) \approx \xi\right) = 1$ for all $\delta \in]0, \delta_1[$.*

The idea is to apply the reconstruction algorithm $\mathcal{A}'$ to all the shifted observations $\theta^i(\tilde{\chi})$, $i \geq 0$. By the hypothesis and an ergodicity argument, as k tends to infinity the proportion of sceneries $\mathcal{A}'(\theta^i(\tilde{\chi}))$ for $i \in [0, k[$ which are equivalent to ξ is strictly bigger than the proportion of sceneries which are not equivalent to ξ. Therefore we are able to reconstruct the scenery.

We build the algorithm $\mathcal{A}'$ required by Theorem 2.3.1 by putting together a hierarchy of partial reconstruction algorithms $\mathcal{A}^m$, $m \geq 1$. The algorithm $\mathcal{A}^m$ tries to reconstruct a piece of scenery around the origin of length of order 2^{n_m} with $(n_m; m \in \mathbb{N})$ recursively defined as follows: We choose n_1 as in Section 2.2.1, and we set for $m \geq 1$

$$n_{m+1} := 2^{\lfloor \sqrt{n_m} \rfloor}. \tag{2.3.1}$$

Definition 2.3.1. *For $m \geq 1$ and a measurable map $f : \mathcal{C}^{\mathbb{N}_0} \to \mathcal{C}^{[-3\cdot 2^{n_m}, 3\cdot 2^{n_m}]}$ we define*

$$E^m_{\text{reconst,f}} := \left\{\xi | [-2^{n_m}, 2^{n_m}] \preceq f(\tilde{\chi}) \preceq \xi | [-4 \cdot 2^{n_m}, 4 \cdot 2^{n_m}]\right\}. \tag{2.3.2}$$

$E^m_{\text{reconst,f}}$ is the event that the reconstruction procedure f reconstructs correctly a piece of scenery of length of order 2^{n_m} around the origin. Note that any finite piece of scenery occurs somewhere with probability 1 because the scenery is i.i.d. uniformly colored. Therefore it is crucial to reconstruct a piece of scenery around the origin.

Theorem 2.3.2. *Suppose there exist $\delta_1 > 0$ and a sequence of measurable maps $\mathcal{A}^m : \mathcal{C}^{\mathbb{N}_0} \to \mathcal{C}^{[-3\cdot 2^{n_m}, 3\cdot 2^{n_m}]}$, $m \geq 1$, such that for all $\delta \in]0, \delta_1[$*

$$\liminf_{m\to\infty} E^m_{\text{reconst},\mathcal{A}^m} = \liminf_{m\to\infty} \left(E^m_{\text{reconst},\mathcal{A}^m} \cap E^{m+1}_{\text{center}}\right) \; P_\delta - a.s., \tag{2.3.3}$$

where $E^{m+1}_{\text{center}} := \{\mathcal{A}^{m+1}(\tilde{\chi}) | [-3 \cdot 2^{n_m}, 3 \cdot 2^{n_m}] = \mathcal{A}^m(\tilde{\chi})\}$. Suppose further that

$$P_\delta\left(\bigcup_{m=1}^{\infty} \left(E^m_{\text{reconst},\mathcal{A}^m}\right)^c\right) < 1/2 \quad \text{for all } \delta \in]0, \delta_1[. \tag{2.3.4}$$

Then there exists a measurable map $\mathcal{A}' : \mathcal{C}^{\mathbb{N}_0} \longrightarrow \mathcal{C}^{\mathbb{Z}}$ such that $P_\delta\left(\mathcal{A}'(\tilde{\chi}) \approx \xi\right) > 1/2$ for all $\delta \in]0, \delta_1[$.

In the following, we explain how we construct maps $\mathcal{A}^m$ satisfying the assumptions of Theorem 2.3.2. The task of $\mathcal{A}^1$ is to reconstruct a piece of scenery of length of order 2^{n_1} around the origin with high probability. It is shown by Löwe, Matzinger, and Merkl in [22] that the whole scenery can be reconstructed with probability one in case there are no errors in the observations. They only prove existence of a reconstruction procedure, but do not explicitly construct an algorithm. In [28] we construct an algorithm which even works in polynomial time: A finite piece of scenery around the origin can be reconstructed with high probability from finitely many error-free observations; the number of observations needed is polynomial in the length of the piece of scenery which is reconstructed. We prove:

Theorem 2.3.3. *For infinitely many $n \in \mathbb{N}$ there exists a measurable map $\mathcal{A}^n_{\text{initial}} : \mathcal{C}^{[0,2\cdot 2^{12\alpha n}[} \to \mathcal{C}^{[-3\cdot 2^n, 3\cdot 2^n]}$ such that*

$$\varepsilon_3(n) := P\left(\left\{\xi|[-2^n,2^n] \preceq \mathcal{A}^n_{\text{initial}}\left(\chi|[0,2\cdot 2^{12\alpha n}[\right) \preceq \xi|[-4\cdot 2^n, 4\cdot 2^n]\right\}^c\right)$$

satisfies $\lim_{n\to\infty} \varepsilon_3(n) = 0$.

As an immediate consequence of Theorem 2.3.3 a piece of scenery around the origin can be reconstructed with high probability even if there are errors in the observations. As long as the probability δ to see an error at a particular time is sufficiently small, the probability to see no errors in the first $2\cdot 2^{12\alpha n}$ observations is close to 1. The following corollary makes this precise:

Corollary 2.3.1. *Let $\mathcal{A}^n_{\text{initial}}$ and $\varepsilon_3(n)$ be as in Theorem 2.3.3. There exist $\delta_2(n) > 0$ such that for all $\delta \in]0, \delta_2(n)[$*

$$P_\delta\left(\left\{\xi|[-2^n,2^n] \preceq \mathcal{A}^n_{\text{initial}}\left(\tilde{\chi}|[0,2\cdot 2^{12\alpha n}[\right) \preceq \xi|[-4\cdot 2^n, 4\cdot 2^n]\right\}^c\right) \le 2\varepsilon_3(n).$$

We will choose $\mathcal{A}^1 := \mathcal{A}^{n_1}_{\text{initial}}$. The maps $\mathcal{A}^m$, $m \ge 2$, will be defined inductively. Given a partial reconstruction algorithm $\mathcal{A}^m$ we define stopping times which tell us when the random walker is in some sense "close" to the origin: We compare $\mathcal{A}^m(\tilde{\chi})$ with $\mathcal{A}^m(\theta^t(\tilde{\chi}))$, i.e. we compare the output of $\mathcal{A}^m$ if the input consists of the observations collected by the random walker starting at the origin and the observations starting at time t. If both outputs agree up to equivalence on a sufficiently large subpiece, then with a high chance, the random walker is - on an appropriate scale - close to the origin.

The stopping times constructed from $\mathcal{A}^m$ are used to reconstruct a piece of scenery around the origin of length of order $2^{n_{m+1}}$ which is much larger than the piece of scenery reconstructed by $\mathcal{A}^m$; recall our choice of n_m (2.3.1). Whenever the stopping times indicate that the random walk is "close" to the origin, we collect significant parts of the observations of length $c_1 n_m$. If we have sufficiently many stopping times, the random walk will walk over the same piece of scenery over and over again. This allows us to filter out the errors in the observations. Once this is done, the obtained words are put together like in a puzzle game. The words are used to extend the piece of scenery of length of order 2^{n_m} which has been reconstructed by $\mathcal{A}^m$.

Formally we define stopping times in the following way:

Definition 2.3.2. *For $m \in \mathbb{N}$ and a measurable map $f : \mathcal{C}^{\mathbb{N}_0} \to \mathcal{C}^{[-3\cdot 2^{n_m}, 3\cdot 2^{n_m}]}$ with the property that $f(\tilde{\chi})$ depends only on $\tilde{\chi}|[0, 2\cdot 2^{12\alpha n_m}[$, we define*

$$\mathbb{T}^{m+1}_f(\tilde{\chi}) := \left\{ \begin{matrix} t \in [0, 2^{12\alpha n_{m+1}} - 2\cdot 2^{12\alpha n_m}[\ : \ \exists w \in \mathcal{C}^{[-2^{n_m}, 2^{n_m}]} \\ \text{such that } w \preceq f(\tilde{\chi}) \text{ and } w \preceq f(\theta^t(\tilde{\chi})) \end{matrix} \right\}.$$

Let $t(1) < t(2) < \cdots$ be the elements of $\mathbb{T}_f^{m+1}(\tilde{\chi})$ arranged in increasing order. We define the sequence $T_f^{m+1}(\tilde{\chi}) := \left(T_{f,k}^{m+1}(\tilde{\chi}); k \geq 1\right)$ by

$$T_{f,k}^{m+1}(\tilde{\chi}) := \begin{cases} t(2 \cdot 2^{2n_{m+1}} k) + 2 \cdot 2^{12\alpha n_m} & \text{if } 2 \cdot 2^{2n_{m+1}} k \leq \left|\mathbb{T}_f^{m+1}(\tilde{\chi})\right|, \\ 2^{12\alpha n_{m+1}} & \text{otherwise.} \end{cases}$$

$T_f^{m+1}(\tilde{\chi})$ is a sequence of $\mathcal{G}$-adapted stopping times with values in $[0, 2^{12\alpha n_{m+1}}]$; the stopping times depend only on $\tilde{\chi}|\,[0, 2^{12\alpha n_{m+1}}[$. We define the event that a sequence of stopping times fulfils the task of stopping the random walk "close" to the origin (on a rather rough scale).

Definition 2.3.3. *For $n \in \mathbb{N}$ and a sequence $\tau = (\tau_k; k \geq 1)$ of $\mathcal{G}$-adapted stopping times we define the event $E_{\text{stop}}^{n,\tau} :=$*

$$\bigcap_{k=1}^{2^{\alpha n}} \left\{\tau_k(\tilde{\chi}) < 2^{12\alpha n}, |S(\tau_k(\tilde{\chi}))| \leq 2^n, \tau_j(\tilde{\chi}) + 2 \cdot 2^{2n} \leq \tau_k(\tilde{\chi}) \text{ for } j < k\right\}.$$

The next theorem states that given an appropriate partial reconstruction algorithm f, the stopping times T_f^{m+1} fulfil their task with a high probability. By the definition of T_f^{m+1}, we stop at time $t + 2 \cdot 2^{12\alpha n_m}$ iff $f(\tilde{\chi})$ and $f(\theta^t(\tilde{\chi}))$ agree on a large enough subpiece. Therefore, for the stopping times to stop the random walk close to the origin, it is necessary that $f(\tilde{\chi})$ is a correctly reconstructed piece of scenery around the origin. Since we apply f often to obtain enough stopping times, we need that given a scenery ξ, there is a high enough chance for the random walk on ξ to be stopped correctly, i.e. f must reconstruct correctly with high enough probability conditional on ξ. This is why we need the event $\left\{P_\delta\left[E_{\text{reconst,f}}^m \,|\, \xi\right] \geq \frac{1}{2}\right\}$ in the following theorem.

Theorem 2.3.4. *Let $m \geq 1$, and let $f : \mathcal{C}^{\mathbb{N}_0} \to \mathcal{C}^{[-3\cdot 2^{n_m}, 3\cdot 2^{n_m}]}$ be a measurable map with the property that $f(\tilde{\chi})$ depends only on $\tilde{\chi}|[0, 2 \cdot 2^{12\alpha n_m}[$. We have for all $\delta \in]0,1[$*

$$P_\delta\left(\left(E_{\text{reconst,f}}^m \setminus E_{\text{stop}}^{n_{m+1}, T_f^{m+1}}\right) \cap \left\{P_\delta\left[E_{\text{reconst,f}}^m \,|\, \xi\right] \geq \frac{1}{2}\right\}\right) \leq e^{-n_{m+1}}.$$

The next theorem shows that there exist partial reconstruction algorithms Alg^n (the reader should think of $n = n_m$) with the following properties: Given stopping times which stop the random walk close to the origin, finitely many observations with errors and a small piece of scenery ψ close to the origin, Alg^n reconstructs with high probability a piece of scenery around the origin of length of order 2^n. If the reconstruction is succesful, the output of Alg^n contains ψ in the middle. The reader should think of ψ as a piece of scenery that has been reconstructed before.

Theorem 2.3.5. *For all $n \in \mathbb{N}$ there exists a measurable map*

$$Alg^n : [0, 2^{12\alpha n}]^{\mathbb{N}} \times \mathcal{C}^{2\cdot 2^{12\alpha n}} \times \bigcup_{k \geq c_1 L} \mathcal{C}^{[-kn, kn]} \to \mathcal{C}^{[-3\cdot 2^n, 3\cdot 2^n]}$$

with the following property: There exist constants $c_3, \delta_3, c_6, c_7 > 0$ such that for all $n \geq c_3$, $\delta \in]0, \delta_3[$ and for any sequence $\tau = (\tau_k; k \geq 1)$ of $\mathcal{G}$-adapted stopping times with values in $[0, 2^{12\alpha n}]$

$$P_\delta\left(E_{\text{stop}}^{n,\tau} \setminus E_{\text{reconstruct}}^{n,\tau}\right) \leq c_6 e^{-c_7 n},$$

where $E^{n,\tau}_{\text{reconstruct}} :=$

$$\left\{\begin{array}{l}\text{For all } \psi \in \mathcal{C}^{[-kn,kn]} \text{ with } k \geq c_1 L \text{ and } \psi \preceq \xi|\,[-2^n, 2^n] \text{ we}\\ \text{have } \xi|[-2^n, 2^n] \preceq Alg^n(\tau, \tilde{\chi}|\,[0, 2\cdot 2^{12\alpha n}[\,, \psi) \preceq \xi|[-4\cdot 2^n, 4\cdot\\ 2^n].\end{array}\right\}.$$

Furthermore if $\xi|[-2^n, 2^n] \preceq Alg^n(\tau, \tilde{\chi}|\,[0, 2\cdot 2^{12\alpha n}[\,, \psi) \preceq \xi|[-4\cdot 2^n, 4\cdot 2^n]$ *holds,* $\psi \in \mathcal{C}^{[-kn,kn]}$ *with* $k \geq c_1 L$, $\psi \preceq \xi|\,[-2^n, 2^n]$ *and* $\xi|\,[-2^n, 2^n] \neq (1)_{[-2^n,2^n]}$, *then we conclude that* $Alg^n(\tau, \tilde{\chi}|\,[0, 2\cdot 2^{12\alpha n}[\,, \psi)|[-kn, kn] = \psi$.

To motivate the allowed range for the abstract arguments τ in this theorem, recall that the $T^m_{f,k}(\tilde{\chi})$'s in Definition 2.3.2 take their values in $[0, 2^{12\alpha n_m}]$. We are now able to define $\mathcal{A}^m$, $m \geq 1$, which fulfill the requirements of Theorem 2.3.2.

Definition 2.3.4. *We define* $\mathcal{A}^m : \mathcal{C}^{\mathbb{N}_0} \to \mathcal{C}^{[-3\cdot 2^{n_m}, 3\cdot 2^{n_m}]}$ *and sequences* $T^{m+1} = \left(T^{m+1}_k; k \geq 1\right)$ *recursively for* $m \geq 1$ *in the following way:*

- $\mathcal{A}^1(\tilde{\chi}) := \mathcal{A}^{n_1}_{\text{initial}}(\tilde{\chi}|[0, 2\cdot 2^{12\alpha n_1}[)$ *with* n_1 *as in Section 2.2.1 and* $\mathcal{A}^{n_1}_{\text{initial}}$ *as in Theorem 2.3.3,*
- $T^{m+1}(\tilde{\chi}) := T^{m+1}_{\mathcal{A}^m}(\tilde{\chi})$ *with* $T^{m+1}_{\mathcal{A}^m}$ *as in Definition 2.3.2,*
- $\mathcal{A}^{m+1}(\tilde{\chi}) := Alg^{n_{m+1}}(T^{m+1}(\tilde{\chi}), \tilde{\chi}|[0, 2\cdot 2^{12\alpha n_{m+1}}[, \mathcal{A}^m(\tilde{\chi}))$ *with* $Alg^{n_{m+1}}$ *as in Theorem 2.3.5.*

Theorem 2.3.6. *There exists* $\delta_1 > 0$ *such that the sequence* $(\mathcal{A}^m; m \in \mathbb{N})$ *defined in Definition 2.3.4 fulfils (2.3.3) and (2.3.4) for all* $\delta \in]0, \delta_1[$.

All theorems of this section together yield the proof of our main theorem:

Proof of Theorem 8.1.2. By Theorem 2.3.6, the assumptions of Theorem 2.3.2 are satisfied. Hence the assumptions of Theorem 2.3.1 are satisfied and Theorem 8.1.2 follows. □

2.4 Proofs

In this section, we prove the statements from Section 2.3 with the exception of Theorem 2.3.5 which will be proved in Sections 2.5 and 2.6.

Lemma 2.4.1. *The shift* $\Theta : \Omega \to \Omega$,

$$(\xi, S, X, Y) \mapsto (\xi(\cdot + S(1)), S(\cdot + 1) - S(1), X(\cdot + 1), Y(\cdot + 1))$$

is measure-preserving and ergodic with respect to P_δ *for all* $\delta \in]0, 1[$.

Proof. Let $\delta \in]0, 1[$. By assumption, Y_k, $k \geq 0$, is stationary and ergodic under P_δ. X_k, $k \geq 0$, is i.i.d., hence stationary and ergodic under P_δ. By Lemma 4.1 of [22], $(\xi, S) \mapsto (\xi(\cdot + S(1)), S(\cdot + 1) - S(1))$ is measure-preserving and ergodic with respect to P. The claim follows from these three observations and the fact that (ξ, S, X, Y) are independent. □

Proof of Theorem 2.3.1. Let δ_1 and $\mathcal{A}' : \mathcal{C}^{\mathbb{N}_0} \to \mathcal{C}^{\mathbb{Z}}$ be as in the hypothesis of the theorem, and let $\delta \in]0, \delta_1[$. We define for $k \in \mathbb{N}$ measurable maps $\mathcal{A}'_k : \mathcal{C}^{\mathbb{N}_0} \to \mathcal{C}^{\mathbb{Z}}$ as follows: If there exists $j \in [0, k[$ such that

$$\left|\left\{i \in [0,k[\colon \mathcal{A}'(\theta^i(\tilde{\chi})) \approx \mathcal{A}'(\theta^j(\tilde{\chi}))\right\}\right| > \left|\left\{i \in [0,k[\colon \mathcal{A}'(\theta^i(\tilde{\chi})) \not\approx \mathcal{A}'(\theta^j(\tilde{\chi}))\right\}\right|,$$

then let j_0 be the smallest j with this property, and define $\mathcal{A}'_k(\tilde{\chi}) := \mathcal{A}'(\theta^{j_0}(\tilde{\chi}))$. Otherwise define $\mathcal{A}'_k(\tilde{\chi})$ to be the constant scenery $(1)_{j\in\mathbb{Z}}$. Finally we define $\mathcal{A} : \mathcal{C}^{\mathbb{N}_0} \to \mathcal{C}^{\mathbb{Z}}$ by

$$\mathcal{A}(\tilde{\chi}) := \begin{cases} \lim_{k\to\infty} \mathcal{A}'_k(\tilde{\chi}) & \text{if this limit exists pointwise,} \\ (1)_{j\in\mathbb{Z}} & \text{else.} \end{cases}$$

As a limit of measurable maps, $\mathcal{A}$ is measurable. For $k \in \mathbb{N}$ we define

$$Z_k := \frac{1}{k} \sum_{i=0}^{k-1} 1\left\{\mathcal{A}'(\theta^i(\tilde{\chi})) \approx \xi\right\};$$

here $1B$ denotes the indicator function of the event B. It follows from Lemma 2.4.1 that the sequence $1\left\{\mathcal{A}'(\theta^k(\tilde{\chi})) \approx \xi\right\}$, $k \geq 0$, is stationary and ergodic because it can be written as a measurable function of the sequence $\Theta^k(\xi, S, X, Y)$, $k \geq 0$; note that $\xi \approx \xi(\cdot + S_k)$. Hence we can use the ergodic theorem and our assumption to obtain P_δ-almost surely:

$$\lim_{k\to\infty} Z_k = P_\delta\left(\mathcal{A}'(\tilde{\chi}) \approx \xi\right) > 1/2. \tag{2.4.1}$$

Note that if $Z_k > 1/2$, then $\mathcal{A}'_k(\tilde{\chi}) \approx \xi$. By (2.4.1) there exists a.s. a (random) k_0 such that $Z_k > 1/2$ for all $k \geq k_0$, and hence $\mathcal{A}'_k(\tilde{\chi}) = \mathcal{A}'_{k_0}(\tilde{\chi}) \approx \xi$; recall that we chose the smallest possible j_0 in the definition of $\mathcal{A}'_k$. Thus a.s. $\mathcal{A}(\tilde{\chi}) \approx \xi$. □

Proof of Theorem 2.3.2. We say a sequence $(\zeta^m; m \in \mathbb{N})$ of pieces of sceneries converges pointwise to a scenery ζ if $\liminf_{m\to\infty} \mathrm{domain}(\zeta^m) = \mathbb{Z}$, and for every $z \in \mathbb{Z}$ there is $m_z > 0$ such that $\zeta^m(z) = \zeta(z)$ for all $m \geq m_z$.

Let δ_1 and $\mathcal{A}^m$ be as in the hypothesis of the theorem, and let $\delta \in]0, \delta_1[$. We set $\mathcal{A}'(\tilde{\chi}) := \lim_{m\to\infty} \mathcal{A}^m(\tilde{\chi})$ if this limit exists pointwise on $\mathbb{Z}$; otherwise we set $\mathcal{A}'(\tilde{\chi}) := (1)_{j\in\mathbb{Z}}$. Being a pointwise limit of measurable maps, $\mathcal{A}' : \mathcal{C}^{\mathbb{N}_0} \to \mathcal{C}^{\mathbb{Z}}$ is measurable. We abbreviate $E^m := E^m_{\mathrm{reconst},\mathcal{A}^m}$, and define the events

$$E^m_{1\mathrm{fit}} := \left\{\xi | [-2^{n_m}, 2^{n_m}] \preccurlyeq_1 \xi | [-4 \cdot 2^{n_{m+1}}, 4 \cdot 2^{n_{m+1}}]\right\}.$$

We claim:

1. $\liminf_{m\to\infty} E^m_{1\mathrm{fit}}$ holds P_δ-a.s.,
2. If the event $(\liminf_{m\to\infty} E^m_{1\mathrm{fit}}) \cap \bigcap_{m=1}^\infty E^m$ holds, then $\mathcal{A}'(\tilde{\chi}) \approx \xi$.

Together with the assumption $P_\delta\left[\cup_{m=1}^\infty (E^m)^c\right] < 1/2$ these two statements imply that $P_\delta\left(\mathcal{A}'(\tilde{\chi}) \approx \xi\right) > 1/2$ which yields the claim of the theorem.

Proof of claim 1: We show for any integer intervals $I_1 \neq I_2$ with $|I_1| = |I_2|$

$$P(\xi | I_1 \approx \xi | I_2) \leq 2 \cdot C^{-|I_j|/3}. \tag{2.4.2}$$

First we define $f_j : [0, |I_j|[\to I_j$ for $j = 1, 2$ to be the unique translation which maps $[0, |I_j|[$ onto I_j. An argument similar to the proof of (2.6.26) below shows that there exists a subset $J \subseteq [0, |I_j|[$ of cardinality $|J| \geq |I_j|/3$ with $f_1(J) \cap f_2(J) = \varnothing$. Since ξ_k, $k \in \mathbb{Z}$, are i.i.d. with a uniform distribution, we conclude

$$P(\xi|I_1 \equiv \xi|I_2) \leq P(\xi|f_1(J) = \xi|f_2(J)) = C^{-|J|} \leq C^{-|I_j|/3}.$$

Since $\xi|I_1 \approx \xi|I_2$ means $\xi|I_1 \equiv \xi|I_2$ or $\xi|I_1 \equiv (\xi|I_2)^{\leftrightarrow}$ with $(\xi|I_2)^{\leftrightarrow}$ denoting the piece of scenery obtained from $\xi|I_2$ by reflection, estimate (2.4.2) follows.

We apply (2.4.2) for $I_1 = [-2^{n_m}, 2^{n_m}]$ and all integer intervals $I_2 \subseteq [-4 \cdot 2^{n_{m+1}}, 4 \cdot 2^{n_{m+1}}]$, $I_1 \neq I_2$, of length $|I_1| = |I_2| = 2 \cdot 2^{n_m} + 1$; there are not more than $8 \cdot 2^{n_{m+1}}$ choices for I_2. We obtain

$$P\left((E^m_{\text{1fit}})^c\right) \leq 8 \cdot 2^{n_{m+1}} \cdot 2 \cdot C^{-(2 \cdot 2^{n_m}+1)/3} \leq 16 \cdot 2^{2\sqrt{n_m} - 2 \cdot 2^{n_m}/3},$$

which is summable over m; recall $C \geq 2$ and (2.3.1). Hence by the Borel-Cantelli lemma $(E^m_{\text{1fit}})^c$ occurs P_δ-a.s. only finitely many times ; this proves claim 1.

Proof of claim 2: By the assumption of this claim, there is a (random) M such that the events E^m_{1fit} and E^m hold for all $m \geq M$. By the assumption of Theorem 2.3.2, M can be chosen in such a way that E^{m+1}_{center} holds for all $m \geq M$, too. Consequently, $\mathcal{A}^{m+1}(\tilde{\chi})|\,[-3 \cdot 2^{n_m}, 3 \cdot 2^{n_m}] = \mathcal{A}^m(\tilde{\chi})$ for all $m \geq M$ and it follows that

$$\mathcal{A}'(\tilde{\chi})|[-k, k] = \mathcal{A}^m(\tilde{\chi})|[-k, k] \tag{2.4.3}$$

for all $k \geq 1$ and all m large enough. In particular, $\lim_{m\to\infty} \mathcal{A}^m(\tilde{\chi})$ exists.

Since E^m and E^m_{1fit} hold, $\mathcal{A}^m(\tilde{\chi}) \preceq_1 \xi|[-4 \cdot 2^{n_m}, 4 \cdot 2^{n_m}]$. Hence there exists a unique map $h^m : \mathbb{Z} \to \mathbb{Z}$ of the form $x \mapsto a_m + b_m x$ with $a_m \in \mathbb{Z}$ and $b_m \in \{-1, 1\}$ that maps $\mathcal{A}^m(\tilde{\chi})$ onto a subpiece of $\xi|[-4 \cdot 2^{n_m}, 4 \cdot 2^{n_m}]$. It follows from (2.4.3) that h^m is independent of m and maps $\mathcal{A}'(\tilde{\chi})$ to ξ. This finishes the proof of claim 2. □

Proof of Theorem 2.3.3. By Theorem 1.1 of [28], we know that there exists $\beta > 0$ and for infinitely many $n \in \mathbb{N}$ there exists a measurable map $\mathcal{A}^n_{\text{ini}} : \mathcal{C}^{[0, 2n^7 + 2 \cdot 2^{12\beta n}[} \to \mathcal{C}^{[-5 \cdot 2^n, 5 \cdot 2^n]}$ such that $\lim_{n\to\infty} P\left([E^n_{\text{ini}}]^c\right) = 0$, where

$$E^n_{\text{ini}} := \left\{\xi|\left[-2^{n-1}, 2^{n-1}\right] \preccurlyeq \mathcal{A}^n_{\text{ini}}\left(\chi|[0, 2n^7 + 2 \cdot 2^{12\beta n}[\right) \preccurlyeq \xi|\left[-10 \cdot 2^n, 10 \cdot 2^n\right]\right\}.$$

Small modifications in the proof of Theorem 1.1 in [28] prove our claim. We remark that alternatively, we could work directly with the maps $\mathcal{A}^n_{\text{ini}}$ from [28] without adjusting the constants; all proofs in the remainder of the article go through, but the notation becomes more cumbersome. □

Proof of Corollary 2.3.1. We estimate the probability under consideration by intersecting with the event $B_0 := \{X_k = 0 \text{ for all } k \in [0, 2 \cdot 2^{12\alpha n}[\}$ that there are no errors in the first $2 \cdot 2^{12\alpha n}$ observations: For any $\delta > 0$ we have

$$\begin{aligned}
& 1 - P_\delta\left(\xi|[-2^n, 2^n] \preceq \mathcal{A}^n_{\text{initial}}\left(\tilde{\chi}|[0, 2 \cdot 2^{12\alpha n}[\right) \preceq \xi|[-4 \cdot 2^n, 4 \cdot 2^n]\right) \\
\leq\ & 1 - P_\delta\left(\left\{\xi|[-2^n, 2^n] \preceq \mathcal{A}^n_{\text{initial}}\left(\tilde{\chi}|[0, 2 \cdot 2^{12\alpha n}[\right) \preceq \xi|[-2^{n+2}, 2^{n+2}]\right\} \cap B_0\right) \\
=\ & 1 - \delta(n) P\left(\xi|[-2^n, 2^n] \preceq \mathcal{A}^n_{\text{initial}}\left(\chi|[0, 2 \cdot 2^{12\alpha n}[\right) \preceq \xi|[-2^{n+2}, 2^{n+2}]\right) \\
=\ & 1 - \delta(n)(1 - \varepsilon_3(n));
\end{aligned}$$

with $\delta(n) := (1-\delta)^{2\cdot 2^{12\alpha n}}$ and $\varepsilon_3(n)$ as in Theorem 2.3.3. We choose $\delta_2(n) > 0$ such that the last expression is bounded above by $2\varepsilon_3(n)$ for all $\delta \in]0, \delta_2(n)[$. □

Proof of Theorem 2.3.4. The proof is very similar to the proof of Theorem 3.11 in section 7 of [22] (Our Theorem 2.3.4 is the analogon of their Theorem 3.11 for our setting). The errors in the observations do not require adaptations of their arguments; note that the errors are independent of scenery and random walk and occurences of errors are i.i.d. Bernoulli. □

The remainder of this section is dedicated to the proof of Theorem 2.3.6. Throughout we assume $\mathcal{A}^m$, $m \geq 1$, are as in Definition 2.3.4, and we set $\delta_1 := \min\{\delta_3, \delta_2(n_1)\}$ with δ_3 as in Theorem 2.3.5 and $\delta_2(n_1)$ as in Corollary 2.3.1. We set for $m \geq 1$

$$E^m := E^m_{\text{reconst},\mathcal{A}^m}. \tag{2.4.4}$$

Definition 2.4.1. *For $\delta \in]0, \delta_1[$ we define events of sceneries*

$$\begin{aligned}
\Xi_1^\delta &:= \left\{\xi \in \mathcal{C}^{\mathbb{Z}} : P_\delta\left[(E^1)^c \,\middle|\, \xi\right] \leq (2\varepsilon_3(n_1))^{1/2}\right\},\\
\Xi_2^\delta &:= \bigcap_{m=2}^{\infty}\left\{\xi \in \mathcal{C}^{\mathbb{Z}} : P_\delta\left[E^{m-1}|\xi\right] \geq \frac{1}{2} \Rightarrow P_\delta\left[E^{m-1} \setminus E^{n_m,T^m}_{\text{stop}} \,\middle|\, \xi\right] \leq e^{-\frac{n_m}{2}}\right\}\\
&= \bigcap_{m=2}^{\infty}\left\{\xi \in \mathcal{C}^{\mathbb{Z}} : P_\delta\left[\left[E^{m-1} \setminus E^{n_m,T^m}_{\text{stop}}\right] \cap \left\{P_\delta\left[E^{m-1} \,|\, \xi\right] \geq \frac{1}{2}\right\} \,\middle|\, \xi\right] \leq e^{-\frac{n_m}{2}}\right\},\\
\Xi_3^\delta &:= \bigcap_{m=2}^{\infty}\left\{\xi \in \mathcal{C}^{\mathbb{Z}} : P_\delta\left[E^{m-1} \cap \left(E^{n_m,T^m}_{\text{stop}} \setminus E^m\right) \;\middle|\; \xi\right] \leq (c_6)^{1/2} e^{-\frac{c_7 n_m}{2}}\right\},\\
\Xi^\delta &:= \Xi_1^\delta \cap \Xi_2^\delta \cap \Xi_3^\delta,
\end{aligned}$$

where $\varepsilon_3(n_1)$ is as in Theorem 2.3.3 and c_6 and c_7 are as in Theorem 2.3.5.

Note the similarity between these events and the bounds in Corollary 2.3.1, Theorems 2.3.4 and 2.3.5. The following lemma provides a link between bounds with and without conditioning on the scenery ξ:

Lemma 2.4.2 ([22], Lemma 4.6). *Let A be an event, $r \geq 0$, and let Q be a probability measure on Ω. If $Q(A) \leq r^2$, then $Q\left(Q(A|\xi) > r\right) \leq r$.*

Lemma 2.4.3. *For all $n \in \mathbb{N}$ there exist $\varepsilon_2(n) > 0$ with $\lim_{n\to\infty} \varepsilon_2(n) = 0$ such that $P_\delta\left(\xi \notin \Xi^\delta\right) \leq \varepsilon_2(n_1)$ for all $\delta \in]0, \delta_1[$.*

Proof. Let $\delta \in]0, \delta_1[$. Using Corollary 2.3.1 and Lemma 9.7.2 for $Q = P_\delta$, we obtain

$$P_\delta\left(\xi \notin \Xi_1^\delta\right) \leq (2\varepsilon_3(n_1))^{1/2}. \tag{2.4.5}$$

An application of Theorem 2.3.4 with $f = \mathcal{A}^m$ yields for $m \geq 2$

$$P_\delta\left(\left(E^{m-1} \setminus E^{n_m,T^m}_{\text{stop}}\right) \cap \left\{P_\delta\left[E^{m-1} \,|\, \xi\right] \geq \frac{1}{2}\right\}\right) \leq e^{-n_m}.$$

An application of Lemma 9.7.2 with $Q = P_\delta$ yields

$$P_\delta\left(\xi \notin \Xi_2^\delta\right) \leq \sum_{m=2}^{\infty} e^{-n_m/2} \leq e^{-c_8 n_1} \tag{2.4.6}$$

for some constant $c_8 > 0$, recall our choice of n_m (2.3.1). Let $m \geq 2$, and recall the definition of the event $E^{n_m,T^m}_{\text{reconstruct}}$ from Theorem 2.3.5. By Definition 2.3.4, we have that $\mathcal{A}^m(\tilde{\chi}) = \text{Alg}^{n_m}\left(T^m(\tilde{\chi}), \tilde{\chi}|[0, 2 \cdot 2^{12\alpha n_m}[, \psi\right)$ with $\psi := \mathcal{A}^{m-1}(\tilde{\chi})$. By our choice of n_1, $(|\psi| - 1)/2 = 3 \cdot 2^{n_{m-1}} \geq c_1 n_m L$. If E^{m-1} holds, then $\psi \preceq \xi|[-2^{n_m}, 2^{n_m}]$. Hence the inclusion

$$E^{m-1} \cap \left(E^{m,T^m}_{\text{stop}} \setminus E^m\right) \subseteq E^{m,T^m}_{\text{stop}} \setminus E^{n_m,T^m}_{\text{reconstruct}} \tag{2.4.7}$$

holds. Together with Theorem 2.3.5 the last inclusion implies

$$P_\delta\left(E^{m-1} \cap \left(E^{n_m,T^m}_{\text{stop}} \setminus E^m\right)\right) \leq P_\delta\left(E^{m,T^m}_{\text{stop}} \setminus E^{n_m,T^m}_{\text{reconstruct}}\right) \leq c_6 e^{-c_7 n_m}.$$

Another application of Lemma 9.7.2 yields for some constant $c_9 > 0$

$$P_\delta\left(\xi \notin \Xi_3^\delta\right) \leq \sum_{m=2}^{\infty} (c_6)^{1/2} e^{-c_7 n_m/2} \leq e^{-c_9 n_1}. \tag{2.4.8}$$

The claim of the lemma follows from (2.4.5), (2.4.6), and (2.4.8); recall $\varepsilon_3(n) \to 0$ as $n \to \infty$. □

Lemma 2.4.4. *For all $\delta \in]0, \delta_1[$, $\xi \in \Xi^\delta$, and $m \geq 2$ the following holds for some constants $c_4, c_5 > 0$:*

$$P_\delta(E^{m-1} \mid \xi) \geq 1 - (2\varepsilon_3(n_1))^{1/2} - \sum_{k=2}^{m-1} c_4 e^{-c_5 n_k} \geq \frac{1}{2}, \tag{2.4.9}$$

$$P_\delta(E^{m-1} \setminus E^m \mid \xi) \leq c_4 e^{-c_5 n_m}. \tag{2.4.10}$$

Proof. Let $\delta \in]0, \delta_1[$ and $\xi \in \Xi^\delta$. We prove (2.4.9) and (2.4.10) simultaneously by induction over m: For $m = 2$ it follows from $\xi \in \Xi_1^\delta$

$$P_\delta(E^1 \mid \xi) \;=\; 1 - P_\delta\left[\left(E^1\right)^c \mid \xi\right] \geq 1 - (2\varepsilon_3(n_1))^{1/2} \geq 1/2; \tag{2.4.11}$$

recall our choice of n_1 from Section 2.2.1. Thus (2.4.9) holds for $m = 2$.

Suppose (2.4.9) holds for some $m \geq 2$. Then we have

$$\begin{aligned} P_\delta[E^{m-1} \setminus E^m|\xi] &\leq P_\delta\left[(E^{m-1} \setminus E^m) \cap E^{m,T^m}_{\text{stop}} \,\middle|\, \xi\right] + P_\delta\left[E^{m-1} \setminus E^{m,T^m}_{\text{stop}} \,\middle|\, \xi\right] \\ &\leq (c_6)^{1/2} e^{-\frac{c_7 n_m}{2}} + e^{-n_m/2} \leq c_4 e^{-c_5 n_m} \end{aligned} \tag{2.4.12}$$

for some constants $c_4, c_5 > 0$; for the first term we used $\xi \in \Xi_3^\delta$ and for the second term we used $\xi \in \Xi_2^\delta$ and our induction hypothesis (2.4.9). Using (2.4.12) and our induction hypothesis (2.4.9) we obtain

$$\begin{aligned} P_\delta(E^m \mid \xi) &\geq P_\delta(E^{m-1} \mid \xi) - P_\delta(E^{m-1} \setminus E^m \mid \xi) \\ &\geq 1 - (2\varepsilon_3(n_1))^{1/2} - \sum_{k=2}^{m} c_4 e^{-c_5 n_k} \geq \frac{1}{2}; \end{aligned}$$

for the last inequality we used our choice of n_1. This completes the induction step. □

Proof of Theorem 2.3.6. Let $\delta \in]0, \delta_1[$; recall our choice $\delta_1 = \min\{\delta_3, \delta_2(n_1)\}$. By Theorem 2.3.5 we know that whenever the events E^{m-1} and E^m hold and $\xi|[-2^{n_m}, 2^{n_m}] \neq (1)_{[-2^{n_m}, 2^{n_m}]}$, then E^m_{center} holds. Since P_δ-a.s. $\xi \neq (1)_{\mathbb{Z}}$, relation (2.3.3) holds. Using Lemma 2.4.3 we have

$$\begin{aligned} P_\delta\left(\bigcup_{m=1}^{\infty}(E^m)^{\mathrm{c}}\right) &\leq P_\delta\left(\xi \notin \Xi^\delta\right) + P_\delta\left(\{\xi \in \Xi^\delta\} \cap \bigcup_{m=1}^{\infty}(E^m)^{\mathrm{c}}\right) \\ &\leq \varepsilon_2(n_1) + \int_{\{\xi\in\Xi^\delta\}} P_\delta\left(\bigcup_{m=1}^{\infty}(E^m)^{\mathrm{c}} \,\middle|\, \xi\right) \mathrm{d}P_\delta. \end{aligned} \tag{2.4.13}$$

To bound the integrand, we use Lemma 2.4.4: For all $\xi \in \Xi^\delta$ and $k \geq 1$, we obtain

$$\begin{aligned} P_\delta\left(\bigcup_{m=1}^{k}(E^m)^{\mathrm{c}} \,\middle|\, \xi\right) &\leq P_\delta\left((E^1)^{\mathrm{c}} \mid \xi\right) + \sum_{m=2}^{k+1} P_\delta(E^{m-1} \setminus E^m \mid \xi) \\ &\leq (2\varepsilon_3(n_1))^{1/2} + \sum_{m=2}^{k+1} c_4 e^{-c_5 n_m}, \end{aligned} \tag{2.4.14}$$

and taking limits as $k \to \infty$, we conclude

$$P_\delta\left(\bigcup_{m=1}^{\infty}(E^m)^{\mathrm{c}} \,\middle|\, \xi\right) \leq (2\varepsilon_3(n_1))^{1/2} + \sum_{m=2}^{\infty} c_4 e^{-c_5 n_m}.$$

Together with (2.4.13) the last estimate yields (2.3.4):

$$P_\delta\left(\bigcup_{m=1}^{\infty}(E^m)^{c}\right) \leq \varepsilon_2(n_1) + (2\varepsilon_3(n_1))^{1/2} + \sum_{m=2}^{\infty} c_4 e^{-c_5 n_m} < \frac{1}{2}; \tag{2.4.15}$$

for the last inequality we used that n_1 is chosen as in Section 2.2.1. □

2.5 The key algorithm of the reconstruction

In this section, we define algorithms Alg^n for which Theorem 2.3.5 holds. We fix $n \in \mathbb{N}$.

For two words $w, w' \in \mathcal{C}^*$ of the same length we define their distance

$$d(w, w') := |\{k \in [1, |w|] : w_k \neq w'_k\}|; \tag{2.5.1}$$

$d(w, w')$ is the number of places where w and w' disagree. Clearly, d is a metric.

When the random walk observes a piece of scenery and δ is small, the observations with errors differ "typically" from the errorfree observations in only a small proportion of the letters because the probability to see an error at a particular time is small under P_δ. Since the random walk observes a given piece of scenery very often, we are able to filter out the errors using a majority rule f^*.

The following notions will be used in this context. For $w = w_1 w_2 \dots w_m \in \mathcal{C}^m$ we define $\mathsf{Cut}(w) := w_2 \dots w_{m-1}$; $\mathsf{Cut}(w)$ is obtained from w by cutting off the first and the last letter.

Definition 2.5.1. *Let $W = (w_j; 1 \le j \le K) \in (\mathcal{C}^{c_1 n})^K$ be a vector consisting of K words of length $c_1 n$. For $i \in [1, c_1 n]$ we define $f_i(W)$, the* favorite letter at position i, *to be the element in $\mathcal{C}$ which most of the first $2^{\gamma n}$ words in W have at position i. If there is no unique letter with this property, then we define the favorite letter to be the smallest one. Formally, we set*

$$f_i(W) = k \quad \textit{iff} \quad |\{j \in [1, 2^{\gamma n}] : w_j(i) = k\}| = \max_{k' \in \mathcal{C}} |\{j \in [1, 2^{\gamma n}] : w_j(i) = k'\}|$$

and k is the smallest element in $\mathcal{C}$ satisfying the last equality; here $w_j(i)$ denotes the i^{th} letter of the word w_j. We set $f(W) := f_1(W) f_2(W) \dots f_{c_1 n}(W)$. Furthermore, we define $f^*(W) :=$

$$\begin{cases} \mathsf{Cut}(f(W)), & \textit{if } K \ge 2^{\gamma n} \textit{ and } \max_{j \in [1, 2^{\gamma n}]} d(\mathsf{Cut}(w_j), \mathsf{Cut}(f(W))) \le \varepsilon n \\ (-1)_{[1, c_1 n - 2]}, & \textit{otherwise.} \end{cases}$$

$f^*(W)$ equals the word $\mathsf{Cut}(f(W))$ which is composed of the favorite letters iff the vector W has sufficiently many components and each of the first $2^{\gamma n}$ words in W differs from $f(W)$ in not more than εn letters. In the proof of Lemma 2.6.9 below it will be essential that we use $\mathsf{Cut}(f(W))$ and not $f(W)$ in the definition of $f^*(W)$. Note that $-1 \notin \mathcal{C}$ so that $(-1)_{[1, c_1 n - 2]}$ differs from all words $w \in \mathcal{C}^{c_1 n - 2}$.

The algorithm Alg^n which will be defined below takes input data

$$\tau \in \left[0, 2^{12 \alpha n}\right]^{\mathbb{N}}, \ \eta \in \mathcal{C}^{2 \cdot 2^{12 \alpha n}}, \ \text{and } \psi \in \bigcup_{k \ge c_1 L} \mathcal{C}^{[-kn, kn]}. \tag{2.5.2}$$

First we define the set of all observations of length $3c_1 n$ which are collected within a time horizon of length 2^{2n} after a time $\tau_k, k \in [1, 2^{\alpha n}]$:

Definition 2.5.2. *We define* $\mathsf{Collection}^n(\tau, \eta) :=$

$$\left\{(w_1, w_2, w_3) \in (\mathcal{C}^{c_1 n})^3 : \exists k \in [1, 2^{\alpha n}] \textit{ such that } w_1 w_2 w_3 \sqsubseteq \eta | [\tau_k, \tau_k + 2^{2n}[\right\}.$$

The set $\mathsf{PrePuzzle}^n(\tau, \eta)$ contains only $(w_1, w_2, w_3) \in \mathsf{Collection}^n(\tau, \eta)$ with the following property: If $(w_1', w_2', w_3') \in \mathsf{Collection}^n(\tau, \eta)$ and w_1' and w_3' are "not too different" from w_1 and w_3 respectively, then w_2' is "not too different" from w_2. Formally:

Definition 2.5.3. *We define* $\mathsf{PrePuzzle}^n(\tau, \eta) :=$

$$\left\{ \begin{array}{l} (w_1, w_2, w_3) \in \mathsf{Collection}^n(\tau, \eta) : \textit{If } (w_1', w_2', w_3') \in \\ \mathsf{Collection}^n(\tau, \eta) \textit{ with } d(w_1, w_1') \le 2\varepsilon n \textit{ and } d(w_3, w_3') \le 2\varepsilon n, \\ \textit{then } d(w_2, w_2') \le 2\varepsilon n. \end{array} \right\}.$$

Definition 2.5.4. *For an element $(w_1, w_2, w_3) \in \mathsf{PrePuzzle}^n(\tau, \eta)$ we denote by $\mathcal{S}^n_{\tau,\eta}(w_1, w_2, w_3)$ the sequence of (random) times $s \in \cup_{k=1}^{2^{\alpha n}} [\tau_k, \tau_k + 2^{2n} - 3c_1 n]$ such that $w_1' w_2' w_3' := \eta | [s, s + 3c_1 n[\in \mathsf{PrePuzzle}^n(\tau, \eta)$, $d(w_1, w_1') \le 2\varepsilon n$, and $d(w_3, w_3') \le 2\varepsilon n$; we assume that the elements of the sequence $\mathcal{S}^n_{\tau,\eta}(w_1, w_2, w_3)$ are arranged in increasing order. We define*

$$\mathsf{List}^n_{\tau,\eta}(w_1, w_2, w_3) := \left(\eta | [s + c_1 n, s + 2c_1 n[; s \in \mathcal{S}^n_{\tau,\eta}(w_1, w_2, w_3)\right)$$

to be the sequence with components $\eta | [s + c_1 n, s + 2c_1 n[$ indexed by the set $\mathcal{S}^n_{\tau,\eta}(w_1, w_2, w_3)$. We set

$$\mathsf{PuzzleLists}^n(\tau, \eta) := \left\{ \mathsf{List}^n_{\tau,\eta}(w_1, w_2, w_3) : (w_1, w_2, w_3) \in \mathsf{PrePuzzle}^n(\tau, \eta) \right\}.$$

Clearly, $w_2 \in \mathsf{List}^n_{\tau,\eta}(w_1, w_2, w_3)$. Note that $\mathsf{List}^n_{\tau,\eta}(w_1, w_2, w_3)$ is a sequence, and not a set. If by coincidence observations $\eta|[s + c_1 n, s + 2c_1 n[$ coincide for two different values of s, we want to keep them both. The components of $\mathsf{List}^n_{\tau,\eta}(w_1, w_2, w_3)$ are close to w_2 in d-distance because we assumed $(w_1, w_2, w_3) \in \mathsf{PrePuzzle}^n(\tau, \eta)$.

Definition 2.5.5. *We define* $\mathsf{Puzzle}^n(\tau, \eta) := \{f^*(W) : W \in \mathsf{PuzzleLists}^n(\tau, \eta)\}$.

$\mathsf{Puzzle}^n(\tau, \eta)$ is the set of all words of length $c_1 n - 2$ which are obtained by the majority rule f^* from the lists in $\mathsf{PuzzleLists}^n(\tau, \eta)$. We use the words in $\mathsf{Puzzle}^n(\tau, \eta)$ like the pieces in a puzzle game to reconstruct a piece of scenery. We want the piece of scenery reconstructed by Alg^n to contain in the middle the piece of scenery ψ from the input data of the algorithm.

Definition 2.5.6. *For* $\psi \in \mathcal{C}^{[-kn,kn]}$ *we define* $\mathsf{SolutionPiece}^n(\tau, \eta, \psi) :=$

$$\left\{ \begin{array}{l} w \in \mathcal{C}^{[-3\cdot 2^n, 3\cdot 2^n]} : w|[-kn, kn] = \psi \text{ and for all ladder intervals} \\ I \subseteq [-3 \cdot 2^n, 3 \cdot 2^n] \text{ with } |I| = c_1 n - 2 \text{ we have } (w|I)_{\to} \in \\ \mathsf{Puzzle}^n(\tau, \eta) \end{array} \right\}.$$

We will see in the proof of Lemma 2.6.4 below that under appropriate conditions, there is precisely one element in $\mathsf{SolutionPiece}^n(\tau, \eta, \psi)$.

Definition 2.5.7. *We define*

$$Alg^n : [0, 2^{12\alpha n}]^{\mathbb{N}} \times \mathcal{C}^{2\cdot 2^{12\alpha n}} \times \bigcup_{k \geq c_1 L} \mathcal{C}^{[-kn,kn]} \to \mathcal{C}^{[-3\cdot 2^n, 3\cdot 2^n]}$$

as follows: If $\mathsf{SolutionPiece}^n(\tau, \eta, \psi)$ *is not empty, then we define* $Alg^n(\tau, \eta, \psi)$ *to be its lexicographically smallest element. Otherwise we define* $Alg^n(\tau, \eta, \psi)$ *to be the constant scenery* $(1)_{[-3\cdot 2^n, 3\cdot 2^n]}$.

2.6 The key algorithm reconstructs correctly

In this section, we prove Theorem 2.3.5. Throughout we fix $n \in \mathbb{N}$. We assume that $\tau \in [0, 2^{12\alpha n}]^{\mathbb{N}}$ is a sequence of $\mathcal{G}$-adapted stopping times. Recall that ε was chosen in Section 2.2.1.

2.6.1 Definition of the key events

In this subsection, we collect the definitions of all the "basic" events which we will need to prove the correctness of Alg^n. The event $B^{n,\tau}_{\text{all paths}}$ holds if the random walk traverses all paths of length $3c_1 n$ in the region where we want to do the reconstruction. $B^n_{\text{few mistakes}}$ makes sure that there are not too many mistakes in the words in $\mathsf{Collection}^n(\tau, \eta)$. $B^n_{\text{ladder diff}}$ gives a lower bound for the d-distance of two different ladder words in the neighborhood of the origin. $B^{n,\tau}_{\text{majority}}$ garanties that the majority decision f^* is not corrupted by the errors in the observations. If $B^n_{\text{outside out}}$ holds, then we can distinguish ladder words from the region where we want to reconstruct from observations which are read further outside. B^n_{signals} implies that there are "signal words" which can be read only left from a certain point $z \in \mathbb{Z}$ or only right from a certain $z \in \mathbb{Z}$; this

event allows use to reconstruct all ladder words in a region around the origin. $B^{n,\tau}_{\text{straight often}}$ guarantees that certain ladder paths are traversed often enough.

We arranged the definitions of the events in alphabetical order so that the reader can easily find them while following the proofs in the next two subsections. We suggest to have a quick look at the definitions, and then to skip ahead to the next subsection and look up definitions when needed.

Definition 2.6.1. *For $z \in \mathbb{Z}$ and n such that $c_1 n \in \mathbb{N}$, we denote by $w_{z,\to,n}$ the ladder word of length $c_1 n$ starting at z read from left to right, and by $w_{z,\leftarrow,n}$ the word $w_{z,\to,n}$ read from right to left:*

$$w_{z,\to,n} := (\xi(z+kL); k \in [0, c_1 n[)_{\to} \quad \textit{and} \quad w_{z,\leftarrow,n} := (w_{z,\to,n})_{\leftarrow}.$$

Note that $w_{z-(c_1 n-1)L,\to,n}$ is the ladder word of length $c_1 n$ ending at z.

Definition 2.6.2. *We define*

$$B^{n,\tau}_{\text{all paths}} := \left\{ \begin{array}{l} \textit{For any admissible piece of path } R \in \mathbb{Z}^{[0,3c_1 n[} \\ \textit{with starting point in } [-7\cdot 2^n, 7\cdot 2^n] \textit{ there exists} \\ t \in \cup_{k=1}^{2^{\alpha n}} [\tau_k, \tau_k + 2^{2n} - 3c_1 n] \textit{ such that } R(i) = S(t+i) \\ \textit{for all } i \in [0, 3c_1 n[\end{array} \right\}.$$

Definition 2.6.3. *We define*

$$B^{n}_{\text{few mistakes}} := \left\{ \sum_{k=t-c_1 n+1}^{t} X_k \le \varepsilon n \textit{ for all } t \in \left[c_1 n - 1, 2\cdot 2^{12\alpha n} \right[\right\}.$$

Definition 2.6.4. *We define*

$$B^{n}_{\text{ladder diff}} := \left\{ \begin{array}{l} \forall z_1, z_2 \in [-8\cdot 2^n, 8\cdot 2^n] \textit{ and } \forall i_1, i_2 \in \\ \{\leftarrow, \to\} \textit{ with } (z_1, i_1) \neq (z_2, i_2) \textit{ we have} \\ d(w_{z_1,i_1,n/3}, w_{z_2,i_2,n/3}) \ge 10\varepsilon n \end{array} \right\}.$$

Definition 2.6.5. *Let $\mathcal{I}_L$ denote the set of ladder intervals $I \subseteq [-7\cdot 2^n, 7\cdot 2^n]$ of length $c_1 n$. For $w_1, w_3 \in \mathcal{C}^{c_1 n}$ and $I \in \mathcal{I}_L$, we denote by $\mathcal{S}^{I\to}_{w_1,w_3} := \left(s_i^{I\to}; i \ge 1\right)$ ($\mathcal{S}^{I\leftarrow}_{w_1,w_3} := \left(s_i^{I\leftarrow}; i \ge 1\right)$) the sequence of all times $s \in \cup_{k=1}^{2^{\alpha n}} [\tau_k, \tau_k + 2^{2n} - 3c_1 n]$ such that $S|[s+c_1 n, s+2c_1 n[$ is a straight crossing from left to right (right to left) of I and $d(\tilde{\chi}|[s+(i-1)c_1 n, s+ic_1 n[, w_i) \le 2\varepsilon n$ for $i = 1, 3$. We assume that the components of $\mathcal{S}^{I\to}_{w_1,w_3}$ and $\mathcal{S}^{I\leftarrow}_{w_1,w_3}$ are arranged in increasing order. We define*

$$B^{n,\tau}_{\text{majority}} := \bigcap_{w_1,w_3 \in \mathcal{C}^{c_1 n}} \bigcap_{I \in \mathcal{I}_L} \left(B^{n,\tau,I\to}_{\text{maj}}(w_1,w_3) \cap B^{n,\tau,I\leftarrow}_{\text{maj}}(w_1,w_3) \right) \textit{ with}$$

$$B^{n,\tau,I\to}_{\text{maj}}(w_1,w_3) := \left\{ \begin{array}{l} \textit{If } \left|\mathcal{S}^{I\to}_{w_1,w_3}\right| \ge 2^{\gamma n}, \textit{ then } \forall j \in [1, c_1 n - 1[\\ \textit{the following holds: } \sum_{i=1}^{2^{\gamma n}} X_{s_i^{I\to}+c_1 n+j} < \\ 2^{\gamma n}/2 \end{array} \right\}$$

and $B^{n,\tau,I\leftarrow}_{\text{maj}}(w_1,w_3)$ defined analogously.

Definition 2.6.6. *We define* $B^n_{\text{outside out}} :=$

$$\left\{ \begin{array}{l} \forall z \in [-5\cdot 2^n, 5\cdot 2^n], \text{ for any admissible piece of path } R \in \\ ([-2L\cdot 2^{2n}, 2L\cdot 2^{2n}] \setminus [-6\cdot 2^n,\, 6\cdot 2^n])^{[0,c_1 n/2[} \text{ and } \forall i \in \{\leftarrow, \rightarrow\} \\ \text{we have that } d(\xi \circ R, w_{z,i,n/2}) \geq 3\varepsilon n \end{array} \right\}.$$

Definition 2.6.7. *We define* $B^n_{\text{recogn straight}} :=$

$$\left\{ \begin{array}{l} \text{For any admissible piece of path } R_1 \in [-7\cdot 2^n, 7\cdot 2^n]^{[0,c_1 n[} \text{ which} \\ \text{is not a ladder path there exists an admissible piece of path} \\ R_2 \in [-8\cdot 2^n, 8\cdot 2^n]^{[0,c_1 n[} \text{ with } R_2(0) = R_1(0),\ R_2(c_1 n - 1) = \\ R_1(c_1 n - 1) \text{ and } d(\xi \circ R_1, \xi \circ R_2) \geq 5\varepsilon n \end{array} \right\}.$$

Definition 2.6.8. *We define*

$$\begin{aligned} B^n_{\text{signals}} &:= B^n_{\text{sign,l,}\rightarrow} \cap B^n_{\text{sign,r,}\rightarrow} \cap B^n_{\text{sign,l,}\leftarrow} \cap B^n_{\text{sign,r,}\leftarrow} \quad \text{with} \\ B^n_{\text{sign,l,}\rightarrow} &:= \left\{ \begin{array}{l} \forall z \in [-6\cdot 2^n, 6\cdot 2^n] \text{ and for any admissible piece} \\ \text{of path } R \in [-2L\cdot 2^{2n}, 2L\cdot 2^{2n}]^{[0,c_1 n[} \text{ with } R(c_1 n - \\ 1) > z \text{ we have that } d(\xi \circ R, w_{z-(c_1 n-1)L,\rightarrow,n}) \geq \\ 5\varepsilon n \end{array} \right\}, \\ B^n_{\text{sign,r,}\rightarrow} &:= \left\{ \begin{array}{l} \forall z \in [-6\cdot 2^n, 6\cdot 2^n] \text{ and for any admissible piece} \\ \text{of path } R \in [-2L\cdot 2^{2n}, 2L\cdot 2^{2n}]^{[0,c_1 n[} \text{ with } R(0) < \\ z \text{ we have that } d(\xi \circ R, w_{z,\rightarrow,n}) \geq 5\varepsilon n \end{array} \right\}, \\ B^n_{\text{sign,l,}\leftarrow} &:= \left\{ \begin{array}{l} \forall z \in [-6\cdot 2^n, 6\cdot 2^n] \text{ and for any admissible piece} \\ \text{of path } R \in [-2L\cdot 2^{2n}, 2L\cdot 2^{2n}]^{[0,c_1 n[} \text{ with } R(0) > \\ z \text{ we have that } d(\xi \circ R, w_{z-(c_1 n-1)L,\leftarrow,n}) \geq 5\varepsilon n \end{array} \right\}, \\ B^n_{\text{sign,r,}\leftarrow} &:= \left\{ \begin{array}{l} \forall z \in [-6\cdot 2^n, 6\cdot 2^n] \text{ and for any admissible piece} \\ \text{of path } R \in [-2L\cdot 2^{2n}, 2L\cdot 2^{2n}]^{[0,c_1 n[} \text{ with } R(c_1 n - \\ 1) < z \text{ we have that } d(\xi \circ R, w_{z,\leftarrow,n}) \geq 5\varepsilon n \end{array} \right\}. \end{aligned}$$

Definition 2.6.9. *We denote the collection of ladder intervals* $I \subseteq [-6\cdot 2^n, 6\cdot 2^n]$ *of length* $3c_1 n$ *by* $\mathcal{J}_L$*. For* $I \in \mathcal{J}_L$*, we denote by* $\mathcal{S}_\rightarrow(I)$ *(*$\mathcal{S}_\leftarrow(I)$*) the sequence of all times* $s \in \cup_{k=1}^{2^{\alpha n}} [\tau_k, \tau_k + 2^{2n} - 3c_1 n]$ *such that* $S|[s, s+3c_1 n[$ *is a straight crossing from left to right (right to left) of* I*; we assume that the components of* $\mathcal{S}_\rightarrow(I)$ *and* $\mathcal{S}_\leftarrow(I)$ *are arranged in increasing order. We define*

$$B^{n,\tau}_{\text{straight often}} := \bigcap_{I \in \mathcal{J}_L} \left\{ |\mathcal{S}_\rightarrow(I)| \geq 2^{\gamma n} \text{ and } |\mathcal{S}_\leftarrow(I)| \geq 2^{\gamma n} \right\}.$$

2.6.2 Combinatorics

In this subsection, we prove that Alg^n reconstructs correctly in the sense that the event $E^{n,\tau}_{\text{reconstruct}}$ holds, under the assumption that $E^{n,\tau}_{\text{stop}}$ and all the "basic" events defined in the previous subsection hold. We abbreviate

$$\tilde{\chi}^n := \tilde{\chi}\big|\,[0, 2\cdot 2^{12\alpha n}[\,.$$

The task is split in four parts: Lemma 2.6.1 states a property of the elements in the set $\mathsf{PrePuzzle}^n(\tau, \tilde{\chi}^n)$. Lemma 2.6.2 shows that all words in $\mathsf{Puzzle}^n(\tau, \tilde{\chi}^n)$ which are observed while the random walk is approximately in the region of the scenery which we want to reconstruct, are ladder words. Lemma 2.6.3 states that $\mathsf{Puzzle}^n(\tau, \tilde{\chi}^n)$ contains all the ladder words we need. Finally Lemma 2.6.4 shows that the reconstruction works.

Definition 2.6.10. *We say* $(w_1, w_2, w_3) \in \mathsf{Collection}^n(\tau, \tilde{\chi}^n)$ *is* read while the random walk is walking on $J \subseteq \mathbb{Z}$ *if there exists* $t \in \cup_{k=1}^{2^{\alpha n}} [\tau_k, \tau_k + 2^{2n} - 3c_1 n]$ *such that* $S(t+j) \in J$ *for all* $j \in [0, 3c_1 n[$ *and* $w_1 w_2 w_3 = \tilde{\chi}|[t, t + 3c_1 n[$. *If we know the time* t, *we say that* (w_1, w_2, w_3) *is* read during $[t, t + 3c_1 n[$.

Definition 2.6.11. *We define* $E^{n,\tau}_{\text{pre ladder}} :=$

$$\left\{ \begin{array}{l} \textit{If } (w_1, w_2, w_3) \in \mathsf{PrePuzzle}^n(\tau, \tilde{\chi}^n) \textit{ and there exists } t \in \\ \cup_{k=1}^{2^{\alpha n}} [\tau_k, \tau_k + 2^{2n} - 3c_1 n] \textit{ such that } (w_1, w_2, w_3) \textit{ is read} \\ \textit{during } [t, t + 3c_1 n[\textit{ while the random walk is walking on} \\ [-7 \cdot 2^n, 7 \cdot 2^n], \textit{ then } S|[t + c_1 n, t + 2c_1 n[\textit{ is a ladder path.} \end{array} \right\}.$$

Lemma 2.6.1. *For all* $n \in \mathbb{N}$ *the following holds:*

$$E^{n,\tau}_{\text{pre ladder}} \supseteq B^{n,\tau}_{\text{all paths}} \cap B^n_{\text{few mistakes}} \cap B^n_{\text{recogn straight}}.$$

Proof. Suppose the events $B^{n,\tau}_{\text{all paths}}$, $B^n_{\text{few mistakes}}$, and $B^n_{\text{recogn straight}}$ hold. Let $(w_1, w_2, w_3) \in \mathsf{PrePuzzle}^n(\tau, \tilde{\chi}^n)$, and suppose there exists $t \in \cup_{k=1}^{2^{\alpha n}} [\tau_k, \tau_k + 2^{2n} - 3c_1 n]$ such that the triple (w_1, w_2, w_3) is read during $[t, t + 3c_1 n[$ while the random walk is walking on $[-7 \cdot 2^n, 7 \cdot 2^n]$.

Let $R_i(j) := S(t + (i-1)c_1 n + j)$ for $j \in [0, c_1 n[$ and $i = 1, 2, 3$. Then $|R_i(j)| \leq 7 \cdot 2^n$ for all $j \in [0, c_1 n[$ and

$$d(\xi \circ R_i, w_i) \leq \varepsilon n \quad \text{for } i = 1, 2, 3 \tag{2.6.1}$$

because $B^n_{\text{few mistakes}}$ holds. We have to show that R_2 is a ladder path. Suppose not. Since $B^n_{\text{recogn straight}}$ holds, there exists an admissible piece of path $R_2' \in [-8 \cdot 2^n, 8 \cdot 2^n]^{[0, c_1 n[}$ with the same starting and endpoint as R_2 and

$$d(\xi \circ R_2, \xi \circ R_2') \geq 5\varepsilon n. \tag{2.6.2}$$

Since $B^{n,\tau}_{\text{all paths}}$ holds and the concatenation $R_1 R_2' R_3$ is an admissible piece of path with starting point in $[-7 \cdot 2^n, 7 \cdot 2^n]$, there exists $t' \in \cup_{k=1}^{2^{\alpha n}} [\tau_k, \tau_k + 2^{2n} - 3c_1 n]$ such that $R_1 R_2' R_3(i) = S(t' + i)$ for all $i \in [0, 3c_1 n[$. Using the triangle inequality, we obtain

$$\begin{aligned} d(w_2, \tilde{\chi}|[t' + c_1 n, t' + 2c_1 n[) &\geq d(w_2, \chi|[t' + c_1 n, t' + 2c_1 n[) - \varepsilon n \\ &= d(w_2, \xi \circ R_2') - \varepsilon n \\ &\geq d(\xi \circ R_2, \xi \circ R_2') - d(w_2, \xi \circ R_2) - \varepsilon n \\ &\geq 5\varepsilon n - \varepsilon n - \varepsilon n = 3\varepsilon n; \end{aligned} \tag{2.6.3}$$

for the first inequality we used that $B^n_{\text{few mistakes}}$ holds, and for the last inequality we used (2.6.2) and (2.6.1). The fact that $B^n_{\text{few mistakes}}$ holds together with inequality (2.6.1) yields

$$\begin{aligned} d(w_1, \tilde{\chi}|[t', t' + c_1 n[) &\leq d(w_1, \chi|[t', t' + c_1 n[) + \varepsilon n \\ &= d(w_1, \xi \circ R_1) + \varepsilon n \leq 2\varepsilon n. \end{aligned}$$

By the same argument, $d(w_3, \tilde{\chi}|[t' + 2c_1 n, t' + 3c_1 n[) \leq 2\varepsilon n$. Together with (2.6.3) this contradicts $(w_1, w_2, w_3) \in \mathsf{PrePuzzle}^n(\tau, \tilde{\chi}^n)$. Hence R_2 is a ladder path. □

Definition 2.6.12. *We define*

$$\begin{aligned}
\mathsf{Puzzle}_1^n(\tau,\tilde{\chi}^n) &:= \left\{ \begin{array}{l} f^*\left(\mathsf{List}^n_{\tau,\tilde{\chi}^n}(w_1,w_2,w_3)\right) \in \mathcal{C}^{c_1 n-2} : \\ (w_1,w_2,w_3) \in \mathsf{PrePuzzle}^n(\tau,\tilde{\chi}^n) \text{ and} \\ \exists (w_1',w_2',w_3') \in \mathsf{PrePuzzle}^n(\tau,\tilde{\chi}^n) \text{ such that} \\ d(w_1,w_1') \le 2\varepsilon n,\ d(w_3,w_3') \le 2\varepsilon n \text{ and} \\ (w_1',w_2',w_3') \text{ is read while the random walk} \\ \text{is walking on } \mathbb{Z}\setminus[-6\cdot 2^n, 6\cdot 2^n]. \end{array} \right\}, \\
\mathsf{Puzzle}_2^n(\tau,\tilde{\chi}^n) &:= \mathsf{Puzzle}^n(\tau,\tilde{\chi}^n) \setminus \left[\mathsf{Puzzle}_1^n(\tau,\tilde{\chi}^n) \cup \left\{(-1)_{[1,c_1 n-2]}\right\}\right].
\end{aligned}$$

Note that $\mathsf{Puzzle}_i^n(\tau,\tilde{\chi}^n)$, $i=1,2$, together with $\left\{(-1)_{[1,c_1 n-2]}\right\}$, form a partition of the set $\mathsf{Puzzle}^n(\tau,\tilde{\chi}^n)$. If we are given an element of $\mathsf{Puzzle}^n(\tau,\tilde{\chi}^n)$, we cannot decide to which set of the partition it belongs. Nevertheless the sets $\mathsf{Puzzle}_i^n(\tau,\tilde{\chi}^n)$, $i=1,2$, will be useful in the following.

Definition 2.6.13. *We define*

$$E^{n,\tau}_{\text{only ladder}} := \left\{ \begin{array}{l} \text{If } w_2 \in \mathsf{Puzzle}_2^n(\tau,\tilde{\chi}^n), \text{ then } w_2 \preceq \\ \xi|\,[-7\cdot 2^n, 7\cdot 2^n] \text{ and } w_2 \text{ is a ladder word} \end{array} \right\}.$$

Let $c_{10} > 0$ be chosen in such a way that for all $n \ge c_{10}$

$$3c_1 nL \le 2^n. \tag{2.6.4}$$

Lemma 2.6.2. *For all $n \ge c_{10}$ the following holds:*

$$E^{n,\tau}_{\text{only ladder}} \supseteq E^{n,\tau}_{\text{pre ladder}} \cap B^n_{\text{few mistakes}} \cap B^n_{\text{ladder diff}} \cap B^{n,\tau}_{\text{majority}}.$$

Proof. Let $n \ge c_{10}$, and suppose the events $E^{n,\tau}_{\text{pre ladder}}$, $B^n_{\text{few mistakes}}$, $B^n_{\text{ladder diff}}$ and $B^{n,\tau}_{\text{majority}}$ hold. Let $(w_1,w_2,w_3) \in \mathsf{PrePuzzle}^n(\tau,\tilde{\chi}^n)$ and abbreviate $W := \mathsf{List}^n_{\tau,\tilde{\chi}^n}(w_1,w_2,w_3)$. Suppose $f^*(W) \in \mathsf{Puzzle}_2^n(\tau,\tilde{\chi}^n)$. Let $w_2' \in W$. Then there exist w_1', w_3' such that $(w_1',w_2',w_3') \in \mathsf{PrePuzzle}^n(\tau,\tilde{\chi}^n)$, $d(w_1,w_1') \le 2\varepsilon n$, and $d(w_3,w_3') \le 2\varepsilon n$. By definition of $\mathsf{Puzzle}_2^n(\tau,\tilde{\chi}^n)$, at least once the random walk is in $[-6\cdot 2^n, 6\cdot 2^n]$ while it reads (w_1',w_2',w_3'). Since the random walk jumps at most a distance of L in each step, it can move in $3c_1 n$ steps at most a distance of $3c_1 nL \le 2^n$. Hence (w_1',w_2',w_3') is observed while the random walk is walking on $[-7\cdot 2^n, 7\cdot 2^n]$. Using that $E^{n,\tau}_{\text{pre ladder}}$ holds, we obtain that w_2' is observed while the random walk is walking on a ladder word. Since $B^n_{\text{few mistakes}}$ holds, there exists a ladder word $\widehat{w}_2 \preceq \xi|\,[-7\cdot 2^n, 7\cdot 2^n]$ such that

$$d(w_2', \widehat{w}_2) \le \varepsilon n. \tag{2.6.5}$$

Suppose $w_2'' \in W$. Then by the above argument, there exists a ladder word $\bar{w}_2 \preceq \xi|\,[-7\cdot 2^n, 7\cdot 2^n]$ such that

$$d(w_2'', \bar{w}_2) \le \varepsilon n. \tag{2.6.6}$$

Since $(w_1,w_2,w_3) \in \mathsf{PrePuzzle}^n(\tau,\tilde{\chi}^n)$, we have that $d(w_2',w_2) \le 2\varepsilon n$ and $d(w_2,w_2'') \le 2\varepsilon n$. Hence

$$d(w_2', w_2'') \le 4\varepsilon n. \tag{2.6.7}$$

Using the triangle inequality, (2.6.5), (2.6.7) and (2.6.6) we obtain

$$\begin{aligned}
d(\widehat{w}_2, \bar{w}_2) &\le d(\widehat{w}_2, w_2') + d(w_2', w_2'') + d(w_2'', \bar{w}_2) \\
&\le \varepsilon n + 4\varepsilon n + \varepsilon n = 6\varepsilon n.
\end{aligned} \tag{2.6.8}$$

If $\widehat{w}_2 \neq \bar{w}_2$, then it follows from $B^n_{\text{ladder diff}}$ that $d(\widehat{w}_2, \bar{w}_2) \geq 10\varepsilon n$, which contradicts (2.6.8). Hence $\widehat{w}_2 = \bar{w}_2$.

We have shown that any $w_2' \in W$ is observed while the random walk reads the ladder word $\widehat{w}_2$. Hence for $j \in [0, c_1 n[$, $w_2'(j)$ equals $\widehat{w}_2(j)$ or an error in the observations. Since by assumption, $f^*(W) \neq (-1)_{[1,c_1 n-2]}$, W has at least $2^{\gamma n}$ components; recall the definition of f^* (Definition 2.5.1). An application of $B^{n,\tau,I}_{\text{maj}}(w_1, w_3)$ with I equal to the ladder interval underlying $\widehat{w}_2$ shows that more than half of the first $2^{\gamma n}$ words in W have j^{th} letter equal to $\widehat{w}_2(j)$. Consequently, $f(W) = \widehat{w}_2$, and since $B^n_{\text{few mistakes}}$ holds, $f^*(W) = \mathsf{Cut}(\widehat{w}_2)$. □

Definition 2.6.14. *We define* $E^{n,\tau}_{\text{all ladder}} :=$

$$\{\forall z \in [-5 \cdot 2^n, 5 \cdot 2^n] : \mathit{Cut}(w_{z,\rightarrow,n}), \mathit{Cut}(w_{z,\leftarrow,n}) \in \mathit{Puzzle}^n(\tau, \tilde{\chi}^n)\}\,.$$

Lemma 2.6.3. *For all $n \geq c_{10}$ the following holds:*

$$\begin{aligned} E^{n,\tau}_{\text{all ladder}} \quad \supseteq \quad & B^{n,\tau}_{\text{all paths}} \cap B^n_{\text{few mistakes}} \cap B^{n,\tau}_{\text{majority}} \cap B^n_{\text{signals}} \\ & \cap B^{n,\tau}_{\text{straight often}} \cap E^{n,\tau}_{\text{stop}}. \end{aligned}$$

Proof. Let $n \geq c_{10}$ and $z \in [-5 \cdot 2^n, 5 \cdot 2^n]$. Suppose the events $B^{n,\tau}_{\text{all paths}}$, $B^n_{\text{few mistakes}}$, $B^{n,\tau}_{\text{majority}}$, B^n_{signals}, $B^{n,\tau}_{\text{straight often}}$, and $E^{n,\tau}_{\text{stop}}$ hold. We will prove $\mathsf{Cut}(w_{z,\rightarrow,n}) \in \mathsf{Puzzle}^n(\tau, \tilde{\chi}^n)$. The proof for $w_{z,\leftarrow,n}$ is similar. We define

$$w_1 := w_{z-c_1 nL,\rightarrow,n}, \quad w_2 := w_{z,\rightarrow,n}, \quad w_3 := w_{z+c_1 nL,\rightarrow,n}.$$

Clearly, $w_1 w_2 w_3$ is the ladder word of length $3c_1 n$ starting at $z - c_1 nL$ and ending at $z + (2c_1 n - 1)L$. We define $R : [0, 3c_1 n[\to \mathbb{Z}$ by $R(i) = z - c_1 nL + iL$. Then R is a ladder path with starting point $z - c_1 nL \geq -6 \cdot 2^n$ and endpoint $z + (2c_1 n - 1)L \leq 6 \cdot 2^n$ by our choice of z and n; recall (2.6.4). Furthermore $\xi \circ R = w_1 w_2 w_3$. Since $B^{n,\tau}_{\text{all paths}}$ holds, there exists $t \in \cup_{k=1}^{2^{\alpha n}} [\tau_k, \tau_k + 2^{2n} - 3c_1 n]$ such that $R = S|[t, t + 3c_1 n[$. We set

$$\widehat{w}_{i,t} := \widetilde{\chi}|[t + (i-1)c_1 n, t + ic_1 n[\quad \text{for } i = 1, 2, 3. \tag{2.6.9}$$

Since $B^{n,\tau}_{\text{straight often}}$ holds, there are at least $2^{\gamma n}$ different t's with this property. Fix t. Clearly, $(\widehat{w}_{1,t}, \widehat{w}_{2,t}, \widehat{w}_{3,t}) \in \mathsf{Collection}^n(\tau, \tilde{\chi}^n)$. We want to show $(\widehat{w}_{1,t}, \widehat{w}_{2,t}, \widehat{w}_{3,t}) \in \mathsf{PrePuzzle}^n(\tau, \tilde{\chi}^n)$. The word $\widehat{w}_{i,t}$ differs from w_i only by errors in the observations. Since $B^n_{\text{few mistakes}}$ holds,

$$d(w_i, \widehat{w}_{i,t}) \leq \varepsilon n \quad \text{for } i = 1, 2, 3. \tag{2.6.10}$$

Suppose $(w_1', w_2', w_3') \in \mathsf{Collection}^n(\tau, \tilde{\chi}^n)$ and $d(w_i', \widehat{w}_{i,t}) \leq 2\varepsilon n$ for $i = 1, 3$. Then there exists $t' \in \cup_{k=1}^{2^{\alpha n}} [\tau_k, \tau_k + 2^{2n} - 3c_1 n]$ such that $w_1' w_2' w_3' = \tilde{\chi}|[t', t' + 3c_1 n[$. Using (2.6.10) and the triangle inequality, we obtain

$$d(w_i', w_i) \leq d(w_i', \widehat{w}_{i,t}) + d(\widehat{w}_{i,t}, w_i) \leq d(w_i', \widehat{w}_{i,t}) + \varepsilon n \leq 3\varepsilon n \quad \text{for } i = 1, 3.$$

We set $I_1 := [t', t' + c_1 n[$, $I_3 := [t' + 2c_1 n, t' + 3c_1 n[$. Since $B^n_{\text{few mistakes}}$ holds,

$$\begin{aligned} d(\xi \circ S|I_i, w_i) \quad & \leq \quad d(\xi \circ S|I_i, w_i') + d(w_i', w_i) \\ & \leq \quad \varepsilon n + d(w_i', w_i) \leq 4\varepsilon n \quad \text{for } i = 1, 3. \end{aligned} \tag{2.6.11}$$

Since $E^{n,\tau}_{\text{stop}}$ holds, $|S(\tau_k)| \leq 2^n$, and for all $i \in [0, 2^{2n}[$, $|S(\tau_k + i)| \leq 2^n + L \cdot 2^{2n} \leq 2L \cdot 2^{2n}$ because each jump of the random walk has length $\leq L$. Hence we can use that $B^n_{\text{sign,l},\rightarrow}$

holds for $w_1 = w_{z-c_1nL,\rightarrow,n}$ (note that $|z-L| \le 6\cdot 2^n$) and $S|I_1$ to conclude from (2.6.11) that $S(t'+c_1n-1) \le z-L$. Similarly, we can use that $B^n_{\mathrm{sign,r},\rightarrow}$ holds for $w_3 = w_{z+c_1nL,\rightarrow,n}$ (note that $|z+c_1nL| \le 6\cdot 2^n$) and $S|I_3$ to conclude that $S(t'+2c_1n) \ge z+c_1nL$. The only path of length c_1n+2 from $z-L$ to $z+c_1nL$ is the ladder path which visits precisely the points $z+iL$, $0 \le i \le c_1n-1$. Hence w_2' is observed with errors by the random walk walking on the ladder word w_2. Using the fact that $B^n_{\text{few mistakes}}$ holds and (2.6.10), we obtain

$$d(w_2', \widehat{w}_{2,t}) \le d(w_2', w_2) + d(w_2, \widehat{w}_{2,t}) \le \varepsilon n + \varepsilon n = 2\varepsilon n.$$

Consequently, $(\widehat{w}_{1,t}, \widehat{w}_{2,t}, \widehat{w}_{3,t}) \in \mathsf{PrePuzzle}^n(\tau, \tilde{\chi}^n)$. We set

$$W := \mathsf{List}^n_{\tau,\tilde{\chi}^n}(\widehat{w}_{1,t}, \widehat{w}_{2,t}, \widehat{w}_{3,t}).$$

Clearly, $W \in \mathsf{PuzzleLists}^n(\tau, \tilde{\chi}^n)$. Consider $\widehat{w}_{i,s}$ for $s \ne t$. Recall that there are at least $2^{\gamma n}-1$ different s with this property. By the triangle inequality and (2.6.10), $d(\widehat{w}_{i,s}, \widehat{w}_{i,t}) \le d(\widehat{w}_{i,s}, w_i) + d(w_i, \widehat{w}_{i,t}) \le 2\varepsilon n$ for $i = 1,2,3$. Consequently, $(\widehat{w}_{1,s}, \widehat{w}_{2,s}, \widehat{w}_{3,s}) \in W$, and we conclude that W has at least $2^{\gamma n}$ components.

Suppose $w_2' \in W$. Then there exist w_1', w_3' with $d(w_i', \widehat{w}_{i,t}) \le 2\varepsilon n$ for $i = 1,3$ and $(w_1', w_2', w_3') \in \mathsf{PrePuzzle}^n(\tau, \tilde{\chi}^n)$. We have shown above (after (2.6.10)) that under these conditions, w_2' must be observed while the random walk reads the ladder word w_2. In particular, for $j \in [0, c_1n[$, $w_2'(j) = w_2(j)$ or $w_2'(j)$ is an error in the observations. Since $B^{n,\tau,I}_{\mathrm{maj}}(\widehat{w}_{1,t}, \widehat{w}_{3,t})$ holds for the ladder interval $I = \{z+iL; i \in [0, c_1n[\}$, in more than half of the words in W the j^{th} letter equals $w_2(j)$. Consequently, the j^{th} letter of $f(W)$ equals $w_2(j)$, and we have proved that $\mathsf{Cut}(w_2) \in \mathsf{Puzzle}^n(\tau, \tilde{\chi}^n)$. □

Recall the definition of $E^{n,\tau}_{\text{reconstruct}}$ from Theorem 2.3.5.

Lemma 2.6.4. *For all $n \ge c_{10}$ with c_{10} as in (2.6.4) the following holds:*

$$\begin{aligned} E^{n,\tau}_{\text{reconstruct}} \supseteq\ & E^{n,\tau}_{\text{only ladder}} \cap E^{n,\tau}_{\text{all ladder}} \cap B^n_{\text{few mistakes}} \cap B^n_{\text{ladder diff}} \\ & \cap B^n_{\text{outside out}} \cap E^{n,\tau}_{\text{stop}}. \end{aligned}$$

Proof. Let $n \ge c_{10}$, and suppose all the events $E^{n,\tau}_{\text{only ladder}}$, $E^{n,\tau}_{\text{all ladder}}$, $B^n_{\text{ladder diff}}$, $B^n_{\text{few mistakes}}$, $B^n_{\text{outside out}}$, and $E^{n,\tau}_{\text{stop}}$ hold. Let $\psi \in \mathcal{C}^{[-kn,kn]}$ for some $k \ge c_1L$, and suppose $\psi \preceq \xi|[-2^n, 2^n]$. There exist $a \in [-2^n, 2^n]$ and $b \in \{-1, 1\}$ such that for all $j \in [-kn, kn]$

$$\psi(j) = \xi(a+bj) \quad \text{and} \quad a+bj \in [-2^n, 2^n]\,. \tag{2.6.12}$$

First we show $w := (\xi(a+bj); j \in [-3\cdot 2^n, 3\cdot 2^n]) \in \mathsf{SolutionPiece}^n(\tau, \tilde{\chi}^n, \psi)$. By (8.5.13), $\psi = w|[-kn, kn]$. Let $I \subseteq [-3\cdot 2^n, 3\cdot 2^n]$ be a ladder interval of length c_1n-2. The image of I under the map $j \mapsto a+bj$ is a ladder interval which is contained in $[-4\cdot 2^n, 4\cdot 2^n]$ because $|a| \le 2^n$. Since $E^{n,\tau}_{\text{all ladder}}$ holds, $(w|I)_\rightarrow \in \mathsf{Puzzle}^n(\tau, \tilde{\chi}^n)$. Consequently, $w \in \mathsf{SolutionPiece}^n(\tau, \tilde{\chi}^n, \psi)$, and in particular, $\mathsf{SolutionPiece}^n(\tau, \tilde{\chi}^n, \psi)$ is not empty.

It remains to show that $\xi|[-2^n, 2^n] \preceq w \preceq \xi|[-4\cdot 2^n, 4\cdot 2^n]$ for any element $w \in \mathsf{SolutionPiece}^n(\tau, \tilde{\chi}^n, \psi)$. Let $w \in \mathsf{SolutionPiece}^n(\tau, \tilde{\chi}^n, \psi)$. Then $w|[-kn, kn] = \psi$, and it follows from (8.5.13) that for all $j \in [-kn, kn]$

$$w(j) = \xi(a+bj). \tag{2.6.13}$$

Suppose we prove (8.5.14) for all $j \in [-3\cdot 2^n, 3\cdot 2^n]$. Then we know there is precisely one element in $\mathsf{SolutionPiece}^n(\tau,\tilde\chi^n,\psi)$. Since $\psi \preceq \xi|[-2^n,2^n]$, there are more than $2\cdot 2^n$ letters to the left and to the right of ψ in w, and consequently $\xi|[-2^n,2^n] \preceq w$. On the other hand, in w, there are less than $3\cdot 2^n$ letters to the left and to the right of ψ. Hence $w \preceq \xi|[-4\cdot 2^n, 4\cdot 2^n]$.

Thus, to finish the proof, it suffices to verify (8.5.14) for all $j \in [-3\cdot 2^n, 3\cdot 2^n]$. We have already seen that (8.5.14) holds for all $j \in [-kn, kn]$. Suppose we know that (8.5.14) holds for all $j \in [-s,s]$ for some $s \in [kn, 3\cdot 2^n - 1]$. We set

$$\begin{aligned} w_l &:= (w|I_l)_\rightarrow \text{ with } I_l := (-s-1+iL; i \in [0, c_1 n - 2[)\,, \\ w_r &:= (w|I_r)_\rightarrow \text{ with } I_r := (s+1+(i-c_1n+3)L; i \in [0, c_1 n - 2[)\,; \end{aligned}$$

note that I_l denotes the ladder interval of length c_1n-2 which contains $-s-1$ as leftmost point, and I_r denotes the ladder interval of length c_1n-2 which contains $s+1$ as rightmost point. The words w_l and w_r are well defined because $c_1nL \le |\psi| = 2kn+1$. Since $w \in \mathsf{SolutionPiece}^n(\tau,\tilde\chi^n,\psi)$, we have $w_l, w_r \in \mathsf{Puzzle}^n(\tau,\tilde\chi^n)$. Note that w_l and w_r have both precisely c_1n-3 points in common with $w|[-s,s]$; w_l extends $w|[-s,s]$ one letter to the left, and w_r extends $w|[-s,s]$ one letter to the right.

Suppose $w_l \in \mathsf{Puzzle}_1^n(\tau,\tilde\chi^n)$. Then we have $w_l = f^*(W)$ for some $W = \mathsf{List}^n_{\tau,\tilde\chi^n}(w_1,w_2,w_3)$ and there exists $(w_1',w_2',w_3') \in \mathsf{PrePuzzle}^n(\tau,\tilde\chi^n)$ such that $d(w_i,w_i') \le 2\varepsilon n$, for $i=1,3$ and (w_1',w_2',w_3') is read while the random walk is walking on $\mathbb{Z}\setminus[-6\cdot 2^n, 6\cdot 2^n]$. Thus, there exists $t \in \cup_{k=1}^{2^{\alpha n}}[\tau_k, \tau_k + 2^{2n} - 3c_1n]$ such that $|S(t+j)| > 6\cdot 2^n$ for all $j \in [0, 3c_1n[$ and $w_2' = \tilde\chi|J$ with $J = [t+c_1n, t+2c_1n[$. Using that $E^{n,\tau}_{\text{stop}}$ holds, we know that $|S(\tau_k)| \le 2^n$ for all k. Since the random walk jumps a distance $\le L$ in each step, it follows that $|S(t+j)| \le 2^n + L\cdot 2^{2n} \le 2L\cdot 2^{2n}$ for all $j \in [0,3c_1n[$. For a word $w = w_1w_2\dots w_m \in \mathcal{C}^m$ of length $m \ge c_1n/2$, we define $\mathsf{Last}(w) := w_{m-c_1n/2+1}\dots w_m$ to be the word consisting of the last $c_1n/2$ letters of w. Let $z \in [-5\cdot 2^n, 5\cdot 2^n]$ and $i \in \{\leftarrow,\rightarrow\}$. Since $B^n_{\text{few mistakes}}$ and $B^n_{\text{outside out}}$ hold, we obtain

$$\begin{aligned} d(\mathsf{Last}(\mathsf{Cut}(w_2')), w_{z,i,n/2}) &= d(\mathsf{Last}(\mathsf{Cut}(\tilde\chi|J)), w_{z,i,n/2}) \\ &\ge d(\mathsf{Last}(\mathsf{Cut}(\chi|J)), w_{z,i,n/2}) - \varepsilon n \ge 3\varepsilon n - \varepsilon n = 2\varepsilon n. \end{aligned} \tag{2.6.14}$$

By definition of $f^*(W)$, $d(\mathsf{Cut}(f(W)), \mathsf{Cut}(w)) \le \varepsilon n$ for all $w \in W$. Hence

$$d(\mathsf{Last}(w_l), \mathsf{Last}(\mathsf{Cut}(w_2'))) \le \varepsilon n. \tag{2.6.15}$$

Combining (2.6.14) and (2.6.15), we obtain

$$\begin{aligned} d(\mathsf{Last}(w_l), w_{z,i,n/2}) &\ge d(\mathsf{Last}(\mathsf{Cut}(w_2')), w_{z,i,n/2}) - d(\mathsf{Last}(w_l), \mathsf{Last}(\mathsf{Cut}(w_2'))) \\ &\ge 2\varepsilon n - \varepsilon n = \varepsilon n. \end{aligned} \tag{2.6.16}$$

Recall that w_l is a ladder word of w of length c_1n-2 and the c_1n-3 right-most letters of w_l overlap with $w|[-s,s]$. Using that (8.5.14) holds for all $j \in [-s,s]$ together with $|a| \le 2^n$ and $|s| \le 3\cdot 2^n$, yields $\mathsf{Last}(w_l) \preceq \xi|\,[-4\cdot 2^n, 4\cdot 2^n]$. This contradicts (2.6.16), which implies that $\mathsf{Last}(w_l)$ is different from any ladder word of $\xi|[-4\cdot 2^n, 4\cdot 2^n]$. We conclude $w_l \in \mathsf{Puzzle}_2^n(\tau,\tilde\chi^n)$. Since $E^{n,\tau}_{\text{only ladder}}$ holds, $w_l \preceq \xi|\,[-7\cdot 2^n, 7\cdot 2^n]$, and w_l is a ladder word of ξ.

Suppose (8.5.14) does not hold for $j = -s-1$. Let $I_{l,\xi}$ denote the image of I_l under the map $j \mapsto a+bj$. Then $\xi|I_{l,\xi} \ne w_l$; more precisely, $\xi|I_{l,\xi}$ and w_l disagree in precisely one

point, namely the leftmost point $\xi(a+b(-s-1)) \neq w_l(0)$. Thus we found two ladder words of length $c_1 n - 2$ in $\xi|[-7 \cdot 2^n, 7 \cdot 2^n]$ which disagree in precisely one point. Consequently, there exist $z, z' \in [-8 \cdot 2^n, 8 \cdot 2^n]$, $i, i' \in \{\leftarrow, \rightarrow\}$ with $(z,i) \neq (z',i')$ such that $\xi|I_{l,\xi} = \mathsf{Cut}(w_{z,i,n})$ and $w_l = \mathsf{Cut}(w_{z',i',n})$. Consequently, there exist $z_1, z_2 \in [-8 \cdot 2^n, 8 \cdot 2^n]$, $i_1, i_2 \in \{\leftarrow, \rightarrow\}$ with $(z_1, i_1) \neq (z_2, i_2)$ such that the two ladder words consisting of the last $c_1 n/3$ letters of $\xi|I_{l,\xi}$ and w_l respectively, equal $w_{z_1,i_1,n/3}$, $w_{z_2,i_2,n/3}$, respectively. Since $B^n_{\text{ladder diff}}$ holds, $w_{z_1,i_1,n/3} \neq w_{z_2,i_2,n/3}$ which is a contradiction. We conclude that (8.5.14) holds for $j = -s-1$.

To see that (8.5.14) holds for $j = s+1$, one applies the above argument with $\bar{w}$ defined by $\bar{w}(j) := w(-j)$ for $j \in [-3 \cdot 2^n, 3 \cdot 2^n]$ in place of w. By the induction principle, (8.5.14) holds for all $j \in [-3 \cdot 2^n, 3 \cdot 2^n]$. □

2.6.3 The basic events have high probabilities

In this subsection, we prove that the events $B^n_{\dots}$ defined in Subsection 2.6.1 have a probability which is exponentially small in n. For some events $B^n_{\dots}$ this is only true under the assumption that $E^{n,\tau}_{\text{stop}}$ holds, i.e. if the stopping times stop correctly. We treat the events from Subsection 2.6.1 in alphabetical order.

Recall that unless otherwise stated, constants depend only on the distribution of the random walk increments and the number of colors of the scenery. In particular, the constants c_i in this section do not depend on n.

Lemma 2.6.5. *There exists a constant $c_{11} > 0$ such that for all $n \geq c_{11}$,*

$$P\left(E^{n,\tau}_{\text{stop}} \setminus B^{n,\tau}_{\text{all paths}}\right) \leq e^{-n}.$$

Proof. We have $P(S_0 = S_2 = 0) > 0$ because the random walk has a positive probability to make first a step of maximal length L to the right and then a step of maximal length L to the left. Hence 2 divides the period of the random walk, and the period must be 1 or 2. Therefore there exists $c_{12} > 0$ such that for all $n \geq c_{12}$ and for all $x, z \in [-7 \cdot 2^n, 7 \cdot 2^n]$, the random walk starting at x can reach z with positive probability in 2^{2n-1} or $2^{2n-1}+1$ steps:

$$P_x\left(S(2^{2n-1}) = z \text{ or } S(2^{2n-1}+1) = z\right) > 0. \tag{2.6.17}$$

We denote by $\mathcal{R}$ the set of all admissible pieces of path $R \in \mathbb{Z}^{[0,3c_1n[}$ with starting point in $[-7 \cdot 2^n, 7 \cdot 2^n]$. For $R \in \mathcal{R}$ and $t \in \mathbb{N}_0$, we define the event

$$E(t,R) := \{S(t+i) = R(i)\ \forall i \in [0, 3c_1 n[\text{ or } S(t+1+i) = R(i)\ \forall i \in [0, 3c_1 n[\}\,.$$

Let $n \geq \max\{c_{12}, c_{10}\}$ with c_{10} as in (2.6.4), and let $k \in [1, 2^{\alpha n}]$. We set $t_{k,n} := \tau_k + 2^{2n-1}$ and we define random variables $Y_k(R)$ as follows: If $|S(\tau_k)| \leq 2^n$ and $E(t_{k,n}, R)$ does not hold, then we set $Y_k(R) = 0$. Otherwise we set $Y_k(R) = 1$. Using the definitions of $E^{n,\tau}_{\text{stop}}$ and $B^{n,\tau}_{\text{all paths}}$, we see that

$$E^{n,\tau}_{\text{stop}} \setminus B^{n,\tau}_{\text{all paths}} \subseteq \bigcup_{R \in \mathcal{R}} E^{n,\tau}_{\text{stop}} \cap \left\{\sum_{k=1}^{2^{\alpha n}} Y_k(R) = 0\right\} \subseteq \bigcup_{R \in \mathcal{R}} E_{2^{\alpha n}}(R) \tag{2.6.18}$$

with

$$E_M(R) := \bigcap_{k=1}^{M} \{|S_{\tau_k}| \leq 2^n, \tau_{k-1} + 2 \cdot 2^n \leq \tau_k, Y_k(R) = 0\}$$

for $M \in [1, 2^{\alpha n}]$. Let $R \in \mathcal{R}$. Since $n \geq c_{10}$, we have $3c_1 nL \leq 2^n$ by (2.6.4). Hence $t_{k,n} + 1 + 3c_1 nL = \tau_k + 1 + 2^{2n-1} + 3c_1 nL \leq \tau_k + 2^{2n}$. Consequently, $\{\tau_k + 2 \cdot 2^{2n} < \tau_{k+1}\} \cap E(t_{k,n}, R) \in \mathcal{F}_{\tau_{k+1}}$; here $\mathcal{F}_k := \sigma(S_i, \tilde{\chi}_i; i \in [0,k])$ denotes the natural filtration of random walk and observations with errors. Using the strong Markov property at time τ_M, we obtain

$$\begin{aligned} P[E_M(R)] &= P\left[E_{M-1}(R) \cap \left\{|S_{\tau_M}| \leq 2^n, \tau_{M-1} + 2^{n+1} \leq \tau_M, Y_M(R) = 0\right\}\right] \\ &\leq P\left[E_{M-1}(R) \cap \{|S(\tau_M)| \leq 2^n\} \cap E(t_{M,n}, R)^c\right] \\ &\leq P\left[E_{M-1}(R) \cap \{|S(\tau_M)| \leq 2^n\} P_{S(\tau_M)}(E(2^{2n-1}, R)^c)\right] \\ &\leq P[E_{M-1}(R)] \max_{x \in [-2^n, 2^n]} P_x[E(2^{2n-1}, R)^c]. \end{aligned}$$

An induction argument yields

$$P\left(E_{2^{\alpha n}}(R)\right) \leq \left[\max_{x \in [-2^n, 2^n]} P_x(E(2^{2n-1}, R)^c)\right]^{2^{\alpha n}}. \tag{2.6.19}$$

To estimate the right-hand side of (2.6.19), let $b \in \mathbb{N}$ be minimal and let $h \in \mathbb{N}$ be maximal such that $P(S_1 - S_0 \in b + h\mathbb{Z}) = 1$. We set $\sigma^2 := E[(S_1 - S_0)^2]$, and $\mathcal{L}_m := \{(mb + hy)/\sqrt{m} : y \in \mathbb{Z}\}$. By the local central limit theorem ([6], page 132, Theorem (5.2)),

$$\lim_{m \to \infty} \sup_{y \in \mathcal{L}_m} \left| \frac{\sqrt{m}}{h} P\left(\frac{S_m}{\sqrt{m}} = y\right) - \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{y^2}{2\sigma^2}\right) \right| = 0.$$

We apply this with $m \in \{2^{2n-1}, 2^{2n-1}+1\}$, $y := (R_0 - x)/\sqrt{m}$ and R_0 equal to the starting point of R. Note that $|R_0| \leq 7 \cdot 2^n$ so that $|R_0 - x|/\sqrt{m} \leq 16$ for all $x \in [-2^n, 2^n]$. Hence $\min_{x \in [-2^n, 2^n], R \in \mathcal{R}} \exp\left(-\frac{(R_0 - x)^2}{2m\sigma^2}\right) > 0$. We conclude that there exist constants $c_{13} > 0$ and $c_{14} \geq \max\{c_{12}, c_{10}\}$ such that for all $n \geq c_{14}$

$$\begin{aligned} &\min_{x \in [-2^n, 2^n], R \in \mathcal{R}} P_x\left(S(2^{2n-1}) = R_0 \text{ or } S(2^{2n-1}+1) = R_0\right) \\ &= \min_{x \in [-2^n, 2^n], R \in \mathcal{R}} P\left(\frac{S(2^{2n-1})}{\sqrt{2^{2n-1}}} = \frac{R_0 - x}{\sqrt{2^{2n-1}}} \text{ or } \frac{S(2^{2n-1}+1)}{\sqrt{2^{2n-1}+1}} = \frac{R_0 - x}{\sqrt{2^{2n-1}+1}}\right) \\ &\geq c_{13} 2^{-n} \end{aligned} \tag{2.6.20}$$

We set $\mu_{\min} := \min\{\mu(j) : j \in \mathcal{M}\}$; recall that μ is the distribution of the random walk increments $S_{k+1} - S_k$. The probability that the random walk starting at R_0 follows the path R for the next $3c_1 n - 1$ steps is bounded below by $\mu_{\min}^{3c_1 n - 1}$. Thus, (2.6.20) yields

$$\min_{x \in [-2^n, 2^n], R \in \mathcal{R}} P_x(E(2^{2n-1}, R)) \geq c_{13} 2^{-n} \mu_{\min}^{3c_1 n - 1} = c_{15} 2^{-n} \mu_{\min}^{3c_1 n}$$

with $c_{15} := c_{13} \mu_{\min}^{-1}$. Combining the last inequality with (2.6.18) and (2.6.19), we obtain

$$\begin{aligned} P\left(E_{\text{stop}}^{n,\tau} \setminus B_{\text{all paths}}^{n,\tau}\right) &\leq |\mathcal{R}| \left(1 - c_{15} 2^{-n} \mu_{\min}^{3c_1 n}\right)^{2^{\alpha n}} \\ &\leq (14 \cdot 2^n + 1) |\mathcal{M}|^{3c_1 n - 1} \exp\left(2^{\alpha n} \ln\left(1 - c_{15} 2^{-n} \mu_{\min}^{3c_1 n}\right)\right). \end{aligned} \tag{2.6.21}$$

Note that choosing a path in $\mathcal{R}$ one has $14 \cdot 2^n + 1$ possible starting points and $|\text{supp}(\mu)| = |\mathcal{M}|$ possibilities for each step of the path. Using the estimate $\ln(1 - x) \leq -x$, we obtain

$$(2.6.21) \leq 2^{n+4} |\mathcal{M}|^{3c_1 n} \exp\left[-c_{15} 2^{(\alpha - 1)n} \mu_{\min}^{3c_1 n}\right] = 2^{n+4} |\mathcal{M}|^{3c_1 n} \exp\left[-c_{15} e^{c_{16} n}\right]$$

and the last expression is $\leq e^{-n}$ for all n sufficiently large because $c_{16} = (\alpha - 1)\ln 2 + 3c_1 \ln \mu_{\min} > 0$ by our choice of α. □

Lemma 2.6.6. *There exist $\delta_4 > 0$ such that for all $n \in \mathbb{N}$ and $\delta \in]0, \delta_4[$*

$$P_\delta\left((B^n_{\text{few mistakes}})^c\right) \leq e^{-n}.$$

Proof. Using Definition 2.6.3 and our convention $\varepsilon = c_1\bar{\varepsilon}$ we obtain

$$(B^n_{\text{few mistakes}})^c = \bigcup_{t \in [c_1 n - 1, 2\cdot 2^{12\alpha n}[} \left\{ \sum_{k=t-c_1 n+1}^{t} X_k > c_1 \bar{\varepsilon} n \right\}. \tag{2.6.22}$$

Recall that X_k, $k \geq 0$, are i.i.d. Bernoulli random variables with parameter δ under P_δ. Hence $E_\delta\left[\sum_{k=t-c_1 n+1}^{t} X_k\right] = c_1 \delta n$. By the large deviation principle (see e.g. [3]), we have for all $\delta \in]0, \bar{\varepsilon}[$

$$P_\delta\left(\sum_{k=t-c_1 n+1}^{t} X_k > c_1 \bar{\varepsilon} n\right) \leq \exp\left(-I_\delta(\bar{\varepsilon} - \delta) c_1 n\right) \tag{2.6.23}$$

with rate function

$$I_\delta(x) = (1-x)\log\left(\frac{1-x}{1-\delta}\right) + x \log\left(\frac{x}{\delta}\right), x \in]0,1[. \tag{2.6.24}$$

Combining (2.6.22) with (2.6.23) we obtain for all $\delta \in]0, \bar{\varepsilon}[$

$$P_\delta\left((B^n_{\text{few mistakes}})^c\right) \leq \exp\left([1 + 12\alpha n]\ln 2 - I_\delta(\bar{\varepsilon} - \delta) c_1 n\right).$$

Since

$$\lim_{\delta \to 0} I_\delta(\bar{\varepsilon} - \delta) = \lim_{\delta \to 0} (1 - \bar{\varepsilon} + \delta)\log\left[\frac{1 - \bar{\varepsilon} + \delta}{1 - \delta}\right] + (\bar{\varepsilon} - \delta)\log\left[\frac{\bar{\varepsilon} - \delta}{\delta}\right] = +\infty,$$

there exists $\delta_4 \in]0, \bar{\varepsilon}[$ such that $[1 + 12\alpha]\ln 2 - I_\delta(\bar{\varepsilon} - \delta)c_1 < -1$ for all $\delta \in]0, \delta_4[$. The assertion of the lemma follows. □

We will need the following lemma in the proofs of Lemmas 2.6.8, 2.6.10, and 2.6.13.

Lemma 2.6.7. *There exist $\varepsilon_1, c_{17}(\varepsilon') > 0$ such that for all m with $c_1 m \in \mathbb{N}$, $\varepsilon' \in]0, \varepsilon_1[$, $w \in \mathcal{C}^{[0, c_1 m[}$, and for any admissible piece of path $R \in \mathbb{Z}^{[0, c_1 m[}$ the following holds:*

$$P\left(d(\xi \circ R, w) < c_1 \varepsilon' m\right) \leq c_{17}(\varepsilon')(c_2)^{c_1 m} \max_J P((\xi \circ R) | J = w | J),$$

where the maximum is taken over all subsets $J \subseteq [0, c_1 m[$ with cardinality $|J| = c_1 m - \lfloor c_1 \varepsilon' m \rfloor$ and c_2 is as in Section 2.2.1.

Proof. Let m be such that $c_1 m \in \mathbb{N}$, let $w \in \mathcal{C}^{[0, c_1 m[}$, and let $R \in \mathbb{Z}^{[0, c_1 m[}$ be an admissible piece of path. If $d(\xi \circ R, w) < c_1 \varepsilon' m$, then $c_1 m - \lfloor c_1 \varepsilon' m \rfloor$ letters of $\xi \circ R$ and w agree. Since there are $\binom{c_1 m}{\lfloor c_1 \varepsilon' m \rfloor}$ possibilities of choosing $c_1 m - \lfloor c_1 \varepsilon' m \rfloor$ out of $c_1 m$ letters, we have

$$P\left(d(\xi \circ R, w) < c_1 \varepsilon' m\right) \leq \binom{c_1 m}{\lfloor c_1 \varepsilon' m \rfloor} \max_J P((\xi \circ R) | J = w | J),$$

where the maximum is taken over all subsets $J \subseteq [0, c_1 m[$ with cardinality $c_1 m - \lfloor c_1 \varepsilon' m \rfloor$. By Stirling's formula ([1], p.24, formula (3.9)) we have for $k \in \mathbb{N}$, $k! = \sqrt{2\pi} k^{k+1/2} e^{-k+\theta(k)}$ with $\theta(k) \in]0,1[$ and $\lim_{k\to\infty} \theta(k) = 0$. Thus

$$\binom{c_1 m}{\lfloor c_1 \varepsilon' m \rfloor} \leq c_{17}(\varepsilon') \varphi\left(\frac{\lfloor c_1 \varepsilon' m \rfloor}{c_1 m}\right)^{c_1 m}$$

with $\varphi(x) = x^{-x}(1-x)^{-(1-x)}$ and some constant $c_{17}(\varepsilon') > 0$ independent of m. Note that φ is continuous at 0 with $\varphi(0) = 1$, and recall that $c_2 \in]1, C/(C-1)[$. There exists ε_1 such that $\varphi(x) < c_2$ for all $x \in]0, \varepsilon_1[$. Note that $\lfloor c_1 \varepsilon' m \rfloor / (c_1 m) \leq \varepsilon'$. The claim follows. □

Lemma 2.6.8. *There exists a constant $c_{18} > 0$ such that for all $n \in \mathbb{N}$*

$$P\left((B^n_{\text{ladder diff}})^c\right) \leq c_{18} e^{-n}.$$

Proof. Let

$$\mathcal{J} := \left\{(z_1, i_1, z_2, i_2) \in ([-8 \cdot 2^n, 8 \cdot 2^n] \times \{\leftarrow, \rightarrow\})^2 : (z_1, i_1) \neq (z_2, i_2)\right\}.$$

By Definition 2.6.4,

$$(B^n_{\text{ladder diff}})^c = \bigcup_{(z_1,i_1,z_2,i_2)\in\mathcal{J}} \{d(w_{z_1,i_1,n/3}, w_{z_2,i_2,n/3}) < 10\varepsilon n\}. \tag{2.6.25}$$

Let $(z_1, i_1, z_2, i_2) \in \mathcal{J}$. For $k = 1, 2$ we set $o_k := +1$ if $i_k = \rightarrow$, $o_k := -1$ if $i_k = \leftarrow$, and we set $f_k(j) := z_k + o_k j L$ for $j \in [0, c_1 n/3[$. First we prove that there exists a subset $J \subseteq [0, c_1 n/3[$ of cardinality $|J| \geq c_1 n/9$ such that

$$f_1(J) \cap f_2(J) = \varnothing. \tag{2.6.26}$$

We distinguish two cases. Case $z_1 = z_2$: By assumption, $i_1 \neq i_2$. Hence $o_1 \neq o_2$, and we conclude that (2.6.26) is satisfied for $J =]0, c_1 n/3[$.

Case $z_1 \neq z_2$: We show by induction over $k \in [1, c_1 n/9]$ that there exists J with $|J| \geq k$ such that (2.6.26) holds. For $k = 1$ the set $J = \{0\}$ has the required property. Suppose there exists J' with $|J'| = k \in [1, c_1 n/9 - 1]$ such that (2.6.26) holds. The sets $J'_i := f_i(J')$, $i = 1, 2$, have cardinality $|J'_i| = |J'| \leq c_1 n/9 - 1$. We set

$$\bar{J} := \{j \in [0, c_1 n/3[: f_1(j) \notin J'_1 \cup J'_2, f_2(j) \notin J'_1, \text{ and } f_1(j) \neq f_2(j)\}.$$

Then $|\bar{J}| \geq c_1 n/3 - |J'_1 \cup J'_2| - |J'_1| - 1 = c_1 n/3 - 3(c_1 n/9 - 1) - 1 = 2$; note that there exists at least one j with $f_1(j) \neq f_2(j)$. In particular $\bar{J}$ is not empty. Let $j \in \bar{J}$, and set $J := J' \cup \{j\}$. Since $f_1(j) \notin J'_1$, we have $|J| = |J'| + 1$. It follows from $f_1(j) \notin J'_2 \cup \{f_2(j)\}$ that $f_1(j) \notin f_2(J)$. Similarly, it follows from $f_2(j) \notin J'_1 \cup \{f_1(j)\}$ that $f_2(j) \notin f_1(J)$, and we have proved that (2.6.26) holds for J. By the induction principle, (2.6.26) holds for a set $J \subseteq [0, c_1 n/3[$ of cardinality $|J| = c_1 n/9$.

Let $J \subseteq [0, c_1 n/3[$ with $|J| = c_1 n/9$ such that (2.6.26) holds. Then the words $w_{z_k,i_k,n/3}|f_k(J)$, $k = 1, 2$, are independent. Note that $P(\xi_k = \xi_{k'}) = 1/C$ for $k \neq k'$. We use Lemma 2.6.7 with $m := n/9$, $\varepsilon' := 90\varepsilon/c_1$ and R equal to the ladder path underlying $w_{z_1,i_1,n/3}$ to obtain

$$\begin{aligned} & P(d(w_{z_1,i_1,n/3}, w_{z_2,i_2,n/3}) < 10\varepsilon n) \\ \leq\ & P\left(d(w_{z_1,i_1,n/3}|f_1(J), w_{z_2,i_2,n/3}|f_2(J)) < 10\varepsilon n\right) \\ \leq\ & c_{17}(90\varepsilon/c_1)(c_2)^{c_1 n/9} C^{\lfloor 10\varepsilon n \rfloor - c_1 n/9}. \end{aligned} \tag{2.6.27}$$

Since the intersection in (2.6.25) is taken over $4(16\cdot 2^n+1)^2$ possible pairs (z_1,i_1), (z_2,i_2), it follows from (2.6.27) that

$$P((B^n_{\text{ladder diff}})^c) \le 4(16\cdot 2^n+1)^2 c_{17}(90\varepsilon/c_1)(c_2)^{c_1 n/9} C^{\lfloor 10\varepsilon n\rfloor - c_1 n/9}.$$

Note that $C^{\lfloor 10\varepsilon n\rfloor} \le \exp(10\varepsilon n \ln C)$. Let $c_{18}>0$ be chosen in such a way that $4(16\cdot 2^n+1)^2 c_{17}(90\varepsilon/c_1) \le c_{18}2^{2n}$. Then

$$P((B^n_{\text{ladder diff}})^c) \le c_{18} e^{n[2\ln 2 + 10\varepsilon \ln C + (c_1/9)[\ln c_2 - \ln C]]}.$$

Since $2\ln 2 + 10\varepsilon \ln C + (c_1/9)[\ln c_2 - \ln C] < -1$ by our choice of ε and c_1, the claim follows. □

Lemma 2.6.9. *There exist constants $c_{19}, \delta_5 > 0$ such that for all $n \ge c_{19}$ and $\delta \in]0,\delta_5[$*

$$P_\delta\left(\left(B^{n,\tau}_{\text{majority}}\right)^c\right) \le e^{-n}.$$

Proof. Recall the notation from Definition 2.6.5. Let $w_1, w_3 \in \mathcal{C}^{c_1 n}$, $I \in \mathcal{I}_L$. Let r_i, $i \ge 1$, denote all the times $s \in \cup_{k=1}^{2^{\alpha n}} [\tau_k + c_1 n, \tau_k + 2^{2n} - 2c_1 n]$ such that $S|[r_i, r_i + c_1 n[$ is a straight crossing of I from left to right. Clearly, the intervals $[r_i, r_i + c_1 n[$, $i \ge 1$, are pairwise disjoint. Let $\mathcal{H} := \sigma(r_i, \tau_i; i \ge 1)$. Since S and X are independent, we know that conditioned on $\mathcal{H}$, the random variables X_{r_i+j}, $i \ge 1$, $j \in [0, c_1 n[$, are i.i.d. Bernoulli with parameter δ under P_δ.

We obtain the random variables $s_i^{I\to} + c_1 n$, $i \ge 1$, from r_i, $i \ge 1$, by checking whether $d(\tilde{\chi}|[r_i + (k-2)c_1 n, r_i + (k-1)c_1 n[, w_k) \le 2\varepsilon n$ for $k = 1, 3$. Since at time $r_i + c_1 n - 1$ the random walk is at the right endpoint of I and at time r_{i+1} at the left endpoint of I, the time interval $[r_i + c_1 n - 1, r_{i+1}]$ has length $\ge c_1 n$. Consequently, the time intervals $[r_i, r_i + c_1 n[$, $[r_{i+1}, r_{i+1} + c_1 n[$ have a distance $\ge c_1 n - 2$ from each other. Since ξ, S, Y are independent of X, we conclude that $\tilde{\chi}|[s_i^{I\to} + k c_1 n, s_i^{I\to} + (k+1)c_1 n[$, $k = 0, 2$, $i \ge 1$, is independent of $\sigma(X_{s_i^{I\to}+c_1 n+j}; j \in [1, c_1 n - 1[, i \ge 1)$. Hence conditioned on $\bar{\mathcal{H}} := \sigma\left(s_i^{I\to} + c_1 n, \tau_i, \tilde{\chi}|[s_i^{I\to} + k c_1 n, s_i^{I\to} + (k+1)c_1 n[; i \ge 1, k = 0, 2\right)$ the random variables $X_{s^{I\to}+c_1 n+j}$, $j \in [1, c_1 n - 1[$, are i.i.d. Bernoulli with parameter δ under P_δ.

By the large deviation principle (see e.g. [3]), we have for all $\delta \in]0, 1/2[$ and $n \in \mathbb{N}$ P_δ-almost surely on the set $\left\{|\mathcal{S}^{I\to}_{w_1,w_3}| \ge 2^{\gamma n}\right\}$

$$P_\delta\left(\sum_{i=1}^{2^{\gamma n}} X_{s_i + c_1 n + j} \ge 2^{\gamma n}/2 \,\middle|\, \bar{\mathcal{H}}\right) \le \exp\left(-I_\delta(1/2-\delta)2^{\gamma n}\right) \tag{2.6.28}$$

with rate function I_δ given by (2.6.24). Since

$$\lim_{\delta\to 0} I_\delta(1/2-\delta) = \lim_{\delta\to}$$

$0(1/2+\delta)\log\left[\frac{1/2+\delta}{1-\delta}\right] + (1/2-\delta)\log\left[\frac{1/2-\delta}{\delta}\right] = +\infty$, there exists $\delta_5 > 0$ such that $I_\delta(1/2-\delta) > 1$ for all $\delta \in]0, \delta_5[$. It follows from (2.6.28) that for all $\delta \in]0, \delta_5[$ P_δ-almost surely on the set $\left\{|\mathcal{S}^{I\to}_{w_1,w_3}| \ge 2^{\gamma n}\right\}$

$$P_\delta\left(\sum_{i=1}^{2^{\gamma n}} X_{s_i + c_1 n + j} \ge 2^{\gamma n}/2 \,\middle|\, \bar{\mathcal{H}}\right) \le \exp\left(-2^{\gamma n}\right). \tag{2.6.29}$$

Consequently, $P_\delta\left(\sum_{i=1}^{2^{\gamma n}} X_{s_i+c_1n+j} \geq 2^{\gamma n}/2\right) \leq \exp\left(-2^{\gamma n}\right)$. By Definition 2.6.5, $B^{\mathrm{n},\tau}_{\mathrm{majority}} = B^{n,\tau}_{\mathrm{maj},\to} \cap B^{n,\tau}_{\mathrm{maj},\leftarrow}$ with

$$B^{\mathrm{n},\tau}_{\mathrm{maj},\to} = \bigcap_{w_1,w_3\in\mathcal{C}^{c_1 n}} \bigcap_{I\in\mathcal{I}_L} B^{n,\tau,I\to}_{\mathrm{maj}}(w_1,w_3)$$

and $B^{n,\tau}_{\mathrm{maj},\leftarrow}$ defined analogously. The event $B^{n,\tau,I\to}_{\mathrm{maj}}(w_1,w_3)$ holds if and only if either $|\mathcal{S}^{I\to}_{w_1,w_3}| < 2^{\gamma n}$ or $|\mathcal{S}^{I\to}_{w_1,w_3}| \geq 2^{\gamma n}$ and $\sum_{i=1}^{2^{\gamma n}} X_{s_i^{I\to}+c_1n+j} < 2^{\gamma n}/2$ for all $j \in [1, c_1n-1[$. Thus, if $B^{n,\tau,I\to}_{\mathrm{maj}}(w_1,w_3)$ does not hold, then $|\mathcal{S}^{I\to}_{w_1,w_3}| \geq 2^{\gamma n}$ and there exists $j \in [1, c_1n-1[$ such that $\sum_{i=1}^{2^{\gamma n}} X_{s_i^{I\to}+c_1n+j} \geq 2^{\gamma n}/2$. Hence

$$\left[B^{\mathrm{n},\tau}_{\mathrm{maj},\to}\right]^c \subseteq \bigcup_{w_1,w_3\in\mathcal{C}^{c_1 n}} \bigcup_{I\in\mathcal{I}_L} \bigcup_{j\in[1,c_1n-1[} \left\{|\mathcal{S}^{I\to}_{w_1,w_3}| \geq 2^{\gamma n}, \sum_{i=1}^{2^{\gamma n}} X_{s_i^{I\to}+c_1n+j} \geq \frac{2^{\gamma n}}{2}\right\}.$$

Since there are less than $14\cdot 2^n$ ladder intervals in $\mathcal{I}_L$, it follows that

$$P_\delta\left(\left(B^{n,\tau}_{\mathrm{maj},\to}\right)^c\right) \leq 14\cdot 2^n c_1 n C^{2c_1 n}\exp\left(-2^{\gamma n}\right).$$

We choose $c_{19} > 0$ large enough that $14\cdot 2^n c_1 n C^{2c_1 n}\exp\left(-2^{\gamma n}\right) \leq e^{-n}/2$ for all $n \geq c_{19}$. The claim follows. □

Lemma 2.6.10. *There exist constants $c_{20}, c_{21} > 0$ such that for all $n \geq c_{10}$ (with c_{10} as in (2.6.4))*

$$P\left((B^n_{\text{outside out}})^c\right) \leq c_{21}e^{-c_{20}n}.$$

Proof. We set

$$\mathcal{J} := \left\{\begin{array}{l}(z,i,R) : R \in \left([-2L\cdot 2^{2n}, 2L\cdot 2^{2n}]\setminus[-6\cdot 2^n, 6\cdot 2^n]\right)^{[0,c_1n/2[} \\ \text{admissible piece of path, } z\in[-5\cdot 2^n, 5\cdot 2^n],\ i\in\{\leftarrow,\to\} \\ \}\end{array}\right\}.$$

By Definition 2.6.6,

$$(B^n_{\text{outside out}})^c = \bigcup_{(z,i,R)\in\mathcal{J}} \left\{d(\xi\circ R, w_{z,i,n/2}) < 3\varepsilon n\right\},$$

and consequently,

$$P\left((B^n_{\text{outside out}})^c\right) \leq |\mathcal{J}| \max_{(z,i,R)\in\mathcal{J}} P\left(d(\xi\circ R, w_{z,i,n/2}) < 3\varepsilon n\right). \tag{2.6.30}$$

Let $(z,i,R)\in\mathcal{J}$, and let $n \geq c_{10}$. The piece of scenery $\xi\circ R$ depends only on $\xi|[-2L\cdot 2^{2n}, 2L\cdot 2^{2n}]\setminus[-6\cdot 2^n, 6\cdot 2^n]$, whereas $w_{z,i,n/2}$ depends only on $\xi|[-5\cdot 2^n - c_1nL/2, 5\cdot 2^n + c_1nL/2]$. Since $n \geq c_{10}$, $c_1nL/2 \leq 2^n$ by (2.6.4), and therefore $w_{z,i,n/2}$ depends only on $\xi|[-6\cdot 2^n, 6\cdot 2^n]$. Since the scenery ξ is i.i.d. uniformly colored, $\xi\circ R$ and $w_{z,i,n/2}$ are independent and $P(\xi_j = \xi_{j'}) = 1/C$ for $j \neq j'$. Thus

$$P\left(\xi(R(j)) = w_{z,i,n/2}(j)\ \ \forall j\in J\right) = C^{\lfloor 3\varepsilon n\rfloor - c_1 n/2}$$

for any subset $J \subseteq [0, c_1 n/2[$ with cardinality $|J| = c_1 n/2 - \lfloor 3\varepsilon n \rfloor$. Applying Lemma 2.6.7 with $\varepsilon' = 6\varepsilon/c_1$ and $m = n/2$, we obtain

$$P\left(d(\xi \circ R, w_{z,i,n/2}) < 3\varepsilon n\right) \leq c_{17}(6\varepsilon/c_1)(c_2)^{c_1 n/2} C^{\lfloor 3\varepsilon n \rfloor - c_1 n/2}. \tag{2.6.31}$$

The cardinality of $|\mathcal{J}|$ satisfies

$$|\mathcal{J}| \leq 2(10 \cdot 2^n + 1)4L \cdot 2^{2n} (C-1)^{c_1 n/2} \tag{2.6.32}$$

for the following reason: There are $10 \cdot 2^n + 1$ possible values for z, 2 possible values for i and at most $4L \cdot 2^{2n}$ possible starting points for R. An admissible piece of path has at each step at most $|\mathcal{M}| \leq C - 1$ possible steps; recall that there are strictly more colors than possible steps for the random walk. Hence the number of possible paths R is bounded by $4L \cdot 2^{2n} (C-1)^{c_1 n/2}$.

Clearly, $C^{\lfloor 3\varepsilon n \rfloor} \leq e^{(3\varepsilon n \ln C)}$. We choose $c_{21} > 0$ such that $c_{17}(6\varepsilon/c_1)2(10 \cdot 2^n + 1)4L \cdot 2^{2n} \leq c_{21} \cdot 2^{3n}$. Combining (2.6.30), (2.6.31), and (2.6.32), we obtain

$$P\left((B^n_{\text{outside out}})^c\right) \leq c_{21} e^{n(3\ln 2 + 3\varepsilon \ln C)} \left(\frac{c_2(C-1)}{C}\right)^{c_1 n/2}.$$

Finally, we set $c_{20} := -\left(3\ln 2 + 3\varepsilon \ln C + (c_1/2)\ln\left(\frac{c_2(C-1)}{C}\right)\right)$, and the claim follows because $c_{20} > 0$ by our choice of ε and c_1. □

We will need the following lemma in the proof of Lemma 2.6.12.

Lemma 2.6.11. *There exists c_{22} such that for all $n \geq c_{22}$ and for any admissible piece of path $R \in \mathbb{Z}^{[0,c_1 n[}$ with $R(0) \leq R(c_1 n - 1)$ there exists an admissible piece of path $\bar{R} \in \mathbb{Z}^{[0,c_1 n[}$ such that $\bar{R}(0) = R(0)$, $\bar{R}(c_1 n - 1) = R(c_1 n - 1)$, and the first $c_1 n/3$ steps of $\bar{R}$ are steps of maximal length L to the right.*

Proof. Let $R \in \mathbb{Z}^{[0,c_1 n[}$ be an admissible piece of path. We set $x := R(0)$, $y := R(c_1 n - 1)$; note $x \leq y$.

Suppose R contains at least $c_1 n/3$ steps of maximal length L to the right. Then we define $\bar{R} \in \mathbb{Z}^{[0,c_1 n[}$ to be the admissible piece of path starting at x and ending at y obtained from R by permuting the order of the steps in such a way that all the steps of maximal length L to the right are at the beginning.

If R contains less than $c_1 n/3$ steps of maximal length L to the right, then

$$y - x \leq \left(\frac{c_1 n}{3} - 1\right) L + \frac{2c_1 n}{3}(L-1) \leq c_1 n L - \frac{2c_1 n}{3}. \tag{2.6.33}$$

In this case, let $R_1 \in \mathbb{Z}^{[0,t_1[}$ denote the path which starts at x and goes with maximum steps to the right until it reaches the interval $]y - L, y]$. In other words, $R_1(0) = x$, $R_1(t_1 - 1) \in]y - L, y]$, and for all $s \in [0, t_1 - 1[$ we have that $R_1(s+1) - R_1(s) = L$. Let $y' := R_1(t_1 - 1)$ be the endpoint of R_1. We have $(t_1 - 1)L \leq y - x$ and using (2.6.33), we obtain

$$t_1 \leq \frac{y-x}{L} + 1 \leq c_1 n - \frac{2c_1 n}{3L} + 1. \tag{2.6.34}$$

As we noticed already in the proof of Lemma 2.6.5, the random walk has period 1 or 2. Thus there exists c_{23} such that for all $z \in]y - L, y]$ there exists an admissible piece

of path of length $\leq c_{23}$ starting at z and ending at y. If furthermore the random walk is aperiodic, then c_{23} can be chosen in such a way that for all $z \in]y-L, y]$ there exist admissible pieces of path of even and odd length $\leq c_{23}$ starting at z and ending at y. We choose c_{22} such that $\min\left\{\frac{c_1 n}{3} - 2, \frac{2c_1 n}{3L} - 2\right\} > c_{23}$ for all $n \geq c_{22}$.

Case 1: The random walk is periodic (with period 2). Let $R_3 \in \mathbb{Z}^{[0,t_3[}$ be an admissible piece of path starting at y', ending at y with $t_3 \leq c_{23}$. The concatenation $R_1 R_3$ is an admissible piece of path starting at x, ending at y of length $t_1 + t_3 \leq c_1 n - 1$ by (2.6.34). By assumption, R also starts at x and ends at y. Thus by periodicity we have that $l := |R| - |R_1 R_3| \geq 0$ is even. Let R_2 be the admissible piece of path starting and ending at y' which makes first $l/2$ steps of length L to the right and then $l/2$ steps of length L to the left. We set $\bar{R} := R_1 R_2 R_3$. We have $|R_1 R_2| \geq c_1 n - c_{23} \geq 2 + 2c_1 n/3$. Since all steps of R_1 and half of the steps of R_2 are maximum steps to the right, $\bar{R}$ contains at least $c_1 n/3$ steps of maximal length L at the beginning. By construction, $\bar{R}$ starts at x and ends at y.

Case 2: The random walk is aperiodic. Let $R_3 \in \mathbb{Z}^{[0,t_3[}$ be an admissible piece of path starting at y', ending at y of length $t_3 \leq c_{23}$. We may assume that t_3 is even iff $c_1 n - t_1$ is even. Then $c_1 n - t_1 - t_3$ is even, and we can define R_2 as before. The same argument as above shows that $\bar{R} := R_1 R_2 R_3$ fulfills the claim. □

Lemma 2.6.12. *There exists c_{24} such that for all $n \geq c_{24}$*

$$P\left(\left(B^n_{\text{recogn straight}}\right)^c\right) \leq c_{18} e^{-n};$$

c_{18} is specified in Lemma 2.6.8.

Proof. Let $c_{24} := \max\{c_{10}, c_{22}\}$ with c_{22} as in Lemma 2.6.11, and let $n \geq c_{24}$. We will show that the following inclusion holds:

$$B^n_{\text{ladder diff}} \subseteq B^n_{\text{recogn straight}}. \tag{2.6.35}$$

The claim follows then from Lemma 2.6.8.

Suppose the event $B^n_{\text{ladder diff}}$ holds. Let $R_1 \in [-7 \cdot 2^n, 7 \cdot 2^n]^{[0,c_1 n[}$ be an admissible piece of path which is not a ladder path. We set $x := R_1(0)$ and $y := R_1(c_1 n - 1)$. We have to show that there exists an admissible piece of path $R_2 \in [-8 \cdot 2^n, 8 \cdot 2^n]^{[0,c_1 n[}$ with starting point x, endpoint y, and $d(\xi \circ R_1, \xi \circ R_2) \geq 5\varepsilon n$. We assume that $x \leq y$. The case $x > y$ is reduced to this case by considering the reversed path $k \mapsto R_1(c_1 n - 1 - k)$. By Lemma 2.6.11 applied to R_1, there exists an admissible piece of path $R_3 \in \mathbb{Z}^{[0,c_1 n[}$ such that $R_3(0) = x$, $R_3(c_1 n - 1) = y$ and the first $c_1 n/3$ steps of R_3 are steps of maximal length L to the right. Since $y - x \neq (c_1 n - 1)L$, at least one step of R_3 is not a step of maximum length to the right. We construct an admissible piece of path R_4 by permuting the steps of R_3. We set $R_4(0) := x$. The first step of R_4 is the first step of R_3 which is not a step of maximum length to the right. Formally we set $j := \min\{i \in [1, c_1 n[\colon R_3(i) - R_3(i-1) \neq L\}$, and define

$$R_4(i) := \begin{cases} R_3(i), & \text{if } i \in [0, c_1 n[\setminus [1, j] \\ R_3(i-1) + R_3(j) - R_3(j-1), & \text{if } i \in [1, j]. \end{cases}$$

Clearly, R_4 is an admissible piece of path of length $c_1 n$ with $R_4(0) = x$ and $R_4(c_1 n - 1) = y$. Using that R_4 jumps in each step at most a distance of L, we obtain that $|R_4(i)| \leq$

$|R_4(0)| + c_1 nL = x + c_1 nL \leq 8 \cdot 2^n$ for all $i \in [0, c_1 n[$ because $c_1 nL \leq 2^n$ for $n \geq c_{10}$. The same is true for R_3.

Since R_3 starts with $c_1 n/3$ steps of maximum length L to the right, we have that $\xi \circ R_3 | [1, c_1 n/3] = w_{x+L,\to,n/3}$, and by definition of R_4, we have $\xi \circ R_4 | [1, c_1 n/3] = w_{x',\to,n/3}$ with $x' = x + R_3(j) - R_3(j-1)$. By construction, $R_3(j) - R_3(j-1) \neq L$ so that $x + L \neq x'$. Since R_3 and R_4 take only values in $[-8 \cdot 2^n, 8 \cdot 2^n]$, we have that $x + L, x' \in [-8 \cdot 2^n, 8 \cdot 2^n]$. Using that $B^n_{\text{ladder diff}}$ holds, yields $d(w_{x+L,\to,n/3}, w_{x',\to,n/3}) \geq 10\varepsilon n$, and by the triangle inequality, we get that $\xi \circ R_1$ cannot have a distance smaller than $5\varepsilon n$ to both $\xi \circ R_3$ and $\xi \circ R_4$. Hence there exists $i \in \{3, 4\}$ such that $d(\xi \circ R_1, \xi \circ R_i) \geq 5\varepsilon n$. Let $R_2 := R_i$ in the definition of $B^n_{\text{recogn straight}}$. □

Lemma 2.6.13. *There exist constants $c_{25}, c_{26} > 0$ such that for all $n \in \mathbb{N}$*

$$P\left(\left(B^n_{\text{signals}}\right)^c\right) \leq c_{25} e^{-c_{26} n}.$$

Proof. We show that there exist $c_{25}, c_{26} > 0$ such that for all n

$$P\left(\left(B^n_{\text{sign,r},\to}\right)^c\right) \leq \frac{c_{25}}{4} e^{-c_{26} n}. \tag{2.6.36}$$

Analogously, one proves statements for $B^n_{\text{sign,l},\to}$, $B^n_{\text{sign,l},\leftarrow}$, and $B^n_{\text{sign,r},\leftarrow}$. The claim follows from these four inequalities and the definition of B^n_{signals}. We set

$$\mathcal{R} := \left\{ \begin{matrix} (z, R) & : & z & \in & [-6 \cdot 2^n, 6 \cdot 2^n], & R & \in \\ [-2L \cdot 2^{2n}, 2L \cdot 2^{2n}]^{[0, c_1 n[} & & \text{admissible} & & \text{piece} & \text{of} & \text{path} \\ \text{with } R(0) < z & & & & & & \end{matrix} \right\}.$$

By Definition 2.6.8,

$$\left(B^n_{\text{sign,r},\to}\right)^c = \bigcup_{(z,R) \in \mathcal{R}} \{d(\xi \circ R, w_{z,\to,n}) < 5\varepsilon n\}. \tag{2.6.37}$$

Let $(z, R) \in \mathcal{R}$. By Definition 2.6.1, $w_{z,\to,n}(k) = \xi(z + kL)$. Note that $R(k) < z + kL$ for all $k \in [0, c_1 n[$: For $k = 0$ this is true by assumption. Suppose $R(k) < z + kL$ holds for some $k \in [0, c_1 n - 1[$. Since the maximal jump length of R is L, we obtain $R(k+1) \leq R(k) + L < z + (k+1)L$, and the claim follows by induction.

We prove by induction over the cardinality of J, that

$$P((\xi \circ R)|J = w_{z,\to,n}|J) = C^{-|J|} \tag{2.6.38}$$

for any $J \subseteq [0, c_1 n[$: For $J = \{j\}$ we use that $\xi(R(j))$ and $w_{z,\to,n}(j) = \xi(z + jL)$ are independent because $R(j) < z + jL$. Suppose (2.6.38) holds for any $J \subseteq [0, c_1 n[$ with $|J| = k$ for some $k \in [1, c_1 n - 1[$. Let $J' \subseteq [0, c_1 n[$ with $|J'| = k + 1$, and let $j := \max J'$. Then $\xi(z + jL)$ is independent of $\xi(z + j'L)$, $j' \in J' \setminus \{j\}$, and of $\xi(R(j'))$, $j' \in J'$, because $R(j') < z + j'L \leq z + jL$. Hence

$$\begin{aligned} P((\xi \circ R)|J' = w_{z,\to,n}|J') &= C^{-1} P((\xi \circ R)|J' \setminus \{j\} = w_{z,\to,n}|J' \setminus \{j\}) \\ &= C^{-(1+|J' \setminus \{j\}|)} = C^{|J'|}; \end{aligned}$$

for the second but last equality with used the induction hypothesis. We use Lemma 2.6.7 with $\varepsilon' := 5\varepsilon$ and $m := n$ to obtain

$$P\left(d(\xi \circ R, w_{z,\to,n}) < 5\varepsilon n\right) \leq c_{17}(5\varepsilon/c_1)(c_2)^{c_1 n} C^{\lfloor 5\varepsilon n \rfloor - c_1 n}. \tag{2.6.39}$$

It is easy to see that the cardinality of $\mathcal{R}$ is bounded by $(12\cdot 2^n+1)(4L\cdot 2^{2n}+1)(C-1)^{c_1 n}$ Combining this with (2.6.37) and (2.6.39), we obtain

$$P\left(\left(B^n_{\mathrm{sign,r},\to}\right)^c\right) \le c_{17}(5\varepsilon/c_1)(12\cdot 2^n+1)(4L\cdot 2^{2n}+1)C^{\lfloor 5\varepsilon n\rfloor}\left(\frac{c_2(C-1)}{C}\right)^{c_1 n}.$$

We choose c_{25} such that $c_{17}(5\varepsilon/c_1)(12\cdot 2^n+1)(4L\cdot 2^{2n}+1)\le c_{25}2^{3n}/4$ for all $n\in\mathbb{N}$. Then

$$P\left(\left(B^n_{\mathrm{sign,r},\to}\right)^c\right) \le \frac{c_{25}}{4}e^{n[3\ln 2+5\varepsilon\ln C]}\left(\frac{c_2(C-1)}{C}\right)^{c_1 n}.$$

We set $c_{26} := -\left(3\ln 2+5\varepsilon\ln C+c_1\ln\left(\frac{c_2(C-1)}{C}\right)\right)$. Since $c_{26}>0$ by our choice of ε and c_1, the claim follows. □

Lemma 2.6.14. *There exists a constant $c_{27}>0$ such that for all $n\ge c_{27}$*

$$P\left(E^{n,\tau}_{\mathrm{stop}}\setminus B^{n,\tau}_{\mathrm{straight\ often}}\right)\le e^{-n}.$$

Proof. Recall Definition 2.6.9. We will show for all n sufficiently large,

$$P\left(E^{n,\tau}_{\mathrm{stop}}\setminus\left(\bigcap_{I\in\mathcal{J}_L}\{|\mathcal{S}_\to(I)|\ge 2^{\gamma n}\}\right)\right)\le e^{-n}/2. \tag{2.6.40}$$

A similar consideration shows that the same estimate is true if we replace $\mathcal{S}_\to(I)$ by $\mathcal{S}_\leftarrow(I)$, and the claim then follows from the definition of $B^{n,\tau}_{\mathrm{straight\ often}}$. Since the proof is very similar to the proof of Lemma 2.6.5, we will omitt some of the details.

Let $I\in\mathcal{J}_L$. We denote by R^I the ladderpath in $\mathbb{Z}^{[0,3c_1n[}$ which traverses I from left to right. For $t\in\mathbb{N}_0$ we define the event $E(t,I):=$

$$\left\{S(t+i)=R^I(i)\ \forall i\in[0,3c_1n[\ \text{or}\ S(t+1+i)=R^I(i)\ \forall i\in[0,3c_1n[\right\}.$$

Let $n\ge c_{10}$ with c_{10} as in (2.6.4), and let $k\in[1,2^{\alpha n}]$. We set $t_{k,n}:=\tau_k+2^{2n-1}$ and we define random variables $Y_k(I)$ as follows: If $|S(\tau_k)|\le 2^n$ and $E(t_{k,n},I)$ does not hold, then we set $Y_k(I)=0$. Otherwise we set $Y_k(I)=1$. By Definition 2.6.9, we have

$$\begin{aligned}E^{n,\tau}_{\mathrm{stop}}\setminus\left(\bigcap_{I\in\mathcal{J}_L}\{|\mathcal{S}_\to(I)|\ge 2^{\gamma n}\}\right)&\subseteq\bigcup_{I\in\mathcal{J}_L}E^{n,\tau}_{\mathrm{stop}}\cap\left\{\sum_{k=1}^{2^{\alpha n}}Y_k(I)<2^{\gamma n}\right\}\\&\subseteq\bigcup_{I\in\mathcal{J}_L}\bigcup_{j=1}^{2^{\gamma n}}E^{n,\tau}_{\mathrm{stop}}\cap\left\{\sum_{k=(j-1)2^{(\alpha-\gamma)n}+1}^{j\cdot 2^{(\alpha-\gamma)n}}Y_k(I)=0\right\}.\end{aligned} \tag{2.6.41}$$

Using the strong Markov property and induction (see the proof of Lemma 2.6.5, in particular (2.6.19), for a similar argument) we obtain for $n\ge c_{10}$ and $m,M\in[1,2^{\alpha n}]$ with $m\le M$

$$P\left(E^{n,\tau}_{\mathrm{stop}}\cap\left\{\sum_{k=m}^{M}Y_k(I)=0\right\}\right)\le\left[\max_{x\in[-6\cdot 2^n,6\cdot 2^n]}P_x(E(2^{2n-1},I)^c)\right]^{M-m+1}. \tag{2.6.42}$$

By the local central limit theorem, there exist constants $c_{27}, c_{28} > 0$ such that for all $n \geq c_{27}$

$$\min_{x,z\in[-6\cdot 2^n, 6\cdot 2^n]} P_x\left(S(2^{2n-1}) = z \text{ or } S(2^{2n-1}+1) = z\right) \geq c_{28} 2^{-n}. \tag{2.6.43}$$

The probability that the random walk starting at x makes $3c_1 n - 1$ consecutive steps of maximum length to the right equals $\mu(L)^{3c_1 n - 1}$. Since all intervals in $\mathcal{J}_L$ are contained in $[-6 \cdot 2^n, 6 \cdot 2^n]$, we obtain

$$\min_{x\in[-2^n,2^n]} \min_{I\in\mathcal{J}_L} P_x(E(2^{2n-1}, I)) \geq c_{28} 2^{-n} \mu(L)^{3c_1 n - 1} = c_{29} 2^{-n} \mu(L)^{3c_1 n}$$

with $c_{29} := c_{28}\mu(L)^{-1}$. Combining the last inequality with (2.6.42), we obtain

$$P\left(E^{n,\tau}_{\text{stop}} \cap \left\{\sum_{k=m}^{M} Y_k(I) = 0\right\}\right) \leq \left(1 - c_{29} 2^{-n} \mu(L)^{3c_1 n}\right)^{M-m+1}. \tag{2.6.44}$$

From (2.6.41) and (2.6.44) it follows that

$$\begin{aligned} P\left[E^{n,\tau}_{\text{stop}} \setminus \left[\bigcap_{I\in\mathcal{J}_L} \{|\mathcal{S}_\rightarrow(I)| \geq 2^{\gamma n}\}\right]\right] &\leq 2^{4+[1+\gamma]n} \left[1 - c_{29} 2^{-n} \mu(L)^{3c_1 n}\right]^{2^{[\alpha-\gamma]n}} \\ &\leq 2^{4+[1+\gamma]n} \exp\left[2^{(\alpha-\gamma)n} \ln\left[1 - c_{29} 2^{-n} \mu(L)^{3c_1 n}\right]\right] \\ &\leq 2^{4+[1+\gamma]n} \exp\left[-c_{29} 2^{[\alpha-1-\gamma]n} \mu(L)^{3c_1 n}\right] \leq 2^{4+[1+\gamma]n} \exp\left[-c_{29} e^{c_{30} n}\right] \leq e^{-n}/2 \end{aligned}$$

for all n sufficiently large because $c_{30} = (\alpha - 1 - \gamma)\ln 2 + 3c_1 \ln \mu(L) > 0$ by our choice of α. □

2.6.4 Alg^n reconstructs with high probability

Proof of Theorem 2.3.5. Suppose $\xi|[-2^n, 2^n] \preceq \text{Alg}^n(\tau, \tilde{\chi}|\,[0, 2 \cdot 2^{12\alpha n}[\,, \psi) \preceq \xi|[-4 \cdot 2^n, 4 \cdot 2^n]$. Assume $\psi \in \mathcal{C}^{[-kn,kn]}$ with $k \geq c_1 L$, $\psi \preceq \xi|\,[-2^n, 2^n]$, and assume $\xi|\,[-2^n, 2^n] \neq (1)_{[-2^n,2^n]}$. Then $\text{Alg}^n(\tau, \tilde{\chi}|\,[0, 2 \cdot 2^{12\alpha n}[\,, \psi)|[-kn, kn] = \psi$ by the definition of Alg^n (Definition 8.5.3) and the definition of $\mathsf{SolutionPiece}^n$ (Definition 9.7.8).

In order to show that Alg^n reconstructs with high probability, we combine Lemmas 2.6.4, 2.6.3, 2.6.2, and 2.6.1 to obtain

$$\begin{aligned} E^{n,\tau}_{\text{stop}} \setminus E^{n,\tau}_{\text{reconstruct}} \quad\subseteq\quad & \left(E^{n,\tau}_{\text{stop}} \setminus B^{n,\tau}_{\text{all paths}}\right) \cup \left(B^{n}_{\text{few mistakes}}\right)^c \cup \left(B^{n}_{\text{ladder diff}}\right)^c \\ & \cup \left(B^{n,\tau}_{\text{majority}}\right)^c \cup \left(B^{n}_{\text{outside out}}\right)^c \cup \left(B^{n}_{\text{signals}}\right)^c \\ & \cup \left(B^{n}_{\text{recogn straight}}\right)^c \cup \left(E^{n,\tau}_{\text{stop}} \setminus B^{n,\tau}_{\text{straight often}}\right). \end{aligned}$$

The claim follows from Lemmas 2.6.5, 2.6.6, 2.6.8, 2.6.9, 2.6.10, 2.6.12, 2.6.13, and 2.6.14. □

Acknowledgement 2.6.1. *This paper was written while the authors were working at Eurandom. They thank Eurandom for its hospitality.*

References

[1] E. Artin. *The Gamma Function.* Holt, Rinehart and Winston, 1964.

[2] I. Benjamini and H. Kesten. Distinguishing sceneries by observing the scenery along a random walk path. *J. Anal. Math.*, 69:97–135, 1996.

[3] F. den Hollander. *Large deviations.* American Mathematical Society, Providence, RI, 2000.

[4] F. den Hollander and J. E. Steif. Mixing properties of the generalized T, T^{-1}-process. *J. Anal. Math.*, 72:165–202, 1997.

[5] W. Th. F. den Hollander. Mixing properties for random walk in random scenery. *Ann. Probab.*, 16(4):1788–1802, 1988.

[6] R. Durrett. *Probability: Theory and Examples.* Duxbury Press, Second edition, 1996.

[7] M. Harris and M. Keane. Random coin tossing. *Probab. Theory Related Fields*, 109(1):27–37, 1997.

[8] D. Heicklen, C. Hoffman, and D. J. Rudolph. Entropy and dyadic equivalence of random walks on a random scenery. *Adv. Math.*, 156(2):157–179, 2000.

[9] C. D. Howard. Detecting defects in periodic scenery by random walks on $\mathbb{Z}$. *Random Structures Algorithms*, 8(1):59–74, 1996.

[10] C. D. Howard. Orthogonality of measures induced by random walks with scenery. *Combin. Probab. Comput.*, 5(3):247–256, 1996.

[11] C. D. Howard. Distinguishing certain random sceneries on $\mathbb{Z}$ via random walks. *Statist. Probab. Lett.*, 34(2):123–132, 1997.

[12] S. A. Kalikow. T, T^{-1} transformation is not loosely Bernoulli. *Ann. of Math. (2)*, 115(2):393–409, 1982.

[13] M. Keane and W. Th. F. den Hollander. Ergodic properties of color records. *Phys. A*, 138(1-2):183–193, 1986.

[14] H. Kesten. Detecting a single defect in a scenery by observing the scenery along a random walk path. In *Itô's stochastic calculus and probability theory*, pages 171–183. Springer, Tokyo, 1996.

[15] H. Kesten. Distinguishing and reconstructing sceneries from observations along random walk paths. In *Microsurveys in discrete probability (Princeton, NJ, 1997)*, pages 75–83. Amer. Math. Soc., Providence, RI, 1998.

[16] A. Lenstra and H. Matzinger. Reconstructing a 4-color scenery by observing it along a recurrent random walk path with unbounded jumps. Eurandom, 2001. In preparation.

[17] D. Levin and Y. Peres. Random walks in stochastic scenery on $\mathbb{Z}$. Preprint, 2002.

[18] D. A. Levin, R. Pemantle, and Y. Peres. A phase transition in random coin tossing. *Ann. Probab.*, 29(4):1637–1669, 2001.

[19] E. Lindenstrauss. Indistinguishable sceneries. *Random Structures Algorithms*, 14(1):71–86, 1999.

[20] M. Löwe and H. Matzinger. Reconstruction of sceneries with correlated colors. Eurandom Report 99-032, accepted by Stochastic Processes and Their Applications, 1999.

[21] M. Löwe and H. Matzinger. Scenery reconstruction in two dimensions with many colors. *Ann. Appl. Probab.*, 12(4):1322–1347, 2002.

[22] M. Löwe, H. Matzinger, and F. Merkl. Reconstructing a multicolor random scenery seen along a random walk path with bounded jumps. Eurandom Report 2001-030, 2001.

[23] H. Matzinger. *Reconstructing a 2-color scenery by observing it along a simple random walk path with holding.* PhD thesis, Cornell University, 1999.

[24] H. Matzinger. Reconstructing a three-color scenery by observing it along a simple random walk path. *Random Structures Algorithms*, 15(2):196–207, 1999.

[25] H. Matzinger. Reconstructing a 2-color scenery by observing it along a simple random walk path. Eurandom Report 2000-003, 2000.

[26] H. Matzinger and S. W. W. Rolles. Finding blocks and other patterns in a random coloring of $\mathbb{Z}$. Preprint.

[27] H. Matzinger and S. W. W. Rolles. Reconstructing a 2-color scenery in polynomial time by observing it along a simple random walk path with holding. Preprint.

[28] H. Matzinger and S. W. W. Rolles. Reconstructing a random scenery in polynomial time. Preprint.

Chapter 3

Reconstructing a random scenery seen along a simple random walk path

accepted in *Ann. Appl. Probab.*, 2004.

By Heinrich Matzinger

Let $\{\xi(n)\}_{n\in\mathbb{Z}}$ be a 2-color random scenery, that is a random coloration of $\mathbb{Z}$ in two colors, such that the $\xi(i)$'s are i.i.d. Bernoulli variables with parameter $\frac{1}{2}$. Let $\{S(n)\}_{n\in\mathbb{N}}$ be a symmetric random walk starting at 0. Our main result shows that a.s., $\xi \circ S$ (the composition of ξ and S) determines ξ up to translation and reflection. In other words, by observing the scenery ξ along the random walk path S, we can a.s. reconstruct ξ up to translation and reflection. This result gives a positive answer to the question of H. Kesten of whether one can a.s. detect a single defect in almost every 2-color random scenery by observing it only along a random walk path. [1]

3.1 Introduction

A scenery is defined to be a function from $\mathbb{Z}$ to $\{0,1\}$. Let ξ and $\tilde{\xi}$ be two sceneries. We say that ξ and $\tilde{\xi}$ are equivalent iff there exist $a \in \mathbb{Z}$ and $b \in \{-1,1\}$ such that for all $x \in \mathbb{Z}$ we have $\xi(x) = \tilde{\xi}(a+bx)$. In this case we write $\xi \approx \tilde{\xi}$. In other words, two sceneries are equivalent iff they can be obtained from each other by a shift or a reflection. In everything that follows $\{S(k)\}_{k\geq 0}$ will be a simple random walk on $\mathbb{Z}$ starting at the origin. We will denote by $\chi \in \{0,1\}^{\mathbb{N}}$ the color record obtained by observing the scenery ξ along the path of the random walk $\{S(k)\}_{k\geq 0}$:

$$\chi := (\xi(S(0)), \xi(S(1)), \xi(S(2)), \ldots),$$

i.e. $\chi(k) := \xi(S(k))$ for all $k \in \mathbb{N}$. We examine the follwoing question: given an unknown scenery ξ, can we "reconstruct" ξ if we can only observe χ ? Thus, does one path realization of the process $\{\chi(k)\}_{k\geq 0}$ uniquely determine ξ? The answer in those general terms is "no". However, under appropriate restrictions, the answer will become "yes".

[1] *MSC 2000 subject classification:* Primary 60K37, Secondary 60G10, 60J75.
Key words: Scenery reconstruction, jumps, stationary processes, random walk, ergodic theory.

This the main result of this paper. Let us explain these restrictions: First, if ξ and $\tilde{\xi}$ are equivalent, we can in general not distinguish whether the observations come from ξ or from $\tilde{\xi}$. Thus, we can only reconstruct ξ up to equivalence modulo $\approx$. Second, it is clear that the reconstruction will in the best case work only almost surely. If the random walk $\{S(k)\}_{k\geq 0}$ decides to walk only to the left (which it could do with probability zero), then we obtain no information about the right side of the scenery ξ and thus are not able to reconstruct the scenery ξ. Eventually, Lindenstrauss in [12] exhibits sceneries which one can not reconstruct. Thus, not all sceneries can be reconstructed. However, we prove that a lot of "typical" sceneries can be reconstructed up to equivalence and almost surely. For this we take the scenery ξ to be the outcome of a random process which is independent of $\{S(k)\}_{k\geq 0}$ such that the $\xi(k)$'s are i.i.d. Bernoulli with parameter $\frac{1}{2}$. We use the following notation: we write ξ for the (random) scenery: $\xi : k \mapsto \xi(k), \mathbb{Z} \longrightarrow \{0,1\}$. Our main result states that, given only the observation χ, almost every scenery ξ can be reconstructed a.s. up to equivalence. Let us state our main theorem:

Theorem 3.1.1. *Let $\{S(k)\}_{k\geq 0}$ and $\{\xi(k)\}_{k\in\mathbb{Z}}$ be two processes independent of each other such that $\{S(k)\}_{k\geq 0}$ is a simple random walk starting at the origin and such that the $\xi(k)$'s are i.i.d. Bernoulli variables with parameter 1/2. Then a.s. χ determines ξ up to equivalence. In other words, there exists a measurable function $\mathcal{A} : \{0,1\}^{\mathbb{N}} \longrightarrow \{0,1\}^{\mathbb{Z}}$ such that $P(\mathcal{A}(\chi) \approx \xi) = 1$. ("Measurable" means measurable with respect to the σ-algebras induced by the canonical coordinates on $\{0,1\}^{\mathbb{N}}$ and on $\{0,1\}^{\mathbb{Z}}$).*

We will prove the above theorem by explicitly describing how to reconstruct ξ from χ. Hence, our approach is constructive. We explicitly give a construction which produces a (random) scenery $\bar{\xi} : \mathbb{Z} \longrightarrow \{0,1\}$ when applied to the observations χ. The constructed scenery $\bar{\xi}$ is shown to be a.s. equivalent to ξ. In this way $\mathcal{A}$ gets defined: $\mathcal{A}(\chi) := \bar{\xi}$.

Let us now make a few historical comments until the end of this section This paper was motivated by Kesten's question to me of whether one can a.s. distinguish a single defect in almost any two color scenery. Let us explain what the scenery distinguishing problem is . Let $\xi, \eta : \mathbb{Z} \longrightarrow \{0,1\}$ and let $\{S(k)\}_{k\in\mathbb{N}}$ be a symmetric random walk on $\mathbb{Z}$. Let the process $\{\chi(k)\}_{k\in\mathbb{N}}$ be equal to either $\{\xi(S(k))\}_{k\in\mathbb{N}}$ or $\{\eta(S(k))\}_{k\in\mathbb{N}}$. Is it possible by observing only one path realization of $\{\chi(k)\}_{k\in\mathbb{N}}$ to say to which one of the two $\{\xi(S(k))\}_{k\in\mathbb{N}}$ or $\{\eta(S(k))\}_{k\in\mathbb{N}}$, $\{\chi(k)\}_{k\in\mathbb{N}}$ is equal to? (We assume that we know ξ and η.) If yes, we say that it is possible to distinguish between the sceneries ξ and η by observing them along a path of $\{S(k)\}_{k\in\mathbb{N}}$. Otherwise, when it is not possible to figure out almost surely by observing $\{\chi(k)\}_{k\in\mathbb{N}}$ alone whether $\{\chi(k)\}_{k\in\mathbb{N}}$ is generated on ξ or on η, we say that ξ and η are indistinguishable. The problem of distinguishing two sceneries was raised independently by I. Benjamini and by den Hollander and Keane. The motivation came from problems in ergodic theory, such as the T, T^{-1}problem (see Kalikov [7]) and from the study of various aspects of $\{\xi(S(k))\}_{n\in\mathbb{N}}$, where $\{\xi(k)\}_{k\in\mathbb{Z}}$ is random. (See Kesten and Spitzer in [9], Keane and den Hollander in [8], den Hollander in [3]). Benjamini and Kesten showed in [1] that one can distinguish almost any two random sceneries even when the random walk is in $\mathbb{Z}^2$. (They assumed the sceneries to be random themselves, so that the $\xi(k)$'s and the $\eta(n)$'s are i.i.d. Bernoulli.) Kesten in [10] proved that when the random sceneries are i.i.d. and have four colors, i.e., ξ and $\eta : \mathbb{Z} \longrightarrow \{0,1,2,3\}$, and differ only in one point, they can be a.s. distinguished. He asked whether this result might still hold with fewer colors. The main result of this paper directly implies that one can distinguish single defects in almost any scenery. In [14],

we proved for the three color case that one can a.s. reconstruct almost every three color scenery. We also established that this implies, that one can distinguish single defects for almost all three color sceneries. In the two color case, i.e. in the case we consider in this paper, the same thing is true. This means that our result for scenery reconstruction implies that one can distinguish single defects in almost all sceneries. We state the following corollary to our main result without giving a proof. (The proof that our main result implies the following corollary is very similar to the one given in [14] for the three color case.)

Corollary 3.1.1. *Let $\mathbb{B}$ designate the set of all two color sceneries. $\mathbb{B} = \{\xi : \mathbb{Z} \longrightarrow \{0,1\}\} = \{0,1\}^{\mathbb{Z}}$. Let $(\mathbb{B}, \sigma(\mathbb{B}))$ denote the measurable space, where $\sigma(\mathbb{B})$ is the σ-algebra induced by the canonical coordinates on $\mathbb{B}$. Let P denote the probability measure on $(\mathbb{B}, \sigma(\mathbb{B}))$ obtained by assuming that the $\xi(i)$'s are i.i.d. Bernoulli variables with parameter $\frac{1}{2}$. Then there exist a $\sigma(\mathbb{B})$-measurable set $\mathbb{S}$, such that $P(\mathbb{S}) = 1$ and such that for ever scenery $\xi \in \mathbb{S}$ and every scenery η which is equal to ξ everywhere except in one point, we have that ξ and η are distinguishable.*

The above corollary says that there are many sceneries which one can distinguish or, in other words, that sceneries which are typical in a certain sense can be distinguished. However the above result becomes false if one tries to extend it to all pairs of sceneries which are not equivalent. Recently, Lindenstrauss [12] exhibited a non denumerable set of pairs of non-equivalent sceneries on $\mathbb{Z}$ which he proved to be indistinguishable. Before that, Howard proved in [4], [5] and [6] that any two periodical sceneries of $\mathbb{Z}$ which are not equivalent modulo translation and reflection are distinguishable and that one can a.s. distinguish single defects in periodical sceneries. Kesten asked in [11] whether this result would still hold when the random walk would be allowed to jump. He also asked what would happen in the two dimensional case. Löwe and Matzinger in [13] have been able to prove that one can a.s. reconstruct almost every scenery up to equivalence in two dimensions, provided the scenery has a lot of colors. However the problem of the reconstruction of two color sceneries in $\mathbb{Z}$ seen along the random walk path of a recurrent random walk which is allowed to jump remains open. In our opinion, this is a central open problem at present. Eventually we should also mention that the two color scenery reconstruction problem for a scenery which is i.i.d. is equivalent to the following problem: let $\{R(k)\}_{k\in\mathbb{Z}}$ and $\{S(k)\}_{k\geq 0}$ be two independent simple random walks on $\mathbb{Z}$ both starting at the origin and living on the same probability space. (Here we mean that $\{R(k)\}_{k\geq 0}$ and $\{R(-k)\}_{k\geq 0}$are two independent simple random walks both starting at the origin.) Does one path realization of the iterated random walk $\{R(S(k))\}_{k\geq 0}$ uniquely determines the path of $\{R(k)\}_{k\in\mathbb{Z}}$ up to shift and reflection around the origin? If one takes the representation of the scenery ξ as a nearest neighbor walk (which we will define later) for $\{R(k)\}_{k\in\mathbb{Z}}$ then it becomes immediately clear that the two problems are equivalent. We leave it to the reader to check the details. So the main result of this paper is equivalent to the following result for iterated nearest neighbor walks: one path realization of the iterated random walk $\{R(S(k))\}_{k\geq 0}$ a.s. uniquely determines the path of $\{R(k)\}_{k\in\mathbb{Z}}$ up to shift and reflection around the origin. This is a discrete analogous of the result of Burdzy [2] concerning the path of iterated Brownian motion.

3.2 Reconstructing a finite piece of the scenery ξ

To explain a key idea, we first present a solution to a simplified but somewhat unrealistic case:

3.2.1 Simplified example

Assume for a moment that the scenery ξ is non-random, and instead of being a two color scenery, would be a four color scenery, i.e. $\xi : \mathbb{Z} \longrightarrow \{0,1,2,3\}$. Let us imagine furthermore, that there are two integers x, y such that $\xi(x) = 2$ and $\xi(y) = 3$, but outside x and y the scenery has everywhere color 0 or 1, (i.e. for all $z \in \mathbb{Z}$ with $z \neq x, y$ we have that $\xi(z) \in \{0,1\}$.) The simple random walk $\{S(k)\}_{k\geq 0}$ can go with each step one unit to the right or one unit to the left. This implies that the shortest possible time for the random walk $\{S(k)\}_{k\geq 0}$ to go from the point x to the point y is $|x-y|$. When the random walk $\{S(k)\}_{k\geq 0}$ goes in shortest possible time from x to y it goes in a straight way, which means that between the time it is at x and until it reaches y it only moves in one direction. During that time, the random walk $\{S(k)\}_{k\geq 0}$ reveals the portion of ξ lying between x and y. If between time t_1 and t_2 the random walk goes in a straight way from x to y, (that is if $|t_1 - t_2| = |x-y|$ and $S(t_1) = x, S(t_2) = y$), then the word $\chi(t_1), \chi(t_1+1), \ldots, \chi(t_2)$ is a copy of the scenery ξ restricted to the interval $[\min\{x,y\}, \max\{x,y\}]$. In this case, the word $\chi(t_1), \chi(t_1+1), \ldots, \chi(t_2)$ is equal to the word $\xi(x), \xi(x+u), \xi(x+2u), \ldots, \xi(y)$, where $u := (y-x)/|y-x|$. Since the random walk $\{S(k)\}_{k\geq 0}$ is recurrent it a.s. goes at least once, in the shortest possible way from the point x to the point y. Because we are given infinitely many observations we can a.s. figure out what the distance between x and y is: the distance between x and y is the shortest time laps that a "3" will ever appear in the observations χ after a "2". When, on the other hand, a "3" appears in the observations χ in shortest possible time after a "2", then between the time we see that "2" and until we see the next "3", we observe a copy of $\xi(x), \xi(x+u), \xi(x+2u), \ldots, \xi(y)$ in the observations χ. This fact allows us to reconstruct the finite piece $\xi(x), \xi(x+u), \xi(x+2u), \ldots, \xi(y)$ of the scenery. Choose any couple of integers t_1, t_2 with $t_2 > t_1$, minimizing $|t_2 - t_1|$ under the condition that $\chi(t_1) = 2$ and $\chi(t_2) = 3$. A.s. then $\chi(t_1), \chi(t_1+1), \ldots, \chi(t_2)$ is equal to $\xi(x), \xi(x+u), \xi(x+2u), \ldots, \xi(y)$.

A numerical example: Let the scenery ξ be such that: $\xi(-2) = 0$, $\xi(-1) = 2$, $\xi(0) = 0$, $\xi(1) = 1$, $\xi(2) = 1$, $\xi(3) = 3$, $\xi(4) = 0$. Assume furthermore that the scenery ξ has a 2 and a 3 nowhere else then in the points -1 and 3. Imagine that χ the observation given to us would start as follows:

$$\chi = (0,2,0,1,0,1,3,0,3,1,1,1,1,0,2,0,1,1,3,\ldots)$$

By looking at all of χ we would see that the shortest time a 3 occurs after a 2 in the observations is 4. In the first observations given above there is however already a 3 only four time units after a 2. The binary word appearing in that place , between the 2 and the 3 is 011. We deduce from this that between the place of the 2 and the 3 the scenery must look like: 011.

In reality the scenery we want to reconstruct has 2 colors only. So, instead of the 2 and the 3 in the example above we will use a special pattern in the observations which will tell us when the random walk is back at the same spot. One possibility (although not yet the one we will eventually use) would be to use binary words of the form: 001100 and 110011. It is easy to verify that the only possibility for the word 001100, resp. 110011

to appear in the observations, is when the same word 001100, resp. 110011 occurs in the scenery and the random walk reads it. So, imagine (to give another pedagogical example of a simplified case) the scenery would be such that in a place x there occurs the word 001100, and in the place y there occurs the word 110011 , but these two words occur in no other place in the scenery. These words can then be used as markers: In order to reconstruct the piece of the scenery ξ comprised between x and y we could proceed as follows: take in the observations the place where the word 110011 occurs in shortest time after the word 001100. In that place in the observations we see a copy of the piece of the scenery ξ comprised between x and y. The reason why the very last simplified example is not realistic is the following: we take the scenery to be the outcome of a random process itself where the $\xi(k)$'s are i.i.d. variables themselves. Thus any word will occur infinitely often in the scenery ξ. However, if for example the markers in the scenery occur far away from each other, then we can still use the above described reconstruction strategy: The random walk will then be very likely to first cross from x to y in a straight way before meeting another marker and creating some confusion. In the next subsection we explain how to construct the markers which we are eventually going to use.

3.2.2 Representation of the scenery ξ as a nearest neighbor walk

The scenery reconstruction problem contains two main ingredients: A random walk $\{S(k)\}_{k\in\mathbb{N}}$ and a "random environment", that is the scenery ξ. The key idea in this paper is to view the random environment itself as a nearest neighbor walk. In this subsection we explain how to do this, by defining "the representation of the scenery ξ as a nearest neighbor walk". We need the following definitions: Let D be an integer interval, i.e. the intersection between a real interval and the integer numbers $\mathbb{Z}$. We call a function $T : D \to \mathbb{Z}$ a nearest neighbor walk, iff for each $t_1, t_2 \in D$ with $|t_1 - t_2| = 1$, we have that $|T(t_1) - T(t_2)| = 1$. In what follows, we will write S for the path of the process $\{S(k)\}_{k\geq 0}$, that is for $S : k \mapsto S(k), \mathbb{N} \longrightarrow \mathbb{Z}$. Let $\varphi : \mathbb{Z} \longrightarrow \{0,1\}$ be one of the two 4-periodic sceneries with period 0011 and $\varphi(0) = \varphi(1)$. Such a scenery φ has a very particular property: for every point in the scenery φ, one neighboring point has color 0, whilst the other one has color 1. This implies that for any color record φ there exists one and only one nearest neighbor walk T generating φ on the scenery φ once we know where T starts. We can use this fact to represent a color record as a nearest neighbor walk: the nearest neighbor walk representing a sequence of colors is simply defined to be the only nearest neighbor walk generating the color sequence on φ and starting at a given point, in general the origin. (For this to work the starting point must have the right color.)

A numerical example: Let $\varphi = (01011000101010100\ldots)$ be a color record we want to represent as a nearest neighbor walk. Let $\varphi : \mathbb{Z} \longrightarrow \{0,1\}$ be the 4-periodic scenery:

$\varphi(k)$	$\ldots$	0	0	1	1	0	0	1	1	0	0	1	1	$\ldots$
k	$\ldots$	-4	-3	-2	-1	0	1	2	3	4	5	6	7	$\ldots$

Define the nearest neighbor walk representing φ to be the only nearest neighbor walk $T : \mathbb{N} \longrightarrow \mathbb{Z}$ starting at the origin and generating the sequence φ on φ, that is such that $\varphi \circ T = \varphi$. In this example we get:

$T(t)$	0	-1	0	-1	-2	-3	-4	-3	-2	-3	-2	$\ldots$
t	0	1	2	3	4	5	6	7	8	9	10	$\ldots$

The scenery ξ we want to represent as a nearest neighbor walk is however a doubly infinite sequence. We will thus take the sequence $\xi(0), \xi(1), \xi(2), \xi(3), \ldots$ first and define with it the portion of the path of the nearest neighbor walk in positive time. Then we take $\xi(0), \xi(-1), \xi(-2), \xi(-3), \ldots$, and this defines us the part of the nearest neighbor walk in negative time.

An example: Let $\xi : \mathbb{Z} \longrightarrow \{0,1\}$ be a scenery with the following values close to the origin:

$\xi(k)$	$\ldots$	1	0	1	0	0	0	1	1	1	0	0	1	$\ldots$
k	$\ldots$	-4	-3	-2	-1	0	1	2	3	4	5	6	7	$\ldots$

Designate by R the nearest neighbor walk representing ξ. Then the part of ξ to the right of the origin defines the path of R which lies in positive time. In this example above, $(00111001\ldots)$ is responsible for this part of R. We get:

$R(t)$	0	1	2	3	2	1	0	-1	$\ldots$
t	0	1	2	3	4	5	6	7	$\ldots$

In the same way, the part of ξ which lies left to the origin is responsible for the restriction of R to the negative integers. In our example,

$(\ldots 1010)$ defines that part of R. We get:

$R(t)$	$\ldots$	2	1	2	1	0
t	$\ldots$	-4	-3	-2	-1	0

We are ready to define R formally:

Definition 3.2.1.

Let $\varphi : \mathbb{Z} \longrightarrow \{0,1\}$ designate the following 4-periodic (random) scenery:

- When $\xi(0) = 0$, we set $(\varphi(0), \varphi(1), \varphi(2), \varphi(3)) = (0,0,1,1)$.
- When $\xi(0) = 1$, we set $(\varphi(0), \varphi(1), \varphi(2), \varphi(3)) = (1,1,0,0)$.

The nearest neighbor walk $R : \mathbb{Z} \longrightarrow \mathbb{Z}$ representing the scenery ξ is defined to be the only (random) nearest neighbor walk R such that $R(0) = 0$ and $\varphi \circ R = \xi$, i.e. $\varphi(R(k)) = \xi(k)$ for all $k \in \mathbb{Z}$.

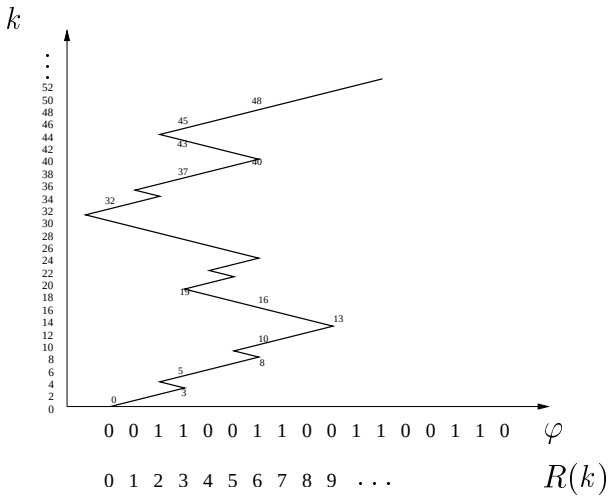
It is easy to verify that when the $\xi(k)$'s are i.i.d. Bernoulli variables with $P(\xi(0) = 0) = P(\xi(0) = 1) = 1/2$, then $\{R(k)\}_{k\in\mathbb{Z}}$ as well as $\{R(-k)\}_{k\in\mathbb{Z}}$ are two independent symmetric random walks starting at the origin.

In figure 1 below, we illustrate the above numerical example by showing a portion of the graph of R. For this we take

$$(\xi(0), \xi(1), \xi(2), \xi(3), \xi(4), \ldots) = (0011100101100011001000010011001001011001001111001\ldots)$$

We get the following picture:

Figure 1:



In the above graph, the label k designates the point $(R(k), k)$.

Next we need a few definitions:

Definition 3.2.2. Let $T : D \longrightarrow \mathbb{Z}$ be a nearest neighbor walk. Let $t_1,\ t_2 \in D$ and $x_1, x_2 \in \mathbb{Z}$, $x_1 \neq x_2$. We call (t_1, t_2) a crossing by T of (x_1, x_2) iff $(T(t_1), T(t_2)) = (x_1, x_2)$ and for all integer t strictly between t_1 and t_2, $T(t)$ is strictly between x_1 and x_2. If $t_2 > t_1$ we say that the crossing (t_1, t_2) is "positive", otherwise we say that it is "negative". If $|t_1 - t_2| = |x_1 - x_2|$ we say that the crossing $(t_1,\ t_2)$ is straight.
Let $(t_3,\ t_4)$ be a crossing by T of (x_3, x_4). Then, we say that $(t_3,\ t_4)$ is the *first crossing* by T of (x_3, x_4) during $(t_1,\ t_2)$ iff $t_3,\ t_4 \in [\min\{t_1, t_2\}, \max\{t_1, t_2\}]$ and $(t_3,\ t_4)$ is the crossing by T of (x_3, x_4) which lies in $[\min\{t_1, t_2\}, \max\{t_1, t_2\}]$ (i.e. $t_3,\ t_4 \in [\min\{t_1, t_2\}, \max\{t_1, t_2\}]$) and is *closest to* t_1.

Let (t_1, t_2) and (s_1, s_2) be two crossings by nearest neighbor walk T of (x_1, x_2). Then either the intervals:

$$]\min\{t_1, t_2\}, \max\{t_1, t_2\}[\quad \text{and} \quad]\min\{s_1, s_2\}, \max\{s_1, s_2\}[$$

are disjoint, or $(t_1, t_2) = (s_1, s_2)$ holds. Thus, we can numerate the crossings by T of (x_1, x_2) in increasing order of appearance. Thus the above definition of " first crossing by

T of (x_3, x_4) during another crossing" makes sense.

In the numerical example of figure 1, we see that between time 0 and time 3 the nearest neighbor walk R crosses from the point 0 to the point 3 in a straight way. In other words, $(0,3)$ is a straight crossing by R of $(0,3)$. Furthermore, R during the time interval $(0,13)$ crosses the interval $(0,9)$. Thus, $(0,13)$ is a crossing by R of $(0,9)$. Because $(0,3) \in (0,13)$ we have that the crossing $(0,3)$ happens during the crossing $(0,13)$. Clearly, $(0,3)$ is the first crossing by R of $(0,3)$ during the crossing $(0,13)$. (In the above example it is also the only one.) The crossing $(0,13)$, unlike $(0,3)$, is not a straight one. $(32,51)$ is a crossing by R of $(0,9)$. This is the second crossing by R of $(0,9)$ after time 0. During the crossing $(32,51)$ there are 2 crossings by R of the $(3,6)$. These are: $(37,40)$ and $(45,48)$.

3.2.3 Localization test

In this subsection, we construct a test to determine at what times the random walk is back at the same location. Combined with the idea of "going in shortest time from x to y", we have the the main ingredients for the reconstruction of a finite piece of the scenery ξ. If we have such a test, we can recognize when the random walk is back at a location x and at which times it is back at locations x and y. We then take a time interval where the random walks visits y in shortest possible time after visiting x.

This "localization test" is based on the representation R of the scenery ξ as a nearest neighbor walk. Recall that R is not observable. The composition of two nearest neighbor walks is again a nearest neighbor walk. Thus, the composition $R \circ S : k \longmapsto R(S(k))$, $\mathbb{N} \longrightarrow \mathbb{Z}$ is a nearest neighbor walk. However, every nearest neighbor walk $T : \mathbb{N} \longrightarrow \mathbb{Z}$ is uniquely determined by $\varphi \circ T$. In the following we set

$$T := R \circ S.$$

We get:

$$\varphi \circ T = (\varphi \circ R) \circ S = \xi \circ S = \chi,$$

i.e. T generates the color record χ on the scenery φ. Furthermore, $T(0) = 0$. Thus T is uniquely determined by the observations χ. Hence T is observable. Thus, although R and S are both not known, their composition $R \circ S$ is observable. We are using the nearest neighborwalk $R \circ S$ to determine when S is back at the same place.

To illustrate the **main idea** of the localization test (and maybe of this paper) we view the random walk S on the graph $k \longmapsto (R(k), k)$ geometrically in two dimensions. This defines a movement in two dimensions:

$$t \longmapsto (R(S(t)), S(t))$$

By projecting this movement along the y-axis on the x-axis we get the known 1-dimensional nearest neighbor walk T. Imagine that the path of R is given, then $t \longmapsto (R(S(t)), S(t))$ can be viewed as a one-dimensional random walk moving in $\mathbb{R}^2$ on the graph of R.

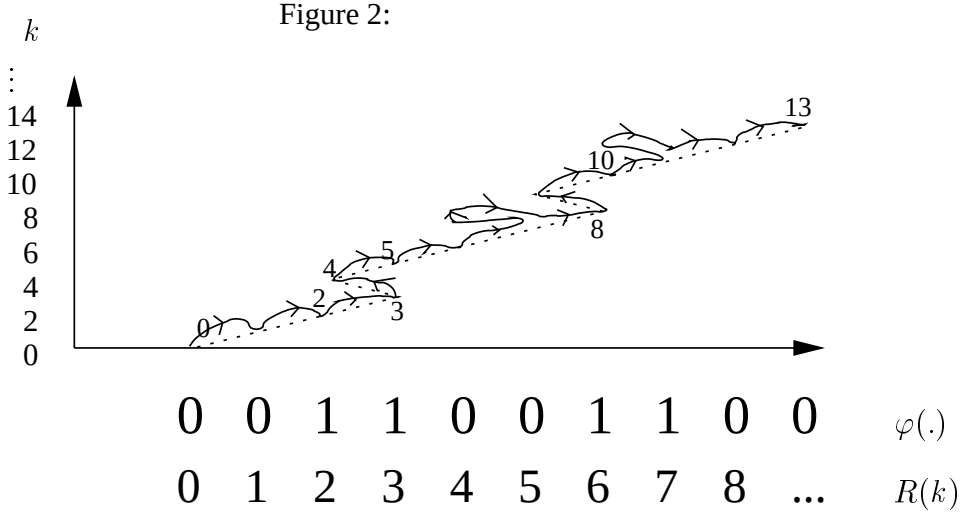
Figure 2 illustrates this situation. The graph of R is drawn as a dotted line, as it is not observable. The hand drawn lines with arrows represent the movement of the random walk S on the graph of G. This is the movement $t \longmapsto (R(S(t), S(t)))$, which is also not observable. However, the projection of this movement onto the horizontal line gives the observable nearest neighbor walk $R \circ S$, which is observable.

Let $\Delta S(k) := S(k+1) - S(k)$. In the example of figure 2 we have that

$$\begin{aligned}&(\Delta S(0), \Delta S(1), \Delta S(2), \ldots)\\ =&(+1,+1,+1,+1,+1,+1,+1,-1,+1,+1,+1,+1,+1,-1,+1,+1,+1,\ldots)\end{aligned}$$

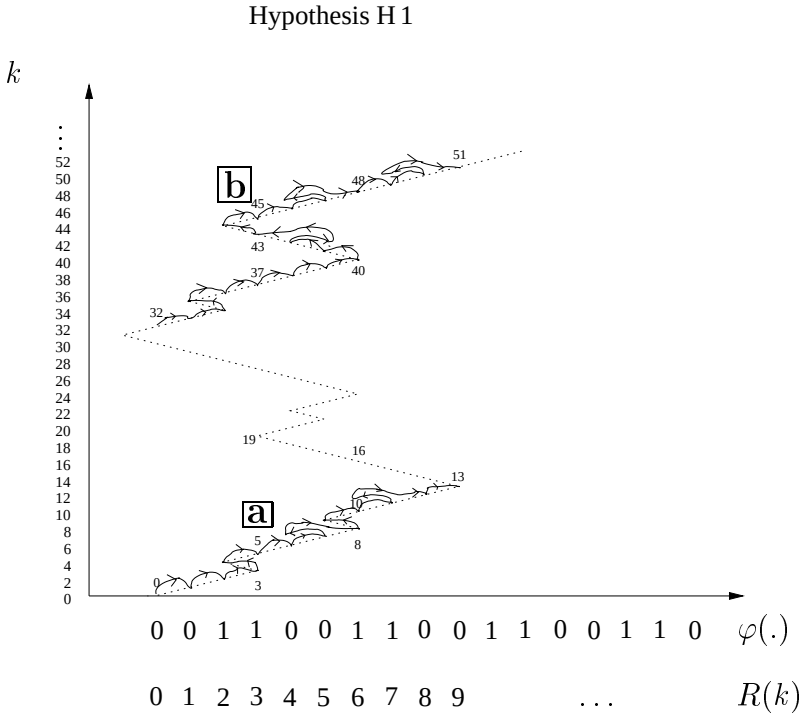
and R takes on the same values as in figure 1.

Figure 2:



Imagine that the dotted line representing the graph of R is made out of invisible glas. The random walk S moves invisibly on that glas line, but its projection onto the x-axis is visible. Seeing only this projection, we want to determine when S has returned to the same place. S has returned exactly when the 2-dimensional movement $t \longmapsto (R(S(t)), S(t))$ has returned to the same place: $S(s) = S(t)$ iff $(R(S(s)), S(s)) = (R(S(t)), S(t))$. Viewing R as fix, this means that S is back at the same place exactly when the random walk S on the graph of R has come back to the same place. As shown below, we can statistically determine this with high precision by counting the number of straight crossings of $R \circ S$ and their location. Let us illustrate the idea with figure 3.

Figure 3:



In figure 3, we show two finite portions of the movement of the random walk S on the graph of R. The first one is designated by the letter a whilst the second one is designated by the letter b. In this example a corresponds to the random walk S making the following first steps:

$$(\Delta S(0), \Delta S(1), \Delta S(2), \ldots) =$$

$$(+1, +1, +1, +1, +1, +1, +1, -1, +1, +1, +1, +1, +1, -1, +1, +1, +1, \ldots)$$

The part b start at time t_b such that $S(t_b) = 32$. Then the random walk S makes the following steps:

$$(\Delta S(t_b), \Delta S(t_b + 1), \Delta S(t_b + 2), \ldots) =$$

$$(+1, +1, +1, +1, +1, +1, +1, +1, +1, +1,$$
$$-1, +1, +1, +1, +1, +1, +1, -1, +1, +1, +1, +1, -1, +1, +1 \ldots)$$

The random walk S from time t_b until time $t_b + 25$ performs a crossing of the interval (32,51). This means that at time t_b the random walk S is at the point 32 and at time $t_b + 25$ it is at the point 51, but strictly in between the time t_b until time $t_b + 25$ the random walk S does not visit the points 32 or 51. In figure 3 if we project the movement b (of the random walk S on the graph of R) onto the horizontal line, we get the movement of the nearest neighbor walk $R \circ S$ during the time interval from time t_b until time $t_b + 25$. This is a crossing as well: during that time $R \circ S$ crosses from the point 0 to the point 9, that is it crosses the interval $(0, 9)$. During that time S on the graph of R, crosses a portion of the graph of R which corresponds itself to a crossing by R. As a matter of fact, between time 32 and time 51 the nearest neighbor walk R crosses the interval $(0, 9)$. Following our convention we say that $(32, 51)$

is a crossing by the nearest neighbor walk R of the interval $(0,9)$. In part a we see the following: $(0,17)$ is a crossing by S of $(0,13)$. On the other hand $(0,13)$ is a crossing by R of $(0,9)$. Eventually, $(0,17)$ is a crossing by $R \circ S$ of $(0,9)$.

The example of figure 3 illustrates one of the 3 main combinatorial facts used in this paper: the composition $T = R \circ S$ performs a crossing iff during that time S performs a crossing of a crossing of R. Let us formulate this as a lemma:

Lemma 3.2.1. *Let $0 < t_1 < t_2$. (t_1, t_2) is a crossing by T of the interval (x_1, x_2) iff there exist $k_1, k_2 \in \mathbb{Z}$ such that (t_1, t_2) is a crossing by S of (k_1, k_2), and (k_1, k_2) is a crossing by R of (x_1, x_2).*

Let us study next the example of figure 3 more: during time $(14,17)$, S performs a straight crossing of the interval $(10,13)$. Furthermore, $(10,13)$ represents itself a straight crossing by R of the interval $(6,9)$. This leads to, that $R \circ S$ performs during the time interval $(14,17)$ a straight crossing of the interval $(6,9)$. On the other hand, during time $(t_b, t_b + 4)$ S performs a straight crossing of the interval $(32,37)$. However $(32,37)$ is a crossing by R, but not a straight one. It follows that $(t_b, t_b + 4)$ is a crossing by $R \circ S$, but not a straight one.

The rule is: on a crossing by R which is not straight it is impossible to get a crossing by $R \circ S$ which is straight. This is the second main combinatorial fact:

Lemma 3.2.2. *Let $0 < t_1 < t_2$. Then (t_1, t_2) is a straight crossing by T of the interval (x_1, x_2) iff there exists $k_1, k_2 \in \mathbb{Z}$ such that (t_1, t_2) is a straight crossing by S of (k_1, k_2) and (k_1, k_2) is a straight crossing by R of (x_1, x_2).*

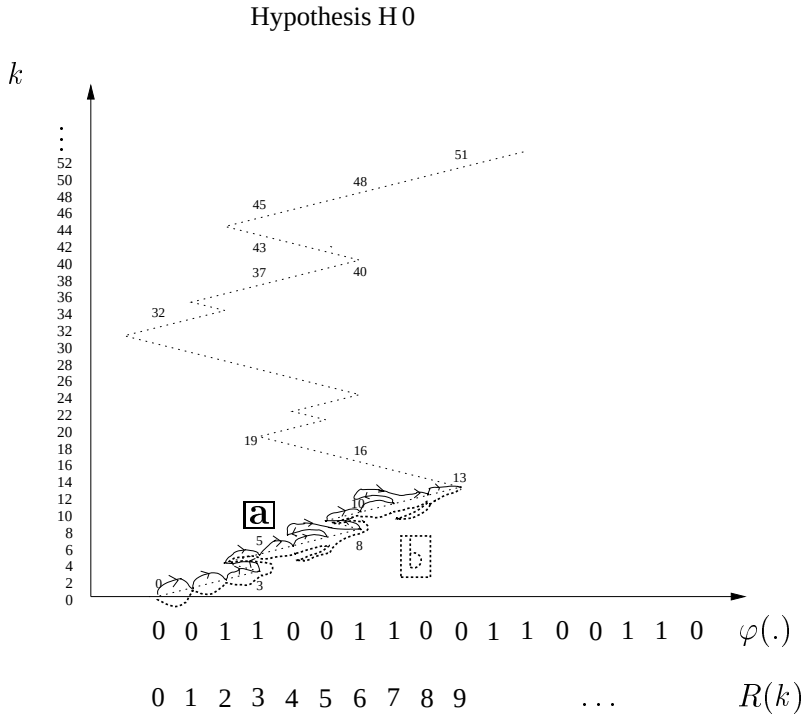
Looking further at figure 3, we see that in portion b of the path of S on the graph of R we have: during the crossing $(32,51)$ the first crossing by R of $(3,6)$ is $(37,40)$ and the last one is $(45,48)$. The first crossing by S of $(37,40)$ during $t_b, t_b + 51)$ is $(t_b + 5, t_b + 8)$. The first crossing during $(t_a, t_a + 25)$ by $R \circ S$ of $(3,6)$ is also $(t_b + 5, t_b + 8)$. Thus, the first crossing during $(t_a, t_a + 25)$ by $R \circ S$ of $(3,6)$ happens when during $(t_a, t_a + 25)$ S crosses for the first time the first crossing by R of $(3,6)$.

We see that a first crossing by $R \circ S$ corresponds to a first crossing by S of a first crossing by R. This yields our third combinatorial fact:

Lemma 3.2.3. *Let $0 < t_1 < t_2 < t_3 < t_4$ and $0 < x_1 < x_2 < x_3 < x_4$. Furthermore, let (t_1, t_4) be a crossing by $R \circ S$ of (x_1, x_4).*
Then (t_2, t_3) is the first crossing during (t_1, t_4) of (x_2, x_3) by $R \circ S$ iff it is the first crossing by S during (t_1, t_4) of (k_2, k_3), where (k_2, k_3) is the first crossing by R of (x_2, x_3) during (k_1, k_4).

To illustrate this, consider figure 3 above and figure 4 below:

Figure 4:



In figure 4, the portion b, of the path of the random walk S is traced on the graph of R as a thick doted line. It is a crossing by S of the crossing $(0, 13)$ by R. The projection onto the horizontal line of this movement is a crossing, too. In figure 4, the crossing b, that is $(t_b, t_b + 18)$, is a crossing by S of $(0, 13)$. Furthermore, $(t_b, t_b + 18)$ is also a crossing by $R \circ S$ of $(0, 9)$.
Figure 4 is identical to figure 3 except for the path b. In figure 3, the crossings a and b by S take place "in different locations", whilst in figure 4 they take place "on the same locations". Given figure 3 and figure 4, one can see if the crossings a and b by S take place "in the same location" or not. However, from the input data of the reconstruction problem, only the projection down onto the horizontal of the path of S on the graph of R is observable: In both cases , we observe two crossings a and b by $R \circ S$ of the same interval $(0, 9)$. Based on the observation of those crossings only, we need to infer if "the crossings occur on the same location" as in figure 3, or "on different locations" as in figure 4.

More generally: Assume we observe two crossings (t_{1a}, t_{2a}) and (t_{1b}, t_{2b}) by $R \circ S$ of an interval $(0, 3n)$; this interval instead of any other interval is chosen for notational convenience. Because of lemma 3.2.1, there exist k_{1a}, k_{2a} such that (k_{1a}, k_{2a}) is a crossing by R of $(0, 3n)$ whilst (t_{1a}, t_{2a}) is a crossing by S of (k_{1a}, k_{2a}). Similarly, there exists k_{1b}, k_{2b} such that (k_{1b}, k_{2b}) is a crossing by R of $(0, 3n)$ whilst (t_{1b}, t_{2b}) is a crossing by S

of (k_{1b}, k_{2b}).

In figure 3, $(t_{1a}, t_{2a}) = (0, 17)$, $(k_{1a}, k_{2a}) = (0, 13)$, $t_{2b} = t_{1b} + 25$, $(k_{1b}, k_{2b}) = (32, 51)$.

We develop a statistical test to determine if the two crossings (t_{1a}, t_{2a}) and (t_{1b}, t_{2b}) by S occur "on the same place" or not. Its input data are two observed crossings (t_{1a}, t_{2a}) and (t_{1b}, t_{2b}) by $R \circ S$ of the same interval. We define the hypotheses of our test:

- **Hypothesis $\mathbf{H_0}$:** during the crossings (t_{1a}, t_{2a}) and (t_{1b}, t_{2b}) the random walk S is on the same crossing of R. More precisely:
  $(S(t_{1a}), S(t_{2a})) = (S(t_{1b}), S(t_{2b}))$.
- **Hypothesis $\mathbf{H_1}$:** $(S(t_{1a}), S(t_{2a})) \neq (S(t_{1b}), S(t_{2b}))$

If H_0 holds, then $S(t_{2a}) = S(t_{2b})$, i.e. the random walk is back at the same place.

To determine if during two crossings by $R \circ S$ the random walk S was at the same place we are going to count the number of common straight crossings on three unit intervals. Let us explain how this is done.

We first, partition the interval $(0, 9)$ in disjoint intervals of length 3. This gives us the 3 intervals: $(0, 3)$, $(3, 6)$ and $(6, 9)$. Then we determine how many of these intervals are crossed in a straight way by $R \circ S$ when they get crossed for the first time during a and and when they get crossed for the first time during b. In figure 3, we see that the first crossing during a of $(0, 3)$ by $R \circ S$ is straight. However, the first crossing during b of $(0, 3)$ by $R \circ S$ is not. Thus, for the interval $(0, 3)$ we don't have a common first straight crossing. Next comes the interval $(3, 6)$. There, the first crossing by $R \circ S$ of $(3, 6)$ during a is not straight. (That first crossing is equal to $(5, 10)$.) On the other hand, the first crossing by $R \circ S$ of $(3, 6)$ during b is straight. (It is the first crossing $(t_b + 5, t_b + 8)$.) Again with the interval $(3, 6)$ we do not observe a common first straight crossing between a and b. Eventually the first crossing by $R \circ S$ of $(6, 9)$ during a is straight, whilst the first crossing by $R \circ S$ of $(6, 9)$ during b is not. So, in total we have zero common straight first crossings between a and b. When we observe few common first straight crossings between two crossings a and b by S, we decide that the crossings a and b took place on different places. In the example of figure 3, the person who only observes $R \circ S$ would thus decide that the crossings a and b by S took place on different places. In the case of figure 4, the first crossing by $R \circ S$ of $(0, 3)$ during a and during b are both straight. So for $(0, 3)$, we have a common first straight crossing. In figure 4 again, the first crossing by $R \circ S$ of $(3, 6)$ during a and during b are both not straight. The first crossing by $R \circ S$ of $(6, 9)$ during a is straight whilst during b it is not. Again for $(6, 9)$ we do not have a common straight crossing. Thus in the case, of figure 4, the total number of "straight common first crossings" equals 1.

General case: Let (t_{1a}, t_{2a}) and (t_{1b}, t_{2b}) be two crossings by $R \circ S$ of the interval $(0, 3n)$. For $0 \leq m < n$, let $w_a(m)$ be equal to 1 if the first crossing by $R \circ S$ of the interval $(3m, 3m + 3)$ during (t_{1a}, t_{2a}) is straight, and be equal to 0 otherwise. Let w_a denote the binary word $w_a(0), w_a(1), w_a(2), \ldots, w_a(n-1)$. In the same manner, define the binary word w_b for the crossing (t_{1b}, t_{2b}). The number of common straight crossings between a and b is defined to be the scalar product:

$$w_a \times w_b := \sum_{m=0}^{n-1} w_a(m) \cdot w_b(m).$$

We use $w_a \times w_b$ as test statistic. What is its distribution under $\mathbf{H_0}$ and under $\mathbf{H_1}$?

Example: To have a first common straight crossing in the H_0-case we need three crossings to be straight whilst in the H_1-case we need four. In order to understand why this is true, look at figure 4 first: we have there for $m = 0$ a first common straight crossing. This means, that when $R \circ S$ crosses during a and during b for the first time $(0, 3)$, we observe in both cases a straight crossing. That we have a common first straight crossing follows from the fact that: the first crossing by R of $(0, 3)$ during $(0, 13)$ is straight and the first crossing during a and during b of $(0, 3)$ are both straight as well. In figure 3, we have that $w_a(0) = 1$ and $w_b(0) = 0$. For $w_a(0) \cdot w_b(0)$ to be equal to 1 in figure one, there is only one thing missing: The first crossing $(32, 37)$ by R of the interval $(0, 3)$ should be straight.

General case: Let $m \in \mathbb{N}$ be such that $m < n$. Let $(k_{1a}, k_{2a}) = (S(t_{1a}), S(t_{2a}))$ and $(k_{1b}, k_{2b}) = (S(t_{1b}), S(t_{2b}))$. Let (k_{1am}, k_{2am}) designate the first crossing by R of $(3m, 3m+3)$ during (k_{1a}, k_{2a}). Let (k_{1bm}, k_{2bm}) designate the first crossing by R of $(3m, 3m+3)$ during (k_{1b}, k_{2b}). In the case of hypothesis H_0 we have $(k_{1a}, k_{2a}) = (k_{1b}, k_{2b})$ and $(k_{1am}, k_{2am}) = (k_{1bm}, k_{2bm})$. We get:

- **Under $\mathbf{H_0}$:** $w_a(m) \cdot w_b(m) = 1$ iff the following *three* *crossings* are straight:
  1. the crossing (k_{1am}, k_{2am}) by R of the interval $(3m, 3m + 3)$
  2. the first crossing by S during (t_{1a}, t_{2a}) of the interval (k_{1am}, k_{2am})
  3. the first crossing by S during (t_{1b}, t_{2b}) of the interval (k_{1am}, k_{2am})
- **Under $\mathbf{H_1}$:** $w_a(m) \cdot w_b(m) = 1$ iff the following *four* *crossings* are straight:
  1. the crossing (k_{1am}, k_{2am}) by R of the interval $(3m, 3m + 3)$
  2. the crossing $(k_{1b}m, k_{2bm})$ by R of the interval $(3m, 3m + 3)$
  3. the first crossing by S during (t_{1a}, t_{2a}) of the interval (k_{1am}, k_{2am})
  4. the first crossing by S during (t_{1b}, t_{2b}) of the interval (k_{1bm}, k_{2bm})

R and S are independent simple random walks. For the simple random walk a crossing of an interval of length 3 is straight with probability 3/4, as is shown below in fact e.5. Under H_0, there are 3 such crossings involved, whilst under H_1 there are 4. This, is why $P(w_a(m) \cdot w_b(m) = 1) = (3/4)^3$ in the case H_0 and $P(w_a(m) \cdot w_b(m) = 1) = (3/4)^4$ in the case H_1. By the Markov property, the variables $w_a(m) \cdot w_b(m)$ for different m's are independent. This gives:

The distribution of the test statistic $w_a \times w_b$ is equal to:

- **Under $\mathbf{H_0}$:** Binomial with parameter n and $(3/4)^3$
- **Under $\mathbf{H_1}$:** Binomial with parameter n and $(3/4)^4$

Let $c := \frac{1}{2}\left(\left(\frac{3}{4}\right)^3 + \left(\frac{3}{4}\right)^4\right)$.

Localization Test with parameter n:

- When $w_a \times w_b > c \cdot n$, we accept H_0.
- When $w_a \times w_b \leq c \cdot n$, we accept H_1.

The above statement about the distribution of the test statistic holds only if we select the pair of crossings $((t_{1a}, t_{2a}), (t_{1b}, t_{2b}))$ in an appropriate manner. For example, if we would choose (t_{1a}, t_{2a}) to be the first crossing by $R \circ S$ of $(0, 3n)$ such that $w_a(m) = 1$ for all $m < n$ and (t_{1b}, t_{2b}) be the first crossing by R of $(0, 3n)$ such that $w_a(m) = 1$ for all $m < n$, then obviously the above statement about the distributions would not hold. In lemma 3.2.4 below, the statement is made rigorous. For this we need to numerate the crossings by $R \circ S$ of $(0, 3n)$, in an appropriate manner. By lemma 3.2.1 we know that any crossing by $R \circ S$ of $(0, 3n)$ can be biewed as a crossing by S of a crossing by R of $(0, 3n)$. A crossing by $R \circ S$ of $(0, 3n)$ can thus be described in a unique manner as the i-th crossing by S of the z-th crossing by R of $(0, 3n)$. We index the crossings by R of $(0, 3n)$ by the set $\mathbb{Z}^* := \mathbb{Z} - \{0\}$. We call z-th crossing by R of $(0, 3n)$:
If $z > 0$, the z-th crossing by $R(k), k \geq 0$ of $(0, 3n)$.
If $z < 0$, the $|z|$-th crossing by $R(k), k \leq 0$ of $(0, 3n)$, where we count in reverse order starting at zero.
Thus, we index the crossings by $R \circ S$ of $(0, 3n)$ by the set $\mathbb{N}^* \times \mathbb{Z}^*$. For $(i, z) \in \mathbb{N}^* \times \mathbb{Z}^*$, the (i, z)-th crossing by $R \circ S$ of $(0, 3n)$, is the crossing which corresponds to the i-th crossing by S of the z-th crossings by R of $(0, 3n)$. Picking (t_{1a}, t_{2a}) and (t_{1a}, t_{2a}) by chosing non-randomly two elements in the index set $\mathbb{N}^* \times \mathbb{Z}^*$, makes the statement about the distribution of the test statistic rigorous. This is the content of the next lemma.

Lemma 3.2.4. *Let $z_a, z_b \in \mathbb{Z}^*$ and let $i_a, i_b \in \mathbb{N}^*$ be non-random numbers. Let (t_{1a}, t_{2a}) and (t_{1b}, t_{2b}) be the two crossings by $R \circ S$ of $(0, 3n)$ for which: (t_{1a}, t_{2a}) is the i_a-th crossing by S of the z_a-th crossing by R of $(0, 3n)$ and (t_{1b}, t_{2b}) is the i_b-th crossing by S of the z_b-th crossing by R of $(0, 3n)$. Then:*

- **H_0-case**, (i.e. case where $z_a = z_b$ and $(S(t_{1a}), S(t_{2a}) = (S(t_{1b}), S(t_{2b}))$): *$w_a \times w_b$ has Binomial distribution with parameter n and $(3/4)^3$*
- **H_1-case**, (i.e. case where $z_a \neq z_b$ and $(S(t_{1a}), S(t_{2a}) \neq (S(t_{1b}), S(t_{2b}))$): *$w_a \times w_b$ has Binomial distribution with parameter n and $(3/4)^4$*

Note that the index in $\mathbb{N}^* \times \mathbb{Z}^*$ of a crossing by $R \circ S$ of $(0, 3n)$ is not observable, (although the crossings by $R \circ S$ of $(0, 3n)$ are themselves observable.) However, by large deviation for the Binomial distribution, lemma 3.2.4 garanties that the probability of an error by our localization test is exponentially small in n, when the crossings compared correspond to two non-random indexes in $\mathbb{N}^* \times \mathbb{Z}^*$. We can not pick crossings by their index in $\mathbb{N}^* \times \mathbb{Z}^*$ for our reconstruction algorithm, since these are not observable. Hence, the crossings we select in an observable manner have slightly different distributions from the distributions

mentionned in lemma 3.2.4. But picking the crossings in an sensible, observable manner modifies the probability of an error only slightly, so that it remains small. Next, we need to mention a few facts which are useful for the proof of lemma 3.2.4:

Fact a Let $M(k)_{k\in\mathbb{N}}$ be a Markov chain with state space $\mathcal{M}$. Let $a_0, a_1, a_2, \ldots$ be a sequence of (non random) elements of $\mathcal{M}$. Let $\eta_{(i+1)}$ denote the first passage time of $M(k)_{k\in\mathbb{N}}$ at $a_{(i+1)}$ after η_i . Recursively: $\eta_0 := \min\{k \geq 0 | M(k) = a_0\}$. Then, $\eta_{i+1} := \min\{k \geq \eta_i | M(k) = a_{i+1}\}$. Let Z_i be the path of M between η_i and $\eta_i + 1$:

$$Z_i := (M(\eta_i), M(\eta_i + 1), M(\eta_i + 2), \ldots, M(\eta_{i+1}))\,.$$

Then, the Z_i's are independent of each other.

Fact b Let X and Z be two random objects living on the same space and independent of each other. Let A be an event that depends only on X, that is $A \in \sigma(X)$. Then conditional on A, X and Z are still independent of each other. Furthermore conditional on A, Z has the same marginal distribution. Thus:

$$\mathcal{L}\,(X, Z|\,A\,) = \mathcal{L}\,(X|\,A\,) \otimes \mathcal{L}\,(Z)\,.$$

Fact c Let $X_0, X_1, \ldots, X_n$ be a collection of random objects that are independent of each other. Let $A_0, A_1, \ldots, A_n$ be a collection of events such that for each $0 \leq i \leq n$, $A_i \in \sigma(X_i)$. Let $A := \cap_{i=0}^n A_i$. Then conditionally on A, the X_i's are still independent of each other:

$$\mathcal{L}\,(X_0, X_1, \ldots, X_n\,|A) = \prod_{i=0}^{n} \mathcal{L}\,(X_i\,|A_i)\,.$$

Fact d Let $X_0, X_1, \ldots, X_n$ be a collection of random objects that are independent of each other. Let $Y_0, Y_1, \ldots, Y_n$ be a collection of random objects satisfying: conditionally on $\sigma(X_m | 0 \leq m \leq n)$, the Y_m's are independent of each other and their distribution depends only on their respective X_m's. More precisely:

$$\mathcal{L}\,(Y_m\,|X_0, X_1, \ldots, X_n\,) = \mathcal{L}\,(Y_m\,|X_m\,)\,.$$

Let $Z_m := (X_m, Y_m)$. Then, the Z_m's are independent of each other.

Fact e Let $l > 0$ be a (non random) natural number. Let κ_l designate the first passage time of S at l. $\kappa_l := \min\{t | S(t) = l\}$. Let κ_0 designate the first recurrence time of S at 0. $\kappa_0 := \min\{t > 0 | S(t) = 0\}$. Let $E_{\text{cross l}}$ designate the event $\{\kappa_l < \kappa_0\}$. Let (j_{1i}, j_{2i}) be an increasing collection of intervals indexed by $i \in \mathbb{N}$ such that the following holds: $j_{1i} < j_{2i} \leq j_{1(i+1)}$. Assume furthermore, that $j_{10} = 0$. Let (s_{1i}, s_{2i}) denote the first crossing by S of (j_{1i}, j_{2i}). For natural numbers $s < t$ let $S(s,t) := (S(s), S(s+1), \ldots, S(t))$. Recall that $\Delta(s) := S(s+1) - S(s)$. Define $\Delta(s,t) := (\Delta(s), \Delta(s+1), \ldots, \Delta(t-1))$. With these definitions the following things hold:

1. The $S(s_{1i}, s_{2i})$'s for various i's are independent of each other. Similarly, the $\Delta(s_{1i}, s_{2i})$'s are independent of each other. **Proof of e.1:** Take the sequence $j_{20}, j_{21}, j_{22}, \ldots$ for the sequence $a_0, a_1, a_2, \ldots$ of fact a. The stopping times of fact a are then equal

to: $\eta_i := s_{2i}$. The crossing (s_{1i}, s_{2i}) happens between time $\eta_{(i-1)}$ and time η_i. By fact a, the pieces of path of S during the time intervals $[\eta_{(i-1)}, \eta_i]$ are independent of each other. Since the crossings (s_{1i}, s_{2i}) for different i's happen during different independent time intervals they are also independent.

2. The distribution of $S(s_{1i}, s_{2i})$ depends only on the length $d_i := j_{2i} - j_{1i}$. It is equal to the distribution of the path of the random walk starting at the point j_{1i} until it reaches j_{2i}, conditioned that it first meets j_{2i} before meeting j_{1i}. In other words, it is conditioned on, that the random walk S makes a crossing of (j_{1i}, j_{2i}). The random walk starting at j_{1i} is defined as: $\{S_i(t) := S(t) + j_{1i}\}_{t\in\mathbb{N}}$. With this notation, the distribution of $S(s_{1i}, s_{2i})$ equals:

$$\mathcal{L}\left(S_i(0, \kappa_{d_i}) \,|E_{\text{cross } d_i}\right)$$

or equivalently:

$$\mathcal{L}\left(S(t,\nu) \;\middle|\; \begin{array}{l} S(t) = j_{1i} \text{ and after time } t,\ S \text{ vis-} \\ \text{its by } S\ j_{2i} \text{ before it return for the} \\ \text{first time to } j_{1i} \end{array}\right)$$

where t designates any non-random time, and ν designates the first visit after t to j_{2i}.

3. The distribution of $\Delta(s_{1i}, s_{2i})$ depends only on the length d_i. It is equal to the distribution of $(\Delta(0), \Delta(1), \ldots, \Delta(\kappa_{d_i}))$ conditional on the event that the random walk first meets d_i before meeting 0. Thus,

$$\mathcal{L}\left(\Delta(s_{1i}, s_{2i})\right) = \mathcal{L}\left(\Delta(0), \Delta(1), \ldots, \Delta(\kappa_{d_i}) \,|E_{\text{cross } d_i}\right).$$

4. The joint distribution of the path of S during the crossings (s_{1i}, s_{2i}) is not changed if we condition on the event that the crossings (s_{1i}, s_{2i}) have to occur during a crossing. More precisely, we are considering the joint distribution of the (s_{1i}, s_{2i})'s for $0 \le i \le n$. We condition under the event that we have a crossing by S of $(0, j_{2n})$ starting at zero. After conditioning we get the same distribution as before:

$$\begin{aligned}\mathcal{L}(S(s_{10}, s_{20}), &S(s_{11}, s_{21}), \ldots, S(s_{1n}, s_{2n})) = \\ &\mathcal{L}\left(S(s_{10}, s_{20}), S(s_{11}, s_{21}), \ldots, S(s_{1n}, s_{2n}) \,|E_{\text{cross } j_{2n}}\right).\end{aligned}$$

Proof of e.4: Let $E^2_{\text{cross}}(i)$ be the event that S does not visit 0 during (s_{1i}, s_{2i}). $E^2_{\text{cross}}(i) := \{S(t) \neq 0, \forall t \in (s_{1i}, s_{2i}]\}$. In a similar manner define: $E^1_{\text{cross}}(i) := \{S(t) \neq 0, \forall t \in (s_{2(i-1)}, s_{1i}]\}$. We get

$$E_{\text{cross } j_{2n}} = \left(\bigcap_{i=0}^{n} E^1_{\text{cross}}(i)\right) \cap \left(\bigcap_{i=0}^{n} E^2_{\text{cross}}(i)\right).$$

The different pieces of paths from the collection:

$$\{S(s_{1i}, s_{2i}) \mid 0 \le i \le n\} \cup \{S(s_{1(i-1)}, s_{1i}) \mid 0 < i \le n\}$$

are independent of one another. Thus, we are exactly in the situation of fact c. Applying c to $\{S(s_{1i}, s_{2i}) \mid 0 \leq i \leq n\}$, we find that

$$\mathcal{L}\left(S(s_{10}, s_{20}), S(s_{11}, s_{21}), \ldots, S(s_{1n}, s_{2n}) \,|E_{\text{cross } j_{2n}}\right) \tag{3.2.1}$$

equals

$$\otimes_{i=0}^{n} \mathcal{L}\left(S(s_{1i}, s_{2i}) \,\middle|E^2_{\text{cross}}(i)\right).$$

However, since (s_{1i}, s_{2i}) is a crossing by S of (j_{1i}, j_{2i}) where $0 \leq j_{1i}, j_{2i}$, it follows that a.s. S during (s_{1i}, s_{2i}) does not visit 0. Thus the event $E^2_{\text{cross}}(i)$ is the almost sure event. Hence:

$$\mathcal{L}(S(s_{1i}, s_{2i}) \mid E^2_{\text{cross}}(i)) = \mathcal{L}(S(s_{1i}, s_{2i})).$$

This proves that the distribution 3.2.1 equals $\otimes_{i=0}^{n} \mathcal{L}(S(s_{1i}, s_{2i}))$. The last expression, by e.1, is however the joint distribution of the "unconditional" $S(s_{1i}, s_{2i})$'s.

5. The probability that a crossing by S of an interval of length 3 is straight equals 3/4. Thus, if $d_i = 3$, we have

$$P\left(s_{2i} - s_{1i} = 3\right) = \frac{3}{4}.$$

Proof of e.5: We need to calculate the probability $P(\kappa_3 = 3|E_{\text{cross } 3})$. $E_{\text{cross } 3}$ is the event that before coming back to zero, the random walk S first visits 3. It can do it in exactly $3, 5, 7, \ldots$ steps. For each given number of steps there is precisely one path. The reason is that when the random walk is in the interval $[0, 3]$, in order to not reach the border, there is always only one possible step. Any path of length $2k+1$ has probability $(1/2)^{2k+1}$. The path of length 3 is the straight path. We find:

$$P\left(\kappa_3 = 3 \,|E_{\text{cross } 3}\right) = \frac{P\left(\kappa_3 = 3\right)}{P\left(E_{\text{cross } 3}\right)} = \frac{\left(\frac{1}{2}\right)^3}{\sum_{k=1}^{\infty}\left(\frac{1}{2}\right)^{2k+1}} = \frac{3}{4}.$$

Note that fact e holds for any simple random walk.

Fact f Let $x_1 < x_2 \leq y_1 < y_2$. Let (t_{1xi}, t_{2xi}) designate the i-th crossing by S of (x_1, x_2). Let (t_{1yi}, t_{2yi}) designate the i-th crossing by S of (y_1, y_2). Then, $(S(t_{1xi}, t_{2xi}))_{i\geq 0}$ is independent of $(S(t_{1yi}, t_{2yi}))_{i\geq 0}$. **Proof of f:** Let ι_j designate the j-th visit by S to the point x_2. This defines a renewal process and a regenerative process. Since the random walk S can not jump, during each renewal period, it can either spend the whole time in $]\infty, x_2[$ or in $]x_2, \infty[$. During the same renewal period, S can not visit both $]\infty, x_2[$ and $]x_2, \infty[$. This implies that a crossing by S of (x_1, x_2) and crossing by S of (y_1, y_2) can never occur during the same renewal period. The renewal periods are independent of each other, i.e. the pieces of path $S(\iota_j, \iota_{j+1})$ are independent for various j's. Since the crossings by S of (x_1, x_2) and the crossings by S of (y_1, y_2), occur during different independent renewal times, it follows that $(S(t_{1xi}, t_{2xi}))_{i\geq 0}$ is independent of $(S(t_{1yi}, t_{2yi}))_{i\geq 0}$.

Fact g Let $x_1 < x_2$ be integer numbers. Let (t_{1xi}, t_{2xi}) designate the i-th crossing by S of (x_1, x_2). Then, the pieces of path $S(t_{1xi}, t_{2xi})$ are independent of each other for various i's. **Proof of g:** Assume without loss of generality that $0 < x_1 < x_2$. Let the sequence $a_0, a_1, a_2, \ldots$ be equal to the alternating sequence $x_1, x_2, x_1, x_2, x_1, \ldots$. Define like in fact a the stopping times η_j. In other words, η_0 designates the firs visit by S to a_0 and $\eta_{(j+1)}$ designates the first visit by S after time η_j to the point $a_{(j+1)}$. The piece of path inbetween stopping times are by fact a independent of each other. In other words, the $S(\eta_j, \eta_{(j+1)})$'s for different j's are independent. However, in each time interval $[\eta_j, \eta_{(j+1)})]$ there can be at most one crossing (t_{1xi}, t_{2xi}). It follows that the $S(t_{1xi}, t_{2xi})$ are independent of each other.

Notations Let $0 \leq m < n$. Let (k_{1z_a}, k_{2z_a}), resp. (k_{1z_b}, k_{2z_b}), designate the z_a-th, resp. z_b-th crossing by R of $(0, 3n)$. Let (k_{1am}, k_{2am}), resp. (k_{1bm}, k_{2bm}) designate the first crossing by R during (k_{1z_a}, k_{2z_a}), resp. (k_{1z_b}, k_{2z_b}), of $(3m, 3m+3)$. Let $w_a^R(m)$, resp. $w_b^R(m)$ designate the Bernoulli variable which is equal to one iff (k_{1am}, k_{2am}), resp. (k_{1bm}, k_{2bm}) is a straight crossing. Let (t_{1am}, t_{2am}), resp. (t_{1bm}, t_{2bm}) designate the first crossing by S during (t_{1a}, t_{2a}), resp. (t_{1b}, t_{2b}) of (k_{1am}, k_{2am}), resp. (k_{1bm}, k_{2bm}). Let $w_a^S(m)$, resp. $w_b^S(m)$ designate the Bernoulli variable which is equal to one iff (t_{1am}, t_{2am}), resp. (t_{1bm}, t_{2bm}) is a straight crossing. With this notation and by lemma 3.2.1, 3.2.2 and 3.2.3, we get $w_a^S(m) \cdot w_a^R(m) = w_a(m)$ and $w_b^S(m) \cdot w_b^R(m) = w_b(m)$. Hence, the test statistic $w_a \times w_b$ is equal to:

$$\sum_{m=0}^{n-1} w_a^S(m) w_a^R(m) w_b^S(m) w_b^R(m).$$

Note that the products $w_a^S(m) w_a^R(m) w_b^S(m) w_b^R(m)$ are Bernoulli random variables. Thus to prove lemma 3.2.4, we only need to prove that these products $w_a^S(m) w_a^R(m) w_b^S(m) w_b^R(m)$ for $m = 0, \ldots, n-1$ are i.i.d. random variables such that:

- Case H_0:

$$P\left(w_a^S(m) w_a^R(m) w_b^S(m) w_b^R(m) = 1\right) = \left(\frac{3}{4}\right)^3 \tag{3.2.2}$$

- Case H_1:

$$P\left(w_a^S(m) w_a^R(m) w_b^S(m) w_b^R(m) = 1\right) = \left(\frac{3}{4}\right)^4 \tag{3.2.3}$$

Proof of lemma 3.2.4 We need to distinguish two cases:

Case H_0: In this case $z_a = z_b$ and $w_a^R(m) = w_b^R(m)$ for all $0 \leq m < n$. Thus,

$$w_a^S(m) w_a^R(m) w_b^S(m) w_b^R(m) = w_a^S(m) w_a^R(m) w_b^S(m).$$

It follows:

$$P(w_a(m) w_b(m) = 1) = P((w_a^S(m) w_b^S(m)) = 1, w_a^R(m) = 1).$$

The right side of the last equality can be written as:

$$P\left(w_a^S(m) w_b^S(m) = 1 \,\middle|\, w_a^R(m) = 1\right) P(w_a^R(m) = 1). \tag{3.2.4}$$

We have that:

$$P\left(w_a^S(m)w_b^S(m)=1\,\middle|w_a^R(m)=1\right)= \\ E\left[\,P\left(\,w_a^S(m)w_b^S(m)=1\middle|\,R(k),k\in\mathbb{Z}\right)\,\middle|\,w_a^R(m)=1\right] \qquad (3.2.5)$$

The crossings (t_{1a},t_{2a}) and (t_{1b},t_{2b}) are crossings by S of the random interval (k_{1z_a},k_{2z_a}). So fact g does not directly apply. However, by conditioning on $\sigma(R(k),k\in\mathbb{Z})$ the interval (k_{1z_a},k_{2z_a}) is no longer random and we can apply fact g: Conditioned on $\sigma(R(k),k\in\mathbb{Z})$, $S(t_{1a},t_{2a})$ and $S(t_{1b},t_{2b})$ are independent of each other. Conditional on $\sigma(R(k),k\in\mathbb{Z})$, $w_a^S(m)$ only depends on $S(t_{1a},t_{2a})$, whilst $w_b^S(m)$ only depends on $S(t_{1b},t_{2b})$. Hence when we condition on R, $w_a^S(m)$ and $w_b^S(m)$ become independent. We get:

$$P\left(w_a^S(m)w_b^S(m)\,=1\,|R(k),k\in\mathbb{Z}\right)= \\ P\left(w_a^S(m)=1\middle|\,R(k),k\in\mathbb{Z}\right)\cdot P\left(w_b^S(m)=1\middle|\,R(k),k\in\mathbb{Z}\right).$$

When $w_a^R(m)=1$, then the crossing (k_{1am},k_{2am}) has length 3, i.e. $|k_{1am}-k_{2am}|=3$. Thus, by fact e.4 and e.5 we find that $P(w_a^S(m)=1|w_a^R(m)=1)=3/4$ and $P(w_b^S(m)=1|w_a^R(m)=1)=3/4$. So, when $w_a^R(m)=1$ holds, we find that

$$P\left(w_a^S(m)w_b^S(m)=1\middle|\,R(k),k\in\mathbb{Z}\right)=\left(\frac{3}{4}\right)^2.$$

This implies that the right side of equality 3.2.5 is equal to
$E[(3/4)^2|w_a^R(m)=1]=(3/4)^2$. Plugging this into 3.2.4, finishes to establish equality 3.2.2. Next we need to demonstrate the independence of the products $w_a^S(m)w_b^S(m)w_a^R(m)$ for $0\le m<n$ in the case H_0. Conditional on $\sigma(R(k),k\in\mathbb{Z})$ all of the following holds:
According to fact g, $S(t_{1a},t_{2a})$ is independent of $S(t_{1b},t_{2b})$. But the $w_a^S(m)$'s for various m's depend only on $S(t_{1a},t_{2a})$ and the $w_b^S(m)$'s for various m's depend only on $S(t_{1b},t_{2b})$. Thus, $(w_a^S(m))_{0\le m<n}$ is independent of $(w_b^S(m))_{0\le m<n}$. Furthermore, by fact e.1, the $w_a^S(m)$'s, resp. the $w_b^S(m)$'s for various m's are independent of each other. This leads to that the products $w_a^S(m)w_b^S(m)$ are independent of each other. (All the last arguments were meant to hold conditionally on $\sigma(R(k),k\in\mathbb{Z})$).
By fact e.1, the $R(k_{1am},k_{2am})$'s are independent among each other for various m's. This puts as in the case of fact d: Take for this $R(k_{1am},k_{2am})$ to be X_m and Y_m to be $w_a^S(m)w_b^S(m)$. Conditional on $(R(k_{1am},k_{2am}))_{0\le m<n}$ the $w_a^S(m)w_b^S(m)$'s are independent of each other and the conditional distribution of $w_a^S(m)w_b^S(m)$ depends only on $R(k_{1am},k_{2am})$. Fact d tells that in this case the random pairs
$(w_a^S(m)w_b^S(m),R(k_{1am},k_{2am}))$ for $0\le m<n$ must be independent. It follows that the products $w_a^S(m)w_b^S(m)w_a^R(m)$ are also independent of each other.

Case H_1: In this case the crossing (k_{1z_a},k_{2z_a}) is different from the crossing (k_{1z_b},k_{2z_b}). Fact g implies that $R(k_{1z_a},k_{2z_a})$ is independent of $R(k_{1z_b},k_{2z_b})$. This implies that $(R(k_{1am},k_{2am}))_{0\le m<n}$ is independent of $(R(k_{1bm},k_{2bm}))_{0\le m<n}$. Conditioned on $\sigma(R(k),k\in\mathbb{Z})$, the crossings (t_{1a},t_{2a}) and (t_{1b},t_{2b}) by S are crossing of non random intervals. Hence, conditional on $\sigma(R(k),k\in\mathbb{Z})$ and by fact f, $S(t_{1a},t_{2a})$ and $S(t_{1b},t_{2b})$ are independent of one another. Fact e.2 implies that conditional on $\sigma(R(k),k\in\mathbb{Z})$, the distribution of $(S(t_{1am},t_{2am}))_{0\le m<n}$, resp.
$(S(t_{1bm},t_{2bm}))_{0\le m<n}$ depends only on $(R(k_{1am},k_{2am}))_{0\le m<n}$,

resp. $(R(k_{1bm}, k_{2bm}))_{0\le m<n}$. Thus, fact d applies, and we get that $((S(t_{1am}, t_{2am}), R(k_{1am}, k_{2am})))_{0\le m<n}$ is independent of $((S(t_{1bm}, t_{2bm}), R(k_{1bm}, k_{2bm})))_{0\le m<n}$. Note that $w_a^S(m)w_a^R(m)$, resp. $w_b^S(m)w_b^R(m)$ is $\sigma((S(t_{1am}, t_{2am}), R(k_{1am}, k_{2am})))$, resp. $\sigma((S(t_{1am}, t_{2am}), R(k_{1am}, k_{2am})$ measurable. Thus,$(w_a^S(m)w_a^R(m))_{0\le m<n}$ is independent of $(w_b^S(m)w_b^R(m))_{0\le m<n}$.
Conditionally on $(R(k_{1bm}, k_{2bm}))_{0\le m<n}$, the crossings by S, (t_{1am}, t_{2am}) for $0 \le m < n$ are crossings of non-random intervals. Hence, fact f applies so that conditionally on $(R(k_{1bm}, k_{2bm}))_{0\le m<n}$ the pieces of paths $S(t_{1am}, t_{2am})$ are independent of each other for various m's. By fact e.2 and e.4, conditionally on $(R(k_{1bm}, k_{2bm}))_{0\le m<n}$, the distribution of $S(t_{1am}, t_{2am})$ depends only on $(R(k_{1am}, k_{2am}))$. However, by fact e.1, the pieces of paths $(R(k_{1am}, k_{2am}))$ are independent of each other for various m's. Thus, we can apply fact d, and get that the pairs: $(R(k_{1am}, k_{2am}), S(t_{1am}, t_{2am}))$ for $0 \le m < n$ are independent of each other. Since, $w_a^S(m)w_a^R(m)$ is $\sigma((R(k_{1am}, k_{2am}), S(t_{1am}, t_{2am}))$-measurable, it follows that the products $w_a^S(m)w_a^R(m)$ for $0 \le m < n$ are independent of each other. In a similar way, one can show that the products $w_b^S(m)w_b^R(m)$ for $0 \le m < n$ are independent of each other. It follows that the products $w_a^S(m)w_a^R(m)w_b^S(m)w_b^R(m)$ for various m's are i.i.d.. By independence of a and b, we have that:

$$P\left(w_a^S(m)\, w_a^R(m) w_b^S(m)\, w_b^R(m) = 1\right) = \\ P\left(w_a^S(m)w_a^R(m) = 1\right) P\left(w_b^S(m)w_b^R(m) = 1\right).$$

The right side of the last equality is equal to $P(w_a^S(m)w_a^R(m) = 1)^2$, because $P(w_a^S(m)w_a^R(m) = 1) = P(w_b^S(m)w_b^R(m) = 1)$. Furthermore:

$$P\left(w_a^S(m)w_a^R(m) = 1\right) = P\left(w_a^S(m) = 1 \,\middle|\, w_a^R(m) = 1\right) P\left(w_a^R(m) = 1\right).$$

By fact e.5, $P\left(w_a^R(m) = 1\right) = 3/4$. When $w_a^R(m) = 1$, then $|k_{1am} - k_{2am}| = 3$. $|t_{1am} - t_{2am}|$ designates the first crossing by S of (k_{1am}, k_{2am}). Thus by fact e.5, $P\left(w_a^S(m) = 1 \,\middle|\, w_a^R(m) = 1\right) = 3/4$. We are done with proving equation 3.2.3.

3.2.4 Details of the reconstruction algorithm

We gave already the main ideas, on how to reconstruct a finite piece of scenery. In this subsection we describe the technical details. Let (k_1^{n+}, k_2^{n+}) be the first crossing after time 0 by R of the interval $(0, 3n)$. In other words: $k_1^{n+}, k_2^{n+} \ge 0$ and for all $s, t \ge 0$ such that (s, t) is a crossing by R of the interval $(0, 3n)$ we have $k_1^{n+} \le s$ and $k_2^{n+} \le t$.
Let (k_1^{n-}, k_2^{n-}) be the last crossing before time 0 by R of the interval $(0, 3n)$. In other words: $k_1^{n-}, k_2^{n-} \le 0$ and for all $s, t \le 0$ such that (s, t) is a crossing by R of the interval $(0, 3n)$ we have $k_1^{n-} \ge s$ and $k_2^{n-} \ge t$.

In the numerical example of figure 1, we have that: $(k_1^{3+}, k_2^{3+}) = (0, 13)$. In other words,$(0, 13)$ is the first crossing after zero by R of $(0, 9)$. The part of the graph $z \mapsto R(z)$ with $z < 0$ is not represented in figure 1, so we can not see there (k_1^{3-}, k_2^{3-}).

The reconstruction algorithm which reconstructs a finite piece of the scenery ξ, reconstructs the word $\xi(k_2^{n-}), \xi(k_2^{n-} + 1), \xi(k_2^{n-} + 2), \ldots, \xi(k_2^{n+})$ or its transpose. It achieves this by recognizing a time interval (r, s) during which the nearest neighbor walk S goes in a straight way:

from the point k_2^{n-} to the point k_2^{n+}

or

from the point k_2^{n+} to the point k_2^{n-}.
(r,s) is thus a straight crossing by S of (k_2^{n-},k_2^{n+}) or of (k_2^{n+},k_2^{n-}). During such a straight crossing (r,s) the observations reveals the piece of the scenery ξ which is comprised between k_2^{n-} and k_2^{n+}: $\chi(r),\chi(r+1),\chi(r+2),\ldots,\chi(s)$ is equal to the word $\xi(k_2^{n-}),\xi(k_2^{n-}+1),\xi(k_2^{n-}+2),\ldots,\xi(k_2^{n+})$ or its transpose. The reconstruction algorithm "for a finite piece of scenery" depends on a parameter n. That is why we will call it the **reconstruction algorithm at level n**. Thus, we have a collection of algorithms indicesed by n. Using these algorithms for increasing n's will allow us to reconstruct increasing finite pieces of the scenery ξ and eventually to reconstruct the whole scenery ξ up to equivalence. (As a limit, after infinite time.) We can already mention here that the reconstruction algorithm at level n does not achieve his goal in 100 percent of the cases: rather it has a small failure probability. However this failure probability is finitely summable over n. This insures that only a finite number of the finite size reconstructions will contain errors. This finite number of errors have no influence on the final total reconstruction, since that one is taken to be a limit.
Next we need a few definitions and notations: let $z_1,z_2\in\mathbb{Z}$ be such that $|z_1-z_2|$ is a multiple of 3, that is there exists $z\in\mathbb{Z}$ such that $z_2-z_1=3z$. Let (s_1,s_2) be a crossing by $R\circ S$ of (z_1,z_2). Let, for $0\le m<|z|$, $w(m)$ be equal to 1 iff the first crossing by $R\circ S$ of $(z_1+3m(z/|z|),z_1+(3m+3)(z/|z|))$ during (s_1,s_2) is straight and equal to zero otherwise. We write $w_{(s_1,s_2)}$ for the binary word:

$$w(0)w(1)w(2)\ldots w(|z|-1)$$

and call it the binary word associated with the crossing (s_1,s_2) by $R\circ S$.
Among the two crossings by R, (k_1^{n+},k_2^{n+}) and (k_1^{n-},k_2^{n-}), let (k_{1a}^n,k_{2a}^{n+}) designate the one of the two which gets crossed first by S. In a similar way, let (k_{1c}^n,k_{2c}^n) designate the other one. In this way, if k_2^{n+} gets visited by S before k_2^{n-}, we have that (k_{1a}^n,k_{2a}^{n+}) equals (k_1^{n+},k_2^{n+}). Otherwise, (k_{1a}^n,k_{2a}^n) equals (k_1^{n-},k_2^{n-}).
Let (t_{1i}^n,t_{2i}^n) designate the i-the crossing by $R\circ S$ or the interval $(0,3n)$. Let w_i^n designate the binary words associated with the crossing (t_{1i}^n,t_{2i}^n). Thus:

$$w_i^n:=w_{(t_{1i}^n,t_{2i}^n)}$$

For $z\neq 0$ with $z\in\mathbb{Z}$, let (k_{1z}^n,k_{2z}^n) designate the z-th crossing by R of $(0,3n)$. (By this we mean that if $z>0$ then (k_{1z}^n,k_{2z}^n) is the z-th crossing after 0 by R of $(0,3n)$. If $z<0$, (k_{1z}^n,k_{2z}^n) designates the $|z|$-last crossing before 0 by R of $(0,3n)$.) Note that with this notation, we have that $(k_{11}^n,k_{21}^n)=(k_1^{n+},k_2^{n+})$ and $(k_{1(-1)}^n,k_{2(-1)}^n)=(k_1^{n-},k_2^{n-})$. Because S starts at the origin, it can not reach any z-th crossing (k_{1z}^n,k_{2z}^n), with $|z|>1$ before it has not crossed (k_1^{n+},k_2^{n+}) or (k_1^{n-},k_2^{n-}) . By lemma 3.2.1, (t_{11}^n,t_{21}^n) is also the first crossing by S of a crossing by R of $(0,3n)$. It follows, that (t_{11}^n,t_{21}^n) is obligatorily a crossing by S of either (k_1^{n+},k_2^{n+}) or (k_1^{n-},k_2^{n-}). Thus, (t_{11}^n,t_{21}^n) is a crossing by S of (k_{1a}^n,k_{2a}^n).
The above discussion suggests a method for constructing stopping times which with high probability will stop the random walk at the point k_{2a}^n. Apply for this the localization test to the two crossings: (t_{11}^n,t_{21}^n) and (t_{1i}^n,t_{2i}^n). If the test decides that (t_{11}^n,t_{21}^n) and (t_{1i}^n,t_{2i}^n) are crossings by S of the same interval (i.e. Hypothesis H_0), decide that $S(t_{2i}^n)=k_{2a}^n$. Let $\tau^n(i)$ designate the i-th stopping time obtained by trying to stop the random walk S at k_{2a}^n . More precisely, $\tau^n(i)$ is equal to the i-th, t_{2j}^n for which:

$$w_j^n\times w_1^n>c\cdot n.$$

The scalar product for binary words of the same length $\times$ is defined in the following way: let $w = w(0)w(1)w(2)...w(k)$ and $v = v(0)v(1)v(2)...v(k)$ be two binary words. $w \times v := \sum_{l=0}^{k} w(l) \cdot v(l)$. We define the relation $\leq$: $w \leq v$ iff for all l with $0 \leq l \leq k$ we have that $w(l) \leq v(l)$. We define the transpose of the word w and write w^* for the word $w^* := w(k)w(k-1)w(k-2)...w(1)$.
Let (t^n_{1a}, t^n_{2a}) denote the first crossing by S of the interval (k^n_{1a}, k^n_{2a}). We have that $(t^n_{1a}, t^n_{2a}) = (t^n_{11}, t^n_{21})$. Let (t^n_{1c}, t^n_{2c}) denote the first crossing by S of the interval (k^n_{1c}, k^n_{2c}). As mentioned, (t^n_{1a}, t^n_{2a}) is also the first crossing by $R \circ S$ of the interval $(0, 3n)$, and thus is observable. Let w^n_a designate the binary word associated with the crossing (t^n_{1a}, t^n_{2a}) by $R \circ S$. Using our notation:

$$w^n_a := w_{(t^n_{1a}, t^n_{2a})}$$

Note that (t^n_{1c}, t^n_{2c}) is also a crossing by $R \circ S$ of the interval $(0, 3n)$. Let w_c denote the binary word associated with the crossing (t^n_{1c}, t^n_{2c}) by $R \circ S$. (t^n_{1c}, t^n_{2c}) and w^n_c are not directly observable. We can only estimate them. We denote by $\hat{w}^n_c$ our estimate for w^n_c and by $(\hat{t}^n_{1c}, \hat{t}^n_{2c})$ our estimate for (t^n_{1c}, t^n_{2c}) . We will explain later how we obtain these estimates.
As already mentioned the goal of the reconstruction algorithm at level n is to reconstruct the finite piece of the scenery ξ:

$$\xi(k^n_{2c}), \xi(k^n_{2c} + u), \xi(k^n_{2c} + 2u), \ldots, \xi(k^n_{2a}).$$

(Here u denotes the signe: $u := (k^n_{2a} - k^n_{2c})/|(k^n_{2a} - k^n_{2c})|$). The reconstruction algorithm at level n achieves this, by constructing a straight crossing (s, r) by S of (k^n_{2c}, k^n_{2a}). When going from k^n_{2c} to k^n_{2a} in a straight way, the random walk S first crosses the interval (k^n_{2c}, k^n_{1c}) in a straight way and then the interval (k^n_{1a}, k^n_{2a}). Crossing (k^n_{2c}, k^n_{1c}), resp. (k^n_{1a}, k^n_{2a}) in a straight way, we get the maximum number of "straight crossings possible by $R \circ S$". Thus, when (s, r) with $s < r$ is a straight crossing by S of (k^n_{2c}, k^n_{2a}) we have that:
there exists $s_2 \leq s_1 \leq r_1 \leq r_2$ with $s_2 = s, r_2 = r$ such that:
(s_2, s_1) is a straight crossing by S of (k^n_{2c}, k^n_{1c}) and
(r_1, r_2) is a straight crossing by S of the interval (k^n_{1a}, k^n_{2a}).
In this case:

$$w_{(s_1, s_2)} \geq w^n_c \tag{3.2.6}$$

and :

$$w_{(r_1, r_2)} \geq w^n_a \tag{3.2.7}$$

The above discussion suggests a method on how to search for straight crossings (s, r) by S of the interval (k^n_{2c}, k^n_{2a}): try to find (s, r) minimizing r-s with $s < r$ under the following constraint:
there exists $s_2 \leq s_1 \leq r_1 \leq r_2$ with $s_2 = s, r_2 = r$ such that:

1. (s_1, s_2) is a crossing by $R \circ S$ of $(0, 3n)$ such that inequality 3.2.6 is satisfied.
2. (r_1, r_2) is a crossing by $R \circ S$ of $(0, 3n)$ such that inequality 3.2.7 is satisfied.

3.2.5 The reconstruction algorithm at level n

Let $\bar{n} := n^{10.89}$ and $\tilde{n} := n^{11}$. We are now ready to define the **reconstruction algorithm at level n** in a precise way:

Algorithm 3.2.1.

- Find (s,r) minimizing r-s with $s<r$ under the following constraint:
  1. There exists $i \le e^{\tilde{n}}$ such that $\tau^{\tilde{n}}(i) \le s < r \le \tau^{\tilde{n}}(i) + n^{220}$.
  2. There exists $s_2 \le s_1 \le r_1 \le r_2$ with $s_2 = s, r_2 = r$ such that:
     (a) (s_1, s_2) is a crossing by $R \circ S$ of $(0, 3n)$ such that $w_{(s_1,s_2)} \ge \hat{w}_c^n$ holds.
     (b) (r_1, r_2) is a crossing by $R \circ S$ of $(0, 3n)$ such that $w_{(r_1,r_2)} \ge w_a^n$ holds.
- The output of the reconstruction algorithm at level n is the binary word which we can read in the observations χ during time (s,r), that is:

$$\chi(s), \chi(s+1), \chi(s+2), \ldots, \chi(r)$$

where (s,r) designates the first ordered pair minimizing $r-s$ under the constraints 1,2.a and 2.b.

Remark 3.2.1.

- w_c^n is not directly observable. Thus, for our reconstruction algorithm we use the estimate $\hat{w}_c^n$ instead of w_c^n.
- The reader might be wondering why the algorithm uses condition 2.a and 2.b instead of the localization test. As a matter of fact, one could imagine to replace condition 2 by the following two conditions:
  - (s_1, s_2) is a crossing by $R \circ S$ of $(0, 3n)$ such that when compared to the crossing $(\hat{t}_{1c}^n, \hat{t}_{2c}^n)$ the localization test decides that the two crossings occurred in the same place (H_0-case).
  - (r_1, r_2) is a crossing by $R \circ S$ of $(0, 3n)$ such that when compared to the crossing (t_{1a}^n, t_{2a}^n) the localization test decides that the two crossings occurred in the same place (H_0-case).

  Replacing conditions 2.a and 2.b by the above conditions 1 and 2 does not work. The reason is the following: typically the points k_{2a}^n and k_{2c}^n are at distance order(n^9) from each other. To get at least one straight crossing by S of an interval of length order(n^9) we need order(2^{n^9}) trials. Thus our algorithm needs to be able to identify correctly order(2^{n^9}) crossings by S of (k_{2c}^n, k_{2a}^n). The localization algorithm (with parameter n) has a positive probability of making an error of order$(e^{-k \cdot n})$ where $k>0$ is a constant not depending on n. With order(2^{n^9}) trials we can be sure that the localization test (with parameter n) will make many errors, and thus can not be used instead of conditions 2.a and 2.b.
- If we perform the localization test with parameter $\tilde{n}$ instead of n, the probability of an error is of order$(e^{-k\tilde{n}})$. This is so small, that with high probability, we can apply it order$(e^{k\tilde{n}})$ times without making a single mistake. This is more then enough trials, to get with high probability one straight crossing by S of an interval of length order(n^9). This is why for condition 1 in the reconstruction algorithm at level n, we construct the stopping times $\tau^{\tilde{n}}(i)$ using the localization algorithm with parameter $\tilde{n}$.

- The conditions 2.a and 2.b can be seen as a modified version of the localization algorithm with parameter n. We will show that with high probability within distance n^{220} of the point $k_{2a}^{\tilde{n}}$ we have:
  only the crossing (k_{1a}^n, k_{2a}^n) is such that a crossing (r_1, r_2) by S of it can satisfy inequality $w_{(r_1,r_2)} \leq w_a$. A similar condition also holds for (k_{1c}^n, k_{2c}^n).
  This implies that as long as we are within distance n^{220} of the point $k_{2a}^{\tilde{n}}$ conditions 2.a and 2.b can never make a mistake at identifying crossings by S of (k_{1a}^n, k_{2a}^n) and of (k_{1c}^n, k_{2c}^n). When, $S(\tau^{\tilde{n}}(i)) = k_{2a}^{\tilde{n}}$, then by definition, a crossing (s, r) satisfying condition 1 of the selection rule of the reconstruction algorithm at level n, is such that $S(s)$ and $S(r)$ are within distance n^{220} of the point $k_{2a}^{\tilde{n}}$. For more details about why the reconstruction algorithm at level n works, see section 3.4.

3.2.6 Construction of $(\hat{t}_{1c}^n, \hat{t}_{2c}^n)$ and of $\hat{w}_c^n$

Recall that a crossing (s, t) is called positive if $s < t$ and negative otherwise. Recall also that from the two crossings (k_{11}^n, k_{21}^n) and $(k_{1(-1)}^n, k_{2(-1)}^n)$ by R of $(0, 3n)$ the one which gets first crossed by S is called (k_{1a}^n, k_{2a}^n) whilst the other one is called (k_{1c}^n, k_{2c}^n). After having crossed from the point k_{1a}^n to the point k_{2a}^n, S first needs to cross back from the point k_{2a}^n to the point k_{1a}^n before being able to cross (k_{1c}^n, k_{2c}^n). More precisely, after a positive crossing by S of (k_{1a}^n, k_{2a}^n) there first needs to be a negative crossing by S of (k_{1a}^n, k_{2a}^n) before there can be a crossing by S of (k_{1c}^n, k_{2c}^n). On the other hand, right after a negative crossing by S of (k_{1a}^n, k_{2a}^n) the random walk S is always located between the points k_{1a}^n and k_{1c}^n. When, the random walk S is located between k_{1a}^n and k_{1c}^n, the next time it crosses an interval (k_{1z}^n, k_{2z}^n) this must be the interval (k_{1a}^n, k_{2a}^n) or (k_{1c}^n, k_{2c}^n). This gives a way to characterize (t_{1c}^n, t_{2c}^n): (Recall that (t_{1c}^n, t_{2c}^n) is the first crossing by S of (k_{1c}^n, k_{2c}^n).)
(t_{1c}^n, t_{2c}^n) is the first crossing by S of an interval (k_{1z}^n, k_{2z}^n) such that the following two conditions are satisfied:

- (t_{1c}^n, t_{2c}^n) is not a crossing by S of (k_{1a}^n, k_{2a}^n)
- the last crossing by S of an interval (k_{1z}^n, k_{2z}^n) before (t_{1c}^n, t_{2c}^n), is a negative crossing by S of (k_{1a}^n, k_{2a}^n)

Note that lemma 3.2.1 implies that the crossings by S of an interval (k_{1z}^n, k_{2z}^n) can be characterized as follows:
(s, t) is a crossing by S of an interval (k_{1z}^n, k_{2z}^n) iff (s, t) is a crossing by $R \circ S$ of $(0, 3n)$.
Applying the last characterization to the above conditions leads to: (t_{1c}^n, t_{2c}^n) is equal to the first crossing (t_{1i}^n, t_{2i}^n) by $R \circ S$ of $(0, 3n)$ with $i > 1$ such that the following two conditions hold:

- (t_{1i}^n, t_{2i}^n) is not a crossing by S of (k_{1a}^n, k_{2a}^n)
- $(t_{1(i-1)}^n, t_{2(i-1)}^n)$ a negative crossing by S of (k_{1a}^n, k_{2a}^n)

Which crossings are crossings by $R \circ S$ of $(0, 3n)$ is observable. That means that the crossings (t_{1i}^n, t_{2i}^n) are known to us. On the other hand, which crossings are crossings by S of (k_{1a}^n, k_{2a}^n) is not directly observable. However, (t_{11}^n, t_{21}^n) is observable and is a crossing by S of (k_{1a}^n, k_{2a}^n). So we can estimate if (t_{1i}^n, t_{2i}^n) is a crossing by S of (k_{1a}^n, k_{2a}^n) or not. For this we ask our localization test to compare the crossings (t_{11}^n, t_{21}^n) and (t_{1i}^n, t_{2i}^n). The

localization test can then estimate if the crossings $(t^n_{11}, t^n_{21)})$ and (t^n_{1i}, t^n_{2i}) of S occur on the same place or not. Our estimate for (t^n_{1c}, t^n_{2c}) will be defined to be the first (t^n_{1i}, t^n_{2i}) for which the above characterizing conditions are estimated to be true:

We define $(\hat{t}^n_{1c}, \hat{t}^n_{2c})$ to be equal to the first (t^n_{1i}, t^n_{2i}) with $i > 1$ for which the following three conditions hold:

- the localization test, when comparing (t^n_{11}, t^n_{21}) with (t^n_{1i}, t^n_{2i}), rejects the H_0-hypothesis.
- $t^n_{1(i-1)} > t^n_{2(i-1)}$
- the localization test, when comparing (t^n_{11}, t^n_{21}) with $(t^n_{1(i-1)}, t^n_{2(i-1)})$, accepts the H_0-hypothesis.

We define $\hat{w}^n_c$ to be the binary word associated with the crossing $(\hat{t}^n_{1c}, \hat{t}^n_{2c})$.

3.3 Assembling the pieces

The reconstruction algorithm at level n tries to reconstructs the finite piece of the scenery ξ:

$$\xi^n := \xi(k^n_{1c}), \xi(k^n_{1c} + u), \xi(k^n_{1c} + 2u), \ldots, \xi(k^n_{1a})$$

where $u := (k^n_{1a} - k^n_{1c})/|k^n_{1a} - k^n_{1c})|$. In this section, we explain how to construct a scenery $\bar{\xi} : \mathbb{Z} \longrightarrow \{0, 1\}$, equivalent to ξ from the collection of finite pieces: $\xi^1, \xi^2, \ldots$. The reconstruction algorithm at level n gives us the binary word ξ^n, but does not tell us where it is located in the scenery ξ. This implies that we need to "assemble" the pieces ξ^n in order to get $\bar{\xi}$.
Let as introduce a few definitions: let $v = v(0)v(1)v(2)\ldots v(i)$
and
$w = w(0)w(1)w(2)\ldots w(j)$ be two binary words. We say that v is contained in w iff there exist $j_1, j_2 \in \{0, 1, 2, \ldots, j\}$ such that v is equal to:

$$v = w(j_1)w(j_1 + u)w(j_1 + 2u)\ldots w(j_2) \tag{3.3.1}$$

, where $u := (j_2 - j_1)/|j_2 - j_1|$. We write $v \preccurlyeq w$ when v is contained in w. We say that v is uniquely contained in w and write $v \preccurlyeq_1 w$, iff there exists exactly one ordered pair (j_1, j_2) in $\{0, 1, 2, \ldots, j\}^2$ such that equation 3.3.1 is satisfied.
Note that the sequence of pieces $\xi^1, \xi^2, \ldots$ is an increasing sequence, in the sense that $\xi^n \preccurlyeq \xi^{n+1}$ for all $n \in \mathbb{N}$. (The reason for this being true is that by definition: $k_2^{n-} > k_2^{(n+1)-}$ and $k_2^{n+} < k_2^{(n+1)+}$ for all $n \in \mathbb{N}$. Thus the interval with the two endpoints k^n_{2c}, k^n_{2a} is contained in the interval with endpoints: k^m_{2c}, k^m_{2a} when $n < m$.) Imagine that not only , $\xi^n \preccurlyeq \xi^{n+1}$, but even $\xi^n \preccurlyeq_1 \xi^{n+1}$ for all $n \in \mathbb{N}$. Then there would be a unique way to assemble the pieces $\xi^1, \xi^2, \xi^3, \ldots$. The situation in this case, is similar to that of a puzzle: for a puzzle once we have decided of the position of one piece, there is a unique way to assemble the whole puzzle. Furthermore when we assemble a puzzle we always get the same image up to an isometric mapping. This is exactly the situation we encounter with the pieces of scenery when $\xi^n \preccurlyeq_1 \xi^{n+1}$ for all $n \in \mathbb{N}$.

Let us illustrate this with a practical example. Let $\xi : \mathbb{Z} \longrightarrow \{0,1\}$ be the scenery from which we show below a finite portion close to the origin:

$\xi(k)$	...	1	0	1	0	0	0	1	1	1	0	0	1	...
k	...	-4	-3	-2	-1	0	1	2	3	4	5	6	7	...

Assume that we would be given the three pieces (of the part of the scenery ξ which is represented above): 11000, 1000111 and 0100011100. In this case the first piece lies in the scenery ξ between the points 3 and -1. The second piece is the piece of ξ which lies between -1 and 4. The last piece lies between the points -3 and 6. We see that the first piece is uniquely contained in the second which itself is uniquely contained in the third piece. To assemble the three piece we first place the first piece anywhere in $\mathbb{Z}$. Then we place the second piece so that it covers the first piece, and that on the first piece it coincides with the first piece. Eventually we place the third piece so that it coincides with and covers the second one. If we place the first piece starting at the origin we get:

$\bar{\xi}(k)$					1	1	0	0	0			
k	-4	-3	-2	-1	0	1	2	3	4	5	6	7

Then we place the second piece so that it covers and coincides with the first piece. For this we have to turn the second piece around. We obtain:

$\bar{\xi}(k)$				1	1	1	0	0	0	1		
k	-4	-3	-2	-1	0	1	2	3	4	5	6	7

Eventually we place the third word and get:

$\bar{\xi}(k)$		0	0	1	1	1	0	0	0	1	0	
k	-4	-3	-2	-1	0	1	2	3	4	5	6	7

If we would go on with more and more pieces, as n tends to infinity we would obtain a scenery $\bar{\xi}$ which is equivalent to ξ.

Let E_0^n denote the event that:

$$E_0^n = \{\xi^n \preccurlyeq_1 \xi^{n+1}\}.$$

We will show that

$$\sum_{n=1}^{\infty} P(E_0^{nc}) < \infty,$$

where E_0^{nc} denotes the complement of E_0^n. From the last inequality above it follows that a.s. for all but a finite number of n's we have that $\xi^n \preccurlyeq_1 \xi^{n+1}$. The assemblage procedure we define below still works if $\xi^n \preccurlyeq_1 \xi^{n+1}$ holds for all but a finite number of n's.
Let us mention an additional problem: each reconstruction algorithm at level n has a small probability of making an error. Thus the output of the reconstruction algorithm at level n is not a.s. equal to ξ^n but is only an estimate of ξ^n. For the output of the reconstruction algorithm at level n, we will thus write $\hat{\xi}^n$ instead of ξ^n . We denote by E^n the event that the algorithm at level n works. That is:

$$E^n := \{\xi^n = \hat{\xi}^n\}.$$

By E^{nc} we denote the complementary event of E^n. In the next section it is shown that

$$\sum_{n=1}^{\infty} P(E^{nc}) < \infty. \tag{3.3.2}$$

From this it follows that almost surely all but a finite number of reconstructions $\hat{\xi}^n$ are correct, i.e. are such that $\xi^n = \hat{\xi}^n$. Our assembling procedure defined below is robust against this kind of problem: if only a finite number of pieces $\hat{\xi}^n$ are wrong it still works. Let us next define in a precise way our **assemblage procedure**:

Algorithm 3.3.1.

- Let $l^n + 1$ designate the length of the word $\hat{\xi}^n$ and let $\hat{\xi}^n(i)$ the i-th bit of the binary word $\hat{\xi}^n$. In this way:
$$\hat{\xi}^n = \hat{\xi}^n(0)\hat{\xi}^n(1)\hat{\xi}^n(2)\dots\hat{\xi}^n(l^n)$$

- Let n_0 designate the smallest natural (random) number such that for all $n \geq n_0$ we have that $\hat{\xi}^n \preccurlyeq_1 \hat{\xi}^{n+1}$ holds.

- We construct the scenery $\bar{\xi}$ by induction on n starting at n_0.
We first place the word $\hat{\xi}^{n_0}$ at the origin.
Once the word $\hat{\xi}^n$ is placed, we place the word $\hat{\xi}^{n+1}$ in the unique manner such that it covers and coincides with $\hat{\xi}^n$ on $\hat{\xi}^n$.
(d_1^n, d_2^n) designates the position of where we placed the word $\hat{\xi}^n$. More precisely :

  - Let $d_1^{n_0} := 0$ and let $d_2^{n_0} := l^{n_0}$. For all $k \in [0, d_2^{n_0}]$ define: $\bar{\xi}(k) := \hat{\xi}^{n_0}(k)$.
  - Once d_1^n, d_2^n are defined and $\bar{\xi}(k)$ is defined for all $k \in [d_1^n, d_2^n]$ let:
d_1^{n+1}, d_2^{n+1} with $d_1^{n+1} \leq d_2^{n+1}$ be the unique ordered pair of integers such that $[d_1^n, d_2^n] \subseteq [d_1^{n+1}, d_2^{n+1}]$ and such that one of the following two cases holds:

    1. For all $k \in [d_1^n, d_2^n]$ we have that:
$$\bar{\xi}(k) = \hat{\xi}^{n+1}(k - d_1^{n+1}).$$

    2. For all $k \in [d_1^n, d_2^n]$ we have that:
$$\bar{\xi}(k) = \hat{\xi}^{n+1}(l^{n+1} - (k - d_1^{n+1})).$$

  For all $k \in [d_1^{n+1}, d_2^{n+1}]$, let $\bar{\xi}(k)$ be equal to:

    1. When case 1 above holds:
$$\bar{\xi}(k) := \hat{\xi}^{n+1}(k - d_1^{n+1}).$$

    2. When case 2 above holds:
$$\bar{\xi}(k) := \hat{\xi}^{n+1}(l^{n+1} - k - d_1^{n+1}).$$

The constructed scenery $\bar{\xi}$ is equivalent to ξ as soon as for all but a finite number of n's we have that $\xi^n \preccurlyeq_1 \xi^{n+1}$ and $\xi^n = \hat{\xi}^n$. This should be obvious and we leave the proof to the reader. It thus only remains to prove that almost surely for all but a finite number of n's, $\xi^n \preccurlyeq_1 \xi^{n+1}$ and $\xi^n = \hat{\xi}^n$ hold.

3.4 Proof that the reconstruction at level n works

In this section we prove that the reconstruction algorithm at level n works with high probability, i.e. we prove equation 3.3.2. For this we decompose E^n into several elementary events. Let us start we some definitions:
We say that (s,r) satisfies the conditions of algorithm 3.2.1 with w_c^n instead of $\hat{w}_c^n$ iff $s<r$ and it satisfies all of the following conditions:

1. Same constraint as constraint 1 of algorithm 9.
2. There exists $s_2 \leq s_1 \leq r_1 \leq r_2$ with $s_2 = s, r_2 = r$ such that:
   (a) (s_1, s_2) is a crossing by $R \circ S$ of $(0, 3n)$ such that $w_{(s_1,s_2)} \geq w_c^n$ holds.
   (b) Same constraint as constraint 2.b of algorithm 9.

Let E_1^n designate the event that if algorithm 3.2.1 is given the real w_c^n instead of the estimate $\hat{w}_c^n$, it produces a straight crossing by S of (k_{2c}^n, k_{2a}^n).

$$E_1^n := \left\{ \begin{array}{l} \text{There exists at least one} \\ \text{pair (s,r), satisfying the} \\ \text{conditions of algorithm} \\ \text{3.2.1 with } w_c^n \text{ instead of} \\ \hat{w}_c^n. \end{array} \right\} \cap \left\{ \begin{array}{l} \text{Any pair } (s,r), \text{ minimizing} \\ r-s \text{ under the conditions} \\ \text{of algorithm 3.2.1 with } w_c^n \\ \text{instead of } \hat{w}_c^n, \text{ is a straight} \\ \text{crossing by } S \text{ of } (k_{2c}^n, k_{2a}^n). \end{array} \right\}$$

Let $E_{t\_c}^n$ be the event that the construction of (t_{c1}^n, t_{c2}^n) works:

$$E_{t\_c}^n := \left\{ (\hat{t}_{c1}^n, \hat{t}_{c2}^n) = (t_{c1}^n, t_{c2}^n) \right\}. \tag{3.4.1}$$

Note that when $E_{t\_c}^n$ holds, then $w_c^n = \hat{w}_c^n$.

$$E_{\text{all correct}}^n := \left\{ \begin{array}{l} \text{All } (s,r) \text{ satisfying the constraints of algorithm} \\ \text{3.2.1 with } w_c^n \text{ instead of } \hat{w}_c^n, \text{ are such that: } S(s) = \\ k_{2c}^n, S(r) = k_{2a}^n. \end{array} \right\}$$

$$E_{\text{at least one}}^n := \left\{ \begin{array}{l} \text{There exists } (s,r) \text{ satisfying the constraints of al-} \\ \text{gorithm 3.2.1 with } w_c^n \text{ instead of } \hat{w}_c^n, \text{ such that} \\ (s,r) \text{ is a straight crossing by } S \text{ of } (k_{2c}^n, k_{2a}^n). \end{array} \right\}$$

Let (t_{1ai}^n, t_{2ai}^n) be the i-th crossing by S of (k_{1a}^n, k_{2a}^n). Let E_{stopping}^n be the event that the stopping times $\tau^n(i)$ stop the random walk at k_{2a}^n:

$$E_{\text{stopping}}^n := \left\{ t_{2ai}^n = \tau^n(i), \forall i \leq \exp(n^{0.99}) \right\}$$

Let:

$$E_{\text{no other a crossing by R}}^n := \left\{ \begin{array}{l} \text{The only crossing } (k_1,k_2) \text{ by } R \text{ of } (0,3n) \text{ with} \\ |k_1 - k_{2a}^{\tilde{n}}|, |k_2 - k_{2a}^{\tilde{n}}| \leq n^{220} \text{ such that } w_{(k_1,k_2)}^R \geq \\ w_a^n \text{ is } (k_{1a}^n, k_{2a}^n). \end{array} \right\}$$

$$E_{\text{no other c crossing by R}}^n := \left\{ \begin{array}{l} \text{The only crossing } (k_1,k_2) \text{ by } R \text{ of } (0,3n) \text{ with} \\ |k_1 - k_{2a}^{\tilde{n}}|, |k_2 - k_{2a}^{\tilde{n}}| \leq n^{220} \text{ such that } w_{(k_1,k_2)}^R \geq \\ w_c^n \text{ is } (k_{1c}^n, k_{2c}^n). \end{array} \right\}$$

$$E^n_{\text{no other crossing by R}} := E^n_{\text{no other a crossing by R}} \cap E^n_{\text{no other c crossing by R}}$$

$$E^n_{\text{straight}} := \left\{ \begin{array}{l} \text{There exists } i \leq e^{\tilde{n}} \text{ and } s, r \text{ with } t^{\tilde{n}}_{2ai} \leq s, r \leq \\ t^{\tilde{n}}_{2ai} + n^{220} \text{ such that } (s,r) \text{ is a straight cross-} \\ \text{ing by } S \text{ of } (k^n_{1c}, k^n_{2a}). \end{array} \right\}$$

Let E^n_{visit} be the event that the random walk S visits the point k^n_{2c} before time $\exp(n^{0.5})$:

$$E^n_{\text{visit}} := \left\{ t^n_{2c} < \exp(n^{0.5}) \right\}$$

Recall that $\tilde{n} := n^{11}$. In subsection 3.4.1 we prove the following inclusions:

$$E^n_1 \cap E^n_{t\_c} \subseteq E^n \tag{3.4.2}$$

$$E^n_{\text{at least one}} \cap E^n_{\text{all correct}} \subseteq E^n_1 \tag{3.4.3}$$

$$E^{\tilde{n}}_{\text{stopping}} \cap E^n_{\text{no other crossing by R}} \subseteq E^n_{\text{all correct}} \tag{3.4.4}$$

$$E^n_{\text{straight}} \cap E^{\tilde{n}}_{\text{stopping}} \subseteq E^n_{\text{at least one}} \tag{3.4.5}$$

$$E^n_{\text{stopping}} \cap E^n_{\text{visit}} \subseteq E_{t\_c} \tag{3.4.6}$$

From the inclusions 3.4.2, 3.4.3, 3.4.4, 3.4.5 and 3.4.6 it follows that:

$$E^n_{\text{straight}} \cap E^{\tilde{n}}_{\text{stopping}} \cap E^n_{\text{stopping}} \cap E^n_{\text{no other crossing by R}} \cap E^n_{\text{visit}} \subseteq E^n$$

Which implies:

$$P(E^{nc}_{\text{straight}}) + P(E^{\tilde{n}c}_{\text{stopping}}) + P(E^{nc}_{\text{stopping}}) + \\ + P(E^{nc}_{\text{no other crossing by R}}) + P(E^{nc}_{\text{visit}}) \geq P(E^{nc}).$$

(Here $E^{nc}_{\text{something}}$ designates the complement of the event $E^n_{\text{something}}$). In subsection 3.4.2 we prove that:

$$P(E^{nc}_{\text{straight}}), P(E^{\tilde{n}c}_{\text{stopping}}), P(E^{nc}_{\text{stopping}}),$$
$$P(E^{nc}_{\text{no other crossing by R}}) \text{ and } P(E^{nc}_{\text{visit}})$$

are all finitely summable over n. Together with the last inequality, this proves that $P(E^{nc})$ is finitely summable over n.

3.4.1 Combinatorics

Proof that $E^n_1 \cap E^n_{t\_c} \subseteq E^n$ holds: When $E^n_{t\_c}$ holds, then $w^n_c = \hat{w}^n_c$. In this case, the event E^n_1 amounts to the same as event E^n. It follows that $E^n_1 \cap E^n_{t\_c} = E^n \cap E^n_{t\_c}$, which implies inclusion 3.4.2.

Proof that $E^n_{\text{at least one}} \cap E^n_{\text{all correct}} \subseteq E^n_1$ holds: Let (s,r) be a pair minimizing $r-s$ under the constraint of algorithm 3.2.1 with w^n_c instead of $\hat{w}^n_c$. Then if $E^n_{\text{all correct}}$ holds, we have that $S(s) = k^n_{2c}, S(r) = k^n_{2a}$. If $E^n_{\text{at least one}}$ also holds, there exists a straight crossing (s', r') by S of (k^n_{2c}, k^n_{2a}) satisfying the constraint of algorithm 3.2.1 with w^n_c instead of $\hat{w}^n_c$. For a straight crossing we have: $r' - s' = |k^n_{2c} - k^n_{2a}|$. Since $r-s$ is minimal under the constraint of algorithm 3.2.1, we get $|r-s| \leq |k^n_{2c} - k^n_{2a}|$. This together with $S(s) = k^n_{2c}, S(r) = k^n_{2a}$ is only possible if (s,r) is a straight crossing by S of (k^n_{2c}, k^n_{2a}). We just proved that when $E^n_{\text{at least one}}$ and $E^n_{\text{all correct}}$ hold, all pair (s,r) minimizing $r-s$ under the constraint of algorithm 3.2.1, is a straight crossing by S of (k^n_{2c}, k^n_{2a}). In this case E^n_1 holds. Thus, together $E^n_{\text{at least one}}$ and $E^n_{\text{all correct}}$ imply E^n_1.

Proof that $E^{\tilde{n}}_{\text{stopping}} \cap E^{n}_{\text{no other crossing by R}} \subseteq E^{n}_{\text{all correct}}$ holds: Let (s,r) satisfy all the constraints of algorithm 3.2.1. Then there exists $s_2 \leq s_1 \leq r_1 \leq r_2$ with $s_2 = s, r_2 = r$ where (r_1, r_2) is a crossing by $R \circ S$ of $(0, 3n)$ such that $w_{(r_1,r_2)} \geq w^n_a$ holds. By lemma 3.2.1 we have that there exists a crossing (k_1, k_2) by R of $(0, 3n)$ such that (r_1, r_2) is a crossing by S of (k_1, k_2). By lemma 3.2.2 and 3.2.3, we have that $w^R_{(k_1,k_2)} \geq w_{(r_1,r_2)}$. Thus, $w^R_{(k_1,k_2)} \geq w^n_a$.
Additional by the constraints of algorithm 3.2.1 there exists $i \leq e^{\bar{n}}$ such that $\tau^{\tilde{n}}(i) \leq s < r \leq \tau^{\tilde{n}}(i) + n^{220}$. If additionally $E^{\tilde{n}}_{\text{stopping}}$ holds, then $S(\tau^{\tilde{n}}(i)) = k^{\tilde{n}}_{2a}$. The random walk S during a time interval of n^{220} time can not walk further than n^{220}. Thus, $|S(r_1) - k^{\tilde{n}}_{2a}|, |S(r_2) - k^{\tilde{n}}_{2a}| \leq n^{220}$. This is equivalent to saying that: $|k_1 - k^{\tilde{n}}_{2a}|, |k_2 - k^{\tilde{n}}_{2a}| \leq n^{220}$. Hence the condition in event $E^{n}_{\text{no other crossing by R}}$ applies to the crossing (k_1, k_2). It follows that if $E^{n}_{\text{no other crossing by R}}$ also holds, then (k_1, k_2) equals (k^n_{1a}, k^n_{2a}). This implies that $S(r) = k^n_{2a}$. We have proven that when $E^{\tilde{n}}_{\text{stopping}}$ and $E^{n}_{\text{no other crossing by R}}$ both hold, then $S(r) = k^n_{2a}$. In a similar way, one can prove that in this case $S(s) = k^n_{2c}$. (We leave that proof to the reader.) Thus, $E^{\tilde{n}}_{\text{stopping}}$ and $E^{n}_{\text{no other crossing by R}}$ jointly imply $E^{n}_{\text{all correct}}$.

Proof that $E^{n}_{\text{straight}} \cap E^{\tilde{n}}_{\text{stopping}} \subseteq E_{\text{at least one}}$ holds: E^{n}_{straight} and $E^{\tilde{n}}_{\text{stopping}}$ jointly imply that there exists $i \leq e^{\bar{n}}$ and s, r with $\tau^{\tilde{n}}(i) \leq s, r \leq \tau^{\tilde{n}}(i) + n^{220}$ such that (s,r) is a straight crossing by S of (k^n_{2c}, k^n_{2a}). Thus, (s,r) already satisfies condition 1 of algorithm 3.2.1. It remains to show that (s,r) also satisfies condition 2. During the time interval (s,r), S crosses from the point k^n_{2c} to the point k^n_{2a} in a straight way. For this, S first needs to cross (k^n_{2c}, k^n_{1c}) in a straight manner and then (k^n_{1a}, k^n_{2a}). Thus, there exists $s_2 \leq s_1 \leq r_1 \leq r_2$ with $s_2 = s, r_2 = r$ such that (s_2, s_1) is a straight crossing by S of (k^n_{2c}, k^n_{1c}) and (r_1, r_2) is a straight crossing by S of (k^n_{1a}, k^n_{2a}). We know by lemma 3.2.1, that a crossing of a crossing is a crossing of the composition. Thus, (s_1, s_2) and (r_1, r_2) are both crossings by $R \circ S$ of $(0, 3n)$. Since the crossing (s_1, s_2) by S is straight, we have by lemma 3.2.2 and 3.2.3 that $w_{(s_1,s_2)} = w^R_{(k^n_{2c},k^n_{1c})}$. By lemma 3.2.2 and 3.2.3 again, we have that $w^R_{(k^n_{2c},k^n_{1c})} \geq w^n_c$. Thus, $w_{(s_1,s_2)} \geq w^n_c$. In a similar way one can show that $w_{(r_1,r_2)} \geq w^n_a$. This proves that (s,r) satisfies the conditions of algorithm 3.2.1 with w^n_c instead of $\hat{w}^n_c$. However, (s,r) is a straight crossing by S of (k^n_{2c}, k^n_{2a}). Thus, $E_{\text{at least one}}$ holds. We just proved that E^{n}_{straight} and $E^{\tilde{n}}_{\text{stopping}}$ together imply $E_{\text{at least one}}$.

Proof that $E^{n}_{\text{stopping}} \cap E^{n}_{\text{visit}} \subseteq E_{t\_c}$ holds: In subsection 3.2.6, we saw that (t^n_{1c}, t^n_{2c}) can be characterized as follows: (t^n_{1c}, t^n_{2c}) is equal to the first crossing (t^n_{1i}, t^n_{2i}) by $R \circ S$ of $(0, 3n)$ with $i > 1$ such that the following two conditions hold:

- (t^n_{1i}, t^n_{2i}) is not a crossing by S of (k^n_{1a}, k^n_{2a})
- $(t^n_{1(i-1)}, t^n_{2(i-1)})$ is a negative crossing by S of (k^n_{1a}, k^n_{2a})

The estimate $(\hat{t}^n_{c1}, \hat{t}^n_{c2})$ is defined to be the first crossing by $R \circ S$ of $(0, 3n)$ for which our localization test decides that the two conditions in the last characterization above hold. Thus, if up to time t^n_{c2}, the localization test gets all the crossings by S of (k^n_{1a}, k^n_{2a}) right, then the reconstruction of (t^n_{1c}, t^n_{2c}) works, i.e. E_{t_c} holds. The event E^{n}_{stopping} tells us that up to t^n_{2ai} with $i = \exp(n^{0.99})$ the localization test makes no errors in recognizing the crossings by S of (k^n_{1a}, k^n_{2a}). However, $\exp(n^{0.99}) \leq t^n_{2ai}$ for $i = \exp(n^{0.99})$, since each crossing lasts at least one time unit. Thus, up to time $\exp(n^{0.99})$ the localization test makes

no errors in recognizing the crossings by S of (k^n_{1a}, k^n_{2a}). However, if E^n_{visit} holds, then the random walk S visits the point k^n_{2c} before time $\exp(n^{0.5})$. Also, $\exp(n^{0.5}) \leq \exp(n^{0.99})$. In that case, before S visits the point k^n_{2c}, no errors occur. This proves that E^n_{stopping} and E^n_{visit} jointly imply E_{t_c}.

3.4.2 Probability bounds

High probability of E^n_{visit} We need a few definitions: Let E^n_2 be the event that the random walk S visits both points n^{10} and $-n^{10}$ before time $\exp(n^{0.5})$. Let:

$$E^n_{\text{k\_a,c}} := \left\{|k^n_{2a}|, |k^n_{2c}| \leq n^{10}\right\}$$

S first needs to visit $|k^n_{2a}|$ and $|k^n_{2c}|$ in order to visit both points n^{10} and $-n^{10}$, when $|k^n_{2a}|, |k^n_{2c}| \leq n^{10}$. (Since S starts at the origin.) Thus,

$$E^n_{\text{k\_a,c}} \cap E^n_2 \subseteq E^n_{\text{visit}}.$$

Thus,

$$P(E^{nc}_{\text{k\_a,c}}) + P(E^{nc}_2) \geq P(E^{nc}_{\text{visit}}).$$

If $P(E^{nc}_{\text{k\_a,c}})$ and $P(E^{nc}_2)$ are both finitely summable over n then $P(E^{nc}_{\text{visit}})$ is also. We prove that $P(E^{nc}_{\text{k\_a,c}})$ is finitely summable and leave the proof that $P(E^{nc}_2)$ is finitely summable to the reader since it is very similar to the other one. Let X^{R+}, resp. X^{R-} be the first passage time of the random walk $R(k)_{k\in\mathbb{N}}$, resp. of $R(-k)_{k\in\mathbb{N}}$ at the point $3n$. Let $E^n_{R+} := \{X^{R+} \leq n^{10}\}$ and let $E^n_{R-} := \{X^{R-} \leq n^{10}\}$. Then, $E^n_{R+} \cup E^n_{R-} = E^n_{\text{k\_a,c}}$. Thus, $P(E^{nc}_{R+}) + P(E^{nc}_{R-}) \geq P(E^{nc}_{\text{k\_a,c}})$. By symmetry, $P(E^{nc}_{R+}) = P(E^{nc}_{R-})$. Thus, $2P(E^{nc}_{R+}) \geq P(E^{nc}_{\text{k\_a,c}})$. Let Z_i denote the first passage time of $\{R(k)\}_{k\in\mathbb{N}}$ at the point i. Let $X_i := Z_i - Z_{i-1}$. Then, $X^{R+} := \sum_1^{3n} X_i$ and $P(E^{nc}_{R+}) = P(\sum_1^{3n} X_i > n^{10}) \leq P((\sum_1^{3n} X_i)^{1/3} > n^3)$. For positive numbers $a_1, a_2, \ldots, a_j$, we always have that $(\sum_{l=1}^j a_i)^3 \geq \sum_{l=1}^j (a_i)^3$. Thus $\sum_{i=1}^{3n}(X_i)^{1/3} \geq (\sum_{i=1}^{3n} X_i)^{1/3}$. It follows: $P(E^{nc}_{R+}) \leq P(\sum_{i=1}^{3n}(X_i)^{1/3} \geq n^3)$. By Chebychev, we get

$$P(E^{nc}_{R+}) \leq \frac{3E[(X_1)^{1/3}]}{n^2}$$

In [20] it is shown that $E[(X_i)^{1/3}]$ is finite. Thus, $P(E^{nc}_{R+})$ is finitely summable over n which finishes this proof.

High probability of E^n_{stopping} Let $E^n_3 := \{\forall i \leq \exp(n^{0.99}), t_{2ai} \leq \exp(n^{0.999})\}$. If up to time t_{2ai} with $i = \exp(n^{0.99})$ the localization test makes no mistake in identifying exactly all the crossings by $R \circ S$ of $(0, 3n)$ which occur in the same place, then E^n_{stopping} holds. Thus, if $t_{2ai} \leq \exp(n^{0.999})$ for $i = \exp(n^{0.99})$ and the localization test makes no mistake of this type up to time $\exp(n^{0.999})$, then E^n_{stopping} holds. Let $E^n_{\text{test correct}}$ be the event that for all $z_a, z_b \in \mathbb{Z}$ with $0 < |z_a|, |z_b| \leq n^{0.999}$ and for all $0 < i_a, i_b \leq n^{0.999}$ the localization test makes no error when comparing the crossings (t_{1a}, t_{2a}) and (t_{1b}, t_{2b}). (Here (t_{1a}, t_{2a}) and (t_{1b}, t_{2b}) are defined like in lemma 3.2.4: (t_{1a}, t_{2a}) is the i_a-th crossing by S of the z_a-th crossing by R of $(0, 3n)$ and (t_{1b}, t_{2b}) is the i_b-th crossing by S of the z_b-th crossing by R of $(0, 3n)$.) Up to time $\exp(n^{0.999})$, S can cross a crossing by R at most $\exp(n^{0.999})$ times. Thus, if (t_{1a}, t_{2a}) and (t_{1b}, t_{2b}) occur before time $\exp(n^{0.999})$, then $0 < i_a, i_b \leq n^{0.999}$. Furthermore, to reach the z-th crossing (k^n_{1z}, k^n_{2z}), S needs first to cross all the crossings

$(k^n_{1z'}, k^n_{2z'})$ with z' strictly between 0 and z. Thus up to time $\exp(n^{0.999})\}$ S can not reach any crossing (k^n_{1z}, k^n_{2z}) with $|z| > \exp(n^{0.999})\}$. If the crossings (t_{1a}, t_{2a}) and (t_{1b}, t_{2b}) occur before time $\exp(n^{0.999})\}$, we hence have that $0 < i_a, i_b \leq n^{0.999}$ and $0 < |z_a|, |z_b| \leq n^{0.999}$. Thus, E^n_3 and $E^n_{\text{test correct}}$ both hold, the localization test makes no mistake in identifying which of $(0, 3n)$ occur in the same place up to time t_{2ai}. In this case, E^n_{stopping} holds. Thus,

$$E^n_3 \cap E^n_{\text{test correct}} \subseteq E^n_{\text{stopping}}.$$

It follows that

$$P(E^{nc}_3) + P(E^{nc}_{\text{test correct}}) \geq P(E^{nc}_{\text{stopping}}).$$

If $P(E^{nc}_3)$ and $P(E^{nc}_{\text{test correct}})$ are both finitely summable over n, then $P(E^{nc}_{\text{stopping}})$ is also. The proof that $P(E^{nc}_3)$ is finitely summable is very similar to the proof for $P(E^{nc}_{\text{k\_a,c}})$, so we leave it to the reader. let: $E^n_{\text{test correct } i_a,i_b,z_a,z_b}$ be the event that the localization test recognizes correctly if with the crossings (t_{1a}, t_{2a}) and (t_{1b}, t_{2b}) we are in the H_0-case or not. By definition:

$$\bigcap E^n_{\text{test correct } i_a,i_b,z_a,z_b} = E^n_{\text{test correct}},$$

where the last intersection is taken over all i_a, i_b, z_a, z_b such that $0 < |z_a|, |z_b| \leq n^{0.999}$ and $0 < i_a, i_b \leq n^{0.999}$. Thus,

$$\Sigma P(E^{nc}_{\text{test correct } i_a,i_b,z_a,z_b}) \geq P(E^{nc}_{\text{test correct}}),$$

where the sum is taken over the same domain as before the union. There are $n^{3.996}$ quadruples (i_a, i_b, z_a, z_b) such that $0 < |z_a|, |z_b| \leq n^{0.999}$ and $0 < i_a, i_b \leq n^{0.999}$. By large deviation principle and lemma 3.2.4, the probability $P(E^{nc}_{\text{test correct } i_a,i_b,z_a,z_b})$ is exponentially small in n. Thus there exist $k > 0$ not depending on n or on (i_a, i_b, z_a, z_b) such that $P(E^{nc}_{\text{test correct } i_a,i_b,z_a,z_b}) \leq \exp(-kn)$. This implies that

$$P(E^{nc}_{\text{test correct}}) \leq n^{3.996} \cdot \exp(-kn).$$

Thus, $P(E^{nc}_{\text{test correct}})$ is finitely summable over n.

High probability of E^n_{straight} Let $\bar{t}^n_{2ai}$ denote the 20.000-th stopping time t^n_{2ai}. Thus, $\bar{t}^n_{2ai} := t^n_{2a(20.000 \cdot i)}$. Let E^n_4 be the event that there exists $i \leq n^{-20.000} \cdot e^{\bar{n}}$ and s, r with $\bar{t}^n_{2ai} \leq s, r \leq \bar{t}^n_{2ai} + n^{220}$ such that (s, r) is a straight crossing by S of (k^n_{1c}, k^n_{2a}). We have that $E^n_4 \subseteq E^n_{\text{straight}}$. Let $E^n_5 := E^n_{\text{k\_a,c}} \cap E^{\tilde{n}}_{\text{k\_a,c}}$. We find that the last inclusion implies:

$$P(E^{nc}_4 \cap E^n_5) + P(E^{nc}_5) \geq P(E^{nc}_{\text{straight}}).$$

We already saw that $P(E^{nc}_5)$ is finitely summable over n. So it only remains to be proven that $P(E^{nc}_4 \cap E^n_5)$ is finitely summable over n. Let X_i be the Bernoulli variable which is equal to one iff there exists s, r with $\bar{t}^n_{2ai} \leq s, r \leq \bar{t}^n_{2ai} + n^{220}$ such that (s, r) is a straight crossing by S of (k^n_{1c}, k^n_{2a}). By the Markov property of the random walk S, we have that conditional under $\sigma(R(k)|k \in \mathbb{Z})$ the variables X_i are i.i.d. Also, E^n_5 is $\sigma(R(k)|k \in \mathbb{Z})$-measurable. We are next going to evaluate the conditional probability: $P(X_1 = 1|R(k), k \in \mathbb{Z})$ when E^n_5 holds. When E^n_5 holds, then $|k^{\tilde{n}}_{2a} - k^n_{2c}| \leq 2\tilde{n}^{10}$. We have $2\tilde{n}^{10} := 2n^{110}$. By definition at any time $t^{\tilde{n}}_{2ai}$ the random walk S is at the point $k^{\tilde{n}}_{2a}$. By the local central limit theorem, when $|k^{\tilde{n}}_{2a} - k^n_{2c}| \leq 2\tilde{n}^{10}$, the probability that S goes from

$k^{\tilde{n}}_{2a}$ to k^n_{2c} in less than $(1/2)n^{220}$ steps is bigger than $k_2 \cdot n^{-110}$. (Here k_2 denotes a constant not depending on n and not depending on R as long as $R \in E^n_5$). Crossing in a straight way to the point k^n_{2a} right after the random walk S is at the point k^n_{2c}, has probability bigger than $(1/2)^{2n^{10}}$, when $|k^n_{2a} - k^n_{2c}| \leq 2n^{10}$. But, when E^n_5 holds, $|k^n_{2a} - k^n_{2c}| \leq 2n^{10}$. All this implies that when E^n_5 holds:

$$P\left(X_1 = 1 \mid R(k), k \in \mathbb{Z}\right) \geq (k_2 n^{-110})(1/2)^{2n^{10}}. \tag{3.4.7}$$

Let $\epsilon_1 := (k_2 n^{-110})(1/2)^{2n^{10}}$. Let $\hat{n} := n^{-20.000} \cdot e^{\bar{n}}$. Note that

$$E^{nc}_4 := \left\{ \sum_{i=1}^{\hat{n}} X_i = 0 \right\}$$

Conditional under $\sigma(R(k)|k \in \mathbb{Z})$ the X_i's are i.i.d. Thus,

$$P(E^{nc}_4 | R(k), k \in \mathbb{Z}) = (1 - P(X_i = 1 | R(k), k \in \mathbb{Z}))^{\hat{n}}.$$

Using inequality 3.4.7 we get for $R \in E^n_5$:

$$P(E^{nc}_4 | R(k), k \in \mathbb{Z}) \leq (1 - \epsilon_1)^{\hat{n}}. \tag{3.4.8}$$

When n goes to infinity, then ϵ_1 tends to zero. Thus, for n big enough we get:

$$(1 - \epsilon_1)^{1/\epsilon_1} \leq e^{-0.5}.$$

Applying this to inequality 3.4.8 leads in the case that $R \in E^n_5$, to:

$$P(E^{nc}_4 | R(k), k \in \mathbb{Z}) \leq e^{-0.5 \hat{n} \epsilon_1}.$$

Integrating the last inequality over E^n_5 leads to:

$$P(E^{nc}_4 \cap E^n_5) \leq e^{-0.5 \hat{n} \epsilon_1}. \tag{3.4.9}$$

Recall that $\bar{n} := n^{10.89}$ and $\tilde{n} := n^{11}$. In $\hat{n}$, the leading term is $e^{\bar{n}}$. In ϵ_1 the leading term is: $e^{\ln(0.5) \cdot 2n^{10}}$. Since $n^{10.89} \gg n^{10}$ we get that $e^{\bar{n}} \gg e^{-\ln(0.5) \cdot 2n^{10}}$. This implies that the leading term in $\hat{n}\epsilon_1$ is $e^{\bar{n}}$. Thus, the term on the right side of inequality 3.4.9 is finitely summable over n.

High probability of $E^n_{\text{no other crossing by R}}$ Let $n^* := n^{110} + n^{220}$. Let (t^n_{111}, t^n_{211}) designate the first crossing by S of (k^n_{11}, k^n_{21}). Let $w^n_{11} := w_{(t^n_{111}, t^n_{211})}$. Define:

$$E^n_{61} := \left\{ \begin{array}{l} \text{The only crossing } (k^n_{1z}, k^n_{2z}) \text{ with} \\ 0 < |z| \leq n^* \text{ such that } w^R_{(k^n_{1z}, k^n_{2z})} \geq \\ w^n_{11} \text{ is } (k^n_{11}, k^n_{21}). \end{array} \right\}$$

Let $(t^n_{1(-1)1}, t^n_{2(-1)1})$ designate the first crossing by S of $(k^n_{1(-1)}, k^n_{2(-1)})$. Let $w^n_{1(-1)} := w_{(t^n_{1(-1)1}, t^n_{2(-1)1})}$. Define:

$$E^n_{6(-1)} := \left\{ \begin{array}{l} \text{The only crossing } (k^n_{1z}, k^n_{2z}) \text{ with} \\ 0 < |z| \leq n^* \text{ such that } w^R_{(k^n_{1z}, k^n_{2z})} \geq \\ w^n_{1(-1)} \text{ is } (k^n_{1(-1)}, k^n_{2(-1)}). \end{array} \right\}$$

If $E^{\tilde{n}}_{\text{k\_a,c}}$ holds, then $|k^{\tilde{n}}_{2a}| \leq n^{110}$. All the crossings (k_1, k_2) concerned by the event $E^n_{\text{no other crossing by R}}$ are such that $|k_1 - k^{\tilde{n}}_{2a}|, |k_2 - k^{\tilde{n}}_{2a}| \leq n^{220}$. Thus, when $E^{\tilde{n}}_{\text{k\_a,c}}$ holds, then all the crossings concerned by $E^n_{\text{no other crossing by R}}$, are within n^* of the origin. When we write those crossings in the form (k^n_{1z}, k^n_{2z}) they must be such that $|z| \leq n^*$. Thus, when $E^{\tilde{n}}_{\text{k\_a,c}}$ holds, the events E^n_{61} and E^n_{62} cover all the crossings involved in the definition of the event $E^n_{\text{no other crossing by R}}$. One of the crossings (k^n_{1a}, k^n_{2a}) and (k^n_{1c}, k^n_{2c}) is equal to (k^n_{11}, k^n_{21}) whilst the other one is equal to $(k^n_{1(-1)}, k^n_{2(-1)})$. Similarly, one of the crossings (t^n_{1a}, t^n_{2a}) and (t^n_{1c}, t^n_{2c}) is equal to (t^n_{111}, t^n_{211}) whilst the other one is equal to $(t^n_{1(-1)1}, t^n_{2(-1)1})$. Eventually, one of the words w^n_a and w^n_c is equal to w^n_{11} whilst the other one is equal to $w^n_{1(-1)}$. This, implies that when $E^{\tilde{n}}_{\text{k\_a,c}}$ holds, the events E^n_{61} and E^n_{62} jointly imply $E^n_{\text{no other crossing by R}}$. Thus,

$$E^n_{61} \cap E^n_{62} \cap E^{\tilde{n}}_{\text{k\_a,c}} \subseteq E^n_{\text{no other crossing by R}}.$$

It follows that:

$$P(E^{nc}_{61}) + P(E^{nc}_{62}) + P(E^{\tilde{n}c}_{\text{k\_a,c}}) \geq P(E^{nc}_{\text{no other crossing by R}}).$$

We already saw that $P(E^{\tilde{n}c}_{\text{k\_a,c}})$ is finitely summable over n. By symmetry:$P(E^{nc}_{61}) = P(E^{nc}_{62})$. Thus, it only remains to prove that $P(E^{nc}_{61})$ is finitely summable over n. Let

$$E^n_{61z} := \left\{ w^R_{(k^n_{1z}, k^n_{2z})} \not\geq w^n_{11} \right\}.$$

We have:

$$E^n_{61} := \bigcap_{0<|z|\leq n^*, z\neq 1} E^n_{61z}.$$

It follows that:

$$P\left(E^{nc}_{61}\right) \leq \sum_{0<|z|\leq n^*, z\neq 1} P\left(E^{nc}_{61z}\right).$$

We saw in the proof of lemma 3.2.4 the distribution of $w^R_{(k^n_{1z}, k^n_{2z})}$ does not depend on z. Thus, the expression on the right side of the last inequality is equal to: $(2n^* - 2)P\left(E^{nc}_{612}\right)$. This yields:

$$P\left(E^{nc}_{61}\right)) \leq (2n^* - 2)P\left(E^{nc}_{612}\right) \tag{3.4.10}$$

We have that $E^{nc}_{612} = \{w^R_{(k^n_{12}, k^n_{22})} \geq w^n_{11}\}$. Hence:

$$E^{nc}_{612} = \bigcap_{m=0}^{n-1} \left\{ w^R_{(k^n_{12}, k^n_{22})}(m) \geq w^n_{11}(m) \right\}.$$

As in the proof of lemma 3.2.4, the bits of the word $w^R_{(k^n_{12}, k^n_{22})}$ are i.i.d. as well as the bits of w^n_{11} and $w^R_{(k^n_{12}, k^n_{22})}$ is independent of w^n_{11}. This gives:

$$P\left(E^{nc}_{612}\right) = \prod_{m=0}^{n-1} P\left(w^R_{(k^n_{12}, k^n_{22})}(m) \geq w^n_{11}(m) \right) = P\left(w^R_{(k^n_{12}, k^n_{22})}(1) \geq w^n_{11}(1) \right)^n.$$

The probability $q := (w^R_{(k^n_{12}, k^n_{22})}(1) \geq w^n_{11}(1))$ is strictly smaller than 1 and does not depend on n. Thus, the bound $(2n^* - 2)q^n$ on the left side of inequality 3.4.10 is finitely summable over n.

3.5 Why the reconstruction of ξ works.

Our reconstruction algorithm constructs a scenery $\bar{\xi}$. The main result of this paper is that a.s. $\bar{\xi}$ is equivalent to ξ. This is also what we need to prove in this section. The reconstruction algorithm we propose, constructs $\bar{\xi}$ by assembling (as explained in section 3.3) the finite reconstructed pieces $\hat{\xi}^n$. The piece $\hat{\xi}^n$ is provided by the reconstruction algorithm at level n. The reconstruction algorithm at level n tries to reconstructs the finite piece of the scenery ξ:

$$\xi^n := \xi(k^n_{1c}), \xi(k^n_{1c}+u), \xi(k^n_{1c}+2u), \ldots, \xi(k^n_{1a})$$

where $u := (k^n_{1a} - k^n_{1c})/|k^n_{1a} - k^n_{1c})|$. We have proven in the last section that $(1 - P(\xi^n = \hat{\xi}^n))$ is finitely summable over n. It follows that a.s. $\xi^n = \hat{\xi}^n$ for all but a finite number of n's. In section 3.3 we have seen that: the constructed scenery $\bar{\xi}$ is equivalent to ξ as soon as for all but a finite number of n's we have that $\xi^n \preccurlyeq_1 \xi^{n+1}$ and $\xi^n = \hat{\xi}^n$. It thus only remains to prove that a.s for all but a finite number of n's, $\xi^n \preccurlyeq_1 \xi^{n+1}$ holds. Define:

$$\xi^n_{\text{inside}} := \xi(-n), \xi(-n+1), \xi(-n+2) \ldots, \xi(n)$$

and

$$\xi^n_{\text{outside}} := \xi(-n^{10}), \xi(-n^{10}+1), \xi(-n^{10}+2) \ldots, \xi(n^{10}).$$

By definition, $|k^n_{2a}|, |k^n_{2c}| \geq n$ from which it follows that $\xi^n_{\text{inside}} \preccurlyeq \xi^n$. On the other hand, if $E^n_{k\_a,c}$ holds, then $|k^n_{2a}|, |k^n_{2c}| \leq n^{10}$ and $\xi^n \preccurlyeq \xi^n_{\text{outside}}$. Recall that $\xi^n \preccurlyeq \xi^{n+1}$ always holds by definition. Summing up: when $E^{n+1}_{k\_a,c}$ holds, we find that

$$\xi^n_{\text{inside}} \preccurlyeq \xi^n \preccurlyeq \xi^{n+1} \preccurlyeq \xi^{n+1}_{\text{outside}}.$$

Next, note that if $\zeta_a, \zeta_b, \zeta_c, \zeta_d \in \cup_{l\in\mathbb{N}}\{0,1\}^l$ with $\zeta_a \preccurlyeq \zeta_b \preccurlyeq \zeta_c \preccurlyeq \zeta_d$ and $\zeta_a \preccurlyeq_1 \zeta_d$ then also $\zeta_b \preccurlyeq_1 \zeta_c$. Thus, when $E^{n+1}_{k\_a,c}$ holds, if $\xi^n_{\text{inside}} \preccurlyeq_1 \xi^{n+1}_{\text{outside}}$ then also $\xi^n \preccurlyeq_1 \xi^{n+1}$. Let:

$$E^n_{\text{unique}} := \{\xi^n_{\text{inside}} \preccurlyeq_1 \xi^{n+1}_{\text{outside}}\}.$$

We have shown that

$$E^n_{\text{unique}} \cap E^{n+1}_{k\_a,c} \subseteq \{\xi^n \preccurlyeq_1 \xi^{n+1}\}.$$

For $(1 - P(\xi^n \preccurlyeq_1 \xi^{n+1}))$ to be finitely summable over n, it is thus enough that $P(E^{nc}_{\text{unique}})$ and $P(E^{c(n+1)}_{k\_a,c})$ both are. We have already proven that the probability of the complement $P(E^{c(n+1)}_{k\_a,c})$ is finitely summable over n. Remains to show that $P(E^{nc}_{\text{unique}})$ also is finitely summable. Let

$$E^n_{\text{unique, }+l} := \{\xi^n_{\text{inside}} \neq (\xi(l), \xi(l+1), \xi(l+2), \ldots, \xi(l+2n))\}$$

and

$$E^n_{\text{unique, }-l} := \{\xi^n_{\text{inside}} \neq (\xi(l), \xi(l-1), \xi(l-2), \ldots, \xi(l-2n))\}.$$

With this notation:

$$\bigcap_{l\neq -n,\ |l|\leq n^{10}} \left(E^n_{\text{unique, }+l} \cap E^n_{\text{unique, }-l}\right) \subseteq E^n_{\text{unique}}.$$

The last inclusion implies:

$$\sum_{l\neq -n,\ |l|\leq n^{10}} P\left(E^{nc}_{\text{unique, }+l}\right) + P\left(E^{nc}_{\text{unique, }-l}\right) \geq P\left(E^{nc}_{\text{unique}}\right) \tag{3.5.1}$$

Because the scenery ξ consists of i.i.d Bernoulli variables with parameter $1/2$, we find that $P(E^{nc}_{\text{unique, }+l}) = P(E^{nc}_{\text{unique, }-l}) = (1/2)^{2n}$. Furthermore, there are less than $2n^{10}$ elements in the set $\{l \neq -n,\ |l| \leq n^{10}\}$. This finishes to prove that the bound on the left side of inequality 3.5.1 if finitely summable over n.

Acknoledgement *I want to express my deepest gratitude to my thesis advisor Professor Harry Kesten, for providing such an interesting question and for all the time he spend helping me.*
I want to thank also Franz Merkl for always pushing me to finish writing up my articles in a readable way and for being a good friend.

References

[1] I. Benjamini and H. Kesten, *Distinguishing sceneries by observing the scenery along a random walk path*, J. Anal. Math. **69,** (1996), 97-135.

[2] K. Burdzy, *Some path properties of iterated Brownian motion*, Seminar on Stochastic processes 1992 K.L., Chung, E. Cinlar and M.J. sharpe, eds., Birkhäuser, Boston, 1993, 67-87.

[3] W.Th.F den Hollander, *Mixing Properties for random walk in random scenery*, Ann. Probab. **16** (1988), 1788-1802.

[4] C.D Howard, *Detecting defects in periodic scenery by random walks on* $\mathbb{Z}$, Random Structures Algorithms **8** (1996), no.1,59-74.

[5] C.D. Howard, *Orthogonality of measures induced by random walks with scenery*, Combin. Probab. Comput. **5** (1996), no. 3, 247-256.

[6] C.D. Howard, *Distinguishing certain random sceneries on* $\mathbb{Z}$ *via random walks*, Statis. Probab. Lett. 34 (1997), no.**2**, 123-132.

[7] S. A. Kalikov, T,T $^{-1}$ *transformation is not loosely Bernoulli*, Annals of Math, **115** (1982), 393-409.

[8] M. Keane and W.Th.F. den Hollander, *Ergodic properties of color records*, Physica 138A (1986), 183-193.

[9] H. Kesten and F. Spitzer, *A limit theorem related to a new class of self similar processes*, Z. Wahrsch. verw. Geb. **50** (1979), 5-25.

[10] H. Kesten, *Detecting a single defect in a scenery by observing the scenery along a random wlak path*, Ito's Stochastic Calculus and Probability theory, Springer, Tokyo, (1996), 171-183.

[11] H. Kesten, *Distinguishing and reconstructing sceneries from observations along a random walk path*, DIMACS; Series in Discrete Mathematics and Theoretical Computer Science, Vol. **41** (1998), 75-83, P. aldous and J. Propp, eds.

[12] E. Lindenstrauss, *Indistinguishable sceneries*, Random Struct. and Alg. **14**, (1999), 71-86.

[13] M. Löwe and H. Matzinger. Scenery reconstruction in two dimensions with many colors. *Ann. Appl. Probab.*, 12(4):1322–1347, 2002.

[14] H. Matzinger. Reconstructing a three-color scenery by observing it along a simple random walk path. *Random Structures Algorithms*, 15(2):196–207, 1999.

[15] H. Matzinger and S. W. W. Rolles. Reconstructing a random scenery observed with random errors along a random walk path. EURANDOM Report 2002-009. Accepted by Probability Theory and Related Fields.

[16] M. Löwe, H. Matzinger, and F. Merkl. Reconstructing a multicolor random scenery seen along a random walk path with bounded jumps. Eurandom Report 2001-030. Submitted., 2001.

[17] M. Löwe and H. Matzinger. Reconstruction of sceneries with correlated colors. Eurandom Report 99-032, accepted by Stochastic Processes and Their Applications, 1999.

[18] D. Levin and Y. Peres. Random walks in stochastic scenery on $\mathbb{Z}$. Preprint, 2002.

[19] D. A. Levin, R. Pemantle, and Y. Peres. A phase transition in random coin tossing. *Ann. Probab.*, 29(4):1637–1669, 2001.

[20] R. Durrett. *Probability: Theory and Examples.* Duxbury Press, Second edition, 1996.

[21] D. Heicklen, C. Hoffman, and D. J. Rudolph. Entropy and dyadic equivalence of random walks on a random scenery. *Adv. Math.*, 156(2):157–179, 2000.

[22] F. den Hollander and J. E. Steif. Mixing properties of the generalized T, T^{-1}-process. *J. Anal. Math.*, 72:165–202, 1997.

Chapter 4

Scenery Reconstruction in Two Dimensions with Many Colors

Ann. Appl. Probab., 12(4):1322–1347, 2002.

By Matthias Löwe, Heinrich Matzinger,

In [7] Kesten observed that the known reconstruction methods of random sceneries seem to strongly depend on the one dimensional setting of the problem and asked whether a construction still is possible in two dimensions. In this paper we answer the above question in the affirmative under the condition that the number of colors in the scenery is large enough. [1]

Short title: Scenery Reconstruction in Dimension 2

4.1 Introduction and the Main Result

The following problem has its roots in ergodic theory but may also be considered interesting in its own rights. Consider a graph (V, E) and color its vertices in an arbitrary way (so we do not only concentrate on proper colorings in the strict sense that any two adjacent vertices need to have a different color). This coloring will be called a scenery on (V, E). Then we run a random walk on (V, E) of which we only know the color record (i.e. the sequence of colors it reads at the vertices) but not where it actually reads them. The question then is: Can we still say anything about how V was colored?

This problem – which at first glance might seem a bit hopeless – was first investigated independently by Benjamini and den Hollander and Keane [5]. From here the problem splits into basically three branches:

[1] *MSC 2000 subject classification:* Primary 60K37, Secondary 60G10, 60J75.
Key words: Scenery reconstruction, jumps, stationary processes, random walk, ergodic theory.

1. Can we distinguish two (known) sceneries by their random walk record? or, more ambitiously:

2. Can we even reconstruct (unknown) sceneries by the observations we obtain from a random walk? and:

3. Are their sceneries which cannot be reconstructed or distinguished by the color record of a random walk?

Basic answers to all of these three question have been already given while other aspects are still wide open. For example Benjamini and Kesten [1] discovered the very strong result that almost surely any two given sceneries on the integer lattices $\mathbb{Z}$ or $\mathbb{Z}^2$ can be distinguished by a simple random walk on these lattices given that the colors are selected by an i.i.d. process. Previous to that Howard [6] had already been able to show that in one dimension a periodic scenery can be distinguished from a periodic scenery with one defect.

Matzinger [10] showed that on $\mathbb{Z}$ even more is true: Almost every i.i.d. two-color scenery can be reconstructed from the color record of a simple random walk (which even might have non zero probability to stand still). This implies Benjamini's and Kesten's result in one dimension as well as the earlier observation by the same author that the same holds true for three and more colors [9]. However, notice that Benjamini's and Kesten's techniques also work in a two dimensional situation or when the random walk is allowed to jump. A remarkable answer to Question 3 has been given by Lindenstrauss [8] who showed that there are still uncountably many sceneries on $\mathbb{Z}$ which cannot be distinguished from the color record of a simple random walk.

To be more specific: In what follows (V, E) will always be the integer lattice $\mathbb{Z}^2$ and a function $\xi : \mathbb{Z}^2 \to \mathbb{Z}$ will be called a two dimensional scenery. For a subset $D \subseteq \mathbb{Z}^2$ we call $\xi : D \to \mathbb{Z}$ a piece of scenery. If the range of ξ contains exactly m elements we will say that ξ has m colors or that it is an m-color scenery. Two sceneries ξ and $\overline{\xi}$ will be called equivalent, if there are $a \in \mathbb{Z}^2$ and

$$M \in \left\{ \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \right\}$$

such that

$$\xi(x) = \overline{\xi}(Mx + a) \qquad \forall x \in \mathbb{Z}^2.$$

Similarly, we call two pieces of scenery $\xi : D \to \mathbb{Z}$ and $\overline{\xi} : \overline{D} \to \mathbb{Z}$ equivalent, if again

$$\xi(x) = \overline{\xi}(Mx + a) \qquad \forall x \in D$$

holds true (a and M as above) and moreover $M(D) + a = \overline{D}$.

In other words ξ and $\overline{\xi}$ are equivalent (in symbols $\xi \sim \overline{\xi}$) if they can be obtained by translation and reflection on the coordinate axes from each other. It is rather obvious that in general we cannot expect to distinguish equivalent sceneries by their color record

and thus also reconstruction will work only up to equivalence. Throughout this paper we will consider ξ's that result from an unbiased i.i.d. random process with m colors (thus we will also say that ξ has m colors), that is the $\xi(v)$ are i.i.d. for all $v \in \mathbb{Z}^2$ and

$$P(\xi(0) = i) = \frac{1}{m}$$

for all colors $i \in \{0, \ldots, m-1\}$. Moreover, let $(S_k)_{k\in\mathbb{N}}$ be simple, symmetric random walk in two dimensions starting at the origin.

The main result of this paper states that if m is large enough the color record of (S_k), i.e.

$$\chi := (\chi(k))_{k\in\mathbb{N}} := (\xi(S_k))_{k\in\mathbb{N}}$$

contains enough information to reconstruct ξ almost surely up to equivalence. Additionally, we will present a well defined algorithm that given the scenery on a finite set reconstructs the whole scenery with probability larger than one half. In the next section we will see why this actually suffices to prove the main theorem. This, in a more mathematical way, is expressed in the following theorem, which states that with sufficiently many colors reconstruction of ξ from χ (up to equivalence) is possible with probability one.

Theorem 4.1.1. *There exists $m_0 \in \mathbb{N}$ such that if $m \geq m_0$, there exists a measurable function (with respect to the canonical σ-fields)*

$$\mathcal{A} : \{0, \ldots, m-1\}^{\mathbb{N}} \to \{0, \ldots, m-1\}^{\mathbb{Z}^2}$$

such that

$$P(\mathcal{A}(\chi) \sim \xi) = 1. \tag{4.1.1}$$

Here the measure P lives on the product space of the outcomes of ξ and the space of all random walk paths.

Remark 4.1.1. *We have not calculated any lower bound for m_0 yet. We are also convinced that the methods presented here will lead to an m_0 which is terribly large and far off any reasonable number and, in particular, any of the "borderline"-cases $m = 4, 5$ for which we have as many colors (or one more color, respectively) as we have neighbors in $\mathbb{Z}^2$ or even $m = 2$ (for which we doubt that Theorem 5.2.1 is valid). This is basically so, since we decided to keep the present proof as simple and transparent as possible and to use as many colors as necessary to this end. The specification of a good bound on m_0 will be subject to further research of the authors.*

This note has two further sections. In Section 2 we present the basic ideas of the algorithm used to reconstruct a random scenery while Section 3 contains the rigorous proof of Theorem 5.2.1.

Acknowledgment: This problem was posed by Harry Kesten. We benefited a lot from electronic discussions with him and private conversation with Mike Keane and Frank den Hollander. We are indebted to all of them for their interest in our work. We are also very grateful to an unknown referee who substantially helped to improve the presentation of this paper.

4.2 The Main Ideas and Basic Notations

The proof of Theorem 5.2.1 crucially bases on an induction argument. Given that we already know the scenery on a finite set A (for a special choice of A) we show how to extend this knowledge to the points sitting next to A. The following three lemmas are the building blocks of this induction. First we see that it suffices to exhibit an algorithm that reconstructs the scenery with probability larger than $1/2$ in order to be able to reconstruct the scenery almost surely.

Lemma 4.2.1. *For all $m \geq 2$ (where m designates the number of colors in ξ), if there exists a measurable map*

$$\overline{\mathcal{A}} : \{0, \ldots, m-1\}^{\mathbb{N}} \to \{0, \ldots, m-1\}^{\mathbb{Z}^2}$$

such that

$$P(\overline{\mathcal{A}}(\chi) \sim \xi) > 1/2$$

then there also exists a measurable

$$\mathcal{A} : \{0, \ldots, m-1\}^{\mathbb{N}} \to \{0, \ldots, m-1\}^{\mathbb{Z}^2}$$

with

$$P(\mathcal{A}(\chi) \sim \xi) = 1.$$

The proof of Lemma 4.2.1 will be given in Section 3.

Lemma 4.2.1 will be useful, since we will soon see that with sufficiently many colors we are able to reconstruct with large probability the scenery on finite regions of $\mathbb{Z}^2$ such as the integer circle of radius n denoted by

$$B^n := \{x \in \mathbb{Z}^2 : ||x|| \leq n\}.$$

Here $||\cdot||$ stands for the standard Euclidean norm in $\mathbb{Z}^2$. Moreover, in the following we will frequently use the following notation: we will write $f|B$ for the restriction of f to a subset B of the domain of definition of f, for example $\xi|B$ will be a piece of scenery (that is the scenery restricted to some subset B of $\mathbb{Z}^2$), while $\chi|B$ will be a part of the observations (here B will be a subset of $\mathbb{N}$).

The next two lemmas will basically contain the induction. Lemma 4.2.2 below is the start of the induction, while Lemma 4.2.3 contains the induction step. So, first we show that we can reconstruct $\xi|B^n$ for each finite n with arbitrary large probability, as long as the scenery contains sufficiently many colors.

Lemma 4.2.2. *Let $n \in \mathbb{N}$ and $\varepsilon > 0$. Then there exists $m_1 \in \mathbb{N}$ such that if $m \geq m_1$ there exists a measurable function*

$$\mathcal{A}^n : \{0, \ldots, m-1\}^{\mathbb{N}} \to \{0, \ldots, m-1\}^{B^n}$$

such that

$$P(\mathcal{A}^n(\chi) \sim \xi|B^n) \geq 1 - \varepsilon.$$

Also Lemma 4.2.2 will be proven in the next section.

The next lemma is the induction step in the sense that it states that we can reconstruct $\xi|B^{n+1}$ with large probability provided we know $\xi|B^n$ up to equivalence and the number of colors is large enough.

Lemma 4.2.3. *There exists $m_2 \in \mathbb{N}$ (random) such that for $m \geq m_2$ there is a sequence of measurable functions $(\tilde{\mathcal{A}}^n)_{n\in\mathbb{N}}$,*

$$\tilde{\mathcal{A}}^n : \bigcup_{a\in\mathbb{Z}^2} \{0,\dots,m-1\}^{B^n+a} \times \{0,\dots,m-1\}^{\mathbb{N}} \to \{0,\dots,m-1\}^{B^{n+1}}$$

such that P-a.s.

$$\tilde{\mathcal{A}}^n(\xi|B^n, \chi) \sim \xi|B^{n+1}$$

occurs for all but finitely many n.

Remark 4.2.1. *Remark that given that m is large enough the critical n in Lemma 4.2.3 from which the algorithms work, i.e. from which*

$$\tilde{\mathcal{A}}^n(\xi|B^n, \chi) \sim \xi|B^{n+1}$$

is random.

Also note that Lemma 4.2.3 implies that for each $\varepsilon > 0$ we can find a number N (non-random) such that the probability that all $\tilde{\mathcal{A}}^n$ work for all $n \geq N$ is bigger than $1-\varepsilon$.

Roughly speaking, Lemma 4.2.3 means that the algorithm obtained by concatenating the different $\tilde{\mathcal{A}}^n$'s works well, in the sense that given $\xi|B^n$ up to equivalence and the observations χ it almost surely fails to reconstruct $\xi|B^{n+1}$ only for finitely many n.

To explain the proof of the induction step, which is crucial to the whole proof of Theorem 1.1, observe that the main difficulty in the reconstruction of sceneries is, of course, that we do not exactly know where the random walk precisely is. This is even more a problem in two dimensions than it is in one dimension as the random walk in one dimensions by time N has returned to the origin about $\sqrt{N}$ times, and therefore produces a lot of information about the neighborhood of the origin. In two dimensions the local time of the origin at time N is only about $\log N$. Thus we have to find an accurate method for guessing when the random walk is close to the origin from the observations χ it produces. This will be achieved by using a set of signal words, i.e. sequences of subsequent colors in B^n. Their frequent appearance in the observations will indicate that we really are in a neighborhood of B^n.

This "guessing that the random walk is inside B^n" is the first step of the reconstruction algorithm. More accurately, these words which will indicate that we are inside B^n (the so called *signal words*) are horizontal, non-overlapping words inside B^n of length proportional to $\log n$. The set of these words will be called $\mathcal{S}^n$. Whenever we read more than n^β words during a time interval of length n^2 whose endpoint is inside $[0, e^{n^\alpha}]$ (α and β some numbers to be specified later), we will "guess" that the walk is inside B^{n^2+n}. The union of these time intervals will be called τ^n and the reconstruction will only take place during τ^n. Note that τ^n designates a random set.

More formally in the sequel let $c_1, c_2, c_3 > 0$ be positive constants (not depending on n) which we will specify later. For convenience we will assume that $c_i \log n \in \mathbb{N}$ for each $i = 1, 2, 3$ (which of course means the c_i slightly depend on n but this dependence is irrelevant). Let

$$\begin{aligned}\mathcal{S}^n &:= \{w = (w_1, \ldots, w_{c_1 \log n}) \Big| \exists k \in \mathbb{Z} \text{ and } (x, y) \in \mathbb{Z}^2 : x = k c_1 \log n \\ &\qquad (x + s, y) \in B^n \text{ and } w_s = \xi((x + s, y)), \quad \forall 0 \le s \le c_1 \log n - 1\}.\end{aligned}$$

In other words $\mathcal{S}^n$ "partitions" $\xi | B^n$ into disjoint horizontal words of length $c_1 \log n$.

Moreover let $1 < \alpha < \beta < 2$ be two real numbers close to two to be specified later,

$$\begin{aligned}\mathcal{I}_{\alpha,\beta} &:= \{I = [t, t + n^2] \big| t \le e^{n^\alpha} - n^2, \\ &\qquad \chi | I \text{ contains more than } n^\beta \text{ different words from } \mathcal{S}^n\},\end{aligned}$$

and

$$\tau^n := \tau^n_{\alpha,\beta} := \bigcup_{I \in \mathcal{I}_{\alpha,\beta}} I.$$

As sketched above, the point is that during the times $k \in \tau^n$ we can be pretty sure that the random walk is "close to B^n", more precisely that it is inside B^{n+n^2}. This will ensure that the reconstruction takes place at the boundary of B^n and not anywhere else.

As a matter of fact, the probability for the random walk to go right through a given signal word is equal to $(1/4)^{c_1 \log n}$. Thus for c_1 very small the random walk when being inside B^n typically reads $n^{2-\varepsilon_1}$ signal words during a time interval of length n^2. Here $\varepsilon_1 > 0$ can be made arbitrarily small. This is basically so, because the random walk typically visits about $n^2 / \log n$ distinct points in a time window of length n^2, and thus during these time steps it would roughly visit about $n^2 / \log n \times (1/4)^{c_1 \log n} \ge n^{2-\varepsilon_1}$ (for c_1 small enough) signal words.

Now, if the number of colors m is large enough we can choose c_1 small and still the signal words will be typical of B^n (that is, the probability to read them in a given ball $B^{n^2}_y$ – the ball of radius n^2 centered in y – is small, as long as the ball does not touch B^n). Indeed, there are less than $\pi n^4 4^{c_1 \log n}$ different paths of length $c_1 \log n$ inside $B^{n^2}_y$. Thus by independence the probability for a given signal word to appear in $B^{n^2}_y \setminus B^n$ is less than $\pi n^4 (4/m)^{c_1 \log n}$, which is as small as we want to, if only m is large enough. As a matter of fact, exploiting the independence of the signals in a large deviations argument we will be able to show, that up to time e^{n^α} the random walk in a time interval of length n^2 will only be able to read more than n^β (α, β as above) signal words if it spends this time in B^{n^2+n} and that the probability of reading so many signals elsewhere is about e^{-n^α}. So, our test, to check when we are back in B^n will not fail until time roughly e^{n^α}. But by that time we will have returned to the origin about $n^{\alpha-\varepsilon_2}$ times ($\varepsilon_2 > 0$, small). If now m were so large that there were only different colors inside B^n this would suffice to reconstruct ξ on the boundary of B^n. We simply would have to follow the walk until it exits B^n and read the first color outside as the color of a boundary point. If all colors were different, we would clearly know where this boundary point was. Moreover, there are order n points in ∂B^n, so $n^{1+\varepsilon_3}$ ($\varepsilon_3 > 0$) returns to the origin would suffice to reconstruct the scenery on the boundary of B^n. As we already saw that we have about n^α of such returns, we would be done.

However, we are not allowed to choose m growing with n, so we cannot assume that all colors inside B^n are different. So we have to employ more subtle methods to reconstruct ξ on the boundary of B^n.

To describe this reconstruction part we have to introduce some more notations. Let

$$\underline{\partial} B^n := \{z \in B^n | \, \exists y \in \mathbb{Z}^2 \setminus B^n \text{ such that } z \text{ and } y \text{ are neighbors } \}$$

be the inner boundary of B^n and

$$\overline{\partial} B^n := B^{n+1} \setminus B^n$$

be its outer boundary. Observe that $\overline{\partial} B^n$ may differ from the outer boundary of B^n in the lattice topology. Indeed, there might points at distance one from B^n without nearest neighbours in B^n. Moreover – using the lattice geometry of $\mathbb{Z}^2$ – it is easily checked that all points in B^{n+1} can be reached from a point in B^n by crossing at most two edges. Since by definition $B^n \cup \overline{\partial} B^n = B^{n+1}$ it clearly suffices to reconstruct $\overline{\partial} B^n$ with large enough probability.

The strategy will be to guess the color of a point v in $\overline{\partial} B^n$ by extending a walk to a neighboring point in $\underline{\partial} B^n$ by two further steps. Of course, we have to be very careful of both, to walk to $v \in \underline{\partial} B^n$ and to extend the walk into the right direction.

The principal idea behind this reconstruction can be described quite easily. Draw a straight (horizontal or vertical) line through v and suppose we knew already the the colors of a line segment of length approximately $\log n$ *inside* B^n and containing v as well as the colors of a line segment of about the same length *outside* B^n at distance 2 from v. Then we could figure out the two missing colors between these two segments by just waiting until the random walk first reads the colors of the segment inside B^n (in the right order) and then after a waiting time of 2 the colors of the segment outside B^n. Except, if the walk is far away from v (which we can exclude by the above arguments) the walk must have followed the straight line supporting the two segments at least partially and thus the missing two colors are the colors read between reading the colors of the two segments. Indeed, the "following partially" part above needs a little more technical work. In fact we could deviate from the above line segment and just accidentally read the right colors. We will get rid of this nuisance by characterising the missing two points as the shortest distance between two cones rather than between two line segments. This idea will be made more precise below.

Now a major difficulty is that we do not know the colors outside B^n. Thus we have to think of another characterisation of the segment outside B^n (supported by the same line as the inner segment). It will turn out that it is useful to think of it as the segment whose colors can be read in shorter time by starting with the inner segment than by starting with any segment parallel to it.

To formalise this idea for $v \in \underline{\partial} B^n$ we define a segment $\sigma(v)$ (the segment associated with v) in the following way: Let $\sigma(v)$ be the horizontal or vertical segment of length $(c_2 + c_3) \log n$ with endpoints v and $\sigma_0(v) \in B^n$, such that the angle between this segment and the tangent to the circle of radius $|v|$ centered in 0 in the point v is at least 45 degrees (the latter is needed to ensure that the objects below are well defined).

The first $c_2 \log n$ lattice points (starting from $\sigma_0(v)$) will be called the root segment of v and abbreviated by $\hat{\sigma}(v)$, the rest of $\sigma(v)$ is called second root segment and will be

denoted by the symbol $\overline{\sigma}(v)$, while the left and right neighboring segments of $\hat{\sigma}(v)$ of the same length $c_2 \log n$ as $\hat{\sigma}(v)$ (or the lower and upper segment next to the root segment of v, if $\sigma(v)$ is a horizontal segment, respectively) are named the side segments of v. For these we reserve the symbols $\lambda(v)$ and $\rho(v)$, and their starting points (next to $\sigma_0(v)$) are denoted by $\lambda_0(v)$ and $\rho_0(v)$, respectively. Finally the segment of length $c_2 \log n$ following $\sigma(v)$ after one step when we keep following the line supporting $\sigma(v)$ will be called the invisible segment associated with v and denoted by $\varphi(v)$. Its endpoints are called v_2 and $\varphi_0(v)$. The words associated with these segments will be called the root word, second root word, side words, and invisible words, respectively. Finally the lattice points we want to guess the color of, that is the points on $\varphi(v)$ of distance one and two to v are named v_1 and v_2.

All this is illustrated in Figure 1 below.

Let us now describe how this reconstruction works.

The idea behind the above setup is that in order to read the color of v_1 and v_2 we take a neighboring vertex $v \in \underline{\partial} B^n$ and read the color of v_1 and v_2 as the next colors when we have read $\sigma(v)$ from $\sigma_0(v)$ to v. To guarantee that indeed we read the color of the right points we require that the algorithm picks a word w of length $c_2 \log n$ satisfying the following conditions

1. w appears in $\chi|\tau^n$ directly (one step) after the word supported by $\sigma(v)$.

2. In $\chi|\tau^n$ the shortest time for w to appear after the root word of v is exactly equal to $c_3 \log n + 1$.

3. In $\chi|\tau^n$ the shortest time for w to appear after the side word of v is exactly $c_3 \log n + 2$.

Condition 2 assures that we do not run backwards after having read the word supported by $\sigma(v)$ while Condition 3 guarantees that we have not deviated from the segment from $\sigma_0(v)$ to v while reading the scenery.

Thus we estimate $\xi(v_2)$ to be the first color of w. The estimate for $\xi(v_1)$ will be the the color between $\sigma(v)$ and w, when they appear in $\chi|[0, e^{n^\alpha}]$ one step apart from each other. If there is no word w satisfying the above conditions we let the algorithm terminate (our conditions imply that this will happen only with extremely small probability).

To realize this idea, that is to actually prove Theorem 5.2.1, we need some more definitions, which we will give now. For $v \in \underline{\partial} B^n$ the half space associated to v – which will be denoted by $\mathcal{H}(v)$ – is the half space separating $\hat{\sigma}(v)$ from $\overline{\sigma}(v)$ orthogonal to $\sigma(v)$ and with $\overline{\sigma}(v)$ in $\mathcal{H}(v)$. The first quart-space $Q_1(v)$ associated with v will be the right-angular cone based in v_2 with bisecting line along $\varphi(v)$ such that the major part of $\varphi(v)$ is inside this cone. The second quart-space $Q_2(v)$ associated with v is the right-angular cone based on the line separating $\mathcal{H}(v)$ from its complement such that $\hat{\sigma}(v)$ is on its bisecting line and $\hat{\sigma}(v)$ is in this cone. The third quart-space $Q_3(v)$ associated with v will be defined as the right-angular cone based on the line separating $\mathcal{H}(v)$ from its compliment such that $\lambda(v)$ is on its bisecting line and $\lambda(v)$ is in this cone. Finally, the fourth quart-space $Q_4(v)$ associated with v will be the right-angular cone based on the line separating $\mathcal{H}(v)$ from its compliment such that $\rho(v)$ is on its bisecting line and $\rho(v)$ is in this cone. The base points of Q_3, Q_2 and Q_4, respectively, are denoted by a, b, and c, respectively.

All this is illustrated in Figure 1. In this figure the points $v, a, b, c, \lambda_0(v), \sigma_0(v)$, and $\rho_0(v)$ are inside B^n, whilst v_1, v_2, and $\varphi_0(v)$ are outside B^n.

As can be seen from there

- $Q_1(v)$ contains the segment $\varphi(v)$ which begins with v_2 and ends in $\varphi_0(v)$
- $Q_2(v)$ contains the segment $\hat{\sigma}(v)$ which begins with $\sigma_0(v)$ and ends in b
- $Q_3(v)$ contains the segment $\lambda(v)$ which begins with $\lambda_0(v)$ and ends in a
- $Q_4(v)$ contains the segment $\rho(v)$ which begins with $\rho_0(v)$ and ends in c
- $\sigma(v)$ consists of $\hat{\sigma}(v)$ and $\overline{\sigma}(v)$
- All of the segments $\varphi(v)$,$\hat{\sigma}(v)$,$\rho(v)$, and $\lambda(v)$ contain $c_2 \log n$ lattice points, while $\overline{\sigma}(v)$ contains $c_3 \log n$ lattice points.

4.3 Proofs

In this section we give the proofs of Theorem 5.2.1 and Lemma 4.2.1, Lemma 4.2.2, and Lemma 4.2.3. Let us start with the proof of Lemma 4.2.1.

Proof of Lemma 4.2.1: Let $X(l)$ be the indicator for the event that the reconstruction algorithm $\overline{\mathcal{A}}$ applied to the observations shifted by l give rise to a scenery which is equivalent to the actual scenery, that is $X(l) = 1$ if $\overline{\mathcal{A}}(\Theta^l(\chi)) \sim \xi$ and $X(l) = 0$ otherwise. Obviously, $(X(l), l \in \mathbb{N})$ is stationary with

$$\mathbb{P}(X(l) = 1) = \mathbb{P}(\overline{\mathcal{A}}(\chi) \sim \xi) > \frac{1}{2}$$

for all l.

Furthermore let

$$\Omega = \{(+1,0),(-1,0),(0,+1),(0,-1)\}^{\mathbb{N}} \times \{0,\ldots,m-1\}^{\mathbb{Z}^2}$$

and let $\mathcal{F}$ be the standard σ-field on Ω. Let $\theta : \Omega \to \Omega$ be defined in the following way. For any

$$\omega = \left((\overline{\Delta}_1, \overline{\Delta}_2, \ldots), \psi\right)$$

where

$$\psi \in \{0,\ldots,m-1\}^{\mathbb{Z}^2}$$

and

$$\overline{\Delta}_i \in \{(+1,0),(-1,0),(0,+1),(0,-1)\} \quad \text{for all} i \in \mathbb{N}$$

we define

$$\theta(\omega) := \left((\overline{\Delta}_2, \overline{\Delta}_3, \ldots), \psi + \overline{\Delta}_1\right).$$

Here $\psi + \overline{\Delta}_1$ stands for $2D$ scenery ψ shifted by $-\overline{\Delta}_1$, i.e.

$$\psi + \overline{\Delta}_1(z) := \psi(\overline{\Delta}_1 + z).$$

Let Δ_i designate the i'th increment of the random walk S, i.e.

$$\Delta_i := S(i) - S(i-1).$$

Let μ be the measure describing the randomness of the object $((\Delta_1, \Delta_2, \ldots), \xi)$. This means $(\Omega, \mathcal{F}, \mu)$ is a probability space. One easily verifies that θ is measure preserving on $(\Omega, \mathcal{F}, \mu)$. Let $Z(l)$ designate the random vector

$$Z(l) = ((\Delta_{l+1}, \Delta_{l+2}, \ldots), \xi + S(l)).$$

Note that $Z(l) = \theta^l(Z(1))$. Since θ is measure preserving the sequence $Z(0), Z(1), Z(2), \ldots$ is measure preserving. Now $X(0), X(1), X(2), \ldots$ is a stationary coding of the sequence $Z(0), Z(1), Z(2), \ldots$. By this we mean, that there exists a measurable function F such that for all $l \in \mathbb{N}$ we have

$$F(Z(l)) = X(l).$$

This implies stationarity of the sequence $X(0), X(1), X(2), \ldots$. Now a stationary coding of an ergodic sequence is ergodic again. Thus in order to prove that $(X(l))_l$ is ergodic we will prove that $Z(0), Z(1), Z(2), \ldots$ is ergodic. To do so we will show that $Z(0), Z(1), Z(2), \ldots$ is actually mixing. For this it is enough to see that for any two $A, B \in \mathcal{F}$ that only depend on finitely many Δ_i we have

$$\lim_{k\to\infty} \mu(\theta^{-k} A \cap B) = \mu(A)\mu(B).$$

Let σ^n denote the σ-algebra

$$\sigma^n = \sigma(\Delta_1, \Delta_2, \ldots, \Delta_n, \xi(z) : z \in B_n\}$$

where

$$B_n := \{z \in \mathbb{Z}^2 : |z| \leq n\}.$$

Eventually let $C_{n,k}$ denote the event that

$$C_{n,k} := \{S(k) \notin B_{2n}\}.$$

Assume that $A, B \in \sigma^n$. Then, conditional on $C_{n,k}$ the events $\theta^{-k}(A)$ and B are independent. Also note that $\theta^{-k}(A)$ and $C_{n,k}$ are independent. Thus we obtain

$$\begin{aligned} \mu(\theta^{-k}(A) \cap B | C_{n,k}) &= \mu(\theta^{-k}(A)|C_{n,k})\mu(B|C_{n,k}) \\ &= \mu(\theta^{-k}(A))\mu(B|C_{n,k}). \end{aligned}$$

Hence

$$\mu(\theta^{-k}(A) \cap B \cap C_{n,k}) = \mu(\theta^{-k}(A))\mu(B \cap C_{n,k}).$$

This implies that

$$\mu(\theta^{-k}(A) \cap B) = \mu(\theta^{-k}(A) \cap B \cap C_{n,k}^c) + \mu(A)(\mu(B) - \mu(B \cap C_{n,k}^c)). \tag{4.3.1}$$

Keeping n fixed and taking k to infinity we obtain

$$\lim_{k\to\infty} \mu(C_{n,k}^c) = 0.$$

Hence also

$$\lim_{k\to\infty} \mu(B \cap C_{n,k}^c) = \lim_{k\to\infty} \mu(\theta^{-k}(A) \cap B \cap C_{n,k}^c) = 0.$$

Thus (4.3.1) implies

$$\mu(\theta^{-k}(A) \cap B) = \mu(A)\mu(B).$$

Hence the shift θ is mixing on $(\Omega, \mathcal{F}, \mu)$ and thus also ergodic. Therefore $Z(0), Z(1), \ldots$ is an ergodic sequence of random variables. Since $X(0), X(1), \ldots$ is a stationary coding of $Z(0), Z(1), \ldots$ it inherits the property of ergodicity.

Hence by the ergodic theorem

$$\frac{X(1) + X(2) + \ldots + X(l)}{l}$$

converges to a limit larger than $1/2$ almost surely. Thus under the assumption that

$$\mathbb{P}(\overline{\mathcal{A}}(\chi) \sim \xi) > \frac{1}{2}.$$

we can identify the equivalence class of ξ as the only equivalence class which eventually is equivalent to the majority of the $\overline{\mathcal{A}}(\Theta^l(\chi))$'s.

□

Let us now prove Lemma 4.2.2.

Proof of Lemma 4.2.2: The principal idea behind the proof of Lemma 4.2.2 is that with enough colors within a large area a certain color is typical of the point underlying it. This will help us to reconstruct the scenery on two basic shapes, which will help to reconstruct the scenery on the points of a three by three square and hence also on any other square. In a final step we will see this already suffices to reconstruct the scenery within a large ball.

To be more precise let

$$E_{01}^n := \bigcap_{x \neq y \in B^n} \{\xi(x) \neq \xi(y)\},$$

and

$$E_{02}^n := \bigcap_{\substack{x,y \in B^n, \\ ||x-y||=1}} \quad \bigcap_{z \in B^n} \quad \bigcap_{\substack{v \notin B^n: \\ \xi(v)=\xi(z)}} \{(S_k)_k \text{ passes from } x \text{ to } y \text{ in one step before visiting } v\}$$

In words the event E_{01}^n says that all colors inside B^n are different, while E_{02}^n states that all edges inside B^n are crossed by $(S_k)_{k\in\mathbb{N}}$ before it visits a point outside B^n having the same color as one of the points inside B^n.

We now show that under the condition that E_{01}^n and E_{02}^n hold true, we can reconstruct the scenery $\xi|B^n$. The reconstruction will be based on the following two important cases.

Case I: Let $x, y, z, v \in B^n$ be the corners of a unit square with x and z (and as well y and v) across the diagonal. Then, if E_{01}^n and E_{02}^n hold, and we know the colors of x, y and z, we can figure out the color of v. As a matter of fact, the color of v is the first color appearing, neighboring both the color of x and the color of z, and different from the

color of y. (Here and in the following we call two colors neighboring if they are read at consecutive times).

Case II: Let $x_1, x_2, x_3, x_4, y \in B^n$ be a "cross" with center y , that is x_1, x_2, x_3, x_4, y are pairwise different and

$$|x_1 - y| = |x_2 - y| = |x_3 - y| = |x_4 - y| = 1.$$

Knowing that E_{01}^n and E_{02}^n hold as well as the colors of x_1, x_2, x_3 and y we can find out the color of x_4 as the only color neighboring $\xi(y)$ different from $\xi(x_1), \xi(x_2)$, and $\xi(x_3)$.

We will now see that these two basic techniques suffice to reconstruct $\xi|B^n$, if E_{01}^n and E_{02}^n hold. Indeed, denoting by Q_j the $2j+1$ by $2j+1$ square with center zero, we can first reconstruct $\xi|Q_1$.

To this end we first recover the color of the origin (which is, of course, trivial) and the colors of $(1,0), (0,1), (-1,0)$, and $(0,-1)$. Indeed, the colors themselves are known from the observations. Note that the only information we need is the relative positions of the colors of $(1,0), (0,1), (-1,0)$, and $(0,-1)$ to each other because we only want to reconstruct up to equivalence. This means we only need to know which of the colors

$$\{\xi((1,0)), \xi((0,1)), \xi((-1,0)), \xi((0,-1))\}$$

are from points across $(0,0)$ and which of them are not. (Here we say that $(1,0)$ and $(-1,0)$ lie across $(0,0)$ as well as $(0,1)$ and $(0,-1)$ lie across $(0,0)$, while the other possible pairs don't.)

Now the following characterization holds:
Pairs from

$$\{\xi((1,0)), \xi((0,1)), \xi((-1,0)), \xi((0,-1))\}$$

lie across $(0,0)$ *if and only if they have exactly one neighbouring color (which is* $\xi(0,0)$*), whilst the other pairs have exactly two neighboring colors.*

Once we know the $\xi|\{(1,0), (0,1), (-1,0), (0,-1)\}$ up to equivalence we can reconstruct the scenery on Q_1 by applying Case I to the four corner points of Q_1.

Now we can proceed inductively. Knowing the $\xi|Q_j \cap B^n$, we want to reconstruct $\xi|Q_{j+1} \cap B^n$, that is we want to find out the color of the boundary points of Q_{j+1} (as far as they are inside B^n). For all points with at least one coordinate different from $j+1, j, -j-1$, or $-j$, this can be done by applying the technique of Case II. Then the color of the points with one coordinate equal to j or $-j$ can be reconstructed by applying the technique of Case I. Finally the same technique yields the color of the corner points of Q_{j+1}.

This shows that under the condition that E_{01}^n and E_{02}^n hold true we can reconstruct $\xi|B^n$ up to equivalence. It remains to understand that both, E_{01}^n and E_{02}^n hold true with arbitrary large probability for fixed n and large enough m. Indeed, this is not very hard to see. For E_{01}^n, note that

$$\mathbb{P}\big((E_{01}^n)^c\big) \leq \text{const}\, n^2 \frac{1}{m},$$

which can be made arbitrarily small by choosing m large.

Similar techniques apply to E_{02}^n. Note that by taking T large enough the random walk $(S_k)_{k \leq T}$ up to time T has visited each point in B^n, at least L times (L some number to

be chosen soon, cf. [11] for similar results). Then the probability that there is an edge in B^n the random walk does not visit up to time T is bounded by

$$\text{const}\, n^2 \left(\frac{3}{4}\right)^L,$$

which is arbitrarily small for L large enough. If we now first choose L, the take T as above, and finally choose m so large that also the probability that all colors in B^T are distinct (by the same techniques as above) is as large as we want to, we see that

$$\mathbb{P}\big((E_{02}^n)^c\big) \leq \varepsilon$$

for each $\varepsilon > 0$ if only m is large enough. This finishes the proof of Lemma 4.2.2. □

Next we will prove Lemma 4.2.3, which is indeed the key ingredient to the proof of Theorem 5.2.1.

Proof of Lemma 4.2.3: Let E^n denote the event that given a piece of scenery ψ with $\psi \sim \xi|B^n$ the "reconstruction algorithm at step n" $\overline{\mathcal{A}}^n$ produces a piece of scenery $\overline{\mathcal{A}}^n(\psi, \chi)$ with

$$\overline{\mathcal{A}}^n(\psi, \chi) \sim \xi|B^{n+1}.$$

We need to show that with probability one E^n holds for all but a finite number of n's (in the following we will also say that an event holds for almost all n if it holds for all n but a finitely many).

To do so see we decompose E^n for $n \in \mathbb{N}$ in such a way that

$$E^n \supset E_1^n \cap E_2^n \cap E_3^n.$$

We will then show that each of E_i^n, $i = 1, 2, 3$ holds for all but finitely many n's.

Whenever in the sequel we will say about some observations κ that "κ appears in A with starting point x", or "κ appears in A with endpoint y", respectively, where $\kappa \in \{0, \ldots, m-1\}^l$ for some l, $A \subseteq \mathbb{Z}^2$, and $x, y \in \mathbb{Z}^2$, we will mean that

$$\chi|T = \kappa$$

for some realisation of the random walk S_n, some discrete time interval $T = [t_0, t_0 + l - 1]$ such that $S_{t_0} = x$ (or $S_{t_0+l-1} = y$, respectively) and $S|T \subseteq A$. In other words κ appears in A with starting point x (or endpoint y) if it can be read inside of A by a nearest neighbor walk starting in x (ending in y). Moreover if, for one of the line segments $\sigma(v), \hat{\sigma}(v), \overline{\sigma}(v), \varphi(v)$ or $\lambda(v)$, we refer to $\xi|\mathcal{L}$ ($\mathcal{L} \in \{\sigma(v), \hat{\sigma}(v), \overline{\sigma}(v), \varphi(v), \lambda(v)\}$) we mean the observations obtained by reading ξ along $\mathcal{L}$ from the center of B^n to the outside of B^n

Now let

$$E_1^n := \bigcap_{x \in B^{\exp(n^\alpha)}} \{\text{There are less than } n^\beta \text{ different words from } \mathcal{S}^n \text{ appearing in } \xi|(B_x^{n^2} \setminus B^n)\},$$

where $B_x^{n^2}$ stands for the discrete ball of radius n^2 centered in x.

Observe that the definition of τ^n implies that on E_1^n we have that $S_k \in B^{n+n^2}$ for all $k \in \tau^n$.

Moreover let

$$E_2^n = E_{21}^n \cap E_{22}^n \cap E_{23}^n \cap E_{24}^n \cap E_{25}^n$$

with

$$E_{21}^n := \bigcap_{v \in \underline{\partial} B^n} \{\xi | \overline{\sigma}(v) \text{ appears in } \xi | B^{n^2+n} \text{ only with end point inside } \mathcal{H}(v)\},$$

$$E_{22}^n := \bigcap_{v \in \underline{\partial} B^n} \{\xi | \hat{\sigma}(v) \text{ appears in } \xi | B^{n^2+n} \text{ only with endpoint } x \in Q_2(v)\},$$

$$E_{23}^n := \bigcap_{v \in \underline{\partial} B^n} \{\xi | \lambda(v) \text{ appears in } \xi | B^{n^2+n} \text{ only with endpoint } x \in Q_3(v)\},$$

$$E_{24}^n := \bigcap_{v \in \underline{\partial} B^n} \{\xi | \rho(v) \text{ appears in } \xi | B^{n^2+n} \text{ only with endpoint } x \in Q_4(v)\},$$

and

$$E_{25}^n := \bigcap_{v \in \underline{\partial} B^n} \{\xi | \varphi(v) \text{ appears in } \xi | B^{n^2+n} \text{ only with starting point } x \in Q_1(v)\}.$$

Finally let

$$E_3^n = E_{3,\sigma}^n \cap E_{3,\lambda}^n \cap E_{3,\rho}^n,$$

where

$$E_{3,\sigma}^n := \bigcap_{v \in \underline{\partial} B^n} \{ \text{ All nearest neighbor walks of length } (2c_2 + c_3) \log n + 1 \text{ initially traversing } \hat{\sigma}(v) \text{ are realized at least once during } \tau^n\},$$

$$E_{3,\lambda}^n := \bigcap_{v \in \underline{\partial} B^n} \{ \text{ All nearest neighbor walks of length } (2c_2 + c_3) \log n + 1 \text{ initially traversing } \lambda(v) \text{ are realized at least once during } \tau^n\},$$

and

$$E_{3,\rho}^n := \bigcap_{v \in \underline{\partial} B^n} \{ \text{ All nearest neighbor walks of length } (2c_2 + c_3) \log n + 1 \text{ initially traversing } \rho(v) \text{ are realized at least once during } \tau^n\}.$$

Before we show that $E_1^n \cap E_2^n \cap E_3^n$ indeed happens for all but a finite number of n's, let us see that this will actually imply the desired result, that is, let us see, that

$$E^n \supset E_1^n \cap E_2^n \cap E_3^n.$$

As a matter of fact, for each event in E_1^n we know that during τ^n we must be close to B^n, more precisely, we know, that during τ^n the walk is inside B^{n^2+n}. Then E_3^n ensures that in this time τ^n we read each sequence of length $(2c_2 + c_3) \log n + 1$ beginning with

either $\xi|\hat{\sigma}(v)$, $\xi|\rho(v)$, or $\xi|\lambda(v)$ for each $v \in \underline{\partial} B^n$ at least once. E_2^n now guarantees that during these times the walk is close to the points a, b, and c (of the appropriate v). Finally E_2^n together with E_3^n ensures some of the walks actually pass the points a, b, and c, correspondingly. Therefore, we are able to read the color of the vertices v_1 and v_2 next to v in direction of $\sigma(v)$.

Let us explain this in detail, since this step is, indeed, the core of the reconstruction step. For fixed $v \in \underline{\partial} B^n$ at the boundary of B^n we need to prove that the reconstruction method works correctly, i.e. that the algorithms we will give below reveals the colors of the corresponding v_1 and v_2 (i.e. $\xi(v_1)$ and $\xi(v_2)$) correctly, if E_1^n, E_2^n and E_3^n hold. Let us now define the reconstruction algorithm properly.

The algorithm is given as input $\xi|B^n$, the scenery restricted to B^n which we assume to know already.

Algorithm to reconstruct $v_1 = v_1(v)$ and $v_2 = v_2(v)$

Step 1: Select all words w of length $c_3 \log n$ in $\xi \circ S|\tau^n$ with the following properties:

a) The shortest number of steps w appears after $\xi|\hat{\sigma}$ in $\xi \circ S|\tau^n$ is $c_3 \log n + 1$.

b) The shortest number of steps w appears after $\xi|\lambda$ in $\xi \circ S|\tau^n$ is $c_3 \log n + 2$.

c) The shortest number of steps w appears after $\xi|\rho$ in $\xi \circ S|\tau^n$ is $c_3 \log n + 2$

d) A word of the form $\xi|\sigma \diamond \upsilon \diamond w$ (where υ is an arbitrary color and the symbol $\diamond$ stands for the concatenation of two words) occurs in $\xi|\tau^n$, i.e. the event that w is read precisely one step after $\xi|\sigma$ occurs in τ^n.

Step 2: Take the first letter of w as an estimator of $\xi(v_2)$.

Step 3: Take an occurrence of a word $\xi|\sigma \diamond \upsilon \diamond w$ in τ^n. Estimate $\xi(v_1)$ with υ.

In order to prove that the above algorithm works and is well defined (Step 3) given that E_1^n, E_2^n and E_3^n hold, we will prove the following: for every word w selected by our algorithm its first letter is read at position v_2. This automatically implies both: that Step 2 of the above algorithm works as well as that Step 3 is well defined and works.

First assume that there is at least one word w selected by the first step of the above algorithm. Call the lattice point at which the first letter of w is read x. Assume that x is in $\mathcal{H}$ but not on the line supporting σ. Then there is a path form either a or c to x which is strictly shorter than any path from any starting point in Q_2 to x (in particular it is shorter than a path from b to x). Now E_3^n holds, so in particular $E_{3,\lambda}^n$ and $E_{3,\rho}^n$ hold. Hence a path first reading λ (or rho), crossing a (or c) and then walking to x in the shortest possible way in order to produce w from there will once be realized. Now E_1^n holds, ensuring that during τ^n the random walk is in B^{n^2+n}. Thus E_{22}^n holds. This guarantees that any time we read $\xi|\hat{\sigma}$ in χ we do this with an endpoint in Q_2. Thus any time we read the word w (and still we assume that x is not on the line supporting σ) a time t' after having read $\xi|\hat{\sigma}$, this time t' will be strictly larger than the time to read w after having read one of $\xi|\lambda$ or $\xi|\rho$. This contradicts our selection criteria.

Thus we can only select words w with a first letter read at $x \in \mathcal{H}$, if x lies on the line supporting σ. Now from $E_{3,\sigma}^n$ we know that all paths of length $(2c_2 + c_3) \log n + 1$ are realized once during τ^n. From this together with the fact that we have selected w such that the *shortest* it appears in the observations after $\xi|\hat{\sigma}$ is $c_3 \log n + 1$, it follows that x is at distance $c_3 \log n + 1$ from b, hence given that $x \in \mathcal{H}$, we conclude that $x = v_2$. It only remains to show that x cannot be in $\mathcal{H}^c$. But this is guaranteed by E_{21}^n.

It remains to show that Step 1 of the above algorithm selects at least one word. But as a consequence of E_1^n, E_{22}^n, E_{23}^n, E_{24}^n, E_{25}^n and $E_{3,\sigma}^n$ Step 1 of the above algorithm will select $\xi|\varphi$. Indeed, the shortest path from Q_3 or Q_4 to Q_1 is $c_3 \log n + 2$ steps long, while Q_1 can be reached from Q_2 in $c_3 \log n + 1$ steps. By E_1^n we know that we are in B^{n^2+n} during τ^n. Thus by E_{23}^n, E_{24}^n, and E_{25}^n we know that the shortest possibility to read $\xi|\varphi$ after $\xi|\rho$ or $\xi|\lambda$ is after $c_3 \log n + 2$ steps, while the shortest possibility to read $\xi|\varphi$ after $\xi|\hat{\sigma}$ is after $c_3 \log n + 1$ step. Finally, $E_{3,\sigma}^n$ ensures that we will observe at least once the sequence $\xi|\sigma\diamond\upsilon\diamond\xi|\varphi$ for some color υ. Thus $\xi|\varphi$ satisfies the selection criteria of Step 1 of the algorithm.

Hence we reconstruct the color of $v_1(v)$ and $v_2(v)$ if $E_1^n \cap E_2^n \cap E_3^n$ is satisfied.

As this works for all $v \in \underline{\partial} B^n$ we are indeed able to reconstruct the scenery on B^{n+1} proving that

$$E^n \supset E_1^n \cap E_2^n \cap E_3^n.$$

It remains to show that $E_1^n \cap E_2^n \cap E_3^n$ is true for all but finitely many n, if we choose α and β in the correct manner.

E_1^n holds for all but finitely many n: Let $\omega \in \mathcal{S}^n$ be any fixed signal word in B^n. By this we mean that ω is the signal word between two fixed starting points; so note that ω although being fixed in this sense, will still be random. Let $y \notin B^n$ be any potential starting point for ω outside B^n. By independence of the colors

$$\mathbb{P}(\omega \text{ appears in } \xi|(\mathbb{Z}^2 \setminus B^n) \text{ with starting point } y) \leq \left(\frac{4}{m}\right)^{c_1 \log n}$$

as there are $4^{c_1 \log n}$ different walks of length $c_1 \log n$ starting in y. Thus for any y

$$\mathbb{P}(\omega \text{ appears in } \xi|(B_y^{n^2} \setminus B^n)) \leq \pi n^4 \left(\frac{4}{m}\right)^{c_1 \log n} = \pi n^{4+c_1(\log 4 - \log m)}$$

as there are πn^4 different points inside B^{n^2}.

Now the indicators I_w for the event that the word $w \in \mathcal{S}^n$ appears in $B_y^{n^2} \setminus B^n$ are conditionally independent (for different w) under $\mathbb{P}$ given $\xi|(B_y^{n^2} \setminus B^n)$ as the different words have mutually disjoint support and therefore are independent. To understand this point correctly it is of importance to recall that $\mathcal{S}^n$ is a random set (under $\mathbb{P}$). The independence claimed above would not be true for any fixed set of words or if we did not condition on knowing $\xi|(B_y^{n^2} \setminus B^n)$.

Hence the number of $w \in \mathcal{S}^n$ appearing in $B_y^{n^2} \setminus B^n$ is stochastically bounded by a Binomial random variable with $N = n^2/c_1 \log n$ different trials and success probability $p = \pi n^{2+c_1(\log 4 - \log m)}$. But for n, m large enough

$$\frac{n^2}{c_1 \log n} \leq n^2$$

as well as

$$p = \pi n^{2+c_1(\log 4 - \log m)} \leq \frac{1}{n^2}.$$

But then the number of $w \in \mathcal{S}^n$ appearing in $B_y^{n^2} \setminus B^n$ is stochastically bounded by a Binomial random variable X with n^2 different trials and success probability $\frac{1}{n^2}$ But by Tschebyschev's exponential inequality

$$\mathbb{P}\left(X \geq n^\beta\right) \leq e^{-n^\beta}\mathbb{E}e^X = e^{-n^\beta}\left(1 + \frac{e-1}{n^2}\right)^{n^2} = \mathcal{O}(e^{-n^\beta}).$$

It follows that

$$\mathbb{P}[(E_1^n)^c] = \mathcal{O}(e^{2n^\alpha - n^\beta})$$

which is ummable for $\beta > \alpha$. This by the Borel–Cantelli Lemma implies that E_1^n holds for all but finitely many n.

E_2^n holds for all but a finite number of n: Since the proofs of that E_{2i}^n holds for almost all n are very similar for each i, we just show it for E_{22}^n and leave the other proofs to the reader.

To this end consider any $v \in \underline{\partial} B^n$ and any oriented connected segment s in $\mathbb{Z}^2$ of length $c_2 \log n$. Note that if the endpoint of s is not in Q_2 the i'th point of $\hat{\sigma}(v)$ is different from all the j'th points of s, $j \leq i$, and thus $\xi(\hat{\sigma}_i(v))$ is a "fresh random variable. Thus by conditional independence the probability of reading $\hat{\sigma}(v)$ along $\xi|s$ is bounded by

$$\mathbb{P}(\xi|s = \xi|\hat{\sigma}(v)) = \left(\frac{1}{m}\right)^{c_2 \log n},$$

and therefore for every fixed $x \in B^{n^2+n} \setminus Q_2(v)$

$$\mathbb{P}(\xi|\hat{\sigma}(v) \text{ appears with endpoint } x) \leq \left(\frac{4}{m}\right)^{c_2 \log n}.$$

As there are at most $\pi(n^2+n)^2$ points in B^{n^2+n} and there are at most $const \times n$ points $v \in \underline{\partial} B^n$, we obtain

$$\mathbb{P}((E_{22}^n)^c) \leq \left(\frac{4}{m}\right)^{c_2 \log n} const \times n\pi(n^2+n)^2 \leq n^{c_2(\log 4 - \log m)+6}$$

The right hand side of this inequality becomes summable if we choose m large enough (depending on c_2 or c_3). More precisely, we choose m such that

$$c_2(\log 4 - \log m) + 6 < -2.$$

Note that this choice does not depend on n. This choice of m will basically be the proof of Theorem 5.2.1. Thus (again by a Borel–Cantelli argument) E_{22}^n holds true for all but finitely many n.

Note that until now we are still free to choose c_1, c_2, c_3.

E_3^n holds for all but finitely many n: Again we only give the proof in detail for one of the events, which will be $E_{3,\sigma}^n$. The proof for the other two events follows the same lines.

We split this proof into several parts.

First let us prove that in a certain (stricter than usual) sense the random walk by time e^{n^α} has returned to the origin more than n^γ times, where $\gamma < \alpha < \beta$. A result like this seems to be very much in the spirit of a result by Erdös and Taylor [2], who showed that almost surely a random walk at time e^n has returned to the origin between $n/(\log n)^{1+\varepsilon}$ and $(1+\varepsilon)n \log\log n$ times for all but finitely many n's and every positive $\varepsilon > 0$. The reason why we cannot simply refer to this result is that we also want these returns to the origin to be apart at least n^2 from each other. So, more precisely let us introduce a sequence ϑ_i^n of stopping times such that $\vartheta_0^n = 0$ for all n and ϑ_{i+1}^n is the time of the first return of the random walk S_k to the origin after time $\vartheta_i^n + n^2$. This will ensure that in the meantime the random walk is able to hit one of the boundary points of B^n. So we want to check that for $\gamma < \alpha < \beta$ (γ appropriately chosen afterwards) the event

$$E_{31}^n := \{\vartheta_{n^\gamma}^n \le e^{n^\alpha}\}$$

happens for all but finitely many n's. Indeed, choosing $\delta = \frac{\alpha-\gamma}{2}$ the result by Erdös and Taylor [2] quoted above states that the event

$$E_{311}^n := \{\text{ Up to time } e^{n^\alpha} \text{ there are more than } n^{\gamma+\delta} \text{ returns to the origin }\}$$

holds true almost surely for all but a finite number of n's. Next we will show that the same is true for the event

$$E_{312}^n := \bigcap_{i=1}^{n^\gamma} \{\text{In the interval } [\vartheta_i^n, \vartheta_i^n + n^2] \text{ there are less than } n^\delta \text{ returns to the origin }\}.$$

As a matter of fact the probability for a simple random walk starting at the origin not returning for t steps is bounded below by $\frac{2\pi}{\log t}$ for t large enough [12, p.167], [3]. Applying this yields

$$\begin{aligned} &\mathbb{P}(\text{In the interval } [\vartheta_i^n, \vartheta_i^n + n^2] \text{ there are more than } n^\delta \text{ returns to the origin }) \\ \le\ &\left(1 - \frac{\pi}{\log n}\right)^{n^\delta} \le e^{-n^{\delta/2}} \end{aligned}$$

for each $i = 1, \ldots, n^\gamma$ and n large enough. Hence by bounding the probability of a union by the sum of the probabilities

$$\mathbb{P}((E_{312}^n)^c) \le n^\gamma e^{-n^{\delta/2}}$$

which is finitely summable. Therefore E_{312}^n holds for all but a finite number of n's. As E_{311}^n and E_{312}^n together imply E_{31}^n also E_{31}^n holds for almost all n.

Next we will show that many of the intervals $[\vartheta_i^n, \vartheta_i^n + n^2]$ above are indeed signal times, that is we will show that we read more than n^β different signals in all of these time intervals. To this end introduce random variables Y_i which are indicators for the event that the interval $[\vartheta_i^n, \vartheta_i^n + n^2]$ is a signal time, that is for the event that there are more than n^β signal words read in $[\vartheta_i^n, \vartheta_i^n + n^2]$. To avoid the dependence among reading different signal words we only concentrate on such words which are "far apart" form each other. To this end we partition the inner part of B^n, that is $B^n \setminus \partial_{(\log n)^3} B^n$ where

$$\partial_{(\log n)^3} B^n := \{x \in B^n, d(x, \underline{\partial} B^n) \le (\log n)^3\}$$

and $d(\cdot,\cdot)$ is the lattice distance in $\mathbb{Z}^2$, into boxes of lengths $c_1 \log n$ and $(\log n)^3$. Let

$$W^n_{k,l} := \{(x,y) \in B^n : c_1 k \log n \le x < c_1(k+1)\log n, l\,(\log n)^3 \le y < (l+1)\,(\log n)^3\},$$

$(k, l \in \mathbb{Z})$.

For $i = 1, \dots, n^\gamma$ consider the following indicators: Let $\mathbb{I}^{1,n}(i)$ be the indicator for the event that $S_{\vartheta_i^n + n^{\frac{4+\beta}{3}}} \in B^{n/\log n}$. $\mathbb{I}^{2,n}(i)$ denotes the indicator for the event that the whole trajectory $(S_k)_{k=\vartheta_i^n,\dots,\vartheta_i^n+n^{\frac{4+\beta}{3}}}$ is contained in B^n. Furthermore, let $\mathbb{I}^{3,n}(i)$ be one if the random walk visits more than $n^{\frac{2(1+\beta)}{3}}/(\frac{2(1+\beta)}{3}\log n)$ distinct points in $[\vartheta_i^n, \vartheta_i^n + n^{\frac{4+\beta}{3}}]$ and zero otherwise.

Moreover let $\mathbb{I}^{4,n}_{k,l}(i)$ be the indicator for the event that in the time interval $[\vartheta_i^n, \vartheta_i^n + n^{\frac{4+\beta}{3}}]$ the walk enters $W^n_{k,l}$ and within $(\log n)^3$ steps after the first entrance time touches one of the lines $x = k\log n$ or $x = (k+1)\log n$, and finally follows the straight line supporting the the word associated with the starting point it touched.

First consider the event $\{\mathbb{I}^{1,n}(i) = 0\}$. By concentration of measure (cf. [13]) we have for every fixed i

$$\mathbb{P}(\mathbb{I}^{1,n}(i) = 0) = \mathbb{P}(||S_{\vartheta_i^n + n^{\frac{4+\beta}{3}}}|| \ge \frac{n}{\log n}) \le e^{-const.\frac{n^{\frac{2-\beta}{3}}}{(\log n)^2}}.$$

Therefore, as $\beta < 2$ as well as $\gamma < 2$

$$\mathbb{P}\left(\left(\bigcap_i \{\mathbb{I}^{1,n}(i) = 1\}\right)^c\right) \le n^2 e^{-const.\frac{n^{\frac{2-\beta}{3}}}{(\log n)^2}}$$

which is finitely summable and thus $\bigcap_i \{\mathbb{I}^{1,n}(i) = 1\}$ holds true for almost all n. Here and in the following $\bigcap_i$ refers to the intersection over $i = 1, \dots, n^\gamma$.

By the same argument

$$\begin{aligned}\mathbb{P}(\mathbb{I}^{2,n}(i) = 0) &= \mathbb{P}(\exists\, t \in [\vartheta_i^n, \vartheta_i^n + n^{\frac{4+\beta}{3}}] \,:\, ||S_t|| \ge n) \\ &\le n^2 \mathbb{P}(||S_{\vartheta_i^n + n^{\frac{4+\beta}{3}}}|| \ge n) \le n^2 e^{-const.\frac{n^{2-\beta}}{3}}.\end{aligned}$$

Thus, also $\bigcap_i \{\mathbb{I}^{2,n}(i) = 0\}$ holds true for all but finitely many n.

To bound the probability that $\mathbb{I}^{3,n}(i)$ is equal to zero, first observe that the number of distinct points D_t visited by a simple symmetric random walk starting at the origin by time t satisfies (cf.[4], [3])

$$ED_t \ge \frac{2t}{\log t}$$

for all t large enough. Moreover such a random walk clearly can only have visited at most t points (i.e. $D_t \le t$) up to time t. Together this implies

$$\mathbb{P}(D_t \ge \frac{t}{\log t}) \ge \frac{1}{\log t}. \tag{4.3.2}$$

Partitioning the interval $[\vartheta_i^n, \vartheta_i + n^{\frac{4+\beta}{3}}]$ into $n^{\frac{2-\beta}{3}}$ intervals of length $n^{2\frac{1+\beta}{3}}$ and applying (4.3.2) with $t = n^{2\frac{1+\beta}{3}}$ (observe that $\log t = 2\frac{1+\beta}{3}\log n$) yields for any fixed i:

$$\mathbb{P}(\mathbb{I}^{3,n}(i) = 0) \leq \left(1 - \frac{const.}{\log n}\right)^{n^{\frac{2-\beta}{3}}} \leq e^{-const.\frac{n^{\frac{2-\beta}{3}}}{\log n}}.$$

Hence by the same summability argument as above $\bigcap_i \{\mathbb{I}^{3,n}(i) = 1\}$ holds for almost all n.

Next let us have a closer look at $\{\mathbb{I}^{4,n}_{k,l}(i) = 1\}$. Suppose that we already know that S_n enters the sector $W^n_{k,l}$ within $[\vartheta_i^n, \vartheta_i^n + n^{\frac{4+\beta}{3}}]$. Considering just the projection of the walk to the x-axes, we see a nearest neighbor random walk on $\mathbb{Z}$ with holding probability $1/2$. The points $k\log n$ and $(k+1)\log n$ obtained by projecting the vertical limiting lines of $W^n_{k,l}$ may be considered absorbing barriers for this random walk. As the expected hitting time of one of these barriers is of order $(\log n)^2$, after time $(\log n)^3$ we will have hit one of the boundaries with a probability bounded away from zero (in n). In other words that is to say, that S_n conditioned on that it will visit $W^n_{k,l}$ at all, will touch one of its left and right boundary lines within $(\log n)^3$ after the first entrance time into this sector with probability bounded away from zero. As the word associated to this boundary point has length $c_1 \log n$ the probability that the walk touches a boundary point and then follows the walk associated to it is bounded by $const.(1/4)^{c_1 \log n}$.

Note that the events $\{\mathbb{I}^{4,n}_{k,l}(i) = 1\}$ are not independent for different choices of (k,l) and the same i and n. First due to the fact that (S_k) is a Markov chain the event $\{\mathbb{I}^{4,n}_{k,l}(i) = 1\}$ increases the chances that we also hit a square close to $W^n_{k,l}$. However, also given that we visit both $W^n_{k,l}$ and $W^n_{(k+1),l}$ for example, the events $\{\mathbb{I}^{4,n}_{k,l}(i) = 1\}$ and $\{\mathbb{I}^{4,n}_{k+1,l}(i) = 1\}$ are dependent since reading a word associated with a boundary point of $W^n_{k,n}$ might easily coincide with touching a boundary point of $W^n_{k+1,n}$ less than $(\log n)^3$ steps after the first entrance time. To cope with this effect we disregard every other square, that is we consider the indicators

$$\hat{\mathbb{I}}^{4,n}_{k,l}(i) := \mathbb{I}^{4,n}_{k,l}(i)\mathbb{I}(k,l)$$

where $\mathbb{I}(k,l)$ is $+1$ if k and l are even and 0 otherwise, instead.

Now observe that on $\{\mathbb{I}^{2,n}(i) = 1\} \cap \{\mathbb{I}^{3,n}(i) = 1\}$ the random walk visits more than $n^{2\frac{1+\beta}{3}}/\frac{2(1+\beta)}{3}\log n$ distinct points within $[\vartheta_i^n, \vartheta_i^n + n^{\frac{4+\beta}{3}}]$ – all of them lying in B^n – and therefore, as each of the $W^n_{k,l}$ has $c_1(\log n)^4$ points, also $n^{2\frac{1+\beta}{3}}/(2c_1\frac{1+\beta}{3}(\log n)^5)$ distinct $W^n_{k,l}$'s. As one fourth of them will have both k and l even $\hat{\mathbb{I}}^{4,n}_{k,l}(i)$ has a chance to become $+1$ for $n^{2\frac{1+\beta}{3}}/(8c_1\frac{1+\beta}{3}(\log n)^5)$ different choices of (k,l). Given the indices (k,l) for which this is true the events $\{\hat{\mathbb{I}}^{4,n}_{k,l}(i) = 1\}$ indeed are independent and have probability at least $const.(1/4)^{c_1 \log n}$. Hence again by moderate deviations or concentration of measure on $\{\mathbb{I}^{2,n}(i) = 1\} \cap \{\mathbb{I}^{3,n}(i) = 1\}$

$$\mathbb{P}(\sum_{k,l} \hat{\mathbb{I}}^{4,n}_{k,l}(i) \leq n^\beta) \leq \exp\left(-const.\frac{n^{\frac{1}{3}(2-\beta)-c_1\log 4}}{(\log n)^{10}}\right) \leq e^{-n^\varepsilon}$$

for some small ε, if c_1 is small enough (depending on how large we have chosen β before). As e^{-n^ε} is finitely summable even after multiplication with the number of different ϑ_i^n we

obtain that on the event $\bigcap_i\{\{\mathbb{I}^{2,n}(i)=1\}\cap\{\mathbb{I}^{3,n}(i)=1\}\}$ we have $\sum_{k,l}\hat{\mathbb{I}}^{4,n}_{k,l}(i)\geq n^\beta$ for all i and all but finitely many n's. As also $\bigcap_i\{\{\mathbb{I}^{2,n}(i)=1\}\cap\{\mathbb{I}^{3,n}(i)=1\}\}$ holds for almost all n

$$\sum_{k,l}\hat{\mathbb{I}}^{4,n}_{k,l}(i)\geq n^\beta$$

also is true for almost all n. As finally also $\bigcap_i\{\mathbb{I}^{1,n}(i)=1\}$ for all but a finite number of n's, we arrive at

$$\bigcap_i\{\{Y_i=1\}\cap\{\mathbb{I}^{1,n}(i)=1\}\}$$

for all n but finitely many. (Recall that the random variables Y_i were the indicators for the event that the interval $[\vartheta^n_i,\vartheta^n_i+n^2]$ is a signal time, that is for the event that there are more than n^β signal words read in $[\vartheta^n_i,\vartheta^n_i+n^2]$.)

Let us summarise what we know already. For almost all n the following holds true: Until time e^{n^α} we have more than n^γ (γ smaller than α) different intervals of length n^2 of signal times. The signals are read in the first $n^{\frac{4+\beta}{3}}$ steps, after which the random walks stops in a distance at most $n/\log n$ from the origin.

Finally we have to show that in these time intervals $[\vartheta^n_i,\vartheta^n_i+n^2]$ we also read all words of length $(2c_2+c_3)\log n$ beginning with either a root word or a side word associated to any of the boundary points. To avoid trouble with independence we will only concentrate on events where this happens in one of the time intervals $J^n_i:=[\vartheta^n_i+n^{\frac{4+\beta}{3}},\vartheta^n_i+n^2]$, $i=1,2,\ldots$.

To this end, first observe that on a time interval of length $|J^n_i|$ the random walk (S_k) deviates form its staring point by the variance of a sum $|J^n_i|$ many independent random variables with variance one. This is immediately computed as

$$\sqrt{n^2-n^{\frac{4+\beta}{3}}}\geq\frac{n}{2}$$

for n large enough. Therefore, and since "in the worst case" $S_{\vartheta^n_i+n^{\frac{4+\beta}{3}}}=0$ with positive probability bounded away from zero (S_k) exits B^n during J_i. This bound will be used to estimate the probability to hit the beginning $\sigma_0(v)$ of a root word for a boundary point $v\in\underline{\partial}B^n$ or the beginning of one of its side words. This probability can be computed as the probability of hitting this point conditioned on that we hit the (discrete) sphere it is contained in, times the probability that we hit this sphere at all. The latter probability is bounded below by a constant away from zero, by the above considerations. On the other hand the probability to hit a certain point in $\underline{\partial}B^n$ conditioned on that we leave B^n is bounded below by $\frac{\varkappa}{n}$ for some constant $\varkappa>0$ no matter where in $B^{n/\log n}$ we started. Of course, it suffices to understand that this is true for large n. But observing that under the scaling $\mathbb{Z}^2\to\frac{1}{n}\mathbb{Z}^2$, the boundary $\underline{\partial}B^n$ converges to the unit sphere, $B^{n/\log n}$ shrinks to the origin and (S_k) converges (after rescaling also the time axes which is irrelevant for our argument) to Brownian motion $W^0(t)$ starting at the origin and moreover taking into account that the harmonic measure on the unit sphere (any sphere centered in zero) with respect to $W^0(t)$ is the uniform distribution on it, shows that the above bound indeed holds. So, as all starting points of root words and side words lie in $B^{n^2}\setminus B^{n^2-(c_2+c_3)\log n}$ we see that the probability of hitting any fixed starting point is bounded from below by

$\frac{\varkappa'}{n}$ for some $\varkappa' > 0$ ($\varkappa'$ results from multiplying $\varkappa$ with the probability of exiting B^{n^2} in a certain J_i).

Now the probability of reading $\hat{\sigma}(v)$ and after that any fixed continuation of length $(c_2 + c_3)\log n$ given that we first read $\sigma_0(v)$ has (for any fixed $v \in \underline{\partial} B^n$) probability

$$\left(\frac{1}{4}\right)^{(2c_2+c_3)\log n} = n^{-(2c_2+c_3)\log 4}.$$

So the (unconditioned) probability of reading $\hat{\sigma}(v)$ and after that any fixed continuation of length $(c_2 + c_3)\log n$ is bounded below by

$$\frac{\varkappa}{n}\left(\frac{1}{4}\right)^{(2c_2+c_3)\log n} = \varkappa n^{-1-(2c_2+c_3)\log 4}.$$

On the other hand there are n^γ different time intervals where we can read such a word. So the probability of not reading $\hat{\sigma}(v)$ and after that any fixed continuation of length $(c_2 + c_3)\log n$ in all of these intervals behaves like

$$\left(1 - \varkappa n^{-1-(2c_2+c_3)\log 4}\right)^{n^\gamma} \leq \exp\left(-\varkappa n^{\gamma-1-(2c_2+c_3)\log 4}\right).$$

As we can choose c_2 and c_3 as small as we want to and $\gamma > 1$ (and still γ, α) this probability is smaller than e^{-n^ε} for some $\varepsilon > 0$. The same holds true for the probability of reading a side word and then any fixed continuation of length $(c_2 + c_3)\log n$ given that we read its first letter. As for fixed n there are only polynomially many of such nearest neighbour walk paths (more precisely, as there less than

$$6\pi n 4^{(c_2+c_3)\log n} = 6\pi n^{1+(c_2+c_3)\log 4}$$

such nearest neighbour walk paths) the probability of not reading all of them is bounded by

$$6\pi n^{1+(c_2+c_3)\log 4} e^{-n^\varepsilon}$$

which is finitely summable in n. Therefore, by the Borel-Cantelli Lemma, also E_n^3 holds for all but finitely many n's. This finishes the proof of Lemma 4.2.3. □

The proof of the main theorem now only consists of choosing the constants in the correct order.

Proof of Theorem 5.2.1: To finish the proof we finally specify the order in which we choose the constants. So first we choose α, β, γ with $2\beta - 2 > \alpha$ (such that right hand side in (**??**) is finitely summable), and $1 < \gamma < \alpha$. Then we choose c_1, c_2 and c_3 to make the last part of the above proof of Lemma 4.2.3 work (note that this part does not depend on the number of colors m). If we now choose m larger than a certain number m_2 (coming from the arguments which guarantee that E_1^n and E_2^n holds for all but a finite number of n's), this procedure ensures that the reconstruction in Lemma 4.2.3 works with probability one for all but a finite number of n's.

Thus for $1/2 > \varepsilon > 0$ we can choose N (non-random) such that the probability that we have

$$\tilde{\mathcal{A}}^n(\xi|B^n, \chi) \sim \xi|B^{n+1}$$

for all $n \geq N$ is bigger than $1 - \frac{\varepsilon}{2}$ given $m \geq m_2$.

Now for N there exists m_1 such that for $m \geq m_1$ the reconstruction algorithm from Lemma 2.2 $\mathcal{A}^N$ ensures that we can reconstruct $\xi | B^N$ with probability larger than $1 - \frac{\varepsilon}{2}$. If we now choose $m \geq \max\{m_1, m_2\}$ and concatenate $\mathcal{A}^N$ from Lemma 2.2 with the different $\mathcal{A}^n$ for $n \geq N + 1$ from Lemma 2.3, we obtain an algorithm $\mathcal{A}$ which reconstructs ξ with probability larger than $1 - \varepsilon$.

In view of Lemma 2.1 this suffices to prove Theorem 5.2.1. □

Figure 4.1: Figure 1

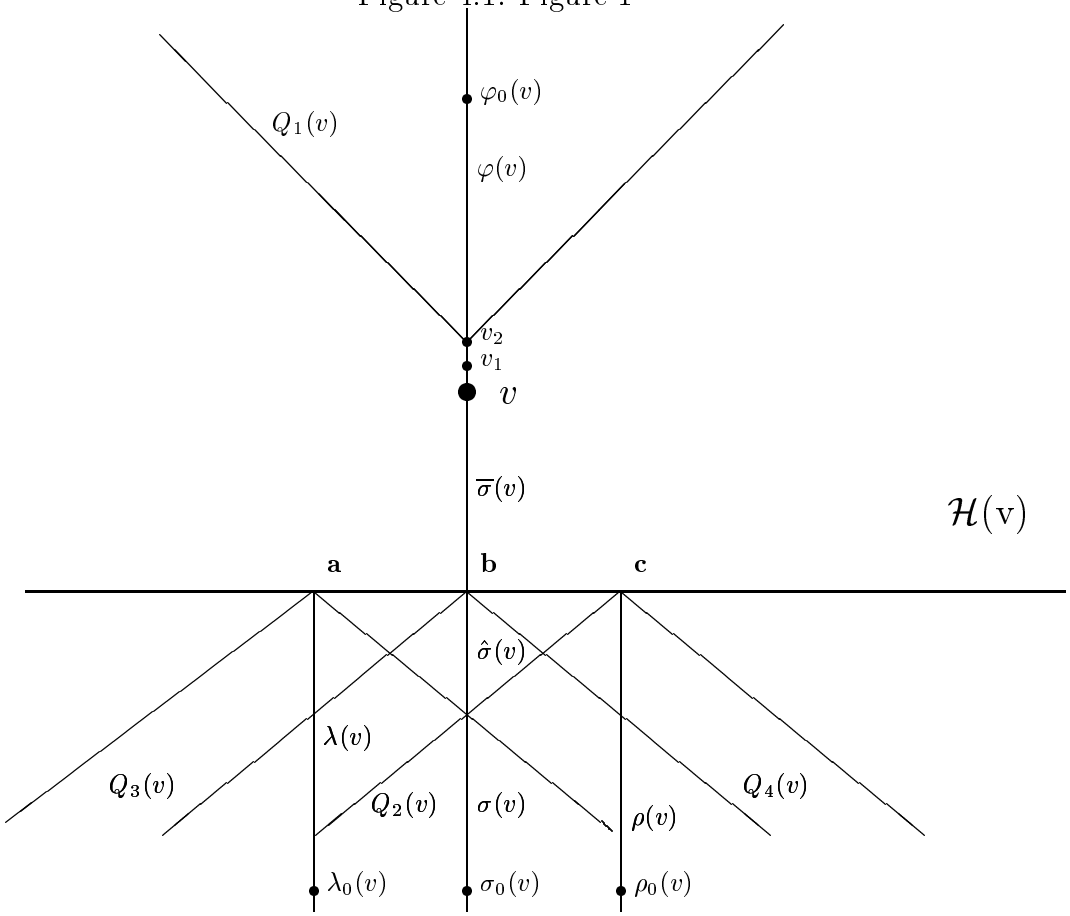


Figure 4.2:

References

[1] I. Benjamini, H. Kesten; *Distinguishing sceneries by observing the sceneries along a random walk path*, J. d'Anal. Math **69**, 97–135 (1996)

[2] P. Erdös, S.J. Taylor; *Some problems concerning the structure of random walk paths*, Acta Math. Acad. Sci. Hung. **11**, 137–162 (1960)

[3] D. Downham, S. Fotopolous; *The transient behaviour of the simple random walk in the plane*, J. Appl. Prob. **25**, 58–69 (1988)

[4] A. Dvoretzky, P. Erdös; *Some problems on random walks in space*, Proc. 2nd Berkeley Math. Stat. Prob., 353–367 (1951)

[5] W. Th. F. den Hollander, M. Keane; *Ergodic properties of color records*, Physica **138A**, 183–193 (1986)

[6] C.D. Howard; *Detecting defects in periodic scenery by random walks in* $\mathbb{Z}$, Random Structures and Algorithms **8**, 59–74 (1996)

[7] H. Kesten; *Distinguishing and reconstructing sceneries from observations from random walk paths*, Microsurveys in discrete probability (D. Aldous and J. Propp, eds.), DIMACS Series in Discrete Mathematics and Theoretical Computer Sciences **41**, 75–83 (1998)

[8] E. Lindenstrauss; *Indistinguishable Sceneries*, random Struct. Alg. **14**, 71–86 (1999)

[9] H. Matzinger; *Reconstructing a 3-color scenery by observing it along a simple random walk*, Preprint, to apper in: Random Struct. Alg. (1997)

[10] H. Matzinger; *Reconstruction of a one dimensional scenery seen along the path of a random walk with holding*, Ph. D. Thesis, Cornell University (1999)

[11] P. Revesz; *Estimates on the largest disc covered by a random walk*, Ann. Prob **18**, 1784–1789 (1990)

[12] F. Spitzer; *Principles of random walk*, Van Nostrand, London (1964)

[13] M. Talagrand; *A new look at independence* Ann. Prob **24**, 1–34 (1996)

Chapter 5

Reconstruction of Sceneries with Correlated Colors

Ann. Appl. Probab., 12(4):1322–1347, 2002.

By Matthias Löwe, Heinrich Matzinger,

In [9] Matzinger showed how to reconstruct almost every three color scenery, that is a coloring of the integers $\mathbb{Z}$ with three colors, by observing it along the path of a simple random walk, if this scenery is the outcome of an i.i.d. process. This reconstruction needed among others the transience of the representation of the scenery as a random walk on the three-regular tree T_3. Den Hollander (private communication) asked which conditions are necessary to ensure this transience of the representation of the scenery as a random walk on T_3 and whether this already suffices to make the reconstruction techniques in [9] work. In this note we answer the latter question in the affirmative. Also we exhibit a large class of examples where the above mentioned transience holds true. Some counterexamples show that in some sense the given class of examples is the largest natural class with the property that the representation of the scenery as a random walk is transient. [1]

5.1 Introduction

The following problems to which this paper will make a contribution were discovered in the context of ergodic theory, for example in connection with the so-called $T - T^{-1}$-problem (see Kalikow [3]), and phrased as statistical questions independently by den Hollander and Keane [2] and Benjamini and Weiss.

For our purposes we will consider the one dimensional lattice $\mathbb{Z}$. Actually, the following problems make sense also for arbitrary graphs, but as there are hardly any results apart from the case when this graph is $\mathbb{Z}^d$ for some $d \in \mathbb{N}$, we immediately concentrate to our object of desire. Assume that $\mathbb{Z}$ is colored with m colors. More precisely, we consider two

[1] *MSC 2000 subject classification:* Primary 60K37, Secondary 60G10, 60J75.
Key words: Scenery reconstruction, jumps, stationary processes, random walk, ergodic theory.

such colorings, that is we consider two functions

$$\eta, \xi : \mathbb{Z} \to \{0, \dots, m-1\}$$

and call these functions m-color sceneries or simply sceneries. Let $(S_k)_{k\in\mathbb{N}_0}$ be symmetric and simple random walk on $\mathbb{Z}$ starting in the origin and walking without holding, that is $S_0 = 0$ and

$$P(S_{k+1} = x+1 | S_k = x) = P(S_{k+1} = x-1 | S_k = x) = \frac{1}{2}$$

for all $x \in \mathbb{Z}$ and $k \in \mathbb{N}_0$. Moreover define $\chi := (\chi_k)_{k\in\mathbb{N}_0}$ to be the color record of $(S_k)_{k\in\mathbb{N}_0}$, that is either

$$\chi_k = \xi(S_k) \quad \text{for all } k \text{ or} \quad \chi_k = \eta(S_k) \quad \text{for all } k$$

depending on which scenery we observe the colors. The question now is: can we just by observing χ (and, of course, without any further knowledge of $(S_k)_{k\in\mathbb{N}_0}$) tell on which of the sceneries ξ or η this color record χ has been produced?

Remarkable answers to this question (even for the higher dimensional case) have been given by Benjamini and Kesten [1], who showed that if ξ and η are produced by an i.i.d. process on $\mathbb{Z}$ (that is to say, if $\xi(z)$ and $\eta(z), z \in \mathbb{Z}$ are i.i.d. random variables), then $\xi(z)$ and $\eta(z)$ can almost surely be distinguished by their color record, if the dimension $d = 1, 2$ and $m \geq 2$ is arbitrary. More precisely in this situation, there exists a test which tells with probability one on which of ξ and η the color record χ has been produced (even with a slightly stronger version of distinguishability excluding trivial solutions such as benefiting from the fact that e.g. $\xi(0) \neq \eta(0)$).

Also Kesten (see [4]) showed that in dimension one, if $m \geq 5$, and ξ is again i.i.d. we can almost surely detect a single defect in ξ from knowing χ, that is, we can almost surely tell, whether χ has been produced on ξ or a scenery η differing from ξ in one vertex $i \in \mathbb{Z}$, only. Here and in the following the notion "almost surely" will refer to a probability measure $\mathbb{P}$ describing both, the randomness in $(S_k)_{k\in\mathbb{N}_0}$, and the randomness in ξ (or in ξ and η, if we are interested in two sceneries) and making $(S_k)_{k\in\mathbb{N}_0}$ and ξ (or $(S_k)_{k\in\mathbb{N}_0}$, η, and ξ, respectively) independent.

Indeed even more is true: In dimension one Matzinger showed in [9], [10] that for arbitrary $m \geq 2$ one can even almost surely reconstruct ξ from $\chi = (\xi(S_k))_{k\in\mathbb{N}_0}$, that is one can reproduce a scenery ξ' from χ which is equal to ξ up to translation and reflection at the origin. This is even true, if the underlying random walk S is allowed to jump (if the jumps are bounded and m is large enough). The latter was shown recently by the authors in collaboration with Merkl [8].

All these results are particularly surprising, since on the other hand it is known that there are uncountably many sceneries which cannot be distinguished by their color record. This has been proven by Lindenstrauss [6].

The analogue to the reconstructability of ξ from χ in two dimensions has been recently proven by the authors under the conditioned that m is large enough (see [7]).

Basically all reconstruction and distinction techniques cited above (with one exception) seem to strongly exploit the fact that the scenery is an i.i.d. process, that is that $\xi(z)$, $z \in \mathbb{Z}$ are i.i.d. random variables. Only the methods employed by Matzinger in [9] are partially combinatorial and therefore seem to allow for a generalization to other sceneries. Indeed, den Hollander (private communication) asked what conditions for the scenery

would be necessary to make the reconstruction ideas in [9] work. We consider an answer to this question interesting in its own rights as it sheds some light on the universality of the solution to the above problem. In particular, the roots of the scenery reconstruction problem in ergodic theory make it attractive to ask for the ergodic properties of the sceneries needed to ensure reconstructability. Moreover, it was pointed out to us, that similar ideas might be useful in the context of DNA reconstruction. As a DNA sequence usually is assumed to be Markov-dependent of some type (at least the assumption of i.i.d. letters is rather far fetched) an analysis of what kind of sceneries are reconstructible might also be helpful concerning possible applications.

This paper is divided into five sections. Section 2 contains a description of the basic setup and the first central result of this paper. In Section 3 we describe the fundamental reconstruction algorithm, while Section 4 contains a proof that the algorithm actually works under the conditions of Theorem 5.2.1. Finally in Section 5 we give the most important examples where Theorem 5.2.1 applies and also cases where it does not apply and actually reconstruction along the ideas of this paper is not possible. Furthermore, we indicate that for these examples not only we can reconstruct a randomly chosen scenery ξ, but also that there is a very powerful test for the initial problem of distinguishing it from any other scenery.

Acknowledgment: We are extremely grateful to an unknown referee who pointed out several weak points in an earlier draft of this paper and thus helped to improve its correctness and readability.

5.2 The Setup and the Main Result

Before we give the first central result of our paper in this section let us quickly introduce and recall the most important notations. If not mentioned otherwise, in what follows ξ will always be a one dimensional 3-color scenery, that is

$$\xi : \mathbb{Z} \to \{0, 1, 2\}.$$

Actually, the generalization to $m \geq 4$ is easy and straightforward, but there will also be examples with $m = 2$ where Theorem 5.2.1 applies. We will always choose ξ randomly from a class of sceneries in such a way that the conditions of Theorem 5.2.1 are fulfilled.

Moreover let $S = (S_k)_{k \in \mathbb{N}_0}$ be symmetric, simple random walk without holding on $\mathbb{Z}$ starting in the origin. The measure $\mathbb{P}$ will denote the product measure on the product space of all sceneries and all random walk paths (with the obvious marginals). Hence we will always assume independence of the walks and the sceneries. The central problem will be to reconstruct ξ from its color record

$$\chi := (\chi((S_k)_{k \in \mathbb{N}_0})) := ((\xi(S_k)_{k \in \mathbb{N}_0})$$

under S (by $\chi|[0, n]$ we will denote the first $n + 1$ observations). This means we want to find a measurable mapping

$$\mathcal{A} : \{0, 1, 2\}^{\mathbb{N}} \to \{0, 1, 2\}^{\mathbb{Z}}$$

such that $\mathbb{P}$-almost surely

$$\mathcal{A}(\chi) \sim \xi.$$

Here for two sceneries $\eta, \xi \in \{0,1,2\}^{\mathbb{Z}}$ we write $\xi \sim \eta$ (and say that they are equivalent), if there are $a \in \mathbb{Z}$ and $s \in \{-1,+1\}$ such that

$$\xi(z+a) = \eta(sz)$$

for all $z \in \mathbb{Z}$, i.e. ξ can be obtained form η by translation and reflection at the origin.

For the idea of the reconstruction algorithm the following representation of ξ as a path on a colored tree is essential. In our case this tree will be the 3-regular tree $T_3 := (V_3, E_3)$, that is the connected, unrooted, infinite tree with all its vertices $v \in V_3$ having degree 3. We choose one (arbitrary) vertex and call it the origin o. We color T_3 (or more precisely V_3) in three different ways φ^0, φ^1 and φ^2. Up to isomorphisms of the tree these colorings are uniquely defined by the color of the origin (or root) o

i)
$$\varphi^i(o) = i$$
for $i = 0,1,2$

and the following construction rule

ii) For each $v \in V_3$ let $\{v_1, v_2, v_3\}$ be the set of its neighboring vertices (according to the graph topology induced by E_3). Then
$$\{\varphi^i(v_1), \varphi^i(v_2), \varphi^i(v_3)\} = \{0,1,2\}$$
for each $i = 0,1,2$ and each $v \in V_3$.

Now we can represent ξ as a nearest neighbor path R on T_3 (taking the randomness in ξ into account this will be a random path, but note that other than in [9] R is not necessarily a random walk on T_3). This will be done in the following way.

a) Choose the coloring φ^i with $i = \xi(0)$. (Note that we know $\xi(0)$ as the random walk S is supposed to start in zero).

b) We let $R = (R(z))_{z \in \mathbb{Z}}$ be the nearest neighbor random path (that is $R(z)$ and $R(z+1)$ are adjacent for each $z \in \mathbb{Z}$) on T_3 colored with $\varphi^{\xi(0)}$ such that
$$R(0) = o$$
and
$$\varphi^{\xi(0)}(R(z)) = \xi(z)$$
for all $z \in \mathbb{Z}$.

Note that given T_3, the choice of o, and the coloring this path R is unique. This representation of ξ as a random path on T_3 colored with $\varphi^{\xi(0)}$ will indeed help us to reconstruct ξ up to equivalence. To this end note that knowing R plus knowing $\xi(0)$ is indeed equivalent to knowing ξ. Unfortunately, we do not know R but only $R \circ S$ (the latter because of

$$\varphi^{\xi(0)} \circ (R \circ S) = (\varphi^{\xi(0)} \circ R) \circ S = \xi \circ S = \chi$$

and thus we can reconstruct $R \circ S$ from χ).

Interestingly, the only knowledge we require about R in order to reconstruct ξ is that it is transient, that is that the random path R (again recall that R is random as ξ is random) visits each vertex $v \in V_3$ only finitely many times almost surely. This is a considerable improvement of Matzinger's previous result [9], who even for i.i.d. sceneries needed some further conditions. A major tool in the proof of the following theorem will consist of reformulation of this transience in terms of *crossings* of some pieces of the tree T_3 by R. For some $v \neq w \in V_3$ we say that a time interval $[s,t]$ (without loss of generality $s < t$ – otherwise we just reverse the order of v and w) is a *crossing* of (v,w), if $R(s) = v$, $R(t) = w$, or $R(s) = w$, $R(t) = v$, and

$$R(s') \neq v \quad \text{and} \quad R(s') \neq w.$$

for all $s < s' < t$. Observe that two crossings of (v,w) either agree or are disjoint (that is the time intervals are disjoint).

Moreover we say that $[s,t]$ is a shortest crossing of (v,w) by R, if

$$t - s = \min\{|t' - s'|, [s',t'] \text{ is a crossing of } (v,w) \text{ by } R\}.$$

Now we are ready to formulate the first central result of this paper (the other will be stated in Section 5).

Theorem 5.2.1. *With the above definitions, assume that the random path R is transient. Then there exists a measurable mapping*

$$\mathcal{A} : \{0,1,2\}^{\mathbb{N}} \to \{0,1,2\}^{\mathbb{Z}}$$

such that

$$\mathbb{P}(\mathcal{A}(\chi) \sim \xi) = 1.$$

5.3 The Algorithm

In this section we are going to present the basic reconstruction scheme, while details will be left to the proofs to follow in the next section. The core of this algorithm consists of stopping the random walk S infinitely many times at the same place. We will see in the next section that this indeed is already enough to be able to reconstruct ξ up to equivalence. Actually this stopping of the random walk is of a different nature, when ξ is essentially symmetric (by this we mean that there is a finite interval $I = (a,b)$ such that $\xi(a - x) = \xi(x - b)$ for all $x \in \mathbb{N}$). Therefore, we first test ξ on essential symmetry. This symmetry can also be expressed in terms of the path R on T_3 which will exploit in the first step of the algorithm.

Step 1 of the Reconstruction Algorithm:

Test whether there are $v \neq w \in V_3$ such that there is only one shortest crossing of (v,w)

¿From Step 1 the algorithm proceeds in two different directions depending on whether it has been successful (that is there are v, w with only one shortest crossing of (v,w) by R) or not. The first case will be called Case A the other one Case B.

Case A (there is at least one pair $v \neq w$ such that there is only one shortest crossing of (v, w)):

Here we proceed by producing infinitely many stopping times all stopping S at the same point.

Step 2 of the Reconstruction Algorithm:
Stop the random walk S infinitely often at the same point.

Finally, we use these stopping times to reconstruct ξ.

Step 3 of the Reconstruction Algorithm: Reconstruct ξ up to equivalence with probability one from these stopping times.

In Case B, where for every $v \neq w \in V_3$ there are at least two shortest crossings of (v, w) by R, we have to apply a slightly different techniques.

Case B (For all $v, w \in V_3$ there are at least two shortest crossings of (v, w):)

Step 2 of the Reconstruction Algorithm:
Stop the random walk S infinitely often at two different points.

Step 3 of the Reconstruction Algorithm: Reconstruct ξ up to equivalence with probability one from these stopping times.

Of course, this is a very rough description of the algorithm. We will fill its different steps with life in the next section, where we prove that it actually works.

5.4 Proof that the Algorithm works

In this section we show that under the condition that R is transient the algorithm actually reconstructs ξ up to equivalence with probability one. This proof is split into different parts. In the first part we show that if the walk is transient, then almost surely there are vertices $v, w \in V_3$ such that there are at most two crossings of (v, w) by R.

Definition 5.4.1. *1. Let W, W' be a set and $f : W \to W'$ be a mapping. Then $\operatorname{Im} f := \{f(w), w \in W\}$ denotes the image of f.*

2. Consider the the 3-regular tree $T_3 := (V_3, E_3)$ and two vertices $v, w \in V_3$. The graph distance $d(v, w)$ is defined as minimum the minimum length of a path (minimum number of connected edges in E_3) to be crossed to get from v to w.

Lemma 5.4.1. *Let R be transient. Then for every fixed $v \in V_3$ and every sequence $v_n \in V_3$ in the image of R with $d(v, v_n) = n$ and $d(v_{n-1}, v_n) = 1$ almost surely the number $N(n)$ of distinct shortest crossings of (v, v_n) by R converges and the limit is $\mathbb{P}$-almost surely either 1 or 2.*

Remark 5.4.1. *Note that due to the transience of R the Image $\operatorname{Im} R$ of the representation of the scenery ξ on T_3 is almost surely infinite. Hence such a point v and a sequence of points v_n as assumed in the above lemma actually exist.*

Proof of Lemma 5.4.1. Without loss of generality we will take v to be the origin o. By transience of R the origin is visited by R only finitely many times with probability one. Thus with probability one the random variables

$$t_{\max} := \max\{t : R(t) = 0\}$$

and

$$t_{\min} := \min\{t : R(t) = 0\}$$

are well defined and obey

$$-\infty < t_{\min} \leq 0 \leq t_{\max} < \infty.$$

In particular, $\Delta t := t_{\max} - t_{\min}$ is finite with probability one. Now take $v_n \in \operatorname{Im} R$ with $d(o, v_n) = n$ as assumed in the above lemma. Then for all n large enough $d(v_n, o) > \Delta t$. Moreover, since $v_n \in \operatorname{Im} R$ there exists $t_n \in \mathbb{Z}$ such that $R(t_n) = v_n$. As $d(v_n, o) > \Delta t$ we conclude that

$$t_n \notin [t_{\min}, t_{\max}].$$

Thus at most two crossings of (o, v_n) can occur, one of the form $(t_{\max}, t_1)$ and another one of the form $(t_2, t_{\min})$, where

$$t_1 := \min\{t > t_{\max} : R(t) = v_n\}$$

and

$$t_2 := \max\{t < t_{\min} : R(t) = v_n\}.$$

Also note that because $v_n \in \operatorname{Im} R$ one of the above two crossings really can be found. □

As the we will see in the next lemma not only we either are in Case A or in Case B, but also is there a test which reveals with probability one (given a full color record) in which of the two situations we are.

Lemma 5.4.2. *For each $v, w \in \operatorname{Im} R$ there exists a test that on the basis of the observations χ decides with probability one whether there is only one shortest crossing from v to w or whether there is more than one such shortest crossing.*

Proof. Let $v, w \in \operatorname{Im} R$ be any two points reached by the random walk R. Note that S as a random walk on the integers is recurrent and hence so is $R \circ S$ as a random path on T. Therefore, $R \circ S$ will pass every finite path in $\operatorname{Im} R$ infinitely often and thus,

$$\tilde{T} := \min\{|s - t|; R(S(s)) = v, R(S(t)) = w\}$$

estimates the time for the shortest crossing of (v, w) by R correctly with probability one.

Also note that the distribution function of the waiting time between two shortest crossings of (v, w) by $R \circ S$ is strictly larger if there is more than one shortest crossing of (v, w) by R than if there is just one such shortest crossing.

To be more specific, let W be the random variable that denotes the first time after which the random path $R \circ S$ has walked from v to w in $\tilde{T}$ steps.

$$W := \min\{n \geq 0, R \circ S(n - \tilde{T}) = v, R \circ S(n) = w\}.$$

Moreover let F the distribution function of W conditioned on starting with S in the point z corresponding to the point w (via the representation R) of the unique shortest crossing of (v, w) by R at time 0, i.e.

$$F := \mathcal{L}(W|S(0) = z).$$

When there are several shortest crossings of (v, w) by R, say $[y_1, z_1], \ldots, [y_k, z_k]$ with $k \geq 2$ is the set of all shortest crossings of (v, w) by R and (for our purposes without loss of generality) $R(z_1) = R(z_2) = \ldots = R(z_k) = w$ we denote with F_i the distribution of W when starting with S in the point w_i at time zero, $i = 1, \ldots, k$.

Note that the distribution function F then will be strictly smaller than each of the the distribution functions F_i (i.e. $F(t) \leq F_i(t)$ for all t and that $F(t) < F_i(t)$ for all $t \geq T_0$ for some finite T_0). Indeed, if there are several shortest crossings of (v, w) by R, then a crossing from v to w by $R \circ S$ in $\tilde{T}$ steps may be caused by crossing one of several intervals $[y_1, z_1], \ldots, [y_k, z_k]$ in $\tilde{T}$ steps. Now the event to cross one of $[y_1, z_1], \ldots, [y_k, z_k]$ in $\tilde{T}$ steps has a higher probability than to cross a fixed interval $[y, z]$ in $\tilde{T}$ steps in case there is just one shortest crossing of (v, w) by R. This will eventually also show up in the distribution function of W.

Moreover, notice that F can be explicitly calculated when when we know that there is only one shortest crossing from v to w and we also know its length. Indeed, denoting by l the length of such a shortest crossing and considering the renewal process, with a renewal after every time where the random path $R \circ S$ has walked from v to w in l steps, we can calculate the probability that a time t is a renewal time. As a matter of fact, this can be done by observing that, if there is only one shortest crossing of (v, w), this crossing corresponds (by the random path R) to two points $z_1, z_2 \in \mathbb{Z}$ with $R(z_1) = v$ and $R(z_2) = w$ and $|z_1 - z_2| = l$. So the probability of having a renewal at time t equals the probability of walking with S from z_2 to z_1 in $t - l$ steps times 2^{-l} (for a straight crossing from z_1 to z_2). By a standard exercise in renewal theory, this also yields the probability that t is the time of a first renewal, hence the distribution function of W.

As we can also estimate the length l by $\tilde{T}$ with probability one correctly, we can, in principal, calculate the distribution function of W in the case where there is only one shortest crossing of (v, w) by R.

Finally, we can also test whether there is only one shortest crossing of (v, w) by R correctly with probability one.

In fact, if $[s_1, t_1], [s_2, t_2], \ldots$ denotes all intervals where the random path $R \circ S$ walks from v to w in $l = \tilde{T}$ steps, so $s_1 < t_1 < s_2 < t_2 < \ldots$ and $R(S(s_i)) = v$ and $R(S(t_i)) = w$ for all $i = 1, 2, \ldots$. Then by the law of large numbers (Glivenko-Cantelli-Lemma) the empirical distribution function of the "first renewal times"

$$\frac{1}{n-1} \sum_{i=2}^{n} \delta_{t_i - t_{i-1}}$$

converges to some distribution function $\overline{F}$ with probability one as n goes to infinity (of course, again, here we exploit the recurrence of S which gives us infinitely many such crossings).

If there is only one shortest crossing of (v, w) by R, the distribution function $\overline{F}$ will equal F with probability one, otherwise $\overline{F}$ will be a mixture of the F_i, hence larger than F.

So, if

$$\lim_{n\to\infty} \frac{1}{n-1} \sum_{i=2}^{n} \delta_{t_i - t_{i-1}}$$

differs from F (which we can calculate as indicated above) we conclude that there is more than one such shortest crossing, otherwise we decide that there is only one shortest crossing of (v, w) by R. As has been shown above this test succeeds in giving the correct number of shortest crossings of (v, w) with probability one. □

In the next steps we will see that the algorithm actually works in Case A. To this end we first have to show that we can indeed stop the random walk S infinitely often at the same place. So, let

$$\mathcal{H}_k := \sigma\{\xi(z), z \in \mathbb{Z}, S(0), \ldots, S(k)\}$$

and define the filtration $\mathcal{H}$ as

$$\mathcal{H} := \{\mathcal{H}_k, k \in \mathbb{N}\}.$$

Lemma 5.4.3. *If there are $v, w \in \operatorname{Im} R$ such that there is only one shortest crossing of (v, w) by R we can stop the random walk infinitely often at the same place, i.e. we are able to construct an infinite sequence of increasing stopping times $\tau(1), \tau(2), \ldots$ with respect to the filtration $\mathcal{H}$ such that*

$$S(\tau(1)) = S(\tau(2)) = \ldots = S(\tau(k)) = \ldots$$

Remark 5.4.2. *Observe that as has been already discussed in the context of Lemma 5.4.1 and in particular when motivating the algorithm in Section 3, Case A is is the relevant case for most distributions of the scenery we might think of. Indeed, whenever the distribution of the scenery exhibits some form of asymptotic independence, for example, the scenery will a.s. not be essentially symmetric and thus we will almost surely be in Case A.*

Proof of Lemma 5.4.3. Let $v, w \in \operatorname{Im} R$ such that there is only one shortest crossing of (v, w) by R. Let the length of this shortest crossing be $\tilde{L}$. This length $\tilde{L}$ can be estimated correctly with probability one by

$$\tilde{T} = \min\{|s - t|; R(S(s)) = v, R(S(t)) = w\}.$$

Thus, whenever we observe that the random walk $R \circ S$ (which can reconstruct from χ) walks from v to w, in time $\tilde{T}$ we know that also with probability one the random walk S must be at the same place when $R \circ S$ has reached w. Hence we can construct a stopping rule and stop S, whenever $R \circ S$ has walked from v to w in time $\tilde{T}$. This rule stops S always at the same place. Now, by recurrence of S with probability one there are infinitely many time intervals of length $\tilde{T}$ where $R \circ S$ walks from v to w, and thus the above rule stops S infinitely often at the same place with probability one.

At first glance it might seem that the sequence of stopping times $\tau(1), \tau(2), \ldots$ thus obtained is not $\mathcal{H}$-adapted in the above sense, since their definition involves $\tilde{T}$ which only is measurable with respect to the whole path $(S(t))_{t\in\mathbb{N}}$. On the other hand, given the scenery ξ, that is in particular given $\tilde{L}$, we are able to construct stopping times $\tau'(1), \tau'(2), \ldots$, such that $\tau'(k)$ stops the random walk $R \circ S$ when it has walked from v to w in $\tilde{L}$ steps for the k'th time. Obviously the $\tau'(k)$ are $\mathcal{H}$-adapted in the above sense. On the other hand, the sequences $\tau(1), \tau(2), \ldots$ and $\tau'(1), \tau'(2), \ldots$ are equal $\mathbb{P}$-almost surely. It follows that the sequence $\tau(1), \tau(2), \ldots$ is $\mathcal{H}$-adapted. □

Finally, we shall see that a rule that stops S infinitely often at the same place, actually is helpful to reconstruct ξ.

Lemma 5.4.4. *If we can create a stopping rule (that is an infinite sequence of $\mathcal{H}$-adapted stopping times) that stops S infinitely often at the same place, we can also reconstruct ξ restricted to any finite interval (up to equivalence) with probability one.*

Proof. We will proof this lemma by induction.

Say, we stop the random walk infinitely often in the point $z \in \mathbb{Z}$. Of course, we then know the color of z. To find out the color of $z-1$ and $z+1$ we let the random walk S run one further step (after we have stopped it in z) and read off the color of the next point. As we have infinitely many such stopping times we will eventually see both, the color of $z+1$ and the color of $z-1$ with probability one. Since we are only interested in reconstruction up to shifts and reflection of the scenery this knowledge suffices to reconstruct ξ on $[z-1, z+1]$. This is the beginning of the induction.

For the induction step assume we already have reconstructed ξ up to shifts and reflection on the interval $[z-n, z+n]$. First assume that ξ is not symmetric under reflection at z on $[z-n, z+n]$, that is $(\xi(z-1), \dots, \xi(z-n)) \neq (\xi(z+1), \dots, \xi(z+n))$. The other case will be treated similarly at the end of this proof.

First we introduce the set of all nearest neighbor paths of length $n+1$ that starting in z in the first n steps read the same color sequence as a straight walk to the right:

$$\begin{aligned}\mathcal{S}_n := \{\rho \ : \ & \{0, \dots, n+1\} \to \mathbb{Z} \, : \, \rho(0) = z, |\rho(t+1) - \rho(t)| = 1, \, \forall t = 0, \dots, n \\ & \text{and } \xi(\rho(t)) = \xi(z+t) \, \forall t = 0, \dots, n\}\end{aligned}$$

and its subset where we exclude the straight walk to the right (the straight walk to the left is automatically excluded as we have already assumed that ξ is non-symmetric with respect to reflection at z)

$$\mathcal{S}'_n := \{\rho \in \mathcal{S}_n : \rho(n) \neq z+n\}.$$

With the help of the sets $\mathcal{S}_n$ and $\mathcal{S}'_n$ we construct two measures on the space $\{0,1,2\}$:

$$\pi_n(\cdot) := \frac{1}{|\mathcal{S}_n|} \sum_{\rho \in \mathcal{S}_n} \delta_{\xi(\rho(n+1))}(\cdot)$$

and

$$\pi'_n(\cdot) := \frac{1}{|\mathcal{S}'_n|} \sum_{\rho \in \mathcal{S}'_n} \delta_{\xi(\rho(n+1))}(\cdot)$$

(if $\mathcal{S}'_n = \varnothing$ we simply set $\pi'_n \equiv 0$.)

Now the following three observations are crucial: First note that the desired color of $z+n+1$ is the only color with a higher value (probability) under π_n than under π'_n, hence

$$\xi(z+n+1) = \operatorname{supp}((\pi_n(\cdot) - \pi'_n(\cdot))^+)$$

where supp denotes the support of a function and for a real number a we write a^+ for $\sup\{a, 0\}$.

Second, observe that from knowing $\xi|[z-n, z+n]$ (which we know by our induction hypotheses), we can construct $\mathcal{S}'_n$ and therefore also calculate π'_n.

Finally, we also have an arbitrarily good approximation for π_n. Indeed, let us denote by ϑ^T the set of all times $t \leq T$ where we stop the random walk S in the point z and read and read the colors $\xi(z+i)$ in the next n steps. More precisely:

$$\vartheta^T := \{t \leq T : S(t) = z \text{ and } \xi(S(t+i)) = \xi(z+i), i = 0, \ldots, n\}.$$

Then by the strong Markov property of the stopping times and the law of large numbers

$$\tilde{\pi}_n^T(\cdot) := \frac{1}{|\vartheta^T|} \sum_{t \in \vartheta^T} \delta_{\xi(S(t+n+1))}(\cdot)$$

converges to π_n, when T becomes large. Thus with probability one

$$\lim_{T \to \infty} \operatorname{supp}((\pi_n^T - \pi_n')^+)$$

consists of precisely one element and reveals the color of $z+n+1$.

The same technique can be applied to reconstruct $\xi(z-n-1)$. To this end we simply replace $\mathcal{S}_n$ by $\overline{\mathcal{S}}_n$ defined as

$$\overline{\mathcal{S}}_n := \{\rho \quad : \quad \{0, \ldots, n+1\} \to \mathbb{Z} \,|\, \rho(0) = z, |\rho(t+1) - \rho(t)| = 1, \forall t = 0, \ldots, n \\ \text{and } \xi(\rho(t)) = \xi(z-t) \, \forall t = 0, \ldots, n\}$$

and $\mathcal{S}_n'$ by

$$\overline{\mathcal{S}}_n' := \{\rho \in \overline{\mathcal{S}}_n : \rho(n) \neq z - n\}$$

and proceed as above.

If finally, ξ restricted to $[z-n, z+n]$ is symmetric under reflection at z, the support of $(\pi_n - \pi_n')^+$ (π_n and π_n' defined as above) may either consist of one or of two elements. More precisely, it will be one-elementary, if also $\xi(z-n-1) = \xi(z+n+1)$, in which case we simply assign this color to each of the two vertices $z-n-1$ and $z+n+1$. If $\xi(z-n-1) \neq \xi(z+n+1)$ indeed

$$\operatorname{supp}((\pi_n(\cdot) - \pi_n'(\cdot))^+)$$

and also

$$\lim_{T \to \infty} \operatorname{supp}((\pi_n^T - \pi_n')^+)$$

will consist of two elements (with the notation introduced above the latter will be the "right colors" with probability one, again). As we only aim at reconstructing ξ up to translations and reflections we do not to need to care about to which of $z-n-1$ and $z+n+1$ we assign which of the two colors.

This finishes the proof of the lemma.

□

Finally, we remark that Lemma 5.4.4 implies that we can reconstruct ξ up to equivalence with probability one

Corollary 5.4.1. *If we can create a stopping rule (that is an infinite sequence of $\mathcal{H}$-adapted stopping times) that stops S infinitely often at the same place, we can also find a mapping $\mathcal{A} : \{0,1,2\}^{\mathbb{N}} \to \{0,1,2\}^{\mathbb{Z}}$ with*

$$\mathcal{A}(\chi) \sim \xi.$$

Proof. Just paste the different pieces of scenery together. □

So the above steps show that the algorithm proposed in Section 3 works in Case A. The next lemmata will show that the same holds true in Case B.

To this end one strategy would be to give the equivalent of Lemma 5.4.2 in the sense that given we know that for each $v, w \in \operatorname{Im} R$ there at least two shortest crossings of (v, w) by R, we want to test whether there are precisely two such crossings or if there are more than two. Such a test may be very difficult to find. Indeed, recall that every shortest crossing of (v, w) by R corresponds to two points $z_1, z_2 \in \mathbb{Z}$ with $R(z_1) = v$ and $R(z_2) = w$. Now it may be very hard to decide at first glance from the empirical distribution function of the waiting time between two shortest walks from v to w by $R \circ S$ whether we have two such intervals which are far apart from each other or whether we have three (or more) of them which are rather close.

To overcome this difficulty we apply another strategy. We first demonstrate how to reconstruct ξ (up to equivalence) if we know that there are $v, w \in V_3$ for which there are precisely two shortest crossings of (v, w) by R. We then see in a final step that in view of Lemma 4.1 this technique already suffices to find a general reconstruction algorithm.

The situation where for each pair $v, w \in \operatorname{Im} R$ there at least two shortest crossings of (v, w) by R can again be split into two different cases. In the first case there are $v, w \in \operatorname{Im} R$ such that there is a shortest crossing of (v, w) by R with a color sequence different of all other shortest crossing. By this we mean, that there is a shortest crossing of (v, w) by R, say the first shortest crossing, such that the sequence of colors read by R when going form v to w say, is different from the corresponding sequence of colors for all other shortest crossing of (v, w) by R. We will call such shortest crossings *distinct* as opposed to the *non-distinct* shortest crossings, where each sequence of colors read by R when following such a shortest crossing of (v, w) agrees with the sequence of colors of another shortest crossing of (v, w).

The case of distinct shortest crossings is quite similar to Case A, and we will also apply Lemma 5.4.4 for the reconstruction.

Before doing so we show that we really can find out whether there are distinct or non-distinct shortest crossings of (v, w) by R.

Lemma 5.4.5. *Assume that for each $v, w \in V_3$ there are at least two shortest crossings of (v, w) by R, then there is a test which decides with probability one whether these crossings are distinct or not.*

Proof. Note that due to the recurrence of S (and hence of $R \circ S$) the random path $R \circ S$ will follow each shortest crossing of (v, w) by R infinitely often. These "direct" passages from v to w can be determined as above, since they are the only ones happening in the "shortest observed time"

$$\tilde{T} = \min\{|s - t|; R(S(s)) = v, R(S(t)) = w\}$$

with probability one. So comparing the color record (that is the color read during such a fastest passage) of these shortest crossings form v to w, we see whether there is only one such color record or whether there are different ones. In the first case the shortest crossings are definitely non-distinct while in the latter case we can test whether the

limiting empirical distribution function of any fixed of these color records is different from the distribution function of a color record of length $\tilde{T}$, given that it is produced on one shortest crossing between two points at distance $\tilde{T}$. Exactly as in the proof of Lemma 5.4.3 we conclude that the crossings are distinct if the two distribution functions above agree, otherwise we deduce that the crossings are non-distinct. As in Lemma 5.4.3 this test works with probability one. □

Next we see that we can reconstruct ξ if we can find (v, w) with distinct shortest crossings of (v, w) by R.

Lemma 5.4.6. *Assume that there are $v, w \in V_3$ such that there are at least two shortest crossings of (v, w) by R and assume that these shortest crossings are distinct. Then we can find a mapping $\mathcal{A} : \{0, 1, 2\}^{\mathbb{N}} \to \{0, 1, 2\}^{\mathbb{Z}}$ with*

$$\mathcal{A}(\chi) \sim \xi.$$

(Note that in this case $\mathcal{A}$ will depend in general on R and (v, w)).

Proof. Recall that we call shortest crossings of (v, w) by R distinct if there is one shortest crossing, say $[s, t]$, of (v, w) by R such that the sequence of colors read by R when going from v to w say, in time $[s, t]$, is different from the corresponding sequence of colors for all other shortest crossing of (v, w) by R. Also note that by the representation of ξ as a random path R on T_3 there is an unique interval $[z_1, z_2] \subseteq \mathbb{Z}$ corresponding to this unique shortest crossing $[s, t]$ of (v, w) by R (without loss of generality $R(z_2) = w$).

Hence, whenever the random path $R \circ S$ walks from v to w in time $\tilde{T}$ and produces the color record characteristic for the unique shortest crossing $[s, t]$ of (v, w) by R we know that the random walk S is in a certain point, namely that it is in z_2. By recurrence of S this will happen infinitely often with probability one, thus we have a stopping rule which allows us to stop S infinitely often at the same point z_2 with probability one. Thus we can apply Lemma 5.4.4 together with Corollary 5.4.1 to prove the statement of the lemma. □

Next we will see what to do, if we know that there are $v, w \in V_3$ for which there are exactly two non-distinct crossings.

Lemma 5.4.7. *Assume that there $v, w \in V_3$ for which there are precisely two crossings of (v, w) by R. Then we can find a mapping $\mathcal{A} : \{0, 1, 2\}^{\mathbb{N}} \to \{0, 1, 2\}^{\mathbb{Z}}$ with*

$$\mathcal{A}(\chi) \sim \xi$$

for $\mathbb{P}$-almost all walks S and fixed R (note that $\mathcal{A}$ will depend on v and w).

Proof. Note that we only need to prove this theorem in the case where the two crossings have the same length τ and are non-distinct, otherwise there is one shortest crossing or Lemma 5.4.6 applies.

Again we will prove this lemma in two steps. In the first step we will show how to reconstruct every finite piece of scenery under the assumptions of the lemma. In the second (short) part we then will prove that this already suffices to reconstruct the whole scenery (up to equivalence).

So let us assume we know that for two fixed points $v \neq w \in V_3$ that there are precisely two crossings of (v, w) by R. By the representation R of ξ these two crossings correspond to two intervals, say $[z_1, z_2]$ and $[z_1', z_2']$ (without loss of generality $z_2 < z_1'$ with

$$|z_2 - z_1| = |z_2' - z_1'| = \tau,$$

$R(z_1) = R(z_2') = v$, $R(z_2) = R(z_1') = w$ or $R(z_1) = R(z_2') = w$, $R(z_2) = R(z_1') = v$ (other situations for the z_1, z_2, z_1', z_2' cannot occur due to our assumption that there are only two crossings of (v, w) by R), and such that $\xi(z_1 + x) = \xi(z_2' - x)$ for all $0 \leq x \leq \tau$.

Now, first of all note that we can estimate τ by

$$\tilde{T} = \min\{|s - t|; R(S(s)) = v, R(S(t)) = w\}$$

(and this estimate is correct with probability one).

Moreover, we can also give an accurate estimate for $z_1' - z_2$. Indeed, observe that the empirical distribution function of the observed walks from v to w converges. More precisely let $[s_1, t_1], [s_2, t_2], \ldots$ denote all intervals where the random path $R \circ S$ walks from v to w in $\tau = \tilde{T}$ steps, so $s_1 < t_1 < s_2 < t_2 < \ldots$ and $R(S(s_i)) = v$ and $R(S(t_i)) = w$ for all $i = 1, 2, \ldots$. Then by the law of large numbers (Glivenko-Cantelli-Lemma) the empirical distribution function of the "first renewal times"

$$\frac{1}{n-1} \sum_{i=2}^{n} \delta_{t_i - t_{i-1}}$$

converges to some distribution function $\overline{F}$ with probability one as n goes to infinity. Then with probability one $\overline{F}$ will be different from the distribution function F we would have, if there were only one crossing from v to w. Note also, as already remarked in the proof of Lemma 5.4.2 that we can actually calculate F. Denote by θ_1 the smallest t where F and $\overline{F}$ differ, that is

$$\theta := \inf\{t : F(t) \neq \overline{F}(t)\}.$$

As illustrated in Figure 1 below (where for a moment we assume that $R(z_1) = R(z_2') = w$, $R(z_2) = R(z_1') = v$) the first possible t for which $F(t)$ and $\overline{F}(t)$ can differ is exactly $z_1' - z_2 + 2\tau$. This is true because for $t = z_1' - z_2 + 2\tau$ the walk, instead of walking from z_1' to z_2' in τ steps then back again to z_1' in another τ steps to create another short crossing of (w, v) from there (which amounts in a renewal time of 2τ), may walk directly from z_1' over z_2' to z_2 and create a shortest crossing of (w, v) from there. Also, if the shortest crossing is from v to w in $t = z_1' - z_2 + 2\tau$ steps the walk may e.g. walk from z_1 to z_2' directly and create a shortest crossing of (v, w) from z_2. As these events have a positive probability to occur after $t = z_1' - z_2 + 2\tau$ steps, this shows that

$$\theta := \inf\{t : F(t) \neq \overline{F}(t)\}.$$

is a good estimator of $z_1' - z_2 + 2\tau$. Moreover $\tilde{T}$ estimates τ with probability one correctly, thus with probability one

$$\theta - 2\tilde{T} = z_1' - z_2.$$

Moreover, we can also deduce in which of the two situations $R(z_1) = R(z_2') = v$, $R(z_2) = R(z_1') = w$ or $R(z_1) = R(z_2') = w$, $R(z_2) = R(z_1') = v$ we are.

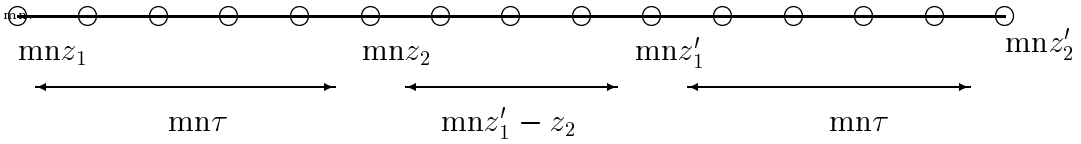


Figure 5.1: The intervals for the shortest crossings

Indeed, with probability one: if $R(z_1) = R(z_2') = v$ and $R(z_2) = R(z_1') = w$ then the only way to observe a crossing from w to v created on the interval $[z_1, z_2]$ exactly θ steps after a crossing from w to v created on the interval $[z_1', z_2']$ is by observing a shortest crossing from v to w in time steps $\tau + 1$ to 2τ. On the other hand one might well observe a from v to w created on the interval $[z_1, z_2]$ exactly θ steps after a crossing from v to w created on the interval $[z_1', z_2']$ is without observing a shortest crossing from w to v in time steps $\tau + 1$ to 2τ. Thus in the case that $R(z_1) = R(z_2') = v$ and $R(z_2) = R(z_1') = w$ we have that

$$\mathbb{P}(G_{wv}^{\theta}) < \mathbb{P}(G_{vw}^{\theta}).$$

Here

$$G_{wv}^{\theta} := \{\text{There are shortest crossing from } w \text{ to } v \text{ which are exactly } \theta \text{ steps apart and there is } \mathbf{no} \text{ shortest crossing from } v \text{ to } w \text{ in steps } \tau + 1 \text{ to } 2\tau\}$$

and

$$G_{vw}^{\theta} := \{\text{There are shortest crossing from } v \text{ to } w \text{ which are exactly } \theta \text{ steps apart and there is } \mathbf{no} \text{ shortest crossing from } w \text{ to } v \text{ in steps } \tau + 1 \text{ to } 2\tau\}.$$

If on the other hand $R(z_1) = R(z_2') = w$ and $R(z_2) = R(z_1') = v$, then we have

$$\mathbb{P}(G_{wv}^{\theta}) > \mathbb{P}(G_{vw}^{\theta}).$$

Now the probabilities $\mathbb{P}(G_{wv}^{\theta})$ and $\mathbb{P}(G_{vw}^{\theta})$ can be arbitrarily well approximated by their corresponding empirical probabilities. Hence we have a test that decides with probability one correctly in which of the two situations $R(z_1) = R(z_2') = v$, $R(z_2) = R(z_1') = w$ or $R(z_1) = R(z_2') = w$, $R(z_2) = R(z_1') = v$ we are. For the rest of this proof we will without loss of generality assume that $R(z_1) = R(z_2') = w$, $R(z_2) = R(z_1') = v$.

Next we will reconstruct $\xi|[z_2, z_1']$. To this end we stop the random walk S, whenever $R \circ S$ has walked from w to v in time $\tilde{T}$ (which will happen infinitely often with probability one, again, since S is recurrent). According to the above we then know with probability one that S is either in z_2 or in z_1', both with positive probability. Just as in the proof of Lemma 5.4.4 the empirical distribution of the first color read after $R \circ S$ has passed from w to v in time $\tilde{T}$ then reveals the color of the points neighboring v but outside $[z_1, z_2]$ and $[z_1', z_2']$.

To understand this in greater detail, first consider the Markov chain on $\{z_2, z_1'\}$, that enters z_2 after each time S has walked from z_1 to z_2 in time τ and that enters z_1' after each time S has walked from z_2' to z_1' in time τ and otherwise stays where it is. It is easy

to see, that this Markov chain has the uniform distribution (charging each of z_2', z_1' with probability $1/2$) as its invariant measure. Hence also, the empirical distribution of the colors read in the next step after each time the path $R \circ S$ has walked from w to v will converge to a distribution that assigns probability $1/4$ to each of $\xi(z_2-1), \xi(z_2+1), \xi(z_1'-1)$, and $\xi(z_1'+1)$ (of course, some of these colors will agree, in this case the probability of this color is just the sum of the above probabilities). Now note that we indeed know $\xi(z_2-1) = \xi(z_1'+1)$. Hence we are able to figure out the colors of $\xi(z_2+1)$ and $\xi(z_1'-1)$. As a matter of fact, if the limiting empirical distribution π of the colors read in the next step after each time the path $R \circ S$ has walked from w to v, satisfies $\pi(\xi(z_2-1)) = 1$, then

$$\xi(z_2+1) = \xi(z_1'-1) = \xi(z_2-1) = \xi(z_1'+1).$$

If $\pi(\xi(z_2-1)) = 3/4$, then there will be exactly one $i \in \{0,1,2\} \setminus \{\xi(z_2-1)\}$ with $\pi(i) \neq 0$ and the colors of $\xi(z_2+1)$ and $\xi(z_1'-1)$ will be this i and $\xi(z_2-1)$. If finally $\pi(\xi(z_2-1)) = 1/2$ there will be one or two i's with $i \in \{0,1,2\} \setminus \{\xi(z_2-1)\}$ and $\pi(i) \neq 0$ the color(s) of $\xi(z_2+1)$ and $\xi(z_1'-1)$ will be this i (resp. these i's). Actually, $\pi(\xi(z_2-1))$ cannot be less that $1/2$ since $\xi(z_2-1) = \xi(z_1'-1)$.

Following these ideas and the ideas already presented in the proof of Lemma 5.4.4 we are then able to reconstruct ξ inductively on $[z_2, z_1']$ (note that due to symmetry we do not need to care about whether we reconstructed ξ from z_2 to z_1' or from z_1' to z_2).

So up to now, we know $\xi|[z_1, z_2']$ (up to reflection symmetry). It remains to reconstruct ξ on $\mathbb{Z} \setminus [z_1, z_2']$. To this end recall that we are in the situation, where between every two points $v_1 \neq v_2 \in \operatorname{Im} R$ there are at least two shortest crossings of (v_1, v_2) by R and that there are exactly two crossings of (v, w) by R. Now let us take a sequence of vertices $(v_n)_{n\in\mathbb{N}_0}$ in V_3 such that such that $v_0 = w$, $d(v_n, v_{n+1}) = 1$, and such that $d(v_n, v)$ increases. Note, that then there can be at most two crossings of (v_n, v) by R. According to the above we can then reconstruct $\xi|[z_1^{(n)}, z_2^{(n)}]$, where $z_1^{(n)}, z_2^{(n)} \in \mathbb{Z}$ are associated to v_n via the representation of ξ as a random path on T_3 (more precisely the intervals $[z_1^{(n)}, z_2]$ and $[z_1', z_2^{(n)}]$ are the intervals associated to the crossings of (v, v_n) by R). As $d(v, v_n) \to \infty$ also $|z_1^{(n)}| \to \infty$ and $|z_2^{(n)}| \to \infty$. Hence we can find an algorithm that reconstructs $\xi|[a, b]$ for each finite interval $[a, b]$.

The last step is just as in Corollary 5.4.1 to concatenate the different reconstruction to reconstruct ξ up to equivalence on all of $\mathbb{Z}$.

□

The last step in proving that the algorithm proposed in Section 3 really works consists of showing that Lemma 5.4.7 implies that it works in Case B.

Lemma 5.4.8. *In Case B we can find a mapping $\mathcal{A} : \{0,1,2\}^{\mathbb{N}} \to \{0,1,2\}^{\mathbb{Z}}$ with*

$$\mathcal{A}(\chi) \sim \xi.$$

Proof. All what is left to show is, how we can get into the situation of Lemma 5.4.7, in particular, how we can guarantee the existence of two points $v, w \in V_3$ such that there are precisely two crossings of (v, w) by R.

To this end take any sequence $v_n \in \operatorname{Im} R$ such that $d(v_n, v_{n+1}) = 1$ and $d(o, v_n)$ is increasing to infinity (o the origin of T_3). Then, according to Lemma 5.4.1, the number of crossings N_n of (o, v_n) by R with probability one converges to a limit which is either 1

or 2. As we are in Case B this limit can only be 2. As N_n is always an integer this means that N_n equals 2 for all but a finitely number of n's. Hence if we apply the reconstruction algorithm proposed in Lemma 5.4.7 to $v = o$ and $w = v_n$, we will get the correct scenery with probability one for all but a finite number of n's. Thus, if we denote by $\mathcal{A}^n$ the reconstruction proposed in Lemma 5.4.7 based on the points $v = o$ and $w = v_n$, we know that

$$\mathcal{A}(\chi) := \lim_{n\to\infty} \mathcal{A}^n(\chi)$$

exists with probability one (up to equivalence) in any reasonable topology as the sequence will be essentially constant (constant for all but a finite number of n's). With probability one this limit will agree (up to equivalence) with ξ.

□

This finishes the proof of the fact that the algorithm proposed in Section 3 works.

5.5 Examples

In this section we shall discuss situations where the only assumption of Theorem 5.2.1, the transience of the representation R of the scenery ξ is satisfied and also such examples, where this assumption is violated, although the distribution of the colors is stationary and ergodic.

Before we start with these examples, let us remark that, of course, the situation where the colors are the output of an i.i.d. experiment, that is the situation where $\xi(z)$ are i.i.d. random variables for $z \in \mathbb{Z}$ and

$$\min\{\mathbb{P}(\xi(0) = 0), \mathbb{P}(\xi(0) = 1), \mathbb{P}(\xi(0) = 2)\} > 0,$$

is covered by Theorem 5.2.1. As a matter of fact, this was already shown by one of the authors in [9] and has been the starting point of the present paper.

Before presenting a large class of examples where the condition of the transience of R is satisfied, we first discuss three counterexamples, which will motivate the conditions in this main class of examples given in Theorem 5.9 below. The counterexamples will show that in a certain sense the class of distribution we give below is the largest "natural class" for Theorem 5.2.1 to hold.

The first example will be one, where R trivially cannot be transient.

Example 5.5.1. *Consider a distribution of ξ produced by the following mechanism: Take a time-homogeneous Markov chain X_n on the set of colors $\{0, 1, 2\}$ with the following transition probabilities*

$$P(X_{n+1} = 0|X_n = 0) = P(X_{n+1} = 1|X_n = 0) = P(X_{n+1} = 2|X_n = 0) = \frac{1}{3},$$

and

$$P(X_{n+1} = 0|X_n = 1) = P(X_{n+1} = 0|X_n = 2) = 1.$$

This Markov chain is irreducible and aperiodic (due to the holding in 0), and hence ergodic (even mixing of any kind). Now choose a coloring of the integers $\mathbb{Z}$ according to X_n, by,

for example, attaching the color 0 *to* $0 \in \mathbb{Z}$*, and then first coloring the positive integers* $\mathbb{Z}_+$ *according to* X_n *and then* Z_- *independently according to the same distribution. Then, of course, the distribution of the colors inherits the properties of the the Markov chain* X_n*, in particular, it admits a stationary distribution, is ergodic and mixing of any kind.*

Note however, that $\operatorname{Im} R$ *consists of 6 points only. Thus, of course,* R *as a random path on* T_3*, cannot be transient and hence the main assumption of Theorem 5.2.1 is not fulfilled.*

The above example illustrated that R might be recurrent, although all nice ergodic properties (such as stationarity and mixing properties) are fulfilled. The reason, of course, is that, as demonstrated above, R despite of fulfilling all these nice properties, may still have a finite image. The next example shows that on the other hand also, $\operatorname{Im} R$ may be infinite and R is still not transient.

Example 5.5.2. *Probably the easiest example where* $\operatorname{Im} R$ *is infinite and still not transient, is that of a one-dimensional walk. More precisely, we choose*

$$\mathbb{P}(\xi(0) = 0) = \mathbb{P}(\xi(0) = 1) = 1/2,$$

and, of course, $\mathbb{P}(\xi(0) = 2) = 0$ *and let the* $\xi(z), z \in \mathbb{Z}$ *be i.i.d. Then* R *is equivalent to a one-dimensional random walk on the integers* $\mathbb{Z}$ *without drift and holding. As, of course, such a random walk is recurrent, so is* R *and hence the main assumption of Theorem 5.2.1 is again not fulfilled.*

This example might, of course, not be too surprising, as it is well-known that a one-dimensional random walk on the integers $\mathbb{Z}$ without drift and holding is recurrent. However, we gave this example, as it is the building block of the following example, which is definitely more surprising. It basically states, that the condition of transience of R might even be violated, when $\operatorname{Im} R$ is infinite and the distribution of the colors has nice ergodic properties. Even more is true: in the example below $\operatorname{Im} R$ will be as "truly two-dimensional" as possible, in the sense that there are three infinite branches in $\operatorname{Im} R$.

Example 5.5.3. *Consider the distribution of* ξ *produced by the following random mechanism. Take the following set of words (by which we mean a sequence of colors)*

$$\mathcal{S} := \bigcup_{l \geq 0} \{(x_0, x_1, \ldots, x_{2l}) : \\ x_0 \in \{0,1\},\, x_1, \ldots, x_l \in \{0,1,2\},\, x_{l+x} = x_{l-x},\, x = 1, \ldots l\}$$

and introduce the following probability distribution π *on* $\mathcal{S}$*:*

$$\pi((x_0, x_1, \ldots, x_{2l})) = \frac{1}{2^{l+2}3^l}$$

for a $(x_0, x_1, \ldots, x_{2l}) \in \mathcal{S}$*.*

Moreover, let us choose a random scenery ξ *according to* π *in the following way. We choose two independent sequences of independent words according to* π*. Moreover, with probability* $1/2$ *we choose the staring point of the scenery to be either 0 or 1. If the starting point is* 1 *we attach the first of the two random sequences to the positive integers starting*

with 1, that is we attach the first word of, say L letters, to $1, \ldots, L \in \mathbb{N}$, after that the next word of, say L' letters, to the points $L+1, \ldots, L+L'$, and so on. After that we do the same thing for the second sequence and the non-positive integers $\mathbb{Z}_0^-$. If the starting point is 0, we attach attach the first sequence in the same way to $\mathbb{Z}_0^+$ and the second sequence to $\mathbb{Z}^-$.

Discussion of Example 5.5.3:

First note that π is indeed a probability distribution on $\mathcal{S}$. Indeed, selecting an element from $\mathcal{S}$ according to π corresponds to first choosing its length $2l+1$ (note that $\mathcal{S}$ only consists of vectors of an odd length) with probability $2^{-(l+1)}$ (which works as $\sum_{l\geq 0} 2^{-l-1} = 1$) and then selecting one of the elements of $\mathcal{S}$ of length $2l+1$ with uniform probability. Note that, as a matter of fact, there are 2×3^l different choices for $(x_0, x_1, \ldots, x_{2l}) \in \mathcal{S}$. Hence π is indeed a probability on $\mathcal{S}$.

With the help of renewal theory we now show that the sequence of colors produced by this mechanism is stationary and ergodic. Indeed, if we consider the renewal process, such that there is a renewal time, whenever a word from $\mathcal{S}$ is finished, then the greatest common divisor of these renewal times is one (since all words have odd length) and the mean renewal time is finite (which follows immediately from the definition of π). Hence it follows from renewal theory that there exists a stationary measure for the renewal times. Hence also the corresponding distribution of the colors on $\mathbb{Z}$ inherits this stationarity property.

To see that this distribution also is mixing and hence ergodic we have to understand that the shift is ergodic under the distribution induced by π. So let Θ be the right-shift on $\mathbb{Z}$. We have to prove that for any two measurable events A and B

$$\lim_{t\to\infty} \mathbb{P}(\Theta^t(B)|A) \to \mathbb{P}(B).$$

First of all observe that for every $A, B \in \sigma(\xi_i, i \in \mathbb{Z})$ the probabilities $\mathbb{P}(A)$ and $\mathbb{P}(B)$ can be arbitrarily well approximated by $\mathbb{P}(A_n)$ and $\mathbb{P}(B_n)$ where

$$A^n, B^n \in \sigma(\xi_i, i \in \{-n, \ldots, n\})$$

for some $n \in \mathbb{N}$ large enough.

Indeed,let $\varepsilon > 0$ be given. By a standard exercise in measure theory there exists an $n \in \mathbb{N}$ and an event

$$A_n(\varepsilon) \in \sigma(\xi(i), i \in \{-n, \ldots, n\})$$

such that

$$\mathbb{P}(A \Delta A_n(\varepsilon)) < \varepsilon$$

(where for two sets A, A' we denote by $A\Delta A'$ the symmetric difference between A and A').

By the same arguments, for a given set $B \in \sigma(\xi(i))$ there exists $B_n(\varepsilon) \in \sigma(\xi(i), i \in \{-n, \ldots, n\})$ such that

$$\mathbb{P}(B \Delta B_n(\varepsilon)) < \varepsilon.$$

Hence by stationarity there also exists

$$B_n^s(\varepsilon) \in \sigma(\xi(i), i \in \{-n+s, \ldots, n+s\})$$

with

$$\mathbb{P}(\Theta^s(B)\Delta B_n^s(\varepsilon)) < \varepsilon.$$

(and indeed $B_n^s(\varepsilon)) = \Theta^s(B_n(\varepsilon)))$.

Again by stationarity (we may shift the whole situation by n), it suffices to assume that

$$A_n(\varepsilon), B_n(\varepsilon) \in \sigma(\xi(i), i \in \{0, \ldots, n\})$$

for some n large enough and hence that Hence we may without loss of generality assume that

$$B_n^s(\varepsilon) \in \sigma(\xi(i), i \in \{s, \ldots, n+s\}).$$

Then for $\varepsilon \leq \frac{1}{2}\mathbb{P}(A)$ we have

$$\mathbb{P}(A_n) = \mathbb{P}(A_n \cap A) + \mathbb{P}(A_n \setminus A) \leq \mathbb{P}(A) + \frac{1}{2}\mathbb{P}(A) = \frac{3}{2}\mathbb{P}(A)$$

and therefore

$$\begin{aligned} |\mathbb{P}(\Theta^t(B)|A) - \mathbb{P}(B)| &= \frac{|\mathbb{P}(\Theta^t(B) \cap A) - \mathbb{P}(B)\mathbb{P}(A)|}{\mathbb{P}(A)} \\ &\leq \frac{|\mathbb{P}(B_n^t(\varepsilon) \cap A_n(\varepsilon)) - \mathbb{P}(B_n(\varepsilon))\mathbb{P}(A_n(\varepsilon)| + 4\varepsilon}{\frac{2}{3}\mathbb{P}(A_n)} \\ &\leq \frac{3}{2}|\mathbb{P}(\Theta^t(B_n(\varepsilon))|A_n(\varepsilon)) - \mathbb{P}(B_n(\varepsilon))| + 6\varepsilon. \end{aligned}$$

Thus if we can show that $|\mathbb{P}(\Theta^t(B_n(\varepsilon))|A_n(\varepsilon)) - \mathbb{P}(B_n(\varepsilon))| \to 0$ we also know that $|\mathbb{P}(\Theta^t(B)|A) - \mathbb{P}(B)| \to 0$.

Thus it suffices to assume that

$$A, B \in \sigma(\xi_i, i \in \{0, \ldots, n\})$$

for some $n \in \mathbb{N}$ large enough.

Moreover let us decompose $\mathbb{P}(B)$ in the following way

$$\begin{aligned} \mathbb{P}(B) = \sum_{s \geq 0} \quad & \mathbb{P}(B| \text{ the last renewal before 0 is at time } -s) \\ & \times \mathbb{P}(\text{ the last renewal before 0 is at time } -s). \end{aligned} \tag{5.5.1}$$

Similarly,

$$\begin{aligned} \mathbb{P}(\Theta^t(B)|A) = \sum_{s \geq 0} \quad & \mathbb{P}(\Theta^t(B)| \text{ the last renewal before } t \text{ is at time } t-s, A) \\ & \times \mathbb{P}(\text{ the last renewal before } t \text{ is at time } t-s|A). \end{aligned} \tag{5.5.2}$$

Now

$$\begin{aligned}
& \mathbb{P}(\Theta^t(B)|\text{ the last renewal before } t \text{ is at time } t-s, A) \\
= \; & \mathbb{P}(\Theta^t(B) \cap \text{ there is a renewal between times } n \text{ and } t| \\
& \quad \text{the last renewal before } t \text{ is at time } t-s, A) \\
& + \mathbb{P}(\Theta^t(B) \cap \text{ there is no renewal between times } n \text{ and } t| \\
& \quad \text{the last renewal before } t \text{ is at time } t-s, A) \\
= \; & \mathbb{P}(B \cap \text{ there is a renewal between times } n-t \text{ and } 0| \\
& \quad \text{the last renewal before } 0 \text{ is at time } -s) \\
& + \mathbb{P}(\Theta^t(B) \cap \text{ there is no renewal between times } n \text{ and } t| \\
& \quad \text{the last renewal before } t \text{ is at time } t-s, A) \\
= \; & \mathbb{P}(B|\text{ the last renewal before } 0 \text{ is at time } -s) \\
& - \mathbb{P}(B \cap \text{ there is no renewal between times } n-t \text{ and } 0| \\
& \quad \text{the last renewal before } 0 \text{ is at time } -s) \\
& + \mathbb{P}(\Theta^t(B) \cap \text{ there is no renewal between times } n \text{ and } t| \\
& \quad \text{the last renewal before } t \text{ is at time } t-s, A)
\end{aligned}$$

where we have used the stationarity of $\mathbb{P}$.

Now we have a finite expected renewal time implying that as $t \to \infty$

$$\begin{aligned}
& \mathbb{P}(B \cap \text{ there is no renewal between times } n-t \text{ and } 0| \\
& \quad \text{the last renewal before } 0 \text{ is at time } -s) \to 0
\end{aligned}$$

as well as

$$\begin{aligned}
& \mathbb{P}(\Theta^t(B) \cap \text{ there is no renewal between times } n \text{ and } t| \\
& \quad \text{the last renewal before } t \text{ is at time } t-s, A) \to 0.
\end{aligned}$$

This establishes equality between the first factors in (5.5.1) and (5.5.2).

For the second factors in (5.5.1) and (5.5.2) observe that due the same arguments by which we established the existence of a stationary measure

$$\mathbb{P}(\text{ the last renewal before } t \text{ is at time } t-s|A)$$

converges independently of A (by which we mean that the limit is independent of A) to a number, which actually is

$$\mathbb{P}(\text{ the last renewal before } 0 \text{ is at time } -s).$$

Hence the right hand sides of (5.5.1) and (5.5.2) converge to each other as $t \to \infty$ yielding that

$$\mathbb{P}(\Theta^t(B)|A) \to \mathbb{P}(B)$$

as $t \to \infty$. Thus the distribution of the colors induced by π is stationary, mixing and hence ergodic.

Moreover, note that all v in V_3 have a positive probability of being in $\operatorname{Im} R$. However, still R is not transient. To understand why, observe that the scenery ξ in this example

may be considered as a scenery drawn according to the distribution considered in Example 5.5.2, modified by random excursions of length $2l$. As already remarked in Example 5.5.2, R there is equivalent to a one-dimensional symmetric random walk. In terms of this random walk the excursions of length $2l$ may be interpreted as a holding. As the expected holding time is $\sum_{l\geq 0} 2l(1/2)^{l+1}$ and hence finite this holding does not spoil the recurrence of the random walk. Thus also in this example R is recurrent and therefore the condition of Theorem 5.2.1 is not fulfilled again.

At this point a little remark seems to be due.

Remark 5.5.1. *Note that although in the examples above (in particular Examples 5.5.2 and 5.5.3) R is not transient and hence our main result Theorem 5.2.1 is not applicable, this does not mean that these sceneries can not be reconstructed at all. As a matter of fact, the scenery in Example 5.5.2 has been proven to be reconstructible in [10] by completely different methods and the same might hold true for the scenery in Example 5.5.3.*

Before we give our main class of examples, let us quickly mention that there indeed are examples of two color sceneries that can be reconstructed with the help of Theorem 5.2.1.

Example 5.5.4. *Like in Example 5.5.2 the easiest example of a two color scenery that can be reconstructed with the help of Theorem 5.2.1 is that of an i.i.d. biased scenery. More precisely, we choose*

$$\mathbb{P}(\xi(0)=0)=p \quad \text{and} \quad \mathbb{P}(\xi(0)=1)=1-p,$$

(and, of course, $\mathbb{P}(\xi(0)=2)=0$) for some $p\in(0,1)$ with $p\neq 1/2$, and let the $\xi(z), z\in\mathbb{Z}$ be i.i.d. Then R is equivalent to a one-dimensional random walk on the integers $\mathbb{Z}$ with drift. As, of course, such a random walk is transient, so is R and hence the main assumption of Theorem 5.2.1 is fulfilled and thus ξ can be reconstructed.

Let us now give our main class of examples. We will avoid the troubles we had in Example 5.5.3 by assuming that ξ is generated by a hidden Markov chain on a finite state space. To also avoid the troubles we had in Example 5.5.2 we additionally have to require that ξ is "essentially tree-like". Let us define this notion first.

Definition 5.5.1. *We will call a class of sceneries essentially tree-like, if for their representation R the following holds:*

$$\{v\in V_3 : P(v\in \operatorname{Im} R)>0\}$$

consists of three distinct infinite branches. Thus, more precisely, we will call a class of sceneries essentially tree-like, if there is a vertex $v_0\in V_3$ such that

$$\{v\in V_3 : P(v\in \operatorname{Im} R)>0\}\setminus\{v_0\}$$

has three infinite connected components.

Let us first show that Definition 5.5.1 is not empty:

Example 5.5.5. *Let the random variables $\{\xi(z), z \in \mathbb{Z}\}$ be i.i.d. with*

$$P(\xi(1) = 0) > 0, P(\xi(1) = 1) > 0 \text{ and } P(\xi(1) = 2) > 0.$$

Then $\{v \in V_3 : P(v \in \operatorname{Im} R) > 0\} = V_3$ and thus the class of sceneries is essentially tree-like (one e.g. take the origin as v_0).

Remark 5.5.2. *The notion essentially tree-like should not be confused with that a fixed scenery ξ has a representation $R = R(\xi)$ that has three distinct infinite branches. Indeed, the latter is never the case, because these branches would correspond to distinct, infinite, connected subset of $\mathbb{Z}$. Obviously, there are only two such subsets.*

The counterexamples above show that the class of distribution we give below is the largest "natural class" for Theorem 5.2.1 to hold.

Theorem 5.5.1. *Consider the distribution of ξ produced by the following random mechanism. Take an aperiodic, irreducible, recurrent and stationary Markov chain $(X_n)_{n\in\mathbb{Z}}$ on a finite state space X. Let*

$$f : X \to \bigcup_{l\geq 1}\{(\zeta_1, \ldots, \zeta_l) : \zeta_i \in \{0, 1, 2\}, i = 1, \ldots, l\}$$

be a mapping from X the set of all words of finite length. Now select a scenery according to X_n and f. By this we mean, that we take a realization of X_n, and place $f(X_0)$ to the integers $0, 1, \ldots, |f(X_0)| - 1$, then we place $f(X_1)$ to the next integers and so on placing one word after the other. In the same way we color the negative integers by $(f(X_n)_{n\in\mathbb{Z}^-})$.

If then ξ is essentially tree-like, R is almost surely transient.

Example 5.5.6. *This example (the simplest one can probably give) shows that the class of defined in Theorem 5.9 above is not empty: Every i.i.d. scenery, i.e. every scenery with i.i.d. colors, falls into the class described in Theorem 5.9. Indeed we simply take $X = \{0, 1, 2\}$ and $f(x) = x$ for $x = 0, 1, 2$.. Moreover take the "independent" Markov chain on X, i.e. the Markov chain X_n with $\mathbb{P}(X_i = x) = \mathrm{P}_x \in (0, 1)$ for all $x \in X$ and $n \in \mathbb{Z}$. Then it follows immediately that the corresponding scenery is essentially tree-like and obeys the conditions of Theorem 5.9.*

Before we start to prove Theorem 5.9 let us define:

Definition 5.5.2. *In the tree T_3 let us define the ball and the sphere of radius $r > 0$ centered in some vertex $v \in V_3$, respectively., as*

$$B(v, r) := \{w \in V_3 : d(v, w) \leq r\}$$

and

$$S(v, r) := \{w \in V_3 : d(v, w) = r\}$$

Theorem 5.9 will be proved after the following lemma which justifies the notion "essentially tree-like" in the sense that the number of points that can possibly be visited by the scenery grows exponentially.

Lemma 5.5.1. *Under the conditions of Theorem 5.5.1 assume that* $\operatorname{Im} R$ *is essentially tree-like. Then there exists a constant* $\kappa > 0$ *such that for each* $v \in V_3$ *and* $r \in \mathbb{N}$ *the ball of radius* r *centered in* v*, i.e.* $B(v,r) \cap \{v \in V_3 : P(v \in \operatorname{Im} R) > 0\}$*, contains at least* $e^{\kappa r}$ *vertices.*

Proof. To show this lemma we will prove that under these conditions $\operatorname{Im} R$ is indeed an infinitely branching tree by exploiting the essential self-similarity of $\operatorname{Im} R$. By the latter we mean that given two vertices $v_1, v_2 \in \operatorname{Im} R$ such that both are, for example, colored by reading the endpoint of a word $f(x), x \in X$, then the neighborhoods of v_1 and v_2 are isomorphic.

So, if $\operatorname{Im} R$ is essentially tree-like, by definition, it will contain three disjoint, infinite branches b_1, b_2, b_3 and without loss of generality we can assume that $v_0 = o$. i.e. the split point is assumed to be the origin. For convenience let $X = \{x_0, \dots, x_l\}$, $f(x_0) = (\zeta^1, \dots, \zeta^\lambda)$ where $\zeta^i \in \{0,1,2\}$, $(i = 1, \dots, \lambda)$, and assume that the color of o is produced by reading ζ^1. This means we read the color of the origin is read in the first letter of the word belonging to x_0.

Define $L = \sum_{i=0}^{l} |f(x_i)|$.

Now $(X_n)_{n\in\mathbb{N}}$ is a stationary and recurrent Markov chain on X. This immediately implies that

$$\mathbb{P}(X_n = x_j | X_1 = x_i) > 0 \quad \text{for some } n \leq |X| = l + 1$$

and all $x_i, x_j \in X$ (because have to be able to come back to x_i from some point in X). In particular,

$$\mathbb{P}(X_n = x_i | X_1 = x_i) > 0 \quad \text{for some } n \leq |X| = l + 1$$

and all $x_i \in X$.

Recall that each x_i in X produces a word $f(x_i)$. Say, we find this word in the scenery, starting in $z_0 \in \mathbb{Z}$. The above considerations imply that we have a positive probability to see $f(x_i)$ again the latest every L steps. This implies that for all $v_1 \in \operatorname{Im} R$ there is a $v_2 \neq v_1 \in \operatorname{Im} R$ such that $d(v_1, v_2) \leq L$ and the color of v_2 as read at the same position of the same word as the color of v_1. Otherwise the random path R would return to v_1 every L steps contradicting the assumption that it is infinite. Thus for every point $v \in V_3$ every possible situation, i.e. every color read at any position of any of the words, occurs within a ball of radius L.

We will apply these considerations to the origin o. We take two auxiliary points $a_1, a_2, a_3 \in V_3$ with $a_i \in b_i$ (recall that b_i was the i'th branch) and $d(o, a_i) = 2L$ for $i = 1, 2, 3$. Applying the above shows that there are vertices $v_1, v_2, v_3 \in V_3$ with $v_i \in b_i$ and $d(v_i, a_i) \leq L$ for $i = 1, 2, 3$, such that the color of v_i is read by R at ζ^1. Obviously, $v_i \neq o$ for each i. On the other hand, the situation at v_i is the same as at o, that means, in particular, at v_i there are three different infinite branches for all i.

Continuing inductively yields the desired result. □

Now we are ready to prove that R is transient.

Proof of Theorem 5.5.1. The basic idea of the proof will be to show that, if R were recurrent, then for any fixed vertex $v \in \operatorname{Im} R$ the distance $d(v, R(n))$ would be stochastically bounded below by a random walk with positive drift. This, of course, is a contradiction, since a random walk with positive drift is transient and thus R would also have to be transient.

The way to derive this contradiction is to first analyze a Markov chain that is obtained form R by stopping it, when is has moved to the next point a certain distance apart from the previous point. We will see that this Markov chain has the tendency to move away from the points it has previously visited. By comparing this chain to a (transient) random walk with drift on the line we see that it is transient as well. But R and this chain are never far apart from each other. Hence also R is transient.

More precisely in what follows take $d_1 < d_2 \in \mathbb{N}$ and typically we will think of d_1 as being much smaller than d_2 (a more precise description will be given below).

Again let $X = \{x_0, \dots, x_l\}$ and $f(x_0) = (\zeta^1, \dots, \zeta^\lambda)$ where $\zeta^i \in \{0, 1, 2\}$, $(i = 1, \dots, \lambda)$, and take $L = \sum_{i=0}^{l} |f(x_i)|$. Consider the following Markov chains induced by R, X_n, and ξ. Take

$$\Omega = V_3 \times X \times \{1, \dots, \max_{i \in \{0,\dots,l\}} |f(x_i)|\}$$

and

$$\tilde{X}_n := (X_{n,1}, X_{n,2}, X_{n,3}) := (R(n), w, k)$$

where $w \in X$ is the word where $R(n)$ is read, and k is the position in w where $R(n)$ is read. Note that $\tilde{X}_n$ is again a Markov chain. Moreover, introduce a sequence of stopping times $(t_n)_{n \in \mathbb{Z}}$ such that $t_0 = 0$ and

$$t_n := \inf\{t > t_{n-1}, d(R(t), R(t_{n-1})) = d_2\}, \qquad n \in \mathbb{N}$$

and

$$t_{-n} := \sup\{t < t_{-n+1}, d(R(t), R(t_{-n+1})) = d_2\}, \qquad n \in \mathbb{N}.$$

Let $\tilde{Y}_t = (Y_{t,1}, Y_{t,2}, Y_{t,3})$ where for $t \in [t_n, t_{n+1})$ $Y_{t,1} = X_{t_n,1}$, $Y_{t,2} = X_{t,2}$ and $Y_{t,3} = X_{t,3}$. Note that $\tilde{Y}_t$ inherits the Markov property from $\tilde{X}_t$. We will prove that $\tilde{Y}_t$ is transient. This also implies that $R(t)$ is transient since

$$d(R(t), Y_{t,1}) \leq d_2$$

and d_2 is independent of t.

Let for any $n \in \mathbb{Z}$ denote $y_n := Y_{t_n,1}$. The transience of y_n will be proved by showing that, if y_n were recurrent, for any fixed vertex $v_0 \in V_3$ the increment $d(v_0, y_{n+1}) - d(v_0, y_n)$ would be stochastically bounded below by the increment of a random walk with positive drift. That random walk can go by a distance d_2 to the left with probability $1/4$ and go by $d_2/2$ to the right with probability $3/4$. Let $n_0 \in \mathbb{N}$ be fixed and without loss of generality in the following we will assume that the color of y_{n_0} (that is $\varphi(y_{n_0})$) is read at ζ_l (that is, it is equal to ζ_l and read at the last position of $f(x_0)$).

Let us assume that $4|d_2$ and take the unique point z_0 such that $d(y_{n_0}, z_0) = \frac{1}{4}d_2$ and $d(v_0, z_0) = d(v_0, y_{n_0}) - \frac{1}{4}d_2$ (see Figure 2 below).

The set Ξ

Denote

$$\Xi := \{v \in S(y_{n_0}, d_2) \setminus S(z_0, \frac{3}{4}d_2)\}.$$

We have that for such $v \in \Xi$

$$d(v, v_0) \geq d(y_{n_0}, v_0) + d_2/2.$$

To understand this and the following, one should keep in mind, that Figure 2 illustrates the tree geometry in T_3. Thus a path from v_0 to $v \in \Xi$, has to follow the path from v_0 to y_{n_0} at least until z_0. On the other hand for

$$v \in S(y_{n_0}, d_2) \setminus \Xi$$

we have by the triangle inequality

$$d(v, v_0) \geq d(y_{n_0}, v_0) - d_2.$$

Hence, if we can, for example, show that

$$\mathbb{P}(y_{n_0+1} \in S(y_{n_0}, d_2) \cap S(z_0, \frac{3}{4}d_2)) \leq \frac{1}{4}$$

we are done, since then $d(y_{n_0+1}, v_0)$ can be smaller than $d(y_{n_0}, v_0)$ by d_2 with probability at most $1/4$. On the other hand, it will increase by $d_2/2$. Thus, the increment $d(y_{n_0+1}, v_0) - d(y_{n_0}, v_0)$ is bounded below by the increment of a random walk, which at each step can do the following: go to the left by a distance d_2 with probability $1/4$ or go to the right by a distance $d_2/2$ with probability $3/4$. We thus get:

$$\mathbb{E}(d(y_{n_0+1}, v_0) - d(y_{n_0}, v_0)) \geq \left(-\frac{1}{4} + \frac{1}{2} \cdot \frac{3}{4}\right) d_2 = \frac{1}{8}d_2.$$

In order to bound $\mathbb{P}(y_{n_0+1} \in S(y_{n_0}, d_2) \cap S(z_0, \frac{3}{4}d_2))$ we will use Lemma 5.5.1 above. The idea is that Lemma 5.5.1 tells us that in a neighborhood of y_{n_0} there are many points with the same color as y_{n_0} and, where this color is read at the same position of the same position as y_{n_0}. Hence the situation in any of these points is the same as in y_{n_0}. On the other hand (as we will prove) most of these points belong to disjoint "first exit regimes" of $S(y_{n_0}, d_2)$. Since the situation in all of the points is the same, the probabilities to leave the circle $S(y_{n_0}, d_2)$ via a particular of the corresponding segments are about the same. In particular, the probability to leave $S(y_{n_0}, d_2)$ via $S(y_{n_0}, d_2) \setminus \Xi$ cannot be too large.

More precisely, in order to bound $\mathbb{P}(y_{n_0+1} \in S(y_{n_0}, d_2) \cap S(z_0, \frac{3}{4}d_2))$ we consider the ball $B(y_{n_0}, d_1)$ with $d_1 \ll d_2$. As we have shown in Lemma 5.5.1 this ball contains exponentially many points (in d_1). By the same argument one can show that there exists a sphere inside $B(y_{n_0}, d_1)$ containing $M := e^{\kappa d_1}$ many points (for some $\kappa > 0$) the color of which is read at the same point ζ^λ. Let us call these points $v_1, \ldots, v_M$.

Now let us assume, that our proposition was wrong and R was recurrent. Then for any given $\varepsilon > 0$ we could choose d_2 large enough such that with probability larger than $1 - \varepsilon$ the random path R will visit each point inside a ball of radius d_1 around its starting point before first exiting a ball of radius $d_2 - d_1$ around the starting point.

Instead of bounding now

$$\mathbb{P}(y_{n_0+1} \in S(y_{n_0}, d_2) \cap S(z_0, \frac{3}{4}d_2))$$

we will rather bound

$$\Delta := \max_{z \in S(y_{n_0}, \frac{1}{4}d_2)} \mathbb{P}(y_{n_0+1} \in S(y_{n_0}, d_2) \cap S(z, \frac{3}{4}d_2))$$

This means, we will bound the maximum of the probabilities to first exit $S(y_{n_0}, d_2)$ via a segment similar to $S(y_{n_0}, d_2) \cap S(z, \frac{3}{4}d_2)$. In order to do so, we will bound

$$\Delta' := \max_{z \in S(y_{n_0}, \frac{1}{4}d_2)} \mathbb{P}(y'_{n_0+1} \in S(y_{n_0}, d_2 - d_1) \cap S(z, \frac{3}{4}d_2 - d_1)),$$

where y'_{n_0+1} is the vertex where the path R first exits $S(y_{n_0}, d_2 - d_1)$ when starting in y_{n_0}. Note that both maxima are actually attained and let $z_{\max}$ be such that

$$\Delta' := \mathbb{P}(y'_{n_0+1} \in S(y_{n_0}, d_2 - d_1) \cap S(z_{\max}, \frac{3}{4}d_2 - d_1)).$$

Note that the stretch $\overline{y_{n_0} z_{\max}}$ is uniquely defined by its corresponding sequence of colors and its starting point y_{n_0}. Since the situation is completely identical in each of the points $v_1, \ldots, v_M$ and in y_{n_0} we can find $z_1, \ldots, z_M$ that correspond to $z_{\max}$ when we replace y_{n_0} by $v_1, \ldots, v_M$. More precisely the points $z_1, \ldots, z_M$ are defined by the fact that the stretch $\overline{v_i z_i}$ has the same color sequence as $\overline{y_{n_0} z_{\max}}$, (for each $i = 1, \ldots, M$).

The points $v_1, \ldots v_M$ and $y_1, \ldots y_M$

Obviously, the probability of first exiting $S(v_i, d_2 - d_1)$ via $S(v_i, d_2 - d_1) \cap S(z_i, \frac{3}{4}d_2 - d_1)$ when starting in v_i also equals Δ'. Let us call z_i and z_j equivalent, if the stretches $\overline{v_i z_i}$ and $\overline{v_j z_j}$ intersect inside $B(y_{n_0}, d_1)$. Note that the stretches $\overline{v_i z_i}$ and $\overline{v_j z_j}$ always decrease there distance to $S(y_{n_0}, d_2)$. Hence, if z_i and z_j are not equivalent, then the stretches $\overline{v_i z_i}$ and $\overline{v_j z_j}$ do not intersect at all and z_i and z_j and $S(v_i, d_2 - d_1) \cap S(z_i, \frac{3}{4}d_2 - d_1)$ and $S(v_j, d_2 - d_1) \cap S(z_j, \frac{3}{4}d_2 - d_1)$ have distance at least $2d_2 - 2d_1$. Moreover, observe that for a given $i \in \{1, \ldots, M\}$ at most $2d_1$ of the $\{z_1, \ldots, z_M\}$ can be equivalent to z_i. This follows from the fact that two stretches $\overline{v_i z_i}$ and $\overline{v_j z_j}$ can only intersect at different "times" (that is their distance from v_i – or v_j, respectively – must be different). As each stretch has to leave $B(y_{n_0}, d_1)$ after at most $2d_1$ steps, there, indeed can be at most $2d_1$ of the z_j equivalent to z_i. As M is an exponential in d_1, but $2d_1$ is of course only linear, we can have as many non-equivalent z_i's as we wish. We will denote the number of non-equivalent z_i's by $M' \in \mathbb{N}$.

Now, once we are first exiting $S(y_{n_0}, d_2 - d_1)$ via $S(y_{n_0}, d_2 - d_1) \cap S(z_{\max}, \frac{3}{4}d_2 - d_1)$ we are at distance at d_1 from $S(y_{n_0}, d_2)$. Define

$$\mathcal{Z} := \{z \in V_3 : d(y_{n_0}, z) = \frac{1}{4}d_2, d(z_{\max}, z) = d_1\}.$$

By construction of d_1 the probability of visiting each point in a ball of radius d_1 before leaving a ball of radius $d_2 - d_1$ for the first time is at least $1 - \varepsilon$. Hence, once we are in $S(y_{n_0}, d_2 - d_1) \cap S(z_{\max}, \frac{3}{4}d_2 - d_1)$ the probability to leave $S(y_{n_0}, d_2 - d_1)$ for the first time via

$$\bigcup_{z \in \mathcal{Z}} S(y_{n_0}, d_2) \cap S(z, \frac{3}{4}d_2)$$

is at least $1 - \varepsilon$. Since ε can be arbitrarily small this shows in particular that

$$\Delta(1 - \varepsilon) \leq \Delta'.$$

Moreover introduce a random variable N, which is equal to $i \in \{1, \ldots M'\}$ if $R(t)$ visits $S(v_i, d_2 - d_1) \cap S(z_i, \frac{3}{4}d_2 - d_1)$ before visiting any of the $S(v_j, d_2 - d_1) \cap S(z_j, \frac{3}{4}d_2 - d_1), j \neq i$

and exiting $S(y_{n_0}, d_2)$ for the first time. If $R(t)$ does not visit any of the $S(v_i, d_2 - d_1) \cap S(z_i, \frac{3}{4}d_2 - d_1)$ before exiting $S(y_{n_0}, d_2)$ for the first time we set N equal to 0. Now, once $R(t)$ is in v_i the probability that $P(N_i = i)$ is at least Δ' (by construction of the v_i and z_i. On the other hand $R(t)$ hits with probability at least $1 - \varepsilon$ any of the v_i before exiting $S(y_{n_0}, d_2 - d_1)$ for the first time, i.e. before the value of N can be determined. This implies

$$1 \geq \sum_{i=1}^{M'} P(N = i) \geq M'(1 - \varepsilon)\Delta'.$$

Since M' can be made as large as we wish, Δ' and hence also Δ are as small as we wish, for example less than $1/4$. Therefore, under the assumption that R is recurrent,

$$\mathbb{E}(d(y_{n_0+1}, y_{n_0-1})) \geq \frac{9}{8}d_2.$$

But this implies that $Y_{n,1}$ is transient. Indeed, by the tree structure of T_3 the distance of $Y_{n,1}$ to any fixed point v' is stochastically larger than the distance to the origin of a a homogeneous random walk Z_n on the integer line $\mathbb{Z}$ with jump length d_2 and

$$\mathbb{P}(Z_{n+1} = z + d_2 | Z_n = z) = 1 - \mathbb{P}(Z_{n+1} = z - d_2 | Z_n = z) = \frac{3}{4}$$

for all $n \in \mathbb{N}$ and $z \in \mathbb{Z}$. This is a contradiction.

Hence $Y_{n,1}$ is transient and therefore also R is transient, which is what we claimed. □

The next theorem basically states that under the conditions of Theorem 5.5.1 the random path R (the representation of ξ as a random path on T_3) is not only transient but also has positive speed with probability exponentially close to one. Theorem 5.5.2 will be the basis for an exceptionally good test for distinguishing two sceneries.

Theorem 5.5.2. *Under the conditions of Theorem 5.5.1 (in particular we also assume that ξ is essentially tree-like) there exist constants $c_0, c_1, c_2 > 0$ such that for all $n \in \mathbb{N}$ and every fixed $v_0 \in V_3$*

$$\mathbb{P}(\min\{d(v_0, R(n)), d(R(-n), v_0)\} \leq c_2 n) \leq c_0 e^{-c_1 n}.$$

Proof. The proof is intrinsically related to the proof of Theorem 5.5.1. We will make use of the notations introduced there. Let $Y'_m := X_{t_m,1}$. Let us (for a moment) assume that for any fixed vertex v_0 the distance $(d(v_0, Y'_m))_m$ can be stochastically bounded below by the distance to the origin of a random walk on the line with drift. Hence the existence of constants $c'_0, c'_1, c'_2 > 0$ such that

$$\mathbb{P}(\min\{d(v_0, Y'_m) \leq c'_2 m) \leq c'_0 e^{-c'_1 m} \tag{5.5.3}$$

follows immediately from an exponential estimate for this dominating random walk.

In order to conclude the desired result from (5.5.3), we need to understand that there exist constants $c_3, c_4 > 0$ such that

$$\mathbb{P}(t_n \geq c_3 n) \leq e^{-c_4 n} \tag{5.5.4}$$

(recall that t_n was the n'th stopping time). But this follows form decomposing t_n into

$$t_n = (t_n - t_{n-1}) + (t_{n-1} - t_{n-2}) + \ldots + (t_1 - t_0).$$

By the Markov property of $\tilde{X}_t$ and $\tilde{Y}_t$ the random variables $(t_m - t_{m-1})$ are stochastically independent. Also, since the state space X of the Markov chain X_n is finite all these $(t_m - t_{m-1})$ can be stochastically dominated from above by a random variable $T_{\max}$ (describing the exit time of $\tilde{X}_t$ from a circle of radius d_2 when starting in the state with the longest exit time). Now $T_{\max}$ has a finite moment generating function

$$\mathbb{E}e^{\theta T_{\max}} < \infty. \tag{5.5.5}$$

This follows, since X is finite. Therefore we can find constants $L_1 < \infty$ and $p > 0$ such that independent of where we start with $R(n)$ we have hit the sphere of radius d_2 centered in the starting point after L_1 steps with probability at least $p > 0$. This implies (5.5.5) and therefore by large deviation estimates also (5.5.4).

It remains to show that for any fixed vertex v_0 the distance $(d(v_0, Y'_m))_m$ can be stochastically bounded below by the distance to the origin of a random walk on the line with drift.

Lemma 5.5.2 below states that the maximum probability for R to ever reach a point at distance d from v_0 when starting in v_0 tends to zero as d goes to infinity. More precisely, it says that for every (t_0, v_0, w_0, k_0) with

$$\mathbb{P}(\tilde{X}_{t_0} = (v_0, w_0, k_0)) > 0$$

we have that

$$\lim_{d\to\infty} \max_{v:d(v,v_0)=d} \mathbb{P}(\exists s > 0 : R(t_0 + s) = v | \tilde{X}_{t_0} = (v_0, w_0, k_0)) = 0. \tag{5.5.6}$$

Assuming that Lemma 5.5.2 below is true, we choose d_2 large enough such that for $d = \frac{1}{4}d_2$ we have

$$\max_{v:d(v,v_0)=d} \mathbb{P}(\exists s > 0 : R(t_0 + s) = v | \tilde{X}_{t_0} = (v_0, w_0, k_0)) \leq \frac{1}{10}.$$

Let Z_n be a random walk on the line that in each step either steps to the left by d_0 (this happens with probability $\frac{1}{10}$) or it steps to the right by $\frac{1}{2}d_0$ (this happens with probability $\frac{9}{10}$). Then $(d(v_0, Y'_m))_m$ can be stochastically bounded below by the distance of Z_n to the origin. Indeed, for fixed n_0 let z_0 denote the unique vertex in V_3 at distance d from y_{n_0} and at distance $3d$ from y_{n_0-1}. If, after time t_{n_0} the random path R never visits z_0 the point y_{n_0+1} is at distance at least $2d = \frac{d_2}{2}$ from y_{n_0}. This happens with probability $9/10$. By the tree structure of T_3 this shows that indeed $(d(v_0, Y'_m))_m$ can be stochastically bounded below by the distance to the origin of a random walk on the line with drift.

Thus the statement of the theorem follows. □

It remains to show

Lemma 5.5.2. *For all (t_0, v_0, w_0, k_0) with*

$$\mathbb{P}(\tilde{X}_{t_0} = (v_0, w_0, k_0)) > 0$$

we have that

$$\lim_{d\to\infty} \max_{v:d(v,v_0)=d} \mathbb{P}(\exists s > 0 : R(t_0 + s) = v | \tilde{X}_{t_0} = (v_0, w_0, k_0)) = 0.$$

Proof. Assume Lemma 5.5.2 was wrong. Then there exists a sequence $v_0, v_1, v_2 \ldots \in V_3$ such that v_i, v_{i+1} are neighbors (for all $i \in \mathbb{N}_0$) but $v_i \neq v_{i+2}$ (for all $i \in \mathbb{N}_0$) such that

$$\lim_{n\to\infty} \mathbb{P}(\exists s > 0 : R(t_0 + s) = v_n | \tilde{X}_{t_0} = (v_0, w_0, k_0)) > 0.$$

Hence for

$$A := \bigcap_{n\in\mathbb{N}} \{\exists s > 0 : R(t_0 + s) = v_n\}$$

it holds

$$\mathbb{P}(A|\tilde{X}_{t_0} = (v_0, w_0, k_0)) > 0.$$

But this implies that

$$\mathbb{P}(A|\tilde{X}_{t_0} = (v_0, w_0, k_0)) = 1.$$

Indeed, otherwise $\mathbb{P}(A|\tilde{X}_{t_0} = (v_0, w_0, k_0)) \in (0, 1)$ and hence already

$$\mathbb{P}(\bigcap_{n\leq N} \{\exists s > 0 : R(t_0 + s) = v_n\}|\tilde{X}_{t_0} = (v_0, w_0, k_0)) =: p \in (0, 1)$$

for some $N \in \mathbb{N}$. But then we could stop R the first time that $(\tilde{X}_{t,2}, \tilde{X}_{t,3}) = (w_0, k_0)$ (which happens infinitely often since the original Markov chain (X_n) on X is recurrent), after $R(t)$ has visited v_N. This first segment of points has probability at most p. But at $(\tilde{X}_{t,2}, \tilde{X}_{t,3}) = (w_0, k_0)$ the situation is the same as in (v_0, w_0, k_0), so again we find a finite segment of the sequence of v_i's with probability at most p and so on. This shows that if $\mathbb{P}(A|\tilde{X}_{t_0} = (v_0, w_0, k_0)) \in (0, 1)$ we already know that $\mathbb{P}(A|\tilde{X}_{t_0} = (v_0, w_0, k_0)) = 0$.

Moreover there only can be one sequence $v_0, v_1, v_2 \ldots \in V_3$ such that v_i, v_{i+1} are neighbors (for all $i \in \mathbb{N}_0$) but $v_i \neq v_{i+2}$ (for all $i \in \mathbb{N}_0$) and $\mathbb{P}(A|\tilde{X}_{t_0} = (v_0, w_0, k_0)) = 1$. This follows from the tree structure of T_3. Indeed, if there were two such sequences they eventually needed to be on disjoint branches of $T3$. But then R in order to visit both sequences with probability one, needs to visit the bifurcation point of the two sequences infinitely often (contradicting its transience).

Eventually we show that also $\mathbb{P}(A|\tilde{X}_{t_0} = (v_0, w_0, k_0)) = 1$ cannot hold. Again we exploit that R is essentially tree-like.

Let us call a sequence essentially k-periodic, if it is k-periodic up to a finite number of elements. By definition the quasi-period k of an essentially k-periodic sequence is the period of the periodic part that sequence. Now recall that the colors $\varphi(v_0), \varphi(v_1), \varphi(v_2), \ldots$ are produced by a Markov chain on *finite* state space X. Since moreover each word $f(x), x \in X$ had a finite length, there is just a finite number of possible positions for the second and third coordinate of the process $\tilde{X}_t$. Since the situation is the same, whenever the second and third coordinate of the process $\tilde{X}_t$ are in the same point and there is just one sequence $v_0, v_1, v_2 \ldots \in V_3$ satisfying the above conditions, the sequence $\varphi(v_0), \varphi(v_1), \varphi(v_2), \ldots$ is essentially periodic. Since we do not care for a finite number of elements we can and will assume that it is periodic.

Now R is essentially tree like. This means in particular that we can find a point $v_0' \in V_3$ with the following properties.

- $v_0' \notin \{v_0, v_1, v_2, \ldots\}$.
- $\mathbb{P}(\exists s > 0 : R(t_0 + s) = v_0'|\tilde{X}_{t_0} = (v_0, w_0, k_0)) > 0$.

- $\varphi(v_0) = \varphi(v_0')$ and moreover the color of v_0 and v_0' are read in the same word at the same position.
- $d(v_0', \{v_0, v_1, v_2, \ldots\}) \geq 3L$, where $L := \sum_{x \in X} |f(x)|$. This latter condition ensures that for all but possibly the first L colors the sequence $\varphi(v_0'), \varphi(v_1'), \varphi(v_2'), \ldots$ follows the same period as $\varphi(v_0), \varphi(v_1), \varphi(v_2), \ldots$.

Let us take such a v_0'. Since the situation in v_0' is completely identical to the situation in v_0, there exists a sequence $v_0', v_1', v_2', \ldots$ such that v_i', v_{i+1}' are neighbors (for all $i \in \mathbb{N}_0$) but $v_i' \neq v_{i+2}'$ (for all $i \in \mathbb{N}_0$) such that

$$\lim_{n \to \infty} \mathbb{P}(\exists s > 0 : R(t_0 + s) = v_n' | \tilde{X}_{t_0} = (v_0', w_0, k_0)) = 1.$$

But then the sequences $v_0, v_1, v_2 \ldots$ and $v_0', v_1', v_2', \ldots$ have to merge. Otherwise (since v_0' van be reached from v_0 with positive probability) from v_0 there would be two disjoint sequences that are both visited with probability one in contradiction to what was shown above. Say the sequences $v_0, v_1, v_2 \ldots$ and $v_0', v_1', v_2', \ldots$ merge in $v_n \in \{v_0, v_1, v_2 \ldots\}$. Then for some n' and all $j \in \mathbb{N}$

$$v_{n+j} = v_{n'+j}'$$

and also

$$\varphi(v_{n+j}) = \varphi(v_{n'+j}').$$

However,

$$\varphi(v_{n-1}) \neq \varphi(v_{n'-1}'),$$

since v_n can only have one neighbor of the color $\varphi(v_{n-1})$. Hence $n \neq n'$. Hence by the above the sequence $\varphi(v_0'), \varphi(v_1'), \varphi(v_2'), \ldots$ has quasi-period $n - n'$. On the other hand, since we have chosen $d(v_0', \{v_0, v_1, v_2, \ldots\}) \geq 3L$, the quasi-periodic sequence $\varphi(v_0'), \varphi(v_1'), \varphi(v_2'), \ldots$ in $n' - 1$ is already in its "periodic part". But then its period is the same as that of $\varphi(v_0), \varphi(v_1), \varphi(v_2), \ldots$, in particular

$$\varphi(v_{n-1}) = \varphi(v_{n'-1}'),$$

This is a contradiction. Hence the lemma is true.

□

Theorem 5.5.2 is extremely useful when we try to attack one of the original problem of this area, that is the scenery distinguishing problem, where we have to tell from the color record on which of two sceneries the this color record has been produced. Of course, in principal, already Theorem 5.5.1 shows that we can distinguish two sceneries drawn independently and at random from a distribution satisfying the conditions of Theorem 5.5.1. On the other hand, for all possible applications this test is not practicable. We now show that, indeed, as a consequence of Theorem 5.5.2 there is a practicable test based on very little information that works "exponentially well".

More precisely, we show that given any scenery ξ randomly drawn from a distribution satisfying the conditions of Theorems 5.5.1 and 5.5.2 and another scenery η of which we know nothing at all, there is a test that works exponentially well for a set of sceneries with probability exponentially close to one. Here the notions "exponentially well" and "with probability exponentially close to one" stand for the following. Given that we know

$\chi|[0,n]$ (so the first $n+1$ observations) and, for example, two points in the representation of ξ, namely

$$v = R(m^+) \quad \text{where} \quad m^+ := \min\{m \in \mathbb{N} : d(o, R(m)) \geq n^{1/3}\}$$

and

$$w = R(m^-) \quad \text{where} \quad m^- := \max\{m \in \mathbb{Z}_- : d(o, R(m)) \geq n^{1/3}\}$$

(where, as above, $R(\cdot) = R(\xi, \cdot)$ is the representation of ξ on T_3) we can find a test, which for a subset of ξ's of probability larger than $1 - k_0 e^{-k_1 n^{1/3}}$ has failure probability less than $e^{-k_2 n^{1/3}}$. Let us formalize this in a theorem:

Theorem 5.5.3. *Let ξ be randomly drawn from a distribution satisfying the conditions of Theorems 5.5.1 and 5.5.2 and let η be another scenery of which we know nothing at all. Assume that we know $\chi|[0,n]$,*

$$v = R(m^+) \quad \textit{where} \quad m^+ := \min\{m \in \mathbb{N} : d(o, R(m)) \geq n^{1/3}\},$$

and

$$w = R(m^-) \quad \textit{where} \quad m^- := \max\{m \in \mathbb{Z}_- : d(o, R(m)) \geq n^{1/3}\},$$

(where, as above, $R(\cdot) = R(\xi, \cdot)$ is the representation of ξ on T_3). Then we can find a test

$$T : \{0,1,2\}^{n+1} \times V_3 \times V_3 \to \{\xi, \eta\}$$

constants $k_0, k_1, k_2, k_3 > 0$ and a set $\overline{\Xi} \subseteq \{0,1,2\}^{\mathbb{Z}}$ with

$$\mathbb{P}(\overline{\Xi}) \geq 1 - k_0 e^{-k_1 n^{1/3}}$$

such that, whenever $\xi \in \overline{\Xi}$

$$\mathbb{P}(T(\chi|[0,n], v, w) = \xi | \, \chi \textit{ has been produced on } \eta) = 0.$$

and

$$\mathbb{P}(T(\chi|[0,n], v, w) = \eta | \, \chi \textit{ has been produced on } \xi) \leq k_2 e^{-k_3 n^{1/6}}.$$

Remark 5.5.3. *Note that we are able to improve the known tests for scenery distinguishing in the following important features:*

1. *ξ can be drawn from a large class of distributions admitting correlations between the color of different sites.*
2. *η can be arbitrary.*
3. *No knowledge is required about η.*
4. *Only very limited knowledge about ξ is required*

Proof of Theorem 5.5.3. For fixed η we propose is the following test T:

Whenever $R \circ S$ reaches $v = v(\xi)$ or $w(\xi)$ within $[0, n]$ we say that the scenery is ξ, otherwise we say that it is η.

$\overline{\overline{\Xi}}$ will be the set of sceneries ξ such that we have a fair chance to see v and w in the first n observations and that $v(\xi)$ and $w(\xi)$ are different from any $v(\eta)$ and $w(\eta)$. Formally we define the following set of sceneries. For a scenery ξ let

$$m^+ := m^+(\xi) := \min\{m \in \mathbb{N} : d(o, R(m)) \geq n^{1/3}\}$$

and

$$m^- := m^+(\xi) := \max\{m \in \mathbb{Z}_- : d(o, R(m)) \geq n^{1/3}\}$$

then for some constant $\kappa > 0$ (note that R depends on ξ).

$$\overline{\overline{\Xi}} := \{\xi : \max m^+, |m^-| \leq \kappa n^{1/3}, R(i) \notin \{v(\eta), w(\eta)\} \text{ for } i = -n, -n+1, \ldots, n\}. \quad (5.5.7)$$

Here $R(i)$ is the images of the scenery ξ in the lattice point $i \in \mathbb{Z}$. Then for κ large enough and some constants $k_0, k_1 > 0$ it holds

$$\mathbb{P}(\overline{\overline{\Xi}}) \geq 1 - k_0 e^{-k_1 n^{1/3}}. \quad (5.5.8)$$

Indeed, applying Theorem 5.5.2 to the origin $v_0 = o$ shows that

$$\mathbb{P}(\{\xi : \max m^+, |m^-| \geq \kappa n^{1/3}\}) \leq k_0' e^{-k_1' n^{1/3}} \quad (5.5.9)$$

for κ large enough and some constants $k_0', k_1' > 0$. On the other hand, trivially

$$\mathbb{P}(R(i) \notin \{v(\eta), w(\eta)\}) = 1$$

for $i = -n^{1/3} + 1, \ldots, n^{1/3} - 1$ and any ξ. Moreover, applying Theorem 5.5.2 to the $v_0 = v(\eta)$ gives

$$\mathbb{P}(R(i) = v(\eta)) \leq k_0'' e^{-k_1'' n^{1/3}}$$

for $|i| \geq n^{1/3}$ and similarly

$$\mathbb{P}(R(i) = w(\eta)) \leq k_0'' e^{-k_1'' n^{1/3}}$$

for $|i| \geq n^{1/3}$ for some constants $k_0'', k_1'' > 0$. So altogether

$$\begin{aligned} \mathbb{P}(\xi : \exists i \in \{-n, -n+1, \ldots, n\} : R(i) \in \{v(\eta), w(\eta)\}) &\leq 2n k_0'' e^{-k_1'' n^{1/3}} \\ &\leq k_0''' e^{-k_1''' n^{1/3}} \end{aligned}$$

for some constants $k_0''', k_1''' > 0$. Together with (5.5.9) this implies (5.5.8).

If now $\xi \in \overline{\overline{\Xi}}$, then indeed

$$\mathbb{P}(T(\chi|[0, n], v, w) = \xi \,|\, \chi \text{ has been produced on } \eta) = 0,$$

since $T(\chi|[0, n], v, w) = \xi$ if and only if we read $v(\xi)$ or $w(\xi)$ and by construction this cannot happen on η. On the other hand $T(\chi|[0, n], v, w) = \eta$ while χ has been produced

on ξ can only happen, if the random walk S does not reach neither m^- nor m^+ in $[0,n]$. Hence for $\xi \in \overline{\overline{\Xi}}$

$$\begin{aligned} & \mathbb{P}(T(\chi|[0,n],v,w) = \eta | \, \chi \text{ has been produced on } \xi) \\ = \; & \mathbb{P}(S \text{ does not reach neither } m^- \text{ nor } m^+ \text{ in } [0,n]) \\ = \; & \mathbb{P}(|S(i)| \leq \kappa n^{1/3} \text{ for all } i = 0, \ldots, n) \end{aligned}$$

Now indeed there are positive constants $k_2, k_3 > 0$ such that

$$\mathbb{P}(|S(i)| \leq \kappa n^{1/3} \text{ for all } i = 0, \ldots, n) \leq k_2 e^{-k_2 n^{1/3}}.$$

To see why, just observe that by the local Central Limit Theorem, for each time interval $I[t_0, t_1]$ of length $n^{2/3}$ we have

$$\mathbb{P}(|S(i)| \leq \kappa n^{1/3} \text{ for all } i \in I || S(t_0)| \leq \kappa n^{1/3}) \leq k_4 < 1$$

for a positive constant $k_4 > 0$. Since there are $n^{1/3}$ disjoint intervals of length $n^{2/3}$ in $[0,n]$ this gives (by conditioning)

$$\begin{aligned} & \mathbb{P}(|S(i)| \leq \kappa n^{1/3} \text{ for all } i = 0, \ldots, n) \\ = \; & \mathbb{P}(|S(i)| \leq \kappa n^{1/3} \text{ for all } i = 0, \ldots, n^{2/3}) \times \\ & \quad \times \mathbb{P}(|S(i)| \leq \kappa n^{1/3} \text{ for all } i = n^{2/3}+1, \ldots 2n^{2/3} || S(n^{2/3})| \leq \kappa n^{1/3}) \times \ldots \\ \leq \; & k_4^{n^{1/3}} \leq k_2 e^{-k_2 n^{1/3}}. \end{aligned}$$

This finishes the proof. □

References

[1] I. Benjamini, H. Kesten; *Distinguishing sceneries by observing the sceneries along a random walk path*, J. d'Anal. Math **69**, 97–135 (1996)

[2] W. Th. F. den Hollander, M. Keane; *Ergodic properties of color records*, Physica **138A**, 183–193 (1986)

[3] S.A. Kalikow; *$T - T^{-1}$ transformation is not loosely Bernoulli*, Ann. Math. **115**, 393-409 (1982)

[4] H. Kesten; *Detecting a single defect form observing the scenery along a random walk path*, Ito's stochastic Calculus and Probability Theory, Springer, Tokyo, 171–183 (1996)

[5] H. Kesten; *Distinguishing and reconstructing sceneries from observations from random walk paths*, Microsurveys in discrete probability (D. Aldous and J. Propp, eds.), DIMACS Series in Discrete Mathematics and Theoretical Computer Sciences **41**, 75–83 (1998)

[6] E. Lindenstrauss; *Indistinguishable Sceneries*, Random Struct. Alg. **14**, 71–86 (1999)

[7] M. Löwe, H. Matzinger; *Scenery Reconstrution in two dimensions with many colors*, Annals of Applied Probability 12, 1322–1347 (2002)

[8] M. Löwe, H. Matzinger, F. Merkl; *Reconstructing a multicolor random scenery seen along a random walk with bounded jumps*, submitted;

[9] H. Matzinger; *Reconstructing a 3-color scenery by observing it along a simple random walk*, Random Structures Algorithms 15, 196–207 (1999).

[10] H. Matzinger; *Reconstruction of a one dimensional scenery seen along the path of a random walk with holding*, Ph. D. Thesis, Cornell University (1999)

[11] F. Spitzer; *Principles of random walk*, Van Nostrand, London (1964)

Chapter 6

Information recovery from a randomly mixed up message-text

(submitted)

By Jüri Lember and Heinrich Matzinger

This paper is concerned with finding a fingerprint of a sequence. As input data one uses the sequence which has been randomly mixed up by observing it along a random walk path. A sequence containing order $\exp(n)$ bits receives a fingerprint with roughly n bits information. The fingerprint is characteristic for the original sequence. With high probability the fingerprint depends only on the initial sequence, but not on the random walk path.
[1]

6.1 Introduction and Result

6.1.1 The information recovery problem

Let $\xi : \mathbb{Z} \to \{0, 1\}$ designate a double-infinite message-text with 2 letters. Such a coloring of the integers is also called a (2-color) scenery. Let $S = \{S(t)\}_{t\in\mathbb{N}}$ be a recurrent random walk on $\mathbb{Z}$ starting at the origin. In this paper we allow the random walk S to jump, i.e. $P(|S(t+1) - S(t)| > 1) > 0$. We use S to mix up the message-text ξ. For this we assume that ξ is observed along the path of S: At each point in time t, one observes $\chi(t) := \xi(S(t))$. Thus, χ designates the mixed up message-text, which is also the scenery ξ seen along the path of S.
The *information recovery problem* can be described as follows: observing only one path realization of the process χ, can one retrieve a certain amount of information contained in ξ?
A special case of the information recovery problem is when one tries to reconstruct the whole ξ. This problem is called *the scenery reconstruction problem* . In many cases being able to reconstruct a finite quantity of the information contained in ξ, already implies that one can reconstruct all of ξ. This paper is concerned with the information recovery

[1] *MSC 2000 subject classification:* Primary 60K37, Secondary 60G10, 60J75.
Key words: Scenery reconstruction, jumps, stationary processes, random walk, ergodic theory.

problem in the context of a 2-color scenery seen along a random walk with jumps. The methods which exist so far seem useless for this case: Matzinger's reconstruction methods [Mat99a, Mat00] do not work when the random walk may jump. Furthermore, it seems impossible to recycle the method of Matzinger, Merkl and Löwe [LMM01] for the 2-color case with jumps. The reason is that their method, requires more than 2-colors. Hence, the fundamentally new approach presented in this paper.

6.1.2 Main assumptions

Let us explain the assumptions which remain valid throughout this paper:

- $\xi = \{\xi(z)\}_{z\in\mathbb{Z}}$ is a collection of i.i.d. Bernoulli variables with parameter 1/2. The path realization $\xi : z \mapsto \xi(z)$ is the scenery from which we want to recover some information. Often the realization of the process $\{\xi(z)\}_{z\in\mathbb{Z}}$ is also denoted by ψ.

- $S = \{S(t)\}_{t\in\mathbb{N}}$ is a symmetric recurrent random walk starting at the origin, i.e. $P(S(0) = 0) = 1$. We assume that S has bounded jump length $L < \infty$, where

$$L := \max\{z | P(S(1) - S(0) = z) > 0\}.$$

  We also assume that S has positive probability to visit any point in $\mathbb{Z}$, i.e. for any $z \in \mathbb{Z}$ there exists $t \in \mathbb{N}$, such that $P(S(t) = z) > 0$.

- ξ and S are independent.

- $m = m(n)$ designates a natural number depending on n, so that:

$$\boxed{\tfrac{1}{4}\exp\left(\tfrac{\alpha n}{\ln n}\right) \le m < \exp(2n)}$$

  where α is a positive constant not depending on n.

- For all $t \in \mathbb{N}$, let $\chi(t) := \xi(S(t))$. Let

$$\chi := (\chi(0), \chi(1), \ldots)$$

  designate the observations made by the random walk S of the random scenery ξ. Hence χ corresponds to the scenery ξ seen along the path of the random walk S.

We need also a few notations:

- For every $k \in \mathbb{N}$, let $\xi_0^k := (\xi(0), \xi(1), \ldots, \xi(k))$ and let $\xi_0^{-k} := (\xi(0), \xi(-1), \ldots, \xi(-k))$.

- Let $f : D \to I$ be a map. For a subset $E \subseteq D$ we shall write $f|E$ for the restriction of f to the set E.

  Thus, when $[a, b] \in \mathbb{Z}$ is an integer interval and ξ is a scenery, then $\xi|[a, b]$ stands for the vector $(\xi(a), \ldots, \xi(b))$. We also write ξ_a^b for $\xi|[a, b]$ and ψ_a^b for $\psi|[a, b]$. The notation $\chi_0^{m^2} := (\chi(0), \chi(1), \chi(2), \ldots, \chi(m^2))$ is often used.

- Let $a = (a_1, \ldots, a_N)$, $b = (b_1, \ldots, b_{N+1})$ be two vectors with length N and $N+1$, respectively. We write $a \sqsubseteq b$, if

$$a \in \{(b_1, \ldots, b_N), (b_2, \ldots b_{N+1})\}.$$

  Thus, $a \sqsubseteq b$ holds if a can be obtained from b by "removing the first or the last element".

6.1.3 Main result

The 2-color scenery reconstruction problem for a random walk with jumps is solved in two phases:

1. Given a finite portion of the observations χ only, one proves that it is possible to reconstruct a certain amount of information contained in the underlying scenery ξ.

2. If one can reconstruct a certain amount of information, then the whole scenery ξ can a.s. be reconstructed. This is proven in the second phase.

This paper solves the first of the two problems above. Imagine that we want to transmit the word ξ_0^m. During transmission the lector head gets crazy and starts moving around on ξ following the path of a random walk. At time m^2, the lector head has reached the point m. Can we now, given only the mixed up information $\chi_0^{m^2}$, retrieve any information about the underlying code ξ_0^m? The main result of this paper theorem 6.1.1, shows that with high probability a certain amount of the information contained in ξ_0^m can be retrieved from the mixed up information $\chi_0^{m^2}$. This is the fingerprint of ξ_0^m, refered to in the abstract. Here is the main result of this paper:

Theorem 6.1.1. *There exists a constant $c > 0$ not depending on n such that :*
For every $n > 0$ big enough, there exist two maps

$$g : \{0,1\}^{m+1} \to \{0,1\}^{n^2+1}$$
$$\hat{g} : \{0,1\}^{m^2+1} \to \{0,1\}^{n^2}$$

and an event $E^n_{\text{cell\_OK}} \in \sigma(\xi(z) | z \in [-cm, cm])$ such that all the following holds:

1) $P(E^n_{\text{cell\_OK}}) \to 1$ *when* $n \to \infty$.

2) *For any scenery $\psi \in E^n_{\text{cell\_OK}}$, we have:*

$$P\left(\hat{g}(\chi_0^{m^2}) \sqsubseteq g(\xi_0^m) \,\middle|\, S(m^2) = m, \xi = \psi\right) > 3/4.$$

3) *$g(\xi_0^m)$ is a random vector with (n^2+1) components which are i.i.d. Bernoulli variables with parameter* $1/2$.

The mapping g can be interpreted as a coding that compresses the information contained in ξ_0^m; the mapping $\hat{g}$ can be interpreted as a decoder that reads the information $g(\xi_0^m)$ from the mixed-up observations $\chi_0^{m^2+1}$. The vector $g(\xi_0^m)$ is the fingerprint of ξ_0^m mentioned in the abstract. We call it the *g-information*. The function $\hat{g}$ will be referred

to as the *g-information reconstruction algorithm.*
Let us explain the content of the above theorem more in detail. The event:

$$\left\{\hat{g}(\chi_0^{m^2}) \sqsubseteq g(\xi_0^m)\right\}$$

is the event that $\hat{g}$ reconstructs the information $g(\xi_0^m)$ correctly (up to the first or last bit), based on the observations $\chi_0^{m^2}$. The probability that $\hat{g}$ reconstructs $g(\xi_0^m)$ correctly is large conditional on the event $E^n_{1,\mathrm{S}} := \{S(m^2) = m\}$. The event $E^n_{1,\mathrm{S}}$ is needed to make sure the random walk S visits the entire ξ_0^m up to time m^2. Obviously, if S does not visit ξ_0^m, we can not reconstruct $g(\xi_0^m)$.
Condition **3**) of our main theorem ensures that the content of the reconstructed information is large enough.
The reconstruction of the *g-information* works with high probability, but conditional on the event that the scenery is nicely behaved. The event $E^n_{\text{cell\_OK}}$ that the scenery is nicely behaved, has large probability for big n. Suppose ψ is a (non-random) scenery such that $E^n_{\text{cell\_OK}}$ holds whenever $\xi|[-cm, cm] = \psi|[-cm, cm]$. In this situation we write $\psi \in E^n_{\text{cell\_OK}}$. In a sense, $E^n_{\text{cell\_OK}}$ contains " typical" (pieces of) sceneries. These are the sceneries, for which the g-information reconstruction algorithm works with high probability.

6.1.4 History and related problems

A coloring of the integers $\xi : \mathbb{Z} \to \{0, 1, \ldots, C-1\}$ is called a C-color scenery. The *scenery reconstruction problem* is: given only one path realization of $\{\chi(t)\}_{t\in\mathbb{N}}$, can we a.s. reconstruct ξ? In other words, does one path realization of χ a.s. uniquely determine ξ? In general, it does not: in many cases it is not possible to distinguish a scenery from a shifted one. Furthermore, Lindenstrauss proved [Lin99] the existence of sceneries which can not be reconstructed. However, one can reconstruct "typical" sceneries: Matzinger takes ξ randomly, independent of S and shows that one can reconstruct a.s. the scenery up to shift and reflection. In [Mat00] and [Mat99a], he proves this for 2-color sceneries observed along the path of simple random walk or a simple random walk with holding. In [Mat99b], he reconstructs 3-color i.i.d. sceneries observed along a simple random walk path. The two cases require very different methods.
The scenery reconstruction problem varies greatly in difficulty depending on the number of colors and the properties of the random walk. In general, when there are less colors and the random walk is allowed to jump, the problem gets more difficult. Kesten [Kes98] noticed, that Matzinger's reconstruction methods [Mat99a] and [Mat00] do not work when the random walk is allowed to jump. Matzinger, Merkl and Loewe [LMM01] showed that it is possible to reconstruct a.s. a scenery seen along the path of a random walk with jumps, provided the scenery contains enough colors. However, with more colors the system is completely differently behaved. This implies that the method of Matzinger, Merkl and Loewe is not useful for the 2-color case with jumps. For a well-written overview of the scenery reconstruction and scenery distinguishing areas, we recommend Kesten's review paper [Kes98]. Scenery reconstruction belongs to the field which investigates the properties of a color record obtained by observing a random media along the path of a stochastic process. The TT^{-1}-problem as studied by Kalikow [Kal82] is one motivation. The ergodic properties of observations have been investigated by Keane and den Hollander

[KdH86], den Hollander [dH88], den Hollander and Steiff [dHS97] and Heicklen, Hoffman and Rudolph [HHR00].
A related problem is the *scenery distinguishing problem*. It can be described as follows: let ψ^a and ψ^b be two non-equivalent sceneries which are known to us. Assume that we are only given one realization of the observations $\chi := \psi \circ S$, where $\psi \in \{\psi^a, \psi^b\}$. Can we a.s. find out whether ψ is equal to ψ^a or ψ^b? If yes, we say that the sceneries ψ^a and ψ^b are distinguishable. Kesten and Benjamini [BK96] considered the case where the sceneries ψ^a and ψ^b are drawn randomly. They take ψ^a to be an i.i.d. scenery which is independent of ψ^b. In this setting, they prove that almost every couple of sceneries is distinguishable even in the two dimensional case and with only 2 colors. Before that Howard [How97] had shown that any two periodical non-equivalent sceneries are distinguishable; he also showed that periodical sceneries which differ only in one element can be distinguished [How96]. The problem of distinguishing two sceneries which differ only in one element is called the *single defect detection problem*. Kesten [Kes96] showed that one can a.s. detect single defects in the case of 5-color i.i.d. sceneries. A generalization of the scenery distinguishing problem is the *scenery distinguishing problem for error-corrupted observations*. The error process $\nu_t, t \in \mathbb{N}$ is supposed to be a sequence of i.i.d. Bernoulli variables with parameter strictly smaller than $1/2$. Furthermore, it is assumed that the error process $\nu_t, t \in \mathbb{N}$ is independent of the (random) sceneries and the random walk. Instead of observing $\chi = \psi \circ S$ we see the error-corrupted observations $\hat{\chi}$. The observations with errors satisfy: $\hat{\chi}_t = \chi_t$ if and only if $\nu_t = 0$. (In other words, $\nu_t = 1$ signals an error at time t.) Knowing ψ^a and ψ^b, can we a.s. decide if $\psi = \psi^a$ or if $\psi = \psi^b$ based on one path realization of the process $\hat{\chi}$ only?
Another closely related question is the *Harris-Keane coin tossing problem*. Here $\{\chi(t)\}_{t\in\mathbb{N}}$ designates a color record obtained in one of the following two ways:

1. A coin with sides 0 and 1 is tossed independently infinitely many times;

2. The same coin as in case 1. is flipped, except at renewal times when a coin with a different bias is flipped. Typical renewal times are the passage times at the origin of a recurrent random walk on $\mathbb{Z}$ or $\mathbb{Z}^2$.

The question is whether it is a.s. possible to determine if χ is drawn according to **1.** or **2.**. For this we suppose that we are given only one path realization of $\{\chi(t)\}_{t\in\mathbb{N}}$. Let u_n denote the probability of a renewal at time n. Harris and Keane [HK97] showed that if $\sum_{n=1}^{\infty} u_n^2 = \infty$, then it is a.s. possible to determine if the observations are drawn according to **1.** or **2.**. They prove that this is not possible if $\sum_{n=1}^{\infty} u_n^2 < \infty$ and θ is small enough. Levin, Pemantle and Peres [LPP01] showed that θ is important for the distinguishing. They prove the existence of a phase transition: there exists a critical parameter θ_c such that for $|\theta| > \theta_c$ the cases 1. and 2. can be distinguished a.s., whilst for $|\theta| < \theta_c$ this is not possible.
A generalization [LPP01] of the Harris-Keane coin tossing problem consists in replacing the renewal times by stopping times. For this we use a scenery: $\psi \in \{0,1\}^{\mathbb{N}}$. Every-time t, the random walk is at a location z, where $\psi(z) = 1$ we throw a coin with a different bias form the coin used to generate the i.i.d. sequence. This generalization of the Harris-Keane coin tossing problem, is very similar to the scenery distinghuishing and reconstruction problems in the presence of random errors. Matzinger and Rolles [MRc] showed that almost every random scenery seen with random errors can be reconstructed

a.s. when it contains a lot of colors. However, their method cannot be used for the case of error corrupted 2-color sceneries.
The question of Kesten whether one can detect a single defect in 2-color sceneries lead Matzinger in [Mat99b, Mat00, Mat99a] to investigate the scenery reconstruction problem. Later Kesten [Kes98] asked, whether one can also reconstruct two dimensional random sceneries. Loewe and Matzinger [LM02] give a positive answer provided the scenery contains many colors. Another question was formulated first by Den Hollander: to which extent can sceneries be reconstructed when they are not i.i.d. in distribution. Loewe and Matzinger [LM99] characterize those distributions for which Matzinger's 3-color reconstruction works. Yet another problem comes from Benjamini: Is it possible to reconstruct a finite piece of a scenery close to the origin in polynomial time? We take for this polynomially many observations in the length of the piece we try to reconstruct. Matzinger and Rolles [MRb, MRa] provide a positive answer. One of the recent developements in this field, is due to Levin and Peres [LP02]. They prove that every scenery which has only finite many one's can a.s. be reconstructed up to shift or reflection when seen along the path of a symmetric random walk. They prove this result in the more general frame of stochastic sceneries. A *stochastic scenery* is a map $\eta : \mathbb{Z} \to \mathcal{P}(\mathbb{R})$, where $\mathcal{P}(\mathbb{R})$ is the set of all probability measures on $\mathbb{R}$. The observations are generated as follows : if at time t the random walk is at z, then a random variable with distribution $\eta(z)$ is observed. Hence, at time t, we observe $\chi(t)$, where: $\mathcal{L}(\chi(t)|S(t) = z) = \eta(z)$.

6.1.5 Organization of the paper

In order to explain the main ideas behind the g-information reconstruction algorithm, we first consider a simplified example in subsection 6.1.6. In this example, ξ is a 3-color i.i.d. scenery instead of a 2-color scenery. The 2's are pretty rare in the scenery ξ: $P(\xi(z) = 2)$ is of negative exponential order in n. The one's and zero's have equal probability: $P(\xi(z) = 0) = P(\xi(z) = 1)$. The (random) locations $\bar{z}_i$ of the 2's in ξ are called signal carriers. For each signal carrier $\bar{z}_i$, we define the *frequency of ones* at $\bar{z}_i$. The frequency of one's at $\bar{z}_i$ is a weighted average of ξ in the neighborhood of $\bar{z}_i$. The g-information $g(\xi_0^m)$ if a function of the different frequencies of ones of the signal carriers which are located in the interval $[0, m]$. The vector of frequencies works as a fingerprint for ξ_0^m. The reading of this fingerprint works as follows: Typically, the signal carriers are apart from each other by a distance of order $o(e^n)$. Suppose that S visits a signal carrier. Before moving to the next signal carrier, it returns to the same signal carrier many times with high probability. By doing this, S generates many 2's in the observations at short distance from each other. This implies: when in the observations we see a cluster of 2's, there is a good reason to believe that they all correspond to the same 2 in the underlying scenery. In this manner we can determine many return times of S to the same signal carrier. This enables us to make inference about ξ in the neighborhood of that signal carrier. In particular, we can precisely estimate the frequencies of ones of the different signal carriers visited. This allows us to estimate $g(\xi_0^n)$. The estimator $\hat{g}$ is the desired decoder. The details are explained in Subsection 6.1.6. However, it is important to note, that between this simplified example and our general case there is only one difference: the signal carriers. In the general case we can no longer rely on the 2's and the signal carriers need to be constructed in a different manner. Everything else – from the definition of g and $\hat{g}$ up to the proof that the g-information reconstruction algorithm works with high

probability – is exactly the same. (Note that the solution to our information recovery problem in the simplified 3-color case requires only five pages!)
For the general case with a 2-color scenery and a jumping random walk, the main difficulty consists in the elaboration of the signal carriers. In Section 2, we define many concepts which are subsequently used for the definition of the signal carriers. Also there, some technical results connected to the signal carriers are proved. The signal carriers are defined in Section 3.
The main goal of the paper is to prove that the g-reconstruction algorithm works with high probability (i.e. that the estimator $\hat{g}$ is precise). For this, we define two sets of events: the random walk dependent events and the scenery dependent event. All these events describe typical behavior of S or ξ. In Section 3, we define the scenery dependent events and prove that they have high probability. In Section 4 the same is done for the events that depend on S.
In section 5, we prove that if all these events hold, then the g-information reconstruction algorithm works, i.e. the event

$$E^n_{\text{g\_works}} := \{\hat{g}(\chi_0^{m^2}) \sqsubseteq g(\xi_0^m)\}$$

holds. The results of Section 3 and Section 4 then guarantee that the g-information reconstruction algorithm works with high probability. This finishes the proof of Theorem 6.1.1.

6.1.6 3-color example

In this subsection, we solve the scenery reconstruction problem in a simplified 3-color case. We do not change the assumptions on S.

Setup

Recall that ξ_0^m and $\chi_0^{m^2}$ denote the piece of scenery $\xi|[0,m]$ and the first m observations $\chi|[0,m]$, respectively. Recall also that we want to construct two functions $g : \{0,1\}^{m+1} \to \{0,1\}^{n^2+1}$ and $\hat{g} : \{0,1\}^{m^2+1} \to \{0,1\}^{n^2}$ such that

1) with high probability

$$P\left(\hat{g}(\chi_0^{m^2}) \sqsubseteq g(\xi_0^m) \,\middle|\, S(m^2) = m\right).$$

2) $g(\xi_0^m)$ is i.i.d. binary vector where the components are Bernoulli random variables with parameter $\frac{1}{2}$.

In other words, **1)** states that, with high probability, we can reconstruct $g(\xi_0^m)$ from the observations, provided that random walk S goes in m^2 steps from 0 to m. (Remember that $\hat{g}(\chi_0^{m^2}) \sqsubseteq g(\xi_0^m)$ means that $\hat{g}(\chi_0^{m^2})$ and $g(\xi_0^m)$ are equal up to one bit.) Thus the function $\hat{g}$ represents a "reconstruction algorithm" which tries to reconstruct the information $g(\xi_0^m)$. (For the simplified case, we prove that the g-information can be reconstructed with high probability when $S(m^2) = m$. This differs slightly from the main case where we prove that that the reconstruction of the g-information works with high conditional probability, conditional on the scenery. This is a detail of no great importance.)

Since this is not yet the real case in which we are interested in this paper, during the present subsection we will not be very formal. For this subsection only, let us assume that the scenery ξ has three colors instead of two. Moreover, we assume that $\{\xi(z)\}$ satisfies all of the following three conditions:

a) $\{\xi(z) : z \in \mathbb{Z}\}$ are i.i.d. variables with state space $\{0, 1, 2\}$,

b) $\exp(n/\ln n) \leq 1/P(\xi(0) = 2) \leq \exp(n))$,

c) $P(\xi(0) = 0) = P(\xi(0) = 1)$.

We define $m = n^{2.5}(1/P(\xi(0) = 2))$. Because of **b)** this means

$$n^{2.5} \exp(n/\ln n) \leq m(n) \leq n^{2.5} \exp(n).$$

The so defined scenery distribution is very similar to our usual scenery except that sometimes (quite rarely) there appear also 2's in this scenery.
We now introduce some necessary definitions.

Let $\bar{z}_i$ denote the i-th place in $[0, \infty)$ where we have a 2 in ξ. Thus $\bar{z}_1 := \min\{z \geq 0 | \xi(z) = 2\}$, $\bar{z}_{i+1} := \min\{z > \bar{z}_i | \xi(z) = 2\}$. We make the convention that $\bar{z}_0$ is the last location before zero where we have a 2 in ξ. For a negative integer $i < 0$, $\bar{z}_i$ designates the $i + 1$-th point before 0 where we have a 2 in ξ. The random variables $\bar{z}_i$-s are called *signal carriers*. For each signal carrier, $\bar{z}_i$, we define the *frequency of ones* at $\bar{z}_i$. By this we mean the (conditional on ξ) probability to see 1 exactly after $e^{n^{0.1}}$ observations having been at $\bar{z}_i$. We denote that conditional probability by $h(\bar{z}_i)$ and will also write $h(i)$ for it. Formally:

$$h(i) := h(\bar{z}_i) := P\Big(\xi(S(e^{n^{0.1}}) + \bar{z}_i) = 1 \Big| \xi\Big).$$

It is easy to see that the frequency of ones is equal to a weighted average of the scenery in a neighborhood of radius $Le^{n^{0.1}}$ of the point $\bar{z}_i$. That is $h(i)$ is equal to:

$$h(i) := \sum_{\substack{z \in [-Le^{n^{0.1}}, Le^{n^{0.1}}] \\ z \neq \bar{z}_i}} \xi(z) P\big(S(e^{n^{0.1}}) + \bar{z}_i = z\big) \tag{6.1.1}$$

(Of course this formula to hold assumes that there are no other two's in $[\bar{z}_i - Le^{n^{0.1}}, \bar{z}_i + Le^{n^{0.1}}]$ except the two at $\bar{z}_i$. This is very likely to hold, see event $E^n_{6\_2}$ below).
Let

$$g_i(\xi_0^m) := I_{[0,0.5)}(h(i)).$$

We now define some events that describe the typical behavior of ξ.

* Let $E^n_{6\_2}$ denote the event that in $[0, m]$ all the signal carriers are further apart than $\exp(n/(2 \ln n))$ from each other as well as from the points 0 and m. By the definition of $P(\xi(i) = 2)$, the event $P(E^n_{6\_2}) \to 1$ as $n \to \infty$.

* Let $E^n_{1\_2}$ be the event that in $[0, m]$ there are more than $n^2 + 1$ signal carrier points. Because of the definition of m, $P(E^n_{1\_2}) \to 1$ as $n \to \infty$.

When $E^n_{1\_2}$ and $E^n_{6\_2}$ both hold, we define $g(\xi_0^m)$ in the following way:

$$g\left(\xi_0^m\right) := \left(g_1\left(\xi_0^m\right), g_2\left(\xi_0^m\right), g_3\left(\xi_0^m\right), \ldots, g_{n^2+1}\left(\xi_0^m\right)\right)$$

Conditional on $E^n_{1\_2} \cap E^n_{6\_2}$ we get that $g(\xi^m)$ is an i.i.d. random vector with the components being Bernoulli variables with parameter $1/2$. Here the parameter $1/2$ follows simply by symmetry of our definition [to be precise, $P(g_i(\xi^m_i) = 1) = 1/2 - P(h(i) = 1/2)$, but we disregard this small error term in this example] and the independence follows from the fact that the scenery is i.i.d. [indeed, $g_i(\xi^m_0)$ depends only on the scenery in a radius $Le^{n^{0.1}}$ of the point $\bar{z}_i$ and, due to E_{6_2}, the points $\bar{z}_i$ are further apart than $\exp(\frac{n}{2\ln n}) > L\exp(n^{0.1})$]. Hence, with almost no effort we get that when $E^n_{1\_2}$ and $E^n_{6\_2}$ both hold, then condition **2)** is satisfied. To be complete, we have to define the function g such that **2)** holds also outside $E^n_{1\_2} \cap E^n_{6\_2}$. We actually are not interested in g outside $E^n_{1\_2} \cap E^n_{6\_2}$ - it would be enough that we reconstruct g on $E^n_{1\_2} \cap E^n_{6\_2}$. Therefore, extend g in any possible way, so that $g(\xi^m_0)$ depends only on ξ^m_0 and its component are i.i.d.

$\hat{g}$-algorithm

We show, how to construct a map $\hat{g} : \{0,1\}^{n^2} \mapsto \{0,1\}^n$ and an event $E^n_{OK} \in \sigma(\xi)$ such that $P(E^n_{OK})$ is close to 1 and for each scenery belonging to E^n_{OK} the probability

$$P\Big(\hat{g}(\chi^{m^2}_0) \sqsubseteq g(\xi^m_0) | S(m^2) = m\Big) \tag{6.1.2}$$

is also high. Note, when the scenery ξ is fixed, then the probability (6.1.2) depends on S. The construction of $\hat{g}$ consists of several steps. The first step is the estimation of the frequency of one's $h(i)$. Note: due to E_{6_2} we have that in the region of our interest we can assume that all the signal carriers are further apart form each other than $\exp(n/(2\ln n))$. In this case we have that all the 2's observed in a time interval of length $e^{n^{0.3}}$ must come from the same signal carrier. We will thus take time intervals T of length $e^{n^{0.3}}$ to estimate the frequency of one's.
Let $T = [t_1, t_2]$ be a (non-random) time interval such that $t_2 - t_1 = e^{n^{0.3}}$. Assume that during time T the random walk is close to the signal carrier $\bar{z}_i$. Then every time we see a 2 during T this gives us a stopping time which stops the random walk at $\bar{z}_i$. We can now use these stopping times to get a very precise estimate of $h(i)$. In order to obtain the independence (which makes proofs easier), we do not take all the 2's which we observe during T. Instead we take the 2's apart by at least $e^{n^{0.1}}$ from each other.
To be more formal, let us now give a few definitions:

***** Let $\nu_{t_1}(1)$ denote the first time $t > t_1$ that we observe a 2 in the observations χ after time t_1. Let $\nu_{t_1}(k+1)$ be the first time after time $\nu_{t_1}(k) + e^{n^{0.1}}$ that we observe a 2 in the observations χ. Thus $\nu_{t_1}(k+1)$ is equal to $\min\{t | \chi(t) = 2, t \geq \nu_{t_1}(k) + e^{n^{0.1}}\}$. We say that T is such that we can significantly estimate the frequency of one's for T, if there are more than $e^{n^{0.2}}$ stopping times $\nu_{t_1}(k)$ during T. In other words, we say that we can significantly estimate the frequency of one's for T, if and only if $\nu_{t_1}(e^{n^{0.2}}) \leq t_2 - e^{n^{0.1}}$.

***** Let $\hat{X}_{t_1}(k)$ designate the Bernoulli variable which is equal to one if and only if $\chi(\nu_{t_1}(k) + e^{n^{0.1}}) = 1$. When $\nu_{t_1}(e^{n^{0.2}}) \leq t_2 - e^{n^{0.1}}$ we define $\hat{h}_T$ the estimated frequency of one's during T in the following obvious way:

$$\hat{h}_T := \frac{1}{e^{n^{0.2}}} \sum_{k=1}^{e^{n^{0.2}}} \hat{X}_{t_1}(k).$$

Suppose we can significantly estimate the frequency of one's for T. Assume $E^n_{6\_2} \cap E_{1\_2}$ hold. Then all the stopping times $\nu_{t_1}(e^{n^{0.2}})$ stop the random walk S at one signal carrier, say $\bar{z}_i$. Because of the strong Markov property of S we get then that, conditional on ξ, the variables $X_{t_1}(k)$ are i.i.d. with expectations h_i. Now use the Höffding inequality to see

$$P(|\hat{h}_T - h(i)| > e^{-n^{0.2}/4}) \leq \exp(-(2e^{n^{0.2}/2})).$$

Hence, with high probability, $\hat{h}_T$ is a precise estimate for $h(i)$.
The obtained preciseness of $\hat{h}_T$ is of the great importance. Namely, it is of smaller order than the typical variation of $h(i)$. In other words, with high probability $|h(i) - h(j)|$ is of much bigger order than $\exp(-n^{0.2}/4)$, $i \neq j$. To see this, consider (6.1.1). Note that, for each z, $\mu_i(z) := P(S(e^{n^{0.1}}) + \bar{z}_i = z)$ is constant, and, conditional under the event that in the radius of $L\exp(n^{0.1})$ are no more 2's in the scenery than $\bar{z}_i$, we have that $\xi(\bar{z}_i + z)$ are iid Bernoulli variables with parameter $\frac{1}{2}$. Hence

$$Var[h(i)] \leq \sum_{[-Le^{n^{0.1}}, Le^{n^{0.1}}]} \frac{1}{4}\Big(\mu_{0.2}(z)\Big)^2.$$

Since our random walk is symmetric we get that $\sum_{z\in[-Le^{n^{0.1}}, Le^{n^{0.1}}]} \frac{1}{4}\left(\mu_{0.2}(z)\right)^2$ is equal to $1/4$ times the probability that the random walk is back at the origin after $2e^{n^{0.1}}$ time. By the central local theorem that probability is of order $e^{-n^{0.1}/2}$. This is much bigger than the order of the precision of the estimation of the frequencies of one's, $e^{-n^{0.2}/4}$. Since $h(i)$ is approximately normal, it is possible to show that with high probability all frequencies $h(0), h(1), \ldots, h(n^2+1)$ are more than $\exp(-n^{0.11})$ apart from $\frac{1}{2}$. Moreover, by the similar argument it is possible to show: if $\{\bar{z}_i\}_{i\in I}$ is the set of signal carriers that S encounters during the time $[0, m^2]$, then for each pair $i, j \in I$, the frequencies of ones satisfy $|h(i) - h(j)| > \exp(-n^{0.11})$. Let $E^n_{3\_2}$ be the set on which both statements holds.

Define $E_{OK} := E^n_{1\_2} \cap E^n_{3\_2} \cap E^n_{6\_2}$. From now on we assume that E_{OK} hold and we describe the $\hat{g}$-construction algorithm in this case :

Phase I) Determine the intervals $T \subseteq [0, m^2]$ containing more than $e^{n^{0.2}}$ two's (in the observations.) Let T_j designate the j-th such interval. Recall that these are the intervals where we can significantly estimate the frequency of one's. Let K designate the total number of such time-intervals in $[0, m^2]$.
Let $\pi(j)$ designate the index of the signal carrier $\bar{z}_i$ the random walk visits during time T_j (due to $E^n_{6\_2}$, the visited signal carriers are further apart than $Le^{n^{0.2}}$ from each other and, hence, there is only one signal carrier that can get visited during time T_j. Thus the definition of $\pi(j)$ is correct.)

Phase II) Estimate the frequency of one's for each interval T_j, $j = 1, \ldots, K$. Obtain thus, based on the observations $\chi_0^{m^2}$ only, the vector $(\hat{h}_{T_1}, \ldots, \hat{h}_{T_K}) = \Big(\hat{h}(\pi(1)), \hat{h}(\pi(2)), \ldots, \hat{h}(\pi(K))\Big)$.
Here, $\hat{h}(i)$ denotes the estimate of $h(i)$, obtained by time interval T_j, with $\pi(j) = i$.

The further construction of the $\hat{g}$-reconstruction algorithm bases on an important property of the mapping $\pi : \{1, \ldots, K\} \to \mathbb{Z}$ - with high probability π is a skip free walk, i.e. $|\pi(j) - \pi(j+1)| \leq 1$. Hence, the random walk during time $[0, m^2]$ is unlikely to go from one signal carrier to another without signaling all those in-between. By signaling those in-between, we mean producing in the observations for each signal carrier $\bar{z}_i$ a time

intervals of length $e^{n^{0.3}}$ for which one can significantly estimate the frequency of one's $h(i)$. In particular, the skip-freeness implies that $\pi(1) \in \{0,1\}$. The skip-freeness of π is proved in Theorem 5.2.
Let $\pi_* := \min\{\pi(j) : j = 1,\dots,K\}$. Now $\pi_* \le 1$. Let $\pi^* := \max\{\pi(j) : j = 1,\dots,K\}$. If $S(m^2) = m$, then, by $E^n_{1\_2}$, $\pi^* > n^2$.

Phase III) Apply clustering to the vector $(\hat{h}_{T_1}, \hat{h}_{T_2}, \dots, \hat{h}_{T_K})$, i.e. define

$$C_i := \{\hat{h}_{T_j} : |\hat{h}_{T_j} - \hat{h}_{T_i}| \le 2\exp(-n^{0.12})\},\ \hat{f}_i := \frac{1}{|C_i|}\sum_{j\in C_i}\hat{h}_{T_j}, \quad i = 1,\dots,K.$$

By $E^n_{3\_2}$, we have $5\exp(-n^{0.12}) < \exp(-n^{0.11}) < |h(i) - h(j)|$, if n is big enough. Hence, $\hat{h}_{T_j} \in C_i$ if and only if $\pi(i) = \pi(j)$. Thus, for each different i, j either $C_i = C_j$ or $C_i \cap C_j = \varnothing$. Hence, $\hat{f}_j$ is the average of all estimates of $h(\pi(j))$ and, therefore, $\hat{f}_j$ is a good estimate of $h(\pi(j))$. Obviously,

$$\hat{f}_i = \hat{f}_j \quad \text{if and only if} \quad \pi(i) = \pi(j). \tag{6.1.3}$$

Thus, we can denote $\hat{f}(\bar{z}_i) := \hat{f}_j$, if $\pi(j) = i$ and (6.1.3) implies $\hat{f}(\bar{z}_i)) \ne \hat{f}(\bar{z}_j)$, if $i \ne j$.
After phrase III we, therefore, end up with a sequence of estimators $\hat{f}(\bar{z}_{\pi(1)}), \dots, \hat{f}(\bar{z}_{\pi(K)})$ that correspond to the sequence of frequencies $h(\pi(1)), \dots, h(\pi(1))$. Or, equivalently, $j \mapsto \hat{f}_j$ is a path of a skip-free random walk π on the set of different reals $\{\hat{f}(\bar{z}_{\pi_*}), \dots, \hat{f}(\bar{z}_{\pi^*})\}$. The problem is that the estimates, $\hat{f}(\bar{z}_{\pi(1)}), \dots, \hat{f}(\bar{z}_{\pi(K)})$ are in the wrong order, i.e. we are not aware of the values $\pi(j)$, $j = 1,\dots,K$. But having some information about the values $\pi(j)$ is necessary for estimating the frequencies $h(1), \dots, h(n^2+1)$. So the question is: How can get from the sequence $\hat{f}(\bar{z}(_{\pi(1)}), \dots, \hat{f}(\bar{z}_{\pi(K)})$ the elements $\hat{f}(\bar{z}_1), \dots, \hat{f}(\bar{z}_{n^2+1})$? Or, equivalently: after observing the path of π on $\{\hat{f}(\bar{z}_{\pi_*}), \dots, \hat{f}(\bar{z}_{\pi^*})\}$, how can we deduce $\hat{f}(\bar{z}_1), \dots, \hat{f}(\bar{z}_{n^2+1})$?

Real scenery reconstruction algorithm

We now present the so-called *real scenery reconstruction algorithm* - $\mathcal{A}^{\mathbb{R}}_n$. This algorithm is able to answer to the stated questions up to the (swift by) one element.
The algorithm works due to the particular properties of π and $\{\hat{f}(\bar{z}_{\pi_*}), \dots, \hat{f}(\bar{z}_{\pi^*})\}$. These properties are:

A1) $\pi(1) \in \{0,1\}$, i.e. the first estimated frequency of one's, $\hat{f}_1$ must be either an estimate of $h(1)$ or of $h(0)$. Unfortunately there is no way to find out which one of the two signal carriers $\bar{z}_0$ or $\bar{z}_1$ was visited first. This is why our algorithm can reconstruct the real scenery up to the first or last bit, only;

A2) $\pi(K) > n^2$. This is true, because we condition on $S(m^2) = m$ and we assume that there are at least $n^2 + 1$ 2-s in $[0, m]$ (event $E^n_{1\_2}$);

A3) π is skip-free (it does not jump);

A4) $\hat{f}(\bar{z}_i) \ne \hat{f}(\bar{z}_j)\ \forall j \ne i, \quad i, j \in \{\pi_*, \dots, \pi^*\}$.

Algorithm 6.1.1. *Let $\varkappa = (\varkappa_1, \varkappa_2, \dots, \varkappa_K)$ be the vector of real numbers such that the number of different reals in $\varkappa$ is at least n^2+1. The vector $\varkappa$ constitutes the input for $\mathcal{A}_n^{\mathbb{R}}$.*
Define $\mathcal{R}_1 := \varkappa_1$. From here on we proceed by induction on j : once $\mathcal{R}_j$ is defined, we define $\mathcal{R}_{j+1} : \varkappa_s$, with $s := 1 + \max\{j : \varkappa_j = \mathcal{R}_j\}$. Proceed until $j = n^2+1$ and put

$$\mathcal{A}_n^{\mathbb{R}}(\varkappa) := \big(\mathcal{R}_2, \mathcal{R}_3, \dots, \mathcal{R}_{n^2+1}\big).$$

The idea of the algorithm is very simple: take the first element $\varkappa_1$ of $\varkappa$ and consider all elements of the input vector $\varkappa$ that are equal to $\varkappa_1$ and find the one with the biggest index (the last $\varkappa_1$). Let j_1 be this index. Then take $\varkappa_{j_1+1}$ as the first output and look for the last $\varkappa_{j_1+1}$. Let the corresponding index be j_2 and take $\varkappa_{j_2+1}$ as the second output. Proceed so n^2+1 times.
Let us proof that the algorithm $\mathcal{A}_n^{\mathbb{R}}$ works. In our case the input vector is $\hat{f} := (\hat{f}_1, \dots, \hat{f}_K)$.

Proposition 6.1.1. *Let $\{\hat{f}(\bar{z}_{\pi_*}), \dots, \hat{f}(\bar{z}_{\pi^*})\}$ and π satisfy A1), A2), A3), A4). Then*

$$\mathcal{A}_n^{\mathbb{R}}(\hat{f}) \in \{(\hat{f}(\bar{z}_1), \dots, \hat{f}(\bar{z}_{n^2})), (\hat{f}(\bar{z}_2), \dots, \hat{f}(\bar{z}_{n^2+1}))\}, \quad \text{i.e.} \quad \mathcal{A}_n^{\mathbb{R}}(\hat{f}) \sqsubseteq (\hat{f}(\bar{z}_1), \dots, \hat{f}(\bar{z}_{n^2+1})).$$

Proof. By A1) we have that the first element of the input vector, $\hat{f}_1$, is either $\hat{f}(\bar{z}_1)$ or $\hat{f}(\bar{z}_0)$. Consider the first case. Thus $\mathcal{R}_1 = \hat{f}(\bar{z}_1)$. Proceed by induction: suppose that $\mathcal{R}_j = \hat{f}(\bar{z}_j)$, $j < n^2+1$. Let $i(j)$ be the index of the last $\hat{f}(\bar{z}_j)$ in vector $\hat{f}$. By A2), $i(j) < K$. Since π is skip-free and ends to the right of n^2, we have that after the last visits of $\hat{f}(\bar{z}_j)$, the next observation must be $\hat{f}(\bar{z}_{j+1})$. Hence, in this case, $(\mathcal{R}_1, \dots, \mathcal{R}_{n^2+1}) = (\hat{f}(\bar{z}_1), \dots, \hat{f}(\bar{z}_{n^2+1}))$ and $\mathcal{A}_n^{\mathbb{R}}(f) = (\hat{f}(\bar{z}_2), \dots, \hat{f}(\bar{z}_{n^2+1}))$.
Similarly, if the first element of the $\hat{f}$ is $\hat{f}(\bar{z}_0)$, then $(\mathcal{R}_1, \dots, \mathcal{R}_{n^2+1}) = (\hat{f}(\bar{z}_0), \dots, \hat{f}(\bar{z}_{n^2}))$ and $\mathcal{A}_n^{\mathbb{R}}(f) = (\hat{f}(\bar{z}_1), \dots, \hat{f}(\bar{z}_{n^2}))$. □

Phase IV) Apply $\mathcal{A}_n^{\mathbb{R}}$ to $\hat{f}$. Denote the output $\mathcal{A}_n^{\mathbb{R}}(\hat{f})$ by $(f_1, \dots, f_{n^2})$. By Proposition6.1.1 we know

$$(f_1, \dots, f_n) \sqsubseteq (\hat{f}(\bar{z}_1), \dots, \hat{f}(\bar{z}_{n^2+1}). \tag{6.1.4}$$

Now recall that we are interested in reconstructing the $g_i(\xi_0^m) := I_{[0,5)}(h(i))$ rather than $\hat{h}(i)$. Thus, having estimates for $h(\bar{z}_i)$, namely $\hat{f}(\bar{z}_i)$, we use the obvious estimator for g_i: $I_{[0,0.5)}(f_i)$.

Phase V) Define the final output of $\hat{g}$

$$\hat{g}(\chi_0^{m^2}) := \Big(I_{[0.5,1]}(f_1), \dots I_{[0.5,1]}(f_{n^2})\Big).$$

Recall that because of $E_{3\_2}^n$, with high probability all random variables $h(1), \dots, h(n^2+1)$ are more than $\exp(-n^{0.11})$ apart from $\frac{1}{2}$. Since $\exp(-n^{0.11})$ is much bigger than the preciseness of our estimate, with high probability we have $\hat{f}(\bar{z}_i) < 0.5$ if and only if $h(\bar{z}_i) < 0.5$. By (6.1.4) this means

$$\hat{g}(\chi_0^{m^2}) = \Big(I_{[0.5,1]}(f_1), \dots I_{[0.5,1]}(f_n^2)\Big) \sqsubseteq \Big(I_{[0.5,1]}(h(\bar{z}_1)), \dots I_{[0.5,1]}(h(z_{n^2+1}))\Big) = g(\xi_0^m).$$

Hence, when E_{OK} holds, then $\hat{g}$ is properly defined and the probability (6.1.2) is high. Since we are not interested in $\hat{g}$ when E_{OK} does not hold, we extend the definition of $\hat{g}$ arbitrary to E_{OK}^c.

6.2 Whole truth about signal probabilities

In the previous section we considered the case where the scenery has three colors: $\{0,1,2\}$. The locations of the 2's where called signal carriers. The i-th such place was denoted by $\bar{z}_i$. In reality we have only two colors 0 and 1. Thus, we need to show that with 2 colors we also manage to define signal carriers $\bar{z}_i$ in such a way that all of the following holds:

a) Whenever the random walk passes by a signal carrier, we can recognize that the random walk is close to a signal carrier by looking at the observations (with high probability).

b) The probability to be induced in error by the observations, so that one infers that at a certain time one is close to a signal carrier when one is not, is small. This type of mistake never happens up to time m^2.

c) When we pass a signal carrier we are able to estimate its frequency of one's with high precision (with high probability).

In the present section, we define and investigate an important concept that leads to the signal carriers: Markov signal probability.

6.2.1 Definitions

In this subsection, we define the main notions of the section: delayed signal probability, strong signal probability and Markov signal probability. We also give a few equivalent characterizations of these concepts, and we try to explain their meaning. In the end of the subsection we give a formal definition of the frequency of ones.

***** Let $D \subseteq \mathbb{Z}$ and let $\zeta : D \longrightarrow \{0,1\}$. For example, ζ can be the scenery, ξ or the observations, χ.
Let $T = [t_1, t_2] \subseteq D$ be an integer interval of length at least 3. Then we say that T is a *block* of ζ if and only if we have that

$$\zeta(t_1) = \zeta(t_2) \neq \zeta(t), \forall t \in]t_1, t_2[$$

We call $t_2 - t_1$ the length of the block T. The point t_1 is called the beginning of the block. For example, T is a block of ζ with length 4, if $\zeta | T = 01110$.

***** Let $T = T(\chi) \subseteq \mathbb{N}$ be a time interval, possibly depending on the observations. For example, T can be a block of χ or $T = [t, t+n^2]$ can be a time interval of length n^2+1 such that $\chi(t) = \chi(t+1) = \cdots = \chi(t+n^2)$. Let $I \subseteq \mathbb{Z}$ be an integer interval (a location set). We say that T *was generated* (by S) *on* I, if and only if $\forall t \in T, S(t) \in I$.

***** We now define the delayed signal probability. To simplify the notations afterwards, define

$$M = M(n) := n^{1000} - n^2, \quad \tilde{M} := n^{1000} - 2n^2.$$

Fix $z \in \mathbb{Z}$ and let S_z denote the random walk translated by z, i.e. for all $t \in \mathbb{N}$, $S_z(t) := S(t) + z$. We define the random variable δ_z^d in the following way:

$$\delta_z^d := P\left(\xi(S_z(M)) = \cdots = \xi\left(S_z\left(n^{1000}-1\right)\right) = \xi\left(S_z\left(n^{1000}\right)\right) \middle| \xi\right). \tag{6.2.1}$$

In other words, δ_z^d is the conditional probability (conditional on ξ) to observe only one color in the time interval $[n^{1000}-n^2, n^2]$ if the random walk starts at z. We shall call δ_z^d *delayed signal probability at* z.
During time n^{1000} the random walk can not move more than Ln^{1000}. Thus, δ_z^d depends only on the scenery ξ in the interval $[z-Ln^{1000}, z+Ln^{1000}]$. Let, for each $z \in \mathbb{Z}$

$$I_z := [z-Ln^{1000}, z+Ln^{1000}]. \tag{6.2.2}$$

We have that δ_z^d is a random variable which is measurable with respect to $\sigma(\xi(s)|s \in I_z)$. Since the distribution of ξ is translation invariant, the distribution of δ_z^d does not depend on z.

* For some technical reason, we need a stronger version of the delayed signal probability. Again, let $z \in \mathbb{Z}$. We define the *strong signal probability* at z, $\tilde{\delta}_z^d$, as follows

$$\tilde{\delta}_z^d := P\Big(\xi(S_z(M)) = \cdots = \xi(S_z(n^{1000})), \quad S_z(M+1), S_z(2), \ldots, S_z(n^{1000}) \in [z-L\tilde{M}, z+L\tilde{M}]\Big|\xi\Big).$$

Note that $\tilde{\delta}_z^d$ is measurable with respect to the sigma algebra $\sigma(\xi(s)|s \in [z-L\tilde{M}, z+L\tilde{M}])$. Also note that, obviously, $\delta_z^c \geq \tilde{\delta}_z^c$. However, the difference is not too big. Indeed, Höffding inequality states that for some constant $d > 0$

$$\begin{aligned}\delta_z^d - \tilde{\delta}_z^d &= P\Big(\xi(S_z(M)) = \cdots = \xi(S_z(n^{1000})), \quad \exists s \in \{M, \ldots, n^{1000}\} : |z - S_z(s)| > L\tilde{M}\Big|\xi\Big) \\ &\leq P\Big(|S(M)| > L(\tilde{M}-n^2)\Big) \leq \exp(-dn^{999}).\end{aligned} \tag{6.2.3}$$

* Next we define the Markov signal probability at z.

Let $z \in \mathbb{Z}$. Roughly speaking, the Markov signal probability at z, denoted by δ_z^M, is the conditional (on ξ) probability to have (at least) n^2+1 times the same color generated on I_z exactly $n^{1000}-n^2$ after we observe n^2+1 times the same color generated on I_z. In this formulation the part "after we observe a string of n^2+1 times the same color generated on I_z" needs to be clarified. The explanation is the following: every time there is in the observations n^2+1 times the same color generated on I_z, we introduce a stopping time $\nu_z(i)$. The position of the random walk at these stopping times defines a Markov chain with state space I_z. As we will prove later, this Markov chain $\{S(\nu_z(k))\}_{k\geq 1}$ converges very quickly to a stationary measure, say μ_z. So, by "M after we observe n^2+1 times the same color generated on I_z" we actually mean: "M time after starting the random walk from an initial position distributed according to μ_z". Since the distribution of $S(\nu_z(i))$ converges quickly to μ_z, δ_z^M is close to the probability of observing n^2+1 times the same color generated on I_z exactly M time after time $\nu_z(i)$. In other words, δ_z^M is close to the conditional (on ξ) probability of the event that we observe only one color in the time interval $[\nu_z(i)+n^{1000}-n^2, \nu_z(i)+n^{1000}]$ and that during that time interval the random walk S is in I_z. Thus (for k big enough) δ_z^M is close to:

$$P\Big(\chi(\nu_z(i)+M) = \cdots = \chi(\nu_z(i)+n^{1000}) \quad \text{and} \quad S(\nu_z(i)+M), \ldots, S(\nu_z(i)+n^{1000}) \in I_z|\xi\Big). \tag{6.2.4}$$

The ergodic theorem then implies that on the long run the proportion of stopping times $\nu_z(i)$ which are followed after M by n^2+1 observations of the same color generated on I_z converges a.s. to δ_z^M. Actually, to make some subsequent proofs easier, we do not take

a stopping time $\nu_z(i)$ after each n^2+1 observations of the same color generated on I_z. Rather we ask that the stopping times be apart by at least $e^{n^{0.1}}$.
In order to prove how quickly we converge to the stationary measure, we also view the explained notions in terms of a regenerative process. The renewal times will be defined as the stopping times, denoted by $\vartheta_z(k)$, which stop the random walk at the point $z-2Le^{n^{0.1}}$. To simplify some proofs, we also require that there is at least one stopping $\nu_z(i)$ between $\vartheta_z(k)$ and $\vartheta_z(k+1)$. Thus $\vartheta_z(0)$ denotes the first visit by the random walk S to the point $z-2Le^{n^{0.1}}$. We define $\nu_z(1)$ to be the first time after $\vartheta_z(0)$ where there happens to be n^2+1 times the same color generated on I_z. Then, $\vartheta_z(1)$ is the first return of S to $z-2Le^{n^{0.1}}$ after $\nu_z(1)$ and so on. Let us give the formal definitions of all introduced notions.

* Let $\vartheta_z(0)$ denote the first visit of S to the point $z-2Le^{n^{0.1}}$. Thus

$$\vartheta_z(0)=\min\{\,t\geq 0|\,S(t)=z-2Le^{n^{0.1}}\}.$$

* Let $\nu_z(1)$ designate the first time after $\vartheta_z(0)$ where we observe n^2+1 zero's or one's in a row, generated on I_z. More precisely:

$$\nu_z(1):=\min\left\{t>\vartheta_z(0)\;\middle|\;\begin{array}{c}\chi\,(t)=\chi\,(t-1)=...=\chi\,(t-n^2)\\ \text{and } S(t-n^2),S(t-n^2+1),\ldots,S(t)\in I_z\end{array}\right\}.$$

Once $\nu_z(i)$ is well defined, define $\nu_z(i+1)$ in the following manner:

$$\nu_z(i+1):=\min\left\{t>\nu_z(i)+e^{n^{0.1}}\;\middle|\;\begin{array}{c}\chi\,(t)=\chi\,(t-1)=...=\chi\,(t-n^2)\\ \text{and } S(t-n^2),S(t-n^2+1),\ldots,S(t)\in I_z\end{array}\right\}.$$

* Let $\vartheta_z(k)$ denote the consecutive visits of S to the point $z-2Le^{n^{0.1}}$ provided that between two visits random walk S generates (at least once) $n+1$ consecutive 0-s or 1-s on I_z. More precisely,

$$\vartheta_z(k+1):=\min\{t>\vartheta_z(k)|S(t)=z-2Le^{n^{0.1}},\exists j:\vartheta_z(k)<\nu_z(j)<t\},\quad k=1,2\ldots.$$

Basically, the definition above says: if $\vartheta_z(k)$ is defined, we wait until we observe n^2+1 same colors generated on I_z. Since $S(\vartheta_z(k))=z-2Le^{n^{0.1}}$, then the first n^2+1 same colors generated on I_z can not happen earlier than $e^{n^{0.1}}$ times after $\vartheta_z(k)$. This means, the first n^2+1 same colors generated on I_z can not happen earlier than $e^{n^{0.1}}$ times after last stopping time ν_z, say $\nu_z(i)$ (this happens before $\vartheta_z(k)$). Thus, the first n^2+1 same colors generated on I_z is actually $\nu_z(i+1)$. Observing $\nu_z(i+1)$, we just wait for the next visit of S to the $z-2Le^{n^{0.1}}$. This defines $\vartheta_z(k+1)$.

* Let $X_{z,i}$, $i=1,2,\ldots$ designate the Bernoulli variable which is equal to one if exactly after time M the stopping time $\nu_z(i)$ is followed by a sequence of n^2+1 one's or zero's generated on I_z. More precisely, $X_{z,i}=1$ if and only if

$$\chi(\nu_z(i)+M)=\chi(\nu_z(i)+M+1)=\cdots=\chi(\nu_z(i)+n^{1000})$$

and

$$S(j)\in I_z\quad\forall j=\nu_z(i)+M,\ldots,\nu_z(i)+n^{1000}$$

* Define $\kappa_z(0) := 0$. Let $\kappa_z(k)$ designate the number of stopping times $\nu_z(k)$ occurring during the time from $\vartheta_z(0)$ to $\vartheta_z(k)$. Thus $\kappa_z(k)$ is defined by the inequalities:

$$\nu_z(\kappa_z(k)) \leq \vartheta_z(k) < \nu_z(\kappa_z(k)+1).$$

For all k, $S(\vartheta_z(k)) = z - 2Ln^{1000}$. Hence, for all i, $\vartheta_z(k) \neq \nu_z(i)$ and the inequalities above are strict.

* Define the following variables:

$$\mathcal{X}_z(k) = \sum_{i=\kappa(k)+1}^{\kappa(k)} X_{z,i}, \quad \mathcal{Z}_z(k)=\kappa(k) - \kappa(k-1), \quad k = 1, 2, \ldots$$

Thus, $\mathcal{Z}_z(k)$ is the number of stopping times occurring during the time interval from time $\vartheta_z(k-1)$ to time $\vartheta_z(k)$. Note that $\mathcal{Z}_z(k) \geq 1$, $\forall k$. The random variable $\mathcal{X}_z(k)$ designates the number of such stopping times which, during the same time interval, have been followed exactly after time M by a sequence of n^2+1 0's or 1's generated on I_z. Note that conditional on ξ the variables $\mathcal{X}_z(1)$, $\mathcal{X}_z(2), \ldots$ are i.i.d. and the same holds for $\mathcal{Z}_z(1)$, $\mathcal{Z}_z(2), \ldots$.

* We define:

$$\delta_z^M := \frac{E\,[\,\mathcal{X}_z(1)|\,\xi]}{E\,[\,\mathcal{Z}_z(1)|\,\xi]}. \tag{6.2.5}$$

We call δ_z^M *Markov signal probability* at z.

In the following we give some equivalent forms of (6.2.5).
Note that conditional on ξ, $X_{z,i}$ is a regenerative process with respect to the renewal $\kappa_z(k)$. Hence, conditioning on ξ, we have

$$\lim_{r\to\infty} \sum_{i=1}^{r} \frac{X_{z,i}}{r} = \lim_{k\to\infty} \sum_{i=1}^{\kappa(k)} \frac{X_{z,i}}{\kappa(k)} = \lim_{k\to\infty} \frac{\sum_{i=1}^{k} \mathcal{X}_z(i)}{\sum_{i=1}^{k} \mathcal{Z}_z(i)} = \frac{E\,[\,\mathcal{X}_{z,1}|\,\xi]}{E\,[\,\mathcal{Z}_{z,1}|\,\xi]}. \quad \text{a.s.} \tag{6.2.6}$$

We count (up to time r) all sequences of length n^2+1 of one's or zero's, generated on the interval I_z according to the stopping times $\nu_z(i)$, $k = 1, 2, \ldots$. Among such sequences, the proportion of those sequences which are followed after exactly time M by another sequence of n^2+1 zero's or one's generated on the interval I_z converges a.s. to δ_z^M, as r goes to infinity.

On the other hand, the limit in (6.2.6) can be represented as follows. Fix ξ and z. Let $Y_i := S(\nu_z(i))$, $i = 1, 2, \ldots$ denote the Markov chain obtained by stopping the random walk S by $\nu_z(i)$. The state space of Y_i is I_z. Because of the nature of S, Y_i is finite, irreducible aperiodic and, therefore, an ergodic Markov chain.

Let μ_z denote the stationary distribution of $\{Y_k\}$. In the present section z is fixed, so we write μ. The measure μ is a discrete probability measure on I_z, so $\mu = (\mu(j))_{j\in I_z}$. For each state, $j \in I_z$ define the hitting times $\tau_j(l)$, $l = 1, 2, 3, \ldots$. Formally,

$$\tau_j(1) := \min\{i \geq 1 : Y_i = j\}, \quad \tau_j(l) := \min\{i > \tau_j(l-1) : Y_i = j\}, \quad l = 2, 3 \ldots.$$

Hence,

$$\frac{1}{r}\sum_{i=1}^{r} X_{z,i} = \sum_j \frac{N_j(r)}{r}\frac{1}{N_j(r)}\sum_{l=1}^{N_j(r)} X_{z,\tau_j(l)},$$

where $N_j(r) := \max\{l : \tau_j(l) \le r\}$, $r = 1, 2, 3, \ldots$. Since $\tau_j(l)$, $l = 1, 2, 3, \ldots$ is a (delayed) renewal process with the corresponding renewal numbers $N_j(r)$ and with the expected renewal time $\frac{1}{\mu(j)}$ we get

$$\frac{N_j(r)}{r} \to \mu(j) \quad \text{a.s..}$$

On the other hand, $X_{z,i}$ is a regenerative process with respect to each $\tau_j(l)$, $l = 1, 2, 3, \ldots$. Hence

$$\frac{1}{N_j(r)}\sum_{l=1}^{N_j(r)} X_{z,\tau_j(l)} \to E[X_{z,\tau_j(2)}], \quad \text{as} \quad r \to \infty \quad \text{a.s..}$$

Since $E[X_{z,\tau_j(2)}] = P(X_{z,\tau_j(2)} = 1)$. The latter equals

$$P\Big(S_j(M), S_j(M+1), \cdots, S_j(n^{1000}) \in I_z \quad \text{and} \quad \xi(S_j(M)) = \xi(S_j(M+1)) = \cdots = \xi(S_j(n^{1000})\Big).$$

This can be rewritten as

$$\sum_{l \in I_z} P(j,l)\delta_z(l),$$

where $P(j,l) := P(S(M) = j - l)$ and

$$\delta_z(l) := P\Big(S_l(0), S_l(1), \ldots, S_l(n^2) \in I_z \quad \text{and} \quad \xi(S_l(0)) = \xi(S_l(1)) = \ldots = \xi(S_l(n^2))\Big) \tag{6.2.7}$$

Hence

$$\delta_z^M = \sum_{j \in I_z} \mu(j) P\Big(S_j(M), S_j(M+1), \cdots, S_j(n^{1000}) \in I_z, \quad \xi(S_j(M)) = \cdots = \xi(S_j(n^{1000})\Big) \tag{6.2.8}$$

or

$$\delta_z^M = \sum_{j,l \in I_z} \mu(j) P(j,l)\delta_z(l). \tag{6.2.9}$$

Using the same notation, we have an equivalent form of writing the delayed signal probability

$$\delta_z^d = \sum_{l = I_z} P(z,l)\delta_z(l). \tag{6.2.10}$$

Formula (6.2.9) can be interpreted as follows: let U be a random variable with distribution μ_z and let S be a random walk, independent of U. Let S_U denote the translation of S by U, i.e., for each t, $S_U(t) = U + S(t)$. Then (6.2.9) states

$$\delta_z^M = P\Big(\xi(S_U(M)) = \cdots = \xi(S_U(n^{1000})) \quad \text{and} \quad S_U(M), \cdots, S_U(n^{1000}) \in I_z | \xi\Big). \tag{6.2.11}$$

Thus, δ_z^M is the limit-version of (6.2.4) when $i \to \infty$.

* We now define the frequency of ones. To obtain the consistency with the Markov signal probability, we formally define the frequency of ones in terms of regenerative processes. However, we also derive the analogue of (6.2.11), which explains the meaning of the notion. Let $U_{z,i} = \xi(S(\nu_z(i) + e^{n^{0.1}}))$ and define

$$\mathcal{U}_z(k) := \sum_{i=\kappa(k)+1}^{\kappa(k)} U_{z,i}.$$

Now, let

$$h(z) := \frac{E(\mathcal{U}_z(1)|\xi)}{E(\mathcal{Z}_z(1)|\xi)}.$$

The random variable $h(z)$ is $\sigma(\xi(i) : i \in [z - L(n^{1000} + e^{n^{0.1}}), z + L(n^{1000} + e^{n^{0.1}})])$-measurable; $h(z)$ is called as *frequency of ones* at z. As in (6.2.6), conditioning on ξ, we have

$$\lim_{r\to\infty} \sum_{i=1}^{r} \frac{\mathcal{U}_{z,i}}{r} = h(z) \quad \text{a.s..}$$

With the same argument as above, we get

$$\lim_{r\to\infty} \frac{1}{r} \sum_{i=1}^{r} U_{z,i} = \lim_{r\to\infty} \sum_j \frac{N_j(r)}{r} \frac{1}{N_j(r)} \sum_{l=1}^{N_j(r)} U_{z,\tau_j(l)} = \sum_j \mu(j) E(U_{z,\tau_j(2)}).$$

Now,

$$E(U_{z,\tau_j(2)}) = \sum_{i=j-Le^{n^{0.1}}}^{i=j+Le^{n^{0.1}}} \xi(i) P(S_j(i))$$

and, therefore

$$h(z) = \sum_{j=I_z} \mu(j) \sum_{i=j-Le^{n^{0.1}}}^{j+Le^{n^{0.1}}} \xi(i) P(S_j(i)) = \sum_{i=z-L(n^{1000}+e^{n^{0.1}})}^{z+L(n^{1000}+e^{n^{0.1}})} \xi(i) \sum_{j=I_z} \mu(j) P(S_j(e^{n^{0.1}}) = i). \quad (6.2.12)$$

Now, it is easy to see that in terms of U and S as in (6.2.11), i.e. U and S are independent, U has law μ_z, we have

$$h(z) = P(\xi(U + S(e^{n^{0.1}})) = 1|\xi) = E[\xi(U + S(e^{n^{0.1}}))|\xi], \quad (6.2.13)$$

6.2.2 Auxiliary results

In the present section we investigate the relations between δ_z^M and δ_z^d. Note that they only depend on the scenery ξ in the interval $[z - Ln^{1000}, z + Ln^{1000}]$. In other words,

$$\delta_z^M, \delta_z^d \in \sigma\Big(\xi(j)|j \in [z - Ln^{1000}, z + Ln^{1000}]\Big).$$

The distribution of both δ_z^M and δ_z^d does not depend on particular choice of z. Hence, w.l.o.g., in the following we consider the point $z = 0$, only.

Define $p_M := \max\{P(S(M) = z)|z \in \mathbb{Z}\}$.
We call a block *big*, if its length is bigger than $\frac{n}{\ln n}$.

Proposition 6.2.1. *For any $c_\delta \in [p_M, 2p_M]$ we have that the following holds:*

a $P(\delta_z^d \geq c_\delta) \leq \exp(-\alpha n/\ln n)$, where $\alpha := \ln(1.5)$

b $P\left(\delta_z^d \geq c_\delta\right) \geq (0.5)^n > \exp\left(-n\right)$

c If all blocks of $\xi|[z - Ln^{1000}, z + Ln^{1000}]$ are shorter than $n/\ln n + 1$, then $\delta_z^d < c_\delta$. Formally:

$$\left\{\delta_z^d \geq c_\delta\right\} \subseteq \left\{ \ [z - Ln^{1000}, z + Ln^{1000}] \text{ contains a big block of } \xi \ \right\}$$

d Conditional on $\left\{\delta_z^d \geq c_\delta\right\}$ it is likely that $[z-Ln^{1000}, z+Ln^{1000}]$ contains at most $0.5 \ln n$ big blocks of ξ. More precisely:

$$P\left(E_{\delta,z}^c \middle| \, \delta_z^d \geq c_\delta\right) \leq \left(2Ln^{1000}\right)^{0.5 \ln n} (0.5)^{-0.5n}$$

where

$$E_{\delta,z} := \left\{ \ [z - Ln^{1000}, z + Ln^{1000}] \text{ has less than } 0.5 \ln n \text{ big blocks of } \xi \ \right\}$$

In order to prove Proposition 12.4.9, we use the following lemma. The proof of it can be found in [LMM01].

Lemma 6.2.1. *There exists a constant $a > 0$ such that for each $t, r \in \mathbb{N}$, for each subset $I \subseteq \mathbb{Z}$, and for each $j \in I$ and for every mapping $\zeta : \mathbb{Z} \longrightarrow \{0, 1\}$ we have the following implication:*

if all blocks of ζ in I are shorter or equal to r, then

$$P\left(\begin{array}{c} \zeta\left(S_j\left(0\right)\right) = \zeta\left(S_j\left(1\right)\right) = \cdots = \zeta\left(S_j\left(t\right)\right) \\ \text{and } S_j\left(0\right), S_j\left(1\right), ..., S_j\left(t\right) \in I \end{array}\right) \leq \exp\left(-\frac{at}{r^2}\right).$$

Proof that c holds: W.l.o.g. assume $z = 0$. Suppose that the length of all blocks of $\xi|[-Ln^{1000}, Ln^{1000}]$ is at most $n/\ln n$. Let $I := [-Ln^{1000}, Ln^{1000}]$. Denote $\delta(l) = \delta_0(l)$, where $\delta_0(l)$ is as in (6.2.7). If the all the blocks in I are not longer than $n/\ln n$ we get by Lemma 6.2.1 that for all $j \in I$

$$\delta(j) \leq \exp\Big(-\frac{an^2}{(n/\ln n)^2}\Big) = n^{-a \ln n}.$$

By (6.2.10) we get that

$$\delta_0^d = \sum_{j=-Ln^{1000}}^{Ln^{1000}} P(0,j)\delta(j) \leq \sum_{j=-Ln^{1000}}^{Ln^{1000}} P(0,j) n^{-a \ln n} \leq n^{-a \ln n} \tag{6.2.14}$$

The expression on the right side of the last inequality is of smaller order than any negative polynomial order in n. By the local central limit theorem p_M is of order $n^{-\frac{M}{2}}$. Thus, for n big enough

$$\delta_0^d < p_M \leq c_\delta.$$

Proof that a holds: W.l.o.g. assume $z = 0$. Define the event

$$E_z := \{\xi(z) = \xi(z+1) = \cdots = \xi(z + \frac{n}{\ln n})\}$$

Part c states that

$$\{\delta_0^d \geq c_\delta\} \subseteq \bigcup_{z \in [-Ln^{1000}, Ln^{1000}]} E_z.$$

Thus,

$$P\left(\delta_0^d \geq c_\delta\right) \leq \sum_{z=-Ln^{1000}}^{Ln^{1000}} P(E_z).$$

Now, clearly

$$P(E_z) = \exp\left(-\frac{\ln(2)n}{\ln n}\right).$$

So,

$$P(\delta_0^d \geq c_\delta) \leq 2Ln^{1000} \exp\left(-\frac{\ln(2)\, n}{\ln n}\right). \tag{6.2.15}$$

The dominating term in the product on the right side (6.2.15) is $\exp(-\ln(2)\, n/\ln n)$. Hence, for n big enough, the expression on the right side of (6.2.15) is smaller than $\exp(-\frac{\ln(1.5)n}{\ln n})$.

Proof that b holds: It suffices to prove that

$$P(\delta_z^d \geq 2p_M) \geq (0.5)^n.$$

W.l.o.g assume $z = 0$. Define $E := \{\xi(0) = \xi(1) = \cdots = \xi(n)\}$. We are going to show that

$$E \subseteq \{\delta_0^d \geq 2p_M\} \quad \text{and} \quad P(E) \geq \exp(-n).$$

Recall the definition of $\delta(j)$. If E holds, then for any $j \in [0, n]$ we have

$$\delta(j) \geq P\Big(S_j(t) \in [0, n], \forall t \in [0, n^2]\Big)$$

Now, because of the central limit theorem, there is a constant $b > 0$ not depending on n, such that for all $j \in [n/3, 2n/3]$ we have:

$$P\Big(S_j(t) \in [0, n], \forall t \in [0, n^2]\Big) > b.$$

By the local central limit theorem, again, for all $j \in [n/3, 2n/3]$ we have, for n big enough, that

$$P(0, j) \geq \frac{p_M}{2}. \tag{6.2.16}$$

Using (6.2.10) and (6.2.16) we find that when E holds, then

$$\delta_0^d \geq \sum_{j=\frac{n}{3}}^{\frac{2n}{3}} bP(0, j) \geq \frac{bnp_M}{6}. \tag{6.2.17}$$

For n big enough, obviously the right side of (6.2.17) is bigger than $2p_M$. This proves $E \subseteq \{\delta_0^d \geq 2p_M\}$. Furthermore, we have that $P(E) = 0.5^n$. The inequality $0.5^n > \exp(-n)$ finishes the proof.

Proof that d holds: W.l.o.g. assume $z = 0$. For a block T, the point $\inf T$ is called the beginning of the block. Let $t_1, t_2, \ldots$ denote the beginnings of the consecutive big blocks in $[-Ln^{1000}, \infty)$. Define $t_0 := -Ln^{1000}$ and $g_i := t_i - t_{i-1}$, $i = 1, 2, \ldots$. So, g_i measures the distances between consecutive big blocks. Clearly, g_i-s are i.i.d. Note,

$$E_{\delta,0}^c \subseteq \Big\{ \sum_{i=1}^{0.5 \ln n} g_i \leq 2Ln^{1000} \Big\} \subseteq \cap_{i=1}^{0.5 \ln n} \Big\{ g_i < 2Ln^{1000} \Big\}.$$

Note

$$P(g_1 < 2Ln^{1000}) \leq \sum_{z=t_0}^{Ln^{1000}-1} P(\text{a big block begins at } z) \leq 2Ln^{1000}(0.5)^{\frac{n}{\ln n}}.$$

Hence,

$$P(E_{\delta,0}^c) \leq P(g_i \leq 2Ln^{1000})^{0.5 \ln n} = \left(2Ln^{1000}\right)^{0.5 \ln n}(0.5)^{0.5n}.$$

Combining this with b, we get

$$P(E_{\delta,0}^c | \delta_0^d > c_\delta) \leq \frac{P(E_{\delta,0}^c)}{P(\delta_0^d > c_\delta)} \leq \left(2Ln^{1000}\right)^{0.5 \ln n}(0.5)^{-0.5n} \to 0.$$

Lemma 6.2.2.

$$P\left(\delta_z^d \geq c_\delta\right)\left(2Ln^{1000}\right)^{-0.5 \ln n} \leq 2P\left(\delta_z^d \wedge \delta_z^M \geq c_\delta(1 - O(M^{-\frac{1}{2}}))\right).$$

6.2.3 Proof of Lemma 7.2.9

In the present subsection we prove Lemma 7.2.9. To the end of the section we assume $z = 0$. At first we define fences.

Fences

***** An interval $[t, t + 4L - 1] \subseteq D$ is called a *fence* of ζ, if

$$0 = \zeta(t) = \zeta(t+1) \cdots = \zeta(t+L-1) \neq \zeta(t+L) = \cdots = \zeta(t+2L-1) \neq \zeta(t+2L) = \cdots = \zeta(t+3L-1) \neq \zeta(t+3L) = \cdots = \zeta(t+4L-1)$$

The point $t + 2L$ is the *breakpoint* of the fence. So, T is a fence of ζ corresponding to the $L = 3$, if and only if $\zeta|T = 000111000111$.

Let $z_0 := -Ln^{1000}$ and let z_k, $k = 1, 2, \ldots$ be defined inductively: z_k denotes the breakpoint of the first fence of scenery ξ in $[z_k + 4L, \infty)$. We call the points z_k the breakpoints of consecutive fences (of scenery ξ). Define $l_i := z_i - z_{i-1}$, $i = 1, 2, \ldots$ and $N := \max\{k : z_{k-1} \leq Ln^{1000}\} < Ln^{1000}$. The random variables l_i measure the distances between the breakpoints of consecutive fences, they are i.i.d. Let $l := Ln^{1000} - z_N$, $l \leq l_{N+1}$. The

moment generating function of l_1, say $M(\lambda)$, does not depend on n and it is finite, if $\lambda > 0$ is small enough. Let $M := \exp(\lambda l_1) < \infty$ and choose $C > 1$ such that $\lambda C > 1$. Now define the event

$$E_b := \{l, l_i \le Cn, \quad i = 1, 2, \ldots, N\}$$

and apply the large deviation inequality to see $P(l_1 > Cn) = P(\lambda l_1 > \lambda Cn) < Me^{-\lambda Cn}$. Now,

$$P(E_b^c) \le \sum_{i=1}^{Ln^{1000}} P(l_i > Cn) = Ln^{1000} P(l_1 > Cn) < Ln^{1000} Me^{-\lambda Cn}.$$

Applying b, we get

$$P(E_b^c | \delta_0^d \ge c_\delta) \le \frac{P(E_b^c)}{P(\delta_0^d \ge c_\delta)} \le Ln^{1000} Me^{(1-\lambda C)n} \to 0. \tag{6.2.18}$$

Mapping

Let $\mathcal{O}$ denote the set of all possible pieces of sceneries in $I := [-Ln^{1000}, Ln^{1000}]$, i.e. $\mathcal{O} := \{0,1\}^I$. The random variables δ_0^d, δ_0^M as well as the events $\{\delta_0^d > c_\delta\}$, $E_{\delta,0}$, E_b depend on the restriction of the scenery to I, only. Hence they can be defined on the probability space $(\mathcal{O}, 2^{\mathcal{O}}, P)$, where P stands for the normalized counting measure.
Define

$$\mathcal{C} := \{\delta_0^d > c_\delta\} \cap E_{\delta,0} \cap E_b \subseteq \mathcal{O}.$$

Hence $\mathcal{C}$ consists of all pieces of sceneries, η, with the following properties: $\delta_0^d(\eta)$ is bigger than c_δ, the number of big blocks is less than $0.5 \ln n$ and the gaps between the breakpoints of the consecutive fences in I is at most Cn.

Let $\eta \in \mathcal{C}$ and let $z_0, z_1, \ldots, z_N$ be the breakpoints of consecutive fences (restricted to I) of η. Since $\eta \subseteq E_b$, we have $N \ge 2Ln^{999}$. Now partition the interval I as follows:

$$I = I_1 \cup I_2 \cup \cdots \cup I_N \cup I_{N+1}, \tag{6.2.19}$$

where $I_k := [z_{k-1}, z_k - 1]$, $k = 1, \ldots, N, I_{N+1} := [I_N, Ln^{1000}]$. Let $l(I_k) := z_k - k_{k-1}$ denote the length of I_k. We shall call the partition (6.2.19) the fence-partition corresponding to η. The fences guarantee that any block of η, that is longer than L is a proper subset of one interval I_k. Since $\eta \in \{\delta_0^d > c_\delta\} \cap E_{\delta,0}$, there is at least one and at most $0.5 \ln n$ big blocks. Let I_k^*, $k = 1, \ldots N^*$, $N^* \le 0.5 \ln n$ denote the $k-th$ interval containing at least one big block. Similarly, let I_k^o, $k = 1, \ldots, N + 1 - N^*$ denote the $k-th$ interval with no big blocks. Clearly, most of the intervals I_k are without big blocks, in particular $\sum_k l(I_k^o) > Ln^{1000}$. Define

$$j^o := \min\{j : \sum_{k=1}^{j} l(I_k^o) > Ln^{1000}\}.$$

To summarize - to each $\eta \in \mathcal{C}$ corresponds an unique fence-partition, an unique labelling of the interval according to the blocks, and, therefore, unique j^o. We now define a mapping $B : \mathcal{C} \to \mathcal{O}$ as follows:

$$B(\eta) := (\eta|I_1^o, \eta|I_2^o, \ldots, \eta|I_{j^o}^o, \eta|I_1^*, \ldots, \eta|I_{N^*}^*, \eta|I_{j^o+1}^o, \ldots, \eta|I_{N+1-N^*}^o).$$

We also define the corresponding permutation:

$$\Pi_\eta : I \to I, \quad \Pi_\eta(I) = (I_1^o, I_2^o, \dots, I_{j^o}^o, I_1^*, \dots, I_{N^*}^*, I_{j^o+1}^o, \dots, I_{N+1-N^*}^o).$$

Thus, $B(\eta) = \eta \circ \Pi_\eta$.
Since all big blocks of η are contained in the intervals I_k, the mapping B keeps all big blocks unchanged, and just moves them closer to the origin.
The mapping B is clearly not injective. However, $B(\eta_1) = B(\eta_2)$ implies that the fence-partitions corresponding to η_1 and η_2 consists of the same intervals, with possibly different order. Also the intervals with big blocks (marked with star) are the same, but possibly differently located. Moreover, the ordering of the similarly marked blocks corresponding to η_1 and η_2 are the same (i.e. if the 8-th interval, I_8, of the partition corresponding to η_1 is the 20-th interval, I_{20}, of the partition corresponding to η_2, then their marks are the same. If I_8 in its partition is the seventh interval with o ($I_8 = I_7^o$ in the partition corresponding to the η_1), then the same block in the second partition must be also the seventh interval with o ($I_{20} = I_7^o$ in the partition corresponding to η_2). Therefore, the partition corresponding to η_1 and η_2 differ on the location of the star-intervals, only. Since the number of intervals is smaller than $2Ln^{1000}$ and the number of star-intervals is at most $0.5 \ln n$, the number of different partitions with the properties described above, is less than $(2Ln^{1000})^{0.5 \ln n}$. This means

$$|B(\mathcal{C})|(2Ln^{1000})^{0.5 \ln n} > |\mathcal{C}|. \tag{6.2.20}$$

Proof of Lemma 7.2.9: Because of the counting measure and (6.2.20) we get

$$\frac{P(B(\mathcal{C}))}{P(\mathcal{C})} = \frac{|B(\mathcal{C})|}{|\mathcal{C}|} > (2Ln^{1000})^{-0.5 \ln n}.$$

By Propositions 12.4.10 and 6.2.3,

$$P(B(\mathcal{C})) \le P\Big(\delta_0^d \wedge \delta_0^M \ge c_\delta(1 - O(M^{-\frac{1}{2}}))\Big).$$

By (6.2.18) and d) of Proposition 12.4.9, we get:

$$\frac{P(\mathcal{C})}{P(\delta_0^d > c_\delta)} = P(E_{\delta,0} \cap E_b | \delta_0^d \ge c_\delta) > 0.5,$$

provided n is big enough. These relations yield:

$$P\Big(\delta_0^d \wedge \delta_0^M \ge c_\delta(1 - O(M^{-\frac{1}{2}}))\Big) \ge (2Ln^{1000})^{-0.5 \ln n} \cdot 0.5 \cdot P(\delta_0^d > c_\delta).$$

The lemma is proved.

Proposition 6.2.2. *For any $\varsigma \in B(\mathcal{C})$ we have*

$$\delta_0^d(\varsigma) \ge c_\delta[1 - O(M^{-\frac{1}{2}})].$$

Proof. Let $\varsigma \in B(\mathcal{C})$. Choose $\eta \in B^{-1}(\varsigma)$. Let $\{I_k\}$ be the fence-partition corresponding to η. Let $\delta_z^\eta(l)$, $\delta_z^\varsigma(l)$ denote the probabilities defined in (6.2.7), with ξ replaced by η and ς, respectively. As already noted, because of the fencing-structure, any sequence of consecutive one's or zero's can be generated on the one interval I_k, only. More precisely, if $l \in I_k$, then

$$\delta_0^\eta(l) = P\big(S_l(0), \dots, S_l(n^2) \in I_k\,,\, \eta(S_l(0)) = \dots = \eta(S_l(n^2))\big). \tag{6.2.21}$$

By the argument of the proof of c of Proposition 12.4.9, we get that each interval without big blocks, I_k^o, has the property: the probability of generating n^2+1 consecutive zeros or ones is smaller than $n^{-a\ln n}$. In other words $\delta_0^\eta(l) \le n^{-a\ln n}, \quad \forall l \in I^o$, where $I^o := \cup_k I_k^o$. Denote $I^* := \cup_k I_k^*$. Now, by (6.2.10) and (6.2.21) we have

$$\begin{aligned}
\delta_0^d(\eta) &= \sum_{l\in I} P(0,l)\delta_0^\eta(l) = \Big(\sum_{l\in I^o} + \sum_{l\in I^*}\Big) P(0,l)\delta_0^\eta(l) \\
&\le \sum_{l\in I^o} P(0,l) n^{-a\ln n} + \sum_{l\in I^*} P(0,l)\delta_0^\eta(l) \\
&\le n^{-a\ln n} + \sum_{l\in I^*} P(0,l)\delta_0^\eta(l) \le n^{-a\ln n} + p_M \sum_{l\in I^*} \delta_0^\eta(l).
\end{aligned}$$

Since $\eta \in \mathcal{C}$, $\delta_0^d(\eta) \ge c_\delta \ge p_M$, we have

$$\sum_{l\in I^*} \delta_0^\eta(l) \ge \frac{c_\delta - n^{-a\ln n}}{p_M} \ge 1 - \frac{n^{-a\ln n}}{p_M} = 1 - O\Big(\frac{\sqrt{M}}{n^{a\ln n}}\Big), \tag{6.2.22}$$

Clearly $O(\frac{\sqrt{M}}{n^{a\ln n}}) = o(n^{-\alpha})$, for all $\alpha \ge 0$.

Now consider $\varsigma = M(\eta)$. Let $J_1, J_2, \dots J_{N+1}$ denote the new location of the intervals I_i after applying the mapping Π_η to I. Fix an $j \in I$ and let $j \in J_k$. The equation $\varsigma|J_k = \eta|I_k$ and (6.2.21) imply

$$\begin{aligned}
\delta_0^\varsigma(j) &= P\big(S_j(0), \dots, S_j(n^2) \in I, \varsigma(S_j(0)) = \dots = \varsigma(S_j(n^2))\big) \\
&\ge P\big(S_j(0), \dots, S_j(n^2) \in J_k, \varsigma(S_j(0)) = \dots \varsigma(S_j(n^2))\big) \\
&= P\big(S_l(0), \dots, S_l(n^2) \in I_k, \eta(S_l(0)) = \dots = \eta(S_l(n^2))\big) = \delta_0^\eta(l),
\end{aligned}$$

where $l = \Pi(j) \in I_k$. This means $\delta_0^\varsigma(j) \ge \delta_0^\eta(\Pi_\eta(j)), \quad \forall j \in I$. In particular,

$$\sum_{j\in J_k} \delta_0^\varsigma(j) \ge \sum_{l\in I_k} \delta_0^\eta(j) \tag{6.2.23}$$

If $I_1 = J_1$ and $I_{N+1} = J_{N+1}$, i.e. the first and last intervals do not contain big blocks, then, obviously, (6.2.23) is an equation.
Let $J^* = \Pi_\eta(I^*)$, i.e. J^* is the union of all intervals with big blocks in the new location. The length of I^* (and, therefore, that of J^*) is at most $0.5Cn\ln n$. Thus, J^* is at most $Cn + 0.5Cn\ln n$ from the origin. Let n be so big, that $Cn + 0.5Cn\ln n \le n^2$. Then, $j \le n^2$ for each for each $j \in J^*$. Denote by:

$$p_o = \min\{P(S(M) = i) : |i| \le n^2\}.$$

Now from (6.2.22) and (6.2.23) we get

$$\begin{aligned}\delta_0^d(\varsigma) &= \sum_j P(0,j)\delta_0^\varsigma(l) \geq \sum_{j\in J^*} P(0,j)\delta_0^\varsigma(j) \geq \sum_{l\in I^*} P(0,j)\delta_0^\eta(l)\\ &\geq p_o \sum_{l\in I^*} \delta_0^\eta(l) \geq (c_\delta - n^{-a\ln n})\frac{p_o}{p_M} = c_\delta(1 - \frac{p_M - p_o}{p_M} - \frac{n^{-a\ln n}p_o}{c_\delta p_M})\\ &= c_\delta[1 - O(M^{-\frac{1}{2}})] - O(\frac{\sqrt{M}}{n^{a\ln n}}) = c_\delta[1 - O(M^{-\frac{1}{2}})].\end{aligned}$$

Proposition 6.2.3. *For any $\varsigma \in B(\mathcal{C})$ we have*

$$\delta_0^M(\varsigma) \geq c_\delta[1 - O(M^{-\frac{1}{2}})].$$

Proof. We use the notation and the results of the previous proof. By the representation (6.2.8) we have

$$\delta_0^M(\varsigma) = \sum_{i,j\in I} \mu(i)P(i,j)\delta_0^\varsigma(j) \geq \sum_{i,j\in J^*} \mu(i)P(i,j)\delta_0^\varsigma(j) \tag{6.2.24}$$

where $\mu = \{\mu(i)\}_{i\in I}$ is the stationary measure of $Y_k = S(\nu_0(k))$, $k = 1, 2, \ldots$.

Use local central limit theorem (CLT in the sequel) to estimate

$$\begin{aligned}\min_{i,j\in J^*} P(j,i) &\geq \min\{P(i,j) : |i-j| \leq n^2\} \geq \frac{c}{\sqrt{M}}\exp\big(-\frac{dn^2}{M}\big) - O(M^{-1})\\ &= \frac{c}{\sqrt{M}}\Big(1 - O\big(\frac{n^2}{M}\big)\Big) - O(M^{-1}) = p_M\Big(1 - O(\frac{1}{\sqrt{M}})\Big).\end{aligned} \tag{6.2.25}$$

with d, c being constants not depending on n.
Hence, because of (6.2.24), (6.2.22) and (6.2.25)

$$\begin{aligned}\delta_0^M(\varsigma) \geq& \mu(J^*)[p_M\big(1 - O(\frac{1}{\sqrt{M}})\big)]\frac{c_\delta - n^{-a\ln n}}{p_M}\\ =& \mu(J^*)\big(1 - O(\frac{1}{\sqrt{M}})\big)(c_\delta - n^{-a\ln n}) = \mu(J^*)\big(1 - O(\frac{1}{\sqrt{M}})\big)c_\delta.\end{aligned} \tag{6.2.26}$$

We now estimate $\mu(J^*)$. We shall show that

$$P(Y_{k+1} \in J^* | Y_k = j) \geq 1 - o(M^{-1}) \quad \forall j \in I.$$

Then $\mu(J^*) = \sum_j P(Y_{k+1} \in J^* | Y_k = j)\mu(j) \geq 1 - o(M^{-1})$ and, by (6.2.26)

$$\delta_0^M(\varsigma) \geq \mu(J^*) \geq (1 - o(M^{-1}))c_\delta[1 - O(M^{-\frac{1}{2}})] = c_\delta[1 - O(M^{-\frac{1}{2}})].$$

Estimation of $\mu(J^*)$

Fix an $j \in I$ and define ν as the first time after $e^{n^{0.1}}$ when n^2+1 consecutive 0-s or 1-s are generated on I. Formally,

$$\nu := \min\left\{ t \geq e^{n^{0.1}} \,\middle|\, \begin{array}{c} \chi(t) = \chi(t-1) = ... = \chi(t-n^2) \\ \text{and } S_j(i) \in I, \forall i = t-n^2, \ldots, t \end{array} \right\}$$

where $\chi = \varsigma \circ S_j$. Clearly

$$P(S_j(\nu) \in J^*) = P(Y_{k+1} \in J^* | Y_k = j).$$

Thus, it suffices to estimate $P(S_j(\nu) \in J^*)$.

At first note that by (6.2.22) and (6.2.23,) we get $\sum_{j \in J^*} \delta_0^\eta(j) \to 1$. Since $|J^*| \leq n^2$ (and n is big enough), we deduce the existence of $j^* \in J^*$ such that

$$\delta_0^\eta(j^*) > \frac{1}{n^3}. \tag{6.2.27}$$

Then, because of the fences we have:

$$\{S_j(\nu) \notin J^*\} = \{S_j(\nu-n^2), \ldots, S_j(\nu) \in I \backslash J^*, \;\; \chi(\nu-n^2) = \cdots = \chi(\nu)\}.$$

Now, let τ_k be the k-th visit after time $e^{n^{0.1}} - n^2$ to the interval I. Let τ_k^* be the k-th visit after time $e^{n^{0.1}} - n^2$ to the point j^*. Define the events

$$F_k := \{S_j(\tau_k - n^2), \ldots, S_j(\tau_k) \in I \backslash J^*, \;\; \chi(\tau_k - n^2) = \cdots = \chi(\tau_k)\}, \;\; k = 1, 2 \ldots$$

$$F_k' = \cup_{i=0}^{n^{2000}-1} \{S_j(\tau_k + i) = j^*\}, \quad k = 1, 2, \ldots$$

$$F_k^* = \{\chi(\tau_k^*) = \cdots = \chi(\tau_k^* + n^2)\}, \;\; k = 1, 2, \ldots$$

We consider the events

$$E_1 := \{\nu > \tau_{n^{2020}}\} \cup \{S_j(\nu) \in J^*\}, \;\; E_2 := \{\tau_{n^{10}}^* \leq \tau_{n^{2020}} - n^2\}, \;\; E_3 := \cup_{k=1}^{n^{10}} F_k^*$$

The event E_1 ensures that within the first n^{2020} visits of S_j to I no consecutive 0's or 1's were generated on $I \backslash J^*$. The event E_2 ensures that before time $\tau_{n^{2020}} - n^2$ the random walk visits at least n^{10} times the point j^*. Finally, the event E_3 ensures that during these n^{10} visits of j^*, at least one of them is a beginning of n^2 consecutive 0's or 1's. If these events hold, then $\nu \leq \tau_{n^{2020}}$ and $S_j(\nu) \in J^*$. Thus

$$E_1 \cap E_2 \cap E_3 \subseteq \{S_j(\nu) \in J^*\}.$$

Next, we give upper bounds for the probabilities $P(E_1), P(E_2), P(E_3)$.

1) Note that: $E_1^c \subseteq \cup_{k=1}^{n^{2020}} F_k$, implies: $P(E_1^c) \leq \sum_{k=1}^{n^{2020}} P(F_k)$. For each k,

$$P(F_k) = \sum_{l \in I \backslash J^*} P[S_l(0), \ldots, S_l(n^2) \in I \backslash J^*, \varsigma(S_l(0)) = \cdots = \varsigma(S_l(n^2))] \times$$
$$\times P(S_j(\tau_k - n^2) = l).$$

There is no big blocks in $I\backslash J^*$, hence by the argument of c:

$$P[S_l(0),\dots,S_l(n^2)\in I\backslash J^*,\varsigma(S_l(0))=\dots=\varsigma(S_l(n^2))]\le n^{-a\ln n},$$

implying that:

$$P(E_1^c)\le n^{2020-a\ln n}.$$

2) To estimate $P(E_2)$ we use the Höffding inequality. By central limit theorem there exists a constant $p>0$ not depending on n such that $P(F_k')\ge p$. Also note that F_k' and F_l' are independent if $|k-l|\ge n^{2000}$. Hence, the set $\{F_k'\}$, $k=1,\dots,n^{2020}$ contains a subset $\{F_{k_i}'\}$ $i=1,\dots n^{20}$ consisting of independent events. Let $X_i:=I_{F_{k_i}'}$. Now, $\tau_{n^{2018}}+n^{2000}\le\tau_{n^{2019}}\le\tau_{n^{2020}}-n^2$, if n is big enough. This means

$$\Big\{\sum_{i=1}^{n^{18}}X_i\ge n^{10}\Big\}\subseteq E_2.$$

Now, when n is big enough, we have

$$\begin{aligned}P(E_2^c)&\le P\Big(\sum_{i=1}^{n^{18}}X_i<n^{10}\Big)=P\Big(\sum_{i=1}^{n^{18}}(X_i-EX_i)<n^{10}-\sum_{i=1}^{n^{18}}EX_i\Big)\\&\le P\Big(\sum_{i=1}^{n^{18}}(X_i-EX_i)<-(n^{18}p-n^{10})\Big)\le P\Big(\sum_{i=1}^{n^{18}}(X_i-EX_i)<-n^{17}\Big)\le\\&\le\exp(-\frac{2n^{34}}{n^{18}})=\exp(-2n^{16}).\end{aligned}$$

3) Note F_l^*,F_k^* are independent, if $|k-l|>n^2$ Let $\{F_{k_i}^*\}$, $i=1,2,\dots,n^7$ be a subset of $\{F_k^*\}$ consisting on independent events, only. By (6.2.27), $P(F_k^*)>\frac{1}{n^3}$, $\forall k$. Now

$$P(E_3^c)\le P(\cap_{i=1}^{n^7}F_{k_i}^*)=\prod_{i=1}^{n^7}(1-P(F_{k_i}^*))\le\Big(1-\frac{1}{n^3}\Big)^{n^7}.\tag{6.2.28}$$

The right side of (6.2.28) is smaller than $(0.5)^{n^4}$ if n is big enough.
Thus,

$$\begin{aligned}P(S_j(\nu)\in J^*)&\ge 1-[n^{2020-a\ln n}+\exp(-2n^{16})+(0.5)^{n^4}]\\&=1-O(n^{-2020+a\ln n})=1-o(M^{-1}).\end{aligned}$$

6.2.4 Corollaries

We determine the critical value c_r. Since we choose it within the interval $[p_M,2p_M]$, it has all properties stated in Proposition 12.4.9 and Lemma 7.2.9. However, we also have to ensure that with high probability the signal probabilities δ_z^d and δ_z^M are significantly away from c_r. By "significantly" we mean that the difference between these probabilities and c_r is bigger than a polynomially small quantity in n. This polynomially small quantity will be denoted by Δ. Thus, c_r must be properly chosen and that will be done with the help of Corollary 6.2.2.
At first, some preliminary observations.

Proposition 6.2.4. *For any $j > 2$, there exists an interval $[a, b] \subseteq [p_M, 2p_M]$ of length $p_M/(n^{j+2})$ such that*

$$P(\delta_0^d < b | \delta_0^d \geq a) \leq \frac{1}{n^j} \tag{6.2.29}$$

Proof. We do the proof by contradiction. Assume on the contrary that there exists no interval $[a, b] \subseteq [p_M, 2p_M]$ of length $l := p_M/n^{j+2}$ such that (6.2.29) is satisfied. Let $a_i := p_M + il$, $i = 0, \ldots, n^{j+2}$. Since $[a_i, a_{i+1}] \subseteq [p_M, 2p_M]$ is an interval of length l, by assumption:

$$P(\delta_0^d \geq a_{i+1} | \delta_0^d \geq p_M + a_i) \leq \left(1 - \frac{1}{n^j}\right), \quad i = 1, \ldots, n^j - 1.$$

Now, by b) of Proposition 12.4.9:

$$e^{-n} < P(\delta^d \geq 2p_M) = \prod_{i=0}^{n^{j+2}-1} P(\delta_0^d \geq a_{i+1} | \delta_0^d \geq a_i) \leq \left(1 - \frac{1}{n^j}\right)^{n^{j+2}}. \tag{6.2.30}$$

Since $(1 - \frac{1}{n^j})^{n^j} < e^{-1}$, we have $(1 - \frac{1}{n^j})^{n^{j+2}} < e^{-n^2}$. Thus, (6.2.30) implies $e^{-n} < e^{-n^2}$ - a contradiction. □

Corollary 6.2.1. *Let $[x, y] \subseteq [p_M, 2P_M]$ be an interval of length l. Then there exists an subinterval $[u, v] \subseteq [x, y]$ of length $\frac{l}{e^{2n}}$ such that*

$$P(\delta_0^d < v | \delta_0^d > u) \leq \frac{1}{e^n}. \tag{6.2.31}$$

Proof. The proof of the corollary follows the same argument that the proof of Proposition 6.2.4: (6.2.31) together with the statement b) of Proposition 12.4.9 yield the contradiction: $\exp(-n) < P(\delta_0^d \geq 2p_M) \leq P(\delta_0^d \geq v) \leq \left[\left(1 - \frac{1}{e^n}\right)^{e^n}\right]^{e^n} < \exp(-e^n)$. □

The next proposition proves the similar result for $\delta_0^M \wedge \delta_0^d$. Since we do not have the analogue of b) of Proposition 12.4.9, we use Lemma 7.2.9, instead.

Proposition 6.2.5. *Let $[a, b] \subseteq [p_M, 2p_M]$ be such that $2p_M - b > p_M O(M^{-\frac{1}{2}})$. For any $i > 2$ there exists an interval $[x, y] \subseteq [a, b]$ with length $(b - a)/n^{i+2}$ such that, for n big enough*

$$P(\delta_0^M < y | \delta_0^M \wedge \delta_0^d > x) \leq P(\delta_0^M \wedge \delta_0^d < y | \delta_0^M \wedge \delta_0^d > x) \leq \frac{1}{n^i}. \tag{6.2.32}$$

Proof. Suppose that such a (sub)interval does not exists. Then follow the argument of the previous proof to get

$$P\left(\delta_0^M \wedge \delta_0^d \geq 2p_M(1 - O(M^{-\frac{1}{2}}))\right) \leq P\left(\delta_0^M \wedge \delta_0^d \geq b\right) \leq \left(1 - \frac{1}{n^i}\right)^{n^{i+2}} < \exp(-n^2). \tag{6.2.33}$$

By Lemma 7.2.9 and b) of Proposition 12.4.9

$$P\left(\delta_0^M \wedge \delta_0^d \geq 2p_M(1 - O(M^{-\frac{1}{2}}))\right) \geq 0.5(2Ln^{1000})^{-0.5 \ln n} \exp(-n). \tag{6.2.34}$$

For n big enough, the right side of (6.2.34) is bigger than e^{-2n}. This contradicts (7.1.1). □

The following corollary specifies c_r and Δ.

Corollary 6.2.2. *Let* $\Delta := \big(p_M/8\big)n^{-10054}$, $\tilde{\Delta} = \Delta e^{-2n}$. *Then there exists* $c_r \in [p_M + \Delta, 2p_M - \Delta]$ *such that, for* n *big enough, simultaneously,*

$$P\left(\delta_0^d \geq c_r - \Delta\right) \leq \exp((\ln n)^3) P\left(\delta_0^d \wedge \delta_0^M \geq c_r - \Delta\right); \tag{6.2.35}$$

$$P(\delta_0^M < c_r + \Delta | \delta_0^M \wedge \delta_0^d \geq c_r - \Delta) \leq n^{-10000} \tag{6.2.36}$$

and

$$P(\delta_0^d < c_r - \Delta + \tilde{\Delta} | \delta_0^d \geq c_r - \Delta) \leq \exp(-n). \tag{6.2.37}$$

Proof. By Proposition 6.2.4 there exists an interval $[a, b] \subseteq [p_M, 2p_M]$ of length p_M/n^{52} such that

$$\frac{P(\delta_0^d \geq b)}{P(\delta_0^d \geq a)} = P(\delta_0^d \geq b | \delta_0^d \geq a) > 1 - \frac{1}{n^{50}} > 0.5. \tag{6.2.38}$$

We now consider the interval $[a, \frac{a+b}{2}]$. Note that:

$$2p_M - \frac{a+b}{2} \geq b - \frac{b+a}{2} = \frac{b-a}{2} = \frac{p_M}{2n^{52}} > p_M O(M^{-\frac{1}{2}}).$$

Now use Proposition 6.2.5 with $i = 10000$ to find a subset $[x, y] \in [a, \frac{a+b}{2}]$ with length $l := \frac{b-a}{2} n^{-10002} = \frac{p_M}{2} n^{-10054}$ such that (6.2.32) holds.
Let us now take $z = x + \frac{l}{4}$. By Corollary 6.2.1, there exists an subinterval $[u, u+\tilde{\Delta}] \in [x, z]$ with length $\frac{l}{4e^{2n}}$ such that

$$P(\delta_0^d < u + \tilde{\Delta} | \delta^d > u) \leq \exp(-n). \tag{6.2.39}$$

Now take $\Delta := \frac{l}{4} = \big(p_M/8\big)n^{-10054}$, $c_r := u + \Delta$. Since $[c_r - \Delta, c_r + \Delta] \subseteq [x, y]$, we have that

$$P(\delta_0^M < c_r + \Delta | \delta_0^M \wedge \delta_0^d > c_r - \Delta) \leq P(\delta_0^M \wedge \delta_0^d < c_r + \Delta | \delta_0^M \wedge \delta_0^d > c_r - \Delta) \leq$$
$$P(\delta_0^M \wedge \delta_0^d < y | \delta_0^M \wedge \delta_0^d > c_r - \Delta) = \frac{P(\Delta - c_r < \delta_0^M \wedge \delta_0^d < y)}{P(\delta_0^M \wedge \delta_0^d > \Delta - c_r)} \leq$$
$$\frac{P(y > \delta_0^M \wedge \delta_0^d > x) - P(x \leq \delta_0^M \wedge \delta_0^d \leq c_r - \Delta)}{P(\delta_0^M \wedge \delta_0^d > x) - P(x < \delta_0^M \wedge \delta_0^d \leq c_r - \Delta)} \leq \frac{P(y > \delta_0^M \wedge \delta_0^d > x)}{P(\delta_0^M \wedge \delta_0^d > x)} =$$
$$P(\delta_0^M \wedge \delta_0^d < y | \delta_0^M \wedge \delta_0^d > x) \leq \frac{1}{n^{10000}}.$$

Hence, (6.2.36) holds.
Since $u = c_r - \Delta$, we also have that (6.2.37) holds.
It only remains to show that the chosen c_r also satisfies (6.2.35).
Clearly $\Delta > 2p_M O(M^{-\frac{1}{2}}) > c_r O(M^{-\frac{1}{2}})$. That implies:

$$P\Big(\delta_0^d \wedge \delta_0^M \geq c_r(1 - O(M^{-\frac{1}{2}})\Big) \leq P(\delta_0^d \wedge \delta_0^M \geq c_r - \Delta).$$

Combine this with Lemma 7.2.9 to get

$$P(\delta_0^d \geq c_r) 0.5 (2Ln^{1000})^{-0.5 \ln n} \leq P(\delta_0^d \wedge \delta_0^M \geq c_r - \Delta) \tag{6.2.40}$$

Since $[c_r - \Delta, c_r + \Delta] \subseteq [a, b]$ we have

$$P(\delta_0^d \geq a) \geq P(\delta_0^d \geq c_r - \Delta) \geq P(\delta_0^d \geq c_r) \geq P(\delta_0^d \geq b).$$

Now, by (6.2.38)

$$\frac{P(\delta_0^d \geq c_r)}{P(\delta_0^d \geq c_r - \Delta)} \geq \frac{P(\delta_0^d \geq b)}{P(\delta_0^d \geq a)} > 0.5.$$

The last inequality above, together with (6.2.40) implies

$$P(\delta_0^d \geq c_r - \Delta) \leq 0.25(2Ln^{1000})^{0.5 \ln n} P(\delta_0^d \wedge \delta_0^M \geq c_r - \Delta) \tag{6.2.41}$$

Now, the relation

$$0.25(2Ln^{1000})^{0.5 \ln n} \leq \exp((\ln n)^3)$$

together with (6.2.41) establishes (6.2.35). □

6.3 Scenery-dependent events

In the present section we define and investigate the signal points and Markov signal points. We show that with high probability the location of the signal points follows certain clustering structure. This structure gives us the desired signal carriers in the 2-color case.

6.3.1 Signal points

We are now going to define the Markov signal points, strong signal points and signal points – these are the location points, where the corresponding signal probabilities are above the critical value c_r. The Markov signal points form the core of the signal carriers, the (strong) signal points will be used in our proofs. In an oversimplified way, we could say that the Markov signal points are places in the scenery ξ where the conditional probability to see in the observations some rare unusual pattern is above c_r. The unusual pattern is basically a string of n^2, zero's or one's.
In the present subsection, with the help of the signal points, we define many other important notions, and we also investigate their properties.

In the following, Δ and c_r are as in Corollary 6.2.2. In particular, $\Delta = \frac{p_M}{8} n^{-10054}$.

* A (location) point $z \in \mathbb{Z}$ is called *signal point*, if $\delta_z^d > c_r - \Delta$.

* A (location) point $z \in \mathbb{Z}$ is called *strong signal point*, if $\tilde{\delta}_z^d > c_r - \Delta$.

* A (location) point $z \in \mathbb{Z}$ is called *Markov signal point*, if

$$\delta_z^d > c_r - \Delta \quad \text{and} \quad \delta_z^M > c_r - \Delta.$$

* We call a Markov signal point z *regular*, if $\delta_z^M > c_r + \Delta$.

* Let $\bar{z}_1$ be the first Markov signal point in $[0, \infty)$. Let $\bar{z}_k$ be defined inductively: $\bar{z}_k$ is the first Markov signal point in $[\bar{z}_{k-1} + 2Ln^{1000}, \infty)$. Let $\bar{z}_0$ be the Markov signal point

in $(-\infty, 0]$ which laies closest to the origin. Let $\bar{z}_{-k}$ be defined inductively: $\bar{z}_{-k}$ is the right-most Markov signal point in $(-\infty, \bar{z}_{-(k-1)} - 2Ln^{1000}]$. Thus $\ldots, \bar{z}_{-2}, \bar{z}_{-1}, \bar{z}_0, \bar{z}_1, \bar{z}_2, \ldots$ is a sequence of ordered random variables which we call *signal carrier points.*

* For given z, the set

$$\mathcal{N}_z := [z - L(n^{1000} + e^{n^{0.3}}), z - L(n^{1000})] \cup [(z + Ln^{1000}, z + L(n^{1000} + e^{n^{0.3}})]$$

is called the *neighborhood* of z.

* We say that the neighborhood of z is *empty*, if $\mathcal{N}_z$ does not contain any block of ξ longer than $n^{0.35}$.
Thus, $\{\mathcal{N}_z \text{ is empty } \} \subseteq \sigma(\xi_i, i \in \mathcal{N}_z)$.

* We say that z has *empty border*, if the set $I_z - [z - \tilde{M}, z + \tilde{M}]$ does not contain any block of ξ longer than $n^{0.35}$. Thus, $\{\mathcal{N}_z \text{ is empty } \} \subseteq \sigma(\xi_i, i \in I_z - [z - \tilde{M}, z + \tilde{M}])$.

* Let p, $\tilde{p}$ and p^d be the probability, that a fixed point is a Markov signal point, a strong signal point or a signal point, respectively.
From (6.2.3), part a) of Proposition 12.4.9 and by (6.2.35) of Corollary 6.2.2 we know

$$p^d - \exp(-dn^{999}) < \tilde{p} \leq p^d; \tag{6.3.1}$$

$$p \leq p^d \leq \exp(-\frac{\alpha n}{\ln n}); \tag{6.3.2}$$

$$\frac{p^d}{p} \leq \exp((\ln n)^3). \tag{6.3.3}$$

* We now define a construction, which we are going to use later:
For each $j = 0, 1, 2, \ldots, 2Ln^{1000}$ partition the set $\mathbb{Z} \cap [-Ln^{1000} + j, \infty)$ into adjacent integer intervals of length $2Ln^{1000}$. Let $I_{k,j}$ denote the k-th interval of the partition who's first interval starts at $-Ln^{1000} + j$. Thus,

$$I_{1,j} = [j - Ln^{1000}, j + Ln^{1000}], \quad I_{2,j} = [j + Ln^{1000} + 1, j + 3Ln^{1000} + 1],$$

$$I_{3,j} = [j + 3Ln^{1000} + 2, j + 5Ln^{1000} + 2],$$

$$\ldots$$

$$I_{k,j} = [j + kLn^{1000} + k - 1, j + (k + 2)Ln^{1000} + k - 1].$$

Let $z_{j,k}$ denote the midpoints of $I_{k,j}$. Hence

$$z_{j,1} = j, \quad z_{j,2} = j + 2Ln^{1000} + 1, \quad \ldots, \quad z_{j,k} = j + 2kLn^{1000} + (k - 1).$$

For, each j, the intervals $I_{k,j}$, $k = 1, 2, \ldots$ are disjoint. Thus, the events

$$\{z_{k,j} \quad \text{is} \quad \text{a} \quad \text{Markov} \quad \text{signal} \quad \text{point}\}, \quad k = 1, 2, \ldots$$

are independent with the same probability p.

Let k' denote the integer valued random variable that shows the index of the first interval $I_{k,0}$ which has its midpoint being a Markov signal point. By such a counting we disregard the first interval. Thus, $k' > 1$ and, formally, k' is defined by the relations

$$\delta_{z_{2,0}} \wedge \delta^M_{z_{2,0}} \le c_r - \Delta, \ \ldots \quad \delta^M_{z_{k'-1,0}} \wedge \delta^d_{z_{k'-1,0}} \le c_r - \Delta, \ \delta^M_{z_{k',0}} \wedge \delta^d_{z_{k',0}} > c_r - \Delta$$

Clearly, $k'-1$ is a geometrical random variable with parameter p and, hence, $Ek' = \frac{1}{p} + 1$.

* Let Z be the location of the first Markov signal point after $2Ln^{1000}$. Recall $\bar{z}_1$ is the location of the first Markov signal point after 0. Note, that for each $i \ge 0$, we have

$$P(\bar{z}_1 \le i) < P(\cup_{j=0}^{i}\{i \text{ is a Markov signal point}\}) \le pi \tag{6.3.4}$$

and

$$P(Z \le i) \le p(i - 2Ln^{1000}), \quad i \ge 2Ln^{1000}. \tag{6.3.5}$$

¿From (6.3.4) and (6.3.2) we get

$$P(\bar{z}_1 \le 2Ln^{1000}) \le p2Ln^{1000} \le 2Ln^{1000} \exp(-\frac{\alpha n}{\ln n}) \to 0. \tag{6.3.6}$$

* We now estimate EZ. For this note: $Z \le z_{k',0} = 2k'Ln^{1000} + k' - 1$ and

$$EZ \le (\frac{1}{p} + 1)2Ln^{1000} + \frac{1}{p} \le \frac{3}{p}Ln^{1000}. \tag{6.3.7}$$

¿From (6.3.3) we get

$$EZp^d \le 3\frac{p^d}{p}Ln^{1000} \le 3Ln^{1000} \exp((\ln n)^3). \tag{6.3.8}$$

On the other hand by (6.3.5) we have, for each x, $EZ \ge xP(Z \ge x) \ge x(1 - px)$. Now, take $x = (2p)^{-1}$ and use (6.3.2) to get

$$EZ \ge \frac{1}{4p} \ge \frac{1}{4}\exp(\frac{\alpha n}{\ln n}). \tag{6.3.9}$$

* Take $m(n) = \ulcorner n^{2.5} EZ \urcorner$.

By (6.3.3) and b) of Proposition 12.4.9 we get

$$n^{2.5}EZ \le \frac{3Ln^{1002.5}}{p^d} \exp((\ln n)^3) \le 3Ln^{1002.5} \exp((\ln n)^3 + n) < \exp(2n),$$

implying

$$\frac{1}{4}\exp(\frac{\alpha n}{\ln n}) \le m < \exp(2n), \tag{6.3.10}$$

provided n is big enough.

* Next, we define the random variables which we are using later:

$$X_z := I_{\{\delta^d_z > c_r - \Delta, \quad \delta^M_i > c_r - \Delta\}}, \quad z = 0, 1, 2, \ldots.$$

Thus, X_z indicates, whether z is a Markov signal point or not. The random variables X_z are identically distributed with mean p.
We estimate the number of Markov signal points in $[0, cm]$, where $c > 1$ is a fixed integer, not depending on n . For this define:

$$E_0 := \Big\{ \sum_{z=0}^{cm} X_z \le n^{10000} \Big\}.$$

Thus, when E_0 holds, the interval $[0, cm]$ contains at most n^{10000} Markov signal points.
To estimate $P(E_0)$ we use the Markov inequality and (6.3.7)

$$\begin{aligned} P(E_0^c) = P\Big(\sum_{i=0}^{cm} X_i > n^{10000} \Big) &< \frac{(cm+1)p}{n^{10000}} \le \frac{c(n^{2.5}EZ+1)p+1}{n^{10000}} \\ &< c3Ln^{1002.5-10000} + (c+1)n^{-10000} = o(1). \end{aligned}$$

***** Finally, define $Z_0 < Z_1 < \cdots < Z_k < \cdots$ as follows:

$Z_0 := 0$, $Z_1 := Z$, and, let Z_{k+1} be the first Markov signal point that is greater than $2Ln^{1000} + Z_k$.

Note the differences: Z, $Z_2 - Z_1$, $Z_3 - Z_2$, ..., $Z_{k+1} - Z_k$, ... are i.i.d. Also:

$$\{\text{No Markov signal points in } [0, 2Ln^{1000}]\} = \{Z_i = \bar{z}_i \text{ for all } i\} := E_s^n. \tag{6.3.11}$$

From (6.3.6) we know that

$$P(E_s^n) \to 1. \tag{6.3.12}$$

6.3.2 Scenery-dependent events

Next, we describe the typical behavior of the signal points in the interval $[0, cm]$. Here $c > 1$ is a fixed integer, not depending on n. Among others we show that, with high probability, for each signal carrier point $\bar{z}_i$ in $[0, cm]$, the corresponding frequency of ones, $h(\bar{z}_i)$, vary more than $e^{-n^{0.11}}$ (events $\bar{E}_3^n$ and $\bar{E}_4^n$ below). We also show that, with high probability, all signal points in $[0, cm]$ have empty neighborhood.
All the properties listed below depend on the scenery ξ only. Therefore we refer to them as the *scenery dependent events.*
We now define all scenery dependent events, $\bar{E}_1^n, \ldots, \bar{E}_9^n$ and prove the convergence of their probabilities. All the events will be defined on the interval $[0, cm]$, where $c > 1$ is a fixed integer. Thus, if a point z is such that $\mathcal{N}_z \not\subset [0, cm]$, by the neighborhood of z, we mean $\mathcal{N}_z \cap [0, cm]$. This means $\bar{E}_i^n \in \sigma(\xi_z : z \in [0, cm])$. The exact value of c will be defined in the next chapter (in connection with the event $E_{2,S}^n$). During this chapter, c is assumed to be any fixed integer bigger than 1.
At first, we list the events of interest:

$\bar{E}_1^n := \{\bar{z}_{n^2+1} \leq m\}$;

$\bar{E}_2^n := \{\text{every signal point in } [0, cm] \text{ has an empty neighborhood}\}$;

$\bar{E}_3^n := \{\text{every pair } \bar{z}_1, \bar{z}' \text{ of signal carrier points in } [0, cm] \text{ satisfies : } |h(\bar{z}) - h(\bar{z}')| \geq e^{-n^{0.11}} \text{ if } \bar{z} \neq \bar{z}'\}$;

$\bar{E}_4^n := \{\text{every signal carrier point } \bar{z}, \text{ in } [0, cm] \text{ satisfies : } |h(\bar{z}) - \frac{1}{2}| \geq e^{-n^{0.11}}\}$;

$\bar{E}_5^n := \{\text{every signal point } z \in [0, cm] \text{ satisfies } \delta_z^M \notin [c_r - \Delta, c_r + \Delta]\}$;

$\bar{E}_6^n := \{\text{for all signal carrier points } \bar{z}_i \text{ in } [0, cm] \text{ we have } EZn^{11001} \geq |\bar{z}_i - \bar{z}_{i+1}| \geq EZn^{-11001}\}$;

$\bar{E}_7^n := \{\text{no signal carrier points in } [m - EZn^{-11001}, m + EZn^{-11001} \wedge cm] \cup [0, EZn^{-11001}]\}$;

$\bar{E}_8^n := \{\text{every strong signal point in } [0, cm] \text{ has empty border}\}$;

$\bar{E}_9^n := \{\text{every signal point in } [0, cm] \text{ is a strong signal point}\}$.

Proof that $P(\bar{E}_1^n) \to 1$

If $\bar{E}_1^n$ holds, then in $[0, m]$ we have more than n^2 signal carrier points .

Define the random variables $Z_0 < Z_1 < \cdots < Z_k < \cdots$ as in (6.3.11). Let $E_{1a}^n := \{Z_{n^2+1} \leq m\}$. Since $E_s \cap E_{1a}^n \subseteq \bar{E}_1^n$, it suffices to show that $P(E_{1a}^n) \to 1$. To see this, we use the Markov inequality:

$$P(E_{1a}^{nc}) = P(Z_{n^2+1} > m) \leq \frac{EZ_{n^2+1}}{m} \leq \frac{(n^2+1)}{n^{2.5}} \to 0.$$

Proof that $P(\bar{E}_2^n) \to 1$

$$\bar{E}_2^{nc} = \{\text{there exists a signal point in } [0, cm] \text{ with non} - \text{empty neighborhood}\}.$$

Clearly,

$$\bar{E}_2^{nc} = \cup_{z=0}^{cm} E_2(z), \quad \text{where} \quad E_2(z) := \{z \text{ is a signal point and } \mathcal{N}_z \text{ is not empty}\}.$$

For each z, the events $\{\mathcal{N}_z \text{ is empty}\}$ and $\{\delta_z > c_r - \Delta\}$ are independent. Thus, for each z,

$$P(E_2(z)) = P(\delta_z > c_r - \Delta) P(\mathcal{N}_z \text{ is empty}) = p^d P(\mathcal{N}_z \text{ is not empty}).$$

We obviously have $P(\mathcal{N}_z \text{ is empty}) = P(\mathcal{N}_o \text{ is empty})$ and

$$P(\mathcal{N}_o \text{ is not empty}) =$$
$$P(\mathcal{N}_o \text{ contains at least one block longer than } n^{0.3}) < 2L \exp(n^{0.3}) 2^{-n^{0.35}}.$$

Hence, from (6.3.8):

$$P(\bar{E}_2^{nc}) \le cmp^d 2L\exp(n^{0.3})(\frac{1}{2})^{n^{0.35}} \le 6cn^{2.5}L^2n^{1000}\exp((\ln n)^3+n^{0.3})2^{-n^{0.35}}$$
$$= 6cL^2n^{1002.5}\exp(n^{0.3}+(\ln n)^3)2^{-n^{0.35}} \to 0,$$

if $n \to \infty$.

Proof that $P(\bar{E}_8^n) \to 1$

For each z, the events $\{\delta_z^d > c_r - \Delta\}$ and $\{z$ has empty border $\}$ are independent. Now use the same argument as in the previous proof.

Proof that $P(\bar{E}_5^n) \to 1$

Note

$$\bar{E}_5^{nc} = \{\text{there exists a non} - \text{regular Markov signal point } z \in [0, cm]\}.$$

As in the previous proof, write:

$$\bar{E}_5^n = \cup_{z=0}^{cm} E_5(z), \quad \text{where} \quad E_5(z) := \{z \text{ is a non} - \text{regular Markov signal point}\}.$$

For each z,

$$P(E_5^c(z)) = P(\delta_z^M \wedge \delta_z^d > c_r - \Delta)P(\delta_z^M \le c_r + \Delta | \delta_z^M \wedge \delta_z^d > c_r - \Delta)$$
$$= pP(\delta_z^M \le c_r + \Delta | \delta_z^M \wedge \delta_z^d > c_r - \Delta).$$

¿From (6.2.36) of Corollary 6.2.2 we have:

$$P(\delta_z^M \le c_r + \Delta | \delta_z^M \wedge \delta_z^d > c_r - \Delta) \le n^{-10^5}.$$

Thus, from (6.3.7) $P(\bar{E}_5^{nc}) \le cmpn^{-10^5} \le c(n^{2.5}EZ+1)pn^{-10^5} = c3Ln^{1002.5-100000} + cpn^{-10^5} \to 0$, as $n \to \infty$.

Proof that $P(\bar{E}_9^n) \to 1$

We use the same argument as in the previous proof. Note

$$\bar{E}_9^{nc} = \{\text{there exists a signal point } z \in [0, cm] \text{ that is not a strong signal point}\}.$$

As in the previous proof, write

$$\bar{E}_9^{nc} = \cup_{z=0}^{cm} E_9(z), \quad \text{where} \quad E_9(z) := \{z \text{ is a non} - \text{strong signal point}\}.$$

Recall (6.2.3): $\tilde{\delta}_z^d > \delta_z^d - \exp(-dn^{999})$. Since, for n big enough, $\exp(-dn^{999}) < \tilde{\Delta} = \Delta\exp(-2n)$, we get

$$\tilde{\delta}_z^d > \delta_z^d - \tilde{\Delta}.$$

Now, for each z,

$$P(E_9(z)) = P(\delta_z^d > c_r - \Delta)P(\tilde{\delta}_z^d \le c_r - \Delta | \delta_z^d > c_r - \Delta)$$
$$= p^d P(\tilde{\delta}_z^d \le c_r - \Delta | \delta_z^d > c_r - \Delta) \le p^d P(\delta_z^d - \tilde{\Delta} \le c_r - \Delta | \delta_z^d > c_r - \Delta)$$
$$\le p^d P(\delta_z^d \le c_r - \Delta + \tilde{\Delta} | \delta_z^d > c_r - \Delta).$$

By (6.2.37) of Corollary 6.2.2 we now have:

$$P(E_9(z)) \leq p^d \exp(-n).$$

Hence, by (6.3.8):

$P(\bar{E}_9^{nc}) \leq cmp^d \exp(-n) \leq p^d c(EZn^{2.5}+1)\exp(-n) \leq c3Ln^{1000} \exp(\ln n)^3 \exp(-n)+o(1) = o(1).$

Proof that $P(\bar{E}_6^n) \to 1$

Consider random variables $Z_0 < Z_1 < \cdots < Z_k < \cdots$ as in (6.3.11). Let $N = \max\{i : Z_i \leq cm\}$. Define

$$E_{6b}^n := \{Z_i - Z_{i-1} \leq EZn^{10001}, \quad i = 1, 2, \ldots, n^{1000}\} \tag{6.3.13}$$
$$\bar{E}_{6c}^n := \{Z_i - Z_{i-1} \geq EZn^{-11001}, \quad i = 1, 2, \ldots, n^{1000}\} \tag{6.3.14}$$

and note that:

$$E_s \cap E_{6b}^n \cap E_{6a}^n \cap \{N \leq n^{10000}\} \subseteq \bar{E}_6^n.$$

Since $E \subseteq \{N \leq n^{10000}\}$, we get $P(N \leq n^{10000}) \to 1$. We also know that $P(E_s) \to 1$. Thus, it suffices to show that $P(E_{6b}^{nc}), P(E_{6c}^{nc}) \to 0$ as $n \to \infty$. Now, by the Markov inequality, (6.3.5) and (6.3.7):

$$P(E_{6b}^{nc}) = P(\exists 1 \leq i \leq n^{10000} \text{ such that : } Z_i - Z_{i-1} > EZn^{10001})$$

$$\leq \sum_{i=1}^{n^{10000}} P(Z_i - Z_{i-1} > EZn^{10001}) = n^{10000} P(Z > EZn^{10001}) \leq$$

$$n^{10000} \tfrac{EZ}{EZn^{10001}} = \tfrac{1}{n};$$

$$P(E_{6c}^{nc}) = P(\exists\ 1 \leq i \leq n^{10000} \text{ such that : } Z_i - Z_{i-1} < EZn^{-11001})$$

$$\leq \sum_{i=1}^{n^{10000}} P(Z_i - Z_{i-1} < EZn^{-11001}) \leq n^{10000} P(Z < EZn^{-11001}) <$$

$$pEZn^{-1001} \leq 3Ln^{1000-1001} = \tfrac{3L}{n}.$$

Proof that $P(\bar{E}_7^n) \to 1$

Consider the event

$$\{\text{there is no signal carrier points in } [0, EZn^{11001}]\}.$$

Every signal carrier point is a Markov signal point. Hence, for the proof, it suffices to show, that with high probability there is no Markov signal points in the interval $[0, EZn^{11001}]$. Now, by (6.3.4) and (6.3.7)

$$P(\text{No Markov signal points in } [0, EZn^{11001}]) =$$

$$P(Z^o > EZn^{-11001}) \leq pEZn^{-11001} \leq 3Ln^{-11001+1000} = o(1).$$

Thus $P(\text{No Markov signal points in } [0, EZn^{-11001}]) \to 1$.
Now repeat the same argument for the intervals $[m, m-EZn^{-11001}]$ and $[m, m+EZn^{-11001}]$.

6.3.3 Proof of $P(\bar{E}_3^n) \to 1$ and $P(\bar{E}_4^n) \to 1$

The proof relies on the rate of convergence in the local central limit theorem (LCLT in sequel). In the next subsection we present some technical preliminaries related to the proof.

Some preliminaries

Let S be the symmetric random walk with span 1. Define: $p_N(k) = P(S(N) = k)$. The random walk S has lattice $+\backslash - z$,$z \in Z$; its variance is σ^2.
Local CLT ([?], page 197):

$$\sup_k \left| \sigma\sqrt{N} p_N(k) - \frac{1}{\sqrt{2\pi}} \exp\{-\frac{k^2}{2\sigma^2 N}\} \right| = O(\frac{1}{\sqrt{N}}) \tag{6.3.15}$$

or

$$\sup_k \left| p_N(k) - \frac{1}{\sigma\sqrt{N}\sqrt{2\pi}} \exp\{-\frac{k^2}{2\sigma^2 N}\} \right| = O(\frac{1}{N}).$$

Denote

$$q_N(k) := \frac{1}{\sigma\sqrt{N}\sqrt{2\pi}} \exp\{-\frac{k^2}{2\sigma^2 N}\} \quad |k| \le LN.$$

Let $t_N := (\ln N)^b$, $b > 1$.
We estimate:

$$\begin{aligned} |p_N^2(k) - q_N^2(k)| &\le (p_N(k) + q_N(k)) \sup_k |p_N(k) - q_N(k)| \\ &\le [2q_N(k) + O(\frac{1}{\sqrt{N}})]O(\frac{1}{N}) = O(\frac{1}{\sqrt{N}N}) \end{aligned}$$

and

$$\sum_{k>t_N+j}^{L\sqrt{N}} [p_N^2(k) - q_N^2(k)] \le (L\sqrt{N})O(\frac{1}{\sqrt{N}N}) = O(\frac{1}{N}), \quad j = -t_N, \cdots, t_N.$$

Estimate:

$$\begin{aligned} &\frac{p_N^2(k)}{\sum_{k>t_N+j} p_N^2(k)} \le \frac{p_N^2(k)}{\sum_{k>t_N+j}^{L\sqrt{N}} p_N^2(k)} \le \frac{q_N^2(k) + O(\frac{1}{N})}{\sum_{k>t_N+j}^{L\sqrt{N}} [p_N^2(k) - q_N^2(k)] + \sum_{k>t_N+j}^{L\sqrt{N}} q_N^2(k)} \\ &\le \frac{O(\frac{1}{N})}{\sum_{k>t_N+j}^{L\sqrt{N}} q_N^2(k) - O(\frac{1}{N})}, \end{aligned}$$

for all k and $j = -t_N \ldots, t_N$.
Now,

$$\sum_{k>t_N+j}^{L\sqrt{N}} q_N^2(k) = \frac{1}{2\sigma^2\pi N} \sum_{k>t_N+j}^{L\sqrt{N}} \exp(-\frac{k^2}{\sigma^2 N})$$

and

$$\sum_{k>t_N+j}^{L\sqrt{N}} \exp(-\frac{k^2}{\sigma^2 N}) \ge \sum_{k>2t_N}^{L\sqrt{N}} \exp(-\frac{k^2}{\sigma^2 N}) > \sum_{k>2t_N}^{L\sqrt{N}} \exp(-\frac{L^2}{\sigma^2}) = M(L\sqrt{N} - 2t_N).$$

Thus, for each $j = -t_N, \dots, t_N$,

$$\sup_k \frac{p_N^2(k)}{\sum_{k>t_N+j} p_N^2(k)} \le \frac{O(\frac{1}{N})}{\frac{K}{N}(L\sqrt{N} - 2t_N) - O(\frac{1}{N})} = l\frac{K_4}{K_1\sqrt{N} - K_2 t_N - K_3} = O(\frac{1}{\sqrt{N}}) \tag{6.3.16}$$

where K, K_1, K_2, K_3, K_4 are constants.

Let μ be a probability distribution on $\{-t_N, -t_N+1, \dots, 0, \dots, t_N-1, t_N\}$. Consider the convolutions

$$u_N(k) = \sum_{j=-t_N}^{t_N} p_N(k-j)\mu_j, \quad k = -(LN - t_N), \dots, LN + t_N. \tag{6.3.17}$$

If $p_N(k) \ge p_N(k+1)$ for all $k \ge 0$, then for each $k > t_N$, we have the bounds

$$p_N(k + t_N) \le u_N(k) \le p_N(k - t_N). \tag{6.3.18}$$

In this case,

$$\sum_{k>t_N}^{t_N+LN} u_N(k) \ge \sum_{l>2t_N}^{N} p_N(l).$$

And from (6.3.16), taking $j = t_N$ we may deduce that:

$$\sup_{t_N<k} \frac{u_N^2(k)}{\sum_{k>t_N} u_N^2(k)} \le \sup_{0<k} \frac{p_N^2(k)}{\sum_{k>2t_N} p_N^2(k)} \le O(\frac{1}{\sqrt{N}}). \tag{6.3.19}$$

Generally, choose an atom $\lambda := \mu_j > 0$. Then

$$u_N(k) \ge \lambda p_N(k+j), \quad u_N^2(k) \ge \lambda^2 p_N^2(k+j)$$

and

$$\sum_{k>t_N}^{t_N+LN} u_N^2(k) \ge \lambda^2 \sum_{k>t_N+j}^{N} p_N^2(k). \tag{6.3.20}$$

Since $\sup_{k>t_N} u_N^2(k) \le \sup_{k>0} p_N^2(k)$, we get from (6.3.16):

$$\sup_{t_N \le k} \frac{u_N^2(k)}{\sum_{k>t_N} u_N^2(k)} \le \sup_k \frac{p_N^2(k)}{\lambda^2 \sum_{k>t_N+j} p_N^2(k)} = O(\frac{1}{N^{\frac{1}{4}}}). \tag{6.3.21}$$

In particular, from (6.3.21) follows:

$$\frac{\sum u_N^3(k)}{\sum u_N^2(k)\sqrt{\sum u_N^2(k)}} \le \max_k u_N(k) \frac{\sum u_N^2(k)}{\sum u_N^2(k)\sqrt{\sum u_N^2(k)}} \le \max_k \frac{u_N(k)}{\sqrt{\sum u_N^2(k)}} \le O\Big(\frac{1}{N^{\frac{1}{4}}}\Big). \tag{6.3.22}$$

Suppose that arrays $u_k := u_N(k)$ and $v_k := v_N(k)$, $t_N < k \le LN + t_N$ both satisfy (6.3.22). Then

$$\frac{\sum(u_k^3 + v_k^3)}{\sum(u_k^2 + v_k^2)\sqrt{\sum(u_k^2 + v_k^2)}} \le \max\{u_k, v_k\} \frac{\sum(u_k^2 + v_k^2)}{\sum(u_k^2 + v_k^2)\sqrt{\sum(u_k^2 + v_k^2)}} \tag{6.3.23}$$

$$\le \max\{\max_k \frac{u_k}{\sqrt{\sum u_k^2}}, \max_k \frac{v_k}{\sqrt{\sum v_k^2}}\} = O(N^{-\frac{1}{4}}) \tag{6.3.24}$$

Let us make one more observation. Since $\exp(\frac{-9t_N^2}{2\sigma^2 N}) \to 1$, there exists a $c' > 0$ such that

$$\exp(\frac{-9t_N^2}{2\sigma^2 N}) > c'$$

for each N big enough. Thus, there exists a constant $c > 0$ such that

$$p_N(k) > \frac{c}{\sqrt{N}}, \quad \forall |k| \le 3t_N.$$

Take λ as previously. Then

$$u_N(k) \ge p(k+j)\lambda \ge \frac{c\lambda}{\sqrt{N}}.$$

Hence there exists $C > 0$: $u(l) \ge \frac{C}{\sqrt{N}}$ $\forall l$ such that $|l + j| \le 3t_N$.
In particular

$$u_N(k) \ge \frac{C}{\sqrt{N}}, \quad -2t_N \ge k \le 2t_N. \tag{6.3.25}$$

Proof that $P(\bar{E}_3^n) \to 1$

Define the random variables z_1, z_2, ... as follows: z_1 is the first Markov signal point in $[0, \infty)$, z_k is the first Markov signal point in $[z_{k-1} + e^{n^{0.3}}, \infty)$. Note that a.s. there are infinitely many such points.
¿From the signal carrier part we know that, if each Markov signal point in $[0, cm]$ has empty neighborhood, i.e. $\bar{E}_2^n$ holds, then they form clusters which have radius at most $2Ln^{1000}$ and lie at least $e^{n^{0.3}}$ apart from each other. In this case all signal carrier points in $[0, cm]$ coincide with the z_i's defined above. We define the event:

$$E_{3a}^n := \Big\{\text{for each} \quad i, j \le n^{10000}, i \ne j \text{ we have } |h(z_i) - h(z_j)| \ge \exp(-n^{0.11})\Big\}.$$

Then:

$$E_{3a}^n \cap \bar{E}_2^n \cap E_0 \subseteq \bar{E}_3^n.$$

Since $P(E_{3a}^n \cap E_0) \to 1$, it suffices to show that $P(E_{3a}^n) \to 1$ as $n \to \infty$.

Let z_i, z_j, $i \ne j$. For simplicity denote them as z and z' Let

$$\epsilon_n := \exp(-n^{0.11}).$$

Consider the event:

$$E_n(i,j) := \{|h(z) - h(z')| \ge \epsilon_n\}.$$

For each $y \in Z$, define the random vector:

$$\xi_n(y) := \Big(\xi(y - Ln^{1000} - e^{n^{0.1}}), \xi(y - Ln^{1000} - e^{n^{0.1}} + 1), \ldots, \xi(y + Ln^{1000})\Big).$$

Now, let $\xi_n := \xi_n(z)$ and $\xi_n' := \xi_n(z')$. They are independent.

$$f_n := \sum_{k=z+Ln^{1000}+1}^{z+L(n^{1000}+e^{n^{0.1}})} u_n(k)\xi(k), \quad f_n' := \sum_{k=z'+Ln^{1000}+1}^{z'+L(n^{1000}+e^{n^{0.1}})} u_n'(k)\xi(k),$$

where

$$u_n(k) := \sum_{i=z-Ln^{1000}}^{z+Ln^{1000}} P(S_i(e^{n^{0.1}}) = k)\mu_i,\ u'_n(k) := \sum_{i=z'-Ln^{1000}}^{z'+Ln^{1000}} P(S_i(e^{n^{0.1}}) = k)\mu'_i$$

and μ_i, $i = z - Ln^{1000}, \cdots, z + Ln^{1000}$ and μ'_i, $i = z' - Ln^{1000}, \cdots, z' + Ln^{1000}$ denote the atoms of the stationary measure corresponding to z and z', respectively.
Recall that by (6.2.13)

$$h(z) := \sum_{k=z-L(n^{1000}+e^{n^{0.1}})}^{z+L(n^{1000}+e^{n^{0.1}})} u_n(k)\xi(k), \quad f'_n := \sum_{k=z'-L(n^{1000}+e^{n^{0.1}})}^{z'+L(n^{1000}+e^{n^{0.1}})} u'_n(k)\xi(k).$$

Note that conditioning on ξ_n, the coefficients $u_n(k)$ become constants.
(More precisely, f_n has the same distribution as

$$\tilde{f}_n := \sum_{k>Ln^{1000}}^{L(n^{1000}+e^{n^{0.1}})} \tilde{u}_n(k)\xi(k),$$

with

$$\tilde{u}_n(k) := \sum_{j=-Ln^{1000}}^{Ln^{1000}} P(S_j(e^{n^{0.1}}) = k)\tilde{\mu}_j = \sum_{j=-Ln^{1000}}^{Ln^{1000}} P(S(e^{n^{0.1}}) = k - j)\tilde{\mu}_j,$$

with $\tilde{\mu} := \{\tilde{\mu}_j\} := \{\mu_{z+j}\}$, $-Ln^{1000} \le j \le Ln^{1000}$ being a random probability measure independent of $\xi_{Ln^{1000}+1}, \ldots \xi_{e^{n^{0.1}}}$. In this setup, conditioning on ξ_n means conditioning on $\tilde{\mu}$.)
Hence

$$P\Big(\frac{f_n - Ef_n}{\sqrt{Df_n}} \le x|\xi_n\Big) = P\Big(\frac{\sum_{k>Ln^{1000}}^{L(e^{n^{0.1}}+N^{1000})} u_n(k)(\xi_(k) - \frac{1}{2})}{\frac{1}{2}\sqrt{\sum_{k>Ln^{1000}}^{L(e^{n^{0.1}}+N^{1000})} u_n^2(k)}} \le x|\xi_n\Big),$$

where $(u_n(k))$ are the fixed coefficients of type (6.3.17) (with $N = e^{n^{0.1}}$, $b = 10000$). Now the Berry-Esseen inequality for independent random variables (see, [?], Thm 3, p.111) states:

$$\sup_x \Big|P\Big(\frac{\sum u_n(k)(\xi(k) - \frac{1}{2})}{\frac{1}{2}\sqrt{\sum u_n^2(k)}} \le x|\xi_n\Big) - \Phi(x)\Big| \le A\frac{\sum u_n^3(k)}{\sum u_n^2(k)\sqrt{\sum u_n^2(k)}}, \tag{6.3.26}$$

with some constant A not depending on n and $u_n(k)$-s. By (6.3.22) (with $N = e^{n^{0.1}}$, $b = 10000$), the right side of (6.3.26) is bounded by $O(e^{\frac{-n^{0.1}}{4}})$. Here Φ stands for the standard normal distribution function.

By similar argument, conditioning on (ξ_n, ξ'_n) and using (11.5.18) instead of (6.3.22) yields:

$$\sup_x \Big|P\big(\frac{f_n - f'_n - \mu_n}{\sigma_n} \le x|\xi_n, \xi'_n\big) - \Phi(x)\Big| = O(e^{\frac{-n^{0.1}}{4}}), \tag{6.3.27}$$

with $\mu_n := E(f_n - f'_n)$, $\sigma_n := \sqrt{Df_n + Df'_n}$ where (f_n and f'_n are independent.)

Let $g_n := h_n - f_n$, $g'_n := h'_n - f'_n$. The event $E_n(i,j)$ can be written as:

$$E_n^c(i,j) := \{f_n - f'_n \in g_n - g'_n + [-\epsilon_n, \epsilon_n]\}.$$

Given ξ_n and ξ'_n, the random variable $g_n - g'_n$ is a constant. By (11.5.15) we have

$$P(E_n^c(i,j)|\xi_n, \xi'_b) = P\Big(\tfrac{f'_n - f_n - \mu_n}{\sigma_n} \in \tfrac{g_n - g'_n + [-\epsilon_n, \epsilon_n] - \mu_n}{\sigma_n} | \xi_n, \xi'_n\Big) \leq$$

$$2\sup_x \Big| P\big(\tfrac{f'_n - f_n - \mu_n}{\sigma_n} \leq x | \xi_n, \xi'_n\big) - \Phi(x)\Big| + \sup\Big\{\Phi(a) - \Phi(b)\Big| a - b = \tfrac{2\epsilon_n}{\sqrt{2\pi}\sigma_n}\Big\} \leq$$

$$O(e^{\frac{-n^{0.1}}{4}}) + \sqrt{\tfrac{2}{\pi}}\tfrac{\epsilon_n}{\sigma_n}.$$

Next, we estimate the standard deviation σ_n. For that note: because of (6.3.25) $u_n^2(z + Ln^{1000} + 1) \geq C^2 e^{-n^{0.1}}$, ${u'}_n^2(z' + Ln^{1000} + 1) \geq C^2 e^{-n^{0.1}}$ if n is big enough. Thus,

$$\sigma_n = \sqrt{Df_n + Df'_n} = \frac{1}{2}\sqrt{\sum u_N^2(k) + \sum {u'}_N^2(k)} > \frac{1}{2}\sqrt{2C^2 e^{n^{0.1}}} = \sqrt{2} C e^{\frac{-n^{0.1}}{2}}.$$

Hence, for n big enough there exists a constant $C_2 < \infty$ such that

$$\sqrt{\frac{2}{\pi}}\frac{\epsilon_n}{\sigma_n} \leq \frac{1}{\sqrt{\pi}} \exp(-n^{0.11} + \frac{n^{0.1}}{2}) \leq C_2 \exp(-n^{0.05}). \tag{6.3.28}$$

Thus, (6.3.28), (11.5.15) give:

$$P(E^n(i,j)) \leq O(e^{\frac{-n^{0.11}}{4}}) + O(e^{-n^{0.05}}) = O(e^{-n^{0.05}}).$$

By definition

$$E_{3a}^n = \cap_{i,j,i\neq j}^{n^{10000}} E^n(i,j)$$

and

$$P(E_{3a}^{nc}) \leq \sum_{i,j,i\neq j}^{n^{10000}} P(E^{nc}(i,j)) < n^{20000} O(e^{-n^{0.05}}) = o(1).$$

Outline of the proof that $P(\bar{E}_4^n)$ is close to one

Denote the Use (6.3.26) to get:

$$\begin{aligned}
P(\bar{E}_4^{nc}|\xi_n) &= P(|f_n + g_n - 0.5| \leq \epsilon_n | \xi_n) = P(f_n + g_n \in [0.5 - \epsilon_n, 0.5 + \epsilon_n] | \xi_n) \\
&= P(f_n \in [(0.5 - g_n) - \epsilon_n, 0.5 - g_n + \epsilon_n] | \xi_n) \\
&= P\Big(\frac{f_n - Ef_n}{\sqrt{Df_n}} \in \Big[\frac{0.5 - Ef_n - g_n - \epsilon_n}{\sqrt{Df_n}}, \frac{0.5 - Ef_n - g_n + \epsilon_n}{\sqrt{Df_n}}\Big] | \xi_n\Big) \\
&\leq 2\sup_x P\Big(\frac{f_n - Ef_n}{\sqrt{Df_n}} \leq x | \xi_n\Big) + \sup\Big\{\Phi(a) - \Phi(b)\Big| a - b = \sqrt{\frac{2}{\pi}}\frac{\epsilon_n}{\sqrt{Df_n}}\Big\} \\
&\leq O(e^{\frac{-n^{0.1}}{4}}) + \sqrt{\frac{2}{\pi}}\frac{\epsilon_n}{\sqrt{Df_n}} = O(e^{-n^{0.05}}),
\end{aligned}$$

because $\sqrt{Df_n} > C\exp(-\frac{n^{0.1}}{2})$. The rest of the proof goes as previously.

* In the following we consider the scenery dependent events defined on $[-cm, cm]$. For this, we define the events $\tilde{E}_i^n$, $i = 1, \ldots, 9$, where $\tilde{E}_i^n$ is defined exactly as $\bar{E}_i^n$, with $[-cm, 0]$ instead of $[0, cm]$.
Finally, we define the events:

$$E_i^n := \tilde{E}_i^n \cap \bar{E}_i^n.$$

The results of the present section show that $\forall\ i = 1, \ldots, 9$,

$$P(E_i^n) \to 0, \quad n \to \infty.$$

6.3.4 What is a signal carrier?

Let us briefly summarize the main ideas of the previous sections.
A signal carrier is a place in the scenery, where the probability to generate a block of $n^2 + 1$ times the same color is high. However, it is clear that such a place can not be too small. In the 3-color example the signal carrier depends on only one bit of the scenery. In the 2-color case, it takes many more bits to make the scenery (locally) atypical. We saw in Proposition 12.4.9, that for z to be a signal point, it is necessary that the interval I_z contains at least one big (longer than $n/\ln n$) block of ξ. Thus, if a point z is a (Markov, strong) signal point or not, depends on $\xi|I_z$.
If z is a signal point, then the scenery ξ is atypical in the interval I_z: δ_z^d is high. Thus, signal points would be our candidates for the signal carriers, if, for each z, we could estimate δ_z^d. The latter would be easy, if we knew when the random walk visits z. Then just take all such visits and consider the proportion of those visits that were followed by $n^2 + 1$ same colors after M steps. Unfortunately, we do not know when the random walk S visits z. But we do know (we observe) when S generates blocks with length at least n^2. Thus we can take these observations (times) as the visits of (the neighborhood of) z and estimate the probability of generating n^2+1 times the same color, M steps after previously observing $n^2 + 1$ times the same color. This idea yields the Markov signal probability. The problem now is to localize the area where the random walk (during a given time period) can generate n^2+1 times the same colors in the observations. If this area was too big, we could neither estimate the Markov signal probability nor understand where we are. To localize the described area, we showed (event E_2^n) that signal points have empty neighborhood. In the next section we shall see that the probability to generate a block of $n^2 + 1$ times the same color on the empty neighborhood is very small. This means, if S is close to a signal point z, then, with high probability, (and during a certain time period) all strings of $n^2 + 1$ times the same colors in the observations are be generated on I_z. The fact that all signal points have also empty borders (events E_8^n and E_9^n) makes the latter statement precise. Thus, a Markov signal point seems to be a reasonable signal carrier. But which one? Note, if z is a Markov signal point, i.e. I_z contains at least one big block, then, very likely the point $z + 1$ is a Markov signal point, too. In other words, Markov signal points come in clusters. However, when E_2^n holds, then each point in such a cluster has empty neighborhood. On the other hand, for z to be a Markov signal point, it is necessary to have at least one big block of ξ in I_z. This means that the diameter of every cluster of Markov signal points is at most $2Ln^{1000}$. The distances between the clusters are at least $L(e^{n^{0.3}} - n^{1000})$. Hence, in 2-color case one can think of signal carriers as clusters of

Markov signal points (provided E_2^n holds, but this holds with high probability). However, to make some statements more formal, for each cluster we have one representator, namely the signal carrier point. Since the diameters of the clusters are at most $2Ln^{1000}$, our definition of signal carrier points ensures that different signal carrier points belong to different clusters. If the cluster is located in $[0,\infty)$, then the signal carrier point is the left most Markov signal point in the cluster; if the cluster is located in $(-\infty,0)$, then the signal carrier point is the right most Markov signal point in the cluster. The event E_7^n ensures that there are no Markov signal points in the $2Ln^{1000}$-neighborhood of 0, so $\bar{z}_1$ and $\bar{z}_0$ belong to different clusters, too. The exact choice of a signal carrier point is irrelevant. However, it is important to note that given a cluster, everything that makes this cluster a signal carrier cluster (namely, the big blocks of scenery) is inside the interval $I_{\bar{z}}$, where $\bar{z}$ is the signal carrier point corresponding to the cluster. In particular, all blocks in the observations that are longer than n^2 will be generated on $I_{\bar{z}}$. This means that the signal carrier points, $\bar{z}_i$ (or the corresponding intervals $I_{\bar{z}_i}$) serve as signal carriers as well. At least, if we are able to estimate $\delta^M_{\bar{z}_i}$ with great precision. This is the subject of the next section.

6.4 Events depending on random walk

In the previous section we saw: if all scenery dependent events hold, then the signal carrier points are good candidates for the signal carriers. In this case the signal is an untypically high Markov signal probability. Hence, to observe this signal, we must be able to estimate the Markov signal probability. In the present section we define these estimators and in the next section we will see that they perform well, if the random walk S behaves typically. We describe the typical behavior of S in terms of several events depending on S. The main objective of the present section is to show that the (conditional) probability of such events tends to 1 as n tends to infinity.

6.4.1 Some preliminaries

As argued in Subsection 6.3.4, the main idea of the estimation of the Markov signal probability is very simple - given a time interval T, consider all blocks in the observations $\chi|_T$ that are bigger than n^2. Among these observations calculate the proportions of such blocks, that after exactly M steps, are followed by another such block. The time interval used for this estimation must be big enough to get a precise estimate but, on the other hand, it must be in correspondence with the size of an (empty) neighborhood. Recall that the neighborhood $\mathcal{N}_z$ consisted of two intervals of length $Le^{n^{0.3}}$. Hence, the optimal size of the interval T is $e^{n^{0.3}}$.
We now define the necessary concepts related to the described estimate - stopping times (that stop when a string of at least n^2+1 times the same color is observed) and the Bernoulli variables that show whether the stopping times are followed (after M step) by another such string or not. For technical reasons after stopping the process, we wait at least $e^{n^{0.1}}$ steps until we look for the next block.

***** Let $t>0$ and let $\hat{\nu}_t(1)$ be the smallest $s\geq t$ such that:

$$\chi(t)=\chi(t-1)=\cdots=\chi(t-n^2). \tag{6.4.1}$$

We define the stopping times $\hat{\nu}_t(i)$, $i = 2, 3, \ldots$ inductively: $\hat{\nu}_t(i)$ is the smallest $t \geq \hat{\nu}_t(i-1) + e^{n^{0.1}}$ such that (6.4.1) holds.

***** Let $X_{t,i}$ be the Bernoulli random variable that is one iff:

$$\chi(\hat{\nu}_t(i) + M) = \chi(\hat{\nu}_t(i) + M + 1) = ... = \chi(\hat{\nu}_t(i) + M + n^2).$$

Let $T = T(t) := [t, t + e^{n^{0.3}}]$. Define:

$$\hat{\delta}_T^M = \begin{cases} \frac{1}{e^{n^{0.2}}} \sum_{i=1}^{e^{n^{0.2}}} X_{t,i} & \text{if } \hat{\nu}_t(e^{n^{0.2}}) < t + e^{n^{0.3}} - e^{n^{0.1}} \\ 0 & \text{otherwise.} \end{cases} \tag{6.4.2}$$

***** We define some analogues of $\hat{\nu}_t$ and X_t.
Let $z \in \mathbb{Z}$ and $t \in \mathbb{N}$.

Let $\nu_{z,t}(1)$ designate the first time after t where we observe n^2 zero's or one's in a row, generated on the interval I_z. More precisely:

$$\nu_{z,t}(1) := \min \left\{ s > 0 \;\middle|\; \begin{array}{c} \chi(s) = \chi(s-1) = \cdots = \chi(s-n^2) \\ S(j) \in I_z, \ \forall j = s - n^2, \ldots, s \end{array} \right\}.$$

Once $\nu_{z,t}(i)$ is well defined, define $\nu_{z,t}(i+1)$ in the following manner:

$$\nu_{z,t}(i+1) := \min \left\{ t \geq \nu_{z,t}(i) + e^{n^{0.1}} \;\middle|\; \begin{array}{c} \chi(s) = \chi(s-1) = ... = \chi(s-n^2) \\ S(j) \in I_z, \ \forall j = s - n^2, \ldots, s \end{array} \right\}.$$

***** Let $X_{z,t,i}$, $i = 1, 2, \ldots$ designate the Bernoulli variable which is equal to one if exactly after time M the stopping time $\nu_{z,t}(i)$ is followed by a sequence of $n^2 + 1$ one's or zero's generated on I_z. More precisely, $X_{z,t,i} = 1$ iff

$$\chi(\nu_{z,t}(i) + M) = \chi(\nu_{z,t}(i) + M + 1) = \cdots = \chi(\nu_{z,t}(i) + n^2) \quad \text{and}$$
$$S(\nu_{z,t}(i) + M), \ldots, S(\nu_{z,t}(i) + n^{1000}) \in I_z.$$

Define

$$\hat{\delta}_{z,t}^M := \frac{1}{e^{n^{0.2}}} \sum_{i=1}^{e^{n^{0.2}}} X_{z,t,i}.$$

As argued in Subsection 7.3.14, $\{S(\nu_{z,t,i})\}$ is an ergodic Markov process with state space I_z and with the stationary measure I_z. Hence,

$$\frac{1}{j} \sum_{i=1}^{j} X_{z,t,i} \to \delta_z^M, \quad \text{a.s.}$$

Now we can apply some large deviation inequality to see that if $j \geq \exp(n^{0.2})$, then $\hat{\delta}_{z,t}^M$ gives a very precise estimate of δ_z^M.
The problem is that the random variables $X_{z,t,i}$ and, hence, the estimate $\hat{\delta}_{z,t}^M$ is *a priori* not observable. This is because we cannot observe whether a string of $n^2 + 1$ times the same color in the observations is generated on I_z or not. Thus, we can not observe neither

$\nu_{t,z}(i)$ nor $X_{t,z,i}$. However, the event $E^n_{3,S}$, stated below, ensures that with high probability $\hat{\delta}^M_{z,t}$ is the same as $\hat{\delta}^M_T$, provided that during the time interval T, the random walk S is close to z (the sense of closeness will specified later).

* We define the estimates for the frequency of ones. Again, we define a general, observable, estimate: $\hat{h}_t$ and its theoretical, *a priori* not-observable counterpart: $\hat{h}_{z,t}$.

Define

$$\hat{h}_t := \begin{cases} \frac{1}{e^{n^{0.2}}} \sum_{i=1}^{e^{n^{0.2}}} \chi(\nu_t(i) + e^{n^{0.1}}) & \text{if, } \hat{\nu}_t(e^{n^{0.2}}) < t + e^{n^{0.3}} - e^{n^{0.1}} \\ 0 & \text{otherwise.} \end{cases},$$

$$\hat{h}_{z,t} := \frac{1}{e^{n^{0.2}}} \sum_{i=1}^{e^{n^{0.2}}} \chi(\nu_{z,t}(i) + e^{n^{0.1}}).$$

* Finally, we define the stopping time that stop the walk, when a new signal carrier is visited.

Let $\ldots, \bar{z}_{-1}, \bar{z}_0, \bar{z}_1, \ldots$ denote the signal carrier-points in $\mathbb{R}$. Denote $I_i := I_{z_i}$ and let $\rho(k)$ denote the time of the k-th visit of S to one of the intervals I_i in the following manner: when an interval I_i is visited, then the next stop is on a different interval.
More precisely, let $\rho(0)$ be the first time $t \geq 0$ such that $S(t) \in \cup_i I_i$. Denote $I(\rho(k))$ the interval I_i visited at time $\rho(k)$. Then define $\rho(k)$ inductively:

$$\rho(k+1) = \min\{t > \rho(k) | S(t) \in \cup_i I_i, \quad S(t) \notin I(\rho(k))\}.$$

6.4.2 Random walk-dependent events

In this section, we define the events that characterize the typical behavior of the random walk S on the typical scenery on the interval $[-cm, cm]$. The (piece of) scenery $\xi|[-cm, cm]$ is typical if it satisfies all the scenery-dependent events E^n_i, $i = 1, \ldots, 9$. Recall, that the events E^n_i are the same as the events $\bar{E}^n_i$ defined in Section 4.2 with $[0, cm]$ replaced by $[-cm, cm]$. Also recall that $c > 1$ is an arbitrary fixed constant not depending on n, and $m = \ulcorner n^{2.5} EZ \urcorner$.
Hence, throughout the section we consider the sceneries belonging to the set:

$$E_{\text{cell\_OK}} := \cap_{i=1}^9 E^n_i. \tag{6.4.3}$$

Clearly, $E_{\text{cell\_OK}}$ depends on n. We know that $P(E_{\text{cell\_OK}}) \to 1$ if $n \to \infty$.

Let $\psi : \mathbb{Z} \to \{0, 1\}$ be a (non random) scenery. Let $P_\psi(\cdot)$ designate the measure obtained by conditioning on $\{\xi = \psi\}$ and $\{S(m^2) = m\}$. Thus,

$$P_\psi(\cdot) := P(\cdot | \xi = \psi, S(m^2) = m). \tag{6.4.4}$$

Let $P(\cdot|\psi)$ denote $P(\cdot|\xi = \psi)$.
We now list the events that describe the typical behavior of S. The objective of the section is to show: if n is big and $\psi_n :=: \psi \in E_{\text{cell\_OK}}$ then all listed events have big conditional

probabilities P_ψ. The events depending on the random walk are:

$$E^n_{1,S} := \{S(m^2) = m\};$$

$$E^n_{2,S} := \{\forall t \in [0, m^2] \text{ we have that } S(t) \in [-cm, cm]\};$$

$$E^n_{3,S} := \{\forall t \in [0, m^2], \text{ it holds : } \hat{\delta}^M_T \le c_r, \text{ if } \delta^d_{S(s)} \le c_r - \Delta \ \forall s \in T(t)\};$$

$$E^n_{4,S} := \{\rho(n^{25000}) \ge m^2\};$$

$$E^n_{5,S} := \{\forall k \le n^{25000} \text{ we have: if } \rho(k) \le m^2 \text{ then } \hat{\nu}_{\rho(k)}(e^{n^{0.2}}) \le \rho(k) + e^{n^{0.3}} - e^{n^{0.1}}\};$$

$$E^n_{6,S} := \left\{ \begin{array}{c} \text{for any } t \in [0, m^2] \text{ satisfying } \chi(t) = \cdots = \chi(t + n^2) \\ \text{there exists } s \in [t, t + n^2] \text{ such that } S(s) \\ \text{is contained in a block of } \xi \text{ bigger than } n^{0.35} \end{array} \right\};$$

$$E^n_{7,S}(z,t) = \left\{ \left| \hat{\delta}^M_{z,t} - \delta^M_z \right| < e^{-n^{0.12}} \right\}, \quad z \in \mathbb{Z}, \ t > 0;$$

$$E^n_{7,S} := \cap_{z=-cm}^{cm} \cap_{t=0}^{m^2} E^n_{13,S}(z,t);$$

$$E^n_{8,S}(z,t) = \left\{ \left| \hat{h}_{z,t} - h(z) \right| < e^{-n^{0.12}} \right\}, \quad z \in \mathbb{Z}, \ t > 0;$$

$$E^n_{8,S} := \cap_{z=-cm}^{cm} \cap_{t=0}^{m^2} E^n_{8,S}(z,t);$$

We now estimate the conditional probabilities of all listed events. In most cases we prove statements like $P_\psi(E^n_{j,S}) \to 1$. This means: for an arbitrary sequence $\psi_n \in E^n_{\text{cell\_OK}}$, we have:

$$\lim_{n\to\infty} P(E^n_{j,S} | S(m^2) = m, \xi = \psi_n) = 1.$$

6.4.3 Proofs

At first note that by LCLT, we have:

$$P(E_{1,S}) = \frac{1}{m} + O(\frac{1}{m^2}).$$

Clearly, $E_{1,S}$ does not depend on ξ, i.e. $P(E_{1,S}|\psi) = P(E_{1,S})$. Using (6.3.10) we get:

$$P(E_{1,S}|\psi) \ge \exp(-2n) - O(\exp(-4n)) \ge \exp(-3n). \tag{6.4.5}$$

¿From (6.4.5) follows that for any event E,

$$P_\psi(E) = \frac{P(E, S(m^2) = m|\psi)}{P(S(m^2) = m|\psi)} \le \frac{P(E|\psi)}{\exp(-3n)}. \tag{6.4.6}$$

Proposition 6.4.1. *For each $\epsilon > 0$ there exists $c(\epsilon)$, independent of n, such that for each ψ, $P_\psi(E^n_{2,S}) \geq 1 - \epsilon$, provided n is big enough.*

Proof. At first note, that, for each n, the event $E^n_{2,S}$ is independent of the scenery ψ. Thus,

$$P_\psi(E^n_{2,S}) = P(E^n_{2,S} | S(m^2) = m).$$

Define

$$E^n_a(c) = \{\forall t \in [0, m^2] \text{ we have that } S(t) \leq cm\}$$

$$E^n_b(c) = \{\forall t \in [0, m^2] \text{ we have that } S(t) \geq -cm\}$$

Clearly,

$$E^n_{2,S} = E^n_a(c) \cap E^n_b(c).$$

We now find c, not depending on n such that $P_\psi(E^{nc}_a(c)), P_\psi(E^{nc}_b(c)) \leq \frac{\epsilon}{2}$.
Let us define the stopping time ϑ:

$$\vartheta := \min\{t | S(t) > cm\}.$$

Let for all $j \in 1, \ldots, L$:

$$p_j := P\Big(S(m^2) = m,\ \vartheta \leq m^2 \text{ and } S(\vartheta) = cm + j\Big)$$

We have that

$$P\left(E^{nc}_a(c) \cap E^n_{1,S}\right) = \sum_{j=1}^{L} p_j.$$

Our random walk S is symmetric. By the reflection principle, for all $j \in 1, \ldots, L$, we have:

$$p_j = P(S(m^2) = cm + j + (cm + j - m) = 2cm + 2j - m,\ \vartheta \leq m^2 \text{ and } S(\vartheta) = cm + j).$$

Thus $p_j \leq P\left(S\left(m^2\right) = 2cm - m + 2j\ \right)$ and

$$P\left(E^{nc}_a(c) \cap E^n_{1,S}\right) \leq \sum_{j=1}^{L} P\left(S(m^2) = m(2c-1) + 2j\right). \tag{6.4.7}$$

By LCLT, for big m, the right side of (6.4.7) can be made arbitrary small in comparison to $P\left(S\left(m^2\right) = m\right)$ by taking c big enough. In other words, there exists c , not depending on n such that:

$$\frac{\sum_{j=1}^{L} P\left(S\left(m^2\right) = 2cm + m + 2j\ \right)}{P\left(S\left(m^2\right) = m\right)} \leq \frac{\epsilon}{2}.$$

This means

$$\frac{P\left(E^{nc}_a(c) \cap E^n_{1,S}\right)}{P\left(E^n_{1,S}\right)} = P_\psi\left(E^{nc}_a(c)\right) \leq \frac{\epsilon}{2}.$$

Similar argument gives $P_\psi\left(E^{nc}_b(c)\right) \leq \frac{\epsilon}{2}$. □

* Note, that the choice of c does not depend on n. From now on, we fix c such that Proposition 6.4.1 holds with $\epsilon > \frac{1}{8}$. This particular c is used in the definition of all scenery-dependent events and, therefore, in the definition of $E_{\text{cell\_OK}}$ as well as in the definitions $E^n_{4,S}$, $E^n_{5,S}$.

* In what follows, we often use these versions of the Hoeffding inequality:
Let $X_1, \ldots, X_N$ be independent random variables with range in $[a, b]$. Let S_N denote their sum. Then:

$$\begin{aligned} P(|S_N - ES_N| \geq \epsilon) &\leq 2\exp(-2\frac{\epsilon^2}{N(b-a)^2}) \leq \exp(-\frac{d'\epsilon^2}{N}); \\ P(\frac{1}{N}|S_N - ES_N| \geq \epsilon) &\leq 2\exp(-2\frac{\epsilon^2 N}{(b-a)^2}) \leq \exp(-d'\epsilon^2 N). \end{aligned} \tag{6.4.8}$$

For our random walk, this gives:

$$\begin{aligned} P(|S(N)| \geq \epsilon) &\leq 2\exp(-\frac{\epsilon^2}{4L^2N}) \leq \exp(-\frac{d\epsilon^2}{N}) \\ P(|\frac{S(N)}{N}| \geq \epsilon) &\leq 2\exp(-\frac{\epsilon^2 N}{4L^2}) \leq \exp(-d\epsilon^2 N), \end{aligned} \tag{6.4.9}$$

for some $d', d > 0$.

We also use the following results: let $X_1, \ldots, X_N$ be i.i.d. random variables with mean 0 and finite variance σ^2. Let $M_n^+ = \max_{i=1,\ldots,N} S_i$, $M_n = \max_{i=1,\ldots,N} |S_i|$. Then

$$\frac{M_N^+}{\sigma\sqrt{N}} \Rightarrow \sup_{0\leq t\leq 1} W_t, \quad \text{and} \quad \Big(\frac{M_N}{\sigma\sqrt{N}}, \frac{S(N)}{\sigma\sqrt{N}}\Big) \Rightarrow (\sup_{0\leq t\leq 1} |W_t|, W(1)), \tag{6.4.10}$$

where W_t is standard Brownian motion. It is well-known that $\forall x > 0$, $P(\sup_{0\leq t\leq 1} W_t \leq x) = 2\Phi(x) - 1$.

Proof that $\liminf_n P_\psi(E^n_{4,S}) \geq 1 - \frac{1}{8}$.

For each n, fix an arbitrary $\psi_n \in E^n_{\text{cell\_OK}}$. Since $\psi_n \in E^n_{\text{cell\_OK}} \subseteq E^n_6$, we have that for every signal carrier point $\bar{z}_i \in [-cm, cm]$:

$$\bar{z}_{i+1} - \bar{z}_i, \bar{z}_i - \bar{z}_{i-1} \geq EZn^{-11001}. \tag{6.4.11}$$

For this proof, let $\mu := EZ$ and $N(n) := \mu^2 n^{-24000}$. Since $m \leq n^{2.5}\mu + 1$, we have $n^{25000} \times N = n^{25000} \times \mu^2 n^{-24000} = \mu^2 n^{1000} > m^2$. Hence, if $E^n_{4,S}$ fails, then there must be at least one $k \in \{0, \ldots, n^{25000} - 1\}$ such that $\rho(k+1) - \rho(k) < N$. Moreover, if $E^n_{4,S}$ fails, then for each $k \in \{0, \ldots, n^{25000} - 1\}$ it holds $\rho(k) \leq m^2$. We formalize this observation. Let:

$$\begin{aligned} E_{a,4}(k) &:= \{\rho(k+1) - \rho(k) \geq N, \quad \rho(k) \leq m^2\} \\ E_{a,4} &:= \cap_{k=0}^{n^{25000-1}} E_{a,4}(k). \end{aligned} \tag{6.4.12}$$

We have:

$$E^{nc}_{4,S} \subseteq E^c_{a,4}. \tag{6.4.13}$$

By Proposition 6.4.1, for n big enough, $P_\psi(E^n_{2,S}) \le \frac{1}{8}$. Thus,

$$P_\psi(E^{nc}_{4,S}) \le P_\psi(E^c_{4,a} \cap E^n_{2,S}) + P_\psi(E^{nc}_{2,S}) \le \frac{1}{8} + \sum_{k=0}^{n^{25000-1}} P_\psi(E^c_{a,4}(k) \cap E^n_{2,S}). \tag{6.4.14}$$

We now bound $P_\psi(E_{a,4}(k))$.

Suppose $E^n_{2,S}$ holds. Then $\rho(k) \le m^2$ implies that the signal carrier visited at time $\rho(k)$ is in $[-cm, cm]$. By (6.4.11) this means that the closest signal carrier point is at least at distance μn^{-11001}. Let I_i be $I(\rho(k))$. Then

$$\inf\{|t-s| : t \in I_i, s \in I_j\} \ge \mu n^{-11001} - 2Ln^{1000}, \tag{6.4.15}$$

where $j \in \{i-1, i+1\}$. By (6.3.9), $\mu^2 > n^{25000}$. Then $\mu > n^{12500} \ge 2Ln^{12002}$, implying

$$\mu n^{-11001} - 2Ln^{1000} \ge \mu n^{-11002}. \tag{6.4.16}$$

We consider the event

$$E_{a,4}(k)^c \cap E^n_{2,S} \subseteq \{\rho(k+1) - \rho(k) < N, \quad S(\rho(k)) \in [-cm, cm]\}.$$

¿From (6.4.15) and (6.4.16) it follows that:

$$\begin{aligned}
&P\Big(\rho(k+1) - \rho(k) < N, \quad S(\rho(k)) \in [-cm, cm]\Big|\psi_n\Big) \le \\
&P\big(\rho(k+1) - \rho(k) < N\Big|S(\rho(k)) \in [-cm, cm], \xi = \psi_n\big) \le \\
&P\big(\sup_{l \le N} |S(l)| > \mu n^{-11001} - 2Ln^{1000}\Big) \le P(\sup_{l \le N} |S(l)| > \mu n^{-11002}\}.
\end{aligned}$$

Use the following maximal inequality:

$$P(\max_{l \le N} |S(l)| > \mu n^{-11002}) \le 3 \max_{l \le N} P\Big(|S(l)| > \frac{\mu}{3} n^{-11002}\Big). \tag{6.4.17}$$

By the Hoeffding inequality, for each $l \le N$:

$$\begin{aligned}
P\Big(|S(l)| > \frac{\mu}{3} n^{-11002}\Big) &\le \exp\big(-\frac{d\mu^2 n^{-22004}}{9l}\big) \le \exp\big(-\frac{d\mu^2 n^{-22004}}{9N}\big) \\
&\le \exp\big(-\frac{dn^{24000-22004}}{9}\big) = \exp\big(-\frac{dn^{1996}}{9}\big).
\end{aligned}$$

Hence,

$$P(E_{a,4}(k)|\psi) \le \exp\big(-\frac{dn^{1996}}{9}\big), \ P(E_{a,4}|\psi) \le n^{25000} \exp\big(-\frac{dn^{1996}}{9}\big).$$

By (6.4.6), we get

$$P_\psi(E^{nc}_{a,4}) \le n^{25000} \exp\big(3n - \frac{dn^{2996}}{9}\big).$$

The right side of the last inequality tends to 0 if $n \to \infty$. Relation (6.4.13) now finish the proof.

Proof that $P_\psi(E^n_{3,S}) \to 1$

Let $t \geq 0$ be an integer and define the stopping times $\hat{\nu}^o_t(1), \hat{\nu}^o_t(2), \ldots$ as follows: $\hat{\nu}^o_t(1)$ is the smallest time $s \geq t + e^{n^{0.1}}$ such that:

$$\chi(s - n^2) = \chi(s - n^2 + 1) = \cdots = \chi(s) \text{ and } \delta^d_{S(s)} \leq c_r - \Delta. \tag{6.4.18}$$

Once $\hat{\nu}^o_t(k)$ is well defined, define $\hat{\nu}^o_t(k+1)$ to be the smallest time $s \geq \hat{\nu}^o_t(k) + e^{n^{0.1}}$ such that (6.4.18) holds.

Let $X^o_{t,k}$ be the Bernoulli variable which is equal to one if and only if

$$\chi(\hat{\nu}^o_t(k) + M) = \chi(\hat{\nu}^o_t(k) + M + 1) = \cdots = \chi(\hat{\nu}^o_t(k) + M + n^2).$$

Finally define:

$$\hat{\delta}^M_{o,t} := \frac{1}{e^{n^{0.2}}} \sum_{k=1}^{e^{n^{0.2}}} X^o_{t,k}.$$

Let

$$E^n_{3,S}(t) := \left\{ \hat{\delta}^M_{o,t} < cr \right\}.$$

Clearly,

$$\bigcap_{t \in 0,\ldots,m^2} E^n_{3,S}(t) \subseteq E^n_{3,S}, \quad \text{imlpying} \quad P(E^{nc}_{3,S}|\psi) \leq \sum_{t=0}^{m^2} P(E^{nc}_{3,S}(t)|\psi), \tag{6.4.19}$$

where ψ is an arbitrary fixed scenery.

Note, for any fixed scenery ψ, the random variables $X^o_{t,1}, X^o_{t,2}, \ldots$ are clearly independent (but not necessarily identically distributed). However, for each i, $E(X^o_{t,i}|\psi) \leq c_r - \Delta$, implying that

$$c_r - \frac{1}{e^{n^{0.2}}} \sum_{i=1}^{e^{n^{0.2}}} E(X^o_{t,i}|\psi) \geq \Delta.$$

Recall $\Delta = \frac{p_M}{n^{10054}}$. We know that $\Delta \geq n^{-\beta}$, where β is an integer bigger than 11000. Thus, by (6.4.8)

$$\begin{aligned} P(E^{nc}_{3,S}(t)|\psi) = P(\hat{\delta}^M_{o,t} \geq c_r|\psi) &= P\Big(\frac{1}{e^{n^{0.2}}} \sum_{i=1}^{e^{n^{0.2}}} X^o_{t,i} \geq c_r|\psi\Big) \\ &\leq P\Big(\frac{1}{e^{n^{0.2}}} \sum_{i=1}^{e^{n^{0.2}}} (X^o_{t,i} - EX^o_{t,k}) \geq \Delta|\psi\Big) \leq \exp(-d'\Delta^2 e^{n^{0.2}}) \\ &\leq \exp\big(-(d' n^{-2\beta} e^{n^{0.2}})\big). \end{aligned}$$

Now, use (6.4.6), (6.4.19) and (6.3.10) to get

$$P_\psi(E^{nc}_{3,S}) \leq m^2 \exp(-d' n^{-2\beta} e^{n^{0.2}} + 3n) \leq \exp\big(7n - (d' n^{-2\beta} e^{n^{0.2}})\big) \to 0,$$

as $n \to \infty$.

Proof that $P_\psi(E^n_{6,S}) \to 1$

Let:

$$E^n_{6,S}(t) = \left\{ \begin{array}{c} \text{if } \chi(t) = \chi(t+1) = \cdots = \chi(t+n^2) \\ \text{then } \exists s \in [t, t+n^2] \text{ such that} \\ S(s) \text{ is contained in a block of } \xi \text{ longer than } n^{0.35} \end{array} \right\}.$$

We have that

$$E^n_{6,S} = \bigcap_{t \in [0,m^2]} E^n_{6,S}(t)$$

and thus

$$P_\psi(E^{nc}_{6,S}) \leq \sum_{t=0}^{m^2} P_\psi(E^{nc}_{6,S}(t)).$$

Note:

$$E^{nc}_{6,S}(t) = \left\{ \begin{array}{c} \forall s \in [t, t+n^2] \text{ the random walk } S(s) \\ \text{is contained in a block of } \xi \text{ with length at most } n^{0.35} \\ \text{and } \chi(t) = \chi(t+1) = \cdots = \chi(t+n^2) \end{array} \right\}.$$

Fix a scenery ψ. Let $I = \mathbb{Z}/ \cup B(\psi)$, where $B(\psi_n)$ is a block of ψ bigger than $n^{0.35}$ and the union is taken over all such blocks. Note $I = \cup_k I_k$, where I_k are disjoint intervals, at least $n^{0.35}$ far from each other. Thus, if $S(t) \in I_k$, then $S(t+s) \notin I_l$ for each $l \neq k$ and for each $s = 1, \ldots, n^2$.
Hence

$$P(E^{nc}_{6,S}(t)|\psi) = \sum_{j \in I} P\Big(S(t), \ldots, S(t+n^2) \in I \text{ and } \chi(t) = \cdots = \chi(t+n^2) | S(t) = j\Big) P(S(t) = j)$$

$$\sum_k \sum_{j \in I_k} P\Big(S_j(0), \ldots, S(n^2) \in I_k \text{ and } \chi(t) = \cdots = \chi(t+n^2)\Big) P(S(t) = j).$$

By Lemma 6.2.1, there exists a constant $a > 0$ not depending on n such that, for each j,

$$P\Big(S_j(0), \ldots, S(n^2) \in I_k \text{ and } \chi(t) = \cdots = \chi(t+n^2)\Big) \leq \exp\Big(-\frac{an^2}{n^{0.7}}\Big). \qquad (6.4.20)$$

Then,

$$P(E^{nc}_{6,S}(t)|\psi) \leq \exp(-an^{1.3}).$$

Thus, by (6.4.6):

$$P_\psi(E^{nc}_{6,S}(t)) \leq \exp\left(-an^{1.3} + 3n\right) \to 0$$

and by (6.3.10)

$$m^2 \exp(-an^{1.2} + 3n) \leq e^{7n - an^{1.3}} \to 0.$$

Proof that $P_\psi(E^n_{7,S}) \to 1$

Preliminaries

Recall that the definitions of stopping times involved:

a) $\vartheta_z(k)$, $k = 0, 1, \dots$ stands for consecutive visits of S to the point $z - 2Le^{n^{0.1}}$, provided that between $\vartheta_z(k)$ and $\vartheta_z(k+1)$ at least once $n^2 + 1$ same colors have been generated on I_z;

b) $\nu_z(1)$ ($\nu_z(i)$, $i = 2, 3, \dots$) is the first time after $\vartheta_z(0)$, (after $\nu_z(k-1) + e^{n^{0.1}}$) observing $n^2 + 1$ times the same color generated on I_z.

In Section 7.3.14 the stopping times $\vartheta_z(k)$, $\nu_z(i)$ as well as the random variables $X_{z,i}$ were used to define the random variables $\kappa_z(k)$, $\mathcal{X}_z(k)$ and $\mathcal{Z}_z(k)$. The latter were used to define δ_z^M.

We fix an arbitrary time t and define the counterparts of all the above-mentioned stopping times and random variables starting from t.

In Section 12.5.8 we already defined the t-counterpart of $\nu_z(i)$ and $X_{z,i}$, namely $\nu_{z,t}(i)$, and $X_{z,t,i}$, $i = 1, 2, \dots$.

Recall that in the definition of $\nu_{z,t}(1)$, the starting point $\vartheta_z(0)$ was replaced by t, the induction for $\nu_{z,t}(i)$ is the same as the one for $\nu_z(i)$, $i = 2, 3, \dots$.

The Bernoulli random variables $X_{z,t,i}$ were defined exactly as $X_{z,i}$ with the stopping times $\nu_{z,t}(i)$ instead of the $\nu_z(i)$'s.

We define the t-counterpart of $\vartheta_z(k)$, $k = 0, 1, \dots$.

***** Let $\vartheta_{z,t}(0) = t$ and let

$$\vartheta_{z,t}(k) := \{\min s > \vartheta_{z,t}(k-1) : S(s) = z - 2Le^{n^{0.1}},\ \exists j : s > \nu_{z,t}(j) > \vartheta_{z,t}(k-1)\}.$$

We use $\vartheta_{z,t}(k)$ to define the t-analogues of κ_z, $\mathcal{Z}_z$ and $\mathcal{X}_z$.

***** More precisely, let $\kappa_{z,t}(0) = 0$ and let $\kappa_{z,t}(k)$ be defined by the inequalities

$$\nu_{z,t}(\kappa_{z,t}(k)) < \vartheta_{z,t}(k) < \nu_{z,t}(\kappa_{z,t}(k) + 1).$$

The definition of $\mathcal{Z}_{z,t}$ and $\mathcal{X}_{z,t}$ is straightforward:

$$\mathcal{X}_{z,t}(k) = \sum_{i=\kappa_{z,t}(k-1)+1}^{\kappa_{z,t}(k)} X_{z,t,i}, \quad \mathcal{Z}_{z,t}(k) = \kappa_{z,t}(k) - \kappa_{z,t}(k-1), \quad k = 1, 2, \dots$$

Note that, if ξ is fixed, then, for all $t > 0$, the random variables $\mathcal{X}_{z,t}(1), \mathcal{X}_{z,t}(2), \dots$ are independent and the random variables $\mathcal{X}_{z,t}(2), \mathcal{X}_{z,t}(3), \dots$ are i.i.d. with the same distribution as $\mathcal{X}_z(k)$. The same holds for $\mathcal{Z}_{z,t}(1)$, $\mathcal{Z}_{z,t}(2), \dots$. Also note, that $\mathcal{Z}_{z,t}(k) \geq 1$, $k = 1, 2, \dots$.

Hence, for all $t > 0$,

$$\delta_z^M = \delta_z^M(\xi) = \frac{E(\mathcal{X}_{z,t}(2)|\xi)}{E(\mathcal{Z}_{z,t}(2)|\xi)} = \lim_{k\to\infty} \frac{\sum_{i=1}^k \mathcal{X}_{z,t}(i)}{\sum_{i=1}^k \mathcal{Z}_{z,t}(i)}.$$

We are now going to show that for each ξ, t, z, the first $e^{n^{0.2}}$ observations of $X_{z,t,i}$ are enough to estimate $\delta_z^M(\xi)$ very precisely, i.e. $\hat{\delta}_{z,t}^M$ is close to δ_z^M.

Fix z, t, ψ and define:

$$\mathcal{Z}_k := \mathcal{Z}_{z,t}(k), \quad \mathcal{X}_k := \mathcal{X}_{z,t}(k), \quad X_i := X_{k,t,i}, \quad E\mathcal{X} = E(\mathcal{X}_2|\psi), \quad E\mathcal{Z} = E(\mathcal{Z}_2|\psi), \quad P(\cdot) = P(\cdot|\psi).$$

Thus:

$$\delta_z^M = \delta_z^M(\psi) = \frac{E\mathcal{X}}{E\mathcal{Z}}.$$

Let $a = \ulcorner e^{3n^{0.1}} \urcorner$ and define:

$$\mathcal{Z}_k^a = \mathcal{Z}_k \wedge a, \quad \mathcal{X}_k^a = \mathcal{X}_k \wedge a, \quad E\mathcal{X}^a := E(\mathcal{X}_2^a|\psi), \quad E\mathcal{Z}^a := E(\mathcal{Z}_2^a|\psi).$$

Finally, define:

$$\Delta := e^{-\frac{n^{0.2}}{4}}.$$

We consider the events:

$$E_{7,a} = \left\{\mathcal{Z}_k \le a, \quad k = 1, 2, \dots, e^{n^{0.2}}\right\},$$
$$E_{7,b} = \left\{\left|\frac{\mathcal{X}_1^a + \dots + \mathcal{X}_k^a}{k} - E\mathcal{X}^a\right| \le \frac{\Delta}{3}, \quad \forall k \in [\frac{e^{n^{0.2}}}{a}, e^{n^{0.2}}]\right\} \text{ and}$$
$$E_{7,c} = \left\{\left|\frac{\mathcal{Z}_1^a + \dots + \mathcal{Z}_k^a}{k} - E\mathcal{Z}^a\right| \le \frac{\Delta}{3}, \quad \forall k \in [\frac{e^{n^{0.2}}}{a}, e^{n^{0.2}}]\right\}.$$

First step
First we show that:

$$E_{7,a} \cap E_{7,b} \cap E_{7,c} \subseteq E_{7S}^n(z,t). \tag{6.4.21}$$

Let $\bar{\imath}$ be (random) number defined by the inequalities:

$$\mathcal{Z}_1 + \dots + \mathcal{Z}_{\bar{\imath}} \le e^{n^{0.2}} < \mathcal{Z}_1 + \dots + \mathcal{Z}_{\bar{\imath}+1}. \tag{6.4.22}$$

Since $\mathcal{Z}_k \ge 1$, we have $\bar{\imath} \le e^{n^{0.1}}$. Let $\bar{k} := \mathcal{Z}_1 + \dots + \mathcal{Z}_{\bar{\imath}}$. Now,

$$\hat{\delta}_{z,t}^M = \frac{\sum_{i=1}^{e^{n^{0.2}}} X_i}{e^{n^{0.2}}} = \frac{\sum_{k=1}^{\bar{\imath}} \mathcal{X}_k + \sum_{i=\bar{k}+1}^{e^{n^{0.2}}} X_i}{\bar{k} + e^{n^{0.2}} - \bar{k}} = \frac{\frac{1}{\bar{\imath}}\sum_{k=1}^{\bar{\imath}} \mathcal{X}_k + \frac{1}{\bar{\imath}}\sum_{i=\bar{k}+1}^{e^{n^{0.2}}} X_i}{\frac{\bar{k}}{\bar{\imath}} + \frac{e^{n^{0.2}} - \bar{k}}{\bar{\imath}}}.$$

Denote

$$\Delta_I := E(\mathcal{X}^a - \mathcal{X}) + \frac{1}{\bar{\imath}}\sum_{i=1}^{\bar{\imath}}(\mathcal{X}_i - E\mathcal{X}^a) + \frac{1}{\bar{\imath}}\sum_{i=\bar{k}+1}^{e^{n^{0.2}}} X_i,$$

$$\Delta_{II} := E(\mathcal{Z}^a - \mathcal{Z}) + \frac{1}{\bar{\imath}}\sum_{i=1}^{\bar{\imath}}(\mathcal{Z}_i - E\mathcal{Z}^a) + \frac{1}{\bar{\imath}}\sum_{i=\bar{k}+1}^{e^{n^{0.2}}} Z_i.$$

Thus,

$$\hat{\delta}_{z,t}^M = \frac{E\mathcal{X} + \Delta_I}{E\mathcal{Z} + \Delta_{II}}.$$

Suppose now, that E_{7a} holds. Then, for each $i = 1, \dots, e^{n^{0.2}}$, we have $\mathcal{Z}_i = \mathcal{Z}_i^a$, $\mathcal{X}_i = \mathcal{X}_i^a$. From (6.4.22) then follows that $e^{n^{0.2}} \le \bar{\imath}a$, i.e.

$$e^{n^{0.2}} \ge \bar{\imath} \ge \frac{e^{n^{0.2}}}{a}. \tag{6.4.23}$$

When $\bar{\imath} = e^{n^{0.2}}$, then $e^{n^{0.2}} - \bar{k} = 0$, otherwise $e^{n^{0.2}} - \bar{k} \le \mathcal{Z}_{i+1} \le a$. Since $\sum_{i=\bar{\imath}+1}^{e^{n^{0.2}}} X_i \le e^{n^{0.2}} - \bar{k}$, we get

$$\frac{1}{\bar{\imath}} \sum_{i=\bar{k}+1}^{e^{n^{0.2}}} X_i \le \frac{e^{n^{0.2}} - \bar{k}}{\bar{\imath}} \le \frac{a}{\bar{\imath}} \le a^2 e^{-n^{0.2}} = \exp(6n^{0.1} - n^{0.2}) < \frac{\Delta}{6}, \tag{6.4.24}$$

provided n is big enough.
Hence, by (6.4.23) we have (recall that we assumed $E_{7,a}$)

$$\Big\{\Big|\frac{1}{\bar{\imath}}\sum_{k=1}^{\bar{\imath}}(\mathcal{X}_k - E\mathcal{X}^a)\Big| \le \frac{\Delta}{3}\Big\} = \Big\{\Big|\frac{1}{\bar{\imath}}\sum_{k=1}^{\bar{\imath}}(\mathcal{X}_k^a - E\mathcal{X}^a)\Big| \le \frac{\Delta}{3}\Big\} = \bigcup_{l=\frac{e^{n^{0.2}}}{a}}^{e^{n^{0.2}}} \Big\{\Big|\frac{1}{l}\sum_{k=1}^{l}(\mathcal{X}_k^a - E\mathcal{X}^a)\Big| \le \frac{\Delta}{3}, \bar{\imath} = l\Big\}$$

$$\supset \Big\{\Big|\frac{1}{l}\sum_{k=1}^{l}(\mathcal{X}_k^a - E\mathcal{X}^a)\Big| \le \frac{\Delta}{3}, l = \frac{e^{n^{0.2}}}{a}, \dots, e^{n^{0.2}}\Big\} = E_{7,b}.$$

Similarly,

$$\Big\{\Big|\frac{1}{\bar{\imath}}\sum_{k=1}^{\bar{\imath}}(\mathcal{X}_k - E\mathcal{X}^a)\Big| \le \frac{\Delta}{3}\Big\} \supset E_{7,c}.$$

Thus, by (6.4.24) on $E_{7a} \cap E_{7b} \cap E_{7c}$ we have

$$|\Delta_I| \le |E\mathcal{X}^a - E\mathcal{X}| + 2\frac{\Delta}{3} = E(\mathcal{X} - \mathcal{X}^a) + 2\frac{\Delta}{3}$$
$$|\Delta_{II}| \le |E\mathcal{Z}^a - E\mathcal{Z}| - 2\frac{\Delta}{3} = E(\mathcal{Z} - \mathcal{Z}^a) + 2\frac{\Delta}{3}.$$

Fix $k = 1, 2, \dots$. Denote by $n_0, n_1, n_2, \dots$ integers that satisfy $n_0 = 0$, $e^{2n^{0.1}} + 1 \ge n_i - n_{i-1} \ge e^{2n^{0.1}}$, $\forall i$. Let Y_j, $j = 0, 1, \dots$ denote a Bernoulli random variable which is equal to 1 if and only if between time $\nu(\vartheta(k)+1+n_j)$ and $\nu(\vartheta(k)+1+n_{j+1})$ the random walk does not visit the point $z^* := z - 2Le^{n^{0.1}}$. The random variables Y_j are independent. By definition, $\nu(i+1) - \nu(i) \ge e^{n^{0.1}}$. Hence, $\nu(\vartheta(k)+1+n_{j+1}) - \nu(\vartheta(k)+1+n_j) \ge e^{3n^{0.1}}$. At time $\nu(\vartheta(k)+1)$, the random walk is located on I_z and, therefore, no more than $3e^{n^{0.1}}$ from z^*. By (6.4.10), the probability to visit the point z^* within time $e^{3n^{0.1}}$ starting from the $3e^{n^{0.1}}$-neighborhood of z^* goes to 1 as $n \to \infty$. Hence, $\sup_j P(Y_j = 1) \to 0$. Let n be so big, that $P(Y_j = 1) \le e^{-1}$, for all j. This means, for each

$$P(\mathcal{Z}_k \ge te^{2n^{0.1}}) \le P(Y_j = 1, j = 0, \dots, \ulcorner t \urcorner - 1) \le \exp(-\ulcorner t \urcorner) \le \exp(-t), \quad k = 1, 2, \dots \tag{6.4.25}$$

Now,

$$E(\mathcal{Z} - \mathcal{Z}^a) = \int_{\{\mathcal{Z} \ge a\}} \mathcal{Z} dP - aP(\mathcal{Z} \ge a) = aP(\mathcal{Z} \ge a) + \int_{(a,\infty)} P(\mathcal{Z} > x)dx - aP(\mathcal{Z} \ge a) = \int_{(a,\infty)} P(\mathcal{Z} > x)dx.$$

By (6.4.25):

$$\int_{(a,\infty)} P(\mathcal{Z} > x)dx \le \int_a^\infty \exp(-xe^{-2n^{0.1}})dx \le e^{2n^{0.1}} \exp(-ae^{-2n^{0.1}})) \le e^{2n^{0.1}} \exp(-e^{n^{0.1}}).$$

Thus, for n big enough:

$$E(\mathcal{Z} - \mathcal{Z}^a) \le e^{2n^{0.1}} \exp(-e^{n^{0.1}}) \le \frac{\Delta}{3}.$$

Since, $\mathcal{X} \le \mathcal{Z}$, we get:

$$E(\mathcal{X} - \mathcal{X}^a) = \int_{(a,\infty)} P(\mathcal{X} > x)dx \le \int_{(a,\infty)} P(\mathcal{Z} > x)dx \le \frac{\Delta}{3}.$$

Thus, on $E_{7a} \cap E_{7b} \cap E_{7c}$ we have:

$$|\Delta_I|, |\Delta_{II}| \le \Delta. \tag{6.4.26}$$

Recall that we have

$$\hat{\delta}^M_{z,t} = \frac{E\mathcal{X} + \Delta_I}{E\mathcal{Z} + \Delta_{II}}.$$

Hence, by (6.4.26):

$$\frac{E\mathcal{X} - \Delta}{E\mathcal{Z} + \Delta} \le \hat{\delta}^M_{z,t} \le \frac{E\mathcal{X} + \Delta}{E\mathcal{Z} - \Delta}.$$

By Taylor's formula,

$$\frac{E\mathcal{X} - \Delta}{E\mathcal{Z} + \Delta} = \frac{E\mathcal{X}}{E\mathcal{Z}} - \Big(\frac{E\mathcal{X} + E\mathcal{Z}}{(E\mathcal{Z})^2}\Big)\Delta + o(\Delta).$$

Since $1 \le E\mathcal{X} \le E\mathcal{Z}$, the latter means (for Δ small enough)

$$\Big|\frac{E\mathcal{X} - \Delta}{E\mathcal{Z} + \Delta} - \frac{E\mathcal{X}}{E\mathcal{Z}}\Big| \le \Big(\frac{E\mathcal{X} + E\mathcal{Z}}{(E\mathcal{Z})^2}\Big)\Delta + o(\Delta) \le 2\Delta + o(\Delta) < 3\Delta.$$

Similarly

$$\Big|\frac{E\mathcal{X} + \Delta}{E\mathcal{Z} - \Delta} - \frac{E\mathcal{X}}{E\mathcal{Z}}\Big| < 3\Delta.$$

Now, $\delta^M_{z,t} = \frac{E\mathcal{X}}{E\mathcal{Z}}$ implying that

$$|\delta^M_z - \hat{\delta}^M_{z,t}| < 3\Delta < e^{-n^{0.12}}.$$

Thus, (6.4.21) holds.

Second step
We now show that $P(E^c_{7,a})$, $P(E^c_{7,b})$ and $P(E^c_{7,c})$ are of order $o(\exp(-n^{1000}))$.
Taking $t = e^{n^{0.1}}$ (6.4.25) yields:

$$P(\mathcal{Z}_k > a) \le \exp(-e^{n^{0.1}}), \quad k = 1, 2, \ldots.$$

Thus:

$$P(E^c_{7,a}) \le \exp(n^{0.2}) \exp(-e^{n^{0.1}}) = \exp(n^{0.2} - e^{n^{0.1}}) < \exp(-n^{1000}). \tag{6.4.27}$$

To estimate $P(E_{7,b})$ and $P(E_{7,c})$ we use the Hoeffding inequality. Fix $l \in [\frac{e^{n^{0.2}}}{a}, e^{n^{0.2}}]$. By (6.4.8) we have:

$$P\Big(\Big|\frac{1}{l}\sum_{k=1}^{l}(\mathcal{X}_k^a - E\mathcal{X}_k^a)\Big| \geq \frac{\Delta}{6}\Big) \leq \exp\Big(-2l\big(\frac{\Delta}{a6}\big)^2\Big).$$

On the other hand, since $\mathcal{X}_k^a$, $k \geq 2$ are i.i.d., we have:

$$\Big|\frac{1}{l}\sum_{k=1}^{l} E\mathcal{X}_k^a - E\mathcal{X}^a\Big| = \frac{1}{l}|E\mathcal{X}^a - E\mathcal{X}_1^a| \leq \frac{2a}{l} \leq 2a^2 e^{-n^{0.2}} = 2\exp(6n^{0.1} - n^{0.2}) < \frac{\Delta}{6}.$$

Thus,

$$P\Big(\Big|\frac{1}{l}\sum_{k=1}^{l}\mathcal{X}_k^a - E\mathcal{X}^a\Big| \geq \frac{\Delta}{3}\Big) \leq P\Big(\Big|\frac{1}{l}\sum_{k=1}^{l}(\mathcal{X}_k^a - E\mathcal{X}^a)\Big| \geq \frac{\Delta}{6}\Big) \leq \exp\Big(-2l\big(\frac{\Delta}{a6}\big)^2\Big) \leq \exp(-Ke^{n^{0.2}}\frac{\Delta^2}{a^3}),$$

where $K = \frac{2}{36}$. Now,

$$e^{n^{0.2}}\frac{\Delta^2}{a^3} = \exp(n^{0.2} - \frac{1}{2}n^{0.2} - 9n^{0.1}) = \exp(\frac{1}{2}n^{0.2} - 9n^{0.1}) > \exp(\frac{n^{0.2}}{4})$$

and

$$P\Big(\Big|\frac{1}{l}\sum_{k=1}^{l}\mathcal{X}_k^a - E\mathcal{X}^a\Big| \geq \frac{\Delta}{3}\Big) \leq \exp(-Ke^{\frac{n^{0.2}}{4}}).$$

Finally

$$P(E_{7,b}^c) \leq \sum_{l=\frac{e^{n^{0.2}}}{a}}^{e^{n^{0.2}}} P\Big(\Big|\frac{1}{l}\sum_{k=1}^{l}\mathcal{X}_k^a - E\mathcal{X}^a\Big| \geq \frac{\Delta}{3}\Big) < e^{n^{0.2}}\exp(-Ke^{\frac{n^{0.2}}{4}}) < \exp(-e^{n^{0.1}}) < \exp(-n^{1000}). \tag{6.4.28}$$

The same bound holds for $P(E_{7,c}^c)$.

Because of (6.4.21), (7.4.3) and (6.4.28) we get:

$$P(E_{7S}^{nc}(c,t)) \leq 3\exp(-n^{1000}). \tag{6.4.29}$$

The bound in (6.4.29) do not depend on the choice of z, t and ψ. Note that on $[-cm, cm] \times [0, m^2]$, there are no more than $(cm)^3$ values of (z,t). Hence

$$P\left(E_{7,S}^{nc}\right) \leq \Sigma_{z\in[-cm,cm],t\in[0,m^2]} P\left(E_{7,S}^{nc}(z,t)\right).$$

¿From (6.4.29) it follows

$$P\left(E_{7,S}^{nc}\right) \leq (cm)^3 3\exp(-n^{1000}). \tag{6.4.30}$$

Recall that by (6.3.10): $(cm)^3 \leq c^3 e^{6n}$. Hence, the right side of (6.4.30) is less than $3c^3\exp(6n - n^{1000})$. This is of order $o(\exp(-3n))$. By (6.4.6), we therefore have:

$$P_\psi(E_{7S}^{nc}) \to 0.$$

Outline of the proof that $P_\psi(E^n_{8,S}) \to 1$

Note that in the previous proof the exact nature of $X_{z,i}$, $\mathcal{X}_z(k)$ as well as $X_{z,t,i}$, $\mathcal{X}_{z,t}(k)$ was not used. Hence, the proof holds, if they were replaced by $U_{z,i}$, $\mathcal{U}_z(k)$, $\chi(\nu_{z,t}(i)+e^{n^{0.1}})$ and

$$\sum_{\kappa(k-1)+1}^{\kappa(k)} \chi(\nu_{z,t}(i) + e^{n^{0.1}}),$$

respectively. By (6.2.12) this proves that $P_\psi(E^n_{8,S}) \to 1$.

Proof that $P_\psi(E^n_{5,S}) \to 1$.

Fix $\psi_n \in E^n_{\text{cell\_OK}}$.

For each $k = 0, 1, 2, \ldots$, let $\tau_k(0) := \rho(k)$ and for each $j = 1, 2, \ldots$, let $\tau_k(j)$ be the smallest time $t > \tau_k(j-1) + 2e^{n^{0.1}}$ for which $S(t) \in I(\rho(k))$.

Let $X_k(j)$ be the Bernoulli random variable which is equal to one if and only if during time $[\tau_k(j), \tau_k(j) + (n^{3000} + n^2)]$ we observe $n^2 + 1$ consecutive 0's or 1's. That is $X_k(j) = 1$ if and only if $\exists t \in [\tau_k(j), \tau_k(j) + n^{3000}]$ such that $\chi(t) = \chi(t+1) = \cdots = \chi(t+n^2)$.
Clearly, for each k, the random variables $X_k(j)$, $j = 0, 1, 2, \ldots$ are independent

At first we show that there exists a constant $a > 0$, not depending on n, such that for each k and j,

$$P(X_k(j) = 1) \geq n^{-a \ln n} = e^{-a \ln^2 n}. \tag{6.4.31}$$

Fix $k = 0, 1, \ldots$ and let $I := I(\rho(k))$. Let $\bar{z}$ be the signal carrier point such that $I_{\bar{z}} = I$. Since $\bar{z}$ is a signal carrier point, then, by Corollary 6.2.2 and c) of Proposition 12.4.9, I contains at least one big block of ψ_n. Let $T = [a, b] \subseteq I$ be that block. Now, let $a < a^* < b^* < b$ be such that $a^* - a, b^* - a^*, b - b^* \geq \frac{|T|}{3} \geq \frac{\ln n}{3n}$. Let $T^* = [a^*, b^*]$. Now,

$$P(X_k(j) = 1) \geq P(S(\tau_k(j) + n^{3000}) \in T^*) P(\chi(t) = \chi(t+1) = \cdots = \chi(t+n^2) | S(t) \in T^*).$$

Now, by LCLT:

$$P(S(\tau_k(j) + n^{3000}) \in T^*) \geq \frac{1}{cn^{1500}} - O(\frac{1}{n^{3000}}) \geq n^{-1501},$$

provided that n is big enough.
Let $N = (\frac{n}{\ln n})^2$ (w.l.o.g we assume that this is an integer) and estimate:

$$P(\chi(t) = \chi(t+1) = \cdots = \chi(t+n^2) | S(t) = j \in T^*) \geq P(S_j(i) \in T, \quad \forall i = 1, 2, \ldots, n^2) \geq$$
$$P\Big(\max_{i=1,\ldots,N} |S_j(i)| \leq \frac{|T|}{3}, S_j(N) \in T^*\Big)^{\ln^2 n} = P\Big(\max_{i=1,\ldots,N} \frac{|S_j(i)|}{\sqrt{N}} \leq \frac{1}{3}, \frac{S_j(N)}{\sqrt{N}} \in \frac{T^*}{\sqrt{N}}\Big)^{\ln^2 n}. \tag{6.4.32}$$

Note: $|T^*| \geq \sqrt{N}$. By (6.4.10):

$$P\Big(\max_{i=1,\ldots,N} \frac{|S_j(i)|}{\sqrt{N}} \leq \frac{1}{3}, \frac{S_j(N)}{\sqrt{N}} \in \frac{T^*}{\sqrt{N}}\Big) \to P(\sup_{0 \geq t \leq 1} |W_t| \leq \frac{1}{3\sigma}, W_1 \in I) > \gamma > 0.$$

Thus, for n big enough there exists $a < \infty$ such that the right side of (6.4.32) is bigger than $(\frac{1}{a})^{\ln^2 n} = n^{-c \ln n}$, with $c > 0$. Hence, (6.4.31) holds with $a = c + 1$.

Define the following events:

$$E_a(k) = \left\{ \begin{array}{c} \text{if } \rho(k) \leq m^2 \\ \text{then during the time } [\rho(k), \rho(k) + e^{n^{0.3}} - e^{n^{0.1}}] \\ S \text{ visits } I(\rho(k)) \text{ more than } e^{n^{0.22}} \text{ times} \end{array} \right\} \quad k = 0, 1, \ldots$$

and

$$E_a := \cap_{k=1}^{25000} E_a(k).$$

Also define

$$E_b(k) := \Big\{ \sum_{j=0}^{e^{n^{0.21}}} X_k(j) \geq e^{n^{0.2}} \Big\}, \quad E_b := \cap_{k=0}^{n^{25000}} E_b(k).$$

Now, clearly, on $E_a(k)$ we have $\tau_k(e^{n^{0.21}}) \leq \rho(k) + e^{n^{0.3}} - 2e^{n^{0.1}}$. Thus $E^n_{5,S}$ holds, if

$$\sum_{j=0}^{e^{n^{0.21}}} X_k(j) \geq e^{n^{0.2}}.$$

Hence

$$E_{5,S} \supset E_a \cap E_b \quad \text{and} \quad P_\psi(E^c_{5,S}) \leq P_\psi(E^c_a) + P_\psi(E^c_b).$$

We are now proving that $P_\psi(E^c_a) \to 0$ and $P_\psi(E^c_b) \to 0$.

Proof that $P_\psi(E^c_b) \to 0$
By (6.4.6) it is enough to show that:

$$P(E^c_b | \psi_n) = o(e^{-3n}). \tag{6.4.33}$$

Note that for big n, $\exp(n^{0.2} - n^{0.21}) < EX_k(j)$, $\forall j$. Thus,

$$\exp(n^{0.2} - n^{0.21}) < \frac{1}{e^{n^{0.21}}} \exp(-n^{0.21}) \sum_{j=0}^{e^{n^{0.21}}} E(X_k(j)) =: \bar{m}.$$

By the Hoeffding inequality we obtain that for a constant $K > 0$:

$$P(E^c_b(k) | \psi_n) = P\Big(\frac{1}{e^{n^{0.21}}} \sum_{j=0}^{e^{n^{0.21}}} X_k(j) < \exp(n^{0.2} - n^{0.21})\Big) \leq P\Big(\frac{1}{e^{n^{0.21}}} \sum_{j=0}^{e^{n^{0.21}}} X_k(j) < \frac{\bar{m}}{2}\Big) =$$

$$P\Big(\frac{1}{e^{n^{0.21}}} \sum_{j=0}^{e^{n^{0.21}}} \big(X_k(j) - EX_k(j)\big) < -\frac{\bar{m}}{2}\Big) \leq \exp(-K\bar{m}^2 e^{n^{0.21}}) \leq \exp(-K e^{n^{0.21} - 2a \ln^2 n}).$$

Hence,

$$P(E^c_b | \psi_n) \leq n^{25000} \exp(-K e^{n^{0.21} - 2a \ln^2 n}) = o(e^{-3n}).$$

Proof that $P_\psi(E^c_a) \to 0$
This proof is a little tricky because unlike the other proofs we have that $P(E_a | \psi_n)$ is much bigger than $P(S(m^2) = m)$.

Let $L = n^{100000}$ and consider the event:

$$C = \Big\{ S\left(m^2(1-n^{-3L})\right) \in \left[m(1-n^{-L}), m(1+n^{-L})\right] = [m - \frac{m}{n^L}, m + \frac{m}{n^L}] \Big\}.$$

Here and in the rest of the proof we assume (w.l.o.g.) that all ratios and exponents are integers. Also define

$$E_c(k) = \left\{\rho(k) \notin [m^2(1-n^{-3L}), m^2]\right\}, \quad k = 0, 1, \ldots, \quad E_c := \cup_{k=1}^{25000} E_c(k).$$

The event E_c means that no stopping time $\rho(k)$ occurs in the time-interval $[m^2(1 - n^{-3L}), m^2]$, the event $E_a \cap E_c$ satisfies

$$E_a \cap E_c = E_a^* := \cap_{k=1}^{25000} E_a^*(k),$$

where

$$E_a^*(k) = \left\{ \begin{array}{c} \text{if } \rho(k) \le m^2(1-n^{-L}) \\ \text{then during the time } [\rho(k), \rho(k) + e^{n^{0.3}} - e^{n^{0.1}}] \\ S \text{ visits } I(\rho(k)) \text{ more than } e^{n^{0.22}} \text{ times} \end{array} \right\}.$$

We show that the probability $P(E_a | E_{1,S}^n, \psi_n)$, can be very well approximated by the probability $P(E_a^* | C, \psi_n)$ and the latter goes to 0 when $n \to \infty$. We proceed in three steps.

1) At first note: since

$$C^c \cap E_{1,S}^n = \{S(m^2(1-n^{-3L})) \notin [m(1-n^L), m(1+n^L)], S(m^2) = m\},$$

we get, by the Hoeffdig inequality

$$\begin{aligned} P(C^c \cap E_{1,S}^n | \psi_n) =& P(C^c \cap E_{1,S}^n) = P(E_{1,S}^n | C^c) P(C^c) \le P(E_{1,S}^n | C^c) \\ =& P\Big(\Big| S\Big(\frac{m^2}{n^{3L}}\Big) \Big| \ge \frac{m}{n^L} \Big) \le \exp(-dn^L) = o(n^{-3n}). \end{aligned}$$

The latter implies

$$P_\psi(C^c) = o(1). \tag{6.4.34}$$

2) Second, use the inequalities:

$$P(E_a^{*c} \cap E_{1,S}^n \cap C | \psi_n) \le P(E_a^c \cap E_{1,S}^n \cap C | \psi_n) \le P(E_a^{*c} \cap E_{1,S}^n \cap C | \psi_n) + P(E_c^c \cap E_{1,S}^n | \psi_n).$$

Since $\psi \in E_7^n$, it has no signal carrier points in $[m - EZn^{-11001}]$. Hence, $E_c^c \cap E_{1,S}^n$ can hold only, if during time interval $[m^2(1-n^{-3L}), m^2]$ the random walk covers a distance of at least $EZn^{-11001} - Ln^{1000}$. Thus,

$$P(E_c^c \cap E_{1,S}^n | \psi_n) \le P\Big(\max_{l=1,\ldots,\frac{m^2}{n^{3L}}} |S(l)| \ge EZn^{-11001} - Ln^{1000} \Big) \le P\Big(\max_{l=1,\ldots,\frac{m^2}{n^{3L}}} |S(l)| \ge \frac{m}{n^{11003}} - Ln^{1000} \Big).$$

Now use the maximal inequality (6.4.17) together with the Hoeffding inequality to estimate

$$\begin{aligned} P\Big(\max_{l=1,\ldots,\frac{m^2}{n^{3L}}} |S(l)| \ge \frac{m}{n^{11003}} - Ln^{1000} \Big) \le& \max_{l=1,\ldots,\frac{m^2}{n^{3L}}} 3P\Big(|S(l)| \ge \frac{1}{3} \frac{m}{n^{12000}} \Big) \\ \le& 3\exp(-dn^{3L-12000}) = o(e^{-3n}). \end{aligned}$$

This implies:

$$\frac{P(E_a^c \cap C \cap E_{1,S}^n|\psi_n) - P(E_a^{*c} \cap C \cap E_{1,S}^n, |\psi_n)}{P(E_{1,S}^n|\psi_n)} = P_\psi(E_a^c \cap C) - P_\psi(E_a^{*c} \cap C) = o(1). \tag{6.4.35}$$

3) Finally, note that:

$$P(E_a^{*c} \cap E_{1,S}^n \cap C|\psi_n) = P(E_a^{*c} \cap C|\psi_n)P(E_{1,S}^n|E_a^{*c} \cap C, \psi_n) = P(E_a^{*c} \cap C|\psi_n)P(E_{1,S}^n|C, \psi_n).$$

On the other hand,

$$P(E_{1,S}^n|\psi_n) \geq P(E_{1,S}^n \cap C|\psi_n) = P(E_{1,S}^n|C, \psi_n)P(C|\psi_n).$$

Hence,

$$P_\psi(E_a^{*c} \cap C) = \frac{P(E_a^{*c} \cap E_{1,S}^n \cap C|\psi_n)}{P(E_{1,S}^n|\psi_n)} \leq \frac{P(E_a^{*c} \cap C|\psi_n)P(E_{1,S}^n|C, \psi_n)}{P(E_{1,S}^n|C, \psi)P(C|\psi_n)} = P(E_a^{*c}|C, \psi_n). \tag{6.4.36}$$

By CLT, $P(C|\psi_n) = P(S\left(m^2(1-n^{-3L}\right) \in [m - \frac{m}{n^L}, m + \frac{m}{n^L}])$ is of order $\frac{1}{n^K}$ for some big $K > 0$. We estimate the probability $P(E_a^{*c}|\psi_n)$.
To do this, fix k and let $T_1, T_2, \ldots$ denote the waiting times of S between visits of the point $S(\rho(k))$ (when we start at the time $\rho(k)$). Although $ET_i = \infty$, it is known that $ET_i^{\frac{1}{3}} =: K' < \infty$ (see, e.g. [LMM01]). The number K', does not depend on n. Thus, by the Markov inequality we have

$$\begin{aligned} P(E_a^{*c}) &\leq P\Big(\sum_{i=1}^{e^{n^{0.22}}} T_i > e^{n^{0.3}} - e^{n^{0.1}}\Big) = P\left(\Big(\sum_{i=1}^{e^{n^{0.22}}} T_i\Big)^{\frac{1}{3}} > \left(e^{n^{0.3}} - e^{n^{0.1}}\right)^{\frac{1}{3}}\right) \\ &\leq P\Big(\sum_{i=1}^{e^{n^{0.22}}} T_i^{\frac{1}{3}} > \left(e^{n^{0.3}} - e^{n^{0.1}}\right)^{\frac{1}{3}}\Big) \leq \frac{e^{n^{0.22}}K'}{\left(e^{n^{0.3}} - e^{n^{0.1}}\right)^{\frac{1}{3}}} \leq e^{-n^{0.25}}. \end{aligned}$$

Thus, $P(E_{a*}^c) \leq n^{25000}e^{-n^{0.25}} = o(n^{-K})$ implying that

$$P(E_{a*}^c|C, \psi) \leq \frac{P(E_{a*}^c|\psi_n)}{P(C|\psi_n)} = o(1). \tag{6.4.37}$$

To complete the proof, use (6.4.34), (6.4.35), (6.4.37), (6.4.37) to get

$$\begin{aligned} P_\psi(E_a^c) &\leq P_\psi(E_a^c \cap C) + P_\psi(C^c) = P_\psi(E_a^{*c} \cap C) + P_\psi(E_a^c \cap C) - P(E_a^{*c} \cap C) + o(1) \\ &\leq P(E_a^{*c}|C, \psi_n) + o(1) = o(1). \end{aligned}$$

6.5 Combinatorics of g and $\hat{\mathbf{g}}$

In this section we show: if all scenery dependent events and random walk dependent events hold, then our estimates $\hat{\delta}_T^M$ and $\hat{h}_t$ are precise. This means, we can observe our signals and, just like in our 3-color example, we can estimate the g-function.
Let us first give the definition of the g-function in the 2-colors case.

6.5.1 Definition of g

In this subsection we give a formal definition of the function

$$g : \{0,1\}^{m+1} \mapsto \{0,1\}^{n^2+1}.$$

The function g depends on n. When n is fixed, we choose $m = \ulcorner n^{2.5}EZ \urcorner$, where the random variable Z is the location of the first Markov signal point after $2Ln^{1000}$ in ξ. We consider the signal carrier points $\bar{z}_1, \bar{z}_2, \ldots$, in $[0, m]$. Define the following subset of $\{0,1\}^{m+1}$:

$$E^* := \{\psi \in \{0,1\}^{m+1} : \bar{z}_1(\psi) \geq L(e^{n^{0.1}} + n^{1000}), \bar{z}_{n^2+1} \leq m - L(e^{n^{0.1}} + n^{1000})\}.$$

Here, $\bar{z}_i(\psi) = \infty$, if the piece of scenery ψ has less than i signal carrier points.
Clearly $E^n_{\text{cell\_OK}} \subseteq E^*$. If $\psi \in E^*$, then for each $\bar{z}_i(\psi)$ we define the vector of the frequency of ones $h(i)$, $i = 1, \ldots, n^2 + 1$. Recall from (6.2.13) that:

$$h(i) = h(\bar{z}_i(\psi)) = P(\psi(U + S(e^{n^{0.1}})) = 1),$$

where U is a random variable with distribution $\mu(\bar{z}_i)$.
Now, if $\psi \in E^*$, let:

$$g_i(\psi) = \begin{cases} 1 & \text{, if } h(i) > 0.5 \\ 0 & \text{, if } h(i) < 0.5 \\ \bar{z}_i(\psi) & \text{otherwise.} \end{cases} \tag{6.5.1}$$

When $\psi \notin E^*$, define

$$g_i(\psi) = \psi(i), \quad i = 2, 3, \ldots, n^2 + 2. \tag{6.5.2}$$

Definition 6.5.1. *$g(\psi) = (g_1(\psi), \ldots, g_{n^2+1}(\psi))$, where $g_i(\psi)$ is (6.5.1), if $\psi \in E^*$ and $g_i(\psi)$ is (6.5.2), if $\psi \notin E^*$.*

Definition 6.5.1 ensures that $g(\psi)$ depends only on ξ_0^m, and that $(g_1(\xi), \ldots, g_{n^2+1}(\xi))$ is an i.i.d. random vector, with the components being Bernoulli random variables with parameter $\frac{1}{2}$.

6.5.2 Definition of $\hat{g}$

Next, we formalize the construction of the $\hat{g}$-function. The function $\hat{g} : \{0,1\}^{m^2+1} \mapsto \{0,1\}^{n^2}$ aims to estimate the (non-observable) function g. The argument of $\hat{g}$ is the vector of observations $\chi_0^{m^2} := (\chi(0), \ldots, \chi(m^2))$, and the estimate is given up to the first or last bit. In other words, $\hat{g}$ aims to achieve $\hat{g}(\chi^{m^2}) \sqsubseteq g(\xi|[0, m])$.
The algorithm for computing $\hat{g}$ has 5 phases and it differs from the $\hat{g}$-reconstruction algorithm for the 3-color case (Subsection 6.1.6) by the first step, only. The rest of the construction is the same.

1. For all $T = [t, t+e^{n0.3}] \subseteq [0, m^2]$ compute the estimate of the Markov signal probability $\hat{\delta}_T^M$. Select all intervals $T_1 = [t_1, t_1+e^{n^{0.3}}], T_2 = [t_2, t_2+e^{n^{0.3}}], \ldots, T_K = [t_K, t_K+e^{n^{0.3}}]$, $t_1 < t_2 < \cdots < t_K$, where the estimated Markov signal probability are higher than c_r. Here K stands for the number of such intervals.

2. For all selected intervals, estimate the frequency of ones. Obtain the estimates $\hat{h}_{T_1}, \dots, \hat{h}_{T_K}$, $i = 1, \dots, K$.

3. Define clusters:

$$C_i := \{\hat{h}_{T_j} : |\hat{h}_{T_j} - \hat{h}_{T_i}| \le 2\exp(-n^{0.12})\},\ \hat{f}_i := \frac{1}{|C_i|}\sum_{j \in C_i} \hat{h}_{T_j}, \quad i = 1, \dots, K.$$

4. Apply the real scenery construction algorithm $\mathcal{A}_n^{\mathbb{R}}$ (see Subsection 6.1.6) to the vector $(\hat{f}_1, \dots, \hat{f}_K)$. Denote the output, $\mathcal{A}_n^{\mathbb{R}}(\hat{f}_1, \dots, \hat{f}_K)$, by

$$(f_1, \dots, f_{n^2}). \tag{6.5.3}$$

If the number of different reals in $(\hat{f}_1, \dots, \hat{f}_K)$ is less than n^2 (e.g. $K \le n^2$), then complete the vector (6.5.3) arbitrarily.

5. Define the final output of $\hat{g}$ as follows

$$\hat{g}(\chi^{m^2}) := (I_{[0.5,1]}(f_1), \dots, I_{[0.5,1]}(f_{n^2})).$$

6.5.3 Main proof

Next, we prove the main result: when all previously stated events hold, then the $\hat{g}$-algorithm *works*, i.e. $\hat{g}(\chi_0^{m^2}) \sqsubseteq g(\xi_0^m)$.
Recall $E^n_{\text{cell\_OK}} = \cap_{i=1}^9 E_i^n$. Similarly define the intersection of the random walk dependent events: $E_S^n := \cap_{i=1}^8 E_{i,S}^n$. Finally, let $E_{\text{g-works}}$ be the event that $\hat{g}$ works, i.e.:

$$E_{\text{g-works}} := \left\{\hat{g}(\chi_0^{m^2}) \sqsubseteq g(\xi_0^m)\right\}. \tag{6.5.4}$$

At first we show that step 1 in the definition of $\hat{g}$ works properly, i.e. a time interval T is selected (i.e. $\hat{\delta}_T^M > c_r$) only if during the time T the random walk is close to a unique signal carrier point $\bar{z}$. The closeness is defined in the following sense: we say that during time period T, the random walk S is close to z, if there exists $s \in T$ such that $S(s) \in I_z$.

Proposition 6.5.1. *Suppose $E^n_{\text{cell\_OK}} \cap E_S^n$ holds. Let $T = [t, t + e^{n^{0.3}}] \subseteq [0, m]$. If during T, the random walk is close to a signal point z, and $\hat{\nu}_t(e^{n^{0.2}}) \le t + e^{n^{0.3}} - e^{n^{0.1}}$, then $\hat{\delta}_T^M = \hat{\delta}_{z,t}^M$ and $\hat{h}_T = \hat{h}_{z,t}$.*

Proof. Since ξ and S are independent, we fix $\xi = \psi \in E^n_{\text{cell\_OK}}$ and show that the claim of the proposition holds.
Let S be close to the signal point z. By $E_2^n \cap E_8^n \cap E_9^n$, the point z has empty neighborhood and empty borders. Hence, in the area

$$([z - L(n^{1000} + e^{n^{0.3}}), z + L(n^{1000} + e^{n^{0.3}})] - [z - L\tilde{M}, z + L\tilde{M}]) \cap [-cm, cm]$$

there are no blocks that are bigger than $n^{0.35}$. Recall that $\tilde{M} = n^{1000} - 2n^2$. Since $2n^{0.35} < n^{0.4} < n^2$, this means: all blocks with length at least $n^{0.4}$ must lay inside the interval $[z - L(n^{1000} - n^2), z + L(n^{1000} - n^2)]$. In particular, this implies - if, during the

time T the random walk S visits a block bigger than $n^{0.4}$, then during the n^2 step before and after that visit, it must stay in the interval I_z. Formally: if $\exists s \in T : S(s) \in B$, then

$$S(s-n^2), S(s-n^2+1), \ldots, S(s+n^2-1), S(s+n^2) \in I_z. \tag{6.5.5}$$

Here B stands for a block of ψ with length at least $n^{0.4}$.
We now take advantage of the event $E^n_{6,S}$: the random walk cannot generate n^2+1 times the same color, if it does not visit a block bigger than $n^{0.4}$. By (6.5.5) this means that all n^2+1 same colors must be generated on I_z. Hence, inside the time interval T, the stopping times $\hat{\nu}_t(i)$ are equal to the stopping times $\nu_{z,t}(i)$. Similarly, $X_{t,i} = X_{z,t,i}$, provided $\hat{\nu}_t(i) + n^{1000} \leq t + e^{n^{0.3}}$.
By assumption, there are at least $e^{n^{0.2}}$ stopping times $\hat{\nu}_t(i)$ in $[t, t+e^{n^{0.3}} - e^{n^{0.1}}]$ These stopping times are then equal to $\nu_{z,t}(i)$. Similarly, $X_{t,i} = X_{z,t,i}$, $i = 1, \ldots, e^{n^{0.2}}$. The latter means that the observable estimates $\hat{\delta}^M_T$ and $\hat{h}_T$ equals the non-observable estimates $\hat{\delta}^M_{z,t}$ and $\hat{h}_{z,t}$, respectively. □

Corollary 6.5.1. *Suppose $E^n_{\text{cell\_OK}} \cap E^n_S$ holds. Let $T = [t, t+e^{n^{0.3}}] \subseteq [0,m]$. If during T the random walk is close to a signal point z, then $\hat{\delta}^M_T > 0$ implies that $\hat{h}_T = \hat{h}_{z,t}$ and $\hat{\delta}^M_T = \hat{\delta}^M_{z,t}$.*

Proof. By definition, $\hat{\delta}^M_T > 0$ if in the time interval $[t, t+e^{n^{0.3}} - e^{n^{0.1}}]$ there are at least $e^{n^{0.2}}$ stopping times $\hat{\nu}_t(i)$. Now Proposition 6.5.1 applies. □

Lemma 6.5.1. *Suppose $E^n_{\text{cell\_OK}} \cap E^n_S$ holds. Let $T = [t, t+e^{n^{0.3}}] \subseteq [0,m]$ be such that $\hat{\delta}^M_T > c_r$. Then there exists an unique signal carrier point $\bar{z} \in [-cm, cm]$ such that S is close to $\bar{z}$ during T and $\hat{\delta}^M_T = \hat{\delta}^M_{\bar{z},t}$.*

Proof. Fix $\xi = \psi \in E^n_{\text{cell\_OK}}$. Note that, since E^n_2 holds, all signal points in $[-cm, cm]$ have empty neighborhood. Together with d) of Proposition 12.4.9 this means that all signal points in $[-cm, cm]$ are in clusters with diameter less than $2Ln^{1000}$. The distance between any two clusters, i.e. the distance between closest signal points in these clusters, is bigger than $e^{n^{0.3}}$. Moreover, by $E^n_8 \cap E^n_9$, all signal points have empty borders.
If $E^n_{2,S}$ holds, then during time $[0, m^2]$, our random walk stays in $[-cm, cm]$. Together with the clustering structure of the signal points, this means: if during the time interval $T \subseteq [0, m^2]$ of length $e^{n^{0.3}}$ the random walk S is close to some signal points, then they all belong to the same cluster. Hence, during T, S can be close to at most one signal carrier point (recall, every cluster has one representant, the signal carrier point). We have to show that if $\hat{\delta}^M_T > c_r$, then there exists at least one signal carrier point $\bar{z}$ such that, (during T) S is close to $\bar{z}$.
During T, the random walk S has 3 options :

- S is not close to any signal point
- S is close to the signal points that are not Markov signal points
- S is close to a Markov signal point.

If S is not close to any signal point, then by $E^n_{3,S}$, $\hat{\delta}^M_T \leq c_r$. This excludes the first possibility. Hence, $\hat{\delta}^M_T > c_r$ cannot happen, if during T, S is not close to any signal point.

Suppose now that there exists a signal point z such that (during T) S is close to z. By assumption we have $\hat{\delta}_T^M > c_r > 0$. By Corollary 6.5.1 we have that $\hat{\delta}_T^M = \hat{\delta}_{z,t}^M$. Now we reap benefit from the events E_5^n and $E_{7,S}^n$. The event E_5^n ensures that z is regular, i.e. $|\delta_z^M - c_r| \geq \Delta > e^{-n^{0.12}}$ (recall, Δ is polynomially small). On the other hand, the event $E_{7,S}^n$ ensures $|\hat{\delta}_T^M - \delta_z^M| = |\hat{\delta}_{z,t}^M - \delta_z^M| \leq \exp(-n^{0.12})$. Thus on $E_5^n \cap E_{7,S}^n$ we have:

$$\hat{\delta}_T^M > c_r \quad \text{if and only if} \quad \delta_z^M > c_r - \Delta. \tag{6.5.6}$$

Suppose that we have the second possibility – S is close to some signal points, but not close to any Markov signal points. Then z is not a Markov signal point. Hence, (7.2.7) ensures that $\hat{\delta}_T^M \leq c_r$. This contradicts our assumption that $\hat{\delta}_T^M > c_r$. Hence, z must be a Markov signal point and our third option holds.
Thus $\hat{\delta}_T^M > c_r$ implies that during T, the random walk S is close to a Markov signal point. By clustering structure we know that S is close to a cluster of signal points with at least one Markov signal points. In Subsection 6.3.4 we argued that such a cluster serves as the signal carrier. However, to complete the proof we must show that, during T, S is also close to the corresponding signal carrier point, say $\bar{z}$.
The points $\bar{z}$ and z belong to the same cluster, i.e. $|\bar{z} - z| < 2Ln^{1000}$. Consider the interval

$$J_z := [z - L(\exp(n^{0.3}), z + L(\exp(n^{0.3})] \cap [-cm, cm].$$

This is the region, where the random walk S stays during time T. We know that the intervals I_z and $I_{\bar{z}}$ both have empty neighborhood and empty borders. Thus all blocks of $\psi|J_z$ that are longer than $n^{0.4}$ must lie in $I_z \cap I_{\bar{z}}$ (by c of Proposition 12.4.9, in $I_z \cap I_{\bar{z}}$ there is at least one big block of ψ). Argue as in the proof of Proposition 6.5.1: because of $E_{6,S}^n$, to generate n^2+1 consecutive 0's or 1's, S must visit a block with length at least $n^{0.4}$. To have $\hat{\delta}_T^M > 0$, during T, S must have at least $e^{n^{0.2}}$ such visits. All those blocks are in $I_z \cap I_{\bar{z}} \subseteq I_{\bar{z}}$. Thus, when $\hat{\delta}_T^M > 0$, then during T, S visits $\bar{z}$ at least $e^{n^{0.2}}$ times. This means that during T, S is close to $\bar{z}$. By Corollary 6.5.1, we get $\hat{\delta}_T^M = \hat{\delta}_{z,t}^M$. □

Theorem 6.5.1. *If $E_{\text{cell\_OK}}^n$ and E_S^n both hold, then, for n big enough, $\hat{g}$ works. In other words,*

$$E_{\text{cell\_OK}}^n \cap E_S \subseteq E_{\text{g-works}}. \tag{6.5.7}$$

Proof. Suppose $E_{\text{cell\_OK}}^n \cap E_S^n$ hold. Fix $\xi = \psi \in E_{\text{cell\_OK}}^n$ and let

$$g(\psi) = (g_1(\psi), \ldots, g_{n^2+1}(\psi)).$$

We have to show: if E_S^n holds, then given the observations $\chi_0^{m^2}$, the function

$$\hat{g}(\chi_0^{m^2}) := (I_{[0.5,1]}(f_1), \ldots, I_{[0.5,1]}(f_{n^2}))$$

is equal to $\hat{g}(\psi)$ up to the first or last bit.
Let $\chi_0^{m^2}$ be the observations. Apply the $\hat{g}$-construction algorithm.
1) At the first step we pick the intervals $T_1 = [t_1, t_1 + e^{n^{0.1}}], \ldots, [t_K, t_K + e^{n^{0.1}}]$ such that for each j, $\hat{\delta}_T^M > c_r$, $j = 1, \ldots, K$. By Lemma 6.5.1 we know that each interval T_j corresponds to exactly one signal carrier point, say $\bar{z}_{\pi(j)}$.

Let us investigate the mapping $\pi : \{1, \ldots, K\} \mapsto \mathbb{Z}$, where $\pi(j)$ is the index of the signal carrier corresponding to the interval T_j. We now show that π posses the properties A1), A2), A3) that are familiar from the Subsection 6.1.6

A1) $\pi(1) \in \{0, 1\}$

A2) $\pi(K) \geq n^2 + 1$

A3) π is skip-free, i.e. $\forall j,\ |\pi(j) \pm \pi(j)| \leq 1$.

All these properties hold because of $E^n_{4,S} \cap E^n_{5,S}$. Indeed, during the time interval $[0, m^2]$ the random walk starts at 0 and, according to the event $E^n_{1,S}$, ends at m. Let $\bar{z}_1 \ldots, \bar{z}_u$ denote all signal carrier points of ψ in $[0, m]$. By E^n_1, $u > n^2$. The maximal length of a jump of S is L and, therefore, on its way, S visits all intervals $I_{\bar{z}_1}, \ldots I_{\bar{z}_u}$. Recall that the stopping times $\rho(k)$ denote the first visits of the new interval (the first visit of the next interval, not necessarily new for the past). By $E^n_{4,S} \cap E^n_{5,S}$, for each k such that $\rho(k) < m^2$ we have: there is at least $e^{n^{0.2}}$ stopping times $\hat{\nu}_{\rho(k)}(i)$ in $T := [\rho(k), \rho(k) + e^{n^{0.3}} - e^{n^{0.1}}]$. Let $\bar{z}$ be the signal carrier point such that $S(\rho(k)) \in I_{\bar{z}}$. Thus the assumptions of Proposition 6.5.1 hold and $\hat{\delta}^M_T = \hat{\delta}^M_{\bar{z},t}$. Moreover, by (7.2.7) we have that $\hat{\delta}^M_T > c_r$, i.e. the interval T will be selected in the first step of the $\hat{g}$-reconstruction.
To summarize: the random walk starts at 0, by convention the first signal carrier point in $[0, \infty)$ is $\bar{z}_1$, the biggest signal carrier point in $(-\infty, 0]$ is $\bar{z}_0$. From Lemma 6.5.1 we know - during T_1, S must be close to a signal carrier point. On the other hand $[\rho(0), \rho(0) + e^{n^{0.3}}]$ is the first time interval, during which S is close to a signal carrier point. We know that this interval will be selected. Hence $\pi(1) \in \{0, 1\}$.
On its way S visits all signal carrier interval $I_{\bar{z}_1}, \ldots I_{\bar{z}_u}$. Right after the first visit of a new signal carrier, $\rho(k)$, the random walk produces an interval $T = [\rho(k), \rho(k) + e^{n^{0.3}}]$ that will be selected. Together with Lemma 6.5.1 the latter yields that π is skip-free.
Recall that $\bar{z}_u$ is the last signal carrier point in $[0, m]$. Thus, the last signal carrier interval S visits during $[0, m^2]$ is $\bar{z}_u$ or $\bar{z}_{u+1}$. By E^n_7 we know that $\bar{z}_u$ lays in $[0, m - Le^{n^{0.3}}]$. Hence, if $S(\rho(k)) \in I_{\bar{z}_u}$, then $[\rho(k), \rho(k) + e^{n^{0.3}}]$ will be selected. We get that the last selected interval corresponds to the signal carrier that is at least $\bar{z}_{n^2+1}$. Thus $\pi(K) \geq n^2 + 1$.
Let $\pi_* := \min\{\pi(j) : j = 1, \ldots, K\}$, $\pi^* := \max\{\pi(j) : j = 1, \ldots, K\}$. We just saw that $\pi_* \leq 1$, $\pi^* \geq n^2 + 1$ and π is a skip-free random walk on $\{\pi_*, \pi_* + 1, \ldots, \pi^*\}$.
The rest of the algorithm was already explained in Subsection 6.1.6. However, in the following we give a bit more formal explanation.

2) At the second step we calculate $\hat{h}_{T_1}, \ldots, \hat{h}_{T_K}$. By Lemma 6.5.1, we know that, for each $j = 1, \ldots, K$

$$\hat{h}_{T_j} = \hat{h}_{\bar{z}_{\pi(j)}, t_j}.$$

3) Since $E^n_{8,S}$ holds, we know that, for each $j = 1, \ldots, K$,

$$|\hat{h}_{T_j} - h(\bar{z}_{\pi(j)})| = |\hat{h}_{\bar{z}_{\pi(j)}, t_j} - h(\bar{z}_{\pi(j)})| < \exp(-n^{0.12}).$$

This means: if $\pi(i) = \pi(j)$ then $|\hat{h}_{T_i} - \hat{h}_{T_j}| \leq 2\exp(-n^{0.12})$.
On the other hand, by E^n_3 we know that $\pi(i) \neq \pi(j)$ implies

$$|h(\bar{z}_{\pi(j)}) - h(\bar{z}_{\pi(i)})| \geq \exp(-n^{0.11}). \tag{6.5.8}$$

We assume n to be big enough to satisfy $\exp(-n^{0.12}) < 5\exp(-n^{0.11})$. Hence $\pi(i) \neq \pi(j)$ implies that $|\hat{h}_{T_i} - \hat{h}_{T_j}| > 2\exp(-n^{0.12})$. Thus, if $E^n_{8,S} \cap E^n_3$, then for each $i, j = 1, \ldots, k$ we have

$$\hat{h}_j \in C_i \quad \text{if and only if} \quad \pi(i) = \pi(j). \tag{6.5.9}$$

Hence the clusters C_i and C_j are either identical or disjoint; $C_i = C_j$ if and only if $\pi(j) = \pi(i)$. The same, obviously, holds for the averages:

$$\hat{f}_j = \hat{f}_i \quad \text{if and only if} \quad \pi(i) = \pi(j).$$

Let for each $i = \{\pi_*, \pi_* + 1, \ldots, \pi^*\}$, $\hat{f}(\bar{z}_i) = \hat{f}_j$, if $\pi(j) = i$. Hence, $\hat{f}(\bar{z}_i)$ is the estimate of $h(\bar{z}_i)$ and

$$\hat{f}_j = \hat{f}(\bar{z}_{\pi(j)}), \quad j = 1, \ldots, K.$$

Hence, $j \mapsto \hat{f}_j$ can be considered as the observations of the skip-free random walk π on the different reals $\{\hat{f}(\bar{z}_{\pi_*}), \hat{f}(\bar{z}_{\pi_*+1}), \ldots \hat{f}(\bar{z}_{\pi^*})\}$.

4) The real scenery construction algorithm $\mathcal{A}_n^{\mathbb{R}}$ is now able to reconstruct the numbers $\hat{f}(z_1), \ldots, \hat{f}(z_{n^2+1})$ up to the first or last number. Thus

$$(f_1, \ldots, f_{n^2}) = \mathcal{A}^{\mathbb{R}}(\hat{f}_1, \ldots, \hat{f}_K) \sqsubseteq (\hat{f}(\bar{z}_1), \ldots, \hat{f}(\bar{z}_{n^2+1})).$$

5) By E_4^n, we have that $|h(\bar{z}_i) - 0.5| \leq \exp(-n^{0.11})$. ¿From (6.5.8) and (6.5.9), it follows:

$$|\hat{f}_i - h(\bar{z}_{\pi(i)})| \leq \exp(-n^{0.12}).$$

The latter implies:

$$\hat{f}(\bar{z}_i) \geq 0.5 \quad \text{if and only if} \quad h(\bar{z}_i) \geq 0.5.$$

Hence , for each $i = 1, \ldots, n^2 + 1$, we have that $I_{[0.5,1]}(\hat{f}(\bar{z}_i)) = I_{[0.5,1]}(h(\bar{z}_i))$. Thus:

$$\hat{g}(\chi_0^{m^2}) = \Big(I_{[0.5,1]}(f_1), \ldots I_{[0.5,1]}(f_n^2)\Big) \sqsubseteq \Big(I_{[0.5,1]}(h(\bar{z}_1)), \ldots I_{[0.5,1]}(h(z_{n^2+1}))\Big) = g(\psi).$$

□

Proof of Theorem 6.1.1 Fix $c > 0$ such that Proposition 6.4.1 holds for $\epsilon = \frac{1}{8}$. Use this particular c to define all scenery dependent events as well as all random walk-dependent vents.

The intersection of all scenery-dependent events is $E^n_{\text{cell\_OK}}$. In Section 6.3.2, we proved that $P(E^n_{\text{cell\_OK}}) \to 1$. Hence **1)** holds.

Now consider the event E_S^n. Use Theorem 6.5.1 to find the integer $N_1 < \infty$ such that for each $n > N_1$, (6.5.4) hold. Then, for each $n > N_1$, $\psi_n \in E^n_{\text{cell\_OK}}$ we have

$$P(g(\chi_0^{m^2}) \sqsubseteq g(\xi_0^m) | S(m^2) = m, \xi = \psi_n) \geq P(E_S^n | S(m^2) = m, \xi = \psi_n) = P_\psi(E_S^n).$$

In Section 6.4.3, we proved that $\liminf_n P_\psi(E_S^n) \geq 1 - \frac{1}{8}$. Let N_2 be so big that $P_\psi(E_n) > \frac{3}{4}$ $\forall n > N_1$. Take $N := N_1 \vee N_2$. With such N, **2)** holds.

Finally, the statement **3)** follows from the definition of g in Section 5.1.

References

[BK96] I. Benjamini and H. Kesten. Distinguishing sceneries by observing the scenery along a random walk path. *J. Anal. Math.*, 69:97–135, 1996.

[dH88] W. Th. F. den Hollander. Mixing properties for random walk in random scenery. *Ann. Probab.*, 16(4):1788–1802, 1988.

[dHS97] F. den Hollander and J. E. Steif. Mixing properties of the generalized T, T^{-1}-process. *J. Anal. Math.*, 72:165–202, 1997.

[HHR00] D. Heicklen, C. Hoffman, and D. J. Rudolph. Entropy and dyadic equivalence of random walks on a random scenery. *Adv. Math.*, 156(2):157–179, 2000.

[HK97] M. Harris and M. Keane. Random coin tossing. *Probab. Theory Related Fields*, 109(1):27–37, 1997.

[How96] C. D. Howard. Orthogonality of measures induced by random walks with scenery. *Combin. Probab. Comput.*, 5(3):247–256, 1996.

[How97] C. D. Howard. Distinguishing certain random sceneries on $\mathbb{Z}$ via random walks. *Statist. Probab. Lett.*, 34(2):123–132, 1997.

[Kal82] S. A. Kalikow. T, T^{-1} transformation is not loosely Bernoulli. *Ann. of Math. (2)*, 115(2):393–409, 1982.

[KdH86] M. Keane and W. Th. F. den Hollander. Ergodic properties of color records. *Phys. A*, 138(1-2):183–193, 1986.

[Kes96] H. Kesten. Detecting a single defect in a scenery by observing the scenery along a random walk path. In *Itô's stochastic calculus and probability theory*, pages 171–183. Springer, Tokyo, 1996.

[Kes98] H. Kesten. Distinguishing and reconstructing sceneries from observations along random walk paths. In *Microsurveys in discrete probability (Princeton, NJ, 1997)*, pages 75–83. Amer. Math. Soc., Providence, RI, 1998.

[Lin99] E. Lindenstrauss. Indistinguishable sceneries. *Random Structures Algorithms*, 14(1):71–86, 1999.

[LM99] M. Löwe and H. Matzinger. Reconstruction of sceneries with correlated colors. Eurandom Report 99-032, accepted by Stochastic Processes and Their Applications, 1999.

[LM02] M. Löwe and H. Matzinger. Scenery reconstruction in two dimensions with many colors. *Ann. Appl. Probab.*, 12(4):1322–1347, 2002.

[LMM01] M. Löwe, H. Matzinger, and F. Merkl. Reconstructing a multicolor random scenery seen along a random walk path with bounded jumps. Eurandom Report 2001-030, 2001.

[LP02] D. Levin and Y. Peres. Random walks in stochastic scenery on $\mathbb{Z}$. Preprint, 2002.

[LPP01] D. A. Levin, R. Pemantle, and Y. Peres. A phase transition in random coin tossing. *Ann. Probab.*, 29(4):1637–1669, 2001.

[Mat99a] H. Matzinger. *Reconstructing a 2-color scenery by observing it along a simple random walk path with holding.* PhD thesis, Cornell University, 1999.

[Mat99b] H. Matzinger. Reconstructing a three-color scenery by observing it along a simple random walk path. *Random Structures Algorithms*, 15(2):196–207, 1999.

[Mat00] H. Matzinger. Reconstructing a 2-color scenery by observing it along a simple random walk path. Eurandom Report 2000-003, 2000.

[MRa] H. Matzinger and S. W. W. Rolles. Reconstructing a 2-color scenery in polynomial time by observing it along a simple random walk path with holding. Preprint.

[MRb] H. Matzinger and S. W. W. Rolles. Reconstructing a random scenery in polynomial time. Preprint.

[MRc] H. Matzinger and S. W. W. Rolles. Reconstructing a random scenery observed with random errors along a random walk path. Preprint.

Chapter 7

Retrieving the exact sequence

(submitted)

By Jüri Lember and Heinrich Matzinger,

We consider a sequence of observations obtained along a random walk from an infinite binary code (scenery). The scenery reconstruction problem is concerned with trying to retrieve the scenery, given only the observations. For a scenery with sufficiently many colors the problem was solved in [20]. This proof does not apply in two color case, and the question of reconstructing the two-color scenery, observed along a random walk with bounded jumps, has not been answered. In this paper we use the main result of [15] to show that, given some preliminary information, a long piece of two-color scenery can be reconstructed with high probability. This is the main ingredient (so-called zag-step) for the whole algorithm for two-color scenery reconstruction.

7.1 Introduction and Result

7.1.1 Introduction

A (one dimensional) *scenery* ξ is a coloring of the integers $\mathbb{Z}$ with C_0 colors $\{1, \ldots, C_0\}$. Two sceneries ξ, ξ' are called *equivalent*, $\xi \approx \xi'$, if one of them is obtained from the other by a translation or reflection. Let $(S(t))_{t\geq 0}$ be a recurrent random walk on the integers. Observing the scenery ξ along the path of this random walk, one sees the color $\xi(S(t))$ at time t. The *scenery reconstruction problem* is concerned with trying to retrieve the scenery ξ, given only the sequence of observations $\chi := (\xi(S(t)))_{t\geq 0}$. Quite obviously retrieving a scenery can only work up to equivalence. For an overview about scenery reconstruction we refer the reader to an excellent survey in [13].

The research in scenery reconstruction was first motivated by the work on the properties of the color record χ by Keane and den Hollander [11], [3]. They investigated the ergodic properties of χ, this study was motivated (among others) by the work of Kalikow [10] and den Hollander, Steif [4] in ergodic theory.
In particular, the research on scenery reconstruction started with the scenery distinguishing problem. The question was raised independently by Benjamini and Kesten in [1] and [12] as well as by den Hollander and Keane in [11]. These questions motivated many researchers to work in the areas concerning randomly observed scenery, let us just mention

Harris [5], Heicklen [6], Burdzy [2], Hoffman [6], Howard [9], [8], [7], Kesten and Spitzer [14], Levin [17], Lindenstauss [18], Rudolph [6], Pemantle [17], Peres [17].

In [12], Kesten asked whether one can recognize a single defect in a random scenery. In order to provide an answer to this question, Matzinger in his Ph.D. thesis [21] proved a somewhat stronger result: typical sceneries can be reconstructed a.s. up to equivalence. The sceneries in Matzinger's setup are independent uniformly distributed random variables. He showed that almost every scenery can be almost surely reconstructed. In [13], Kesten noticed that Matzinger's proof in [21] heavily relies on the skip-free property of the random walk. He asked whether the result might still hold in the case of a random walk with jumps. Merkl, Matzinger and Loewe in [20] gave a positive answer to Kesten's question under a particular assumption: there are strictly more colors than possible single steps for the random walk.

In the present paper we consider the following problem: can a two-color scenery be reconstructed, if it is observed along a random walk with jumps. Among others, this question was asked by H. Kesten in [13]. It turns out that the two color case ($C_0 = 2$) is more difficult than the case investigated by Merkl, Matzinger and Loewe in [20]. Although several arguments in [20] do not use the fact that there are more than two colors, the central idea hopelessly fails in the two-color case. To overcome the problem, the existence of certain test becomes crucial. The aim of the tests is to provide some information about the localization of random walk. As explained later, this kind of information makes the scenery reconstruction possible.
The existence of such kind of test was proved in [15]. This was the first important step towards the whole two-color scenery reconstruction. The present paper provides the second step of two-color scenery reconstruction. We construct an algorithm that, given some general information about the origin (stopping times) as well as a small piece of original scenery, retrieves a (long) piece of scenery with exponentially small error. With this result in hand, one can use the method described in [20] to reconstruct the whole scenery. In the terminology of [20], the constructed algorithm provides the "zag"-procedure of overall scenery reconstruction; in fact, "zag"-procedure is the core of scenery reconstruction. The whole scenery reconstruction shall be given in a follow-up paper.

7.1.2 Main notations and assumptions

We define the main concepts of the paper: scenery, random scenery random walk and observations. Also, some general notations will be introduced.

*** Scenery** is an element of $\{0,1\}^{\mathbb{Z}}$.
For every $I \subseteq \mathbb{Z}$, the elements of $\{0,1\}^I$ are called **pieces of scenery**. Given a piece of scenery $\varphi \in \{0,1\}^I$, and a subset $I' \subseteq I$, the piece of scenery $(\varphi(i))_{i\in I'}$ is denoted by $\varphi|I'$. Two pieces of scenery $\varphi \in \{0,1\}^I$ and $\varphi' \in \{0,1\}^{I'}$ are **equivalent**, $\varphi \approx \varphi'$, if φ is obtained by some translation and reflection of φ', i.e. $I' = aI + b$, for some $a \in \{-1,+1\}$, $b \in \mathbb{Z}$ and $\varphi(i) = \varphi'(ai+b)$, $\forall i \in I$. If φ is obtained from φ' by translation, i.e. $\varphi(i) = \varphi'(b+i)$, then φ and φ' are called **strongly equivalent**, we denote this $\varphi \equiv \varphi'$. If φ is obtained from φ' by reflection i.e. $\varphi(i) = \varphi'(-i)$, $\forall i \in I$, we write $\varphi = \varphi'^{-}$. By definition, $\varphi \sqsubseteq \varphi'$ means that $\varphi \equiv \varphi'|J$ for some $J \subseteq I'$. In this case φ is equal to $\varphi'|J$ up to the translation,

only.
For a piece of scenery $\varphi|[x,y]$, where $[x,y]=(x,\dots,y)\subseteq\mathbb{Z}$ is an integer interval, we often write φ_x^y. If $x=0$, then it is skipped, i.e. $\varphi|[x,y]$ is written as φ^y.

* **Random scenery** $\xi=\{\xi(z)\}_{z\in\mathbb{Z}}$ is a family of i.i.d. Bernoulli random variables with parameter $1/2$. We use ψ for a realization of ξ.
The notations defined above is valid for random sceneries. For example, ξ_x^y stands for random piece of scenery $\xi|[x,y]$, ξ^y means $\xi|[0,y]$ etc. etc.

* In this paper, $S=\{S(t)\}_{t\in\mathbb{N}}$ is a recurrent **random walk** that visits every integer z with positive probability. We assume S starts at origin, i.e. $S(0)=0$. For a $z\in\mathbb{Z}$ we denote $S_z=S+z$. An important assumption is that S has only a finite number of steps ("bounded jumps"). More precisely, we assume that the set $\{z:P(S(1)-S(0)=z)>0\}$ is finite. Throughout this paper we denote

$$L:=\max\{z:P(S(1)-S(0)=z)>0\}.$$

Thus L stands for length of the maximum jump.
We also define

$$p_L:=P(S(L)-S(0)),\quad p_{min}:=\min_i\{P(S(1)-S(0)=i)>0\}.$$

To simplify some proofs we also assume that S is symmetric (however, we do not believe that the symmetricity is necessary).

* We realize (ξ,S) as canonical projections of $\Omega=\{0,1\}^{\mathbb{Z}}\times\Omega$ endowed with product σ-algebra and probability measure $B(1,\frac{1}{2})^{\mathbb{Z}}\times Q_o$, where $\Omega_2\subseteq\mathbb{Z}^{\mathbb{N}}$ is the set of all possible paths S, Q denotes the law of S and $B(1,\frac{1}{2})$ is the Bernoulli $\frac{1}{2}$-distribution. Hence, the random walk S and scenery ξ are **independent**.
For a fixed scenery $\psi\in\{0,1\}^{\mathbb{Z}}$ (a realization of ξ), we write $P_\psi=\delta_\psi\times Q=P(\cdot|\xi=\psi)$.
We define the filtrations $\mathcal{F}:=(\mathcal{F}_n)_{n\in\mathbb{N}}$, where $\mathcal{F}_n:=\sigma(\xi,S(k):k=0,\dots,n)$ and $\mathcal{G}:=(\mathcal{G}_n)_{n\in\mathbb{N}}$, where $\mathcal{G}_n=\sigma(\chi(1),\dots\chi(n))$.

* We denote by χ the **observations** :

$$\chi:=\xi\left(S\left(0\right)\right),\xi\left(S\left(1\right)\right),\xi\left(S\left(2\right)\right),\dots$$

and we interpret χ as a random piece of scenery $\{0,1\}^{\mathbb{N}}$, so that $\chi(k):=\xi\left(S\left(k\right)\right)$ for all $k\in\mathbb{N}$.
For any $z\in\mathbb{Z}$, we denote $\chi_z(k)=\xi(S_z(k))$. The notation introduced in connection with sceneries are used with observations; in particular, for time interval $[x,y]\subseteq\mathbb{N}$ we denote

$$\chi_z|[x,y]=:\chi_{z,x}^y:=(\chi_z(x),\chi_z(x+1),\dots,\chi_z(y)),\quad\chi_z^y:=\chi_{z,0}^y,\quad\chi^y:=\chi_{0,0}^y.$$

* **Words** are the binary vectors $(w(1),\dots,w(n))$, $w(i)\in\{0,1\}$, $n\in\mathbb{N}$. Formally, words are just the pieces of sceneries φ_1^N. Therefore, all definitions introduced in connection with sceneries hold for words as well. In particular, two words w and w' can be equivalent (requires the same length) or they can satisfy the relation $w\sqsubseteq w'$. Similarly, a word w

can be equivalent to a piece of scenery $\varphi \in \{0,1\}^I$, $w \approx \varphi$ or w and φ can satisfy $w \sqsubseteq \varphi$. If $I = [a, a+n]$ for some n and $\varphi \in \{0,1\}^I$, then the equivalence $\varphi \approx w$ means that $\varphi(a) = w(1), \ldots, \varphi(a+n) = w(n+1)$ or $\varphi(a) = w(n+1), \ldots, \varphi(a+n) = w(1)$.
We shall also use the reflected words w^-. Hence, for a word $w = (w(1), \ldots, w(N))$, $w^- = (w(N), \ldots, w(1))$.
Let $I = [x, y]$. The piece of scenery $\varphi|I$ (where φ is usually ξ or χ) as a mapping consists of domain I as well as from the image. The term "word" is usually used in connection with images only. So, we consider a piece of scenery as a word, if the domain is not important or needs not to be specified (although, formally every word has a domain $(1, \ldots, N)$). Hence, we can state that "the piece φ_x^y is the word w", meaning that the image of $\varphi|I$ is w or, equivalently, $\psi_x^y \equiv w$. Depending on φ, we shall call w as the observation- or scenery-word.

7.1.3 The theorem

The aim of the paper is to show that that, for every natural number l_1 that is big enough, there exists an algorithm $\mathcal{A}^1$ which is capable with high probability to reconstruct a finite piece of ξ of length $4e^{l_1}$ *around the origin*. For that, the algorithm $\mathcal{A}^1$ uses first $\exp^{12\alpha l_1}+1$ observations, $\chi^{12\alpha l_1}$, only. Throughout the paper $\alpha > 0$ is a fixed constant that does not depend on l_1. We need α to be big enough and we specify it in Subsection 7.3.6. Since $\mathcal{A}^1$ is supposed to reconstruct the scenery around the origin, it becomes necessary to get some additional information about the location of S around the origin. In other words, besides the observations, the algorithm $\mathcal{A}^1$ should receive some signals telling him that a particular observation was generated when S was sufficiently close to the origin. To get such information, $\mathcal{A}^1$ is given $\exp(\alpha l_1)$ $\mathcal{G}$-adapted stopping times $\tau = (\tau(1), \ldots, \tau(\exp(\alpha l_1)))$ as an additional input. The stopping times are assumed to satisfy the conditions:

$$\tau(k)-\tau(k-1) \geq 2\exp(2l_1), \quad k = 2,3,\ldots,\exp(\alpha l_1)+1, \qquad \text{where } \tau(\exp(\alpha l_1)+1) := \exp[12\alpha l_1]. \tag{7.1.1}$$

The aim of τ is to show when S is at most $\exp(l_1)$ from origin. Thus, they do well, if the following event holds

$$E^1_{\text{stop}}(\tau) := \{|S(\tau(k))| \leq \exp(l_1), \quad k = 1, \ldots, \exp(\alpha l_1)\}.$$

The condition (7.1.1) states that all stopping times are sufficiently far from each other and they depend on first $\exp(12\alpha l_1)$ observation $\chi^{\exp[12\alpha l_1]}$, only. In particular, for each $\tau(k)$, the algorithm $\mathcal{A}^1$ can use $2\exp(2l_1)$ observations starting from $\tau(k)$. On $E^1_{\text{stop}}(\tau)$, all these observations are generated by S being at most $\exp(l_1) + 2\exp(2l_1)$ from origin. These are the observations that are actually used by $\mathcal{A}^1$. The information provided by τ is essential for the algorithm $\mathcal{A}^1$, which is supposed to work on $E^1_{\text{stop}}(\tau)$, only.
We shall not define the stopping times in this paper. The construction of τ such that the probability of $E^1_{\text{stop}}(\tau)$ is sufficiently big is the so-called zig-step of overall scenery reconstruction (see Chapter 3 in [20]).

Besides the observations and the stopping times, $\mathcal{A}^1$ is given the third input: a (small) piece ψ^o of original scenery. Formally, $\psi^o = \psi|I^o$, where I^o is an integer interval and ψ is the underlying scenery (the realization of ξ.) The length of ψ^o (i.e. the length of I^o) is

at least l_1c_1L, moreover, we assume $I^o \subseteq [-\exp(l_1), \exp(l_1)]$. Here c_1 is a fixed constant not depending on l_1 (see Section 7.3.6).

The output of $\mathcal{A}^1$ is a word of length $4\exp(l_1)$. Hence, formally $\mathcal{A}^1$ is the mapping

$$\mathcal{A}^1 : \{0,1\}^{[0,\exp(12\alpha l_1)]} \times [0, \exp(12\alpha l_1)]^{[1,\exp(\alpha l_1)]} \times \Big(\bigcup_{k=2c_1l_1L+1}^{2\exp(l_1)+1} \{0,1\}^k \Big) \mapsto \{0,1\}^{[-2\exp(l_1), 2\exp(l_1)]},$$

where the first input stands for observations $\chi^{12\alpha_1 l_1}$, the second for stopping times τ and the third for ψ^o.

The aim of $\mathcal{A}^1$ is to produce a piece of original scenery that lies between $\psi|[-\exp(l_1), \exp(l_1)]$ and $\psi|[-3\exp(l_1), 3\exp(l_1)]$. Recall that ψ is the realization of ξ. Thus, $\mathcal{A}^1$ does well, if the following event holds

$$E^1_{\text{alg works}}(\tau, I^o) := \Big\{ \xi|[-\exp(l_1), \exp(l_1)] \sqsubseteq \mathcal{A}^1\big(\chi^{\exp(12\alpha l_1)}, \tau, \xi|I^o\big) \sqsubseteq \xi|[-3\exp(l_1), 3\exp(l_1)] \Big\}. \tag{7.1.2}$$

Obviously the event (7.1.2) depends on τ as well as on the chosen interval I^o. In the following we do not know exactly the interval I^o. Hence, we want that $\mathcal{A}^1$ works with any given interval I^o. The corresponding event is

$$E^1_{\text{alg works}}(\tau) := \bigcap_{I^o \subseteq [-\exp(l_1), \exp(l_1)]} E^1_{\text{alg works}}(\tau, I^o).$$

The description and formal definition of $\mathcal{A}^1$ is given in Subsection 7.3.3. The main result of the paper, Theorem 7.1.1 states that the definition of $\mathcal{A}^1$ is successful: given $E^1_{\text{stop}}(\tau)$ holds, the conditional probability of $E^1_{\text{alg works}}(\tau)$ is big.

Theorem 7.1.1. *There exists a constant $k > 0$ not depending on l_1 such that, for l_1 big enough*

$$P\Big(E^1_{\text{stop}}(\tau) \cap \big(E^1_{\text{alg works}}(\tau)\big)^c \Big) \le e^{-kl_1}. \tag{7.1.3}$$

The use of τ and ψ^o might seem unrealistic - one would like to reconstruct (a piece of) scenery without any additional help. In Chapter 3 of [20], a general description of such a scenery reconstruction procedure is given. This procedure is based on repeated use of algorithms $\mathcal{A}^1$, where in every stage a longer and longer piece of scenery around origin is constructed (l_1 is increasing). In this procedure, the output of $\mathcal{A}^1$ in a lower level (for small l_1) is used to define stopping times τ in higher level (for big l_1) such that with high probability the event $E^1_{\text{stop}}(\tau)$ holds. Also the output in lower level is used as an input ψ^o for $\mathcal{A}^1$ in higher level. In the perspective of such a feedback, the result of the present paper becomes necessary; in fact, this is the core of the overall scenery reconstruction.

7.1.4 Preview

Let us briefly introduce some main ideas behind the construction of $\mathcal{A}^1$. We begin with the description of a ladder word. Let $x, y \in \mathbb{Z}$ be two location points such that $y = x + c_1l_1L$, where c_1 is a fixed constant, specified in Section 7.3.6. A ladder word w is the piece

of observations that S generates by moving from x to y as quickly as possible. Since the length of the maximum step of S is L, then for $\xi = \psi$ the described ladder word is obviously the vector

$$\Big(\psi(x), \psi(x+L), \dots, \psi(x+(c_1 l_1 - 1)L)), \psi(y)\Big). \tag{7.1.4}$$

The importance of the ladder words in scenery reconstruction comes form the fact that they can be sometimes recognized (with high probability). Indeed, suppose we "see x and y in χ", i.e. looking at the observations, we know exactly when S is in location x and in location y. In this case, we can almost surely identify (7.1.4): just look at all occurrences of x and y in χ with minimal distances. The words occurring in χ between x any y are (a.s.) always the same and equal to (7.1.4). The formal definition of ladder words is given in Section 7.3.1.

The algorithm $\mathcal{A}^1$ consists of two phases. In the first phase, $\mathcal{A}^1$ builds a collection of ladder words, $\mathcal{W}^1$. For this, we introduce a *selection rule*: an observation-word w passes the selection and will be collected as a ladder word, if it satisfies certain criterions. In the second phase, $\mathcal{A}^1$ assembles the words of $\mathcal{W}^1$ to produce a word of length $4\exp(2l_1)$ as the output. The assembling-rule of the second phase is straightforward: we start with the given piece ψ^o, and we attach a ladder word $w \in \mathcal{W}^1$ with it only if w has an overlap with ψ^o at least $\frac{c_1 l_1}{4}$. Thus, the second phase looks like a puzzle playing. The role of ψ^o becomes now obvious – ψ^o is the starting piece (the "seed") for our puzzle. For the second phase to works, it is clearly necessary that every ladder word of length $\frac{c_1 l_1}{4}$ occurs only once in $\xi|[-e^{3l_1}, e^{3l_1}]$. It turns out that for c_1 big enough, the latter holds with high probability (Proposition 7.3.1). Clearly, it is necessary that $\mathcal{W}^1$ contains enough ladder words. On the other hand, for $\mathcal{A}^1$ to work, it is also necessary that $\mathcal{W}^1$ contains only ladder words. This means that the selection rule for $\mathcal{W}^1$ must be balanced – it cannot be neither too strict nor to weak. To construct such a selection rule is the most difficult part of the scenery reconstruction.

Simplified selection rule

The selection rule is based on the fact that (with high probability) some location pairs (x, y) such that $y = c_1 l_1 L + x$ can be seen from observations. This is done by the *location tests*. Roughly speaking, a location test for y is the procedure that allows us to take decision, whether a particular observation $\chi(t)$ was generated on y (i.e. $S(t) = y$) or not. As explained before, with such information in hand, one can easily "collect" the ladder word (7.1.4).

Let us briefly introduce the main ideas behind the location test for y. For tutorial reason, we start with a very unrealistic and oversimplified version of the tests and then, step by step, we approach to the real tests.

Let $\xi = \psi$. We consider a long piece of scenery $\psi|[y, y + lm]$, where l, m are sufficiently big constants; and we aim to define a (name) function $g(\psi|[y, y + lm]) =: g_y(\psi)$ as well as a (name reading) function $\hat{g}(w)$, $w \in \{0,1\}^{lm^2+1}$ such that the following holds

1 If $S(t) \geq y$, then $\hat{g}(\chi|[t, t + lm^2])$ is able to reproduce $g_y(\xi)$ with certain positive probability;

2 If $S(t) < y$, then the probability that $\hat{g}(\chi|[t, t + lm^2])$ reproduces $g_y(\xi)$ is negligible.

In other words, we try to define the name function g and the name-reader $\hat{g}$ such that $\hat{g}(\chi|[t,t+lm^2])$ reads $g_y(\psi)$ only if the piece of observation $\chi|[t,t+lm^2]$ satisfies $S(t)\geq y$. Similarly, to get a location test for x, we define the name function $g^*(\psi|[x-lm,x]) =: g^*_x(\psi)$ and the (name reading) function $\hat{g}^*(w)$, $w\in\{0,1\}^{lm^2+1}$ such that the following holds

1* If $S(t)\leq x$, then $\hat{g}^*(\chi|[t-lm^2,t])$ is able to reproduce $g^*_x(\xi)$ with certain positive probability;

2* If $S(t)>x$, then the probability that $\hat{g}^*(\chi|[t-lm^2,t])$ reproduces $g^*_x(\xi)$ is negligible.

It is easy to see that g^* and $\hat{g}^*$ can be deduced from g and $\hat{g}$ – just define $g^*(w):=g(w^-)$ and $\hat{g}^*(w):=\hat{g}(w^-)$.

Suppose, for a moment, that we have a working location tests for a pair (x,y), with $y=x+c_1l_1L$. Moreover, suppose that "being able to reproduce" above just means equalities $\hat{g}(\chi|[t,t+lm^2])=g_y(\psi)$, $\hat{g}^*(\chi|[t,t+lm^2])=g^*_x(\psi)$ and "is negligible" means being zero. In this case, the reconstruction (or collecting) of the word (7.1.4) is rather straightforward. Indeed, for each $t\geq 0$ define the observation words

$$w^1(t):=\chi|[t-lm,t],\quad w^2(t):=\chi|[t,t+c_1l_1],\quad w^3(t):=\chi|[t+c_1l_1,t+c_1l_1+lm^2] \tag{7.1.5}$$

and apply the name-reading functions $\hat{g}^*(w^1(t))$ and $\hat{g}(w^3(t))$. Because S is recursive, a.s. there exists a t such that $\hat{g}^*(w^1(t))=g^*_x(\psi)$ and $\hat{g}(w^3(t))=g_y(\psi)$. In particular, this implies that

$$S(t)\leq x\quad\text{and}\quad S(t+c_1l_1)\geq y. \tag{7.1.6}$$

On the other hand, during c_1l_1 steps, the random walk S cannot move more than c_1l_1L. But this is exactly the distance between x and y. Hence, the only possibility for (7.1.6) to hold is that both inequalities are equalities. In this case, $w^2(t)$ equals the ladder word (7.1.4).

The example above is unrealistic in many respect. It is obvious that a necessary condition for the location test to work is that there is no $z<y$ such that $\psi|[z,z+lm]=\psi|[y,y+lm]$. But from the definition of ξ it follows that for almost all realizations such a z exists (any finite pattern occurs infinitely many times in ξ). Therefore, it is more realistic to assume that the word $\psi|[y,y+lm]$ is unique in a certain piece of $\psi|I_1$, only. Since we are interested in reconstructing the scenery around the origin, from now on, we define

$$I_1:=[-\exp(3l_1),\exp(3l_1)]$$

and we consider the pairs (x,y) in I_1, only. Thus the conditions **2** and **2*** are replaced by

$$P\Big(\hat{g}(\chi|[t,t+lm^2])=g_y(\psi),\quad S(t)\in[-\exp(3l_1),y]\Big)=0 \tag{7.1.7}$$

$$P\Big(\hat{g}^*(\chi|[t-lm^2,t])=g^*_x(\psi),\quad S(t)\in[x,\exp(3l_1)]\Big)=0. \tag{7.1.8}$$

Since the above-described selection rule now works only on I_1, we have to modify the construction of (7.2.1) such that $S(t),S(t+c_1l_1)\in I_1$. For this we use the stopping times $\tau(j)$. Define times

$$T^1(j):=\tau(j)+\exp(2l_1)+lm^2,\quad T^3(j):=T^1(k)+c_1l_1,\quad j=1,\dots,\exp(\alpha l_1). \tag{7.1.9}$$

Note that on $E_{\text{stop}}(\tau)$ it holds $S(T^1(j)), S(T^3(j)) \in I_1$, provided l_1 is big enough. Now the words defined by $T^1(j)$ and $T^3(j)$ can be used. More precisely, we define

$$\begin{aligned} w^1(j) &:= \chi|[T^1(j) - lm^2, T^1(j)] \\ w^2(j) &:= \chi|[T^1(j), T^3(j)] \\ w^3(j) &:= \chi|[T^3(j), T^3(j) + lm^2] \end{aligned}$$

and we use the same selection rule as previously, with $w^1(j), w^2(j), w^3(j)$ instead of $w^1(t), w^2(t), w^3(t)$. Note that a necessary condition for this rule is that the probability in **1** and **1*** is so big that among $\exp(\alpha l_1)$ stopping times most likely there is at least one j such that $\hat{g}^*(w^1(j)) = g_x^*(\psi)$ and $\hat{g}(w^3(j)) = g_y(\psi)$. Also note that the $T^1(j)$ is not defined right after $\tau(j)$, but after $\tau(j) + \exp(2l_1)$, instead. The reason for this is following: we are interested in reconstructing the a piece of scenery with length $4\exp(l_1)$ around origin (recall the definition of $E^1_{\text{alg works}}$). This means that we have to collect also these ladder words that are about $2\exp(l_1)$ from origin. The stopping times $\tau(j)$ stop S at most $\exp(l_1)$ from origin (on $E_{\text{stop}}(\tau)$). Hence, for S to reach to the ladder words that are are about $2\exp(l_1)$ from origin, some additional time is needed.

The rule in the previous example requires that we know the names $g_x^* := g_x^*(\psi)$ and $g_y := g_y(\psi)$. They depend on ψ that is unknown. However, by conditions **1** and **1***, the names g_x^* and g_y can be red with positive probability. We now modify the selection rule to take into consideration that g_x^* and g_y are not known. The modification is based on the fact that the probability to read g_x^* and g_y is so big that among $\exp(\alpha l_1)$ pairs $\hat{g}^*(w^1(j)), g(w^3(j))$ there is at least $\exp(\gamma l_1)$ pairs such that $\hat{g}^*(w^1(j) = g_x^*$ and $g(w^3(j)) = g_y$ (with high probability, of course). Here $0 < \gamma < \alpha$ is a properly chosen proportion. If the latter holds, then there exists a pair of names g_1^*, g_3 such that the number of stopping times satisfying $\hat{g}^*(w^1(j) = g_1^*$ and $g(w^3(j)) = g_3$ is more than $\exp(\gamma l_1)$. Unfortunately, there can be many pairs having the same property. To choose the right pair, we riep benefit from the conditions (7.1.7) and (7.1.8). Due to these condition, the right pair of names g_x^*, g_y has an important characteristic – for every j such that $\hat{g}^*(w^1(j)) = g_x^*$ and $\hat{g}(w^3(j)) = g_y$, the word $w^2(j)$ must be (7.1.4) and, therefore, the same. Our modified rule is the following:

Simplified selection: The word w is taken as (7.1.4), if there exists a pair of names g_1^*, g_3 such that the following holds:

a) there exists more than $\exp(\gamma l_1)$ stopping times such that

$$\hat{g}^*(w^1(j)) = g_1^*, \quad \hat{g}(w^3(j)) = g_3; \tag{7.1.10}$$

b) for every j satisfying (7.1.10), it holds $w^2(j) = w$.

Avoiding non-ladder words

In the selection rule above, the right choice of γ is crucial: if γ is too big, then the probability that the true ladder word passes the criterion **a)** becomes too small. On the other hand, if γ is too small, then the probability that a non-ladder word passes the selection rule becomes too big. Let us briefly introduce the basic argument used to find

a suitable lower bound for γ.
Suppose $z, z' \in I_1$ such that $|z - z'| < Lc_1l_1$. Consider the possible observation-words that S generates by going from z to z' in c_1l_1 steps. If c_1 is big enough, then the probability that all these words are the same, is small (Proposition 7.3.1). In Section 7.3.1 we define the event $B^1_{\text{recon straight}}$ which states that for every $z, z' \in I_1$ there are at least two possible observation-words that S can generate during its way from z to z' with c_1l_1 steps. Any path of S that consists of c_1l_1 steps has the probability at least $(p_{min})^{c_1l_1}$.
Suppose w passes the selection rule. Hence, there exists a set $J \subseteq \{1, \dots, \exp(\alpha l_1)\}$ such that at least $|J| \geq \exp(\gamma l_1)$ and for each $j \in J$ the following holds : $|S(T^3(j)) - S(T^1(j))| < Ll_1c_1$ and $w^2(j) = w$. Let $Y_k := 1 - I_{w^2(j_1)}(w^2(j_k))$, where $j_1, j_2, \dots$ are the elements of J. This means that $\sum_{k=2}^{\exp(\gamma l_1)} Y_k = 0$. Suppose now that w is a non-ladder word. If the event $E^1_{\text{stop}} \cap B^1_{\text{recon straight}}$ holds, then, for each $k \geq 2$, the probability that $Y_k = 1$ cannot be smaller than $(p_{min})^{c_1l_1}$. Given $S(T^1(j_k))$ and $S(T^3(j_k))$ the random variables Y_k are independent. Now the Höffding's inequality can be used to estimate (see (7.3.21))

$$P\Big(\sum_{k=2}^{\exp(\gamma l_1)} Y_k = 0 \Big| E^1_{\text{stop}} \cap B^1_{\text{recon straight}}\Big) \leq \exp[-2\exp((\gamma + 2c_1 \ln p_{min})l_1)].$$

The right side of the previous display is exponentially small in exponentially small quantity of l_1, if $\gamma > -2c_1 \ln p_{min}$ (see 7.3.31). Using the obtained bound, it is not hard to see that the probability that a non-ladder word passes the selection rule is exponentially small in l_1 (Proposition 7.3.2)..
Note that in the foregoing argument we did not use any properties of g and $\hat{g}$. Hence, the argument applies also for the final selection rule given in Subsection 7.1.4.

The names

In this subsection, we explain the nature of the functions g and $\hat{g}$ (recall that $\hat{g}^*$ and g^* are practically the same). The construction of these function is based on the following theorem proved in [15]

Theorem 7.1.2. *There exists constants $c > 0$ (not depending on n), $N < \infty$, $m(n) > n$, the maps*

$$g : \{0,1\}^{m+1} \mapsto \{0,1\}^{n^2+1}$$
$$\hat{g} : \{0,1\}^{m^2+1} \mapsto \{0,1\}^{n^2}$$

and the sequence of events $B_{\text{cell\_OK}}(n) \in \sigma(\xi(z) | z \in [-cm, cm])$ such that:

1) $P(B_{\text{cell\_OK}}(n)) \to 1$

2) *For all $n > N$ and $\psi_n \in B_{\text{cell\_OK}}(n)$:*

$$P_{\psi_n}\Big(\hat{g}(\chi_0^{m^2}) \sqsubseteq g(\psi_0^m) \Big| S(m^2) = m\Big) = P\Big(\hat{g}(\chi_0^{m^2}) \sqsubseteq g(\psi_0^m) \Big| S(m^2) = m, \xi = \psi_n\Big) > 3/4.$$

3) *$g(\xi_0^m)$ is an i.i.d. binary vector where the components are Bernoulli with parameter $1/2$.*

(Note the abuse of notation: in [15] the sign "$\preccurlyeq$" was used instead of "$\sqsubseteq$".)
From now on we assume that $n > N$ and $m(n)$ are fixed constant. We specify them in Section 7.3.6. Theorem 7.1.2 provides a test that uses m^2 observations $\chi_t^{t+m^2}$ to test the hypotheses:

$$H_o : S(t) = y,$$
$$H_1 : S(t) < y - Lm^2$$

given $S(t+m^2) = S(t) + m$ and $\xi \in B_{\text{cell\_OK}}(n)$. Indeed, it $S(t) < y - Lm^2$, then $\chi_t^{t+m^2}$ is independent of $g(\xi_y^{y+m})$. By the properties of ξ,

$$P\Big(\hat{g}(\chi_t^{t+m^2}) \sqsubseteq g(\xi_y^{y+m})\Big) = \Big(\hat{g}(\chi_t^{t+m^2}) \sqsubseteq g(\xi_0^m)\Big) \le \Big(\frac{1}{2}\Big)^{n^2-1}.$$

On the other hand, if $\psi \in E_{\text{OK}}$, then conditional on $A := \{\xi \in B_{\text{cell\_OK}}(n), S(t+m^2) = m, S(t) = y\}$ it holds

$$P\Big(\hat{g}(\chi_t^{t+m^2}) \sqsubseteq g(\xi_y^{y+m})\Big|A\Big) > \frac{3}{4}.$$

The functions g and $\hat{g}$ look like the desired name and name-reading procedures. Indeed, there is certainly a positive probability that $\hat{g}(w^3(j))$ "reproduces" $g(\psi|[y, y+m])$, where "reproducing" now means the relation $\hat{g}(w^3(j)) \sqsubseteq g_y$ (note that in this case "$\sqsubseteq$" actually means the equality to the first or last bit). On the other hand, the following modification of the (7.1.7) holds

$$P\Big(\hat{g}(\chi|[t, t+m^2]) \sqsubseteq g(\psi|[y, y+m]), \quad S(t) \in [-\exp(3l_1), y - Lm^2)\Big) = \Big(\frac{1}{2}\Big)^{n^2-1}. \quad (7.1.11)$$

So, taking n big enough, we can make the right side of (7.1.11) as small as we want.
Unfortunately, for several reasons, the functions from Theorem 7.1.2 is not good enough. Recall that we want the mistake (7.1.3) to be exponentially small in l_1. The right side of (7.1.11) does not depend on l_1. To handle this, we apply Theorem 7.1.2 repeatedly. This procedure is called *iteration* and it is the subject of Section 7.2. Let us briefly introduce the main ideas behind the iteration.

From now on, we define

$$l := l_1 \cdot l_2, \quad \text{where } l_2 \text{ is fixed positive integer, specified in Section 7.3.6.}$$

We shall apply the functions g and $\hat{g}$ from Theorem 7.1.2 l times consecutively. Let $w = (w(0), \dots, w(lm)) \in \{0,1\}^{lm+1}$. We define l sub-words, called *cells*

$$w_i = \big(w((i-1)m), \cdots, w(im)\big), \quad i = 1, \dots, l.$$

Note that w_i and w_{i+1} are not disjoint. Using the sub-words w_i, we naturally extend the definition of g to the words in $\{0,1\}^{lm+1}$. We define

$$g : \{0,1\}^{lm+1} \mapsto \{0,1\}^{l(n^2+1)}, \quad g(w) = (g(w_1), \dots, g(w_l)).$$

Note that we denote by g the function in Theorem 7.1.2 as well as its extension (they coincide if $l = 1$).

Similarly, let $v = (v(0), \dots, v(lm^2)) \in \{0,1\}^{lm^2+1}$. We define cells

$$v_i = \big(v((i-1)m^2), \dots, v(im^2)\big), \quad i = 1, \dots, l.$$

Using the sub-words v_i, we extend the definition of $\hat{g}$ to the words in $\{0,1\}^{lm^2+1}$. We define

$$\hat{g} : \{0,1\}^{lm^2+1} \mapsto \{0,1\}^{ln^2}, \quad \hat{g}(v) = (\hat{g}(v_1), \dots, \hat{g}(v_l)).$$

We now give a more accurate interpretation to the phrase "to reproduce" in the description **1**. Since the "name-reading" or "reproducing" procedure is based on Theorem 7.1.2, it is natural to expect that $\hat{g}(\chi|[t, t+m^2 l])$ reproduces $g(\psi|[y, y+ml])$, if the relation $\sqsubseteq$ holds cell-wise, i.e. $\hat{g}(\chi|[t+(i-1)m^2, t+im^2]) \sqsubseteq g(\psi|[y+(i-1)m, y+im])$ for each $i = 1, \dots, l$. Note that Theorem 7.1.2 gives lower bound to the probability

$$P_\psi\Big(\hat{g}(\chi|[t+(i-1)m^2, t+im^2]) \sqsubseteq g(\psi|[y+(i-1)m, y+im])\Big),$$

only if the piece of scenery $\psi|[y+(i-1)m - cm, y+(i-1)m+cm]$ belongs to the set $E^n_{\text{cell\_OK}}$. If this is the case, we say that the cell $\psi|[y+(i-1)m, y+im]$ is *OK*.
For each (long) piece of scenery $\psi|[y, y+lm]$ we now correspond the index set $\mathcal{I}(\psi|[y, y+lm]) =: \mathcal{I}_y(\psi) \subseteq \{1, \dots, l\}$ of OK-cells. Similarly, we define $\mathcal{I}^*(\psi|[x-lm, x]) := \mathcal{I}((\psi|[x-lm, x])^-)$ (the reader should be warned that now we only give a simplified definition of $\mathcal{I}$ and $\mathcal{I}^*$; the final definition is given in Section 7.2.1).
Although $E^n_{\text{cell\_OK}}$ has the probability close to one, since l is big, we expect a proportion of cells not to be OK, i.e $\mathcal{I}_y \neq \{1, \dots, l\}$. We say that $\psi|[y, y+lm]$ is OK, if at least $l(1-3\epsilon)$ cells are OK, i.e $|\mathcal{I}_y(\psi)| \geq l(1-3\epsilon)$. We say that $\psi|[x-lm, x]$ is OK*, if $(\psi|[x-lm, x])^-$ is OK. Equivalently, $\psi^-|[-x, -x+lm]$ is OK. We denote by $B^1_{\text{intervals OK}}$ the set of sceneries that satisfy: $\psi|[y, y+lm]$ is OK and $\psi|[x-lm, x]$ is OK* for every pair $(x, y) \in I_1$. In particular, if $\psi \in B^1_{\text{intervals OK}}$, then $|\mathcal{I}_y(\psi)|, |\mathcal{I}^*_x(\psi)| \geq (1-\epsilon)l$. The proportion ϵ is chosen such that $P(B^1_{\text{intervals OK}})$ is sufficiently big (Theorem 7.2.1 and the estimation (7.3.10)).
For not OK cells, the statement **2)** of Theorem 7.1.2 needs not hold, and the cell-wise reproducing might fail. Hence, we relax the requirement of the full cell-wise reproducing to the requirement that the OK cells are reproduced. More formally, for any subset $I \subseteq \{1, \dots, l\}$, we define $\hat{g}(w) \sqsubseteq_I g(v)$, if $\hat{g}(w_i) \sqsubseteq g(v_i)$, $\forall i \in I$. Now we say that $g(\chi|[t, t+m^2 l])$ reproduces $g_y(\psi)$, if

$$g(\chi|[t, t+m^2 l]) \sqsubseteq_{\mathcal{I}(\psi)} g_y(\psi).$$

If $\psi \in B^1_{\text{intervals OK}}$, then the latter means that cell-wise reproduction holds for at least $l(1-3\epsilon)$ cells.

Getting selected

Let us now give some insight, how do we show that the probability for a ladder word (7.1.4) to pass the selection is sufficiently high. What follows, is a simplified version of Proposition 7.3.2. In the present subsection we assume that

$$|x|, |y| \leq 4\exp(l_1).$$

Define

$$E_j(x,y) := \left\{ \begin{array}{c} S(T^1(j) - lm^2) = x - lm \\ S(T^1(j)) = x,\ S(T^3(j)) = y, \\ \hat{g}^*(w^1(j)) \sqsubseteq_{\mathcal{I}_x^*(\xi)} g_x^*(\xi), \\ \hat{g}(w^3(j)) \sqsubseteq_{\mathcal{I}_y(\xi)} g_y(\xi) \end{array} \right\}, \quad Y_j := I_{E_j}, \quad j = 1, \ldots, e^{\alpha l_1}.$$

Clearly (7.1.4) passes the selection if

$$\Big\{ \sum_{j=1}^{e^{\alpha l_1}} Y_j > e^{\gamma l_1} \Big\}.$$

Now, by the Markov property of S, for each ψ

$$\begin{aligned} P_\psi\big(Y_j = 1 | E_{\text{stop}}(\tau)\big) = & P_\psi\Big(S(T^1(j) - lm^2) = x - lm \Big| E_{\text{stop}}(\tau)\Big) \\ & \times P_\psi\Big(S(T^1(j)) = x,\ \hat{g}^*(w^1(j)) \sqsubseteq_{\mathcal{I}_x^*(\psi)} g_x^*(\psi) \Big| S(T^1(j) - lm^2) = x - lm\Big) \\ & \times P_\psi\Big(S(T^3(j)) = y \Big| S(T^1(j)) = x\Big) \\ & \times P_\psi\Big(\hat{g}(w^3(j)) \sqsubseteq_{\mathcal{I}_y(\psi)} g_y(\psi) \Big| S(T^3(j)) = y\Big). \end{aligned}$$

Recall that $T^1(j) - lm^2 = \tau(j) + \exp(2l_1)$. By $E_{\text{stop}}(\tau)$, $|S(\tau(j))| \le \exp(l_1)$. We use the local central limit theorem (LCLT) to estimate

$$\begin{aligned} P_\psi\big(S(\tau(j) + e^{2l_1}) = x - lm \big| E_{\text{stop}}(\tau)\big) & \ge \inf_{z:|z| \le 4\exp(l_1)} P_\psi\big(S(\tau(j) + e^{2l_1}) = z \big| E_{\text{stop}}(\tau)\big) \\ & \ge \inf_{z:|z| \le 5\exp(l_1)} P_\psi\big(S(e^{2l_1}) = z) \ge \exp(-1.5 l_1), \end{aligned}$$

provided l_1 is big enough. By the definitions of $w^1(j)$, $\hat{g}^*$ and $\mathcal{I}^*$, we have

$$\begin{aligned} & P_\psi\Big(S(T^1(j)) = x,\ \hat{g}^*(w^1(j)) \sqsubseteq_{\mathcal{I}_x^*(\psi)} g_x^*(\psi) \Big| S(T^1(j) - lm^2) = x - lm\Big) = \\ & P_\psi\Big(S(T^1(j)) = x,\ \hat{g}^*(\chi | [T^1(j) - lm^2, T^1(j)]) \sqsubseteq_{\mathcal{I}(\psi|[x-lm,x])^-} g\big((\psi|[x-lm,x])^-\big) \Big| \\ & S(T^1(j) - lm^2) = x - lm\Big) = \\ & P_\psi\Big(S(lm^2) = x,\ \hat{g}^*(\chi_{x-lm} | [0, lm^2]) \sqsubseteq_{\mathcal{I}(\psi^-|[-x,-x+lm])} g(\psi^- | [-x, -x+lm])\Big) = \\ & P_\psi\Big(S(lm^2) = x,\ \hat{g}\big((\chi_{x-lm}|[0,lm^2])^-\big) \sqsubseteq_{\mathcal{I}_{-x}(\psi^-)} g_{-x}(\psi^-)\Big) \end{aligned}$$

By symmetry of S, for each set $\mathcal{V} \subseteq \{0,1\}^{ln^2}$, we have

$$P_\psi\Big(S(lm^2) = x,\ \hat{g}\big((\chi_{x-lm}|[0,lm^2])^-\big) \in \mathcal{V}\Big) = P_\psi\Big(S(lm^2) = x - lm,\ \hat{g}\big(\chi_x|[0,lm^2]\big) \in \mathcal{V}\Big).$$

The right side of the previous display equals

$$P_{\psi^-}\Big(S(lm^2) = -x + lm,\ \hat{g}\big(\chi_{-x}|[0,lm^2]\big) \in \mathcal{V}\Big).$$

Hence,

$$P_\psi\Big(S(T^1(j))=x,\ \hat{g}^*(w^1(j))\sqsubseteq_{\mathcal{I}^*_x(\psi)} g^*_x(\psi)\Big|S(T^1(j)-lm^2)=x-lm\Big)=$$
$$P_{\psi^-}\Big(S(lm^2)=-x+lm,\ \hat{g}\big(\chi_{-x}|[0,lm^2]\big)\sqsubseteq_{\mathcal{I}_{-x}(\psi^-)} g_{-x}(\psi^-)\Big)=$$
$$P_{\psi^-}\Big(S(T^3(j)+lm^2)=-x+lm,\ \hat{g}\big(\chi|[T^3(j),T^3(j)+lm^2]\big)\sqsubseteq_{\mathcal{I}_{-x}(\psi^-)} g_{-x}(\psi^-)\Big|S(T^3(j)=-x)\Big)=$$
$$P_{\psi^-}\Big(S(T^3(j)+lm^2)=-x+lm,\ \hat{g}\big(w^3(j)\big)\sqsubseteq_{\mathcal{I}_{-x}(\psi^-)} g_{-x}(\psi^-)\Big|S(T^3(j)=-x)\Big).$$

Suppose $\psi\in B^1_{\text{intervals OK}}$. Then the probability in the previous display has the lower bound

$$\inf_{\psi:\psi|[y,y+lm]\text{ is OK}} P_\psi\Big(S(T^3(j)+lm^2)=y+lm,\ \hat{g}(w^3(j))\sqsubseteq_{\mathcal{I}_y(\psi)} g_y(\psi)\Big|S(T^3(j))=y\Big). \tag{7.1.12}$$

Indeed, (7.1.12) does not depend on y any more. It is not very hard to see now that by **2)** of Theorem 7.1.2, (7.1.12) can be bounded below by

$$\inf_{\psi:\psi|[y,y+lm]\text{ is OK}}\ \prod_{i\in\mathcal{I}(\psi)} P_\psi\Big(\hat{g}(\chi^{im^2}_{(i-1)m^2})\sqsubseteq g(\psi^{im}_{(i-1)m})\Big|S(im^2)=S((i-1)m^2)+m\Big)\geq\Big(\frac{3}{4}\Big)^l.$$

Finally, for every ψ,

$$P_\psi\Big(S(T^3(j))=y\Big|S(T^1(j))=x\Big)=(p_L)^{c_1l_1}.$$

Hence, if $\psi\in B^1_{\text{intervals OK}}$, we have

$$P_\psi\big(Y_j=1|E_{\text{stop}}(\tau)\big)\geq\exp(-1.5l_1)\Big(\frac{3}{4}\Big)^l(p_L)^{c_1l_1}\Big(\frac{3}{4}\Big)^l=\exp[-(1.5-2\ln(\frac{3}{4})l_2-c_1\ln(p_L))l_1]. \tag{7.1.13}$$

Conditional on E_{stop} and ψ, the random variables Y_j are independent. Using Höffding's inequality, it is now not difficult to show that α and γ can be chosen such that

$$P\Big(\sum_{j=1}^{e^{\alpha l_1}}Y_j\leq e^{\gamma l_1},B^1_{\text{intervals OK}}\cap E_{\text{stop}}(\tau)\Big)$$

is exponentially small in l_1. Since $P(B^1_{\text{intervals OK}})$ is big (7.3.10), we obtain that the the probability of selecting (7.1.4) is sufficiently big.

Avoiding mistakes

In the previous subsections we saw how the selection rule works if "being negligible" in **2** means "equal to zero". The latter is unrealistic and cannot be guaranteed. We now modify the selection rule such that the the probability in **2** is considerably small in comparison with the (modified version of the) right side of (7.1.13) (which also goes to zero as l_1 grows). To explain the meaning of the additional modification, we consider the events

$$E_{z,I}:=\big\{\forall i\in I\text{ we have that }S_z(m(i-1))<m(i-1)-Lm^2\big\},\quad I\subseteq\{1,\dots,l\}. \tag{7.1.14}$$

Suppose $E_{z,I}$ holds. Then, for each cell $i \in I$, the random variables $\chi_z|[(i-1)m, im]$ and $\xi|[(i-1)m, im]$ are independent. By **3** of Theorem 7.1.2, we then have $P(\chi_z|[(i-1)m, im] \sqsubseteq \xi|[(i-1)m, im]) = (0.5)^{n^2-1}$. This implies $P(\hat{g}(\chi_z^{lm^2}) \sqsubseteq_I g_y(\xi)) \le (0.5)^{(n^2-1)|I|}$ and, for l big enough the latter yields

$$P\left(B^1_{\text{intervals OK}} \cap \{\hat{g}(\chi_z^{lm^2}) \sqsubseteq_{\mathcal{I}_y(\xi)} g_y(\xi)\} \cap E_{z,\mathcal{I}(\xi)}\right) \le \exp[-(0.3n)l]. \tag{7.1.15}$$

(Corollary 7.2.1). Recall that on $B^1_{\text{intervals OK}}$. Since n can be chosen very big, the right side of (7.1.13) can be as many times bigger than $\exp[-(0.3n)l]$ as we want. This property together with the fact that $P(B^1_{\text{intervals OK}})$ is big makes the selection rule work.

We now define an additional characteristic of $\psi|[y, y+lm]$, denoted by $q(\psi|[y, y+lm]) =: q_y(\psi)$, and corresponding "reading function" $\hat{q}(w)$, $w \in \{0,1\}^{lm^2+1}$ such that for each j, we have

3 If $S(T^3(j)) \ge y$, then $\hat{q}(w^3(j))$ reproduces $q_y(\xi)$ with certain probability,

4 If $S(T^3(j)) < y$, then $\hat{q}(w^3(j))$ reproduces $q_y(\xi)$ only if $E_{z,\mathcal{I}(\xi)}$ holds.

Denote $z = T^3(j)$. Note the difference with **1** and **2**: if $z \ge y$, then $\hat{q}$ and q must fulfill the requirement like **1**. Of course, the meaning of "reproduction" is now different, we shall call it q-reproduction. For $z < y$, the requirements for q and $\hat{q}$ are different from that one in **2** – we do not require that the probability for q-reproduction is small. We require instead that the q-reproducing always implies $E_{z,\mathcal{I}(\xi)}$. And then, as we just saw, the probability that $\hat{g}(w^3(j)) \sqsubseteq_{\mathcal{I}(\xi)} g_y(\xi)$ (the g-reproduction, in the sequel) is exponentially small (at least for $y = 0$, but the case for general y is not different). Hence, we consider g and q together. For a ladder word to be selected, both q-and g-reproduction must simultaneously hold (for $\exp(\gamma l_1)$ stopping times, as usually). In the case $z \ge y$, the additional requirement obviously reduces the probability (7.1.13); however, if the q-reproduction has a relatively big probability, then the lower bound like (7.1.13) might still hold. In the case $z < y$, the q-reproduction of $q_y(\xi)$ (which might hold with rather big probability) implies $E_{z,\mathcal{I}(\xi)}$, and then the probability of g-reproduction is very small.

The idea of q-reproduction is partially based on the fact that we do not need every ladder word (7.1.4) with $x, y \in I_1$ do be collected. So far, we have not restricted our choice of x (y is obviously uniquely determined by x). Now we consider pairs (x, y) that satisfy pair (x, y) that

$$\psi(y-L) = \cdots = \psi(y-1) \ne \psi(y) = \cdots = \psi(y+m^3L) \ne \psi(y+m^3L+1) = \cdots = \psi(y+m^3L+L) \tag{7.1.16}$$

$$\psi(x+L) = \cdots = \psi(x+1) \ne \psi(x) = \cdots = \psi(x-m^3L) \ne \psi(x-m^3L-1) = \cdots = \psi(x-m^3L-L). \tag{7.1.17}$$

Such pairs are called a *barriers*. The barriers are random, they depend on ξ. The event $B^1_{\text{enough barriers}}$, formally defined in Section 7.3.1 states that we have sufficiently many barriers. In Proposition 7.3.1 we show that this event has high probability if l_1 is big enough. To the end of this section we assume $y = 0$ and we skip y from the notation.

Let $\psi|[(2Lm^2-1)m,(2Lm^2)m]$ be the first OK cell of ψ. In terms of cell indexes, $2Lm^2 = i_1 := \min \mathcal{I}(\psi)$.
Let $z < y$. We consider now the random walk S_z, and we want to be able to see from the observations $\chi_z|[0,(i_1-1)m^2]$ whether $S_z((i_1-1)m^2) < (i_1-1)m - Lm^2$, i.e. E_{z,i_1} holds. The number $m(n)$ is certainly so big that $(2Lm^2-1)m - Lm^2 > Lm^3$. Hence E_{z,i_1} holds, if $S_z((i_1-1)m^2) \le m^3L$. The latter obviously holds $S_z(t) \le m^3L\ \forall t \le (i_1-1)m^2$, which, in turn, holds if the observation-word $\chi_z|[0,(i_1-1)m^2]$ has the following property: $\chi_z|[0,(i_1-1)m^2]$ does not contain at least m^3 consecutive same colors *followed by the different color*. Indeed, in order to reach a point $z' > m^3L$, the random walk S_z must generate at least m^3 consecutive same-color observations and then at least one observation of the other color.
Hence, when $\chi_z|[0,(i_1-1)m^2]$ satisfies the mentioned condition, we can be sure that $S_z(t) \le m^3L\ \forall t \le (i_1-1)m^2$, i.e. E_{z,i_1} holds. If the condition is not met, then the word $\chi_z|[0,(i_1-1)m^2]$ is not considered for g-reproduction, it will be filtered out.
Suppose now $z = 0$. In this case we want that $\chi_z|[0,(i_1-1)m^2] = (i_1-1)m$. This gives a big chance for g-reproduction of the i_1-th cell $\hat{g}(\chi|[(i_1-1)m^2, i_1m^2]) \sqsubseteq g(\psi|[(i_1-1)m, i_1m])$. But in this case the observation-word $\chi|[0,(i_1-1)m^2]$ definitely contains m^3 consecutive same colors followed by the different color and such a word will be filtered out. Therefore, we must adjust the described condition to make sure that (with certain probability) the word $\chi|[0,(i_1-1)m^2]$ will be not filtered out. For this note: in order to reach from $z < 0$ to $z' > m^3L$, the random walk must generate (in the observations) at least m^3 consecutive same colors, having *the different color at the beginning and at the end*. On the other hand, to reach from 0 to $z' > m^3L$, the random walk can follow the path that begins with m^3 same colors, hence the word $\chi|[0,(i_1-1)m^2]$ will not necessarily contain least m^3 consecutive same colors with the different color in the beginning (although this event has probability bigger than $\frac{1}{2}$).
A word $(w(0), w(1), \ldots, w(u-1), w(u))$ is called *block* with length u, if $w(0) \ne w(1) = \cdots = w(u-1) \ne w(u)$. Hence the filtering rule is: the word $\chi|[0,(i_1-1)m^2]$ will be filtered out, if it contains a block with length at least m^3. Such blocks are called *big*.
For each block B in ψ, we define *the reading length* of B as the length of the smallest block that the random walk generates in observations by crossing it. If the length of B is Lm^3, then the reading length of B is roughly m^3 (see Section 7.2.3 for the formal definition and examples). Suppose now that $i_1 > 2Lm^2$ and there is one block with B the reading length at least m^3 between $m^3L + L$ and $(i_1-1)m - Lm^2$. Then, to reach $(i_1-1)m$ from y, the random walk necessarily generates at least one big block in observation. To reach $(i_1-1)m$ from $z < 0$, the random walk necessarily generates at least two big block in observations. Hence, the filtering rule in this case is: $\chi|[0,(i_1-1)m^2]$ will be filtered out, if it contains more than one big block.

Generally, we proceed as follows: we define $\mathcal{I}(\psi)$ to be indexes if cells that are not only OK, but have the additional property: if $i \in \mathcal{I}(\psi)$ then $\psi|[(i-1)m - Lm^2, im + lm^2]$ cannot be a part of any block with reading length at least m^3 (see Section 7.2.1). This means that any block B with the reading length at least m^3 must end before $(i-1)m - Lm^2$. This makes our q-reproduction procedure to work. We call a group of blocks with reading length at least m^3 a *big cluster* if the random walk can cross the group by generating only one big block in observations. Note that all big clusters of $\psi|[0,lm]$ are located in the pieces of ψ corresponding to the cells $\{1,\ldots,l\}\backslash\mathcal{I}(\psi) =: \mathcal{I}^c(\psi)$.

For each i we count all big clusters in $\psi|[0,im]$, for each $i=1,2,\dots,l$ and we compare them with the big clusters in $\chi_z|[0,im^2]$ for each i. Formally, e define the functions

$$q:\{0,1\}^{lm+1}\mapsto\mathbb{N}^l,\quad\text{and}\quad\hat{q}:\{0,1\}^{lm^2+1}\mapsto\mathbb{N}^l$$

as follows: $q(w)=(q_1(w),\dots,q_l(w))$, $\hat{q}(v)=(\hat{q}_1(v),\dots,\hat{q}_l(v))$ where

$$\begin{aligned} q_i(w) &:= \text{ number of big clusters contained in sub-vector } (w(0),\dots,w(im))\\ \hat{q}_i(v) &:= \text{ number of big blocks contained in sub-vector } (v(0),\dots,v(im^2)). \end{aligned}$$

As usually we define $q^*(w):=q(w^-)$ and $\hat{q}^*(v)=\hat{q}^*(v^-)$.
We denote

$$\hat{q}(v)\le q(w)\quad\big(\hat{q}^*(v)\le q^*(w)\big)\quad\text{ if and only if }\quad\hat{q}_i(v)\le q_i(w)\quad\big(\hat{q}_i^*(v)\le q_i^*(w)\big)\text{ for all } i.$$

Hence, if $\hat{q}(\chi_z|[0,lm^2])\le q(\psi|[0,ml])=:q(\psi)$, then for each i, the number of big blocks in $\chi_z|[0,im^2])$ is not bigger than the number of big clusters in $\psi|[0,mi]$. The foregoing argument shows that in case $z<y$, this implies that S_z is always "one cluster-end behind" implying $E_{z,\mathcal{I}(\psi)}$.
If $z=0$, then the observation word $\chi|[0,lm^2]$ will be not filtered out if, for each $i\in\mathcal{I}(\psi)$, the S moves from 0 to $(i-1)m$ generating as few big blocks in observations as possible. In Proposition 7.2.1 we show that this event has the probability bigger than

$$(p_{min})^{|\mathcal{I}^c(\psi)|m^2}.$$

This follows from the observation that this particular event restricts the behavior if S_y during its stay on the cells in $\mathcal{I}^c(\psi)$, only. The bound on the previous display is big enough to still have the bound like (7.1.13) (Theorem 7.2.3).

Final selection

We are now ready to define the final version of the selection rule.

* Note, for every $u\in\{0,1\}^{lm+1}$, $q(u)=(q_1,\dots,q_l)$ is vector, such that $q_i=\{0,1,\dots,l\}$, $q_1=0$ and $q_i\le q_{i+1}\le q_i+1$. Any such vector is called a q-**vector**. Hence, for every u, $q(u)$ and $q^*(u)$ are q-vectors.

Recall that, for any $u\in\{0,1\}^{lm+1}$, $g(u)=(g_1,\dots,q_l)$, where $g_i\in\{0,1\}^{n^2+1}$. Any such word is called a g-**word**. Hence, for each u, $g(u)$ and $g^*(u)$ are g-words.

In section 7.2.1 we shall give the formal definition of $\mathcal{I}_y(\xi)$ and $\mathcal{I}_x^*(\xi)$. When $B^1_{\text{intervals OK}}$ holds, then $|\mathcal{I}_y(\xi)|,|\mathcal{I}_x^*(\xi)|\ge(1-3\epsilon)l$ for each pair $x,y\in I_1$.

* We call (I^*,I,q^*,q,g^*,g) **a set of attributes**, if $I^*,I\subseteq\{1,\dots,l\}$, $|I^*|,|I|\ge l(1-3\epsilon(n))$, q,q^* are q-vectors and g^*,g are g-words.

Recall the definition if observation words $w^1(j),w^2(j),w^3(j)$, $j=1,\dots,\exp(\alpha l_1)$. For each set of attributes (I^*,I,q^*,q,g^*,g) we define the set $J(I^*,I,q^*,q,g^*,g)\subseteq[1,\exp(\alpha l_1)]$ as follows:

$j \in J(I^*, I, q^*, q, g^*, g)$ if and only if j satisfies

$$\hat{q}^*(w^1(j)) \le q^*, \quad \hat{g}^*(w^1(j)) \sqsubseteq_{I^*} g^*, \quad \hat{q}(w^3(j)) \le q, \quad \hat{g}(w^3(j)) \sqsubseteq_I g. \tag{7.1.18}$$

As described, the selection rule is based on g- and q-reproduction, and it consists of two parts – getting selected and avoiding non-ladder words. The principle of the final selection is exactly the same as the one of simplified selection described in Subsection 7.1.4.

With g- and q-reproduction, the getting selected part (**a)**) means that (with high probability) for each $x, y \in I_1$, $y - x = Lc_1l_1$ there exists a set of attributes (I^*, I, q^*, q, g^*, g) and at least $\exp(\gamma l_1)$ stopping times $\tau(j)$ with corresponding index set $J(x, y)$ such that for each $j \in J(x, y)$, (7.1.18) hold and the word $w^2(j)$ is the same, say w. Hence the first requirement of selection rule is to check whether there exists a set of attributes (I^*, I, q^*, q, g^*, g) such that $\exists J' \subseteq J(I^*, I, q^*, q, g^*, g)$ such that $|J'| \ge \exp(\gamma l_1)$ and $j \mapsto w^2(j)$ is constant on J'. The existence of such set of attributes and index-set J' can be easily checked.
The second requirement of the selection rule (**b)**) is avoiding the non-ladder words. We already know that if (x, y) form a barrier then (with high probability) the vectors $q^*_x(\xi)$, $q_y(\xi)$ and words $g^*_x(\xi)$ and $g_y(\xi)$ cannot be read somewhere else. Hence, if I^*, I, q^*, q, g^*, g found in the first step are indeed $\mathcal{I}^*_x(\xi), \mathcal{I}_y(\xi)$ $q^*_x(\xi)$, $q_y(\xi)$, $g^*_x(\xi)$, $g_y(\xi)$ as we want them to be, and if w is the word to be selected, then the following must hold: whenever there is a stopping time index j satisfying (7.1.18), then $w^2(j) = w$. Thus, the set J' must actually be $J(I^*, I, q^*, q, g^*, g)$.

We now give the formal definition of the selection rule.

Definition 7.1.1. *We define the set $\mathcal{W} = \mathcal{W}(\chi^{12\alpha l_1}, \tau)$ as follows. A word $w \in \{0,1\}^{c_1l_1+1}$ belongs to $\mathcal{W}$ if and only if there exists a complect of attributes (I^*, I, q^*, q, g^*, g) such that the following conditions are satisfied:*

a) $|J(I^*, I, q^*, q, g^*, g)| > \exp(\gamma l_1)$

b) *if $j \in J(I^*, I, q^*, q, g^*, g)$, then $w^2(j) = w$.*

7.2 Iteration

In this Section, we formalize g- and q-reproduction, described in Subsection 7.1.4. We begin with the definition of the OK-pieces of scenery, and we prove that a long piece of random scenery is typically OK (Theorem 7.2.1). In Subsection 7.2.2, we prove the inequality (7.1.15) (Theorem 7.2.2). In Subsection 7.2.3, we formalize q-reproduction and we found a suitable lower bound for (7.1.12) (Theorem 7.2.3). This is the main ingredient for obtaining the lower bound (7.1.13). Finally, in Subsection 7.2.4 we show how the barriers make the whole name-reading procedure to work.

Throughout the section, n, $m(n)$ and $l > 2Lm^2$ are fixed integer.

7.2.1 OK cells

In Theorem 7.1.2 we defined the set $B_{\text{cell\_OK}}(n) \in \sigma(\xi(z)|z \in [-cm, cm])$ that contains all typical pieces of sceneries in interval $[-cm, cm]$. In this definition, $c > 1$ is a fixed integer not depending on m. Thus, any word $w \in \{0,1\}^{2cm+1}$, regarded as a piece of scenery restricted to $[-cm, cm]$ either belongs to $B_{\text{cell\_OK}}(n)$ or not. We say that such a word w is **completely OK**, if $w \in B_{\text{cell\_OK}}(n)$.

***** Let $w := (w(1), \dots, w(N))$, $w(j) \in \{0,1\}$ be a word. Consider a sub-word w_a^{a+m} of w. We say that w_a^{a+m} is **weak-OK**, if $a - cm \geq 1$, $a + cm \leq N$ and the extension of w, w_{a-cm}^{a+cm} is completely OK.
Thus, any word of length m is weak-OK, if it is a certain sub-word of a larger word of length $2cm$ that is completely OK.

Let $\varphi \in \{0,1\}^I$ be a piece of scenery. Consider a subinterval $[a, a+m] \subseteq I$ such that $[a - cm, a + cm] \subseteq I$ We say that $\varphi|[a, a+m]$ is weak-OK, if $\varphi|[a - cm, a + cm]$ is completely OK. If φ is fixed we skip it from notation and we express the properties in terms of supports: we say that $[a, a+m]$ is weak-OK, if $[a-cm, a+cm]$ is completely OK.

***** Define integer intervals

$$D_i :=: [d_{i-1}, d_i] := (d_{i-1}, \dots, d_i), \quad \text{where} \quad d_i := im, \quad i = 1, 2 \dots .$$

Clearly D_i-s are not disjoint, $D_i \cap D_{i+1} = \{d_i\}$. It is also clear that $D_1 \cup \dots \cup D_l = [0, lm]$.

***** Consider the words $w \in \{0,1\}^{lm+1}$. For each such a word we define l sub-words, called **cells** $w_1, \dots, w_l$ as follows:

$$w_i \in \{0,1\}^{m+1}, w_i := w_{d_{i-1}}^{d_i} = (w(d_{i-1}), \dots, w(d_i)), \quad i = 1, \dots, l. \tag{7.2.1}$$

Hence, when speaking about a cell w_i, we always consider it as a sub-word of a longer word w with the length lm. Regarding w as a mapping, we equivalently define $w_i = w|D_i$.

***** Using the representation (7.2.1) we define the sets of indexes

$$\mathcal{I}_I(w) := \{i \in [2Lm^2, l] : w_i \text{ is weak-OK}\}.$$

Hence $\mathcal{I}_I(w)$ is a set of all indexes bigger than $2Lm^2$ such that w_i is weak-OK.

***** We say that binary word $w = (w(1), \dots, w(N))$ of length at least $N \geq m^{0.9}$ is **empty**, if there is no index j such that $w(j) = w(j+1) = \dots = w(j + m^{0.9})$. We say that a cell w_i has **empty neighborhood** if $d_i + Lm^2 \leq lm$, $d_{i-1} - Lm^2 \geq 0$ and $(w(d_{i-1} - Lm^2), \dots, w(d_i + Lm^2))$ is empty.

***** We say that a word $(w(1), \dots, w(N))$ **contains a fence** if $\exists 1 \leq i \leq N - 2L + 1$ such that

$$w(i) = \dots = w(i + L - 1) \neq w(i + L) = \dots = w(i + 2L - 1).$$

We say that a cell w_i in representation (7.2.1) is **isolated**, if $Lm + 2 \leq i \leq l - Lm - 1$ and both (sub-)words, $w_{i+Lm+1} = (w(d_i + Lm^2), \dots, w(d_i + Lm^2 + m))$ and $w_{i-(Lm+1)} = (w(d_{i-1} - Lm^2 - m), \dots, w(d_{i-1} - Lm^2))$ contain a fence.

In terms of pieces of sceneries.
Let $\varphi \in \{0,1\}^I$ be a piece of scenery. Consider a subinterval $[a, a+m] \subseteq I$. We say that $\varphi|[a, a+m]$ has empty neighborhood, if $\varphi|[a - Lm^2, a + m + Lm^2]$ is empty. When φ is fixed, we also say that $[a, a+m]$ has empty neighborhood (for φ).
We say that $\varphi|[a, a+m]$ is isolated, if $\varphi|[a - Lm^2 - m, a - Lm^2]$ and $\varphi|[a - Lm^2 + m, a + 2m + Lm^2]$ both contain a fence. For fixed φ, we say $[a, a+m]$ is isolated, if $[a - Lm^2 - m, a - Lm^2]$ and $[a - Lm^2 + m, a + 2m + Lm^2]$ both contain a fence (for φ).

***** Let w be as in (7.2.1). Define

$$\begin{aligned}
\mathcal{I}^1_{II}(w) &:= \{i \in [2Lm^2, l] : w_i \text{ is isolated}\} \\
\mathcal{I}^2_{II}(w) &:= \{i \in [2Lm^2, l] : w_i \text{ has empty neighborhood}\} \\
\mathcal{I}_{II}(w) &:= \mathcal{I}^1_{II}(w) \cap \mathcal{I}^2_{II}(w), \quad \mathcal{I}(w) := \mathcal{I}_I(w) \cap \mathcal{I}_{II}(w).
\end{aligned}$$

***** Let $\epsilon(n) := P(B_{\text{cell\_OK}}(n)^c) \vee \exp(-m^{0.7})$. We know, that $\epsilon(n) \to 0$. Consider a word $w \in \{0,1\}^{lm+1}$. We say that w is **OK** if

$$|\mathcal{I}_I(w)| \geq l(1 - 2\epsilon(n)) \quad \text{and} \quad |\mathcal{I}_{II}(w)| \geq l(1 - \exp(-m^{0.7})),$$

Recall the definition $\xi^{ml} := \xi|[0, lm]$ and let us define the events

$$\begin{aligned}
E_{\text{OK}} &:= \{\xi^{ml} \text{ is OK}\} \\
E_{\text{OK}a} &:= \left\{\left|\mathcal{I}_I(\xi^{ml})\right| \geq l(1 - 2\epsilon(n))\right\} \\
E_{\text{OK}b} &:= \left\{\left|\mathcal{I}_{II}(\xi^{ml})\right| \geq l(1 - \exp(-m^{0.7}))\right\}.
\end{aligned}$$

Clearly,

$$E_{\text{OK}} = E_{\text{OK}a} \cap E_{\text{OK}b} \tag{7.2.2}$$

and on E_{OK}

$$|\mathcal{I}(\xi^{ml})| \geq l(1 - 3\epsilon(n)). \tag{7.2.3}$$

The following theorem states that for n big enough, the probability of E^c_{OK} is exponentially decreasing in l. Hence, E_{OK} represents the typical behavior of ξ^{ml}. The proof is based on Höffding's inequalities and we leave it to Appendix.

Theorem 7.2.1. *There exists $N < \infty$ such that for each $n > N$ there exists $a(n) > 0$ not depending on l such that for all l big enough the event E_{OK} is independent on ξ^{Lm^3} and*

$$P(E_{\text{OK}}) \geq 1 - e^{-al}.$$

7.2.2 Iterated g-functions

Recall the function $g : \{0,1\}^{m+1} \mapsto \{0,1\}^{n^2+1}$ and $\hat{g} : \{0,1\}^{m^2+1} \mapsto \{0,1\}^{n^2}$ from Theorem 7.1.2. In the present section we extend these definitions to the sets $\{0,1\}^{lm}$ and $\{0,1\}^{lm^2+1}$.

***** Let $w \in \{0,1\}^{lm+1}$. Using the cell-representation (7.2.1) we extend the definition of g as follows

$$g : \{0,1\}^{lm+1} \mapsto \{0,1\}^{l(n^2+1)}, \quad g(w) := (g(w_1), g(w_2), \ldots, g(w_l)). \tag{7.2.4}$$

Note: by definition w_i and w_{i+1} are not disjoint - they have a common bit. However, by the definition, g does not depend on the first bit. Hence, applied on the scenery ξ^{ml}, the components $g_i(\xi^{ml})$ and $g_j(\xi^{ml})$ are independent.

* Define intervals

$$T_i =: [t_{i-1}, t_i] := (t_{i-1}, \ldots, t_i), \quad \text{where} \quad t_i := im^2, \quad , i = 1, 2 \ldots .$$

So, T_i-s are defined as D_i-s with m^2 instead of m.
Clearly T_i-s are not disjoint, $T_i \cap T_{i+1} = \{t_i\}$. It is also clear that $T_1 \cup \cdots \cup T_l = [0, lm^2]$.

* Consider words $v = (v(1), \ldots, v(lm)) \in \{0,1\}^{lm^2+1}$. For each such a word we define l sub-words, $v_1, \ldots, v_l$ as follows:

$$v_i \in \{0,1\}^{m+1}, v_i := v_{t_{i-1}}^{t_i} = (v(t_{i-1}), \ldots, v(t_i)), \quad i = 1, \ldots, l. \tag{7.2.5}$$

Regarding v as a mapping, we equivalently define $v_i = v|T_i$.
Using the sub-words (7.2.5) we define

$$\hat{g} : \{0,1\}^{lm+1} \mapsto \{0,1\}^{ln^2}, \quad \hat{g}(v) := (\hat{g}(v_1), \hat{g}(v_2), \ldots, \hat{g}(v_l)).$$

* Let $A = (a'_1, \ldots, a'_l)$, $B = (b'_1, \ldots, b'_l)$ with $a_i \in \{0,1\}^p$ and $b_i \in \{0,1\}^r$ $(p \le r)$ be lp and lr dimensional words, respectively. Let $I \subseteq \{1, 2, \ldots, l\}$. We define the following notation:

$$A \sqsubseteq_I B \quad \text{iff for each } i \in I \text{ it holds} \quad a'_i \sqsubseteq b'_i.$$

Recall the definition of $E_{z,I}$ in (7.1.14). The event $E_{z,I}$ states that for each $i \in I$, at time t_{i-1} the random walk S_z is further away than $L(m^2)$ from the point d_{i-1}. In that case, during the time interval T_i the random walk S_z can not visit the (location) set D_i. This, in turn, implies that the observation $\chi_z|T_i$ are independent of $\xi|D_i$. Then, obviously, $\hat{g}(\chi_z|T_i)$ is independent of $g(\xi|D_i)$.

The following theorem yields the bound (7.1.15).

Theorem 7.2.2. *There exists $\alpha_I(n) > 0$ not depending on l, such that for all $z < 0$ the following holds:*

$$P\left(\begin{array}{c} \exists I \subseteq \{1, 2, ..., l\} \text{ with } |I| = l(1 - 3\varepsilon(n)) \text{ such that} \\ E_{z,I} \text{ holds and } \hat{g}(\chi_z^{lm^2}) \sqsubseteq_I g\left(\xi^{ml}\right) \end{array} \right) \le e^{-\alpha_I l}, \tag{7.2.6}$$

provided l and n are both big enough.

Proof. Let $z < 0$. Denote $\xi_i = \xi|D_i$, $\chi_{z,i} := \chi_z|T_i$. Let Y_i, X_i $i = 1, \ldots, l$ be Bernoulli random variables, where

$$\begin{aligned} X_i = 1 \quad &\text{iff} \quad \hat{g}(\chi_{z,i}) \sqsubseteq g(\xi_i) \\ Y_i = 1 \quad &\text{iff} \quad S_z(t_{i-1}) < d_i - Lm^2. \end{aligned}$$

By definition, $g(\xi_i)$ is a $n^2 + 1$ dimensional random vector, with elements being Bernoulli iid with parameter $\frac{1}{2}$. For each fixed n^2-dimensional binary vector w we, therefore, get:

$$P(w \sqsubseteq g(\xi_i)) = (0.5)^{n^2-1} \tag{7.2.7}$$

Note, if $\{Y_i = 1\}$ holds, then $g(\xi_i)$ is independent of $\hat{g}(\chi_{z,i})$. By (7.2.7) then

$$P(X_i = 1|Y_i = 1) = P(\hat{g}(\chi_{z,i}) \sqsubseteq g(\xi_i)|Y_i = 1) = (0.5)^{n^2-1}.$$

Let $I \subseteq \{1, \ldots, l\}$. Consider the probability $P(X_i = 1, i \in I|Y_i = 1, i \in I)$. If $\{Y_i = 1, i \in I\}$ holds, then, $\{X_i, i \in I\}$ are iid random variables, with parameter $(0.5)^{n^2-1}$. Hence

$$P(X_i = 1, i \in I|Y_i = 1, i \in I) = (0.5)^{(n^2-1)|I|}.$$

Thus, for each $I \subseteq \{1, \ldots, l\}$ we have

$$P\Big(E_{z,I} \cap \{\hat{g}(\chi_z^{lm^2}) \sqsubseteq_I g(\xi^{ml})\}\Big) = E(\prod_{i\in I} X_i Y_i) = P(\prod_{i\in I} X_i Y_i = 1) =$$
$$P(X_i = 1, i \in I|Y_i = 1, i \in I)P(Y_i = 1, i \in I) \leq (0.5)^{(n^2-1)|I|} \tag{7.2.8}$$

Using (7.2.8), the probability in (7.2.6) can bound by

$$\sum_{\substack{I\subseteq\{1,2,\ldots,l\},\\ |I|=l(1-3\epsilon(n))}} P\Big(E_{z,I} \cap \big\{\hat{g}(\chi_z^{lm^2}) \sqsubseteq_I g(\xi^{ml})\big\}\Big) \leq \left(\begin{array}{c} l \\ 3l\varepsilon\,(n) \end{array}\right) \left(\frac{1}{2}\right)^{(n^2-1)l(1-3\epsilon(n))}. \tag{7.2.9}$$

Using Stirling's approximation, one can show that for l big enough

$$\left(\begin{array}{c} l \\ 3l\varepsilon\,(n) \end{array}\right) \leq \exp[-l\big((3\varepsilon\,(n)\ln(3\varepsilon\,(n)) + (1-3\varepsilon\,(n))\ln(1-3\varepsilon\,(n))\big)] = \exp(-l\epsilon_2(n)),$$

where $\epsilon_2(n) := 3\varepsilon(n)\ln(3\epsilon(n)) + (1-3\epsilon(n)) \to 0$, as n grows. Hence, if n is big enough, then the sum in (7.2.9) can bounded by

$$\exp(-l\epsilon_2(n))\Big((0.5)^{(n^2-1)l(1-3\epsilon(n))}\Big) \leq \exp(-ln^2\frac{\ln 2}{2}) = \exp(-l\alpha_I(n)),$$

where $\alpha_I(n) = n^2\frac{\ln 2}{2}$. □

7.2.3 Counting blocks

We now give formal definition of block.

***** Let $w = (w(u), \ldots, w(v))$ be a binary word. We say that w is a **block**, if

$$w(u) \neq w(u+1) = w(u+2) = \cdots = w(v-1) \neq w(v).$$

The **length of block** is defined as $v - u$. We call a block **big** if its length is $\geq m^3$. The $w(u)$ and $w(v)$ (or u and v) are the beginning of the block and the end of the block, respectively. The color $w(u+1)$ is called the **color of block**.

Let $\varphi \in \{0,1\}^I$ be a piece of scenery. Let $T = [t_1, t_2] \subseteq I$ be an integer interval of length at least 3. Since $\varphi|T$ can be considered as a binary word, the definition of block applies to $\varphi|T$ as well.
For given φ, we also call a location interval $T = [t_1, t_2]$ a block of φ, if $\varphi|T$ is a block (as word). So, in the following, a block can be a certain pattern (word) or a certain location

(T), where a string φ has a block.

For two blocks, $A = [a_1, a_2], B = [b_1, b_2]$ we denote $A < B$ if $a_1 < b_1$.

Note: although the block basically means many consecutive bits of the same color, by definition the first and last bit of a block must be different. For example, 01110 is a block with length 4, but 00001 is not a block.

* Let $[t_1, t_2] \in \mathbb{N}$ be a (time) interval. We call $R \in \mathbb{Z}^{[t_1,t_2]}$ an **admissible path** of length $t_2 - t_1$, if for all $t \in [t_1, t_2 - 1]$

$$P(S(1) - S(0) = R(t+1) - R(t)) > 0.$$

So, an admissible path is just a possible trajectory of S in time interval $[t_1, t_2]$, starting at $R(t_1)$ and ending at $R(t_2)$. The word "possible" means that the probability of such a trajectory is positive.
Let $\mathcal{R}(n)$ be the set of all admissible paths of length n. Thus

$$\mathcal{R}(n) := \Big\{ R \in \mathbb{Z}^{[0,n]} :,\ P(S(1) - S(0) = R(i+1) - R(i)) > 0,\ i = 0, \dots, n-1 \Big\}.$$

Let $B = [b_1, b_2] \subseteq \mathbb{Z}$ be a block of scenery φ. Define

$$l(B) := \min \left\{ n > 1 \,\middle|\, \begin{array}{c} \exists R \in \mathcal{R}(n) \quad \text{such} \quad \text{that} \ \ \varphi \circ R = \varphi(R(0)), \dots, \varphi(R(n)) \\ \text{is a block}, R(0) \le b_1, R(n) \ge b_2 \end{array} \right\}. \tag{7.2.10}$$

The number $l(B)$ will be called as **the reading-length** of B.
Suppose $l(B) = n$ and $R(0), \dots, R(n)$ is the admissible path that attains the minimum in 7.2.10. Then the points $R(0)$ and $R(n)$ are called **the reading-beginning** and **the reading-end** of B, respectively.
The reading length of a block is, the length of the smallest block in observations, generated under conditions that S crosses B. Clearly, $l(B)$ is approximately $\frac{b_2 - b_1}{L}$, but it depends also on the φ outside the block B. Let us consider some examples.

Examples: 1. If S is a simple random walk (i.e. $L = 1$), then $l(B) = b_2 - b_1$ and reading beginning (reading end) and the beginning (the end) of the block coincide.

2. Let $L = 3$. Consider the word $(w(1), w(2), \dots, w(11)) = 00111111000$. This word contains a block with the length 7. The reading length of this block is, obviously, 3. The beginning of the block is $w(2)$, the end of the block is $w(9)$. The reading beginning is $w(2)$ or $w(1)$ with the reading ends $w(11)$ or $w(10)$, respectively.

3. Let $L = 3$. Consider the word $(w(1), w(2), \dots, w(11)) = 0011111111000$. It contains a block of length 9, the reading length of the block is 3, the reading beginning of the block is $w(2)$, the reading end of the block is $w(11)$.

4. Suppose $L = 4$ and $P(S(1) - S(0) = 2) = P(S(1) - S(0) = 3) = 0$. Consider the word $w(1), \dots, w(18) = 011101111111110111$. This word contains a block of length 10 $B = (w(5), \dots, w(15))$. The reading length of this block is 5.

5. Change the word without changing the block and consider the word 1110111111111000. The reading length of B is now 3, the reading-beginning is $w(4)$, the reading-end is $w(16)$.

6. Consider now the words as in the last 2 examples. Suppose $P(S(1)-S(0)=i)>0$, $i=-4,-3,\dots,3,4$. Then the block has reading length 3 no matter what the neighborhood of the block is.

***** Let $A=[a_1,a_2]$, $B=[b_1,b_2]$ be two blocks of ψ, $A<B$. We say that A and B are **connected** if they are of same color, say 1, and there is an admissible path from A to B such that moving along this path, only the color 1 is observed. Formally, A and B is connected, if there exists an n and $R\in\mathcal{R}(n)$ such that $R(0)\in(a_1,a_2), R(n)\in(b_1,b_2)$ and $\psi\circ R(0)=\psi\circ R(1)=\dots=\psi\circ R(n)$.
In other words, the blocks of the same color are connected, if it is possible to read them as one block.

Let $B_1<B_2<\ldots<B_h$ be blocks of ψ. We say that $B_1\cup\ldots\cup B_h$ is a **big cluster**, if

- B_i has the reading length at least m^3, $i=1,\dots,h$;
- $B_1,\dots,B_h$ are connected;
- there is no more blocks with the reading length at least m^3 connected to B_1.

We define the reading-path of a big cluster in the same way as the reading path of a block (which can be a big cluster consisting of one block) – this is the shortest admissible path to cross the big bluster and producing exactly one block. Formally, for a big cluster $C:=B_1\cup\ldots\cup B_h$ we define the reading length of the big cluster as

$$l(C):=\min\{n>1:\exists R\in\mathcal{R}(n) \text{ such that } \quad \psi(R(0)),\dots,\psi(R(n)) \text{ is a block}, R(0)\le c, R(n)\ge d\},$$

where c is the beginning of B_1 and d is the end of B_h. These points are referred to as the beginning and the end of C, respectively. Clearly, $l(C)\ge m^3$. The reading-path of C is any path that attains the minimum above.

***** Let us fix $\psi\in E_{\rm OK}$. Denote $\mathcal{I}=\mathcal{I}(\psi^{ml})$, $\mathcal{I}_I=\mathcal{I}_I(\psi^{ml})$, $\mathcal{I}_{II}=\mathcal{I}_{II}(\psi^{ml})$.
Consider the set $\mathcal{I}_{II}^c:=[1,l]-\mathcal{I}_{II}$. Clearly $\mathcal{I}_{II}^c$ is an union of disjoint intervals, i.e.

$$\mathcal{I}_{II}^c=[l_1,l_2]\cup[l_3,l_4]\cup\dots\cup[l_{2k-1},l_{2k}], \tag{7.2.11}$$

where $l_1=1,l_2,l_3,\ldots\in[2Lm^2,l]$, $l_j\le l_{j+1}$.
The set of cell-indexes $[l_{2j-1},l_{2j}]$ corresponds to the location-interval (cells) $[(l_{2j-1}-1)m,l_{2j}m]$ or $[d_{l_{2j-1}-1},d_{l_{2j}}]$. Let us denote

$$r_j:=(l_{2j-1}-1)m,\quad s_j=l_{2j}m,\quad j=1,\dots,k. \tag{7.2.12}$$

By definition, S visits every point in $\mathbb{Z}$ i.o.. This means, there exists an integer $k\ge1$ such that $P(S(k)-S(0)=1)$. Let $\bar v:=\inf\{k:P(S(k)-S(0)=1)>0\}$. Thus there is an admissible path $R(0),\dots,R(\bar v)$ such that $R(0)=0$ and $R(\bar v)=1$. Similarly, between points $a<b$ there exists an admissible path $R(0),\dots,R((b-a)\bar v)$ such that $R(0)=a$, $R(\bar v)=a+1$, $R(2\bar v)=a+2,\dots R((b-a)\bar v)=b$. We say that S moves *stepwise* from a to b, if it moves along the path just described. Obviously, $\bar v\ll m$.

In Subsection 7.1.4, we defined big cluster counter $q : \{0,1\}^{lm+1} \mapsto \mathbb{N}^l$ and block counter $\hat{q} : \{0,1\}^{lm^2+1} \mapsto \mathbb{N}^l$.
Define the events

$$F_{min}(1) := \big\{\hat{q}(\chi|[0, s_1 m]) \le q(\psi|[0, s_1]),\ \chi|[s_1 - m\bar{v}, s_1] \text{ contains both colors}, \quad S(s_1 m) = s_1\big\}.$$
$$F_{min}(j) := \big\{\hat{q}(\chi_{r_j}|[0, (s_j - r_j)m]) \le q(\psi|[r_j, s_j]),\ \chi_{r_j}|[0, m\bar{v}] \text{ and } \chi_{r_j}|[(s_j - r_j)m - m\bar{v}, (s_j - r_j)m] \text{ contain both colors } S_{r_j}((s_j - r_j)m) = s_j\big\},\ j = 2, \dots, k-1.$$

For the last interval in (7.2.12) we define $F(k)$ as $F(j)$, $j > 1$, if $s_k < l$. If $r_k = l$, we define

$$F_{min}(k) := \big\{\hat{q}(\chi_{r_k}|[0, (l - r_k)m]) \le q(\psi|[r_k, l]),\ \chi_{r_k}|[0, m\bar{v}] \text{ contains both colors},\ S_{r_k}((l - r_k)m) = l\big\}.$$

Obviously, the events $F_{min}(j)$ depend on the random walk, S, only. Moreover, by definition, the event $F_{min}(j)$ depends on the behavior of the random walk during the time interval $[0, (s_j - r_j)m]$. This means, if for a j, there exists at least one admissible path $R_j \subseteq \mathcal{R}((s_j - r_j)m)$ such that

R1 $R_j(0) = r_j,\ R_j((s_j - r_j)m) = s_j,$

R2 $\hat{q}(\psi \circ R_j) \le q(\psi[s_j, r_j])$

R3 if $r_j \neq 0$ and $s_j \neq l$ then $(\psi \circ R_j)|[0, m\bar{v}]$ and $(\psi \circ R_j)|[(s_j - r_j) - m\bar{v}, (s_j - r_j)]$ have both colors,

then $F_{min}(j) \neq \varnothing$ and $P_\psi\big(F_{min}(j)\big) \ge (p_{min})^{(s_j - r_j)m}$. The following proposition, proved in Appendix, shows that for each j, at least one such admissible path exists.

Proposition 7.2.1. *For each $j = 1, \dots k$ the following holds:*

$$P_\psi\big(F_{min}(j)\big) \ge (p_{min})^{(s_j - r_j)m} = (p_{min})^{(l_{2j} - l_{2j-1} + 1)m^2}. \tag{7.2.13}$$

The next theorem is the main ingredient of the "getting selected" part of the reconstruction. It gives a lower bound for the probability that g- and q-reproduction to work.

Theorem 7.2.3. *There exist constant $\alpha_{II}(n) > 0$ not depending on l, such that for all $\psi \in E_{\rm OK}$ the following holds:*

$$P_\psi\left(\hat{g}(\chi^{lm^2}) \sqsubseteq_{\mathcal{I}} g(\psi^{ml}), \quad \hat{q}(\chi^{m^2 l}) \le q(\psi^{ml}), \quad S(m^2 l) = ml\right) \ge e^{-l\alpha_{II}}. \tag{7.2.14}$$

Proof. For each $i \in [1, l]$ and subset $I \subseteq [1, l]$, we define the events

$$E_S(i) := \{S(t_{i-1}) - S(t_i) = m\}, \quad E_S(I) := \cap_{i \in I} E_S(i)$$
$$E_\sqsubseteq(i) := \{\hat{g}(\chi|T_i) \sqsubseteq g(\psi|D_i)\}, \quad E_\sqsubseteq(I) := \cap_{i \in I} E_\sqsubseteq(i);$$
$$E_{no-block}(i) := \{\text{ the sequence } \chi|T_i \text{ contains both colors }\}, \quad E_{no-block}(I) := \cap_{i \in I} E_{no-block}(i).$$

Use $[r_j, s_j]$, $j = 1, \dots, k$ as in (7.2.12) to define

$$\begin{aligned} E_{min}(1) &:= \{\hat{q}(\chi|[0, s_1 m]) \le q(\psi|[0, s_1]), \quad S(s_1 m) = s_1, \\ &\qquad \chi|[s_1 - m\bar{v}, s_1] \text{ contain both colors}\} \\ E_{min}(j) &:= \{\hat{q}(\chi|[r_j m, s_j m]) \le q(\psi|[r_j, s_j]), \quad S(s_j m) = s_j, \\ &\qquad \chi|[r_j, r_j + m\bar{v}] \text{ contain both colors}, \ \chi|[s_j - m\bar{v}, s_j] \text{ contain both colors}\}, \\ &\qquad j = 2, \dots, k \quad \text{and} \\ E_{min} &:= \cap_{j=1}^{k} E_{min}(j). \end{aligned}$$

If $s_k = l$, then the requirement $\{\chi|[s_k - m\bar{v}, s_k] \text{ contain both colors}\}$ is dropped for the definition of $E_{min}(k)$.

Consider the event $E_{min} \cap E_S(\mathcal{I}_{II})$. Use Proposition 7.2.1 to get

$$\begin{aligned} P_\psi(E_{min} \cap E_S(\mathcal{I}_{II})) &= P_\psi(E_{min}(1) \cap E_S(\mathcal{I}_{II})) \prod_{j=2}^{k} P_\psi(E_{min}(j) | E_{min}(j-1) \cap \dots \cap E_{min}(1) \cap E_S(\mathcal{I}_{II})) \\ &= P(E_S(\mathcal{I}_{II})) P_\psi(E_{min}(1)) \prod_{j=2}^{k} P_\psi(E_{min}(j) | S(m r_j) = r_j) \\ &= \prod_{j=1}^{k} P_\psi(F_{min}(j)) P(E_S(\mathcal{I}_{II})) \ge (p_{min})^{|\mathcal{I}_{II}^c| m^2} P_\psi(E_S(\mathcal{I}_{II})). \end{aligned} \tag{7.2.15}$$

Note: if $E_{min} \cap E_S(\mathcal{I}_{II})$ holds, then, for each $i \in \mathcal{I}_{II}$ we have

$$S(t_{i-1}) = d_{i-1}, \quad S(t_i) = d_i.$$

Hence, given $E_{min} \cap E_S(\mathcal{I}_{II})$, the behavior of S during T_i is independent of the behavior of S outside T_i. In particular, for each $i \in \mathcal{I}_{II}$

$$\begin{aligned} &P_\psi\big(E^c_{no-block}(i) | E_{min} \cap E_S(\mathcal{I}_{II})\big) = \\ &P_\psi\big(E^c_{no-block}(i) | S(t_{i-1}) = d_{i-1}, S(t_i) = d_i\big) = P_\psi\big(E^c_{no-block}(i) | E_S([1, l])\big). \end{aligned} \tag{7.2.16}$$

Let us estimate (7.2.16). If $S(t_{i-1}) = d_{i-1}$ and $S(t_i) = d_i$, then during T_i, the random walk stays in the Lm^2-neighborhood of D_i. But $\psi|D_i$ is isolated and has empty neighborhood. Thus, during T_i, the random walk stays on the area where is no $m^{0.9}$ consecutive colors. In this case, the probability of generating a block of length at least m^2 is, for big m, bounded above by $\exp(-\frac{am^2}{m^{1.8}}) = \exp(-am^{0.2})$, where $a > 0$ is a constant that does not depend on m (see, e.g. Lemma 2.1 in [15]).
Denote

$$p_m := P(S(m^2) = m).$$

Then

$$P(E_S([1, l])) = (p_m)^l. \tag{7.2.17}$$

So, for each $i \in \mathcal{I}_{II}$, it holds

$$P_\psi\Big(E^c_{no-block}(i) \Big| E_{min} \cap E_S(\mathcal{I}_{II})\Big) = P_\psi\Big(E^c_{no-block}(i) \Big| E_S([1, l])\Big) \le \frac{\exp(-am^{0.2})}{(p_m)^l}.$$

Now, by local central limit theorem, p_m is of order $\frac{1}{m}$. Thus, when m is big enough

$$P_\psi(E_{no-block}(i)|E_{min}\cap E_S(\mathcal{I}_{II})) > 0.75, \quad P_\psi(E_{no-block}(\mathcal{I}_{II}\backslash\mathcal{I})|E_{min}\cap E_S(\mathcal{I}_{II})) > (0.75)^{|\mathcal{I}_{II}|-|\mathcal{I}|}. \tag{7.2.18}$$

The second inequality holds because given $E_{min} \cap E_S(\mathcal{I}_{II})$, the events $E_{no-block}(i)$ and $E_{no-block}(j)$ are conditionally independent, provided $j, i \in \mathcal{I}_{II}$.

Suppose now $i \in \mathcal{I} \subseteq \mathcal{I}_{II}$. Then $\psi|D_i$ is weak-OK. By **2)** of Theorem 7.1.2 we now get that

$$P_\psi(E^c_\sqsubseteq(i)|E_{min}\cap E_S(\mathcal{I}_{II})) = P_\psi(E^c_\sqsubseteq(i)|S(t_{i-1}) = d_{i-1}, S(t_i) = d_i) = P_\psi(\hat{g}(\chi|[0, m^2]) \sqsubseteq g(\psi_0^m)) \leq 0.25.$$

This also means that, with $i \in \mathcal{I}$

$$P_\psi\Big(\big(E_{no-block}(i) \cap E_\sqsubseteq(i)\big)^c \Big| E_{min} \cap E_S(\mathcal{I}_{II})\Big) \leq$$
$$P_\psi\Big(E^c_{no-block}(i)\Big|E_{min} \cap E_S(\mathcal{I}_{II})\Big) + P_\psi\Big(E^c_\sqsubseteq(i)\Big|E_{min} \cap E_S(\mathcal{I}_{II})\Big) < 0.5.$$

And, by independence, again

$$P_\psi\Big(E_{no-block}(\mathcal{I}) \cap E_\sqsubseteq(\mathcal{I})\Big|E_{min} \cap E_S(\mathcal{I}_{II})\Big) > (0.5)^{|\mathcal{I}|}. \tag{7.2.19}$$

Finally, by the same independence-argument, (7.2.19) and (7.2.18),

$$P_\psi\Big(E_{no-block}(\mathcal{I}_{II}) \cap E_\sqsubseteq(\mathcal{I})\Big|E_{min} \cap E_S(\mathcal{I}_{II})\Big) =$$
$$P_\psi\Big((E_{no-block}(\mathcal{I}) \cap E_\sqsubseteq(\mathcal{I})) \cap E_{no-block}(\mathcal{I}_{II}\backslash\mathcal{I})\Big|E_{min} \cap E_S(\mathcal{I}_{II})\Big) > (0.5)^l \tag{7.2.20}$$

Consider $[r_j, s_j]$, $j = 1, \ldots, k$ as in (7.2.12). By the definition of $\mathcal{I}_{II}$, $[s_j - Lm^2, r_{j+1} + Lm^2]$ is empty, for each $j = 1, \ldots, k-1$ as well as for $[s_k - Lm^2, l]$, if $s_k < l$. This implies that these intervals do not contain any small block (and, therefore, no big clusters). Also $[s_j - Lm^2 - m, s_j - Lm^2]$ as well as $[r_{j+1} + Lm^2, r_{j+1} + Lm^2 + m]$ $(j = 1, \ldots, k-1)$ and $[s_k - Lm^2 - m, s_j - Lm^2 - m]$, if $s_k < l$, contain a fence. This means that a interval $[s_j - Lm^2, r_{j+1} + Lm^2]$ $(j = 1, \ldots, k-1)$ as well as $[s_k - Lm^2 - m, s_j - Lm^2 - m]$ (if $s_k < l$) is not inside a big cluster (without fences this could be a case even if the interval is empty). The emptiness and the isolation of $[s_j, r_j]$ imply that the cluster-counting vector $q(\psi^{ml})$ is constant on $\mathcal{I}_{II}$.
The event $E_{no-block}(\mathcal{I}_{II})\cap E_{min}$ ensures that the word $\chi|[s_j - m\bar{v}, r_{j+1} + m\bar{v}]$, $j = 1, \ldots, k-1$ does not contain more than $m\bar{v} + m^2$ consecutive colors. The same is true for the word $\chi|[s_k - m\bar{v}, l]$. The event E_{min} also guarantees that all big blocks in observations end before time interval T_i, $i \in \mathcal{I}_{II}$. Hence, the block-counting vector $\hat{q}(\chi^{m^2 l})$ is constant on $\mathcal{I}_{II}$. Thus, $\hat{q}(\chi^{t_l}) \leq q(\psi^{t_l})$ if $\hat{q}_i(\chi^{t_l}) \leq q_i(\psi^{t_l})$ for each $i \in \mathcal{I}^c_{II}$. The latter holds if and only if $\hat{q}(\chi|[r_j m, s_j m]) \leq q(\psi|[r_j, s_j])$ for each $j = 1, \ldots, k$. Hence

$$E_{min} \cap E_{no-block}(\mathcal{I}_{II}) \subseteq \{\hat{q}(\chi^{m^2 l}) \leq q(\psi^{ml})\}.$$

This means

$$E_{min}\cap E_{no-block}(\mathcal{I}_{II})\cap E_\sqsubseteq(\mathcal{I})\cap E_S(\mathcal{I}_{II}) \subseteq \Big\{\hat{g}(\chi^{lm^2}) \sqsubseteq_{\mathcal{I}} g(\psi^{ml}), \quad \hat{q}(\chi^{m^2 l}) \leq q(\psi^{ml}), \quad S(m^2 l) = ml\Big\}. \tag{7.2.21}$$

From (7.2.20), (7.2.17) and (7.2.15) it follows

$$P_\psi\Big(E_{min}\cap E_{no-block}(\mathcal{I}_{II})\cap E_{\sqsubseteq}(\mathcal{I})\cap E_S(\mathcal{I}_{II})\Big)=$$
$$P_\psi\Big(E_{\sqsubseteq}(\mathcal{I})\cap E_{no-block}(\mathcal{I}_{II})\cap\Big|E_{min}\cap E_S(\mathcal{I}_{II})\Big)P_\psi\Big(E_{min}\cap E_S(\mathcal{I}_{II})\Big)>$$
$$(0.5)^l P_\psi\Big(E_{min}\cap E_S(\mathcal{I}_{II})\Big)\geq(0.5)^l(p_{min})^{|\mathcal{I}_{II}^c|m^2}P_\psi(E_S(\mathcal{I}_{II}))\geq$$
$$(0.5)^l(p_{min})^{|\mathcal{I}_{II}^c|m^2}(p_m)^l. \tag{7.2.22}$$

Hence (7.2.21), (7.2.22) and the inequality $|\mathcal{I}_{II}^c|\leq l\exp(-m^{0.7})$ imply

$$P_\psi\Big(\hat{g}(\chi^{lm^2})\sqsubseteq_{\mathcal{I}} g(\psi^{ml}),\quad \hat{q}(\chi^{m^2l})\leq q(\psi^{ml}),\quad S(m^2l)=ml\Big)\geq(0.5)^l(p_{min})^{|\mathcal{I}_{II}^c|m^2}(p_m)^l\geq$$
$$[0.5p_m(p_{min})^{m^2\exp(-m^{0.7})}]^l=\exp[l(\ln(0.5p_m)+m^2\exp(-m^{0.7})\ln(p_{min}))]=\exp[-l\alpha_{II}(m)].$$

□

Let us show that, for n big enough,

$$8\alpha_{II}(n)=-8\ln(0.5p_m)-m(n)^2\exp(-m(n)^{0.7})\ln(p_{min})<n^2\frac{\ln 2}{2}=\alpha_I(n) \tag{7.2.23}$$

By the LCLT, p_m is of order $\frac{1}{m}$, meaning that $-\ln(0.5p_m)$ is of order $\ln 2m$. On the other hand, $m(n)<\exp(2n)$ ([15], (3.10)), implying that $-\ln(0.5p_m)$ is of order n. The expression

$$-m(n)^2\exp(-m(n)^{0.7})\ln(p_{min})$$

is negligible in comparison with $-\ln(0.5p_m)$. So, if n is big enough, it holds $\alpha_{II}(n)<Kn$, for some $K<\infty$. Since $\alpha_{II}(n)$ is of order n^2, for big n, the inequality (7.2.23) clearly holds.

7.2.4 Block at origin

Define the event

$$E_{\text{origin}}:=\{\xi(-L)=\cdots=\xi(-1)\neq\xi(0)=\cdots=\xi(m^3L)\neq\xi(m^3L{+}1)=\cdots=\xi(m^3L{+}L)\}.$$

The reason of block-counting is the following observation. Recall the definition of $E_{z,I}$ given in (7.1.14). The next theorem formalizes the argument explained in Subsection 7.1.4.

Theorem 7.2.4. *If $z<0$ then*

$$E_{origin}\cap\{\hat{q}(\chi_z^{t_l})\leq q(\xi^{ml})\}\subseteq E_{z,\mathcal{I}(\xi^{ml})}. \tag{7.2.24}$$

Proof. Let $\xi=\psi$, $\mathcal{I}=\mathcal{I}(\psi)$. Let $i\in\mathcal{I}$. The interval D_i is isolated and, hence, D_i is not included into any big cluster of ψ, i.e. $q_i(\psi^{d_l})=q_i(\psi^{d_i})$. The interval D_i has empty neighborhood, which together with the isolation implies that the number of big clusters in $[0,d_i]$ is the same as the number of big clusters in $[0,d_i-Lm^2-m)=[0,d_{i-1-Lm})$ or, equivalently,

$$q_i(\psi^{ml})=q_{i-1-Lm}(\psi^{ml}). \tag{7.2.25}$$

Let $z < 0$. By crossing an interval, the random walk cannot produce less big blocks than the number of big clusters in this interval. Hence, the number of big blocks in observations generated by S_z by crossing the interval $[z, d_{i-1-Lm}]$ is at least the number of big clusters in $[z, d_{i-1-Lm}]$. Suppose now that E_{origin} holds. Then the interval $[z, d_{i-1-Lm}]$ contains strictly more big clusters than the interval $[0, d_{i-1-Lm}]$. Therefore, the number of big blocks in observations generated by S_z by crossing the interval $[z, d_{i-1-Lm}]$ is strictly bigger than the number of big clusters in $\psi|[0, d_{i-1-Lm}]$. By (7.2.25), this number equals $q_i(\psi^{ml})$. Hence, if $S_z(t_i) \geq d_{i-1-Lm}$, then $\hat{q}_i(\chi_z^{m^2 l}) > q_i(\psi^{ml})$. Consequently, $E_{\text{origin}} \cap E^c_{z,\mathcal{I}} \subseteq E_{\text{origin}} \cap \{\hat{q}(\chi_z^{t_l}) \leq q(\xi^{ml})\}^c$. This proves the statement. □

Define

$$E_{\text{mistake}}(z) := \left\{\hat{q}(\chi_z^{m^2 l}) \leq q(\xi^{ml})\right\} \cap \left\{\hat{g}(\chi_z^{m^2 l}) \sqsubseteq_{\mathcal{I}(\xi^{ml})} g(\xi^{ml})\right\} \cap E_{\text{origin}}.$$

Corollary 7.2.1. *If $z < 0$, then for n and l big enough*

$$P\big(E_{mistake}(z) \cap E_{OK}\big) \leq \exp(-\alpha_I l). \tag{7.2.26}$$

Proof. By (12.3.8) we have

$$E_{\text{mistake}}(z) \subseteq E_{z,\mathcal{I}(\xi^{ml})} \cap \left\{\hat{g}(\chi_z^{m^2 l}) \sqsubseteq_{\mathcal{I}(\xi^{ml})} g(\xi^{ml})\right\}.$$

Thus

$$E_{\text{mistake}}(z) \cap E_{OK} \subseteq E_{z,\mathcal{I}(\xi^{ml})} \cap \left\{\hat{g}(\chi_z^{m^2 l}) \sqsubseteq_{\mathcal{I}(\xi^{ml})} g(\xi^{ml})\right\} \cap E_{OK}. \tag{7.2.27}$$

Consider the right side of (7.2.27). By E_{OK} and (7.2.3), $|\mathcal{I}(\xi^{ml})| \geq l(1 - 3\epsilon(n))|$. Thus, if the right side of (7.2.27) holds, then there exists a subset $I \subseteq \mathcal{I}(\xi^{ml})$ such that $|I| = |l(1 - 3\epsilon(n))|$, $\left\{\hat{g}(\chi_z^{m^2 l}) \sqsubseteq_I g(\xi^{ml})\right\}$ and $E_{z,I}$ holds. By Theorem 7.2.2, this event has probability not bigger than $\exp(-l\alpha_I)$. □

7.3 Reconstruction at level l_1

In this chapter we prove the main result, Theorem 7.1.1. We start with the formal definitions of many events and notions that were already introduced in Subsection 7.1.4. Using the defined events, the proof of Theorem 7.1.1 can be splitten into two parts: the event-combinatorial part (Subsection 12.3) and the probability estimation part (Subsection 7.3.5). These two parts together establish the proof of Theorem 7.1.1 (Subsection 12.4).

From now on $l := l_1 \cdot l_2$, where l_2 as well as $m(n)$, c_1, α and γ are defined in Subsection 7.3.6. The only variable is l_1. Hence, the statement "l big enough" in the previous chapter must be interpreted as "l_1 big enough".

7.3.1 Some definitions

* A vector $I \in \mathbb{Z}^{[0,n]}$ is **ladder interval** of length n, if $I = (a, a+L, a+2L, \dots, a+nL)$ for some $a \in \mathbb{Z}$. Let $\mathcal{L}(n)$ be the set of all ladder intervals of length n.

Let I be a ladder interval. A piece of scenery $\varphi \in \{0,1\}^I$ is called a **ladder piece**. If $\varphi \in \{0,1\}^D$, $I \subseteq D$ is a ladder interval, we sometimes say that $\varphi|I$ is a ladder piece of φ (or $\varphi|D$).

Hence, a ladder piece of a non-random scenery ψ is any vector $(\psi(a), \psi(a+L), \dots, \psi(a+nL))$, $a \in \mathbb{Z}$, $n \in \mathbb{N}$.

Let $I = (a, a+L, \dots, a+nL)$ be a ladder-interval and let $\varphi \in \{0,1\}^I$ be a ladder piece. We write $\varphi \approx_l w$, if $\varphi(a) = w(1), \dots, \varphi(a+Ln) = w(n+1)$ or $\varphi(a) = w(n+1), \dots, \varphi(a+Ln) = w(1)$. Hence, if $L = 1$, then the relation " $\approx_l$ " is the same as the equivalence " $\approx$ ".

Given a ladder piece $\varphi \in \{0,1\}^I$, $I \in \mathcal{L}(n)$, we say that $w \in \{0,1\}^{n+1}$ is a **ladder word** of φ, if $\varphi \approx_l w$. Hence, any ladder piece has at most two ladder words that are equivalent. Also note that two ladder pieces are equivalent, if and only if their ladder words coincide. (In the notation of [20]), w is a ladder word of φ, if $w \in \{(\varphi)_\rightarrow, (\varphi)_\leftarrow\}$.)

* Recall that $I_1 := [-\exp(3l_1), \exp(3l_1)]$. The following event, $B^1_{\text{unique fit}}$, states that any ladder piece of $\xi|I_1$ of length $\frac{l_1c_1}{4}$ has unique ladder word up to equivalence. Formally,

$$B^1_{\text{unique fit}} := \Big\{\text{if } I, J \in \mathcal{L}(l_1c_1/4), I, J \subseteq I_1 \text{ and } I \neq J \text{ then } \xi|I \not\approx \xi|J\Big\}.$$

* Suppose $x, y \in \mathbb{Z}$, $y = x + (l_1c_1)L$. In this case there is only one admissible path of length c_1l_1 from x to y, i.e. there exist unique $R \in \mathcal{R}(l_1c_1)$ such that $R(c_1l_1) - R(0) = (l_1c_1)L$. Obviously, this path consists of maximum jumps, only, i.e. $R(i+1) - R(i) = L$, $i = 0, 1, \dots, l_1c_1 - 1$.
Suppose now that $x, y \in \mathbb{Z}$, $x < y$ are such that $y < x + (l_1c_1)L$. In this case, this might happen that there is no admissible path going from x to y with exactly l_1c_1 steps. However, if there is one such admissible path, then it is clearly not unique. The following event, $B^1_{\text{recon straight}}$, states that if $x, y \in I_1$, then among these admissible paths, there are at least two that generate different words in the observations. More precisely,

$$B^1_{\text{recon straight}} := \left\{ \begin{array}{c} \text{if } R \in \mathcal{R}(l_1c_1) \text{ such that } R(0), R(l_1c_1) \in I_1 \text{ and } R(l_1c_1) - R(0) < (l_1c_1)L, \text{ then} \\ \exists R' \in \mathcal{R}(l_1c_1) \text{ such that } R(0) = R'(0), R(c_1l_1) = R'(c_1l_1) \text{ and } \xi \circ R \neq \xi \circ R' \end{array} \right\}.$$

* Let ψ be a scenery. We say that $x \in \mathbb{Z}$ is a **left-barrier point** of ψ, if x satisfies (7.1.17). We say that $y \in \mathbb{Z}$ is a **right-barrier point** of ψ, if y satisfies (7.1.16). The pair (x, y) is called a **barrier** of ψ, if x is a left- and y is a right-barrier point. Recall the event E_{origin}. The point y is a right-barrier point of ψ, if the translated scenery $(\psi(i+y))_{i\in\mathbb{Z}}$ belongs to the event E_{origin}. Similarly, x is a left-barrier point, if the translated and reflected scenery $(\psi(x-i))_{i\in\mathbb{Z}}$ belongs to the event E_{origin}.

We consider the barriers of ξ, (x, y) such that $y - x = (c_1l_1)L$. In order to carry on the reconstruction in level l_1, every interval $[z, z + (c_1l_1/4)L]$, $z \in I_1$ should contain enough

left-barrier points of such barriers. This is the meaning of the event $B^1_{\text{enough barriers}}$. More precisely,

$$B^1_{\text{enough barriers}} := \left\{ \begin{array}{c} \text{for any } j=0,\dots,L-1 \text{ and for any } z\in I_1, \\ \text{there exists } x\in[z,z+(c_1l_1/4)L] \text{ such that:} \\ x \mathrm{mod} L = j \text{ and } (x,x+(c_1l_1)L) \text{ is a barrier of } \xi \end{array} \right\}.$$

***** We now define the left-side counterparts of $g, \hat{g}$, q and $\hat{q}$. For a word $u=(u_1,\dots,u_n)$ denote by u^- its reflection, i.e. $u^- := (u_n,\dots,u_1)$. Now let

$$q^* : \{0,1\}^{lm+1} \mapsto \mathbb{N}^l, \quad \hat{q}^* : \{0,1\}^{lm^2+1} \mapsto \mathbb{N}^l, \quad g^* : \{0,1\}^{lm+1} \mapsto \{0,1\}^{ln^2+1}.$$

and

$$\hat{g}^* : \{0,1\}^{lm^2+1} \mapsto \{0,1\}^{ln^2}$$

be as follows

$$q^*(w) = q(w^-), \quad g^*(w) = g(w^-), \quad w\in\{0,1\}^{lm+1} \tag{7.3.1}$$

$$\hat{q}^*(v) = \hat{q}(v^-), \quad \hat{g}^*(v) = \hat{g}(v^-), \quad v\in\{0,1\}^{lm^2+1}. \tag{7.3.2}$$

In Section 7.1.4, we already introduced the following notation:

$$\begin{aligned} q_y(\xi) &:= q(\xi|[y,y+ml]), \quad g_y(\xi) := g(\xi|[y,y+ml]) \\ q^*_x(\xi) &:= q^*(\xi|[x-ml,x]), \quad g^*_x(\xi) := g^*(\xi|[x-ml,x]) \\ \mathcal{I}^*_x(\xi) &:= \mathcal{I}(\xi|[x-ml,x])^-), \quad \mathcal{I}_y(\xi) := \mathcal{I}(\xi|[y,y+ml]). \end{aligned}$$

Recall the definition of a piece of scenery $\psi|[y,y+lm]$ being OK. We say that a piece of scenery $\psi|[x-lm,x]$ is OK*, if $(\psi|[x-lm,x])^-$ is OK. Finally, let

$$B^1_{\text{intervals OK}} := \{\xi|[z,z+ml] \text{ is OK } \forall z\in I_1\} \cap \{\xi|[z-ml,z] \text{ is OK}^* \ \forall z\in I_1\}.$$

7.3.2 Stopping-time events

Recall the definition of $T^1(j)$ and $T^3(j)$ in (7.1.9). Also recall the definition of $w^1(j)$, $w^2(j)$, $w^3(j)$.

***** Let

$$H_1 := [-4\exp(l_1), 4\exp(l_1)].$$

We define the event $E^1_{\text{enough times}}$ that states that all pairs (x,y) in H_1 such that $y-x=(c_1l_1)L$ pass the criterion **a)** of the selection rule.
At first an auxiliary event

$$E^1_{\text{enough times}}(x,y) := \left\{ \begin{array}{c} \text{there exists a set} \quad J(x,y)\subseteq[1,\exp(\alpha l_1)] \quad \text{such that} \\ |J(x,y)| > \exp(\gamma l_1) \quad \text{and for every} \quad j\in|J(x,y)| \\ S(T^1(j))=x, \ S(T^3(j))=y, \\ \hat{q}^*(w^1(j)) \le q^*_x(\xi), \ \hat{g}^*(w^1(j)) \sqsubseteq_{\mathcal{I}^*_x(\psi)} g^*_x(\xi), \\ \hat{q}(w^3(j)) \le q_y(\xi), \ \hat{g}(w^3(j)) \sqsubseteq_{\mathcal{I}_y(\psi)} g_y(\xi) \end{array} \right\}.$$

and now

$$E^1_{\text{enough times}} := \bigcap_{x,y\in H_1, \ x-y=Lc_1l_1} E^1_{\text{enough times}}(x,y).$$

* Recall the attributes defined in Subsection 7.1.4. For a set of attributes (I^*, I, q^*, q, g^*, g), we define the indexes

$$j_1 := \min\left\{ j \geq 0 : \begin{array}{c} \hat{q}^*(w^1(j)) \leq q^*,\ \hat{g}(w^1(j)) \sqsubseteq_{I^*} g^*,\ \hat{q}(w^3(j)) \leq q,\ \hat{g}(w^3(j)) \sqsubseteq_I g, \\ |S(T^3(j)) - S(T^1(j))| < Lc_1 l_1 \end{array} \right\}$$

$$j_k := \min\left\{ j > j_{k-1} : \begin{array}{c} \hat{q}^*(w^1(j)) \leq q^*,\ \hat{g}(w^1(j)) \sqsubseteq_{I^*} g^*,\ \hat{q}(w^3(j)) \leq q,\ \hat{g}(w^3(j)) \sqsubseteq_I g, \\ |S(T^3(j)) - S(T^1(j))| < Lc_1 l_1 \end{array} \right\}. \tag{7.3.3}$$

Here the minimum over empty set is defined to be ∞. Let $\kappa := \max\{k : j_k < \infty\}$.

Clearly the subindexes $j_1, j_2, \ldots$ depend on chosen attributes (I^*, I, q^*, q, g^*, g).

Recall $B^1_{\text{recon straight}}$. The following events are of similar nature. Let

$$E^1_{\text{recon straight}}(I^*, I, q^*, q, g^*, g) :=$$
$$\Big\{\kappa > \exp(\gamma l_1), \quad \exists k \leq \kappa \quad \text{such that} \quad w^2(j_1) \neq w^2(j_k)\Big\} \cup \Big\{\kappa \leq \exp(\gamma l_1)\Big\},$$
$$E^1_{\text{recon straight}} := \bigcap_{I^*, I, q^*, q, g^*, g} E^1_{\text{recon straight}}(I^*, I, q^*, q, g^*, g),$$

where the intersection is taken over all sets of attributes.

* We now define the more refined counterparts of $T^i(j)$ and $w^i(j)$, $i = 1, 3$. These stopping times are used for technical reasons, only.

Let

$$\kappa^i(z, 1) := \min\{j : T^1(j) = z\}, \quad i = 1, 3$$

and, inductively,

$$\kappa^i(z, k) := \min\{j > \kappa^i(z, k-1) : T^i(j) = z\} \quad i = 1, 3.$$

Thus, $\kappa^1(z, k)$ ($\kappa^3(z, k)$) is the index of k-th stopping time $\tau(j)$, for which $S(T^1(j)) = z$ ($S(T^3(j)) = z$).
Define

$$T^i_z(k) := T^i(\kappa^i(z, k)), \quad w^i_z(k) = w^i(\kappa^1(z, k)) \quad i = 1, 3.$$

and let

$$E^1_{\text{mistake-l}}(z, x, k) := \Big\{\hat{q}^*(w^1_z(k)) \leq q^*_x(\xi)\Big\} \cap \Big\{\hat{g}^*(w^1_z(k)) \sqsubseteq_{\mathcal{I}^*_x(\xi)} g^*_x(\xi)\Big\} \cap \Big\{x \text{ is a left-barrier point}\Big\}$$
$$E^1_{\text{mistake-l}} := \bigcup E^1_{\text{mistake-l}}(z, x, k),$$

where the union is taken over all z, x, k such that $x < z, z, x \in I_1$ and $k \leq \exp(\alpha l_1)$.

Similarly, the right side

$$E^1_{\text{mistake-r}}(z, y, k) := \Big\{\hat{q}(w^3_z(k)) \leq q_y(\xi)\Big\} \cap \Big\{\hat{g}(w^3_z(k)) \sqsubseteq_{\mathcal{I}_y(\xi)} g_y(\xi)\Big\} \cap \Big\{y \text{ is a right-barrier point}\Big\}$$
$$E^1_{\text{mistake-r}} := \bigcup E^1_{\text{mistake-r}}(z, y, k),$$

where the union is taken over all z, y, k such that $z < y, z, y \in I_1$ and $k \leq \exp(\alpha l_1)$. Finally, let

$$E^1_{\text{no mistake}} := \left(E^1_{\text{mistake-l}} \cap E^1_{\text{mistake-r}}\right)^c.$$

7.3.3 Algorithm

We are ready to give the precise definition of the algorithm $\mathcal{A}^1$. Recall that the input of $\mathcal{A}^1$ consists of a piece of observations $\chi^{12\alpha l_1}$, stopping times τ and a little piece of true scenery ψ^o. The output of $\mathcal{A}^1$ is a word of length $4\exp(l_1)$.

As described in Subsection 7.1.4, the construction of $\mathcal{A}^1$ consists of two phases.

Phase I Collect the ladder words of $\xi|I^1$. For this, the observation-words triples $(w^1(j), w^2(j), w^3(j))$ are used. The word $w^2(j)$ will be collected as a ladder word, if it passes selection procedure given in Definition 7.1.1. The set of collected works is denoted by $\mathcal{W}^1$.

Phase II We assemble the words from $\mathcal{W}^1$ to get a big word of length $4\exp(l_1)$ as the output. This means the construction of a big word (of length $4\exp(l_1)$) by attaching, one by one, suitable words from $\mathcal{W}^1$. We start from ψ^o, and we attach to it a word from $\mathcal{W}^1$, which has an overlap with ψ^o at least $\frac{c_1 l_1}{4}$. We then attach a word from $\mathcal{W}^1$ to the enlarged ψ^o using the same overlapping-criterion. We proceed so, until the desired length has been achieved.

The description and the formal definition of **Phase I** was given in Subsection 7.1.4. As mentioned, the selection rule is the most crucial part of the whole scenery reconstruction; it must be restrictive enough to ensure that only ladder words of ladder pieces of original scenery ξ can pass it (with high probability). Formally, the following event should hold

$$E^1_{\text{only ladders}} := \{\forall w \in \mathcal{W}^1 \text{ there exists } I \in \mathcal{L}(c_1 l_1) \text{ such that } I \subseteq I_1 \text{ and } \xi|I \approx_l w\}.$$

On the other hand, the selection rule must be flexible enough to ensure that enough ladder words pass it (otherwise the set $\mathcal{W}^1$ is too small). More precisely, the following event should hold

$$E^1_{\text{enough ladders}} :=$$
$$\left\{ \begin{array}{c} \text{for any } j = 0, \ldots, L-1 \text{ and for any } z \in [-3\exp(l_1), 3\exp(l_1)], \text{ there exists} \\ \bar{x} \in [z - (c_1 l_1/4)L, z],\ \bar{y} \in [z, z + (c_1 l_1/4)L] \text{ such that: } \bar{x} \text{mod} L = j,\ \bar{y} \text{mod} L = j \\ \text{and } (\xi(\bar{x}), \xi(\bar{x}+L), \ldots, \xi(\bar{x} + (c_1 l_1)L)), (\xi(\bar{y}), \xi(\bar{y}-L), \ldots, \xi(\bar{y} - (c_1 l_1)L)) \in \mathcal{W}^1 \end{array} \right\}.$$

We now give the precise definition of assembling rule for **Phase II**. This definition completes the definition of $\mathcal{A}^1$.

For a ladder interval I and a set $D \subseteq \mathbb{Z}$ we write $|I \cap D| \geq r$ if there exists a ladder interval $J \in \mathcal{L}(r)$ such that $J \subseteq D \cap I$. Recall that two pieces of scenery φ and φ' are strongly equivalent, $\varphi \equiv \varphi'$, if φ is obtained by some translation of φ'. Let $\psi^o \in \{0,1\}^{k+1}$ be the given piece of original scenery. Thus, $\psi^o \equiv \xi|I^o$ for some interval $I^o \subseteq [-\exp(l_1), \exp(l_1)]$.

Definition 7.3.1. *We say that the piece of scenery $\varphi \in \{0,1\}^{[-2\exp(l_1),2\exp(l_1)]}$ is a solution, formally $\varphi \in \mathcal{S}(\chi^{12\alpha l_1},\tau,\psi^o)$, if and only if there exist $\varphi_i \in \{0,1\}^{D_i}$, $i=1,2,\dots,n$ such that $D_i \subseteq [-3\exp(l_1),3\exp(l_1)]$ and the following conditions are satisfied:*

1. $D_1=[0,k]$, $\varphi_1 \equiv \psi^o$;

2. *for each $i=2,\dots,n$ it holds $\varphi_i|D_{i-1}=\varphi_{i-1}$;*

3. *for each $i=2,\dots,n$ there exists $I_i \in \mathcal{L}(c_1l_1)$ such that*

3a) $D_i = D_{i-1}\cup V_i$;

3b) $|D_{i-1}\cap V_i| \geq \frac{c_1l_1}{4}$;

3c) $\exists w_i \in \mathcal{W}^1(\chi^{12\alpha l_1},\tau)$ *such that* $\varphi_i|V_i \approx_l w_i$;

4. $[-2\exp(l_1),2\exp(l_1)] \subseteq D_n, \quad \varphi=\varphi_n|[-2\exp(l_1),2\exp(l_1)]$.

Finally, the formal definition of $\mathcal{A}^1$. The output is any element of $\mathcal{S}$; we choose one of them, if $\mathcal{S}$ is not empty.

Definition 7.3.2. *We define $\mathcal{A}^1(\chi^{12\alpha l_1},\tau,\psi^o)$ as follows:*

- *If $\mathcal{S}(\chi^{12\alpha l_1},\tau,\psi^o)$ is nonempty, then we define $\mathcal{A}^1(\chi^{12\alpha l_1},\tau,\psi^o)$ to be its lexicographically smallest element;*
- *otherwise, $\mathcal{A}^1(\chi^{12\alpha l_1},\tau,\psi^o) := (1)_{[-2\exp(l_1),2\exp(l_1)]}$.*

7.3.4 Combinatorics for main theorem

The rest of the paper is the proof of Theorem 7.1.1. In this subsection, we prove some useful inclusions.

Lemma 7.3.1. *The following inclusions hold*

$$E^1_{\textit{recon straight}} \cap E^1_{\textit{stop}}(\tau) \subseteq E^1_{\textit{only ladders}}; \tag{7.3.4}$$

$$B^1_{\textit{intervals OK}} \cap E^1_{\textit{stop}}(\tau) \cap E^1_{\textit{no mistake}} \cap E^1_{\textit{enough times}} \cap E^1_{\textit{enough barriers}} \subseteq E^1_{\textit{enough ladders}}; \tag{7.3.5}$$

$$E^1_{\textit{only ladders}} \cap E^1_{\textit{enough ladders}} \cap B^1_{\textit{unique fit}} \subseteq E^1_{\textit{alg works}}(\tau), \tag{7.3.6}$$

provided l_1 is big enough.

Proof. At first note: if $E^1_{\rm stop}(\tau)$ holds, then, for each $j=1,2,\dots,\exp(\alpha l_1)$, it holds

$$|S(T^3(j))| \leq |S(\tau(j)) + L(\exp(2l_1)+lm^2+c_1l_1)| \leq \exp(3l_1), \tag{7.3.7}$$

provided l_1 is big enough. Thus, in this case, during the time interval $[T^1(j),T^3(j)]$, S stays on I_1, $j=1,2,\dots,\exp(\alpha l_1)$. In particular, all words $w^2(j)$ will be collected, when S stays on I_1.

Proof of (12.4.12):
We prove

$$\left(E^1_{\text{only ladders}}\right)^c \cap E^1_{\text{stop}}(\tau) \subseteq \left(E^1_{\text{recon straight}}\right)^c. \tag{7.3.8}$$

Suppose $\left(E^1_{\text{only ladders}}\right)^c \cap E^1_{\text{stop}}(\tau)$ holds. Then there exists a $w \in \mathcal{W}^1$ that is not a ladder word of any ladder piece $\xi|I$ of length l_1c_1 such that $I \subseteq I_1$. However, the word w has passed the selection rule. This means that for a complect of attributes (I^*, I, q^*, q, g^*, g) the conditions **1.** and **2.** of Definition 7.1.1 hold. This means, that

$$|S(T^3(j)) - S(T^1(j))| < c_1 l_1, \quad \forall j \in J(I^*, I, q^*, q, g^*, g). \tag{7.3.9}$$

Indeed, if there were an index $j^* \in J(I^*, I, q^*, q, g^*, g)$ such that (7.3.9) fails, then there would be a ladder interval I of length $c_1 l_1$ such that $\xi|I \approx_l w$. Clearly, during the time interval $[T^1(j^*), T^3(j^*)]$, the random walk S is on I. Since then S is also on I_1, we get $I \subseteq I_1$. This contradicts our assumption on w.
Recall the definition of κ. Since $|J(I^*, I, q^*, q, g^*, g)| > \exp[\gamma l_1]$, we have $\kappa > \exp[\gamma l_1]$. On the other hand, by **b)** of Definition 7.1.1, for each j_k, $k = 1, 2, \ldots, \kappa$, it holds $w(j_k) = w(j_1) = w$. Thus, $E^1_{\text{recon straight}}(I^*, I, q^*, q, g^*, g)$ fails. This completes the proof of (7.3.8).

Proof of (12.4.13):
Let $x, y \in H_1$ and $y - x = (c_1 l_1)L$. Since $B^1_{\text{intervals OK}}$ holds then, by (7.2.3), $I^* = \mathcal{I}(\xi|[x - lm])$ and $I = \mathcal{I}(\xi|[y, y + lm])$ satisfy $|I^*|, |I| \geq l(1 - 3\epsilon(n))$. Since $E^1_{\text{enough times}}(x, y)$ holds, there exists q-vectors $q^* = q^*_x(\xi)$, $q = q_y(\xi)$ and g-words $g^* = g_x(\xi), g = g_y(\xi)$ such that for each $j \in J(x, y)$, (7.1.18) holds. Moreover, $|J(x, y)| > \exp(\gamma l_1)$ and for each $j \in J(x, y)$ it holds $S(T^1(j)) = x$ and $S(T^3(j)) = y$. Then, obviously, $w^2(j) = (\xi(x), \xi(x + L), \ldots, \xi(y))$. Hence, we have a set of attributes (I^*, I, q^*, q, g^*, g) and an index set $J' = J(x, y) \subseteq J(I^*, I, q^*, q, g^*, g)$ such that $|J'| > \exp(\gamma l_1)$ and $w^2(j)$ is constant on J'.
Assume, in addition, that (x, y) is a barrier. Then $J' = J(I^*, I, q^*, q, g^*, g)$. Suppose not. Then there exists $j^* \in J(I^*, I, q^*, q, g^*, g) \backslash J'$. This means that j^* satisfies (7.1.18), but $w^2(j^*) \neq (\xi(x), \xi(x + L), \ldots, \xi(y))$. The latter is possible only, if $S(T^1(j^*)) > x$ or $S(T^3(j^*)) < y$. Let $S(T^1(j^*)) = z > x$. The event $E^1_{\text{stop}}(\tau)$ implies (7.3.7) and then $z \in I_1$. Hence, there is $z \in I_1$ and $k^* \leq j^*$ such that $E^1_{\text{mistake-1}}(z, x, k^*)$ holds. This is a contradiction with $E^1_{\text{no mistake}}$. Hence $J' = J(I^*, I, q^*, q, g^*, g)$ and $(\xi(x), \xi(x + L), \ldots, \xi(y)) \in \mathcal{W}^1$.
Now, let $z \in [-3\exp(l_1), 3\exp(l_1)]$. Then $z - \frac{c_1 l_1}{4} L \in I_1$ and by $B^1_{\text{enough barriers}}$ there exists a barrier (x', y') such that $x' \in [z - \frac{c_1 l_1}{4} L, z]$. Clearly, $(x', y') \in H_1$ (provided l_1 is big enough) and by the foregoing argument, $(\xi(x'), \xi(x' + L), \ldots, \xi(y')) \in \mathcal{W}^1$. Similarly, $z - (c_1 l_1)L \in I_1$ and there exists another barrier (x'', y'') such that $x'' \in [z - c_1 l_1 L, z - \frac{3c_1 l_1}{4} L]$ and, therefore, $y'' \in [z, z + \frac{c_1 l_1}{4} L]$. Again $(x', y') \in H_1$ and $(\xi(x''), \xi(x'' + L), \ldots, \xi(y'')) \in \mathcal{W}^1$. Finally, take $\bar{x} = x''$ and $\bar{y} = y''$.

Proof of (7.3.6):
It suffices to show that $E^1_{\text{only ladders}} \cap E^1_{\text{enough ladders}} \cap B^1_{\text{unique fit}}$ ensures that for each $I^o \subseteq [-e^{l_1}, e^{l_1}]$, it holds $\mathcal{S}(\chi^{12\alpha l_1}, \tau, \xi|I^o)$ consists of one element that satisfies (7.1.2).
Consider the "puzzle-playing" algorithm formalized in Definition 7.3.1. We show that there is an unique way to combine the words from $\mathcal{W}^1$, i.e. the solution set $\mathcal{S}$ is unique. Let $\varphi \in \mathcal{S}$ and let $D_1 \subseteq D_2 \subseteq \cdots \subseteq D_n$ be the sequence of sets ensured by the definition of φ. By **1**, $\varphi|D_1$ is translated from a piece of $\xi|I_1$ by some b satisfying $|b| \leq \exp(l_1)$, i.e. $\xi|I^o = T[\varphi|D_1]$, where $Tz = z + b$ is the translation and $I_o \subseteq [-e^{l_1}, e^{l_1}] \subseteq I_1$. We show: if $\varphi|D_i$ is translated from a piece of $\xi|I_1$ by b, i.e. $\xi|J_i = T[\varphi|D_i]$, for some $J_i \subseteq I_1$, then the same applies for $\varphi|D_{i+1}$. Recall that $\varphi|D_{i+1}$ and $\varphi|D_i$ differ on V_{i+1}, only. By

3c) and $E^1_{\text{only ladders}}$, $\varphi|V_{i+1} \approx \xi|J(w)$ for some $J(w) \subseteq I_1$. Thus, there is an affine T' such that $\xi|J(w) = T'[\varphi|V_{i+1}]$ and, hence, there is a ladder interval $J' \subseteq J(w)$ such that $\xi|J' = T'[\varphi|(V_{i+1} \cap D_i)]$. So, $\varphi|(V_{i+1} \cap D_i)$ is equivalent with some ladder word of $\xi|I_1$ by T'. On the other hand, $\varphi|(V_{i+1} \cap D_i)$ is translated by b, hence it is equivalent with some ladder word of $\xi|I_1$ by T. Let this word be $\xi|J$. Clearly $\xi|J \approx \xi|J'$. By **3b)**, the length of the ladder interval $V_{i+1} \cap D_i$ as well as J' and J is at least $\frac{c_1 l_1}{4}$. If $T \neq T'$, then $J \neq J'$, which contradicts $B^1_{\text{unique fit}}$. Hence, $T' = T$ and $\varphi|V_{i+1}$ is translated from a piece of $\xi|I_1$ by b and $\varphi|D_{i+1}$ is translated from a piece of $\xi|I_1$ by b as well. The same holds for φ, i.e. $\varphi \equiv \xi|I(\varphi)$ for some interval $I(\varphi)$. By **4**, $I(\varphi) = [a_o - 2\exp(l_1), a_o + 2\exp(l_1)]$, where $I_o := [a_o, b_o]$. So, φ is obtained from a fixed piece of scenery $\xi|I(\varphi)$ by a fixed translation, T. Clearly such a φ is unique.
Let us show that φ satisfies (7.1.2). Since $|a_o| \leq \exp(l_1)$, we have that

$$[-\exp(l_1), \exp(l_1)] \subseteq I(\varphi) \subseteq [-3\exp(l_1), 3\exp(l_1)].$$

This means

$$\xi|[-\exp(l_1), \exp(l_1)] \sqsubseteq \varphi \sqsubseteq \xi|[-3\exp(l_1), 3\exp(l_1)],$$

i.e. (7.1.2) holds.
Let us show that $\mathcal{S}$ is not empty. Fix an $i \geq 1$ and let $D = D_i$, be the domain of φ_i. Note that $D = \cup_{j=0}^{L-1} I(j)$, where $I(j)$ is a ladder interval with length at least $c_1 l_1$. Hence, D is an union of disjoint ladder intervals. Let $a_j < b_j$ be the endpoints of $I(j)$. If, for each j, $a_j \leq -2\exp(l_1)$ and $b_j \geq 2\exp(l_1)$, then $[-2\exp(l_1), 2\exp(l_1)] \subseteq D_i$ and there is nothing to prove. Therefore, without loss of generality assume j to be such that $b_j < 2\exp(l_1)$. Obviously, $b_j > 0$. It suffices to show that there exists $V = (v, v + L, \ldots, v + L(c_1 l_1)) \in \mathcal{L}(c_1 l_1)$ and a piece $\varphi_{i+1} \in \{0,1\}^{I(j) \cup V}$ such that:

- $\varphi_{i+1}|I(j) = \varphi_i|I(j)$
- $\exists w \in \mathcal{W}^1$ such that $\varphi_{i+1}|V \approx_l w$
- $b_j = v + kL$, where $\frac{c_1 l_1}{4} \leq k \leq \frac{c_1 l_1}{2}$. This means that $|I(j) \cap V| \geq \frac{c_1 l_1}{4}$ but $|V \backslash D| \geq \frac{c_1 l_1}{2}$.

We know that φ_i is a translation of $\xi|J_i$ for some $J_i \in I_1$. This means that $I(j)$ is a translation of a ladder interval $J(j)$. Let d_j be the endpoint of $J(j)$. We also know that this translation is not more than e^{l_1}. Hence $d_j \in [-\exp(l_1), 3\exp(l_1)]$. Consider the ladder interval

$$\bar{J}(j) := \big(d_j - 2(c_1 l_1/4)L, \ldots, d_j - 1(c_1 l_1/4)L\big).$$

By $E^1_{\text{enough ladders}}$ there exists $\bar{x} \in \bar{I}(j)$ such that a ladder word of $\xi|V(\bar{x})$, with $V(\bar{x}) = \big(\bar{x}, \bar{x} + L, \ldots, \bar{x} + (c_1 l_1)L\big) \in \mathcal{L}(c_1 l_1)$ belongs to $\mathcal{W}^1$. Let this word be $\bar{w}$. Clearly, $d_j = \bar{x} + kL$, where $\frac{c_1 l_1}{4} \leq k \leq \frac{c_1 l_1}{2}$. By $B^1_{\text{unique fit}}$, $\bar{w}$ is not a ladder word of any ladder piece $\varphi_i|V_j$, $j = 1, \ldots, i$. This means that the word $w \in \mathcal{W}^1$ has not been used before. Hence $\bar{w}$ and the translation of $V(\bar{x})$ can be taken as w and V. The same argument applies if $a_j > -2\exp(l_1)$, implying that D_i can be efficiently enlarged in other direction as well. □

7.3.5 Probabilities for main theorem

Scenery-dependent events

At first, estimate the probabilities of B-events. These events depend on ξ, only. Note that all exponential bounds are valid for l_1 being big enough.

Estimate $P\big((B^1_{\text{intervals OK}})^c\big)$

Let

$$E := \{\xi | [z, z+ml] \text{ is OK } \forall z \in I_1\}, \quad E^* := \{\xi | [z-ml, z] \text{ is OK}^* \ \forall z \in I_1\}.$$

Now, by translation invariancy of ξ and Theorem 7.2.1, it holds that for l_1 big enough

$$P(E^c) \le \sum_{z \in I_1} P(\xi | [z, z+ml] \text{ is not OK}) \le 2e^{3l_1} P(E^c_{\text{OK}}) \le 2\exp[3l_1 - al].$$

Similarly,

$$P(E^{*c}) \le \sum_{z \in I_1} P(\xi | [z, z-ml] \text{ is not OK}^*) \le 2e^{3l_1} P(E^c_{\text{OK}}) \le 2\exp[3l_1 - al].$$

Hence, if l_1 is sufficiently big, then

$$P\Big(\big(B^1_{\text{intervals OK}}\big)^c\Big) \le 4\exp[(3 - al_2)l_1]. \tag{7.3.10}$$

The following proposition also specifies the choice of c_1.

Proposition 7.3.1. *There exists constants $C_1(n)$ and $k_1, k_2, k_3 > 0$ not depending on l_1 such that for $c_1 > C_1(n)$ it holds:*

$$P\Big(\big(B^1_{\text{unique fit}}\big)^c\Big) \le \exp[-k_1 l_1] \tag{7.3.11}$$

$$P\Big(\big(B^1_{\text{recon straight}}\big)^c\Big) \le \exp[-k_2 l_1] \tag{7.3.12}$$

$$P\Big(\big(B^1_{\text{enough barriers}}\big)^c\Big) \le \exp[-k_3 l_1], \tag{7.3.13}$$

provided l_1 is big enough.

Proof. It follows from Lemma 6.33 in [20] that for some constants a_1, a_2 depending on L, only, the bound $P\big((B^1_{\text{unique fit}})^c\big) \le a_1 \exp[-a_2 l_1]$ is valid. Also, there is a fixed constant C_r such that $a_2 > 0$ if $c_1 > C_r$. This implies (7.3.11) for l_1 sufficiently big.

Estimate $P\big((B^1_{\text{recon straight}})^c\big)$

Let $\mathcal{R}(l_1c_1)(x,y) := \{\mathcal{R}(l_1c_1)(x,y) : R(0) = x, R(l_1c_1L) = y\}$. Thus $\mathcal{R}(l_1c_1)(x,y)$ is (possibly empty) the set of admissible path from x to y with l_1c_1 steps. Fix x, y such that $|y - x| < (l_1c_1)L$. At first note: if l_1 is big enough, then (for any value of $c_1 \ge 1$) $\mathcal{R}(l_1c_1)(x,y)$ is either empty or has cardinality at least 2. Any admissible path

$R \in \mathcal{R}(l_1c_1)(x,y)$ is a sequence $R = (t_1, \dots, t_{c_1l_1})$ of steps, where $|t_i| \le L$. Hence, there exists a $R = (t_1, \dots, t_{c_1l_1}) \in \mathcal{R}(l_1c_1)(x,y)$ such that $t_i \neq t_1$ for a $i = 2, \dots c_1l_1$ (if no, then $\mathcal{R}(l_1c_1)(x,y)$ would consists of one path, only). Let R be one of such paths. Let $c_1 \ge \lceil \frac{100}{2L+1} \rceil$. The number of possible steps is bounded by $2L+1$. Hence, there is a step t' that occurs in R at least $2k := 100l_1$ times. If $t' = 0$, then there exists a $t \neq 0$ that occurs at least k times Formally, $\exists t \in \{-L, \dots, L\}, t \neq 0$ such that $|\{i = 1, \dots, c_1l_1 : t_i = t\}| \ge 50l_1$. Any rearrangement of the order of steps in R corresponds to another path in $\mathcal{R}(l_1c_1)(x,y)$. We consider two rearrangements of R. The first, R^1, starts with k steps of size t. Thus $R^1 = \{t^1_1, \dots, t^1_{c_1l_1}\} \in \mathcal{R}(l_1c_1)(x,y)$ is such that $t^1_1 = \cdots = t^1_k = t$. Let u be another step if R such that $u \neq t$. The second path, R^2, starts with u, and then is followed by k-steps of size t. Formally, $R^2 = \{t^2_1, \dots, t^2_{c_1l_1}\} \in \mathcal{R}(l_1c_1)(x,y)$ is such that $t^2_1 = u$, $t^2_2 = \cdots = t^2_{k+1} = t$. We now estimate the probability that the paths R^1 and R^2 generate the same word in observation; we estimate

$$
\begin{aligned}
&P(\xi \circ R^1 = \xi \circ R^2) \\
&\le P\Big(\big(\xi(x+t), \dots, \xi(x+kt)\big) = \big(\xi(x+u), \xi(x+u+t), \dots, \xi(x+u+(k-1)t)\big)\Big) \\
&\le P\big(\xi(x+t) = \xi(x+u)\big) P\big(\xi(x+2t) = \xi(x+u+t) | \xi(x+t) = \xi(x+u)\big) \times \\
&\times P\big(\xi(x+3t) = \xi(x+u+2t) | \xi(x+t) = \xi(x+u)), \xi(x+2t) = \xi(x+u+t)\big) \times \cdots \\
&\cdots \times P\big(\xi(x+kt) = \xi(x+u+(k-1)t) | \xi(x+t) \\
&= \xi(x+u), \cdots, \xi(x+(k-1)t) = \xi(x+u+(k-2)t)\big) \\
&\le 2^{-k} = \exp[-50 \ln 2l_1].
\end{aligned}
$$

Now,

$$
E_{\text{recon straight}} = \bigcup_{x,y \in I_1, |x-y| < l_1c_1} E_{\text{recon straight}}(x,y),
$$

$$
P((E_{\text{recon straight}})^c) \le \sum_{x,y \in I_1} P(E_{\text{recon straight}}(x,y)) \le 4 \exp(6l_1) \exp[-50 \ln 2l_1] \le \exp[-25l_1].
$$

Estimate $P\big((B^1_{\text{enough barriers}})^c\big)$

For each z, j define

$$
B^1{}_{\text{enough barriers}}(z,j) := \left\{ \begin{array}{c} \text{there exists } x \in [z, z + (\frac{c_1l_1}{4})L] \text{ such that } x \mod L = j \\ \text{and } (x, x + (c_1l_1))L\} \text{ is a barrier of } \xi \end{array} \right\}.
$$

Define

$$
B(x) := \Big\{(x, x + (c_1l_1)L) \text{ is a barrier of } \xi\Big\}, \quad Y_x := I_{B(x)}.
$$

Note, if $c_1l_1L - 3m^2L \ge \frac{c_1l_1}{4} \ge x' - x \ge 3m^3L =: r$, then, by the definition, the events $B(x)$ and $B(x')$ are independent. Clearly the probability of $B(x)$ does not depend on x, let us denote $p = P(B(x))$. By definition, $p > 2^{-3m^3L}$. Denote $w = \llcorner \frac{c_1l_1}{4r} - \frac{L}{r} \lrcorner > \frac{c_1 - 4L}{4r} l_1$.

Without loss of generality assume $z \bmod L = 0$. By Höffding's inequality,

$$P\Big(\big(B^1_{\text{enough barriers}}(z,j)\big)^c\Big) = P\Big(\sum_{k=1}^{K(j)} Y_{z+(k-1)L+j} = 0\Big) \le P\Big(\sum_{k=1}^{\frac{c_1 l_1}{4r}} Y_{r(k-1)+z+j} = 0\Big)$$
$$\le P\Big(\sum_{k=1}^{w}(Y_{r(k-1)+z+j} - p) \le wp\Big) \le 2\exp[-2wp^2]$$
$$\le 2\exp[-2\frac{c_1 - 4L}{4r}2^{-6m^3L}l_1] = 2\exp[-k_2' l_1],$$

for $k_2' := \frac{c_1-4L}{4r}2^{-(6m^3L+1)}$. Here $K(j) = \frac{c_1 l_1}{4}$, if $j \neq 0$ and $K(0) = \frac{c_1 l_1}{4} + 1$.
Obviously, $k_2' > 0$, if $c_1 > 4L$. Thus

$$P\Big(\big(B^1_{\text{enough barriers}}\big)^c\Big) \le \sum_{z\in I_1, j=\{0,\dots,L-1\}} P\Big(\big(B^1_{\text{enough barriers}}(z,j)\big)^c\Big) \le 8\exp[(6-k_2')l_1] \le \exp[-l_1],$$

if $k_2' \ge 8$. The latter implies $c_1 - 4L \ge r4\cdot 2^{6m^3+6}$ or $c_1 \ge r2^{6m^3+8} + 4L = 3m^3L2^{6m^3+8} + 4L$.

Hence, Proposition 7.3.1 holds with $C_1(n) := \max\{C_r, \ulcorner\frac{100}{2L+1}\urcorner, 3m^3L2^{6m^3+8} + 4L\}$. □

Random-walk depending events

In the present subsection, we estimate the events that also depend on the random walk.

Estimate $P(E^1_{\text{mistake-r}} \cap B^1_{\text{intervals OK}})$.

Fix $y, z \in I_1$, $z < y$ and note

$$E^1_{\text{mistake-r}}(z,y,k) \cap B^1_{\text{intervals OK}} \subseteq E^1_{\text{mistake-r}}(z,y,k) \cap \{\xi | [y, y+lm] \text{ is OK}\}, \quad k = 1,2,\dots. \tag{7.3.14}$$

We now estimate the right side of (7.3.14). Recall the definitions of $T^3_z(k), w^3_z(k)$ and $g_y(\xi)$. Consider the events

$$E^1_{\text{mistake-r}}(y,z,k) \cap \{\xi | [y, y+lm] \text{is OK}\} =$$
$$\Big\{\hat{q}(w^3_z(k)) \le q_y(\xi),\ \hat{g}(w^3_z(k)) \sqsubseteq_{\mathcal{I}_y(\xi)} g_y(\xi),\ y \text{ is a right barrier point},\ \xi | [y, y+lm] \text{ is OK}\}, \tag{7.3.15}$$

Because of (7.1.1), conditionally on ξ the events (7.3.15) are independent and identically distributed. Hence, the events (7.3.15) all have the probability equal to

$$P\Big(\hat{q}(\chi^{m^2l}_z) \le q_y(\xi), \quad \hat{g}(\chi^{m^2l}_z) \sqsubseteq_{\mathcal{I}_y(\xi)} g_y(\xi), \quad y \text{ is a right barrier point}, \quad \xi | [y, y+lm] \text{ is OK}\Big). \tag{7.3.16}$$

The event in (7.3.16) depends on ξ, only. The distribution of ξ is obviously translation invariant. Therefore, by Corollary 7.2.1, (7.3.16) can be estimated

$$P\Big(\hat{q}(\chi^{m^2l}_{z-y}) \le q_0(\xi)), \quad \hat{g}(\chi^{m^2l}_{z-y}) \sqsubseteq_{\mathcal{I}(\xi^{ml})} g_0(\xi), \quad 0 \text{ is a right barrier point }, \quad \xi^{ml} \text{ is OK}\Big) =$$
$$P\Big(\Big\{\hat{q}(\chi^{m^2l}_{z-y}) \le q_0(\xi), \quad \hat{g}(\chi^{m^2l}_{z-y}) \sqsubseteq_{\mathcal{I}(\xi^{ml})} g_0(\xi)\Big\} \cap E_{\text{origin}} \cap E_{OK}\Big) =$$
$$P\Big(E_{\text{mistake}}(z-y) \cap E_{OK}\Big) \le \exp(-l\alpha_I),$$

provided l_1 is big enough. Therefore,

$$P(E^1_{\text{mistake-r}} \cap B^1_{\text{intervals OK}}) \le \sum_{y,z,k} P(E^1_{\text{mistake-r}}(y,z,k) \cap B^1_{\text{intervals OK}}) \le \sum_{y,z,k} \exp(-l\alpha_I) < 4\exp[(6+\alpha)l_1 - \alpha_I l]. \tag{7.3.17}$$

The sum here is taken over all $z, y \in I_1$, $z < y$ and $k = 1, \dots, \exp(\alpha l_1)$.

Estimate $P(E^1_{\text{mistake-l}} \cap B^1_{\text{intervals OK}})$.

We need some additional notations. Recall $T^1_z(k)$. Now fix $x' \in I_1$ and define $T^1_z(k_i)$, $i = 1, 2, \dots, N(x')$. as the i-th stopping time $T^1_z(k)$, for which $S(T^1_z(k) + \exp(2l_1)) = x'$. The indexes k_i depend on chosen x'. Define now

$$E^1_{\text{mistake-l}}(z,x,i,x') := \Big\{\hat q^*(w^1_z(k_i)) \le q^*_x(\xi)\Big\} \cap \Big\{\hat g^*(w^1_z(k_i)) \sqsubseteq_{\mathcal{I}^*_x(\xi)} g^*_x(\xi)\Big\} \cap \Big\{x \text{ is a left barrier point}\Big\},$$

$i = 1, 2, \dots, N(x')$.
Clearly, for each k there exist i, x' such that $E^1_{\text{mistake-l}}(z,x,k) = E^1_{\text{mistake-l}}(z,x,i,x')$. The counterpart of (7.3.14) is

$$E^1_{\text{mistake-l}}(z,x,i,x') \cap B^1_{\text{intervals OK}} \subseteq E^1_{\text{mistake-r}}(z,x,i,x') \cap \{\xi | [x-lm, x] \text{ is OK}^*\} =: E(i,x'),$$

$i = 1, 2, \dots, N(x')$.
As previously, we observe that $P(E(i,x'))$ is equal to

$$P\Big(\hat q^*(\chi^{m^2 l}_{x'}) \le q^*_x(\xi),\ \hat g^*(\chi^{m^2 l}_{x'}) \sqsubseteq_{\mathcal{I}^*_x(\xi)} g^*_x(\xi),\ S_{x'}(m^2 l) = z,\ x \text{ is a left b. p. },\ \xi|[x-lm, x] \text{ is OK}^*\Big). \tag{7.3.18}$$

To calculate (7.3.18), at first note the following. Let $R(i)$, $i = 0, 1, \dots, k$ be an admissible path such that $R(0) = x'$, $R(k) = z$. Thus, for any scenery ψ, the observation $\chi|[0,k]$ equals $\psi(R(i))$, $i = 0, \dots, k$. This means, $\big(\chi|[0,k]\big)^- = \psi(R^-(i))$, where $R^-(i) = -R(k-i)$, $i = 0, \dots, k$. By symmetry of S, any admissible path $R[0,k]$ has the same probability as its reverse $R^-[0,k]$. This means that for any $u \in \{0,1\}^{k+1}$ and for any fixed scenery ψ we have

$$P_\psi\Big(\big(\chi|[0,k]\big)^- = u,\ S_{x'}(k) = z\Big) = P_\psi\Big(\chi|[0,k] = u,\ S_z(k) = x',\Big)$$

or

$$P_\psi\Big(\big(\chi^k_{x'}\big)^- = u,\ S_{x'}(k) = z\Big) = P_\psi\Big(\chi^k_z = u,\ S_z(k) = x'\Big).$$

By symmetry, again, the right side of last equality equals

$$P_{\psi^-}\Big(\chi^k_{-z} = u,\ S_{-z}(k) = -x'\Big).$$

In particular, since $(\psi|[x-lm, x])^- = \psi^-|[-x, -x+lm]$

$$P_\psi\Big(\hat q\big((\chi^k_{x'})^-\big) \le q\big((\psi|[x-lm,x])^-\big),\ \hat g\big((\chi^k_{x'})^-\big) \sqsubseteq_{\mathcal{I}((\psi|[x-lm,x])^-)} g\big((\psi|[x-lm,x])^-\big),\ S_{x'}(k) = z\Big) =$$

$$P_{\psi^-}\Big(\hat q\big(\chi^k_{-z}\big) \le q(\psi^-|[-x,-x+lm]), \hat g\big(\chi^k_{-z}\big)\big) \sqsubseteq_{\mathcal{I}(\psi^-|[-x,-x+lm])} g(\psi^-|[-x+lm,-x]),\ S_{-z}(k) = -x'\Big).$$

Recall the definitions of $\hat{q}^*, q^*, \hat{g}^*, g^*$. Clearly x is a left barrier point for ψ if and only if $-x$ is a right barrier point for ψ^- and, by definition, $\psi|[x-lm,x]$ is OK* if and only $(\psi|[x-lm,x])^- = \psi^-|[-x,-x+lm]$ is OK. Let

$$A^*(x) := \{x \text{ is a left barrier point of } \psi, \psi|[x-lm,x] \text{ is OK*}\},$$
$$A(x) := \{x \text{ is a right barrier point of } \psi,\ \psi|[x,x+lm] \text{ is OK }\}.$$

Thus, for each ψ,

$$P_\psi\Big(\hat{q}^*(\chi^k_{x'}) \le q^*_x(\psi),\ \hat{g}^*(\chi^k_{x'}) \sqsubseteq_{\mathcal{I}^*_x(\psi)} g^*_x(\psi),\ S_{x'}(k) = z\Big) I_{A^*(x)}(\psi) =$$
$$P_{\psi^-}\Big(\hat{q}\big(\chi^k_{-z}\big) \le q_{-x}(\psi^-),\ \hat{g}(\chi^k_{-z}) \sqsubseteq_{\mathcal{I}_{-x}(\psi^-)} g_{-x}(\psi^-),\ S_{-z}(k) = -x'\Big) I_{A(-x)}(\psi^-).$$

Finally, integrate over ξ and use the fact that ξ and ξ^- have the same distribution to get

$$P\Big(\hat{q}^*(\chi^k_{x'}) \le q^*_x(\xi),\ \hat{g}^*(\chi^k_{x'}) \sqsubseteq_{\mathcal{I}^*_x(\xi)} g^*_x(\xi),\ S_{x'}(k) = z, \xi \in A^*(x)\Big) =$$
$$P\Big(\hat{q}\big(\chi^k_{-z}\big) \le q_{-x}(\xi),\ \hat{g}(\chi^k_{-z}) \sqsubseteq_{\mathcal{I}_{-x}(\xi)} g_{-x}(\xi),\ S_{-z}(k) = -x',\ \xi \in A(-x)\Big). \qquad (7.3.19)$$

Now take $k = m^2 l$, define $y := -x$, $z := -z$ and sum over x' to obtain that $P(E(i,x'))$ equals

$$P\Big(\hat{q}\big(\chi^{m^2 l}_z\big) \le q_y(\xi),\ \hat{g}(\chi^{m^2 l}_z) \sqsubseteq_{\mathcal{I}_y(\xi)} g_y(\xi),\ y \text{ is a right barrier point, } \xi|[y,y+lm] \text{ is OK}\Big).$$

Hence, $P(E(i,x'))$ equals (7.3.16) and, therefore, it is bounded by $\exp(-l\alpha_I)$. This means

$$P(E^1_{\text{mistake-l}} \cap B^1_{\text{intervals OK}}) \le \sum_{y,z,i,x'} P(E^1_{\text{mistake-r}}(y,z,i,x') \cap B^1_{\text{intervals OK}})$$
$$\le \sum_{y,z,i,x'} \exp(-l\alpha_I) < 8\exp[(9+\alpha)l_1 - \alpha_I l], \qquad (7.3.20)$$

where the sum is taken over all $z, y, x' \in I_1$, $z < y$ and $i = 1, \dots, \exp(\alpha l_1)$.

Estimate $P\big(E^1_{\text{stop}}(\tau) \cap B^1_{\text{recon straight}} \cap (E^1_{\text{recon straight}})^c\big)$

Fix a set of attributes (I^*, I, q^*, q, g^*, g) and consider random indexes $j_1, \dots, j_\kappa$ as in (7.3.3). They depend on chosen attributes. We consider the set E^c, where

$$E := E^1_{\text{recon straight}}(I^*, I, q^*, q, g^*, g).$$

On E^c, the following hold: $\kappa > \exp(\gamma l_1)$ and for every $k = 1, \dots, \exp(\gamma l_1) + 1$, it holds $w^2(j_k) = w^2(j_1)$. Define

$$Y_k := 1 - I_{w^2(j_1)}(w^2(j_k)), \quad k = 2, \dots, \kappa.$$

Hence $Y_k = 1$ if and only if $w^2(j_k) \ne w^2(j_1)$. Therefore, $E^c \subseteq \{\sum_{k=1}^{\exp(\gamma l_1)+1} Y_k = 0\}$. We now consider the following σ-algebra

$$\mathcal{A} := \sigma\Big(\xi(z),\ S(\tau(j)),\ S(T^1(j_k)),\ S(T^3(j_k)),\ z \in \mathbb{Z},\ j = 1, \dots, \exp(\alpha l_1),\ k = 1, \dots, \kappa\Big).$$

Given $\mathcal{A}$, the values of κ as well as $S(T^1(j_k)) = x_k$ and $S(T^3(j_k)) = y_k$, $k = 1, \dots, \kappa$ are known. This means that the random variables $Y_1, \dots, Y_\kappa$ depend on the behavior of S from x_k to y_k during $c_1 l_1$ steps. Hence, given $\mathcal{A}$ the random variables $Y_1, \dots, Y_\kappa$ are independent.
Consider now the events $E^1_{\text{stop}}(\tau)$ and $B^1_{\text{recon straight}}$. Obviously they both belong to $\mathcal{A}$. By (7.3.7), on $E^1_{\text{stop}}(\tau)$ we have that $x_k, y_k \in I_1$, for every $k = 1, \dots, \kappa$. Hence, if in addition also $B^1_{\text{recon straight}}$ holds, then for each $k = 2, \dots, \kappa$ there exists at least one admissible path from x_k to y_k that generates different words in observations. Recall the definition of p_{min} and deduce that on $E^1_{\text{stop}}(\tau) \cap B^1_{\text{recon straight}}$ it holds $P(Y_k = 1|\mathcal{A}) \geq (p_{min})^{c_1 l_1}$, $k = 2, \dots, \kappa$. Hence, by Höffding's inequality on $E^1_{\text{stop}}(\tau) \cap B^1_{\text{recon straight}}$,

$$P(E^c|\mathcal{A}) \leq P\Big(\sum_{k=2}^{\exp(\gamma l_1)+1} Y_k = 0 \Big| \mathcal{A}\Big) \leq \exp[-2\exp((\gamma + 2c_1 \ln p_{min}) l_1)]. \tag{7.3.21}$$

Indeed, for $Y_1, \dots, Y_{e^b}$ independent Bernoulli random variables with $E(X_i) \geq e^a$, the Höffding's inequality states

$$P\Big(\sum_{i=1}^{e^b} Y_i = 0\Big) = P\Big(\sum_{i=1}^{e^b} (Y_i - EY_i) \leq -\sum_{i=1}^{e^b} EY_i\Big) \leq \exp\big[-2e^{-b}\big(\sum_{i=1}^{e^b} EY_i\big)^2\big] \leq \exp[-2e^{b+2a}]$$

Now take $b = \gamma l_1$, $a = c_1 l_1 \ln(p_{min})$ to obtain (7.3.21).
Integrate (7.3.21) over $E^1_{\text{stop}}(\tau) \cap B^1_{\text{recon straight}}$ to obtain

$$P\Big(E^c \cap E^1_{\text{stop}}(\tau) \cap B^1_{\text{recon straight}}\Big) \leq \exp[-2\exp((\gamma + 2c_1 \ln p_{min}) l_1)]. \tag{7.3.22}$$

Finally, estimate

$$\begin{aligned} &P\Big(\big(E^1_{\text{recon straight}}\big)^c \cap E^1_{\text{stop}}(\tau) \cap B^1_{\text{recon straight}}\Big) \\ &\leq \sum_{(I^*, I, q^*, q, g^*, g)} P\Big(E^c \cap E^1_{\text{stop}}(\tau) \cap B^1_{\text{recon straight}}(I^*, I, q^*, q, g^*, g)\Big), \end{aligned}$$

where the sum is taken over all attributes (I^*, I, q^*, q, g^*, g). There are less than $2^{2(n^2 l + l)} l^{4l}$ attributes. Thus, the right side of the previous display is bounded by

$$\begin{aligned} &2^{2(n^2 l + l)} l^{4l} \exp[-d\exp((\gamma + 2c_1 \ln p_{min}) l_1)] = \\ &\exp[2(n^2 l + l)\ln 2 + (4l)\ln l - d\exp((\gamma + 2c_1 \ln p_{min}) l_1)] = \\ &\exp[l_1(2(n^2 l_2 + l_2)\ln 2 + (4l_2)(\ln l_1 + \ln l_2)) - d\exp((\gamma + 2c_1 \ln p_{min}) l_1)]. \end{aligned}$$

So,

$$\begin{aligned} &\big(E^1_{\text{stop}}(\tau) \cap B^1_{\text{recon straight}} \cap (E^1_{\text{recon straight}})^c\big) \leq \\ &\leq \exp[l_1(2(n^2 l_2 + l_2)\ln 2 + (4l_2)(\ln l_1 + \ln l_2)) - d\exp((\gamma + 2c_1 \ln p_{min}) l_1)]. \end{aligned} \tag{7.3.23}$$

Estimate $P\Big(E^1_{\text{stop}}(\tau) \cap (E^1_{\text{enough times}})^c \cap B^1_{\text{intervals OK}}\Big)$

Recall $p_L := P(S(1) - S(0) = L)$ and define

$$p^* := \exp[-(1.5 + 2\alpha_{II} l_2 + c_1 \ln p_L) l_1].$$

Proposition 7.3.2. *If*

$$\exp(\alpha l_1)p^* \geq 2\exp(\gamma l_1), \tag{7.3.24}$$

then

$$P\Big(E^1_{\text{stop}}(\tau)\cap(E^1_{\text{enough times}})^c\cap B^1_{\text{intervals OK}}\Big)\leq 64\exp[(2l_1-2\exp((2\gamma-\alpha)l_1)], \tag{7.3.25}$$

provided l_1 is big enough.

Proof. Recall the definitions of $T^1(j)$, $T^3(j)$, $j=1,\dots\exp(\alpha l_1)$. Let $x,y\in H_1$ be such that $y=x+c_1l_1L$ and define

$$E_j(x,y):=\left\{\begin{array}{c} S(T^1(j)-lm^2)=x-lm \\ S(T^1(j))=x,\ S(T^3(j))=y, \\ \hat{q}^*(w^1(j))\leq q^*_x(\xi),\ \hat{g}^*(w^1(j))\sqsubseteq_{\mathcal{I}^*_x(\xi)} g^*_x(\xi), \\ \hat{q}(w^3(j))\leq q_y(\xi),\ \hat{g}(w^3(j))\sqsubseteq_{\mathcal{I}_y(\xi)} g_y(\xi)\end{array}\right\},\quad Y_j:=I_{E_j},\quad j=1,\dots,e^{\alpha l_1}.$$

Obviously,

$$\Big\{\sum_{j=1}^{e^{\alpha l_1}}Y_j>e^{\gamma l_1}\Big\}\subseteq E^1_{\text{enough times}}(x,y). \tag{7.3.26}$$

For each j and for every scenery ψ, it holds

$$\begin{array}{l}P_\psi(Y_j=1)=P_\psi(S(T^1(j)-lm^2)=x-lm)\times\\ P_\psi(S(T^1(j))=x,\hat{q}^*(w^1(j))\leq q^*_x(\xi),\ \hat{g}^*(w^1(j))\sqsubseteq_{\mathcal{I}^*} g^*_x(\xi)|S(T^1(j)-lm^2)=x-lm)\times\\ P_\psi(S(T^3(j))=y|S(T^1(j)-lm^2)=x-lm,S(T^1(j))=x,\hat{q}^*(w^1(j))\leq q^*_x(\xi),\ \hat{g}^*(w^1(j))\sqsubseteq_{\mathcal{I}^*} g^*_x(\xi))\times\\ P_\psi(\hat{q}(w^3(j))\leq q_y(\xi),\ \hat{g}(w^3(j))\sqsubseteq_{\mathcal{I}} g_y(\xi)|\\ S(T^1(j)-lm^2)=x-lm,\ S(T^1(j))=x,\ S(T^3(j))=y,\hat{q}^*(w^1(j))\leq q^*_x(\xi),\ \hat{g}^*(w^1(j))\sqsubseteq_{\mathcal{I}^*} g^*_x(\xi)).\end{array}$$

Now, by the Markov property of S

$$\begin{array}{l}P_\psi(Y_j=1|E_{\text{stop}}(\tau))=\\ P_\psi(S(T^1(j)-lm^2)=x-lm|E_{\text{stop}}(\tau)))\\ \times P_\psi(S(T^1(j))=x,\hat{q}^*(w^1(j))\leq q^*_x(\xi),\ \hat{g}^*(w^1(j))\sqsubseteq_{\mathcal{I}^*} g^*_x(\xi)|S(T^1(j)-lm^2)=x-lm)\\ \times P_\psi(S(T^3(j))=y|S(T^1(j))=x)\\ \times P_\psi(\hat{q}(w^3(j))\leq q_y(\xi),\ \hat{g}(w^3(j))\sqsubseteq_{\mathcal{I}} g_y(\xi)|S(T^3(j))=y).\end{array}$$

In Subsection 7.1.4, we showed

$$P_\psi\Big(S(T^1(j)-lm^2)=x-lm\Big|E_{\text{stop}}(\tau)\Big)\geq\exp(-1.5l_1)$$

By the same argument as in Subsection 7.1.4, we get for $\psi\in B^1_{\text{intervalsOK}}$,

$$\begin{array}{l}P_\psi\Big(S(T^1(j))=x,\ \hat{q}^*(w^1(j))\leq q^*_x(\psi),\ \hat{g}^*(w^1(j))\sqsubseteq_{\mathcal{I}^*_x(\psi)} g^*_x(\psi)|S(T^1(j)-lm^2)=x-lm^2\Big)\geq\\ \inf\limits_{\psi:\psi|[y,y+lm]\text{ is OK}}P_\psi\Big(S(T^3(j)+lm^2)=y+lm,\ \hat{q}(w^3(j))\leq q_y(\psi),\ \hat{g}(w^3(j))\sqsubseteq_{\mathcal{I}_y(\psi)} g_y(\psi)\Big|S(T^3(j))=y\Big).\end{array}$$

By Theorem 7.2.3, we have that the right side of the previous display is at least $\exp[-l\alpha_{II}]$. Hence, for $\psi \in B^1_{\text{intervals OK}}$, it holds

$$P_\psi\big(S(T^1(j)) = x, \hat{q}^*(w^1(j)) \le q^*_x(\psi),\ \hat{g}^*(w^1(j)) \sqsubseteq_{\mathcal{I}^*_x(\psi)} g^*_x(\psi) | S(T^1(j) - lm^2) = x - lm^2\big) \ge \exp[-l\alpha_{II}]$$

$$P_\psi\big(\hat{q}(w^3(j)) \le q_y(\psi),\ \hat{g}(w^3(j)) \sqsubseteq_{\mathcal{I}_y(\psi)} g_y(\psi) | S(T^3(j)) = y\big) \ge \exp[-l\alpha_{II}].$$

Finally,

$$P_\psi(S(T^3(j)) = y | S(T^1(j)) = x) = (p_L)^{c_1 l_1 L}.$$

This means, for $\psi \in B^1_{\text{intervals OK}}$

$$P_\psi(Y_j = 1 | E_{\text{stop}}(\tau)) \ge \exp[-1.5l_1]\exp[-2l\alpha_{II}](p_L)^{c_1 l_1 L} = p^*. \tag{7.3.27}$$

Conditional on $E_{\text{stop}}(\tau)$ and ψ, the random variables Y_i are independent. That follows from the definition of $E_{\text{stop}}(\tau)$). Hence

$$P_\psi\Big(\sum_{j=1}^{e^{\alpha l_1}} Y_j \le e^{\gamma l_1} \big| E_{\text{stop}}(\tau)\Big) \le P\Big(\sum_{j=1}^{e^{\alpha l_1}} Z_j \le e^{\gamma l_1}\Big) = P\Big(\sum_{j=1}^{e^{\alpha l_1}} (Z_j - p^*) \le e^{\gamma l_1} - e^{\alpha l_1} p^*\Big), \tag{7.3.28}$$

where Z_i are independent Bernoulli random variables with parameter p^*. By (7.3.24), the right side of (7.3.28) is bounded by

$$P\Big(\sum_{j=1}^{e^{\alpha l_1}} (Z_j - p^*) \le e^{\gamma l_1} - e^{\alpha l_1} p^*\Big) \le P\Big(\sum_{j=1}^{e^{\alpha l_1}} (Z_j - p^*) \le -e^{\gamma l_1}\Big).$$

Use Höffding's inequality to get

$$P\Big(\sum_{j=1}^{e^{\alpha l_1}} (Z_j - p^*) \le -e^{\gamma l_1}\Big) \le \exp[-2e^{(2\gamma - \alpha) l_1}].$$

Finally, integrate over $E^1_{\text{stop}}(\tau) \cap B^1_{\text{intervals OK}}$ and use (7.3.26) to deduce

$$P\Big(E^1_{\text{stop}}(\tau) \cap (E^1_{\text{enough times}}(x,y))^c \cap B^1_{\text{intervals OK}}\Big) \le \exp[-2\exp((2\gamma - \alpha) l_1)].$$

Sum over all pairs $(x, y) \in H_1$ to get (7.3.25). □

7.3.6 Tuning parameters

Recall that for big n, $\alpha_I > 8\alpha_{II}$.

- Choose n so big that statements of Theorem 7.2.1, Theorem 7.2.2, relation (7.2.23) and the statement of Corollary 7.2.1 hold.
- Then choose $c_1(n) > C_1(n)$, where $C_1(n)$ is specified in Proposition 7.3.1.

- Then choose $l_2(c_1, n)$ so big that simultaneously

$$\alpha_{II} l_2 > 1.5 + \ln 2 - c_1 \ln p_L \tag{7.3.29}$$
$$(\alpha_I - 7\alpha_{II}) l_2 > 9 \tag{7.3.30}$$
$$4\alpha_{II} l_2 > -2c_1 \ln p_{min} \tag{7.3.31}$$
$$a l_2 > 3 \tag{7.3.32}$$

- Then take $\gamma(n, c_1, l_2) = 4\alpha_{II} l_2$
- Then take $\alpha(n, c_1, l_2) = 7\alpha_{II} l_2$

7.3.7 Proof of the main theorem

Recall Lemma 7.3.1. By (12.4.12), (12.4.13) and (7.3.6), for l_1 big enough, it holds

$$P\Big(\big(E^1_{\text{alg works}}(\tau)\big)^c \cap E^1_{\text{stop}}(\tau)\Big) \leq$$
$$P\Big(\big(E^1_{\text{only ladders}}\big)^c \cap E^1_{\text{stop}}(\tau)\Big) + P\Big(\big(E^1_{\text{all ladders}}\big)^c \cap E^1_{\text{stop}}(\tau)\Big) + P\Big(\big(B^1_{\text{unique fit}}\big)^c\Big); \tag{7.3.33}$$

$$P\Big(\big(E^1_{\text{only ladders}}\big)^c \cap E^1_{\text{stop}}(\tau)\Big) \leq P\Big(\big(E^1_{\text{recon straight}}\big)^c \cap E^1_{\text{stop}}(\tau)\Big) \leq \tag{7.3.34}$$
$$P\Big(\big(E^1_{\text{recon straight}}\big)^c \cap E^1_{\text{stop}}(\tau) \cap B^1_{\text{recon straight}}\Big) + P\Big(\big(B^1_{\text{recon straight}}\big)^c\Big);$$

$$P\Big(\big(E^1_{\text{all ladders}}\big)^c \cap E^1_{\text{stop}}(\tau)\Big) \leq P\Big(\big(B^1_{\text{enough barriers}}\big)^c\Big)$$
$$+ P\Big(\big(E^1_{\text{no mistake}}\big)^c\Big) + P\Big(\big(B^1_{\text{enough times}}\big)^c \cap E^1_{\text{stop}}(\tau)\Big); \tag{7.3.35}$$

$$P\Big(\big(E^1_{\text{no mistake}}\big)^c\Big) \leq P\Big(\big(E^1_{\text{no mistake}}\big)^c \cap B^1_{\text{intervals OK}}\Big) + P\Big(\big(B^1_{\text{intervals OK}}\big)^c\Big); \tag{7.3.36}$$

$$P\Big(\big(B^1_{\text{enough times}}\big)^c \cap E^1_{\text{stop}}(\tau)\Big) \leq P\Big(\big(B^1_{\text{enough times}}\big)^c \cap E^1_{\text{stop}}(\tau) \cap B^1_{\text{intervals OK}}\Big) + P\Big(\big(B^1_{\text{intervals OK}}\big)^c\Big). \tag{7.3.37}$$

Recall the definitions of l_2. The condition (7.3.33) states $7\alpha_{II} l_2 > 4\alpha_{II} l_2 + 1.5 + \ln 2 - c_1 \ln p_L + 2\alpha_{II} l_2$ or, equivalently,

$$\alpha l_1 > (\gamma + 1.5 + \ln 2 - c_1 \ln p_L) l_1 + 2\alpha_{II} l.$$

Taking exponentials,

$$\exp(\alpha l_1) \exp(-1.5 l_1 - 2\alpha_{II} l)(p_L)^{c_1 l_1} > 2 \exp(\gamma l_1).$$

Recall the definition of p^* and note that the inequality in the previous display is (7.3.24). Hence, by Proposition 7.3.2, we have the bound (7.3.25). By definition of α and γ, we get $k_4 := \exp(2\gamma - \alpha) = \exp(\alpha_{II} l_2)$, implying that (7.3.25) is exponentially small in l_1. By

(7.3.32), there exist $k_5 > 0$ such that (7.3.10) is bounded by $4\exp[-k_5 l_1]$. With (7.3.25), we obtain that (7.3.37) is bounded by $68\exp[-(k_4 \wedge k_5)l_1]$.

Use (7.3.20) and (7.3.17) with (7.3.30) to obtain that $P(\big(E^1_{\text{no mistake}}\big)^c \leq 12\exp[(9+\alpha)l_1 - \alpha_I l] = 12\exp[-k_6 l_1]$ for a $k_6 > 0$. Hence, (7.3.36) is bounded by $12\exp[-k_6 l_1] + 4\exp[-k_5 l_1] \leq 16\exp[-(k_6 \wedge k_5)l_1]$.

By (7.3.13), we now get that (7.3.35) is bounded by $68\exp[-(k_4 \wedge k_5)l_1] + 16\exp[-(k_6 \wedge k_5)l_1] + \exp[-k_3 l_1] \leq 85\exp[-k_7 l_1]$ for a $k_7 > 0$.

The requirement (7.3.31) states that $\gamma + 2c_1 \ln p_{min} > 0$ implying that for a $k_8 > 0$

$$\exp\big[l_1[2(n^2 l_2 + l_2)\ln 2 + (4l_2)(\ln l_1 + \ln l_2)] - 2\exp((\gamma + 2c_1 \ln p_{min})l_1)\big] \leq \exp[-k_8 l_1]$$

for l_1 big enough. By (7.3.23), this means (7.3.34) is bounded by $\exp[-k_9 l_1]$ for l_1 big enough.

Finally, we get that (7.3.33) is bounded by $\exp[-kl_1]$, if l_1 is big enough. This proves Theorem 7.1.1.

7.4 Appendix

7.4.1 Proof of Theorem 7.2.1

Recall $m(n) > n$.
For each $i = 1, \ldots, l$ random cells $\xi_i = \xi | D_i = (\xi(d_{i-1}), \ldots, \xi(d_i))$.
Consider the event $E_{\text{OK}a}$. We can rewrite

$$E_{\text{OK}a} = \Big\{ \sum_{i=2Lm^2}^{l} X_i \leq l2\epsilon(n) \Big\},$$

where X_i is Bernoulli random variable that is one if and only iff ξ_i is not weak-OK. Let

$$l_* := Lm^2 + c + 2, \quad l^* = l - c + 1.$$

Then $(l_* - 1)m - cm = Lm^3 + m$ and $(l^* - 1)m + cm = lm$. Clearly $P(X_i = 1) \leq \epsilon(n)$, if $l_* \leq i \leq l^*$. If $i > l^*$, then, by definition, ξ_i cannot be weak-OK and, hence, $X_i = 1$. Now, let n be so big that $l_* \leq 2Lm^2$ i.e. $c + 2 \leq Lm^2$. This means, $E_{\text{OK}a}$ is independent on ξ^{Lm^3}. Then also $l - l^* = c - 1 \leq 2Lm^2$. Let us estimate

$$\begin{aligned}
E^c_{\text{OK}a} &= \Big\{ \sum_{i=2Lm^2}^{l} X_i > l2\epsilon(n) \Big\} \subseteq \Big\{ \sum_{i=2Lm^2}^{l-2Lm^2} X_i > l2\epsilon(n) - 2Lm^2 \Big\} \\
&\subseteq \bigcup_{j=-c+1}^{c} \Big\{ \sum_{k=k_*}^{k^*} X_{ik2c-j} > \frac{l2\epsilon(n) - 2Lm^2}{2c} \Big\} \\
&\subseteq \bigcup_{j=-c+1}^{c} \Big\{ \sum_{k=k_*}^{k^*} X_{k2c-j} - (k^* - k_* + 1)\epsilon(n) > \frac{l2\epsilon(n) - 2Lm^2}{2c} - \frac{l\epsilon(n)}{2c} \Big\}.
\end{aligned}$$

Here $k_* := \ulcorner \frac{2Lm^2+c}{2c} \urcorner$ and $k^* := \llcorner \frac{l-2Lm^2-c+1}{2c} \lrcorner$. Thus $k^*-k_* \le \frac{l-4Lm^2+1}{2c} < \frac{l}{2c}$, $k^*-k_*+1 < l$. Note, by definition $X_i \in \sigma(\xi_j | j = i-c, i-c+1, \dots, i+c-1)$. Thus, X_k and X_{k2c} are independent. This means, for each j we can apply Höffding's inequality. Thus, for each j

$$P\Big(\sum_{k=k_*}^{k^*}(X_{k2c-j}-\epsilon(n)) > \frac{l\epsilon(n)-2Lm^2}{2c}\Big) \le P\Big(\sum_{k=k_*}^{k^*}(X_{k2c-j}-EX_{k2c-j}) > \frac{l\epsilon(n)-2Lm^2}{2c}\Big)$$
$$\le \exp\Big[-\frac{(l\epsilon(n)-2Lm^2)^2}{c(k^*-k_*)}\Big] \le \exp[-\frac{l\epsilon^2(n)}{2c}],$$

provided l is big enough to satisfy $l\epsilon(n) - 2Lm^2 \ge l\frac{\epsilon(n)}{2}$. Hence,

$$P(E_{\mathrm{OK}_a}^c) \le 2c\exp\Big(\frac{-\epsilon^2(n)l}{2c}\Big) \le \exp(-a_1(n)l), \tag{7.4.1}$$

for some $a_1(n) > 0$, provided l is big enough.

We estimate $P(E_{\mathrm{OK}_b}^c)$ by the same argument. Define

$$E_{\mathrm{OK}b}^{i*} := \left\{ \left|\mathcal{I}_{II}^i(\xi^{ml})\right| \ge l(1-\exp(-m^{0.8}))\right\}, \quad i = 1, 2.$$

Clearly, for n big enough,

$$E_{\mathrm{OK}b}^{1*} \cap E_{\mathrm{OK}b}^{2*} \subseteq E_{\mathrm{OK}b} \quad \text{and } \mathrm{P}(E_{\mathrm{OK}}{}^c) \le P(E_{\mathrm{OK}b}^{1*}) + P(E_{\mathrm{OK}b}^{2*}). \tag{7.4.2}$$

Let us estimate $P(E_{\mathrm{OK}b}^{2*})$.
Let Y_i be Bernoulli random variable that is 1 if and only if ξ_i has not empty neighborhood. Let us estimate $P(Y_i = 1)$. If $d_{i-1} - Lm^2 \ge 0$ and $d_i + Lm^2 \le lm$, then

$$P(Y_i = 1) = \big(\exists j \in [d_{i-1}-Lm^2, d_i+Lm^2] : \xi(j) = \cdots = \xi(j+m^{0.9})\big)$$
$$\le (2Lm^2+m+1)(0.5)^{m^{0.9}} \le \exp(-m^{0.85}),$$

in m is big. Otherwise, by definition, $Y_i = 1$. Let N be such that the inequality above holds as well as (7.4.2) if $n > N$. Note that $E_{\mathrm{OK}b}^{2*}$ is independent of ξ^{Lm^3}.
Clearly $Y_i \in \sigma(\xi_{i-Lm}, \dots, \xi_{i+Lm})$. Hence Y_i and $Y_{i+2+2Lm}$ are independent. Let $k = 2(1+Lm)$. Now with $i^* = \llcorner \frac{l-2Lm^2-k+1}{k} \lrcorner$ and $i^* \le \frac{l}{k}$ we get

$$E_{\mathrm{OK}b}^{2*\,c} = \Big\{\sum_{i=2Lm^2}^{l} Y_i > l\exp(-m^{0.8})\Big\} \subseteq \Big\{\sum_{i=2Lm^2}^{l-2Lm^2} Y_i > l\exp(-m^{0.8}) - 2Lm^2\Big\}$$
$$\subseteq \bigcup_{j=0}^{k-1}\Big\{\sum_{i=0}^{i^*} Y_{2Lm^2+j+ik} > \frac{l\exp(-m^{0.8})-2Lm^2}{k}\Big\}$$
$$\subseteq \bigcup_{j=0}^{k-1}\Big\{\sum_{i=0}^{i^*} Y_{2Lm^2+j+ik} - i^*\exp(-m^{0.85}) > \frac{l(\exp(-m^{0.8})-\exp(-m^{0.85}))-2Lm^2}{k}\Big\}$$
$$\subseteq \bigcup_{j=0}^{k-1}\Big\{\sum_{i=0}^{i^*} (Y_{2Lm^2+j+ik} - EY_{2Lm^2+j+ik}) > \frac{l(\exp(-m^{0.8})-\exp(-m^{0.85}))-2Lm^2}{k}\Big\}.$$

Denote $\exp(-m^{0.8}) - \exp(-m^{0.85}) =: e(m)$ and apply Höffdings inequality

$$P\Big(\sum_{i=0}^{i^*}(Y_{2Lm^2+j+ik} - EY_{2Lm^2+j+ik}) \geq \frac{le(m) - 2Lm^2}{k}\Big) \leq \exp\Big[-\frac{2(le(m) - 2Lm)^2}{lk}\Big] \leq \exp[-a_2(m)l],$$

for some $a_2(m) > 0$, if l is sufficiently big. Now, for big l,

$$P(E_{\mathrm{OK}b}^{2*\,c}) \leq 2(k+1)\exp(-a_2(m)l) \leq 2(m+1)\exp(-a_2(m)l) \leq \exp(-a_3(m)l),$$

for some $a_3(m) > 0$.

Similarly we estimate $P(E_{\mathrm{OK}b}^{1*})$.
Let Z_i be Bernoulli random variable that is 1 if and only if ξ_i is not isolated. If $i \geq l - Lm$, then, by definition $Z_i = 1$. Thus

$$E_{\mathrm{OK}b}^{1*\,c} = \{\sum_{i=2Lm^2}^{l} Z_i > l\exp(-m)\Big\} \subseteq \Big\{\sum_{i=2Lm^2}^{l-Lm} Z_i > l\exp(-m) - Lm\Big\}.$$

Again, $E_{\mathrm{OK}b}^{1*}$ is independent on ξ^{Lm^3}. Note, if $\sum_{i+2Lm^2}^{l} Z_i > l\exp(-m) - Lm$, then among the vectors $\{\xi_{2Lm^2-Lm-1}, \xi_{2Lm^2-Lm}, \ldots, \xi_l\}$ there exists at least $\frac{1}{2}(l\exp(-m) - Lm - 1)$ intervals ξ_i without fence.
Let Z_i' Bernoulli random variable that is 1 if and only if the random vector (but not the cell) $\xi|(d_{i-1}, d_i)$ does not contain a fence. Since the intervals (d_{i-1}, d_i) and (d_{j-1}, d_j) $(i \neq j)$ are disjoint, Z_i' are iid. random variables. Hence, with $j^* = 2Lm^2 - Lm - 1$, we get

$$P(E_{\mathrm{OK}b}^{1*\,c}) \leq P\Big(\sum_{j=j^*}^{l} Z_j' > \frac{1}{2}(l\exp(-m) - Lm - 1)\Big).$$

Clearly

$$P(Z_i' = 1) = P(\xi|(d_{i-1}, d_i) \text{ contains no fence}) \leq (1 - (0.5)^{2L-1})^{\frac{m-2}{2L}} < e^{-cm},$$

for some $c > 0$. Now Höffding's inequality yields

$$P\Big(\sum_{j=j^*}^{l} Z_j' \geq \frac{1}{2}(le^{-m^{0.8}} - Lm)\Big) \leq P\Big(\sum_{j=1}^{l} Z_j' - le^{-cm} \geq \frac{1}{2}(le^{-m^{0.8}} - Lm) - le^{-cm}\Big) =$$
$$P\Big(\sum_{j=1}^{l}(Z_j' - EZ_j') > \frac{1}{2}l(e^{-m^{0.8}} - 2e^{-cm}) - \frac{L}{2}m\Big) \leq \exp\Big[-\frac{(l(e^{-m^{0.8}} - 2e^{-cm}) - Lm)^2}{2l}\Big]. \tag{7.4.3}$$

The right side of (7.4.3) is bounded by $\exp(-la_4(m))$, for some $a_4(m) > 0$, provided l is big enough.
Now, there exists $a_5(m) > 0$ such that for big l,

$$P(E_{\mathrm{OK}_b^c}) \leq \exp(-a_3 l) + \exp(-a_4 l) \leq \exp(-a_5 l) \tag{7.4.4}$$

Now, by (7.2.2), (7.4.1), (7.4.4)

$$P(E_{\mathrm{OK}}{}^c) \leq P(E_{\mathrm{OK}_a^c}) + P(E_{\mathrm{OK}_b^c}) \leq \exp(-la_1) + \exp(-la_5) \leq \exp(-la),$$

for some $a(m) > 0$ and big l.

7.4.2 Proof of Proposition 7.2.1

By definition,

$$F_{min}(j) \in \sigma\Big(S(t) - S(t-1)\Big| t \in [1, (s_j - r_j)m]\Big).$$

This means, if $F_{min}(j) \neq \varnothing$, then $P_\psi(F_{min}(j)) \geq (p_{min})^{(s_j - r_j)m}$. We shall show that $F_{min}(j) \neq \varnothing$.
Let $j \in \{1, \dots, k\}$. Let us describe an admissible path $R := R_j \in \mathcal{R}((s_j - r_j)m)$ such that simultaneously satisfies **R1**, **R2**, **R3**. If such a path exists, then (7.2.13) holds.

Consider an arbitrary index-interval $[l_{2j-1}, l_{2j}]$, $j > 1$. It corresponds to the location-interval $[r_j, s_j]$. Let $C_1 < \dots < C_q$ be the big clusters of ψ in $[s_j, r_j]$. Denote by c_i, d_i, $i = 1, \dots, q$ the beginnings and ends of big clusters, respectively. Hence, $C_i \subseteq [c_i, d_i]$. The path R should read the big clusters as one block, i.e. along the reading-path.
Moreover, let $B_1 < B_2 < \dots < B_p$ be the blocks of ψ in the set $[s_j, r_j] \backslash (\cup_{i=1}^{q} [c_i + 2, d_i - 2])$ that are bigger than $m^2/2\bar{v}$. By definition, $l(B_i) < m^3$, $i = 1, \dots, p$. Indeed, if $l(B_i) \geq m^3$, then B_i would be a (part of) big cluster. We refer to a B_i as a **small block**. The small blocks should be crossed as shortly as possible, i.e. along the reading path.
Finally let $A_1 < A_2 < \dots < A_K$, $K = p + q$ be the ordered big clusters and small blocks. Let a_i, b_i denote (an arbitrary) reading-beginning and reading-end of A_i.
Since $j > 1$, it holds $l_{2j-1} \in \mathcal{I}_{II}$. Then D_{2j-1} has empty neighborhood, hence $[r_j, r_j + Lm^2]$ is empty (for ψ) and, therefore, does not contain any small blocks. Also D_{2j-1} is isolated. This implies that there is no point in $[r_j, r_j + Lm^2]$ that is connected with any point in $[r_j + Lm^2 + m, s_j]$. In particular, all objects $A_1, \dots, A_K$ are outside of $[r_j, r_j + Lm^2]$ or, formally, $a_1 > r_j + Lm^2$.
If $s_j - r_j \leq 2Lm^2$, then the interval does not contain blocks that are bigger than $m^{0.9}$. In this case R starts at r_j, i.e. $R(0) = r_j$ and goes to the point s_j with $(l_{2j} - l_{2j-1} + 1)m^2$ step without generating more than $m\bar{v}$ consecutive same colors in observations. This is clearly possible.
If $s_j - r_j > 2Lm^2$, then we define the minimum-blocks path R for interval $[r_j, s_j]$ backwards. More precisely, we define or prescribe a path R^* that starts at s_j and goes to r_j with $(s_j - r_j)m^2$ steps. The prescription of R^* is the following: start at s_j, i.e. $R^*(0) = s_j$. Then move stepwise to b_K (recall, this is a reading-end of the last small block or the last big cluster in $[r_i, s_j]$). Recall $s_j = l_{2j}m$ If $s_j \neq l$, then $l_{2j} \in \mathcal{I}_{II}$ and $[s_j - Lm^2, s_j + m]$ is empty and $[s_j - Lm^2 - m, s_j - Lm^2]$ contains a fence. As explained above, this implies that $b_K \leq s_j - Lm^2$. So, by moving stepwise from s_j to b_K, it is not possible that S generates more than $m^{0.9}\bar{v}$ same colors in the beginning.
After reaching b_K move along the reading path to a_K. Then move stepwise to b_{K-1}. Continue so until a_1 and then go stepwise until $r_j + Lm^2$. Since $a_1 > r_j + Lm^2$, for such a path less than $((s_j - r_j) - Lm^2)\bar{v}$ steps are needed. This means that the path has more than $(s_j - r_j)(m - \bar{v}) + Lm^2\bar{v}$ steps to cover the interval $[r_j, r_j + Lm^2]$ with length Lm^2 without generating more than $m\bar{v}$ consecutive same colors in observations and satisfying $R^*((s_j - r_j)m) = r_j$. This is obviously possible, because the interval does not contain more than $m^{0.9}$ consecutive same colors. Finally define R as R^* backwards, i.e. $R(0) = R^*((s_j - r_j)m) = r_j$, $R(1) = R^*((s_j - r_j)m - 1), \dots, R(i) = R^*((s_j - r_j)m - i), \dots, R((s_j - r_j)m) = R^*(0) = s_j$ (recall, S is symmetric).
Such definition of R_j ensures that **R1** and **R3** are met. Let us show that **R2** holds as well.

Note that the number of big blocks in $\psi \circ R$ is equal with the number of big clusters in $[r_j, s_j]$. Let this number be M. That means

$$\hat{q}_V(\psi \circ r_j) = q_V([r_j, s_j]) = M,$$

where $V := l_{2j} - l_{2j-1} + 1$. Let

$$T(i) := \inf\{k : q_k(\psi|[r_j, s_j]) = i\}, \quad \hat{T}(i) := \inf\{k : \hat{q}_k(\psi \circ r_j) = i\} \quad i = 1, \ldots, M.$$

R2 is violated, if there exists $i \in \{1, \ldots, M\}$ such that $\hat{T}(i) < T(i)$. Fix an $i \in \{1, \ldots, M\}$. The inequality $\hat{T}(i) < T(i)$ means that after reading the i-th big cluster, R has more than $(V - T(i) + 1)m^2$ steps to go to s_j. However, the path R is constructed such that after reaching to the b_i we have at most $(V - T(i) + 1)m\bar{v}$ step to go s_j. That proves **R2**.

Finally consider the first interval $[r_1, s_1] = [0, s_1]$ (obviously, $r_1 = 0$). Since $l_1 = 1 \notin \mathcal{I}_{II}$, the interval $[0, Lm^2]$ is not necessarily empty. And $[Lm^2, Lm^2 + m]$ does not necessarily contain a fence. This means that it might be not possible to go from a_1 to 0 without generating more than $m\bar{v}$ consecutive same colors in observations and satisfying $R^*(s_1 m) = 0$. However, it is clearly possible to go from a_1 to 0 without generating any big block in observations. So, for R_1, the description of reverse-path, R^* ends: go from a_1 to 0 without generating any big block in the observations. For example, if $\psi(0) = \psi(1) = \cdots = \psi(Lm^3) = 1$, then the reverse of the minimum-block path, R^*, states that S goes to 0 (with suitable many steps, satisfying $R^*(s_1 m^2) = 0$) by generating only one's. Thus, if R_1 and $\psi(0) = \psi(1) = \cdots = \psi(Lm^3) = 1$ hold, then $\psi \circ R_1$ starts with at least m^3 consecutive ones but it does not start with a big block. This means that **R2** still holds.
Hence, $F_{min}(j) \neq \varnothing$ for each $j = 1, \ldots, k$.

References

[1] Itai Benjamini and Harry Kesten.
Distinguishing sceneries by observing the sceneries along a random walk path.
J. Anal. Math 69, 97-135, 1996

[2] Krzysztof Burdzy.
Some path properties of iterated Brownian motion.
In *Seminar on Stochastic Processes, 1992 (Seattle, WA, 1992)*,
pages 67–87. Birkhäuser Boston, Boston, MA, 1993.

[3] Frank den Hollander.
Mixing properties for random walk in random scenery.
Ann. Probab. 16(4), 1788–1802, 1988.

[4] Frank den Hollander and Jeffrey E. Steif.
Mixing properties of the generalized T, T^{-1}-process.
J. Anal. Math., 72, 165–202, 1997.

[5] Matthew Harris and Michael Keane.
Random coin tossing.
Probab. Theory Related Fields, 109(1), 27–37, 1997.

[6] Deborah Heicklen, Christopher Hoffman and Daniel J. Rudolph.
Entropy and dyadic equivalence of random walks on a random scenery.
Adv. Math., 156(2), 157–179, 2000.

[7] C. Douglas Howard.
Detecting defects in periodic scenery by random walks on $\mathbb{Z}$.
Random Structures Algorithms, 8(1), 59–74, 1996.

[8] C. Douglas Howard.
Orthogonality of measures induced by random walks with scenery.
Combin. Probab. Comput., 5(3), 247–256, 1996.

[9] C. Douglas Howard.
Distinguishing certain random sceneries on $\mathbb{Z}$ via random walks.
Statist. Probab. Lett., 34(2), 123–132, 1997.

[10] Steven Arthur Kalikow.
T, T^{-1} transformation is not loosely Bernoulli.
Ann. of Math. (2), 115(2), 393–409, 1982.

[11] M. Keane and W. Th. F. den Hollander.

Ergodic properties of color records.

Phys. A, 138(1-2), 183–193, 1986.

[12] Harry Kesten.

Detecting a single defect in a scenery by observing the scenery along a random walk path.

In *Itô's stochastic calculus and probability theory*,

pages 171–183. Springer, Tokyo, 1996.

[13] Harry Kesten.

Distinguishing and reconstructing sceneries from observations along random walk paths.

In *Microsurveys in discrete probability (Princeton, NJ, 1997)*,

pages 75–83. Amer. Math. Soc., Providence, RI, 1998.

[14] H. Kesten and F. Spitzer.

A limit theorem related to a new class of self-similar processes.

Z. Wahrsch. Verw. Gebiete

50(1), 5–25, 1979.

[15] J. Lember and H.Matzinger

A location test for observations.

Eurandom Report 2002-014, Eurandom, 2002.

[16] Arnoud Le Ny and Frank Redig.

Reconstruction of sceneries in the Gibbsian context.

In preparation, Eurandom, 2002.

[17] D.A. Levin, R. Pemantle and Y. Peres.

A phase transition in random coin tossing.

Preprint, 2001.

[18] Elon Lindenstrauss.

Indistinguishable sceneries.

Random Structures Algorithms, 14(1), 71–86, 1999.

[19] M. Löwe and H. Matzinger.

Scenery reconstruction in two dimensions with many colors.

Preprint: Eurandom Report 99-018, Eurandom, 1999.

Submitted to *The Annals of Applied Probability*.

[20] M. Löwe, H. Matzinger and F. Merkl.

Reconstructing a multicolor random scenery seen along a random walk path with bounded jumps.

Eurandom Report 2001-030, Eurandom, 2001.

Submitted to *Transaction of the American Mathematical Society*.

[21] H. Matzinger.

Reconstructing a 2-color scenery by observing it along a simple random walk path with holding.

PhD-thesis, Cornell University, 1999.

[22] H. Matzinger.

Reconstructing a 2-color scenery by observing it along a simple random walk path.

Eurandom Report 2000-003, Eurandom, 2000.

Submitted to *The Annals of Applied Probability*.

[23] H. Matzinger.

Reconstructing a 2-color scenery in polynomial time by observing it along a simple random walk path with holding.

Eurandom Report 2000-002, Eurandom, 2000.

Submitted to *Probability Theory and Related Fields*.

[24] H. Matzinger and S. Rolles.

Reconstructing a random scenery observed with random errors along a random path.

Preprint 2001.

Chapter 8

Retrieving random media

(submitted)

By Heinrich Matzinger and Silke Rolles

Benjamini asked whether the scenery reconstruction methods of Matzinger (see e.g. [19], [20], [18]) can be done in polynomial time. In this article, we give the following answer for a 2-color scenery and simple random walk with holding: We prove that a piece of the scenery of length of the order 3^n around the origin can be reconstructed – up to a reflection and a small translation – with high probability from the first $2 \cdot 3^{10\alpha n}$ observations with a constant $\alpha > 0$ independent of n. Thus, the number of observations needed is polynomial in the length of the piece of scenery which we reconstruct. The probability that the reconstruction fails tends to 0 as $n \to \infty$.

In contrast to [19], [20], and [18], the proofs in this article are all constructive. Our reconstruction algorithm is an algorithm in the sense of computer science. This is the first article which shows that the scenery reconstruction is also possible in the 2-color case *with holding*. The case with holding is much more difficult than [20] and requires a completely different methods. [1]

8.1 Introduction and Result

A *scenery* is a coloring of $\mathbb{Z}$ with finitely many colors. We call two sceneries ξ and ξ' *equivalent*, $\xi \approx \xi'$, if $\xi = \xi' \circ T$ where T is a translation, a reflection, or the composition of both. Let $S := (S_k)_{k \in \mathbb{N}_0}$ be a recurrent random walk on $\mathbb{Z}$. Observing the scenery along the random walk path, we obtain the color record $\chi := (\chi_k := \xi(S_k))_{k \in \mathbb{N}_0}$. The *scenery reconstruction problem* asks the following question: Given the color record χ, can we reconstruct the scenery ξ up to equivalence?

Early questions about random sceneries were raised by Benjamini and Kesten and, independently, by Keane and den Hollander. For the history of the problem we refer the reader to the survey paper of Kesten [12]. Early work on random sceneries include articles of Benjamini and Kesten [1], den Hollander [4], Howard ([7], [8], [9]), Keane and den Hollander [10], Kesten [11], and Lindenstrauss [15]. More recent contributions are due to Burdzy [2], Heicklen, Hoffman, and Rudolph [6], den Hollander and Steif

[1] *MSC 2000 subject classification:* Primary 60K37, Secondary 60G10, 60J75.
Key words: Scenery reconstruction, jumps, stationary processes, random walk, ergodic theory.

[3], Levin, Pemantle and Peres [14], Levin and Peres [13]. We refer the reader to the introductions of [**?**] and [22] for more details. Various contributions to the subject of scenery reconstruction have been made by Matzinger ([19],[20]), Löwe and Matzinger ([17], [16]), Löwe, Matzinger, and Merkl [18], Matzinger and Rolles [22]. In these papers, the scenery is taken random, independent of the random walk, and it is shown that for almost all realizations of the random walk path, almost all sceneries can be reconstructed up to equivalence.

The scenery reconstruction algorithms in [19], [20], [17], [16], [18], and [22] do not work in polynomial time. Benjamini asked whether some of these reconstructions can be done in polynomial time. In this article, we give the following answer to Benjamini's question: Let $\xi := (\xi_k)_{k\in\mathbb{Z}}$ with ξ_k i.i.d. uniform on $\{0,1\}$, and let $S = (S_k)_{k\in\mathbb{N}_0}$ be a simple random walk with holding on $\mathbb{Z}$, independent of ξ. We prove that in order to reconstruct – up to a reflection and a small translation – with high probability a piece of scenery of length of the order 3^n around the origin, we need only the observations up to time $p(3^n)$ with a polynomial p, independent of n.

In order to reconstruct the whole scenery, we need infinitely many observations because the scenery is infinite. In finite time, we can never reconstruct with probability 1 a piece of scenery of length ≥ 2. As a matter of fact, the random walk stays with positive probability at the origin. Hence, we mean by reconstruction in polynomial time that there exist algorithms $\mathcal{A}_n$, $n \geq 1$, with the following properties: $\mathcal{A}_n$ obtains as input finitely many observations, namely $\chi|[0, 2\cdot 3^{10\alpha n}[$ with a constant $\alpha > 0$ and produces an output of length of the order 3^n. The probability that the reconstruction succeeds, in the sense that the output is – up to a reflection and a small translation – a piece of the scenery around the origin, tends to 1 as $n \to \infty$. The number of observations needed is polynomial in the length of the reconstructed piece of scenery. Since the scenery is assumed to be i.i.d., with probability 1 every finite piece of scenery occurs somewhere in the scenery. Thus it is crucial to reconstruct something close to the origin.

Formally, our result can be described as follows: Let $\mathcal{C} := \{0,1\}$ denote the set of colors. For two pieces of scenery ψ and ψ' (not necessarily of the same length), we write $\psi \preceq \psi'$ if ψ is up to a possible reflection contained in ψ'. We prove:

Theorem 8.1.1. *There exist constants $\alpha, c_3, c_4, c_{25} > 0$ and maps $\mathcal{A}_n : \mathcal{C}^{2\cdot 3^{10\alpha n}} \to \mathcal{C}^{[-3\cdot 3^n, 3\cdot 3^n]}$, $n \geq c_3$, which are measurable with respect to the canonical σ-algebras, such that for all $n \geq c_3$ the event*

$$E_n := \left\{\xi|[-3^n, 3^n] \preceq \mathcal{A}_n(\chi|\left[0, 2\cdot 3^{10\alpha n}\right[) \preceq \xi|[-4\cdot 3^n, 4\cdot 3^n]\right\}$$

satisfies $P\left([E_n]^c\right) \leq c_4 \exp\left(-c_{25} n^{0.2}\right)$.

As a consequence of Theorem 8.1.1 the whole scenery can be reconstructed almost surely:

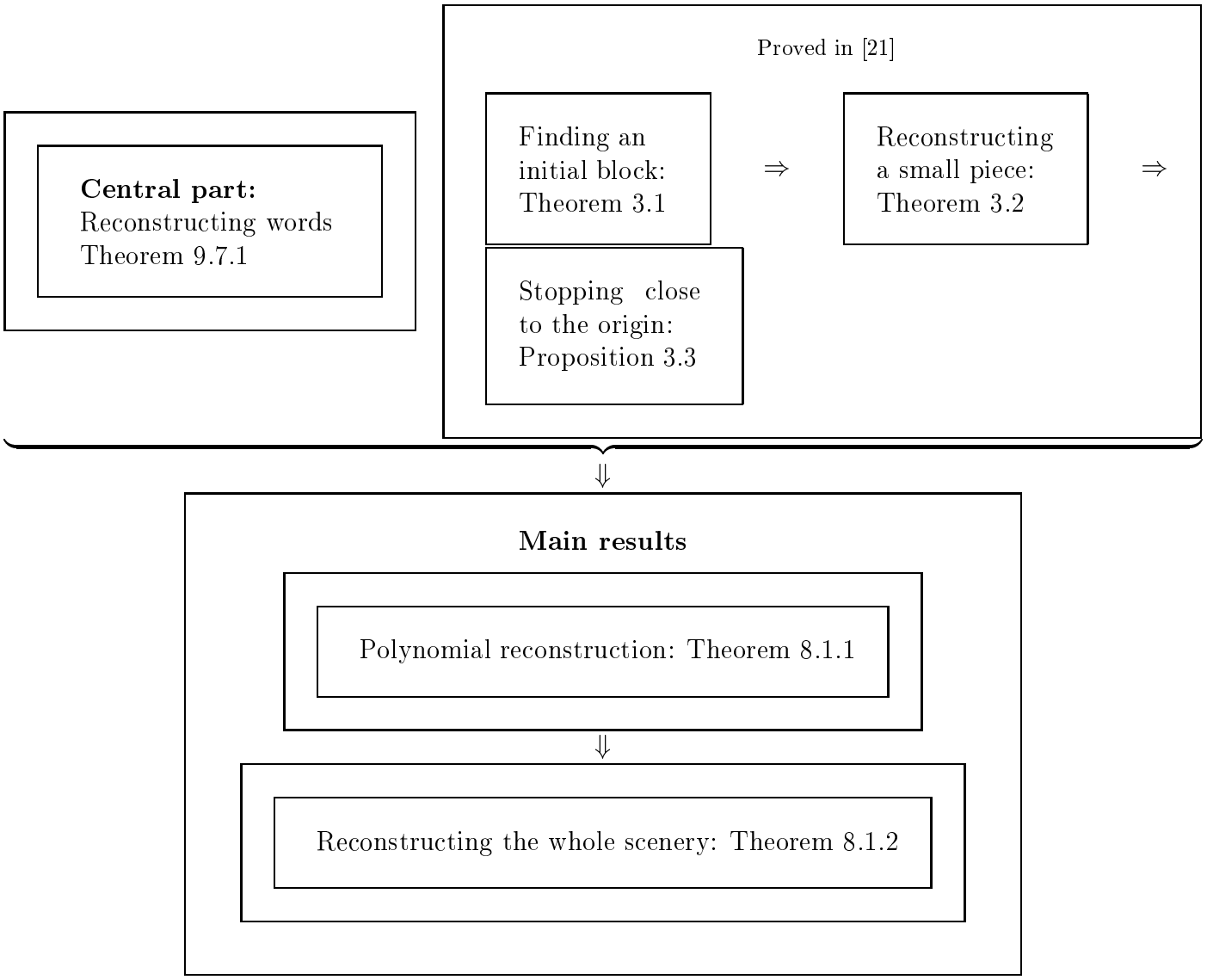
Theorem 8.1.2. *There exists a map $\mathcal{A} : \mathcal{C}^{\mathbb{N}_0} \longrightarrow \mathcal{C}^{\mathbb{Z}}$, which is measurable with respect to the canonical σ-algebras, such that $P\left(\mathcal{A}(\chi) \approx \xi\right) = 1$.*

The present article is the first article which solves the scenery reconstruction problem in the case of two colors and simple random walk with holding. We call this case the *statistical case*. On the piece of scenery 01, the random walker can produce every pattern by jumping back and forth or staying. Thus, at many places in the scenery, the random

walk can produce every possible pattern in the observations. This makes the statistical case much more difficult than the *combinatorial case*, where with high probability words of length $c_1 n$ (with a constant $c_1 > 0$) are characteristic for certain parts of the scenery. Two colors and simple random walk [20], many colors and simple random walk on $\mathbb{Z}^2$ [17], enough colors and random walk on $\mathbb{Z}$ with jumps [18] are examples of the combinatorial case. In the statistical case, it is much more difficult than in the combinatorial case to reconstruct small pieces of the scenery. The methods used below are completely different from the techniques developed in earlier articles.

The remainder of the article is organized as follows: Section 9.2 collects some notation. In Section 8.3, we show how Theorem 8.1.2 follows from Theorem 8.1.1. Since the definition of the maps $\mathcal{A}_n$ which fulfill the claim of Theorem 8.1.1 is quite involved, the construction is split in several steps. In Section 8.3, we state the results needed in the construction of the $\mathcal{A}_n$. The crucial step consists in finding small words in the scenery; this is done in Section 8.4. The second important step is the construction of a partial reconstruction algorithm BigAlg^n which is treated in Section 8.5. In addition, we need a small piece of the scenery to get the reconstruction started and also sequences of stopping times indicating when the random walker is close to the origin. These results are proved in [21]. At the end of Section 8.3, we show how the results of Sections 8.4 and 8.5 together with the results from [21] imply Theorem 8.1.2.

The following diagram is a guide to the proofs of Theorems 8.1.2 and 8.1.1:



8.2 Notation

In this section, we collect some notations and conventions.

Numbers, sets, and functions: We denote by $\mathbb{N} := \{1, 2, 3, \ldots\}$ the set of natural numbers and set $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. If $x \in \mathbb{R}$, we denote by $\lfloor x \rfloor$ the largest integer $\leq x$. We write $x \wedge y$ for the minimum of $x, y \in \mathbb{R}$. For a vector $y = (y_k)_{k\in[1,m]} \in \mathbb{R}^m$ we define the l^1-norm $\|y\|_1 := \sum_{k=1}^m |y_k|$ and the l^2-norm $\|y\|_2 := (\sum_{k=1}^m [y_k]^2)^{1/2}$. The cardinality of a set D is denoted by $|D|$. We write $f|D$ for the restriction of a function f to a set D. An *integer interval* is a set of the form $I \cap \mathbb{Z}$ with an interval $I \subseteq \mathbb{R}$. In this article, intervals are always taken over the integers, e.g. $[a, b] = \{z \in \mathbb{Z} : a \leq z \leq b\}$.

Admissible paths: Let $I = [i_1, i_2]$ be an integer interval. We call $R \in \mathbb{Z}^I$ an *admissible piece of path* if $R_{i+1} - R_i \in \{-1, 0, 1\}$ for all $i \in [i_1, i_2 - 1]$. We call R_{i_1} the starting point, R_{i_2} the endpoint, and $|I|$ the length of R.

Measures: We define δ_x to be the Dirac measure in x. We denote the image of a measure Q under a map F by QF^{-1}.

Sceneries: We denote by $\mathcal{C} := \{0,1\}$ the set of colors. A *scenery* is an element of $\mathcal{C}^{\mathbb{Z}}$. Let $I \subseteq \mathbb{Z}$ be an integer interval. An element of $\mathcal{C}^I$ is a *piece of scenery* or a *word*. If $\psi \in \mathcal{C}^I$, we call $|I|$ the *length* of ψ and denote it by $|\psi|$. We write $(1)_I$ for the piece of scenery in $\mathcal{C}^I$ which is identically equal to 1.

Blocks: Let $a, b \in I$ with $a < b$ and $|a-b| \geq 2$. We define $\psi \in \mathcal{C}^{[a,b]}$ to be a *block* if $\psi_a = \psi_b$ and $\psi_c \neq \psi_a$ for all $c \in]a,b[$. ψ_c is the *color* of the block. We call a the *left endpoint*, b the *right endpoint*, and $|\psi| := b-a-1$ the *blocklength* of ψ. For instance, 01110 is a block of length 3. We set $\partial\psi := \{a,b\}$.

Let $\chi|[t_1,t_2]$ and $\xi|[a,b]$ be blocks. We say that $\chi|[t_1,t_2]$ *is generated by the random walk S on the block $\xi|[a,b]$* if $\{S_{t_1}, S_{t_2}\} \subseteq \{a,b\}$ and $S_t \in]a,b[$ for all $t \in]t_1,t_2[$.

Equivalence of sceneries: Let $\psi \in \mathcal{C}^I$ and $\psi' \in \mathcal{C}^{I'}$ be two pieces of scenery. We say that ψ and ψ' are *equivalent* and write $\psi \approx \psi'$ iff I and I' have the same length and there exists $a \in \mathbb{Z}$ and $b \in \{-1,1\}$ such that for all $k \in I$ we have that $a+bk \in I'$ and $\psi_k = \psi'_{a+bk}$. We call ψ and ψ' *strongly equivalent* and write $\psi \equiv \psi'$ if $I' = a+I$ for some $a \in \mathbb{Z}$ and $\psi_k = \psi'_{a+k}$ for all $k \in I$. We say ψ *occurs in* ψ' and write $\psi \sqsubseteq \psi'$ if $\psi \equiv \psi'|J$ for some $J \subseteq I'$. We write $\psi \preceq \psi'$ if $\psi \approx \psi'|J$ for some $J \subseteq I'$. If the subset J is unique, we write $\psi \preceq_1 \psi'$.

Random walks and random sceneries: Let $\Omega_2 \subseteq \mathbb{Z}^{\mathbb{N}_0}$ denote the set of admissible paths. Let $p, q > 0$ satisfy $2p+q = 1$. We denote by Q_x the distribution on Ω_2 of a random walk $(S_k)_{k\in\mathbb{N}_0}$ starting at x with i.i.d. increments distributed according to $p\delta_{-1}+q\delta_0+p\delta_1$, i.e. S is a simple random walk with holding, and satisfies

$$\begin{aligned} p &= P(S_{k+1}-S_k = 1) = P(S_{k+1}-S_k = -1), \\ q &= P(S_{k+1}-S_k = 0) \end{aligned}$$

for all $k \geq 0$. The scenery $\xi := (\xi_k)_{k\in\mathbb{Z}}$ is i.i.d. with $P(\xi_k = 0) = P(\xi_k = 1) = 1/2$. We assume that ξ and S are independent and realized as canonical projections on $\Omega := \mathcal{C}^{\mathbb{Z}} \times \Omega_2$ with the product σ-algebra generated by the canonical projections and probability measures $P_x := (\frac{1}{2}\delta_0 + \frac{1}{2}\delta_1)^{\otimes\mathbb{Z}} \otimes Q_x$, $x \in \mathbb{Z}$. We abbreviate $P := P_0$. We call $\chi := (\chi_k := \xi(S_k))_{k\in\mathbb{N}_0}$ the *scenery observed along the random walk path*; sometimes we write $\xi \circ S$ instead of χ.

For a fixed scenery $\xi \in \mathcal{C}^{\mathbb{Z}}$ we set $P_{x,\xi} := \delta_\xi \otimes Q_x$, $P_\xi := P_{0,\xi}$. Thus $P_{x,\xi}$ is the canonical version of the conditional probability $P_x(\cdot|\xi)$, the distribution P conditioned on the random walk to start in x and the scenery ξ. We never work with a different version of the conditional probability $P_x(\cdot|\xi)$.

Filtration: We define $\mathcal{G} := (\mathcal{G}_n)_{n\in\mathbb{N}_0}$ with $\mathcal{G}_n := \sigma(\chi_k; k \in [0,n])$ to be the natural filtration of the observations over Ω.

Shifts: We define the shift $\theta : \mathcal{C}^{\mathbb{N}_0} \to \mathcal{C}^{\mathbb{N}_0}$, $\eta \mapsto \eta(\cdot+1)$. We introduce the shift $\Theta : \Omega \to \Omega$, $(\xi, S) \mapsto (\xi(S_1+\cdot), S(1+\cdot)-S_1)$. For a set $A \subseteq \Omega$ and a random time $T \geq 0$ we set $\Theta^{-T}(A) := \{\omega : \Theta^{T(\omega)}(\omega) \in A\}$.

Constants: We denote constants by c_i, $i \geq 1$; they keep their meaning throughout the whole article. Constants c_1, c_2, c_6, c_7, and α play a special role. They are chosen as follows:

- $c_2 > 21$,
- $c_1 \in 4\mathbb{N}$ with $c_1 > \max\{153, 4c_2\}$,
- $c_6 > (c_1+4)\ln 3$,

- $c_7 > \max\{0, 2\ln 3 - 2c_1 \ln p + 2c_6 + 2c_1 \ln[\max_{i\in[1,5]} \|x_i^*\|_2]\}$ with x_i^* as in Definition 9.7.5,
- $\alpha \in \mathbb{N}$ with $\alpha > 1 + 17c_1 + [24c_7 - 3c_1 \ln p]/\ln 3$.

8.3 Overview of the reconstruction

In this section, we show how Theorem 8.1.1 is proved using the results from Sections 8.4 and 8.5 and [21]. First we show how Theorem 8.1.1 implies Theorem 8.1.2.

Proof of Theorem 8.1.2. Let $\mathcal{A}_n : \mathcal{C}^{2\cdot 3^{10\alpha n}} \to \mathcal{C}^{[-3\cdot 3^n, 3\cdot 3^n]}$ be as in Theorem 8.1.1. We say that a sequence of pieces of sceneries $(\zeta_n \in \mathcal{C}^{I_n})_{n\geq c_3}$ converges pointwise to a scenery ζ if for all $z \in \mathbb{Z}$ there exists n_z such that $z \in I_n$ and $\zeta_n(z) = \zeta(z)$ for all $n \geq n_z$. We define

$$\mathcal{A}(\chi) := \begin{cases} \lim_{n\to\infty} \mathcal{A}_n(\chi|\,[0, 2\cdot 3^{10\alpha n}[) & \text{if this limit exists pointwise,} \\ (1)_{\mathbb{Z}} & \text{else.} \end{cases}$$

As a limit of measurable maps, $\mathcal{A}$ is measurable. Theorem 8.1.1 implies $\sum_{n=c_3}^{\infty} P\left([E_n]^c\right) \leq \sum_{n=c_3}^{\infty} c_4 \exp\left(-c_{25} n^{0.2}\right) < \infty$. Hence by the Borel-Cantelli lemma, $P(\cup_{m=c_3}^{\infty} \cap_{n=m}^{\infty} E_n) = 1$. In order to prove $P\left(\mathcal{A}(\chi) \approx \xi\right) = 1$, we use the same arguments as in the proof of Theorem 3.7 of [18]. (One shows $P\left(\cup_{m=c_3}^{\infty} \cap_{n=m}^{\infty} \{\xi|\,[-3^n, 3^n] \preceq_1 \xi|\,[-4\cdot 3^{n+1}, 4\cdot 3^{n+1}]\}\right) = 1$, which implies that the reconstructed pieces of scenery $\mathcal{A}_n(\chi|\,[0, 2\cdot 3^{10\alpha n}[)$ fit uniquely together for all n sufficiently large and yield the scenery ξ.) □

Hence, it suffices to define maps $\mathcal{A}_n$ which fulfill the claim of Theorem 8.1.1. The main ingredient in the construction of $\mathcal{A}_n$ is a map BigAlg^n which obtains as input data the observations collected by the random walk up to time $2\cdot 3^{10\alpha n}$ (as $\mathcal{A}_n$ does). In addition, BigAlg^n needs a sequence of stopping times $\tau := (\tau_k)_{k\in[1,3^{\alpha n}]}$ and a small piece of scenery ψ. BigAlg^n produces as output a piece of scenery $w \in \mathcal{C}^{[-3\cdot 3^n, 3\cdot 3^n]}$ which satisfies $\xi|[-3^n, 3^n] \preceq w \preceq \xi|[-4\cdot 3^n, 4\cdot 3^n]$ with high probability.

The reason why we need the stopping times $(\tau_k)_{k\in[1,3^{\alpha n}]}$ is the following: In order to be able to reconstruct the scenery in the interval $[-3^n, 3^n]$, the random walk must visit this part of the scenery many times. Otherwise, we will not have enough information for the reconstruction. Since $2\cdot 3^{10\alpha n}$ is considerably larger than 3^n, there is a good chance, the random walk visits the interval $[-3^n, 3^n]$ often up to time $2\cdot 3^{10\alpha n}$. However, up to time $2\cdot 3^{10\alpha n}$, only a small fraction of the time is spent in $[-3^n, 3^n]$. The rest of the time, when the random walk is outside of $[-3^n, 3^n]$, the observations do not provide us with useful information. Hence we need to be able to determine which parts of the observations are generated by the random walk on $\xi|[-3^n, 3^n]$. Formally, the task of the stopping times $(\tau_k)_{k\in[1,3^{\alpha n}]}$ is specified by the event $E_{\mathrm{stop}}^{n,\tau}$ defined as follows.

Definition 8.3.1. *For $n \in \mathbb{N}$ and a sequence $\tau = (\tau_k)_{k\geq 1}$ of $\mathcal{G}$-adapted stopping times, we define the event*

$$E_{\mathrm{stop}}^{n,\tau} := \bigcap_{k=1}^{3^{\alpha n}} \left\{\tau_k < 3^{10\alpha n}, |S_{\tau_k}| \leq 3^n, \tau_j + 2\cdot 3^{3n} \leq \tau_k \text{ for } j < k\right\}.$$

Besides stopping times, BigAlg^n obtains as input a piece of scenery ψ of length $\geq 2n^2+1$. Compared to the output of BigAlg^n, which has length of the order 3^n, ψ is very small. If $\psi \preceq \xi|[-3^n, 3^n]$, i.e. if we have with ψ some information about the underlying scenery, and if the event $E^{n,\tau}_{\text{stop}}$ holds, then with high probability, BigAlg^n reconstructs a piece of scenery around the origin. More formally:

Theorem 8.3.1. *There exist $c_8, c_{17}, c_{21} > 0$ and a sequence of measurable maps*

$$\text{BigAlg}^n : [0, 3^{10\alpha n}]^{[1,3^{\alpha n}]} \times \mathcal{C}^{2\cdot 3^{10\alpha n}} \times \bigcup_{k \geq n^2} \mathcal{C}^{[-k,k]} \to \mathcal{C}^{[-3\cdot 3^n, 3\cdot 3^n]}, n \in \mathbb{N},$$

such that for all $n \geq c_8$ and every sequence $\tau = (\tau_k)_{k\in[1,3^{\alpha n}]}$ of $\mathcal{G}$-adapted stopping times

$$P\left(E^{n,\tau}_{\text{stop}} \setminus E^{n,\tau}_{\text{recon Big}}\right) \leq c_{17} e^{-c_{21} n}, \text{ where}$$

$$E^{n,\tau}_{\text{recon Big}} := \begin{Bmatrix} \text{For all } \psi \in \mathcal{C}^{[-k,k]} \text{ with } k \geq n^2 \text{ and } \psi \preceq \xi|\,[-3^n, 3^n] \text{ we have} \\ \xi|[-3^n, 3^n] \preceq \text{BigAlg}^n(\tau, \chi|\,[0, 2\cdot 3^{10\alpha n}[\,, \psi) \preceq \xi|[-4\cdot 3^n, 4\cdot 3^n]. \end{Bmatrix}.$$

Let us explain how BigAlg^n reconstructs a piece of the scenery. Using the stopping times τ together with the observations from its input, BigAlg^n reconstructs with high probability all words of length $c_1 n/2$ in $\xi|[-5\cdot 3^n, 5\cdot 3^n]$; here c_1 is a (large) constant as described in Section 9.2. This is the crucial step in the definition of BigAlg^n. The words cannot be extracted from χ in a simple manner. Instead we need to look at certain empirical distributions of words which then allow us to obtain information about the true distribution and finally about the words themselves. Theorem 9.7.1 below provides a criterion to find words in the scenery. Reconstructing the words is a hard problem under our assumptions on random walk and scenery. In fact, this part of the reconstruction is much more difficult in the present setting than in previously solved scenery reconstruction problems.

Since with high probability, each word of length $c_1 n/4$ occurs at most once in $\xi|[-5\cdot 3^n, 5\cdot 3^n]$, it is possible to reconstruct a piece of scenery containing $\xi|[-3^n, 3^n]$ from the collection of words of length $c_1 n/2$. The assemblage will be done as follows: We start with the small piece of scenery ψ from the input of BigAlg^n. Then we look for a word of length $c_1 n/2$ which overlaps with ψ by at least $c_1 n/4$ letters and extends ψ by at least one letter. We continue the procedure with the extended ψ.

Once we have defined BigAlg^n, we can define the map $\mathcal{A}_n$ in terms of BigAlg^n with suitable stopping times τ and a piece of scenery ψ as input. The initial piece ψ will be a piece of scenery around a long block of ξ close to the origin. Since the ideas for finding words and defining BigAlg^n are central for this paper, we decided to concentrate on these parts. The proofs concerning the stopping times and the initial piece can be found in [21].

Let $\mathsf{block}^{n+} := \xi|[b_l^{n+}, b_r^{n+}]$ designate the leftmost block of ξ of length $\geq n$ with $b_l^{n+} \geq 0$, and let $\mathsf{block}^{n-} := \xi|[b_l^{n-}, b_r^{n-}]$ denote the rightmost block of ξ of length $\geq n$ with $b_r^{n-} \leq 0$. Finally, let $\mathsf{block}^n \in \{\mathsf{block}^{n+}, \mathsf{block}^{n-}\}$ denote the block which is visited first by S.

The map $\mathcal{A}_n$ will reconstruct a piece of scenery around block^n. Thus, first we need to locate block^n. With high probability, in a large neighborhood of block^n there is no large block in the scenery. Hence, up to a certain time horizon, long blocks in the observations χ indicate that the random walker generates the observations on block^n. The following theorem states that with high probability, there is a stopping time that stops the random walk in the set $\partial\mathsf{block}^n$.

Theorem 8.3.2. *([21], Theorem 3.1) For all $n \in \mathbb{N}$, there exists a $\mathcal{G}$-adapted stopping time $\nu^n(0)$, measurable with respect to $\sigma(\chi_k; k \in [0, 3^{10\alpha n}[)$, such that the probability of the event*

$$E^n_{\nu(0)\text{ ok}} := \{S_{\nu^n(0)} \in \partial\mathsf{block}^n\} \cap \{\nu^n(0) \le 2 \cdot 3^{3n}\} \cap \{\partial\mathsf{block}^n \subseteq [-3^n/3, 3^n/3]\}.$$

satisfies the following bound: There exist constants c_{11}, c_{12}, c_{13} such that for all $n \ge c_{11}$

$$P\left(\left[E^n_{\nu(0)\text{ ok}}\right]^c\right) \le c_{12} e^{-c_{13} n^{0.3}}.$$

Next, we reconstruct a piece of scenery around block^n. We show that there is a map SmallAlg^n with the following properties: Given $3^{\lfloor n^{0.3}\rfloor}$ observations collected by the random walker starting in the set $\partial\mathsf{block}^n$, a piece of scenery of length of the order $3^{\lfloor n^{0.2}\rfloor}$ around block^n can be reconstructed with high probability. For our purposes, it is convenient to state this differently: For ξ in a set of probability close to 1, conditioned on the scenery ξ, SmallAlg^n reconstructs with high probability a piece of scenery around block^n.

Theorem 8.3.3. *([21], Theorem 3.2) There exist constants $c_{14}, c_{15}, c_{18}, c_{13}, c_{15} > 0$ and a sequence*

$$\text{SmallAlg}^n : \mathcal{C}^{[0,3^{\lfloor n^{0.3}\rfloor}[} \to \mathcal{C}^{[-3\cdot 3^{\lfloor n^{0.2}\rfloor}, 3\cdot 3^{\lfloor n^{0.2}\rfloor}]}, n \ge c_{14},$$

of measurable maps such that the following holds: We set $H^n_i := \min\{k \ge 0 : S_k = b^n_i\}$ for $i \in \{l, r\}$. If we define

$$E^n_{\text{recon Small}} := \left\{\text{SmallAlg}^n\left(\chi|\left[0, 3^{\lfloor n^{0.3}\rfloor}\right[\right) \preceq \xi|[-3\cdot 3^{\lfloor n^{0.2}\rfloor}, 3\cdot 3^{\lfloor n^{0.2}\rfloor}]\right\} \text{ and}$$
$$\Xi^n := \left\{\xi \in \mathcal{C}^{\mathbb{Z}} : P_\xi\left(\left[\Theta^{-T} E^n_{\text{recon Small}}\right]^c\right) \le e^{-c_{18} n^{0.2}} \text{ for all } T \in \{H^n_l, H^n_r\}\right\},$$

then $P(\xi \notin \Xi^n) \le e^{-c_{18} n^{0.2}}$ for all $n \ge c_{14}$.

In fact, in [21], we heavily use the ideas from the construction of BigAlg^n to define SmallAlg^n. The piece of scenery reconstructed by SmallAlg^n is much smaller than the piece of scenery which $\mathcal{A}_n$ is supposed to reconstruct. The map SmallAlg^n is used to define stopping times $\nu^n(k)$, $k \ge 1$, which indicate when the random walk is in the interval $[-3^n, 3^n]$. Recall that $\mathcal{A}_n$ should reconstruct a piece of scenery of length of the order 3^n which is contained in $\xi|[-4\cdot 3^n, 4\cdot 3^n]$. Hence, it will be useful to have stopping times which stop the random walk in the interval $[-3^n, 3^n]$. We define

$$\psi_n := \text{SmallAlg}^n(\chi|[\nu^n(0), \nu^n(0) + 3^{\lfloor n^{0.3}\rfloor}[), \tag{8.3.1}$$

$$\mathbb{T}^n := \left\{\begin{array}{l} t \in [\nu^n(0), 3^{10\alpha n} - 3^{\lfloor n^{0.3}\rfloor}[: \exists w \in \mathcal{C}^{[-3^{\lfloor n^{0.2}\rfloor}, 3^{\lfloor n^{0.2}\rfloor}]} \text{ such that } w \preceq \psi_n \\ \text{and } w \preceq \text{SmallAlg}^n(\chi|[t, t + 3^{\lfloor n^{0.3}\rfloor}[) \end{array}\right\} \tag{8.3.2}$$

Let $\tilde\nu^n(1) < \tilde\nu^n(2) < \ldots$ denote the points in $\mathbb{T}^n$ in increasing order. We define $\nu^n := (\nu^n(k))_{k\in[1,3^{\alpha n}]}$ by

$$\nu^n(k) := \begin{cases} \tilde\nu^n(2\cdot 3^{3n}k) + 3^{\lfloor n^{0.3}\rfloor} & \text{if } 2\cdot 3^{3n}k \le |\mathbb{T}^n| \\ 3^{10\alpha n} & \text{else.} \end{cases}$$

Note that $\nu^n(k)$ depends only on $\chi|[0,3^{10\alpha n}[$ and is a $\mathcal{G}$-adapted stopping time. In fact, in order to determine whether $t \in \mathbb{T}_n$, we need to look at $\chi|[t, t+3^{\lfloor n^{0.3}\rfloor}[$, but $\nu^n(k)$ is never defined to be t, but only $t+3^{\lfloor n^{0.3}\rfloor}$.

The idea behind the definition of the $\nu^n(k)$'s is the following: With high proability, $\nu^n(0)$ stops the random walk in the set $\partial\mathsf{block}^n$ and ψ_n is up to a possible reflection a piece of scenery of length $6\cdot 3^{\lfloor n^{0.2}\rfloor}+1$ around block^n. The set $\mathbb{T}_n$ consists of times $t \geq \nu^n(0)$ such that $\mathrm{SmallAlg}^n$ applied to the observations starting at time t produces an output which agrees on a large subpiece, namely a piece of length $2\cdot 3^{\lfloor n^{0.2}\rfloor}+1$, with ψ_n. With high probability, ψ_n is typical for the scenery around block^n, and hence the random walker is in the interval $[-3^n, 3^n]$ at time t. (With high probability, block^n can be found in the piece of scenery $\xi|[-3^n/3, 3^n/3]$.) For the construction below, it will be essential that we have sufficiently many $\nu^n(k)$'s which are far enough apart from each other and all bounded by $3^{10\alpha n}$. Formally, the task of the stopping times $\nu^n(k)$ is specified by the event $E^{n,\nu^n}_{\mathrm{stop}}$, see Definition 9.3.1.

Recall the definition of Ξ^n from Theorem 8.3.3. If the event $E^n_{\nu(0)\ \mathrm{ok}} \cap \Theta^{-\nu^n(0)}[E^n_{\mathrm{recon\ Small}}]$ holds and $\xi \in \Xi_n$, then with high probability the stopping times ν^n stop the random walk correctly, in the sense that the event $E^{n,\nu^n}_{\mathrm{stop}}$ holds. This is made precise by the following propositon:

Proposition 8.3.1. *([21], Proposition 3.3) There exist constants c_{19}, c_{20}, c_{21} such that for all $n \geq c_{19}$*

$$P\big(\big[E^n_{\nu(0)\ \mathrm{ok}} \cap \Theta^{-\nu^n(0)}[E^n_{\mathrm{recon\ Small}}] \cap \{\xi \in \Xi^n\}\big] \setminus E^{n,\nu^n}_{\mathrm{stop}}\big) \leq c_{20} e^{-c_{21} n^{0.3}}.$$

Now, we have achieved the following: Using $\mathrm{SmallAlg}^n$,

we can reconstruct a piece of scenery ψ_n around block^n. With high probability, $\psi_n \preceq \xi|[-3^n, 3^n]$. Furthermore, the stopping times $\nu^n(k)$ stop the random walk with high probability in the interval $[-3^n, 3^n]$. Hence, with this input data, the algorithm BigAlg^n reconstructs with high probability a piece of scenery of length of the order 3^n around the origin.

Let $n \geq c_{14}$ with c_{14} as in Theorem 8.3.3, and let ψ_n be as in (9.3.4). We define

$$\mathcal{A}_n(\chi|[0, 2\cdot 3^{10\alpha n}[) := \mathrm{BigAlg}^n(\nu^n, \chi|[0, 2\cdot 3^{10\alpha n}[, \psi_n).$$

Proof of Theorem 8.1.1. We show that the maps $\mathcal{A}_n$ defined above fulfill the claim of Theorem 8.1.1. We have

$$\begin{aligned} P\left([E_n]^c\right) \;\leq\; & P\big(\big[E^{n,\nu^n}_{\mathrm{stop}} \cap E^n_{\nu(0)\ \mathrm{ok}} \cap \Theta^{-\nu^n(0)}[E^n_{\mathrm{recon\ Small}}]\big] \setminus E_n\big) + P([E^n_{\nu(0)\ \mathrm{ok}}]^c) \\ & + P\big(\big[E^n_{\nu(0)\ \mathrm{ok}} \cap \Theta^{-\nu^n(0)}[E^n_{\mathrm{recon\ Small}}] \cap \{\xi \in \Xi^n\}\big] \setminus E^{n,\nu^n}_{\mathrm{stop}}\big) \\ & + P\big([E^n_{\nu(0)\ \mathrm{ok}} \cap \{\xi \in \Xi^n\}] \setminus \Theta^{-\nu^n(0)}[E^n_{\mathrm{recon\ Small}}]\big) + P(\xi \notin \Xi^n). \quad (8.3.3) \end{aligned}$$

If $\Theta^{-\nu^n(0)}[E^n_{\mathrm{recon\ Small}}]$ holds, then $\psi_n \preceq \xi|[S_{\nu^n(0)} - 3\cdot 3^{\lfloor n^{0.2}\rfloor}, S_{\nu^n(0)} + 3\cdot 3^{\lfloor n^{0.2}\rfloor}]$. If in addition $E^n_{\nu(0)\ \mathrm{ok}}$ holds, then $S_{\nu^n(0)} \in \partial\mathsf{block}^n \subseteq [-3^n/3, 3^n/3]$, and consequently, $\psi_n \preceq \xi|[-3^n, 3^n]$ for all n sufficiently large. Hence, using Theorem 8.3.1,

$$P\big(\big[E^{n,\nu^n}_{\mathrm{stop}} \cap E^n_{\nu(0)\ \mathrm{ok}} \cap \Theta^{-\nu^n(0)}[E^n_{\mathrm{recon\ Small}}]\big] \setminus E_n\big) \leq P\big(E^{n,\nu^n}_{\mathrm{stop}} \setminus E^{n,\nu^n}_{\mathrm{recon\ Big}}\big) \leq c_{17} e^{-c_{21} n}$$

for all n sufficiently large. By Theorem 9.3.1, $P([E^n_{\nu(0)\text{ ok}}]^c) \le c_{12}e^{-c_{13}n^{0.3}}$ for all $n \ge c_{11}$. Proposition 9.3.3 states that

$$P\big([E^n_{\nu(0)\text{ ok}} \cap \Theta^{-\nu^n(0)}[E^n_{\text{recon Small}}] \cap \{\xi \in \Xi^n\}] \setminus E^{n,\nu^n}_{\text{stop}}\big) \le c_{20}e^{-c_{21}n^{0.3}}$$

for all $n \ge c_{19}$. Next, we estimate the second but last term in (8.3.3):

$$\begin{aligned} & P\big([E^n_{\nu(0)\text{ ok}} \cap \{\xi \in \Xi^n\}] \setminus \Theta^{-\nu^n(0)}[E^n_{\text{recon Small}}]\big) \\ = & \int_{\{\xi\in\Xi^n\}} P_\xi\big(E^n_{\nu(0)\text{ ok}} \cap \big[\Theta^{-\nu^n(0)}[E^n_{\text{recon Small}}]\big]^c\big)dP \\ \le & \int_{\{\xi\in\Xi^n\}} P_\xi\big(\{S_{\nu^n(0)} = b^n_l\} \cap \big[\Theta^{-\nu^n(0)}[E^n_{\text{recon Small}}]\big]^c\big)dP \\ & + \int_{\{\xi\in\Xi^n\}} P_\xi\big(\{S_{\nu^n(0)} = b^n_r\} \cap \big[\Theta^{-\nu^n(0)}[E^n_{\text{recon Small}}]\big]^c\big)dP. \end{aligned}$$

Using the strong Markov property of the random walk and Theorem 8.3.3, we conclude that the last quantity is $\le 2e^{-c_{18}n^{0.2}}$. Finally, by Theorem 8.3.3, $P(\xi \notin \Xi^n) \le \exp[-c_{18}n^{0.2}]$ for all $n \ge c_{14}$. Combining all these estimates with (8.3.3), the claim follows. □

8.4 How we find words in the observations

In this section, we show that many words of length of the order n can be reconstructed from the observations.

8.4.1 A sufficient criterion

Let us first explain why in the present setting it is so difficult to reconstruct words.

Assume for a moment that the scenery ξ, instead of being a 2-color scenery, is a 4-color scenery, i.e. $\xi \in \{0,1,2,3\}^{\mathbb{Z}}$. Let us assume furthermore, that for two integers x_1, x_2 we have $\xi_{x_1} = 2$ and $\xi_{x_2} = 3$, but $\xi_x \notin \{2,3\}$ for all $x \in \mathbb{Z} \setminus \{x_1, x_2\}$. Then we could reconstruct the portion of the scenery ξ lying between x_1 and x_2: As a matter of fact, since the random walk S is recurrent, it traverses a.s. at least once (and hence infinitely often) the shortest path from x_1 to x_2. Since we are given infinitely many observations χ the distance between x_1 and x_2 is the shortest time lapse that a 3 ever appears in the observations χ after a 2. When the random walk goes in a shortest possible way from x_1 to x_2, it traverses the straight path from x_1 to x_2. During that time, the random walk reveals in the observations the portion of ξ lying between x_1 and x_2. More precisely, if the pair of integers t_1, t_2 with $t_2 > t_1$ minimizes $|t_2 - t_1|$ under the conditions $\xi_{S_{t_1}} = 2$ and $\xi_{S_{t_2}} = 3$, then the piece of the scenery ξ lying in the interval $[\min\{x_1, x_2\}, \max\{x_1, x_2\}]$ is equal to the word $\chi|[t_1, t_2]$ or its transpose.

A related, but much more involved idea can still be used for 2-color sceneries. Let us next explain why the 2-color scenery reconstruction problem is much more difficult for simple random walk *with holding* than for simple random walk. So assume for the moment that S is a simple random walk, i.e. in each step S jumps one to the right or one to the left with probability 1/2. In this case, we can use instead of the extra colors 2 and 3 in the previous paragraph binary words of the form 001100 and 110011: It is easy to verify that the only possibility for the word 001100 to appear in the observations, is when 001100 occurs in the scenery (i.e. $\xi|[x, x+5] = 001100$ for some x) and the random walk traverses the straight path between x and $x+5$. The same is true for the word 110011.

If 001100 occurred in precisely one place x_1 of the scenery and 110011 occurred in precisely one place $x_2 \ne x_1$ of the scenery, then we could reconstruct up to a reflection the piece of scenery occurring between 001100 and 110011. We would just look in the observations where the word 110011 occurs in shortest

time after the word 001100. In between, we see a copy of the piece of the scenery ξ comprised between 001100 and 110011.

Of course, in an i.i.d. scenery, the word 001100 occurs a.s. infinitely often. Nevertheless, a slight modification of this idea was for instance used by Löwe, Matzinger, and Merkl in [18]. They used, that with high probability certain words occur only in certain areas of the scenery, which allowed them to reconstruct the words in between.

For a random walk with holding, the idea of patterns in the observations which tell us when we are back at the same spot like for example 001100 does not work at all. The reason is that if $\xi_z = 1$ and $\xi_{z+1} = 0$, then the random walk with holding can produce any pattern by just moving back and forth between z and $z+1$ and doing holdings. Thus, all patterns can be produced in most places in ξ and are thus not specific for some places in the scenery. However, in the case of a random walk with holding, the same idea of getting in shortest time from a point x_1 to a point x_2 can be applied to the distributions of the observations. In order to make this precise, we need some notation.

Let $\eta \in \bigcup_{k \geq 3^{3n}} \mathcal{C}^{[0,k[}$; the reader should think of η as a piece of observations. We consider $\tilde{\mathrm{O}}_1^n \mathrm{O}_2^n \tilde{\mathrm{O}}_3^n$ where $\tilde{\mathrm{O}}_1^n$ consists of the first $c_1 n$ blocks of η after time 3^{2n}, O_2^n equals the following $c_1 n/2$ observations in η extended until the next block starts, and $\tilde{\mathrm{O}}_3^n$ consists of the following $c_1 n$ blocks of η. (We do not look at the first 3^{2n} bits of η because the random walk needs to have a chance to visit all points in the interval $[-3^n, 3^n]$ before we start collecting information.) We are interested in O_2^n. The words $\tilde{\mathrm{O}}_1^n$ and $\tilde{\mathrm{O}}_3^n$ are used to find those O_2^n which occur in ξ "close to the origin". $\tilde{\mathrm{O}}_1^n$ and $\tilde{\mathrm{O}}_3^n$ play the role of the extra colors 2 and 3 in the argument above. It turns out that it suffices to use instead of $\tilde{\mathrm{O}}_j^n$ $(j = 1, 3)$ the corresponding sequences of truncated block lengths: we replace $\tilde{\mathrm{O}}_j^n$ by the sequence $\mathrm{O}_j^n \in \{1, 2, 3, 4, 5\}^{c_1 n}$ where the ith component equals the minimum of 5 and the length of the ith block of $\tilde{\mathrm{O}}_j^n$. Formally:

Definition 8.4.1. *Let $\eta \in \bigcup_{k \geq 3^{3n}} \mathcal{C}^{[0,k[}$. We abbreviate $\eta^n := \eta | [3^{2n}, 3^{3n}[$. We denote by $B_k(\eta)$ the kth block of η if η possesses at least k blocks; otherwise we set $B_k(\eta) := 101 \in \mathcal{C}^{[3^{3n}, 3^{3n}+3[}$. We denote by $o_l^n(\eta)$ the right end of $B_{c_1 n}(\eta^n)$. Furthermore we denote by $\tilde{o}_r^n(\eta)$ the left end of the first block of $\eta^n | [o_l^n(\eta) + c_1 n/2 - 2, 3^{3n}[$ and set $o_r^n(\eta) := \tilde{o}_r^n(\eta) + 1$. If $\eta^n | [o_l^n(\eta) + c_1 n/2 - 2, 3^{3n}[$ does not contain a block, then we set $o_r^n(\eta) := o_l^n(\eta)$. We define $\mathrm{O}^n := (\mathrm{O}_1^n, \mathrm{O}_2^n, \mathrm{O}_3^n)$ by*

$$\begin{aligned} \mathrm{O}_1^n(\eta) &:= (|B_k(\eta^n)| \wedge 5)_{k \in [1, c_1 n]}, \\ \mathrm{O}_2^n(\eta) &:= \eta | [o_l^n(\eta), o_r^n(\eta)], \\ \mathrm{O}_3^n(\eta) &:= (|B_k(\theta^{\tilde{o}_r^n(\eta)}(\eta))| \wedge 5)_{k \in [1, c_1 n]}. \end{aligned}$$

The letter "O" should remind the reader of "observation". By definition, $|\mathrm{O}_2^n(\eta)| \geq c_1 n/2$ unless $o_r^n(\eta) = o_l^n(\eta)$. The following picture illustrates our definitions for $c_1 n = 6$:

$$\eta = \underbrace{1110\ldots01110010}_{\eta|[0,3^{2n}[}\,0\,\underbrace{01110100000011000\,\boxed{1}\,000}_{\tilde{\mathrm{O}}_1^n(\eta)}\,0\,\underbrace{\boxed{1}\,1100101111000100111110\ldots}_{\tilde{\mathrm{O}}_3^n(\eta)}$$

$\eta_{o_l^n}$ and $\eta_{o_r^n}$ are marked with boxes. In this example, we have $\mathrm{O}_1^n(\eta) = (3, 1, 1, 5, 2, 3)$, $\mathrm{O}_2^n(\eta) = 100001$, $\mathrm{O}_3^n(\eta) = (3, 2, 1, 1, 4, 3)$.

In the following, let $\tau = (\tau_k)_{k \in [1, 3^{\alpha n}]}$ be a sequence of $\mathcal{G}$-adapted stopping times.

Definition 8.4.2. *For $\eta \in \mathcal{C}^{[0,2\cdot 3^{10\alpha n}[}$, we define the empirical distribution of O^n observed after each time τ_k, $k \in [1, 3^{\alpha n}]$:*

$$\hat{\mu}_\eta^{n,\tau} := 3^{-\alpha n} \sum_{k\in[1,3^{\alpha n}]} \delta_{\mathrm{O}^n(\theta^{\tau_k}\eta)}.$$

For $\eta \in \mathcal{C}^{\mathbb{N}_0}$, we set $\hat{\mu}_\eta^{n,\tau} := \hat{\mu}_{\eta|[0,2\cdot 3^{10\alpha n}[}^{n,\tau}$.

Recall that $P_{x,\xi}\left[\mathrm{O}^n(\chi)\right]^{-1}$ denotes the distribution of $\mathrm{O}^n(\chi)$ conditioned on the scenery to be ξ and conditioned on the random walk to start in x.

Definition 8.4.3. *For an admissible path $R \in \mathbb{Z}^{[0,2\cdot 3^{10\alpha n}[}$, let $a_R^{n,\tau}(x)$ be the proportion of $k \in [1, 3^{\alpha n}]$ with $R_{\tau_k} = x$. We define*

$$\begin{aligned} \mu_{\xi,R}^{n,\tau} &:= \sum_{x\in[-3^n,3^n]} a_R^{n,\tau}(x) P_{x,\xi}\left[\mathrm{O}^n(\chi)\right]^{-1} \\ \varepsilon_{\xi,R}^{n,\tau} &:= \hat{\mu}_{\xi\circ R}^{n,\tau} - \mu_{\xi,R}^{n,\tau}. \end{aligned}$$

For an admissible path $R \in \mathbb{Z}^{\mathbb{N}_0}$, we set $\mu_{\xi,R}^{n,\tau} := \mu_{\xi,R|[0,2\cdot 3^{10\alpha n}[}^{n,\tau}$.

Hence, $\mu_{\xi,R}^{n,\tau}$ is a mixture of the distributions $P_{x,\xi}\left[\mathrm{O}^n(\chi)\right]^{-1}$. The signed measure $\varepsilon_{\xi,R}^{n,\tau}$ measures the difference between the empirical measure and $\mu_{\xi,R}^{n,\tau}$. It will be shown in Lemma 8.5.8 that with high P-probability $\varepsilon_{\xi,S}^{n,\tau}$ is small provided the stopping times τ stop correctly. This is used to reconstruct the scenery: $\hat{\mu}_\chi^{n,\tau}$ can be computed from $\chi|[0, 2\cdot 3^{10\alpha n}[$ and τ. It is close to $\mu_{\xi,S}^{n,\tau}$, from which we will extract information about the scenery ξ. Of course, $\mu_{\xi,S}^{n,\tau}$ cannot be obtained from χ and τ only.

By definition, $\mu_{\xi,R}^{n,\tau}$ and $\hat{\mu}_\eta^{n,\tau}$ are measures on the set $\mathrm{obs} := [1,5]^{c_1 n} \times \mathrm{obs}_2 \times [1,5]^{c_1 n}$ with $\mathrm{obs}_2 := \{w \in \mathcal{C}^k : k \geq c_1 n/2, w_{k-1} \neq w_k, w_j = w_{k-1} \text{ for all } j \in [c_1 n/2 - 1, k-1]\}$. We denote by

$$\Pi_2 : \mathrm{obs} \to \mathrm{obs}_2, \quad \Pi_{1,3} : \mathrm{obs} \to [1,5]^{c_1 n} \times [1,5]^{c_1 n}$$

the canonical projections. Furthermore, we introduce the event that an observation $\mathrm{O} \in \mathrm{obs}$ has $\Pi_2(\mathrm{O})$ of length $d \geq c_1 n/2$:

$$E_{\text{block}}^{n,d} := \{\mathrm{O} \in \mathrm{obs} : [\Pi_2(\mathrm{O})]_{d-1} \neq [\Pi_2(\mathrm{O})]_d\}.$$

We order the 2^d elements of $\mathcal{C}^d$ lexicographically and denote them by $v^1, v^2, \cdots, v^{2^d}$. Let $e_{v^k} := (e_{v^k}(i))_{i\in[1,2^d]}$ be defined by $e_{v^k}(i) := \delta_k(i)$; i.e. $\{e_{v^k}; k \in [1, 2^d]\}$ is the canonical basis in $\mathbb{R}^{2^d}$. Let $\{1_{v^k}; k \in [1, 2^d]\}$ be the dual basis, i.e. $1_{v^k}(e_{v^j}) = \delta_k(j)$ for all $j, k \in [1, 2^d]$.

Sometimes it will be convenient to identify a measure λ which is supported on a finite ordered set $\{s_1, s_2, \ldots, s_m\}$ with the vector $(\lambda(\{s_1\}), \lambda(\{s_2\}), \ldots, \lambda(\{s_m\}))$. Similarly, we sometimes identify measures supported on $\mathbb{N}_0$ by one-sided infinite vectors.

Let $w \in \mathcal{C}^d$. For any probability measure λ on $\mathcal{C}^d$ we have $1_w(\lambda) = \lambda(w)$. In particular, if λ gives mass one to w, then $1_w(\lambda) = 1$. We denote by 1 the linear functional defined by $1(\lambda) := \sum_{i=1}^d \lambda_i$. If g_1 and g_2 are two linear functionals we denote by $g_1 \otimes g_2$ their tensor product.

The following theorem gives sufficient conditions for a word $w \in \mathcal{C}^d$ to be contained in the scenery $\xi|[-3^{3n}, 3^{3n}]$. Its proof is postponed to Section 8.4.3. For the definition of *positivity* for a linear functional we refer the reader to Section 8.4.2, in particular Definition 9.7.6.

Theorem 8.4.1. *There exists $c_{11} > 0$ such that for all $n \geq c_{11}$, $d \in [c_1 n/2, c_1 n]$, and $w \in \mathcal{C}^d$ with $w_{d-1} \neq w_d$ the following holds whenever the event $E^{n,\tau}_{\text{stop}}$ holds: Suppose there exist positive linear functionals g_1 and g_3 on $(\mathbb{R}^5)^{\otimes c_1 n}$ with the following properties:*

1. *Case $q \neq 0$:*

$$\begin{aligned}(g_1 \otimes 1_w \otimes g_3)(\hat{\mu}^{n,\tau}_{\xi\circ S}[\cdot \cap E^{n,d}_{\text{block}}]) &> 1 && (8.4.1)\\ (g_1 \otimes 1 \otimes g_3)(\hat{\mu}^{n,\tau}_{\xi\circ S}[\cdot \cap E^{n,d-1}_{\text{block}}]) &\leq 1/(5n^2) && (8.4.2)\\ \|g_1 \otimes g_3\|_2 \cdot \|\varepsilon^{n,\tau}_{\xi,S}\|_1^{1/2} &\leq 1/(2n^2). && (8.4.3)\end{aligned}$$

2. *Case $q = 0$: (9.7.1), (9.7.3), and*

$$(g_1 \otimes 1 \otimes g_3)(\hat{\mu}^{n,\tau}_{\xi\circ S}[\cdot \cap E^{n,d-2}_{\text{block}}]) \leq 1/(5n^2). \quad (8.4.4)$$

Then $w \preceq \xi|\,[-3^{3n}, 3^{3n}]$.

8.4.2 Positive linear functionals

For $m \in \mathbb{N}$, we denote by T_m the first hitting time of $\{-1, m\}$ by the random walk S: $T_m := \min\{k \geq 0 : S_k \in \{-1, m\}\}$. Let $\xi|[a,b]$ be a block of length m and let S^B be a random walk with $S^B_0 \in \{a, b\}$ having its increments distributed as S. Then the P_ξ-distribution of the length of the first block in $\xi \circ S^B$ equals $P[T_m]^{-1}$, the distribution of T_m. We abbreviate

$$\lambda^m_l(\cdot) := P\left(\{T_m \in \cdot\} \cap \{S_{T_m} = -1\}\right), \quad \lambda^m_r(\cdot) := P\left(\{T_m \in \cdot\} \cap \{S_{T_m} = m\}\right).$$

Clearly, $T_m \geq 1$ P-almost surely. We compute

$$\begin{aligned}\lambda^1_r &= (p, pq, pq^2, pq^3, pq^4, \ldots),\\ \lambda^2_r &= (0, p^2, 2p^2q, p^4 + 3p^2q^2, \ldots),\\ \lambda^3_r &= (0, 0, p^3, 3p^3q, 2p^5 + 6p^3q^2, \ldots),\\ \lambda^4_r &= (0, 0, 0, p^4, \ldots),\\ \lambda^5_r &= (0, 0, 0, 0, p^5, \ldots);\end{aligned}$$

here "..." means we are not interested in these values. We define $h : \mathbb{N}_0 \to [1,5]$, $x \mapsto x \wedge 5$. Then

$$\lambda^m_l h^{-1}(\cdot) = P\left(\{T_m \wedge 5 \in \cdot\} \cap \{S_{T_m} = -1\}\right).$$

The measures $\lambda^m_l h^{-1}, \lambda^m_r h^{-1}$ are supported on the set $\{1,2,3,4,5\}$. Hence we can identify them with vectors in $\mathbb{R}^5_+$.

Definition 8.4.4. *We define vectors in $\mathbb{R}^5_+$:*

$$\begin{aligned}\vec{x}_1 &:= (p, pq, pq^2, pq^3, pq^4),\\ \vec{x}_2 &:= \lambda^2_r h^{-1} = (0, p^2, 2p^2q, p^4 + 3p^2q^2, \lambda^2_r([5,\infty[),\\ \vec{x}_3 &:= (0, 0, p^3, 3p^3q, p^5 + 6p^3q^2),\\ \vec{x}_4 &:= \lambda^4_r h^{-1} = (0, 0, 0, p^4, \lambda^4_r([5,\infty[),\\ \vec{x}_5 &:= (0, 0, 0, 0, 1).\end{aligned}$$

Clearly, $\{\vec{x}_i\}_{i\in[1,5]}$ is a basis of $\mathbb{R}^5$. We denote by $\{\vec{x}^*_i\}_{i\in[1,5]}$ the corresponding dual basis.

Remark 8.4.1. *1. For any $m \geq 1$ and $i \in \{l, r\}$ the vector $\lambda_i^m h^{-1}$ can be written as a linear combination with positive coefficients of $\vec{x}_j$, $1 \leq j \leq 5$.*

2. We have $\vec{x}_2^*(\lambda_i^m h^{-1}) \neq 0$ iff $i = r$ and $m = 2$. Furthermore, $\vec{x}_4^*(\lambda_i^m h^{-1}) \neq 0$ iff $i = r$ and $m = 4$.

3. For $i \in \{1, 3, 5\}$, we have $x_i^*(\lambda_r^2) = 0$ and $x_i^*(\lambda_r^4) = 0$.

4. The lower bound $\vec{x}_{m\wedge 5}^*(\lambda_r^m h^{-1}) \geq (m+1)^{-1}$ holds for all $m \geq 1$.

Proof. By definition, $\lambda_r^m h^{-1} = \vec{x}_m$ for $m = 2, 4$. Furthermore, $\lambda_r^1 = \vec{x}_1 + \lambda_r^1(]5, \infty[)\vec{x}_5$ and $\lambda_r^3 = \vec{x}_3 + (p^5 + \lambda_r^3(]5, \infty[))\vec{x}_5$. Finally, we have $\lambda_r^m h^{-1} = P(S_{T_m} = m) \cdot \vec{x}_5 = (m+1)^{-1} \cdot \vec{x}_5$ for $m \geq 5$.

By symmetry, $\lambda_l^1 h^{-1} = \lambda_r^1 h^{-1}$. Furthermore, $\lambda_l^2 h^{-1} = (p, pq, pq^2 + p^3, pq^3 + 3p^3 q, p^5 + 6p^3 q^2 + pq^4 + \lambda_l^2(]5, \infty[)) = \vec{x}_1 + \vec{x}_3 + \lambda_l^2(]5, \infty[) \cdot \vec{x}_5$ and $\lambda_l^3 h^{-1} = (p, pq, pq^2 + p^3, pq^3 + 3p^3 q, 2p^5 + 6p^3 q^2 + pq^4 + \lambda_l^3(]5, \infty[)) = \vec{x}_1 + \vec{x}_3 + \lambda_l^3(]5, \infty[) \cdot \vec{x}_5$. Let $m \geq 4$. Since any path, which starts at 0, hits 3 before hitting -1, and hits -1 before m has ≥ 7 steps, we have $\lambda_l^m h^{-1} = \lambda_l^3 h^{-1}$ for all $m \geq 4$. The claim follows. □

Definition 8.4.5. *We call a function $f : (\mathbb{R}^5)^{\otimes m} \to \mathbb{R}$* positive *if $f(\otimes_{k=1}^m \vec{x}_{n_k}) \geq 0$ for all $n_1, n_2, \ldots, n_m \in \{1, 2, 3, 4, 5\}$.*

Remark 8.4.2. *Let g be a positive linear functional on $(\mathbb{R}^5)^{\otimes c_1 n}$. If $P_{x,\xi}[S_{o_l^n} = y] > 0$, then $g(P_{x,\xi}[\mathrm{O}_1^n \in \cdot | S_{o_l^n} = y]) \geq 0$. If $P_{x,\xi}[S_{o_r^n} = y] > 0$, then $g(P_{x,\xi}[\mathrm{O}_3^n \in \cdot | S_{o_r^n} = y]) \geq 0$.*

Proof. Suppose $P_{x,\xi}(S_{o_l^n} = y) > 0$. By the definition of O_1^n, we can write the probability $P_{x,\xi}[\mathrm{O}_1^n \in \cdot | S_{o_l^n} = y]$ as a linear combination with positive coefficients of tensor products of the $\lambda_i^m h^{-1}$'s. Each $\lambda_i^m h^{-1}$ equals a linear combination with positive coefficients of $\vec{x}_i$, $1 \leq i \leq 5$ by Remark 8.4.1. The estimate $g(P_{x,\xi}[\mathrm{O}_1^n \in \cdot | S_{o_l^n} = y]) \geq 0$ follows because g is positive. The second part of the statement is proved analogously. □

8.4.3 Proof of Theorem 9.7.1

We begin with a lemma, which we need in the proof of Theorem 9.7.1.

Lemma 8.4.1. *There exists $c_7 > 0$ such that for all $n \geq c_7$, for all $d \in]2, c_1 n]$, and all $x \in [0, d[$ the following hold:*

1. If $q \neq 0$, then $P(S_d = x) \leq n^2 P(S_{d-1} = x)$.

2. If $q = 0$ and $P(S_{d-2} = x) > 0$, then $P(S_d = x) \leq n^2 P(S_{d-2} = x)$.

Proof. Let $n \in \mathbb{N}$, $d \in]2, c_1 n]$, $x \in [0, d[$, and suppose $q \neq 0$. We denote by $\Pi_{d,x} \subseteq \mathbb{Z}^{[0,d]}$ the set of all admissible pieces of paths from 0 to x, and we define a map $f : \Pi_{d,x} \to \Pi_{d-1,x}$ as follows: If $\pi \in \Pi_{d,x}$ contains a holding, i.e. $\pi_y = \pi_{y-1}$ for some $y \in]0, d]$, then we define $f(\pi)$ to be the path obtained from π by removing the first holding in π. Otherwise, because of $x < d$, there exists either a step to the left followed by a step to the right or a step to the right followed by a step to the left in π. In this case, we define $f(\pi)$ to be the path obtained from π by replacing the first occurrence of such a pair of steps by a holding. For any $\pi \in \Pi_{d,x}$ we have

$$P(S|[0, d] = \pi) \leq \max\left\{q, p^2 q^{-1}\right\} P(S|[0, d-1] = f(\pi)).$$

Furthermore, any $\pi' \in \Pi_{d-1,x}$ has at most $3d$ pre-images under f. Hence we obtain

$$\begin{aligned} P(S_d = x) &= \sum_{\pi \in \Pi_{d,x}} P(S|[0,d] = \pi) \leq \sum_{\pi' \in \Pi_{d-1,x}} \sum_{\pi \in f^{-1}(\pi')} P(S|[0,d] = \pi) \\ &\leq \sum_{\pi' \in \Pi_{d-1,x}} \left| f^{-1}(\pi') \right| \max\left\{q, p^2 q^{-1}\right\} P(S|[0,d-1] = \pi') \\ &\leq 3d \max\left\{q, p^2 q^{-1}\right\} P(S_{d-1} = x). \end{aligned}$$

Since $d \leq c_1 n$, we have $3d \max\{q, p^2 q^{-1}\} \leq n^2$ for all n sufficiently large and the claim follows in the case $q \neq 0$. The case $q = 0$ is treated similarly. □

Proof of Theorem 9.7.1. Let $q \neq 0$. Let the assumptions of the theorem be satisfied and suppose $w \not\preceq \xi|\,[-3^{3n}, 3^{3n}]$.

By Hölder's inequality, $|\sum_{i=1}^m g_i \lambda_i| \leq [\sum_{i=1}^m g_i^2]^{1/2} [\sum_{i=1}^m \lambda_i^2]^{1/2}$ for all $g_i, \lambda_i \in \mathbb{R}$. Hence we have for any linear functional $g : \mathbb{R}^m \to \mathbb{R}$ and $\lambda \in \mathbb{R}^m$ the estimate $|g(\lambda)| \leq \|g\|_2 \|\lambda\|_2$. Since $\varepsilon_{\xi,S}^{n,\tau}$ is the difference of two probability measures, $\|\varepsilon_{\xi,S}^{n,\tau}[\,\cdot \cap E_{\text{block}}^{n,d}]\|_2^2 \leq \|\varepsilon_{\xi,S}^{n,\tau}\|_1$. Using this together with (9.7.3), we obtain

$$\begin{aligned} |(g_1 \otimes 1_w \otimes g_3)(\varepsilon_{\xi,S}^{n,\tau}[\,\cdot \cap E_{\text{block}}^{n,d}])| &\leq \|g_1 \otimes 1_w \otimes g_3\|_2 \cdot \|\varepsilon_{\xi,S}^{n,\tau}[\,\cdot \cap E_{\text{block}}^{n,d}]\|_2 \\ &\leq \|g_1 \otimes g_3\|_2 \cdot \|\varepsilon_{\xi,S}^{n,\tau}\|_1^{1/2} \leq 1/(2n^2). \end{aligned} \tag{8.4.5}$$

Hence it follows from $\mu_{\xi,S}^{n,\tau} = \hat{\mu}_{\xi \circ S}^{n,\tau} - \varepsilon_{\xi,S}^{n,\tau}$ and (9.7.1) that

$$(g_1 \otimes 1_w \otimes g_3)(\mu_{\xi,S}^{n,\tau}[\,\cdot \cap E_{\text{block}}^{n,d}]) \geq 1 - 1/(2n^2) > 3/4$$

for all $n \geq 2$. Inserting the definition of $\mu_{\xi,S}^{n,\tau}$ and using the linearity of $g_1 \otimes 1_w \otimes g_3$ and the definition of 1_w, yields

$$\begin{aligned} 3/4 &< (g_1 \otimes 1_w \otimes g_3)(\mu_{\xi,S}^{n,\tau}[\,\cdot \cap E_{\text{block}}^{n,d}]) \qquad (8.4.6) \\ &= \sum_{x \in [-3^n, 3^n]} a_S^{n,\tau}(x)(g_1 \otimes 1_w \otimes g_3)(P_{x,\xi}[(\mathrm{O}_1^n(\chi), \mathrm{O}_2^n(\chi), \mathrm{O}_3^n(\chi)) \in \cdot \cap E_{\text{block}}^{n,d}]) \\ &= \sum_{x \in [-3^n, 3^n]} a_S^{n,\tau}(x)(g_1 \otimes 1_w \otimes g_3)(P_{x,\xi}[(\mathrm{O}_1^n(\chi), \mathrm{O}_2^n(\chi) = w, \mathrm{O}_3^n(\chi)) \in \cdot]). \end{aligned}$$

In the following, we omitt dependencies on χ in the notation if there is no risk of confusion. Since we assumed $w \not\preceq \xi|\,[-3^{3n}, 3^{3n}]$, we have for any admissible path $R \in [-3^{3n}, 3^{3n}]^{[0,d[}$ with $\xi \circ R = w$ the estimate $|R_0 - R_{d-1}| < d - 1$. Consequently,

$$P_{x,\xi}[(\mathrm{O}_1^n, \mathrm{O}_2^n = w, \mathrm{O}_3^n) \in \cdot] = \sum_{y,z} P_{x,\xi}[S_{o_l^n} = y, S_{o_r^n} = z, (\mathrm{O}_1^n, \mathrm{O}_2^n = w, \mathrm{O}_3^n) \in \cdot], \tag{8.4.7}$$

where the sum is taken over all $y, z \in [-3^{3n}, 3^{3n}]$ with the property $|y - z| < d - 1$ and $P_{x,\xi}[S_{o_l^n} = y, S_{o_l^n + d - 1} = z] > 0$. We rewrite the summands: On the event $\{\mathrm{O}_2^n = w\}$, we have $o_r^n = o_l^n + d - 1$. Note that O_1^n depends only on the random walk up to time o_l^n,

whereas $(S_{o_r^n}, \mathrm{O}_2^n, \mathrm{O}_3^n)$ depends only on $S_{o_l^n}$ and the random walk increments $S_{o_l^n+t} - S_{o_l^n}$, $t \geq 0$. Therefore, O_1^n and $(S_{o_r^n}, \mathrm{O}_2^n, \mathrm{O}_3^n)$ are independent conditioned on $S_{o_l^n} = y$. Thus

$$(8.4.7) = \sum_{y,z} P_{x,\xi}[S_{o_l^n} = y, \mathrm{O}_1^n \in \cdot] \otimes P_{x,\xi}[S_{o_r^n} = z, (\mathrm{O}_2^n = w, \mathrm{O}_3^n) \in \cdot | S_{o_l^n} = y]$$

$$= \sum_{y,z} P_{x,\xi}[S_{o_l^n} = y, \mathrm{O}_1^n \in \cdot] \otimes P_{x,\xi}[S_{o_l^n+d-1} = z | S_{o_l^n} = y] P_{x,\xi}[(\mathrm{O}_2^n = w, \mathrm{O}_3^n) \in \cdot | A_{y,z}^{d-1}] \tag{8.4.8}$$

with $A_{y,z}^{d-1} := \{S_{o_l^n} = y, S_{o_l^n+d-1} = z\}$. Using again the Markov property of the random walk, we see that O_2^n and O_3^n are independent, conditioned on $A_{y,z} := \{S_{o_l^n} = y, S_{o_r^n} = z\}$. Hence

$$\begin{aligned} P_{x,\xi}[(\mathrm{O}_2^n = w, \mathrm{O}_3^n) \in \cdot | A_{y,z}^{d-1}] &= \frac{P_{x,\xi}[A_{y,z}]}{P_{x,\xi}[A_{y,z}^{d-1}]} P_{x,\xi}[(\mathrm{O}_2^n = w, \mathrm{O}_3^n) \in \cdot | A_{y,z}] \\ &= \frac{P_{x,\xi}[A_{y,z}]}{P_{x,\xi}[A_{y,z}^{d-1}]} P_{x,\xi}[\mathrm{O}_2^n = w \in \cdot | A_{y,z}] \otimes P_{x,\xi}[\mathrm{O}_3^n \in \cdot | A_{y,z}] \\ &= P_{x,\xi}[\mathrm{O}_2^n = w \in \cdot | A_{y,z}^{d-1}] \otimes P_{x,\xi}[\mathrm{O}_3^n \in \cdot | A_{y,z}]. \end{aligned}$$

Consequently, we obtain from (8.4.7) and (8.4.8)

$$\begin{aligned} P_{x,\xi}[(\mathrm{O}_1^n, \mathrm{O}_2^n = w, \mathrm{O}_3^n) \in \cdot] &= \sum_{y,z} P_{x,\xi}[S_{o_l^n} = y, \mathrm{O}_1^n \in \cdot] \otimes P_{x,\xi}[S_{o_l^n+d-1} = z | S_{o_l^n} = y] \\ &\quad P_{x,\xi}[\mathrm{O}_2^n = w \in \cdot | A_{y,z}^{d-1}] \otimes P_{x,\xi}[\mathrm{O}_3^n \in \cdot | A_{y,z}]. \end{aligned} \tag{8.4.9}$$

In view of (8.4.6), the aim is to apply $g_1 \otimes 1_w \otimes g_3$ to the last sum. We observe for $n \geq c_7$ with c_7 as in Lemma 8.4.1

$$\begin{aligned} & P_{x,\xi}[S_{o_l^n+d-1} = z | S_{o_l^n} = y] 1_w (P_{x,\xi}[\mathrm{O}_2^n = w \in \cdot | A_{y,z}^{d-1}]) \\ \leq\ & P_{x,\xi}[S_{o_l^n+d-1} = z | S_{o_l^n} = y] = P_{0,\xi}[S_{d-1} = z - y] \leq n^2 P_{0,\xi}[S_{d-2} = z - y] \\ =\ & n^2 P_{x,\xi}[S_{o_l^n+d-2} = z | S_{o_l^n} = y]; \end{aligned} \tag{8.4.10}$$

here we used the Markov property of the random walk and Lemma 8.4.1. Combining (8.4.9) with (8.4.10) and using Remark 8.4.2 yields

$$\begin{aligned} & (g_1 \otimes 1_w \otimes g_3)(P_{x,\xi}[(\mathrm{O}_1^n, \mathrm{O}_2^n = w, \mathrm{O}_3^n) \in \cdot]) \qquad (8.4.11) \\ \leq\ & \sum_{y,z} g_1[P_{x,\xi}[S_{o_l^n} = y, \mathrm{O}_1^n \in \cdot]] n^2 P_{x,\xi}[S_{o_l^n+d-2} = z | S_{o_l^n} = y] g_3[P_{x,\xi}[\mathrm{O}_3^n \in \cdot | A_{y,z}]]. \end{aligned}$$

We can enlarge the last sum by summing over *all* $y, z \in [-3^{3n}, 3^{3n}]$ with $|y-z| < d-1$ and $P_{x,\xi}(S_{o_l^n} = y, S_{o_r^n} = z) > 0$ and not only over those with $P_{x,\xi}[S_{o_l^n} = y, S_{o_l^n+d-1} = z] > 0$. The terms added in this way are non-negative by Remark 8.4.2. Note that

$$P_{x,\xi}[(\mathrm{O}_1^n, \mathrm{O}_2^n \mathrm{O}_3^n) \in \cdot \cap E_{\mathrm{block}}^{n,d-1}] = \sum_{w'} P_{x,\xi}[\mathrm{O}_2^n = w', (\mathrm{O}_1^n, \mathrm{O}_3^n) \in \cdot],$$

where the sum is taken over all $w' \in \mathcal{C}^{d-1}$ with $w'_{d-1} \neq w'_{d-2}$. We use (8.4.9) for $d-1$ instead of d to obtain

$$(8.4.11) \quad \leq \quad n^2 (g_1 \otimes 1 \otimes g_3)[P_{x,\xi}[(\mathrm{O}_1^n, \mathrm{O}_2^n, \mathrm{O}_3^n) \in \cdot \cap E_{\mathrm{block}}^{n,d-1}]]. \tag{8.4.12}$$

Since the event $E^{n,\tau}_{\text{stop}}$ holds, $\sum_{x\in[-3^n,3^n]} a^{n,\tau}_S(x) = 1$. Consequently, the estimates (8.4.6) and (8.4.12) imply

$$3/(4n^2) \quad < \quad (g_1 \otimes 1 \otimes g_3)(\mu^{n,\tau}_{\xi\circ S}[\cdot \cap E^{n,d-1}_{\text{block}}]). \tag{8.4.13}$$

Similar to (8.4.5), we obtain

$$|(g_1 \otimes 1 \otimes g_3)(\varepsilon^{n,\tau}_{\xi,S}[\cdot \cap E^{n,d-1}_{\text{block}}])| \le \|g_1 \otimes g_3\|_2 \cdot \|\varepsilon^{n,\tau}_{\xi,S}\|_1^{1/2} \le 1/(2n^2).$$

Hence it follows from (8.4.13) and $\hat\mu^{n,\tau}_{\xi\circ S} = \mu^{n,\tau}_{\xi\circ S} - \varepsilon^{n,\tau}_{\xi,S}$

$$(g_1 \otimes 1 \otimes g_3)(\hat\mu^{n,\tau}_{\xi\circ S}[\cdot \cap E^{n,d-1}_{\text{block}}]) > 3/(4n^2) - 1/(2n^2) = 1/(4n^2),$$

which contradicts assumption (9.7.2). Thus $w \preceq \xi|[-3^{3n}, 3^{3n}]$ and the theorem is proved in the case $q \neq 0$. If $q = 0$, one replaces $d-1$ by $d-2$ in the above argument und uses (8.4.4) instead of (9.7.2). □

8.5 Reconstructing a piece of scenery

Let $n \in \mathbb{N}$. The aim of this section is to define a map BigAlg^n which fulfills the claim of Theorem 8.3.1. Special functionals and events are needed in the proof of Theorem 8.3.1; their definitions are stated in Subsection 8.5.2. Subsection 8.5.3 contains the combinatorial part in the proof of Theorem 8.3.1, and Subsection 8.5.4 deals with the probabilistic estimates.

8.5.1 Definition of BigAlg^n

BigAlg^n takes as arguments

$$\tau \in \left[0, 3^{10\alpha n}\right]^{[1,3^{\alpha n}]},\ \eta \in \mathcal{C}^{2\cdot 3^{10\alpha n}},\ \text{and } \psi \in \bigcup_{k\ge n^2} \mathcal{C}^{[-k,k]} \tag{8.5.1}$$

and produces an output $\text{BigAlg}^n(\tau,\eta,\psi) \in \mathcal{C}^{[-3\cdot 3^n, 3\cdot 3^n]}$. The reader should think of τ as a realization of a sequence of $3^{\alpha n}$ stopping times, η stands for $2\cdot 3^{10\alpha n}$ observations, and ψ should be thought of as a small piece of the scenery ξ around which the reconstruction takes place. In the following, we treat τ, η, and ψ as abstract input data of BigAlg^n which need to fulfill (8.5.1) only.

Let τ, η, and ψ satisfy (8.5.1). We use the conditions of Theorem 9.7.1 to define a set $\mathsf{Words}^n(\tau,\eta)$ of building blocks for the scenery which we would like to reconstruct.

Definition 8.5.1. *Let $c_7 > 0$ be chosen as in Section 9.2. We define $\mathsf{Words}^n(\tau,\eta)$ to be the set of all $w \in \mathcal{C}^d$, $d \in [c_1 n/2, c_1 n]$ such that there exist positive linear functionals g_1 and g_3 on $(\mathbb{R}^5)^{\otimes c_1 n}$ with the following properties:*

1. *Case $q \neq 0$:*

$$(g_1 \otimes 1_w \otimes g_3)\,(\hat\mu^{n,\tau}_\eta[\cdot \cap E^{n,d}_{\text{block}}]) \quad > \quad 1 \tag{8.5.2}$$

$$(g_1 \otimes 1 \otimes g_3)(\hat\mu^{n,\tau}_\eta[\cdot \cap E^{n,d-1}_{\text{block}}]) \quad \le \quad 1/(5n^2) \tag{8.5.3}$$

$$\|g_1 \otimes g_3\|_2 \quad \le \quad e^{c_7 n}. \tag{8.5.4}$$

2. *Case $q = 0$: (9.7.10), (9.7.12), and*

$$(g_1 \otimes 1 \otimes g_3)(\hat{\mu}_\eta^{n,d-2,\tau}[\cdot \cap E_{\text{block}}^{n,d-2}]) \;\leq\; 1/(5n^2). \tag{8.5.5}$$

The output of BigAlg^n is supposed to contain ψ in the middle and all subpieces of length $c_1 n/2$ should be contained in a possibly bigger piece of $\mathsf{Words}^n(\tau,\eta)$. Formally:

Definition 8.5.2. *We define* $\mathsf{Output}^n(\tau,\eta,\psi) :=$

$$\left\{\begin{array}{l} w \in \mathcal{C}^{[-3\cdot 3^n, 3\cdot 3^n]} : w|[-k,k] = \psi \text{ for } k = (|\psi|-1)/2 \text{ and for all intervals } I \subseteq \\ [-3\cdot 3^n, 3\cdot 3^n] \text{ with } |I| = c_1 n/2 \text{ there exists } w' \in \mathsf{Words}^n(\tau,\eta) \text{ such that } w|I \sqsubseteq w' \end{array}\right\}.$$

We will see in the proof of Lemma 8.5.2 below that under appropriate conditions, there is precisely one element in $\mathsf{Output}^n(\tau,\eta,\psi)$.

Definition 8.5.3. *We define*

$$\text{BigAlg}^n : [0, 3^{10\alpha n}]^{[1,3^{\alpha n}]} \times \mathcal{C}^{2\cdot 3^{10\alpha n}} \times \bigcup_{k \geq n^2} \mathcal{C}^{[-k,k]} \to \mathcal{C}^{[-3\cdot 3^n, 3\cdot 3^n]}$$

as follows: If $\mathsf{Output}^n(\tau,\eta,\psi) \neq \varnothing$, then we define $\text{BigAlg}^n(\tau,\eta,\psi)$ to be its lexicographically smallest element. Otherwise we set $\text{BigAlg}^n(\tau,\eta,\psi) := (1)_{[-3\cdot 3^n, 3\cdot 3^n]}$.

8.5.2 Definitions of functionals and events

Below we will need some special linear functionals. Recall the definition of $\{\vec{x}_i^*\}_{i\in[1,5]}$ from Definition 9.7.5.

Definition 8.5.4. *Let $\xi \in \mathcal{C}^{\mathbb{Z}}$.*

1. *Let $z \in \mathbb{Z}$ be such that $\xi_z \neq \xi_{z-1}$, and let $B_{i,z}^{\leftarrow}$ denote the ith block of $\xi^{\leftrightarrow}|]-\infty, z]$, where $\xi^{\leftrightarrow}$ denotes the reflected scenery, defined by $\xi_y^{\leftrightarrow} := \xi_{-y}$ for all $y \in \mathbb{Z}$. We set*

$$\tilde{g}_{z,\xi}^{n,l} \;:=\; \bigotimes_{i=1}^{c_1 n} \left(|B_{i,z}^{\leftarrow}| + 1\right) \cdot \vec{x}^*_{|B_{i,z}^{\leftarrow}| \wedge 5}$$

and call $g_{z,\xi}^{n,l} := 3^{2n} p^{-c_1 n - 2} \tilde{g}_{z,\xi}^{n,l}$ the left functional of ξ at z.

2. *Let $z \in \mathbb{Z}$ be such that $\xi_z \neq \xi_{z-1}$, and let $B_{i,z}^{\rightarrow}$ denote the ith block of $\xi|[z-1, \infty[$. We define the* right functional of ξ at z *by*

$$g_{z,\xi}^{n,r} := \bigotimes_{i=1}^{c_1 n} \left(|B_{i,z}^{\rightarrow}| + 1\right) \cdot \vec{x}^*_{|B_{i,z}^{\rightarrow}| \wedge 5}.$$

Clearly, $g_{z,\xi}^{n,l}$ and $g_{z,\xi}^{n,r}$ are positive linear functionals.

Definition 8.5.5. *Let $\xi \in \mathcal{C}^{\mathbb{Z}}$.*

1. *Let $x_1 \in \mathbb{Z}$ such that $\xi_{x_1} \neq \xi_{x_1-1}$. We call a positive linear functional g a* left limiting functional of ξ at x_1 *iff for all $x_2 > x_1$ with $\xi_{x_2-1} \neq \xi_{x_2}$ we have that for all $x \in [-3^n, 3^n]$, $P_{x,\xi}(S_{o_l^n} = x_2) > 0$ implies $g(P_{x,\xi}(\mathrm{O}_1^n \in \cdot | S_{o_l^n} = x_2)) = 0$, whilst $g(P_{x,\xi}(\mathrm{O}_1^n \in \cdot | S_{o_l^n} = x_1)) > 0$.*

2. *Let $y_1 \in \mathbb{Z}$ such that $\xi_{y_1} \neq \xi_{y_1-1}$. We call a positive linear functional g a* right limiting functional of ξ at y_1 *iff for all $y_2 < y_1$ with $\xi_{y_2} \neq \xi_{y_2-1}$ we have that for all $x \in [-3^n, 3^n]$, $P_{x,\xi}(S_{o_r^n} = y_2) > 0$ implies $g(P_{x,\xi}(\mathrm{O}_3^n \in \cdot | S_{o_r^n} = y_2)) = 0$, whilst $g(P_{x,\xi}(\mathrm{O}_3^n \in \cdot | S_{o_r^n} = y_1)) > 0$.*

In the remainder, we abbreviate

$$\chi^n := \chi|[0, 2 \cdot 3^{10\alpha n}[.$$

We define in alphabetical order events which will be needed below. The event $B^n_{\text{blocks bd}}$ holds if the lengths of any $c_1 n$ consecutive blocks are bounded in a certain sense in a region around the origin. $B^n_{\text{functional}}$ is the event that $g^{n,l}_{z,\xi}$ and $g^{n,r}_{z,\xi}$ are limiting functionals for all z not too large. $B^{n,\tau}_{\mathrm{O}_2}$ gives bounds on the length of $\mathrm{O}_2^n(\chi)$. If $B^n_{\text{scen ok}}$ holds, then for every word w of length $c_1 n/2$ there exist blocks to the left and to the right of w which are close to w. $B^n_{\text{unique fit}}$ guarantees that all words of length $c_1 n/4$ in a certain region of the scenery are distinct. Blocks of lengths 2 and 4 play a special role in the arguments below. $B^n_{\text{blocks 2,4}}$ guarantees that there are sufficiently many blocks of lengths 2 and 4 in the scenery. In Definition 8.5.12 we introduce a convenient notation for a sequence of blocks of lengths 2 and 4. B^n_{signals} denotes the event that certain sequences of blocks of lengths 2 and 4 can only be observed to the left or to the right of a point in the scenery. Finally, $E^{n,\tau}_{\text{Words ok}}$ is the event that all words in $\mathsf{Words}^n(\tau, \chi^n)$ are contained (up to a possible reflection) in $\xi|[-3^{3n}, 3^{3n}]$ and $\mathsf{Words}^n(\tau, \chi^n)$ contains sufficiently many words.

Definition 8.5.6. *Let $c_6 > 0$ be as in Section 9.2. Recall the definitions of $B^{\rightarrow}_{i,z}$ and $B^{\leftarrow}_{i,z}$ from Definition 8.5.4. We define $B^n_{\text{blocks bd}} := B^{n,\rightarrow}_{\text{bb}} \cap B^{n,\leftarrow}_{\text{bb}}$ with $B^{n,\rightarrow}_{\text{bb}} :=$*

$$\{\forall z \in \left[-2 \cdot 3^{3n}, 2 \cdot 3^{3n}\right] \text{ we have } \prod_{i=1}^{c_1 n} [|B^{\rightarrow}_{i,z}| + 1] \leq e^{c_6 n} \text{ and } \sum_{i=1}^{c_1 n} [|B^{\rightarrow}_{i,z}| + 2] \leq 8 c_1 n\},$$

and $B^{n,\leftarrow}_{\text{bb}}$ is defined by replacing "$\rightarrow$" by "$\leftarrow$" in the definition of $B^{n,\rightarrow}_{\text{bb}}$.

Definition 8.5.7. *Let c_2 be as in Section 9.2. We define*

$$B^n_{\text{blocks 2,4}} := \left\{ \begin{array}{l} \text{In any sequence of } c_1 n \text{ consecutive blocks of } \xi|\,[-7 \cdot 3^n, 7 \cdot 3^n] \\ \text{there are at least } c_2 n \text{ blocks of length 2 or 4.} \end{array} \right\}.$$

Definition 8.5.8. *Let c_7 be as in Section 9.2. We define $B^{n,\tau}_{\varepsilon} := \left\{ \|\varepsilon^{n,\tau}_{\xi,S}\|_1 \leq e^{-4c_7 n} \right\}$.*

Definition 8.5.9. *We define $B^n_{\text{functional}} := B^{n,l}_{\text{func}} \cap B^{n,r}_{\text{func}}$ with*

$$B^{n,l}_{\text{func}} := \left\{ \begin{array}{l} \text{For all } y \in [-6 \cdot 3^n, 6 \cdot 3^n] \text{ with } \xi_y \neq \xi_{y-1} \text{ the left functional at } y \text{ is} \\ \text{a left limiting functional at } y \end{array} \right\},$$

$$B^{n,r}_{\text{func}} := \left\{ \begin{array}{l} \text{For all } y \in [-6 \cdot 3^n, 6 \cdot 3^n] \text{ with } \xi_y \neq \xi_{y-1} \text{ the right functional at } y \\ \text{is a right limiting functional at } y \end{array} \right\}.$$

Definition 8.5.10. *We define the event $B^{n,\tau}_{\mathrm{O}_2} := B^{n,\tau}_{\mathrm{O}_2 \text{ small}} \cap B^n_{\mathrm{O}_2 \text{ large}}$ with*

$$B^{n,\tau}_{\mathrm{O}_2 \text{ small}} := \{\forall k \in [1, 3^{\alpha n}] : |\mathrm{O}_2^n(\theta^{\tau_k}\chi)| \leq 3^n\},$$

$$B^n_{\mathrm{O}_2 \text{ large}} := \left\{\forall \xi \in \mathcal{C}^{\mathbb{Z}} \text{ and } \forall x \in [-3^n, 3^n] : P_{x,\xi}(|\mathrm{O}_2^n(\chi)| > 3^n) \leq e^{-8c_7 n}\right\}.$$

Definition 8.5.11. *We define* $B^n_{\text{scen ok}} :=$

$$\left\{\begin{matrix}\text{For all intervals } I \subseteq [-5\cdot 3^n, 5\cdot 3^n] \text{ of length } c_1 n/2 \text{ there exist } y,z \in \mathbb{Z} \text{ such}\\ \text{that } |y-z| < c_1 n,\ I \subseteq [y,z],\ \xi_y \neq \xi_{y-1},\ \text{and } \xi_z \neq \xi_{z-1}.\end{matrix}\right\}.$$

Definition 8.5.12. *Let* $n_{2,4}$ *be the number of blocks of length* 2 *and* 4 *in the piece of scenery* $\xi^n := \xi|[-7\cdot 3^n, 7\cdot 3^n]$. *Let* $B^{2,4}_{i,y}$ *be the* i*th block of* $\xi|[y-1,\infty[$ *of length* 2 *or* 4, *and let* $C^{2,4}_{i,y}$ *be its color. We can describe the blocks of length* 2 *and* 4 *of* ξ^n *by* $\text{col}(\xi^n) := (\text{col}_i(\xi^n) := (|B^{2,4}_{i,y}|, C^{2,4}_{i,y}))_{i\in[1,n_{2,4}]}$ *with* $y = -7\cdot 3^n$. *For* $R \in [1,n_{2,4}]^I$ *we have* $\text{col}\circ R = (\text{col}_{R_i})_{i\in I}$. *We set*

$$\hat{w}_{x,c_2n,\rightarrow} := \text{col}(\xi^n)|[x, x+c_2n[, \quad \hat{w}_{x,c_2n,\leftarrow} := (\text{col}_{x-i}(\xi^n); i \in [0,c_2n[) \tag{8.5.6}$$

for all x *where this makes sense. For all other* x, *we set* $\hat{w}_{x,c_2n,\rightarrow}, \hat{w}_{x,c_2n,\leftarrow} := ((1,1))_{i\in[0,c_2n[}$. *We denote by* $\bar{\xi}^n$ *the scenery obtained from* ξ^n *by replacing all* 0*'s by* 1*'s and all* 1*'s by* 0*'s. We define* $\bar{w}_{x,c_2n,\rightarrow}$, $\bar{w}_{x,c_2n,\leftarrow}$ *by replacing* ξ^n *by* $\bar{\xi}^n$ *in (8.5.6).*

Definition 8.5.13. *We call* $R \in \mathbb{Z}^{[a,b]}$ *a nearest-neighbor path if* $R_{i+1} - R_i \in \{-1,+1\}$ *for all* $i \in [a,b[$. *We define* $B^n_{\text{signals}} := B^{n,l}_{\text{sign}} \cap B^{n,r}_{\text{sign}}$ *with*

$$\begin{aligned} B^{n,l}_{\text{sign}} &:= \left\{\begin{matrix}\forall x \in [1,n_{2,4}]\ \forall \text{ nearest-neighbor path } R \in [1,n_{2,4}]^{[0,c_2n[} \text{ with } R_0 > x\\ \text{we have } \text{col}(\xi^n)\circ R \notin \{\hat{w}_{x,c_2n,\leftarrow}, \bar{w}_{x,c_2n,\leftarrow}\}\end{matrix}\right\},\\ B^{n,r}_{\text{sign}} &:= \left\{\begin{matrix}\forall x \in [1,n_{2,4}]\ \forall \text{ nearest-neighbor path } R \in [1,n_{2,4}]^{[0,c_2n[} \text{ with } R_0 < x\\ \text{we have } \text{col}(\xi^n)\circ R \notin \{\hat{w}_{x,c_2n,\rightarrow}, \bar{w}_{x,c_2n,\rightarrow}\}\end{matrix}\right\}. \end{aligned}$$

Definition 8.5.14. *For* $z \in \mathbb{Z}$ *and* $m \in \mathbb{N}$ *we define* $w_{z,m,\rightarrow} := \xi|[z, z+m[$ *to be the word of length* m *starting at* z, *and we denote by* $w_{z,m,\leftarrow}$ *the word obtained by reading* $w_{z,m,\rightarrow}$ *from right to left. We define*

$$B^n_{\text{unique fit}} := \left\{\begin{matrix}\forall z_1, z_2 \in [-3^{3n}, 3^{3n}] \text{ and } \forall i_1, i_2 \in \{\leftarrow,\rightarrow\} \text{ with } (z_1,i_1) \neq\\ (z_2,i_2) \text{ we have } w_{z_1,i_1,c_1n/4} \neq w_{z_2,i_2,c_1n/4}\end{matrix}\right\}.$$

Definition 8.5.15. *We define* $E^{n,\tau}_{\text{Words ok}} := E^{n,\tau}_{\text{only xi}} \cap E^{n,\tau}_{\text{all words}}$ *with*

$$\begin{aligned} E^{n,\tau}_{\text{only xi}} &:= \left\{\text{If } w \in \mathsf{Words}^n(\tau,\chi^n), \text{ then } w \preceq \xi|\left[-3^{3n}, 3^{3n}\right]\right\},\\ E^{n,\tau}_{\text{all words}} &:= \left\{\begin{matrix}\text{If } w \preceq \xi|[-5\cdot 3^n, 5\cdot 3^n] \text{ and } |w| = c_1n/2, \text{ then } \exists w' \in\\ \mathsf{Words}^n(\tau,\chi^n) \text{ with } w \sqsubseteq w'\end{matrix}\right\}. \end{aligned}$$

8.5.3 Combinatorics

Lemma 8.5.1. *There exists* $c_8 > 0$ *such that for all* $n \geq c_8$ *the following inclusion holds:*

$$E^{n,\tau}_{\text{stop}} \cap B^n_{\text{blocks bd}} \cap B^{n,\tau}_{\varepsilon} \cap B^n_{\text{functional}} \cap B^n_{\text{scen ok}} \subseteq E^{n,\tau}_{\text{Words ok}}.$$

Proof. Let $n \in \mathbb{N}$ and suppose the events $E^{n,\tau}_{\text{stop}}$, $B^n_{\text{blocks bd}}$, $B^{n,\tau}_{\varepsilon}$, $B^n_{\text{functional}}$, and $B^n_{\text{scen ok}}$ hold.

First we show that $E^{n,\tau}_{\text{only xi}}$ holds: Let $w \in \mathsf{Words}^n(\tau,\chi^n)$. Then there exist positive linear functionals g_1 and g_3 such that (9.7.10), (9.7.11/8.5.5), and (9.7.12) are fulfilled. Since $B^{n,\tau}_{\varepsilon}$ holds, it follows from (9.7.12) that $\|g_1 \otimes g_3\|_2 \cdot \|\varepsilon^{n,\tau}_{\xi,S}\|_1^{1/2} \leq e^{-c_7 n}$, which is

$\leq 1/(2n^2)$ for all n sufficiently large. Consequently, the assumptions (9.7.1), (9.7.2/8.4.4), and (9.7.3) of Theorem 9.7.1 are satisfied, and Theorem 9.7.1 implies $w \preceq \xi|\,[-3^{3n}, 3^{3n}]$ for all n sufficiently large.

It remains to show that $E^{n,\tau}_{\text{all words}}$ holds: Let $I \subseteq [-5\cdot 3^n, 5\cdot 3^n]$ with $|I| = c_1 n/2$. Since $B^n_{\text{scen ok}}$ holds, there exist y, z such that $|y-z| < c_1 n$, $I \subseteq [y,z]$, $\xi_{y-1} \neq \xi_y$, and $\xi_z \neq \xi_{z-1}$. For n sufficiently large, $|y|, |z| \leq 6\cdot 3^n$. We set $d := z-y+1$, $w := \xi|[y,z]$, $g_1 := g^{n,l}_{y,\xi}$, and $g_3 := g^{n,r}_{z,\xi}$ and claim that w, g_1, and g_3 satisfy (9.7.10), (9.7.11/8.5.5), and (9.7.12) with $\eta = \chi^n$ which implies $w \in \mathsf{Words}^n(\tau, \chi^n)$.

Note that $\|g \otimes g'\|_2 = \|g\|_2 \|g'\|_2$ for any g, g'. Using this together with the fact that $B^n_{\text{blocks bd}}$ holds, we obtain

$$\begin{aligned} \|g_1 \otimes g_3\|_2 &\leq 3^{2n} p^{-c_1 n - 2} \prod_{i=1}^{c_1 n} \left[|B^{\rightarrow}_{i,z}| + 1\right] \cdot \left[|B^{\leftarrow}_{i,y}| + 1\right] [\max_{i\in[1,5]} \|\vec{x}^*_i\|_2]^{2c_1 n} \\ &\leq 3^{2n} p^{-2c_1 n} e^{2c_6 n} [\max_{i\in[1,5]} \|\vec{x}^*_i\|_2]^{2c_1 n} \leq e^{c_7 n} \end{aligned} \tag{8.5.7}$$

because $c_1 n \geq 2$ and $c_7 \geq 2\ln 3 - 2c_1 \ln p + 2c_6 + 2c_1 \ln\left[\max_{i\in[1,5]} \|\vec{x}^*_i\|_2\right]$ (see Section 9.2). Hence (9.7.12) is satisfied.

Next, we verify (9.7.11) in the case $q \neq 0$. By the definition of $\mu^{n,\tau}_{\xi,S}$, we have

$$\mu^{n,\tau}_{\xi,S}[\cdot \cap E^{n,d-1}_{\text{block}}]\Pi^{-1}_{1,3} = \sum_{x\in[-3^n,3^n]} a^{n,\tau}_S(x) P_{x,\xi}[(\mathrm{O}^n_1, \mathrm{O}^n_3) \in \cdot, o^n_r = o^n_l + d - 2].$$

With $A_{u,v} := \{S_{o^n_l} = u, S_{o^n_r} = v, o^n_r = o^n_l + d - 2\}$ the following holds

$$\begin{aligned} &P_{x,\xi}[(\mathrm{O}^n_1, \mathrm{O}^n_3) \in \cdot, o^n_r = o^n_l + d - 2] \\ &= \sum_{\{u,v\in\mathbb{Z}: |u-v|\leq d-2\}} P_{x,\xi}[A_{u,v}]\, P_{x,\xi}[(\mathrm{O}^n_1, \mathrm{O}^n_3) \in \cdot | A_{u,v}] \\ &= \sum_{\{u,v\in\mathbb{Z}: |u-v|\leq d-2\}} P_{x,\xi}[A_{u,v}]\, P_{x,\xi}[\mathrm{O}^n_1 \in \cdot | A_{u,v}] \otimes P_{x,\xi}[\mathrm{O}^n_3 \in \cdot | A_{u,v}]; \end{aligned}$$

for the last equality we used that O^n_1 and O^n_3 are independent conditioned on $S_{o^n_l}$ and $S_{o^n_r}$. Let $|u-v| \leq d-2$ such that $P_{x,\xi}(A_{u,v}) > 0$. We cannot have simultaneously $u \leq y$ and $v \geq z$ because $z - y = d - 1$. Hence $u > y$ or $v < z$. Recall that we chose $g_1 = g^{n,l}_{y,\xi}$ and $g_3 = g^{n,r}_{z,\xi}$. Since the event $B^n_{\text{functional}}$ holds, g_1 and g_3 are left and right limiting functionals at y and z, respectively. Consequently, $g_1(P_{x,\xi}[\mathrm{O}^n_1 \in \cdot | A_{u,v}]) = 0$ or $g_3(P_{x,\xi}[\mathrm{O}^n_3 \in \cdot | A_{u,v}]) = 0$, and we conclude

$$(g_1 \otimes 1 \otimes g_3)(\mu^{n,\tau}_{\xi,S}[\cdot \cap E^{n,d-1}_{\text{block}}]) = 0.$$

Hence, because of $\hat{\mu}^{n,\tau}_{\xi\circ S} = \mu^{n,\tau}_{\xi,S} + \varepsilon^{n,\tau}_{\xi,S}$ and the linearity of $g_1 \otimes g_3$, we obtain

$$\begin{aligned} (g_1 \otimes 1 \otimes g_3)(\hat{\mu}^{n,\tau}_{\xi\circ S}[\cdot \cap E^{n,d-1}_{\text{block}}]) &= g_1 \otimes 1 \otimes g_3(\varepsilon^{n,\tau}_{\xi,S}[\cdot \cap E^{n,d-1}_{\text{block}}]) \\ &\leq \|g_1 \otimes g_3\|_2 \cdot \|\varepsilon^{n,\tau}_{\xi,S}[\cdot \cap E^{n,d-1}_{\text{block}}]\|_2. \end{aligned} \tag{8.5.8}$$

Since $\varepsilon^{n,\tau}_{\xi,S}$ is the difference of two probability measures and $B^{n,\tau}_{\varepsilon}$ holds, $\|\varepsilon^{n,\tau}_{\xi,S}[\cdot \cap E^{n,d-1}_{\text{block}}]\|^2_2 \leq \|\varepsilon^{n,\tau}_{\xi,S}\|_1 \leq e^{-4c_7 n}$. Thus, using (8.5.8) and (8.5.7) yields

$$(g_1 \otimes 1 \otimes g_3)(\hat{\mu}^{n,\tau}_{\xi\circ S}[\cdot \cap E^{n,d-1}_{\text{block}}]) \leq e^{-c_7 n} \leq 1/(5n^2)$$

for all n sufficiently large. Thus (9.7.11) holds.

Finally, we check that (9.7.10) holds for $q \neq 0$. Note that $\|\varepsilon^{n,\tau}_{\xi,S}\|_2 \leq \|\varepsilon^{n,\tau}_{\xi,S}\|_1^{1/2} \leq e^{-2c_7 n}$ because $\varepsilon^{n,\tau}_{\xi,S}$ is the difference of two probability measures and $B^{n,\tau}_{\varepsilon}$ holds. Since $\hat{\mu}^{n,\tau}_{\chi} = \mu^{n,\tau}_{\xi,S} + \varepsilon^{n,\tau}_{\xi,S}$ and $(g_1 \otimes 1_w \otimes g_3)(\varepsilon^{n,\tau}_{\xi,S}) \leq \|g_1 \otimes g_3\|_2 \cdot \|\varepsilon^{n,\tau}_{\xi,S}\|_2 \leq e^{-c_7 n}$ by (8.5.7), we obtain

$$(g_1 \otimes 1_w \otimes g_3)(\hat{\mu}^{n,\tau}_{\chi}[\cdot \cap E^{n,d}_{\text{block}}]) \geq (g_1 \otimes 1_w \otimes g_3)(\mu^{n,\tau}_{\xi,S}[\cdot \cap E^{n,d}_{\text{block}}]) - e^{-c_7 n}. \tag{8.5.9}$$

Since $E^{n,\tau}_{\text{stop}}$ holds, $\sum_{x \in [-3^n, 3^n]} a^{n,\tau}_S(x) = 1$. Hence, by the definition of $\mu^{n,\tau}_{\xi,S}$, it suffices to show that

$$(g_1 \otimes 1_w \otimes g_3)(P_{x,\xi}[(\mathrm{O}^n_1, \mathrm{O}^n_2, \mathrm{O}^n_3) \in \cdot, o^n_r = o^n_l + d - 1]) \geq 2 \tag{8.5.10}$$

for all $x \in [-3^n, 3^n]$ because (8.5.10) and (8.5.9) imply (9.7.10) for all n sufficiently large.

Let y_0 and z_0 denote the right end of the $c_1 n$th block of $\xi^{\leftrightarrow}|]\infty, y]$ and $\xi|[z, \infty[$, respectively; recall $\xi^{\leftrightarrow}_u = \xi_{-u}$. I.e. y_0 is the left end of the $c_1 n$th block in ξ to the left of y. The following picture illustrates this for $c_1 n = 6$. The points y and z are marked with a box.

$$\ldots \underbrace{0}_{y_0} 1101000010 \boxed{1} 1100001 \boxed{0} 1111001000110 \underbrace{1}_{z_0} \ldots$$
$$\xi|[y, z]$$

Let $\hat{o}^n_l$ denote the left end of the $c_1 n$th block of χ before o^n_l (here blocks are counted backwards), and let $\hat{o}^n_r$ denote the right end of the $c_1 n$th block of χ after o^n_r. Recall the definitions of $B^{\leftarrow}_{i,y}$ and $B^{\rightarrow}_{i,z}$ from Definition 8.5.4. We observe

$$\begin{aligned}
& P_{x,\xi}\left[(\mathrm{O}^n_1, \mathrm{O}^n_2, \mathrm{O}^n_3) \in \cdot, o^n_r = o^n_l + d - 1\right] \qquad (8.5.11)\\
\geq\ & P_{x,\xi}\left[(\mathrm{O}^n_1, \mathrm{O}^n_2, \mathrm{O}^n_3) \in \cdot, o^n_r = o^n_l + d - 1, S_{\hat{o}^n_l} = y_0, S_{o^n_l} = y, S_{o^n_r} = z, S_{\hat{o}^n_r} = z_0\right]\\
=\ & p P_{x,\xi}\left[S_{\hat{o}^n_l} = y_0\right] \bigotimes_{i=1}^{c_1 n} \lambda_r^{|B^{\leftarrow}_{i,y}|} h^{-1} \bigotimes \left[p^d \delta_w\right] \bigotimes_{i=1}^{c_1 n} \lambda_r^{|B^{\rightarrow}_{i,z}|} h^{-1}.
\end{aligned}$$

Decomposing (8.5.11) according to the different possible values for $S_{o^n_l}$ and $S_{o^n_r}$ and using Remark 8.4.2, we obtain

$$\begin{aligned}
& (g_1 \otimes 1_w \otimes g_3)(P_{x,\xi}[(\mathrm{O}^n_1, \mathrm{O}^n_2, \mathrm{O}^n_3) \in \cdot, o^n_r = o^n_l + d - 1])\\
\geq\ & (g_1 \otimes 1_w \otimes g_3)\Big(p P_{x,\xi}\left[S_{\hat{o}^n_l} = y_0\right] \bigotimes_{i=1}^{c_1 n} \lambda_r^{|B^{\leftarrow}_{i,y}|} h^{-1} \bigotimes [p^d \delta_w] \bigotimes_{i=1}^{c_1 n} \lambda_r^{|B^{\rightarrow}_{i,z}|} h^{-1}\Big)\\
=\ & 3^{2n} p^{-c_1 n + d - 1} P_{x,\xi}[S_{\hat{o}^n_l} = y_0] \prod_{i=1}^{c_1 n} \Big([|B^{\leftarrow}_{i,y}| + 1] \cdot \vec{x}^*_{|B^{\leftarrow}_{i,y}| \wedge 5}(\lambda_r^{|B^{\leftarrow}_{i,y}|} h^{-1})\Big)\\
& \cdot \prod_{i=1}^{c_1 n} \Big([|B^{\rightarrow}_{i,z}| + 1] \cdot \vec{x}^*_{|B^{\rightarrow}_{i,z}| \wedge 5}(\lambda_r^{|B^{\rightarrow}_{i,z}|} h^{-1})\Big)\\
\geq\ & 3^{2n} p^{-1} P_{x,\xi}[S_{\hat{o}^n_l} = y_0]; \qquad (8.5.12)
\end{aligned}$$

for the last estimate we used $d \leq c_1 n$ and the fact that $\vec{x}^*_{m \wedge 5}(\lambda_r^m h^{-1}) \geq (m+1)^{-1}$ for all $m \geq 1$ by Remark 8.4.1. Recall the definition of o^n_l. We have that $\hat{o}^n_l$ is the left end of

the first block of $\theta^{3^{2n}}(\xi \circ S)$. If $S_{3^{2n}} = y_0$ and $S_{3^{2n}+1} = y_0 + 1$, then $S_{\hat{o}_l^n} = y_0$. (Recall that a block of ξ starts at y_0.) Using this and the local central limit theorem (see e.g. [5] Theorem (5.2), page 132) yields

$$\begin{aligned} 3^{2n} p^{-1} P_{x,\xi}[S_{\hat{o}_l^n} = y_0] &\geq 3^{2n} p^{-1} P_{x,\xi}[S_{3^{2n}} = y_0, S_{3^{2n}+1} = y_0 + 1] \\ &= 3^{2n} P_{x,\xi}[S_{3^{2n}} = y_0] \geq c_{25} 3^{2n} 3^{-n} = c_{25} 3^n \geq 2 \end{aligned}$$

for all $n \geq c_9$ with constants $c_{25}, c_9 > 0$ independent of $x \in [-3^n, 3^n]$ and y_0; recall that $|y_0| \leq 7 \cdot 3^n$ for all n sufficiently large because $B^n_{\text{blocks bd}}$ holds. The estimate (8.5.10) follows from (8.5.12).

In the case $q = 0$, the above proof can be easily adapted. □

Lemma 8.5.2. *There exists $c_{27} > 0$ such that for all $n \geq c_{27}$ the following inclusion holds:*

$$E^{n,\tau}_{\text{Words ok}} \cap B^n_{\text{unique fit}} \subseteq E^{n,\tau}_{\text{recon Big}}.$$

Proof. Let $n \in \mathbb{N}$, and suppose $E^{n,\tau}_{\text{Words ok}}$ and $B^n_{\text{unique fit}}$ hold. Let $\psi \in \cup_{k \geq n^2} \mathcal{C}^{[-k,k]}$ with $\psi \preceq \xi|[-3^n, 3^n]$. There exist $a \in [-3^n, 3^n]$ and $b \in \{-1, 1\}$ such that

$$\psi_j = \xi_{a+bj} \quad \text{and} \quad a + bj \in [-3^n, 3^n] \quad \text{for all } j \in [-k, k]. \tag{8.5.13}$$

We argue that $w := (\xi_{a+bj})_{j \in [-3 \cdot 3^n, 3 \cdot 3^n]} \in \mathsf{Output}^n(\tau, \chi^n, \psi)$: By (8.5.13), $\psi = w|[-k, k]$. Let $I \subseteq [-3 \cdot 3^n, 3 \cdot 3^n]$ be an integer interval with $|I| = c_1 n/2$. Then the image of I under the map $j \mapsto a + bj$ is again an integer interval, which is contained in $[-5 \cdot 3^n, 5 \cdot 3^n]$ for all n sufficiently large because $|a| \leq 3^n$ and $c_1 n/2 \leq 3^n$ for all n sufficiently large. Since $E^{n,\tau}_{\text{all words}}$ holds, there exists $w' \in \mathsf{Words}^n(\tau, \chi^n)$ with $w|I \sqsubseteq w'$. Hence $w \in \mathsf{Output}^n(\tau, \chi^n, \psi)$. In particular, $\mathsf{Output}^n(\tau, \chi^n, \psi) \neq \varnothing$.

It remains to show $\xi|\,[-3^n, 3^n] \preceq w \preceq \xi|[-4 \cdot 3^n, 4 \cdot 3^n]$ for all $w \in \mathsf{Output}^n(\tau, \chi^n, \psi)$. Let $w \in \mathsf{Output}^n(\tau, \chi^n, \psi)$. Then $w|[-k, k] = \psi$, and consequently, by (8.5.13),

$$w_j = \xi_{a+bj} \tag{8.5.14}$$

for all $j \in [-k, k]$. Suppose we prove (8.5.14) for all $j \in [-3 \cdot 3^n, 3 \cdot 3^n]$. Then there is precisely one element in $\mathsf{Output}^n(\tau, \chi^n, \psi)$. Since $\psi \preceq \xi|[-3^n, 3^n]$, there are more than $2 \cdot 3^n$ letters to the left and to the right of ψ in w, and consequently $\xi|[-3^n, 3^n] \preceq w$. On the other hand, in w, there are less than $3 \cdot 3^n$ letters to the left and to the right of ψ. Hence $w \preceq \xi|[-4 \cdot 3^n, 4 \cdot 3^n]$.

Thus, to finish the proof, it suffices to verify (8.5.14) for all $j \in [-3 \cdot 3^n, 3 \cdot 3^n]$. Suppose we know (8.5.14) for all $j \in [-s, s]$ for some $s \in [k, 3 \cdot 3^n - 1]$. This assumption is true for $s = k$. We set $I_l := [-s - 1, -s - 1 + c_1/2[$, $I_r :=]s + 1 - c_1 n/2, s + 1]$, $w_l := w|I_l$, and $w_r := w|I_r$. Note that w_l and w_r have both precisely $c_1 n/2 - 1$ points in common with $w|[-s, s]$; w_l and w_r extend $w|[-s, s]$ one letter to the left and to the right, respectively. The words w_l and w_r are well defined because $c_1 n/2 \leq |\psi| = 2k + 1$ for all n sufficiently large. Since $w \in \mathsf{Output}^n(\tau, \chi^n, \psi)$, there exist $w'_l, w'_r \in \mathsf{Words}^n(\tau, \chi^n)$ with $w_l \sqsubseteq w'_l$, $w_r \sqsubseteq w'_r$. Using that $E^{n,\tau}_{\text{only xi}}$ holds, we see that $w_l, w_r \preceq \xi|\,[-3^{3n}, 3^{3n}]$.

Suppose (8.5.14) does not hold for $j = -s - 1$. Let $I_{l,\xi}$ denote the image of I_l under the map $j \mapsto a + bj$. Then $\xi|I_{l,\xi} \neq w_l$; more precisely, $\xi|I_{l,\xi}$ and w_l disagree in precisely one point, namely the leftmost point. Thus we found two words of length

$c_1 n/2$ in $\xi|[-3^{3n}, 3^{3n}]$ which disagree in precisely one point. Consequently, there exist $z, z' \in [-3^{3n}, 3^{3n}]$, $i, i' \in \{\leftarrow, \rightarrow\}$ with $(z, i) \neq (z', i')$ such that $\xi|I_{l,\xi} = w_{z,i,c_1 n/2}$ and $w_l = w_{z',i',c_1 n/2}$. If we restrict $w_{z,i,c_1 n/2}$ and $w_{z',i',c_1 n/2}$ to the last $c_1 n/4$ letters, we obtain two words of length $c_1 n/4$ in $\xi|\,[-3^{3n}, 3^{3n}]$, and these two words agree. This contradicts the fact that the event $B^n_{\text{unique fit}}$ holds. Thus (8.5.14) holds for $j = -s - 1$.

To see that (8.5.14) holds for $j = s+1$, one applies the above argument with $\bar{w}$ defined by $\bar{w}_j := w_{-j}$ for $j \in [-3 \cdot 3^n, 3 \cdot 3^n]$ in place of w. By the induction principle, (8.5.14) holds for all $j \in [-3 \cdot 3^n, 3 \cdot 3^n]$. □

Lemma 8.5.3. *There exists c_{28} such that for all $n \geq c_{28}$ the following inclusion holds:*

$$B^n_{\text{blocks bd}} \cap B^n_{\text{blocks 2,4}} \cap B^n_{\text{signals}} \subseteq B^n_{\text{functional}}.$$

Proof. The proof will be done by contradiction. Suppose the events $B^n_{\text{blocks bd}}$, $B^n_{\text{blocks 2,4}}$, and B^n_{signals} hold, but $B^n_{\text{functional}} = B^{n,l}_{\text{func}} \cap B^{n,r}_{\text{func}}$ does not hold. Suppose $B^{n,r}_{\text{func}}$ does not hold. Then there exists $y \in [-6 \cdot 3^n, 6 \cdot 3^n]$ with $\xi_y \neq \xi_{y-1}$ such that the right functional at y is not a right limiting functional at y, i.e. there exist $y_1 < y$ with $\xi_{y_1} \neq \xi_{y_1 - 1}$ and $x \in [-3^n, 3^n]$ such that $g^{\xi,n}_{y,r}(P_{x,\xi}(\mathrm{O}^n_3 \in \cdot | S_{o^n_r} = y)) = 0$ or both $P_{x,\xi}(S_{o^n_r} = y_1) > 0$ and $g^{\xi,n}_{y,r}(P_{x,\xi}(\mathrm{O}^n_3 \in \cdot | S_{o^n_r} = y_1)) \neq 0$ hold.

Let R be an admissible piece of path. If $\xi \circ R$ consists of precisely k blocks, we say that R *generates k blocks on ξ*. We denote by $\xi|[b^R_{i,l}, b^R_{i,r}]$ the block of ξ on which the ith block of $\xi \circ R$ is generated. If $R_{b^R_{i,l}} = R_{b^R_{i,r}}$, we set $j^R_i := l$, otherwise we set $j^R_i := r$. We abbreviate $l^R_i := b^R_{i,r} - b^R_{i,l} - 1$. Using this notation, we have

$$P_{x,\xi}(\mathrm{O}^n_3 \in \cdot, S_{o^n_r} = y) = \sum_{(l_i, j_i)} \bigotimes_{i=1}^{c_1 n} \lambda^{l_i} h^{-1}_{j_i}, \tag{8.5.15}$$

where the sum is taken over all $(l_i, j_i)_{i \in [1, c_1 n]} \in (\mathbb{N} \times \{l, r\})^{[1, c_1 n]}$ such that there exists an admissible piece of path R starting at y which generates blocks with $(l^R_i, j^R_i) = (l_i, j_i)$. Since $B^n_{\text{blocks bd}}$ holds, the path which starts at y and walks $6c_1 n$ (which is $\leq 3^n$ for all n sufficiently large) steps to the right generates at least $c_1 n$ blocks on ξ, namely $B^{\rightarrow}_{i,y}$, $i \in [1, c_1 n]$. Consequently, by the definition of the right functional of ξ at y and Remark 8.4.1, we have $g^{\xi,n}_{y,r}(P_{x,\xi}(\mathrm{O}^n_3 \in \cdot, S_{o^n_r} = y)) > 0$. Hence, by our assumption, $P_{x,\xi}(S_{o^n_r} = y_1) > 0$ and $g^{\xi,n}_{y,r}(P_{x,\xi}(\mathrm{O}^n_3 \in \cdot | S_{o^n_r} = y_1)) \neq 0$. Writing $P_{x,\xi}(\mathrm{O}^n_3 \in \cdot, S_{o^n_r} = y_1)$ as a sum as in (8.5.15), we see that for at least one admissible piece of path R starting at y_1 and generating at least $c_1 n$ blocks on ξ we have $g^{\xi,n}_{y,r}(\otimes_{i=1}^{c_1 n} \lambda^{l^R_i}_{j^R_i} h^{-1}) > 0$. Inserting the definition of $g^{\xi,n}_{y,r}$, we obtain

$$0 < g^{\xi,n}_{y,r}\Big(\bigotimes_{i=1}^{c_1 n} \lambda^{l^R_i}_{j^R_i} h^{-1} \Big) = \prod_{i=1}^{c_1 n} [|B^{\rightarrow}_{i,y}| + 1] \cdot \vec{x}^*_{|B^{\rightarrow}_{i,y}| \wedge 5}(\lambda^{l^R_i}_{j^R_i} h^{-1}).$$

By Remark 8.4.1, $\vec{x}^*_2(\lambda^m_i h^{-1}) \neq 0$ iff $i = r$ and $m = 2$, and also, $\vec{x}^*_4(\lambda^m_i h^{-1}) \neq 0$ iff $i = r$ and $m = 4$. Furthermore, $x^*_i(\lambda^2_r) = 0$ and $x^*_i(\lambda^4_r) = 0$ for $i \in \{1, 3, 5\}$. Thus $|B^{\rightarrow}_{i,y}| \in \{2, 4\}$ iff $l^R_i \in \{2, 4\}$ and R crosses the block $B^{\rightarrow}_{i,y}$ from left to right. Since $|y| \leq 6 \cdot 3^n$ and $B^n_{\text{blocks bd}}$ holds, we have $B^{\rightarrow}_{i,y} \sqsubseteq \xi|\,[-7 \cdot 3^n, 7 \cdot 3^n]$ for all n sufficiently large and $i \in [1, c_1 n]$. Using that $B^n_{\text{blocks 2,4}}$ holds, we see that at least $c_2 n$ of the blocks $B^{\rightarrow}_{i,y}$, $i \in [1, c_1 n]$, have length 2 or 4. Hence there are $\geq c_2 n$ blocks with $l^R_i \in \{2, 4\}$.

Clearly, the color of two successive blocks in ξ, and also in the observations, must be different. Hence the colors of the blocks of length 2 or 4 among the first $c_1 n$ blocks of $\xi \circ R$ either all agree with the colors of the blocks $B^{\rightarrow}_{i,y}$, $i \in [1, c_1 n]$, of length 2 or 4 or they have all the opposite color. But this contradicts the fact that $B^{n,r}_{\text{sign}}$ holds. A similar argument shows that the assumption that $B^{n,l}_{\text{func}}$ holds leads to a contradiction. □

8.5.4 Probabilistic estimates

In this section, we prove that the complements of all the basic events $B^n_{\dots}$ defined in Section 8.5.2 have a probability which is exponentially small in n; for some events this is only true under the assumption that $E^{n,\tau}_{\text{stop}}$ holds. We treat the events in alphabetical order.

Lemma 8.5.4. *There exist $c_{29}, c_5 > 0$ such that for all $n \geq c_{29}$*

$$P\left([B^n_{\text{blocks bd}}]^c\right) \leq 2e^{-c_5 n}.$$

Proof. By the definition of $B^n_{\text{blocks bd}} = B^{n,\rightarrow}_{\text{bb}} \cap B^{n,\leftarrow}_{\text{bb}}$,

$$[B^{n,\rightarrow}_{\text{bb}}]^c = \bigcup_{z \in [-2\cdot 3^{3n}, 2\cdot 3^{3n}]} \Big\{ \prod_{i=1}^{c_1 n} [|B^{\rightarrow}_{i,z}| + 1] > e^{c_6 n} \Big\} \cup \Big\{ \sum_{i=1}^{c_1 n} [|B^{\rightarrow}_{i,z}| + 2] > 8 c_1 n \Big\}.$$

For each z, the block lengths $|B^{\rightarrow}_{i,z}|$, $i \geq 1$, are i.i.d. with $P(|B^{\rightarrow}_{i,z}| = k) = 2^{-k}$, $k \geq 1$; in particular $E|B^{\rightarrow}_{i,z}| = 2$. By Chebyshev's inequality, we obtain

$$P\Big(\prod_{i=1}^{c_1 n} [|B^{\rightarrow}_{i,z}| + 1] > e^{c_6 n} \Big) \;\leq\; e^{-c_6 n} E\Big(\prod_{i=1}^{c_1 n} [|B^{\rightarrow}_{i,z}| + 1] \Big) = 3^{c_1 n} e^{-c_6 n}.$$

Furthermore, by the large deviation principle, we have

$$P\Big(\sum_{i=1}^{c_1 n} [|B^{\rightarrow}_{i,z}| + 2] > 8 c_1 n \Big) = P\Big(\sum_{i=1}^{c_1 n} |B^{\rightarrow}_{i,z}| > 6 c_1 n \Big) \leq e^{-c_1 n I(6)}$$

with rate function $I(x) = (x-1)\ln(x-1) + x \ln(2/x)$. Since $I(6) > 1$, we conclude

$$P([B^{n,\rightarrow}_{\text{bb}}]^c) \;\leq\; (4 \cdot 3^{3n} + 1) \left[3^{c_1 n} e^{-c_6 n} + e^{-c_1 n} \right] \leq e^{-c_5 n}$$

for some constant $c_5 > 0$ for all n sufficiently large; here we used that $c_6 - (c_1 + 4)\ln 3 > 0$ by our choice of c_6 and $c_1 > 4 \ln 3$. The same estimate holds for $P([B^{n,\leftarrow}_{\text{bb}}]^c)$. □

Lemma 8.5.5. *There exist $c_{31} > 0$ such that for all $n \in \mathbb{N}$*

$$P\left(\left[B^n_{\text{blocks } 2,4} \right]^c \right) \leq 14 e^{-c_{31} n}.$$

Proof. Recall that for all z, the block lengths $|B^{\rightarrow}_{i,z}|$, $i \geq 1$, are i.i.d. with $P(|B^{\rightarrow}_{i,z}| = k) = 2^{-k}$, $k \geq 1$. Hence $P(|B^{\rightarrow}_{i,z}| \in \{2,4\}) = 2^{-2} + 2^{-4} = 5/16$. Let Y_k, $k \geq 1$, be i.i.d. Bernoulli with parameter $5/16$, and let $J(x) := (1-x)\ln\left(\frac{16(1-x)}{11}\right) + x \ln\left(\frac{16x}{5}\right)$. By the large deviation principle (see e.g. [5]), $P(\sum_{k=1}^{c_1 n} Y_k \leq c_1 n/4) \leq e^{-J(1/4) c_1 n}$. Since $c_2 < c_1/4$ and there are at most $14 \cdot 3^n$ sequences of $c_1 n$ consecutive blocks in $\xi | [-7 \cdot 3^n, 7 \cdot 3^n]$, we have

$$P\left(\left[B^n_{\text{blocks } 2,4} \right]^c \right) \leq 14 \cdot 3^n e^{-J(1/4) c_1 n} \leq 14 e^{-c_{31} n}$$

because $J(1/4) c_1 - \ln 3 > 0$. □

Recall that $3^{\alpha n} a_S^{n,\tau}(x)$ equals the number of stopping times τ_k, $k \in [1, 3^{\alpha n}]$, with $S_{\tau_k} = x$. The following lemma, which will be needed in the proof of Lemma 8.5.8, states that with very high probability, the stopping times stop often in x provided the event $E_{\text{stop}}^{n,\tau}$ holds.

Lemma 8.5.6. *There exists $c_{32} > 0$ such that for all $n \geq c_{32}$*

$$P\Big(E_{\text{stop}}^{n,\tau} \cap \bigcup_{x \in [-3^n, 3^n]} \big\{3^{\alpha n} a_S^{n,\tau}(x) \leq 3^{17 c_1 n} e^{16 c_7 n}\big\}\Big) \leq e^{-n}.$$

Proof. The proof is very similar to the proof of Lemma 6.14 in [22]. In the notation of [22], the estimate holds whenever $\alpha > 1 + \gamma - [3c_1 \ln p]/\ln 3$ with $\gamma := 17c_1 + 16c_7/\ln 3$, which is satisfied by our choice of α (see Section 9.2). □

The following basic large deviation estimate will be needed below.

Lemma 8.5.7. *Let X_i, $i \geq 1$, be i.i.d. Bernoulli with parameter δ, and let $\sigma_m := \sum_{i=1}^m X_i$. There exists a constant $c_{33} > 0$ such that for all $m \in \mathbb{N}$ and all $a > 0$*

$$P(\sigma_m \geq m(a + \delta)) \leq e^{-c_{33} m a^2}.$$

Proof. By the large deviation principle (see e.g. [5]), we have $P(\sigma_m \geq m(a+\delta)) \leq e^{-m I_\delta(a+\delta)}$ with rate function $I_\delta(a) = a \ln\left(\frac{a}{\delta}\right) + (1-a) \ln\left(\frac{1-a}{1-\delta}\right)$. One verifies that $I_\delta(a+\delta) \geq c_{33} a^2$ for all $\delta \in]0,1[$ and $a \in]0, 1-\delta[$ with a constant $c_{33} > 0$ independent of δ and a. □

Lemma 8.5.8. *There exist constants $c_2, c_{35}, c_{36} > 0$ such that*

$$P\left(E_{\text{stop}}^{n,\tau} \setminus B_\varepsilon^{n,\tau}\right) \leq c_{35} e^{-c_{36} n} \quad \text{for all } n \geq c_2.$$

Proof. We define for $x \in [-3^n, 3^n]$

$$\hat{\mu}_{x,\xi\circ S}^{n,\tau} := [3^{\alpha n} a_S^{n,\tau}(x)]^{-1} \sum_{k \in [1, 3^{\alpha n}]} 1\{S_{\tau_k} = x\} \delta_{\mathrm{O}^n(\theta^{\tau_k}\chi)},$$

i.e. $\hat{\mu}_{x,\xi\circ S}^{n,\tau}$ is the empirical distribution of the O^n collected after times τ_k with $S_{\tau_k} = x$. Suppose the event $E_{\text{stop}}^{n,\tau}$ holds. Then $|S_{\tau_k}| \leq 3^n$ for all $k \in [1, 3^{\alpha n}]$, and consequently

$$\varepsilon_{\xi,S}^{n,\tau} = \sum_{x \in [-3^n, 3^n]} a_S^{n,\tau}(x) \left[\hat{\mu}_{x,\xi\circ S}^{n,\tau} - P_{x,\xi}\left[\mathrm{O}^n(\chi)\right]^{-1}\right].$$

By the triangle inequality,

$$\|\varepsilon_{\xi,S}^{n,\tau}\|_1 \leq \sum_{x \in [-3^n, 3^n]} a_S^{n,\tau}(x) \left\|\hat{\mu}_{x,\xi\circ S}^{n,\tau} - P_{x,\xi}\left[\mathrm{O}^n(\chi)\right]^{-1}\right\|_1. \tag{8.5.16}$$

Let $\mathcal{S}$ denote the set of possible states of the random variable $\mathrm{O}^n(\chi)$ if $|\mathrm{O}_2^n(\chi)| \leq 3^n$, and let $\mathcal{S}'$ be the set of possible states of $\mathrm{O}^n(\chi)$ if $|\mathrm{O}_2^n(\chi)| > 3^n$. Recall that $\mathrm{O}^n = (\mathrm{O}_1^n, \mathrm{O}_2^n, \mathrm{O}_3^n)$ where O_1^n, $\mathrm{O}_3^n \in \{1, 2, \ldots, 5\}^{c_1 n}$ and O_2^n is the concatenation of a word of length $< c_1 n/2$ with a block. Consequently, $|\mathcal{S}| \leq 5^{2c_1 n} 2^{c_1 n} 3^n \leq 2^{8 c_1 n}$.

Recall the definition of $B_{\mathrm{O}_2}^{n,\tau}$ from Definition 8.5.10. Clearly,

$$P\left(E_{\text{stop}}^{n,\tau} \setminus B_\varepsilon^{n,\tau}\right) \leq P\left(\left[E_{\text{stop}}^{n,\tau} \cap B_{\mathrm{O}_2}^{n,\tau}\right] \setminus B_\varepsilon^{n,\tau}\right) + P\left(E_{\text{stop}}^{n,\tau} \setminus B_{\mathrm{O}_2}^{n,\tau}\right). \tag{8.5.17}$$

We split the sum in (8.5.16) in two parts. Let

$$J_{\text{seldom}} := \left\{x \in [-3^n, 3^n] : 3^{\alpha n} a_S^{n,\tau}(x) \le 3^n |\mathcal{S}|^2 e^{16 c_7 n}\right\}, \quad J_{\text{often}} := [-3^n, 3^n] \setminus J_{\text{seldom}}.$$

By the definition of J_{seldom}, we have

$$\sum_{x \in J_{\text{seldom}}} a_S^{n,\tau}(x) \left\| \hat{\mu}_{x,\xi\circ S}^{n,\tau} - P_{x,\xi}\left[\mathrm{O}^n(\chi)\right]^{-1} \right\|_1 \le 3^{(1-\alpha)n} 2^{16 c_1 n} e^{16 c_7 n} \le e^{-8 c_7 n}, \tag{8.5.18}$$

where the last inequality follows from our choice of α. Next, we define the event that the contribution to $\|\varepsilon_{\xi,S}^{n,\tau}\|_1$ coming from $\mathrm{O}^n = s \in \mathcal{S}$ is small: We set for $x \in [-3^n, 3^n]$ and $s \in \mathcal{S}$

$$B_{x\ \text{often}}^{n,\tau,s} := \left\{ \text{ If } x \in J_{\text{often}}, \text{ then } \left| \hat{\mu}_{x,\xi\circ S}^{n,\tau}(\{s\}) - P_{x,\xi}\left[\mathrm{O}^n(\chi)\right]^{-1}(\{s\}) \right| \le |\mathcal{S}|^{-1} e^{-8 c_7 n} \right\}.$$

If the event $\cap_{x\in[-3^n,3^n]} \cap_{s\in\mathcal{S}} B_{x\ \text{often}}^{n,\tau,s}$ holds, then

$$\sum_{x \in J_{\text{often}}} a_S^{n,\tau}(x) \sum_{s\in\mathcal{S}} \left| \hat{\mu}_{x,\xi\circ S}^{n,\tau}(\{s\}) - P_{x,\xi}\left[\mathrm{O}^n(\chi)\right]^{-1}(\{s\}) \right| \ \le \ e^{-8 c_7 n}. \tag{8.5.19}$$

If the event $B_{\mathrm{O}_2}^{n,\tau}$ holds, then $\hat{\mu}_{x,\xi\circ S}^{n,\tau}(\{s\}) = 0$ for all $s \in \mathcal{S}'$ and consequently,

$$\begin{aligned} &\sum_{x \in J_{\text{often}}} a_S^{n,\tau}(x) \sum_{s\in\mathcal{S}'} \left| \hat{\mu}_{x,\xi\circ S}^{n,\tau}(\{s\}) - P_{x,\xi}\left[\mathrm{O}^n(\chi)\right]^{-1}(\{s\}) \right| \\ \le &\sum_{x \in J_{\text{often}}} a_S^{n,\tau}(x) P_{x,\xi}(|\mathrm{O}_2^n(\chi)| > 3^n) \le e^{-8 c_7 n}. \end{aligned}$$

Combining the last estimate with (8.5.19) and (8.5.18), we obtain

$$\begin{aligned} E_{\text{stop}}^{n,\tau} \cap B_{\mathrm{O}_2}^{n,\tau} \cap \bigcap_{x\in[-3^n,3^n]} \bigcap_{s\in\mathcal{S}} B_{x\ \text{often}}^{n,\tau,s} &\subseteq E_{\text{stop}}^{n,\tau} \cap B_{\mathrm{O}_2}^{n,\tau} \cap \left\{ \|\varepsilon_{\xi,S}^{n,\tau}\|_1 \le 3 e^{-8 c_7 n} \right\} \\ &\subseteq E_{\text{stop}}^{n,\tau} \cap B_{\mathrm{O}_2}^{n,\tau} \cap B_{\varepsilon}^{n,\tau} \end{aligned}$$

for all n sufficiently large. Hence, using $\Omega = \{x \in J_{\text{seldom}}\} \cup \{x \in J_{\text{often}}\}$, we obtain

$$\begin{aligned} &P\left(\left[E_{\text{stop}}^{n,\tau} \cap B_{\mathrm{O}_2}^{n,\tau} \right] \setminus B_{\varepsilon}^{n,\tau} \right) \le P\Big(E_{\text{stop}}^{n,\tau} \cap B_{\mathrm{O}_2}^{n,\tau} \cap \bigcup_{x\in[-3^n,3^n]} \bigcup_{s\in\mathcal{S}} [B_{x\ \text{often}}^{n,\tau,s}]^c \Big) \qquad (8.5.20) \\ \le\ & P\Big(E_{\text{stop}}^{n,\tau} \cap \bigcup_{x\in[-3^n,3^n]} \{x \in J_{\text{seldom}}\} \Big) + P\Big(\bigcup_{x\in[-3^n,3^n]} \bigcup_{s\in\mathcal{S}} [\{x \in J_{\text{often}}\} \setminus B_{x\ \text{often}}^{n,\tau,s}] \Big) \\ \le\ & P\Big[E_{\text{stop}}^{n,\tau} \cap \bigcup_{x\in[-3^n,3^n]} \{x \in J_{\text{seldom}}\} \Big] + 3^{2n} |\mathcal{S}| \max_{x\in[-3^n,3^n], s\in\mathcal{S}} P\Big[\{x \in J_{\text{often}}\} \setminus B_{x\ \text{often}}^{n,\tau,s} \Big]. \end{aligned}$$

It follows from $|\mathcal{S}| \le 2^{8 c_1 n}$ and Lemma 8.5.6 that for all $n \ge c_{32}$

$$\begin{aligned} P\Big[E_{\text{stop}}^{n,\tau} \cap \bigcup_{x\in[-3^n,3^n]} \{x \in J_{\text{seldom}}\} \Big] &\le P\Big[E_{\text{stop}}^{n,\tau} \cap \bigcup_{x\in[-3^n,3^n]} \left\{ 3^{\alpha n} a_S^{n,\tau}(x) \le 3^{17 c_1 n} e^{16 c_7 n} \right\} \Big] \\ &\le e^{-n}. \end{aligned} \tag{8.5.21}$$

We introduce the stopping times τ_k^x when the random walker is at x: $\tau_1^x := \min\{\tau_i : i \in [1, 3^{\alpha n}], S_{\tau_i} = x\}$, $\tau_{k+1}^x := \min\{\tau_i > \tau_k^x : i \in [1, 3^{\alpha n}], S_{\tau_i} = x\}$. The random variables $\chi|[\tau_k^x + 3^{2n}, \tau_k^x + 3^{3n}[$, $k \in [1, j]$, are i.i.d. conditioned on $E_{\text{stop}}^{n,\tau}$. Hence, by the definition of $\hat{\mu}_{x,\xi\circ S}^{n,\tau}$, $P\left(\{x \in J_{\text{often}}\} \setminus B_{x\ \text{often}}^{n,\tau,s} | E_{\text{stop}}^{n,\tau}\right)$ equals a large deviation probability for sums of Bernoulli random variables and we can apply Lemma 8.5.7 with $m = 3^{\alpha n} a_S^{n,\tau}(x) > 3^n |\mathcal{S}|^2 e^{16 c_7 n}$ and $a = |\mathcal{S}|^{-1} e^{-8 c_7 n}$. Since for this choice, $ma^2 > 3^n$ we obtain

$$P\left(\{x \in J_{\text{often}}\} \setminus B_{x\ \text{often}}^{n,\tau,s}\right) \le \exp(-c_{33} 3^n). \tag{8.5.22}$$

Combining (8.5.20) with (8.5.21), $|\mathcal{S}| \le 2^{8c_1 n}$, and (8.5.22), we conclude

$$P\left(\left[E_{\text{stop}}^{n,\tau} \cap B_{\text{O}_2}^{n,\tau}\right] \setminus B_{\varepsilon}^{n,\tau}\right) \le 2e^{-n} \tag{8.5.23}$$

for all $n \ge c_2$ with some constant $c_2 \ge c_{32}$. The claim of the lemma follows from (8.5.17), (8.5.23), and Lemma 8.5.10. □

Lemma 8.5.9. *There exist $c_{37}, c_{38}, c_{39} > 0$ such that for all $n \ge c_{37}$*

$$P\left([B_{\text{functional}}^n]^c\right) \le c_{38} e^{-c_{39} n}.$$

Proof. By Lemma 8.5.3, $B_{\text{functional}}^n \subseteq [B_{\text{blocks bd}}^n]^c \cup \left[B_{\text{blocks 2,4}}^n\right]^c \cup \left[B_{\text{signals}}^n\right]^c$. The claim follows immediately from Lemmas 8.5.4, 8.5.5, and 8.5.12. □

Lemma 8.5.10. *There exist $c_{40}, c_{41}, c_{42} > 0$ such that for all $n \ge c_{40}$*

$$P\left(E_{\text{stop}}^{n,\tau} \setminus B_{\text{O}_2}^{n,\tau}\right) \le c_{41} e^{-c_{42} n}.$$

Proof. Clearly,

$$P\left(E_{\text{stop}}^{n,\tau} \setminus B_{\text{O}_2}^{n,\tau}\right) \le P\left(\left[E_{\text{stop}}^{n,\tau} \cap B_{\text{blocks bd}}^n\right] \setminus B_{\text{O}_2}^{n,\tau}\right) + P\left([B_{\text{blocks bd}}^n]^c\right). \tag{8.5.24}$$

Recall that $B_{\text{O}_2}^{n,\tau} = B_{\text{O}_2\ \text{small}}^{n,\tau} \cap B_{\text{O}_2\ \text{large}}^n$. By definition,

$$\begin{aligned} &P\left(\left[E_{\text{stop}}^{n,\tau} \cap B_{\text{blocks bd}}^n\right] \setminus B_{\text{O}_2\ \text{small}}^{n,\tau}\right) \le 3^{\alpha n} \max_{x \in [-3^n, 3^n]} P_x(B_{\text{blocks bd}}^n \cap \{|\text{O}_2^n(\chi)| > 3^n\}) \\ &= 3^{\alpha n} \max_{x \in [-3^n, 3^n]} E_x\left[1_{B_{\text{blocks bd}}^n} P_{x,\xi}(|\text{O}_2^n(\chi)| > 3^n)\right]. \end{aligned} \tag{8.5.25}$$

Let $x \in [-3^n, 3^n]$. Suppose the random walk starts at x and $|\text{O}_2^n(\chi)| > 3^n$. Then $\chi|[0, 3^{3n}[$ contains a block of length $\ge 3^n - c_1 n$ and this block must be generated on $\xi|[-2 \cdot 3^{3n}, 2 \cdot 3^{3n}]$. If $B_{\text{blocks bd}}^n$ holds, all blocks of $\xi|[-2 \cdot 3^{3n}, 2 \cdot 3^{3n}]$ have length $\le 6c_1 n$. Consequently, the random walk stays time $t \ge 3^n - c_1 n$ in an interval I of length $\le 6c_1 n$. It is known (see e.g.[21], Lemma 5.2) that

$$P(S_i \in I \text{ for all } i \in [0, t[) \le c_{43} \exp(-c_{44} t / |I|^2)$$

with constants $c_{43}, c_{44} > 0$. Thus it follows from (8.5.25)

$$P\left(\left[E_{\text{stop}}^{n,\tau} \cap B_{\text{blocks bd}}^n\right] \setminus B_{\text{O}_2\ \text{small}}^{n,\tau}\right) \le c_{43} 3^{\alpha n} \exp\left[-\frac{c_{44}[3^n - c_1 n]}{36 c_1^2 n^2}\right] \le e^{-n} \tag{8.5.26}$$

for all n sufficiently large. Furthermore, by the above argument, $\left[E_{\text{stop}}^{n,\tau} \cap B_{\text{blocks bd}}^n\right] \setminus B_{\text{O}_2\ \text{large}}^n = \varnothing$ for all n sufficiently large. Thus $P\left(\left[E_{\text{stop}}^{n,\tau} \cap B_{\text{blocks bd}}^n\right] \setminus B_{\text{O}_2}^{n,\tau}\right) \le e^{-n}$ for all n sufficiently large. The claim follows from (8.5.24) and Lemma 8.5.4. □

Lemma 8.5.11. *There exist $c_{45}, c_4 > 0$ such that for all $n \geq c_{45}$*

$$P\left([B^n_{\text{scen ok}}]^c\right) \leq 12e^{-c_4 n}.$$

Proof. It is not hard to see that for all n sufficiently large, $B^n_{\text{scen ok}}$ contains the event $\{\text{All blocks of } \xi|\,[-6\cdot 3^n, 6\cdot 3^n] \text{ have length } \leq c_1 n/4\}$. Consequently,

$$\begin{aligned} P\left([B^n_{\text{scen ok}}]^c\right) &\leq P\left(\exists \text{ block of } \xi|\,[-6\cdot 3^n, 6\cdot 3^n] \text{ of length } > c_1 n/4\right) \\ &\leq 12\cdot 3^n \cdot 2^{-c_1 n/4}; \end{aligned}$$

here we used that there are $\leq 12\cdot 3^n$ possible left endpoints for a block in $\xi|\,[-6\cdot 3^n, 6\cdot 3^n]$ and that the probability that a block starting at x has length $> c_1 n/4$ equals $2^{-c_1 n/4}$ because the scenery is i.i.d. uniformly colored. The claim follows because $c_1 > 4\ln 3/\ln 2$. □

Lemma 8.5.12. *There exists $c_{47} > 0$ such that for all $n \in \mathbb{N}$*

$$P\left(\left[B^n_{\text{signals}}\right]^c\right) \leq 60e^{-c_{47} n}.$$

Proof. Recall the notation introduced in Definitions 8.5.12 and 8.5.13. Let $y := -7\cdot 3^n$. The sequence $(|B^{2,4}_{i,y}|, C^{2,4}_{i,y})_{i\geq 1}$ is a Markov chain under P with time-homogeneous transition probabilities. The block lengths $(|B^{2,4}_{i,y}|)_{i\geq 1}$ are i.i.d. with $P(|B^{2,4}_{i,y}| = 2) = 2^{-2}/(2^{-2}+2^{-4}) = 4/5$ and $P(|B^{2,4}_{i,y}| = 4) = 1/5$ and independent of the colors $(C^{2,4}_{i,y})_{i\geq 1}$. Note that $C^{2,4}_{i,y} \neq C^{2,4}_{i+1,y}$ iff between $B^{2,4}_{i,y}$ and $B^{2,4}_{i+1,y}$ there are $2k$ blocks of length 1, 3, or 5 for some $k \geq 0$. Recall the definition of $B^{\rightarrow}_{i,y}$ from Definition 8.5.4. Let $p_{2,4} := P(|B^{\rightarrow}_{i,y}| \in \{2,4\}) = 2^{-2}+2^{-4} = 5/16$ and set $q_{2,4} := 1 - p_{2,4} = 11/16$. Then

$$P(C^{2,4}_{i,y} \neq C^{2,4}_{i+1,y}) = \sum_{k=0}^{\infty} q^{2k}_{2,4} p_{2,4} = \frac{p_{2,4}}{1-q^2_{2,4}} = \frac{1}{1+q_{2,4}} = \frac{16}{27}$$

and $P(C^{2,4}_{i,y} = C^{2,4}_{i+1,y}) = 11/27$. Hence the one-step transition probabilities of the Markov chain $\text{col}_i(\xi^n)$, $i \geq 1$, are $\leq \frac{4}{5}\cdot\frac{16}{27} = \frac{64}{135} < \frac{1}{2}$.

Let $x \in [1, n_{2,4}]$, let $R \in [1, n_{2,4}]^{[0,c_2 n[}$ be a nearest-neighbor path with $R_0 < x$, and let $w \in \{\hat{w}_{x,c_2 n,\rightarrow}, \bar{w}_{x,c_2 n,\rightarrow}\}$, $w = (w_i)_{i\in[0,c_2 n[}$. We set $\mathcal{H}_k := \sigma(\text{col}_i(\xi^n); i \in [1,k])$. Clearly, $w_k \in \mathcal{H}_{x+k}$. Since R is a nearest-neighbor path, $R_k < x+k$ for all k; hence $\text{col}_{R_k} \in \mathcal{H}_{x+k-1}$ for all k. Using that w_k, $k \in [0, c_2 n[$, is a Markov chain with the above specified transition probabilities, we obtain

$$\begin{aligned} P(\text{col}(\xi^n)\circ R = w) &= P(\text{col}_i(\xi^n) = w_i \ \forall i \in [0, c_2 n[) \\ &\leq \prod_{i=0}^{c_2 n-2} P(\text{col}_{i+1}(\xi^n) = w_{i+1} | \text{col}_i(\xi^n) = w_i) \leq \left(\frac{64}{135}\right)^{c_2 n-1}. \end{aligned}$$

There are $n_{2,4} \leq 14\cdot 3^n$ possibilities to choose x and $2^{c_2 n-1}$ possibilities to choose R. Thus, by the definition of $B^{n,r}_{\text{sign}}$,

$$P\left(\left[B^{n,r}_{\text{sign}}\right]^c\right) \leq 2\cdot 14\cdot 3^n 2^{c_2 n-1}\cdot\left(\frac{64}{135}\right)^{c_2 n-1} \leq 30\cdot 3^n 2^{c_2 n}\left(\frac{64}{135}\right)^{c_2 n} \leq 30e^{-c_{47} n}$$

for some constant $c_{47} > 0$ because $64/135 < 1/2$ and $c_2 > \ln 3/(\ln(135/128))$. The same estimate holds for $B^{n,l}_{\text{sign}}$, and the claim follows from the definition of $B^n_{\text{signals}} = B^{n,l}_{\text{sign}} \cap B^{n,r}_{\text{sign}}$. □

Lemma 8.5.13. *There exists $c_{48} > 0$ such that for all $n \in \mathbb{N}$*

$$P\left(\left[B^n_{\text{unique fit}}\right]^c\right) \leq 4e^{-c_{48}n}.$$

Proof. Let $z_1, z_2 \in [-3^{3n}, 3^{3n}]$ and $i_1, i_2 \in \{\leftarrow, \rightarrow\}$ with $(z_1, i_1) \neq (z_2, i_2)$. For $k = 1, 2$, we set $o_k := +1$ if $i_k = \rightarrow$, $o_k := -1$ if $i_k = \leftarrow$, and we define $f_k(j) := z_k + o_k j$ for $j \in [0, c_1 n/4[$. As is shown in the proof of Lemma 6.8 of [22], there exists a subset $J \subseteq [0, c_1 n/4[$ of cardinality $|J| \geq c_1 n/12$ such that $f_1(J) \cap f_2(J) = \varnothing$. Consequently,

$$P(w_{z_1,i_1,c_1 n/4} = w_{z_2,i_2,c_1 n/4}) \leq P(w_{z_1,i_1,c_1 n/4} | f_1(J) = w_{z_2,i_2,c_1 n/4} | f_2(J)) = 2^{-c_1 n/12}.$$

Since there are $\leq (2 \cdot 3^{3n} + 1)^2 \leq 3^{8n}$ possibilities to choose z_1 and z_2 and ≤ 4 possibilities to choose i_1 and i_2, we conclude

$$P\left(\left[B^n_{\text{unique fit}}\right]^c\right) \leq 4 \cdot 3^{8n} 2^{-c_1 n/12} \leq 4e^{-c_{48}n}$$

for some constant $c_{48} > 0$ because $c_1 > 96 \ln 3 / \ln 2$. □

Proof of Theorem 8.3.1

Proof of Theorem 8.3.1. Combining Lemmas 8.5.2, 8.5.1, and 8.5.3 we obtain

$$E^{n,\tau}_{\text{stop}} \cap B^n_{\text{blocks bd}} \cap B^n_{\text{blocks 2,4}} \cap B^{n,\tau}_{\varepsilon} \cap B^n_{\text{functional}} \cap B^n_{\text{scen ok}} \cap B^n_{\text{signals}} \cap B^n_{\text{unique fit}} \subseteq E^{n,\tau}_{\text{recon Big}}$$

for all n sufficiently large. Hence

$$E^{n,\tau}_{\text{stop}} \setminus E^{n,\tau}_{\text{recon Big}} \subseteq [B^n_{\text{blocks bd}}]^c \cup \left[B^n_{\text{blocks 2,4}}\right]^c \cup \left[E^{n,\tau}_{\text{stop}} \setminus B^{n,\tau}_{\varepsilon}\right] \cup [B^n_{\text{functional}}]^c \cup [B^n_{\text{scen ok}}]^c \cup \left[B^n_{\text{signals}}\right]^c \cup \left[B^n_{\text{unique fit}}\right]^c.$$

The claim follows from Lemmas 8.5.4, 8.5.5, 8.5.8, 8.5.9, 8.5.11, 8.5.12, and 8.5.13. □

References

[1] I. Benjamini and H. Kesten. Distinguishing sceneries by observing the scenery along a random walk path. *J. Anal. Math.*, 69:97–135, 1996.

[2] K. Burdzy. Some path properties of iterated Brownian motion. In *Seminar on Stochastic Processes, 1992 (Seattle, WA, 1992)*, volume 33 of *Progr. Probab.*, pages 67–87. Birkhäuser Boston, Boston, MA, 1993.

[3] F. den Hollander and J. E. Steif. Mixing properties of the generalized T, T^{-1}-process. *J. Anal. Math.*, 72:165–202, 1997.

[4] W. Th. F. den Hollander. Mixing properties for random walk in random scenery. *Ann. Probab.*, 16(4):1788–1802, 1988.

[5] R. Durrett. *Probability: Theory and Examples*. Duxbury Press, Second edition, 1996.

[6] D. Heicklen, C. Hoffman, and D. J. Rudolph. Entropy and dyadic equivalence of random walks on a random scenery. *Adv. Math.*, 156(2):157–179, 2000.

[7] C. D. Howard. Detecting defects in periodic scenery by random walks on $\mathbb{Z}$. *Random Structures Algorithms*, 8(1):59–74, 1996.

[8] C. D. Howard. Orthogonality of measures induced by random walks with scenery. *Combin. Probab. Comput.*, 5(3):247–256, 1996.

[9] C. D. Howard. Distinguishing certain random sceneries on $\mathbb{Z}$ via random walks. *Statist. Probab. Lett.*, 34(2):123–132, 1997.

[10] M. Keane and W. Th. F. den Hollander. Ergodic properties of color records. *Phys. A*, 138(1-2):183–193, 1986.

[11] H. Kesten. Detecting a single defect in a scenery by observing the scenery along a random walk path. In *Itô's stochastic calculus and probability theory*, pages 171–183. Springer, Tokyo, 1996.

[12] H. Kesten. Distinguishing and reconstructing sceneries from observations along random walk paths. In *Microsurveys in discrete probability (Princeton, NJ, 1997)*, pages 75–83. Amer. Math. Soc., Providence, RI, 1998.

[13] D. Levin and Y. Peres. Random walks in stochastic scenery on $\mathbb{Z}$. Preprint, 2002.

[14] D. A. Levin, R. Pemantle, and Y. Peres. A phase transition in random coin tossing. *Ann. Probab.*, 29(4):1637–1669, 2001.

[15] E. Lindenstrauss. Indistinguishable sceneries. *Random Structures Algorithms*, 14(1):71–86, 1999.

[16] M. Löwe and H. Matzinger. Reconstruction of sceneries with correlated colors. Eurandom Report 99-032, accepted by Stochastic Processes and Their Applications, 1999.

[17] M. Löwe and H. Matzinger. Scenery reconstruction in two dimensions with many colors. *Ann. Appl. Probab.*, 12(4):1322–1347, 2002.

[18] M. Löwe, H. Matzinger, and F. Merkl. Reconstructing a multicolor random scenery seen along a random walk path with bounded jumps. Eurandom Report 2001-030. Submitted., 2001.

[19] H. Matzinger. Reconstructing a three-color scenery by observing it along a simple random walk path. *Random Structures Algorithms*, 15(2):196–207, 1999.

[20] H. Matzinger. Reconstructing a 2-color scenery by observing it along a simple random walk path. Eurandom Report 2000-003, 2000.

[21] H. Matzinger and S. W. W. Rolles. Finding blocks and other patterns in a random coloring of $\mathbb{Z}$. Preprint.

[22] H. Matzinger and S. W. W. Rolles. Reconstructing a random scenery observed with random errors along a random walk path. EURANDOM Report 2002-009. Submitted.

Chapter 9

Finding blocks and other patterns in a random coloring of $\mathbb{Z}$

(submitted)

By Heinrich Matzinger, and Silke Rolles

Let $\xi := (\xi_k)_{k\in\mathbb{Z}}$ be i.i.d. with $P(\xi_k = 0) = P(\xi_k = 1) = 1/2$, and let $S := (S_k)_{k\in\mathbb{N}_0}$ be a symmetric random walk with holding on $\mathbb{Z}$, independent of ξ. We consider the scenery ξ observed along the random walk path S, namely the process $(\chi_k := \xi_{S_k})_{k\in\mathbb{N}_0}$. With high probability, we reconstruct the color and the length of block^n, a block in ξ of length $\geq n$ close to the origin, given only the observations $(\chi_k)_{k\in[0,2\cdot 3^{3n}]}$. We find stopping times that stop the random walker with high probability at particular places of the scenery, namely on block^n and in the interval $[-3^n, 3^n]$. Moreover, we reconstruct with high probability a piece of ξ of length of the order $3^{n^{0.2}}$ around block^n, given only $3^{\lfloor n^{0.3}\rfloor}$ observations collected by the random walker starting on the boundary of block^n. [1]

9.1 Introduction

We call a $\{0,1\}$-coloring of the integers a *scenery*. Let $\xi := (\xi_k)_{k\in\mathbb{Z}}$ be i.i.d. uniformly distributed on $\{0,1\}$; so ξ is a random scenery. Let $S := (S_k)_{k\in\mathbb{N}_0}$ be a symmetric random walk with holding on $\mathbb{Z}$, independent of ξ, i.e. we assume that there exist $p, q > 0$ with $2p + q = 1$ such that for all $k \in \mathbb{N}_0$

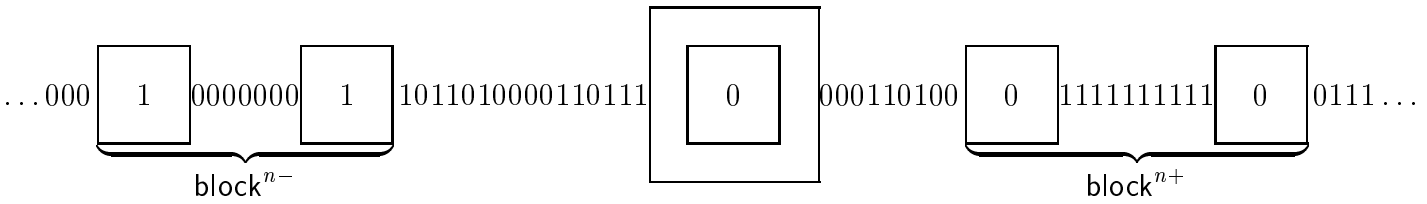
$$\begin{aligned} P(S_{k+1} - S_k = 1) = P(S_{k+1} - S_k = -1) &= p \quad \text{and} \\ P(S_{k+1} - S_k = 0) &= q. \end{aligned}$$

We observe the scenery ξ along the random walk path $(S_k)_{k\in\mathbb{N}_0}$: at time k, we observe $\chi_k := \xi_{S_k}$, the color at location S_k. We denote by $\chi := (\chi_k)_{k\in\mathbb{N}_0}$ the *color record*.

For $k \in \mathbb{N}$, we denote by 0^k and 1^k the element in $\{0,1\}^k$ consisting of k zeros and k ones, respectively. A *block* is a piece of scenery of the form 01^k0 or 10^k1. For $n \in \mathbb{N}$, let block^{n+} designate the leftmost block of ξ of length $\geq n$ to the right of the origin, and let block^{n-} denote the rightmost block of ξ of length $\geq n$ to the left of the origin.

[1] *MSC 2000 subject classification:* Primary 60K37, Secondary 60G10, 60J75.
Key words: Scenery reconstruction, jumps, stationary processes, random walk, ergodic theory.

Furthermore, let $\mathsf{block}^n \in \{\mathsf{block}^{n+}, \mathsf{block}^{n-}\}$ denote the block which is visited first by S. For a block $\xi|[a,b]$, we set $\partial(\xi|[a,b]) := \{a,b\}$. The following picture illustrates our definitions for $n = 7$. The origin is marked with a double box, whereas the points of $\partial\mathsf{block}^{n+}$ and $\partial\mathsf{block}^{n-}$ are marked with one box.



Let $n \in \mathbb{N}$ be large, and suppose we are given $\chi|[0, 2 \cdot 3^{3n}]$ only, i.e. the observations collected by the random walker up to time $2 \cdot 3^{3n}$. We will show that with high probability, we can construct a random time (depending on $\chi|[0, 2 \cdot 3^{3n}]$) when the random walker visits $\partial\mathsf{block}^n$. Our proof is constructive: long blocks in the observations χ turn out to be a good indication that the random walker collects observations on a long block of ξ. Since with high probability, block^n is isolated, up to a certain time horizon, long blocks in the observations are with high probability generated on block^n.

With high probability, we can even reconstruct the color and the length of block^n given only a stopping time T with $S_T \in \partial\mathsf{block}^n$ and $\chi^{n,T} := \chi|[T, T + 3^{n^{\lfloor 0.3 \rfloor}}[$. The idea behind this is the following: If $S_T \in \partial\mathsf{block}^n$, long blocks in $\chi^{n,T}$ are with high probability generated on block^n. The distribution of the length of a block in χ which is generated by the random walk on block^n equals the distribution of the first exit time of the random walk from the interval $[-1, m]$ given the starting point is 0. Explicit calculations with the distribution of the random walk allow us to determine the length of block^n with high probability. The color of block^n agrees with high probability with the color of the first block of $\chi^{n,T}$ of length $\geq n^2$.

Even more is true: It will be shown that there exists a finite sequence of stopping times such that with high probability all of them stop the random walk in $\partial\mathsf{block}^n$. These stopping times are used to define an algorithm that reconstructs with high probability a piece of scenery of length of the order $3^{n^{0.2}}$ around block^n given only $3^{\lfloor n^{0.3} \rfloor}$ observations collected by the random walker starting in $\partial\mathsf{block}^n$. This reconstruction algorithm is used to define a sequence of stopping times that stop the random walk with high probability in the interval $[-3^n, 3^n]$. Our approach is completely constructive: the method described below could be performed by a computer program.

The results of this paper are used in [22] to solve the scenery reconstruction problem for i.i.d. 2-color sceneries and simple random walk with holding using only polynomially many observations. The *scenery reconstruction problem* addresses the question whether one can reconstruct the scenery ξ if one is only given the color record χ (without any additional information about the random walk path $(S_k)_{k \in \mathbb{N}_0}$). More precisely, the question is the following: We define two sceneries ξ and ξ' to be *equivalent* and write $\xi \approx \xi'$ if ξ is obtained from ξ' by a reflection and/or translation. Does there exist a map $\mathcal{A} : \{0,1\}^{\mathbb{N}_0} \to \{0,1\}^{\mathbb{Z}}$, measurable with respect to the σ-algebras generated by the canonical projections, such that $\mathcal{A}(\chi) \approx \xi$ for almost all pairs (ξ, S)? In [22], this question is answered in the affirmative. Furthermore, it is proved that in order to reconstruct with probability $\geq 1 - e^{-cn^{0.2}}$ a piece of the scenery of length of the order 3^n around the origin, we need only the first $2 \cdot 3^{10\alpha n}$ observations collected by the random walker; here $c, \alpha > 0$ are constants.

The scenery reconstruction problem and related questions have attracted considerable attention during the past decade. Among others, the following people contributed to the area: Benjamini and Kesten [1], Burdzy [2], Heicklen, Hoffman, and Rudolph [7], den Hollander [4], den Hollander and Steif [3], Howard ([8], [9], [10]), Keane and den Hollander [11], Kesten ([12], [13]), Levin, Pemantle, and Peres [15], Levin and Peres [14], Lindenstrauss [16], Löwe and Matzinger ([18], [17]), Löwe, Matzinger, and Merkl [19], Matzinger ([20], [21]), Matzinger and Rolles [23].

The remainder of this article is organized as follows: Section 9.2 collects frequently used notations. The results of the article are explained in more detail in Section 9.3. In Section 9.4, we analyze the distribution of the first exit time of the random walker from a finite interval. These results are needed to give a good estimate of the length of block^n. Section 9.5 contains the construction of a stopping time that stops the random walker with high probability in the set $\partial\mathsf{block}^n$. In Section 9.6 we show that there is a whole sequence of stopping times with this property. In Section 9.7, we define an algorithm which reconstructs with high probability a piece of the scenery ξ around block^n. This algorithm is used in Section 9.8 to construct a sequence of stopping times that stop the random walker with high probability in the interval $[-3^n, 3^n]$.

9.2 Frequently used notation

Numbers, sets and functions: We denote by $\mathbb{N} := \{1, 2, 3, \ldots\}$ the set of natural numbers and set $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. If $x \in \mathbb{R}$, we denote by $\lfloor x \rfloor$ the largest integer $\leq x$. We write $x \wedge y$ for the minimum of $x, y \in \mathbb{R}$. For a vector $y = (y_k)_{k\in[1,m]} \in \mathbb{R}^m$ we define the l^1-norm $\|y\|_1 := \sum_{k=1}^m |y_k|$ and the l^2-norm $\|y\|_2 := (\sum_{k=1}^m [y_k]^2)^{1/2}$. The cardinality of a set D is denoted by $|D|$. We write $f|D$ for the restriction of a function f to a set D. An *integer interval* is a set of the form $I \cap \mathbb{Z}$ with an interval $I \subseteq \mathbb{R}$. In this article, intervals are always taken over the integers, e.g. $[a, b] = \{n \in \mathbb{Z} : a \leq n \leq b\}$.

Admissible paths: Let $I = [i_1, i_2]$ be an integer interval. We call $R \in \mathbb{Z}^I$ an *admissible piece of path* if $R_{i+1} - R_i \in \{-1, 0, 1\}$ for all $i \in [i_1, i_2 - 1]$. We call R_{i_1} the starting point, R_{i_2} the endpoint, and $|I|$ the length of R.

Measures: We define δ_x to be the Dirac measure in x. We denote the image of a measure P under a mapping F by PF^{-1}.

Sceneries: We set $\mathcal{C} := \{0, 1\}$. A *scenery* is an element of $\mathcal{C}^{\mathbb{Z}}$, *a piece of scenery* or a *word* is an element of $\mathcal{C}^I$ with an integer interval $I \subseteq \mathbb{Z}$. If $\psi \in \mathcal{C}^I$, we call $|\psi| := |I|$ the *length* of ψ. We write $(1)_I$ for the piece of scenery in $\mathcal{C}^I$ which is identically equal to 1.

Blocks: Let $a, b \in I$ with $a < b$ and $|a - b| \geq 2$. We define $\psi \in \mathcal{C}^{[a,b]}$ to be a *block* if $\psi_a = \psi_b$ and $\psi_c \neq \psi_a$ for all $c \in]a, b[$. ψ_c is the *color* of the block. We call a the *left endpoint*, b the *right endpoint*, and $|\psi| := b - a - 1$ the *blocklength* of ψ. For instance, 01110 is a block of length 3. We set $\partial\psi := \{a, b\}$.

Let $\chi|[t_1, t_2]$ and $\xi|[a, b]$ be blocks. We say that $\chi|[t_1, t_2]$ *is generated by the random walk S on the block $\xi|[a, b]$* if $\{S_{t_1}, S_{t_2}\} \subseteq \{a, b\}$ and $S_t \in]a, b[$ for all $t \in]t_1, t_2[$.

Equivalence of sceneries: Let $\psi \in \mathcal{C}^I$ and $\psi' \in \mathcal{C}^{I'}$ be two pieces of sceneries. We say that ψ and ψ' are *equivalent* and write $\psi \approx \psi'$ iff I and I' have the same length and there exists $a \in \mathbb{Z}$ and $b \in \{-1, 1\}$ such that for all $k \in I$ we have that $a + bk \in I'$ and $\psi_k = \psi'_{a+bk}$. We call ψ and ψ' *strongly equivalent* and write $\psi \equiv \psi'$ if $I' = a + I$ for some $a \in \mathbb{Z}$ and $\psi_k = \psi'_{a+k}$ for all $k \in I$. We say ψ *occurs in* ψ' and write $\psi \sqsubseteq \psi'$ if $\psi \equiv \psi'|J$

for some $J \subseteq I'$. We write $\psi \preceq \psi'$ if $\psi \approx \psi'|J$ for some $J \subseteq I'$.

Random walks and random sceneries: Let $\Omega_2 \subseteq \mathbb{Z}^{\mathbb{N}_0}$ denote the set of admissible paths. Let $p, q > 0$ satisfy $2p + q = 1$. We denote by Q_x the distribution on $(\Omega_2)^{\mathbb{N}_0}$ of a random walk $(S_k)_{k\in\mathbb{N}_0}$ starting at x with i.i.d. increments distributed according to $p\delta_{-1} + q\delta_0 + p\delta_1$, i.e. S is a symmetric random walk with holding. The scenery $\xi := (\xi_k)_{k\in\mathbb{Z}}$ is i.i.d. with ξ_k uniformly distributed on $\mathcal{C} = \{0, 1\}$. We assume that ξ and S are independent and realized as canonical projections on $\Omega := (\mathcal{C}^{\mathbb{Z}}, \Omega_2)$ with the product σ-algebra generated by the canonical projections and probability measures $P_x := (\frac{1}{2}\delta_0 + \frac{1}{2}\delta_1)^{\otimes\mathbb{Z}} \otimes Q_x$, $x \in \mathbb{Z}$. We abbreviate $P := P_0$. We denote the expectation with respect to P by E. We call $\chi := (\chi_k := \xi(S_k))_{k\in\mathbb{N}_0}$ the *scenery observed along the random walk path*. Sometimes, we write $\xi \circ S$ instead of χ.

For a fixed scenery $\xi \in \mathcal{C}^{\mathbb{Z}}$ we set $P_{x,\xi} := \delta_\xi \otimes Q_x$, $P_\xi := P_{0,\xi}$. Thus $P_{x,\xi}$ is the canonical version of the conditional probability $P_x(\cdot|\xi)$; we never work with a different version.

Filtration and shifts: We define a filtration over Ω: $\mathcal{G} := (\mathcal{G}_n)_{n\in\mathbb{N}_0}$ with $\mathcal{G}_n := \sigma(\chi_k; k \in [0, n])$ is the natural filtration of the observations. We define the shift $\theta : \mathcal{C}^{\mathbb{N}_0} \to \mathcal{C}^{\mathbb{N}_0}$, $\eta \mapsto \eta(\cdot + 1)$. We introduce the shift $\Theta : \Omega \to \Omega$, $(\xi, S) \mapsto (\xi(S_1 + \cdot), S(1 + \cdot) - S_1)$. For a set $A \subseteq \Omega$ and a random time $T \geq 0$ we set $\Theta^{-T}(A) := \{\omega : \Theta^{T(\omega)}(\omega) \in A\}$.

Constants are denoted by c_i, $i \geq 1$. They keep their meaning throughout the whole article. Constants c_1, c_7, and α play a special role. They are chosen as follows:

- $c_1 \in 4\mathbb{N}$ with $c_1 > 153$,
- $c_7 > \max\{0, 10\ln 3 + 2c_1(\ln(3/p) + \ln[\max_{i\in[1,5]} \|x_i^*\|_2])\}$ with x_i^* as in Definition 9.7.5,
- $\alpha \in \mathbb{N}$ with $\alpha > 1 + 17c_1 + [24c_7 - 3c_1 \ln p]/\ln 3$.

9.3 Results

Recall the definition of block^n from the introduction. Our first aim is to find a random time which stops the random walk S with high probability on $\partial\mathsf{block}^n$. With high probability, in a large neighborhood of block^n there is no long block in the scenery. Hence, up to a certain time horizon, long blocks in the observations χ indicate that the random walker generates the observations on block^n.

In order to make this precise, we need the following fact: Suppose the random walker generates a block B in the observations χ while walking on a block of ξ of length m. The length of the block B is random and has the same distribution as T_m, the first hitting time of the set $\{-1, m\}$ by the random walk S starting at 0.

It turns out that given the length of an observed block is $\geq (8/p)n^2 \ln n$, the expected length of a block observed on blocks of ξ of length $< n$ on one hand and $\geq n$ on the other hand can be distinguished by a threshold d_n. This statement is made precise by Part 1 of the following lemma. Part 2 shows that the conditional expectations of T_m given $T_m \geq (8/p)n^2 \ln n$ and S_{T_m} are significantly different for m and $m+1$ for $m \in [\lfloor n^{0.4} \rfloor, 2n[$. This will allow us to determine the precise length of block^n with high probability.

Lemma 9.3.1. *Let $i_n := (8/p)n^2 \ln n$. There exists a constant $c_3 > 0$ such that for all $n \geq c_3$ the following holds:*

1. *There exists $d_n > i_n$ such that*

$$E(T_m|T_m \geq i_n, S_{T_m}) \leq d_n - 1 \qquad \text{for all } m \in [1, n[\text{ and}$$
$$E(T_m|T_m \geq i_n, S_{T_m}) \geq d_n + 1 \qquad \text{for all } m \in [n, 2n].$$

2. *For all $m \in [\lfloor n^{0.4} \rfloor, 2n[$ there exist d_m^n such that $m \mapsto d_m^n$ is strictly increasing and*

$$\begin{aligned} E(T_m|T_m \geq i_n, S_{T_m}) &\leq d_m^n - 1 \text{ and} \\ E(T_{m+1}|T_{m+1} \geq i_n, S_{T_{m+1}}) &\geq d_m^n + 1. \end{aligned}$$

Lemma 9.3.1 is used to introduce a stopping time $\nu^n(0)$ which is supposed to stop the random walk with high probability in $\partial\mathsf{block}^n$. We have no chance to detect the first time when $\partial\mathsf{block}^n$ is visited. Typically, at time $\nu^n(0)$, the random walker has visited $\partial\mathsf{block}^n$ already many times.

Let i_n and d_n be as in Lemma 9.3.1. We denote by $\beta_l^n(k)$ and $\beta_r^n(k)$ the left and right endpoint of the kth block of $\chi|[0, 3^{10\alpha n}[$ of length $\geq i_n$, respectively. If there exists no such block, we set $\beta_l^n(k) = \beta_r^n(k) = 3^{10\alpha n}$. We denote by $\nu^n(0)$ the smallest element of the set $\{\beta_r^n(k); k \geq 3^{\lfloor n^{0.2} \rfloor}\}$ with the property that

$$3^{-\lfloor n^{0.2} \rfloor} \sum_{i \in]k - 3^{\lfloor n^{0.2} \rfloor}, k]} [\beta_r^n(i) - \beta_l^n(i) - 1] \geq d_n. \tag{9.3.1}$$

If there exists no such element, we set $\nu^n(0) := \infty$. Note that $\nu^n(0)$ is a $\mathcal{G}$-adapted stopping time.

In other words, we look at the empirical mean of the length of $3^{\lfloor n^{0.2} \rfloor}$ successive blocks of length $\geq i_n$ in the observations. If the empirical mean is $\geq d_n$, then $\nu^n(0)$ is the right end of a block of length $\geq i_n$. Roughly speaking, the idea behind this construction is the following: Typically, there is no block of length $\geq n^{0.4}$ in a neighborhood of block^n. Blocks of length $\geq i_n$ in the observations are with high probability generated on blocks of length $\geq n^{0.4}$, and the first time the random walker visits a block of length $\geq n^{0.4}$, she visits block^n. Moreover, for n large, the empirical mean is close to the conditional expectation $E(T_m|T_m \geq i_n)$. If the empirical mean is $\geq d_n$, then Lemma 9.3.1 suggests that $m \geq n$. This way, we have identified a time when the random walker visits $\partial\mathsf{block}^n$. The following theorem states that indeed with high probability $\nu^n(0)$ stops the random walk in the set $\partial\mathsf{block}^n$.

Theorem 9.3.1. *We define*

$$E^n_{\nu(0)\text{ ok}} := \{S_{\nu^n(0)} \in \partial\mathsf{block}^n\} \cap \{\nu^n(0) \leq 2 \cdot 3^{3n}\} \cap \{\partial\mathsf{block}^n \subseteq [-3^n/3, 3^n/3]\}.$$

There exist constants c_{11}, c_{12}, c_{13} such that for all $n \geq c_{11}$

$$P\left(\left[E^n_{\nu(0)\text{ ok}}\right]^c\right) \leq c_{12} e^{-c_{13} n^{0.3}}.$$

Next, we estimate the color and the length of block^n. We start with an abstract definition: Let $\eta \in \mathcal{C}^{[0, 3^{\lfloor n^{0.3} \rfloor}[}$. Let $\mathsf{color}^n(\eta)$ be the color of the first block of length $\geq n^2$ in η. If there exists no such block, we set $\mathsf{color}^n(\eta) := 1$. Let i_n and d_m^n be as in Lemma 9.3.1. We denote by $\bar{\beta}_l^n(k, \eta)$ and $\bar{\beta}_r^n(k, \eta)$ the left and right endpoint of the kth block of η of

length $\geq i_n$, respectively. If there exists no such block, we set $\bar{\beta}_l^n(k,\eta) = \bar{\beta}_r^n(k,\eta) = 2\cdot 3^{3n}$. Let $\mathsf{length}^n(\eta)$ be the smallest $m \geq n$ such that

$$d_{m+1}^n > 3^{-\lfloor n^{0.2}\rfloor} \sum_{i\in[1,3^{\lfloor n^{0.2}\rfloor}]} \left[\bar{\beta}_r^n(i,\eta) - \bar{\beta}_l^n(i,\eta) - 1\right] \geq d_m^n. \tag{9.3.2}$$

If there exists no m with this property, we set $\mathsf{length}^n(\eta) := 2n+1$. Let $\mathsf{estimated\text{-}block}^n(\eta) \in \cup_{k\geq n+2}\mathcal{C}^k$ be the block of length $\mathsf{length}^n(\eta)$ and of color $\mathsf{color}^n(\eta)$.

Let b_l^n and b_r^n denote the left and right endpoint of block^n, respectively. Then $\partial\mathsf{block}^n = \{b_l^n, b_r^n\}$. For $i \in \{l, r\}$ we define

$$H_i^n := \min\{k \geq 0 : S_k = b_i^n\}. \tag{9.3.3}$$

Proposition 9.3.1. *There exist constants $c_7, c_8, c_9 > 0$ such that for all $n \geq c_7$ and all $T \in \{H_l^n, H_r^n\}$ the following holds:*

$$P(\mathsf{estimated\text{-}block}^n(\chi|[T, T+3^{\lfloor n^{0.3}\rfloor}[) \neq \mathsf{block}^n) \leq c_8 e^{-c_9 n^{0.3}}.$$

It turns out that, once we have a stopping time T with $S_T \in \partial\mathsf{block}^n$, we obtain a whole (finite) sequence of stopping times which stop the random walk in the set $\partial\mathsf{block}^n$: we just look at long blocks in the observations. For $k \geq 1$, let $\tilde{\nu}_b^{n,T}(k)$ denote the right end of the kth block of $\chi[T, T+3^{10\alpha\lfloor n^{0.2}\rfloor}[$ of length $\geq n^2/16$. If there exists no such block, we define $\tilde{\nu}_b^{n,T}(k) := T + 3^{10\alpha\lfloor n^{0.2}\rfloor}$. We define $\nu_b^{n,T} := (\nu_b^{n,T}(k))_{k\in[1,3^{\alpha\lfloor n^{0.2}\rfloor}]}$ by

$$\nu_b^{n,T}(k) := \tilde{\nu}_b^{n,T}(2k\cdot 3^{3\lfloor n^{0.2}\rfloor}).$$

Note that for all k, $\nu_b^{n,T}(k)$ is a $\mathcal{G}$-adapted stopping time. The following proposition states that with high probability, the $\nu_b^{n,T}(k)$'s stop the random walk in the set $\partial\mathsf{block}^n$. It will be essential below that the $\nu_b^{n,T}(k)$'s are sufficiently far apart from each other and that we have sufficiently many of them.

Proposition 9.3.2. *There exist constants $c_{10}, c_{11} > 0$ such that for all $n \geq c_{10}$ and all $T \in \{H_l^n, H_r^n\}$ the following holds: The event*

$$E_{\nu_b\ \mathrm{ok}}^{n,T} := \bigcap_{k=1}^{3^{\alpha\lfloor n^{0.2}\rfloor}} \left\{ \begin{array}{l} \nu_b^{n,T}(k) \in [T, T+3^{10\alpha\lfloor n^{0.2}\rfloor}[,\ S_{\nu_b^{n,T}(k)} \in \partial\mathsf{block}^n, \\ \nu_b^{n,T}(j) + 2\cdot 3^{3\lfloor n^{0.2}\rfloor} \leq \nu_b^{n,T}(k) \text{ for all } j < k \end{array} \right\}.$$

satisfies $P\left(\left[E_{\nu_b\ \mathrm{ok}}^{n,T}\right]^c\right) \leq e^{-c_{11}n^{0.2}}$.

Given a stopping time T that stops the random walk in $\partial\mathsf{block}^n$, we can define a map $\mathrm{SmallAlg}^n$ which reconstructs with high probability a piece of scenery of length of the order $3^{\lfloor n^{0.2}\rfloor}$ around block^n, given only the observations between times T and $T + 3^{\lfloor n^{0.3}\rfloor}$. The probability that the reconstruction fails is small *conditioned* on the scenery ξ, at least for ξ in a set of large probability. The following theorem makes this precise.

Theorem 9.3.2. *There exist constants $c_{14}, c_{18} > 0$ and a sequence*

$$\mathrm{SmallAlg}^n : \mathcal{C}^{[0,3^{\lfloor n^{0.3}\rfloor}[} \to \mathcal{C}^{[-3\cdot 3^{\lfloor n^{0.2}\rfloor}, 3\cdot 3^{\lfloor n^{0.2}\rfloor}]}, n \geq c_{14},$$

of measurable maps such that the following holds: If we define

$$\begin{aligned} E^n_{\text{recon Small}} &:= \left\{\text{SmallAlg}^n\big(\chi|\big[0,3^{\lfloor n^{0.3}\rfloor}\big[\big) \preceq \xi|[-3\cdot 3^{\lfloor n^{0.2}\rfloor}, 3\cdot 3^{\lfloor n^{0.2}\rfloor}]\right\} \text{ and} \\ \Xi^n &:= \left\{\xi\in\mathcal{C}^{\mathbb{Z}} : P_\xi\big(\big[\Theta^{-T}E^n_{\text{recon Small}}\big]^c\big) \le e^{-c_{18}n^{0.2}} \text{ for all } T\in\{H^n_l, H^n_r\}\right\}, \end{aligned}$$

then $P(\xi\notin\Xi^n)\le e^{-c_{18}n^{0.2}}$ *for all* $n\ge c_{14}$.

We use SmallAlg^n to define a sequence of stopping times which stop the random walker with high probability in the interval $[-3^n, 3^n]$. We define

$$\psi_n := \text{SmallAlg}^n(\chi|[\nu^n(0), \nu^n(0)+3^{\lfloor n^{0.3}\rfloor}[), \tag{9.3.4}$$

$$\mathbb{T}^n := \left\{\begin{array}{l} t\in[\nu^n(0), 3^{10\alpha n}-3^{\lfloor n^{0.3}\rfloor}[\colon \exists w\in\mathcal{C}^{[-3^{\lfloor n^{0.2}\rfloor}, 3^{\lfloor n^{0.2}\rfloor}]} \text{ such that } w\preceq\psi_n \\ \text{and } w\preceq \text{SmallAlg}^n(\chi|[t,t+3^{\lfloor n^{0.3}\rfloor}[) \end{array}\right\} \tag{9.3.5}$$

Let $\tilde\nu^n(1)<\tilde\nu^n(2)<\dots$ denote the points in $\mathbb{T}^n$ in increasing order. We define $\nu^n := (\nu^n(k))_{k\in[1,3^{\alpha n}]}$ by

$$\nu^n(k) := \begin{cases} \tilde\nu^n(2\cdot 3^{3n}k)+3^{\lfloor n^{0.3}\rfloor} & \text{if } 2\cdot 3^{3n}k\le|\mathbb{T}^n|, \\ 3^{10\alpha n} & \text{else.}\end{cases}$$

Note that $\nu^n(k)$ is a $\mathcal{G}$-adapted stopping time: in order to determine whether $t\in\mathbb{T}_n$, we need to look at $\chi|[t,t+3^{\lfloor n^{0.3}\rfloor}[$, but $\nu^n(k)$ is never defined to be t, but only $t+3^{\lfloor n^{0.3}\rfloor}$.

The idea behind the definition of the $\nu^n(k)$'s is the following: With high probability, $\nu^n(0)$ stops the random walk in the set $\partial\mathsf{block}^n$, ψ_n is a piece of scenery of length $6\cdot 3^{\lfloor n^{0.2}\rfloor}+1$ containing block^n, and ψ_n is up to a possible reflection contained in $\xi|[-3^n,3^n]$. (With high probability, block^n has length $\le 2n$ and can be found in the piece of scenery $\xi|[-3^n/3, 3^n/3]$.) The set $\mathbb{T}_n$ consists of times $t\ge\nu^n(0)$ such that SmallAlg^n applied to the observations starting at time t produces an output which agrees on a large subpiece, namely a piece of length $2\cdot 3^{\lfloor n^{0.2}\rfloor}+1$, with ψ_n. With high probability, ψ_n is typical for the scenery $\xi|[-3^n,3^n]$, and hence the random walker is in the interval $[-3^n,3^n]$ at time t. The $\nu^n(k)$'s are defined in such a way that they are in some sense far apart from each other and all bounded by $3^{10\alpha n}$. This is needed for our application in [22]. Formally, the task of the stopping times $\nu^n(k)$ is specified by the event $E^{n,\nu}_{\text{stop}}$ defined as follows:

Definition 9.3.1. *For* $n\in\mathbb{N}$, *we define the event*

$$E^{n,\nu}_{\text{stop}} := \bigcap_{k=1}^{3^{\alpha n}} \left\{\nu^n(k)<3^{10\alpha n}, |S_{\nu^n(k)}|\le 3^n, \nu^n(j)+2\cdot 3^{3n}\le\nu^n(k) \text{ for all } j<k\right\}.$$

Recall the definition of Ξ^n from Theorem 9.3.2. If the event $E^n_{\nu(0)\text{ ok}}\cap\Theta^{-\nu^n(0)}[E^n_{\text{recon Small}}]$ holds and $\xi\in\Xi_n$, then with high probability the event $E^{n,\nu}_{\text{stop}}$ holds as well:

Proposition 9.3.3. *There exist constants* c_{19}, c_{20}, c_{21} *such that for all* $n\ge c_{19}$

$$P\Big(\big[E^n_{\nu(0)\text{ ok}}\cap\Theta^{-\nu^n(0)}[E^n_{\text{recon Small}}]\cap\{\xi\in\Xi^n\}\big]\setminus E^{n,\nu}_{\text{stop}}\Big)\le c_{20}e^{-c_{21}n^{0.3}}.$$

9.4 Distinguishing blocks of different lengths

Recall the definition of T_m, $m \in \mathbb{N}$: $T_m := \min\{k \geq 0 : S_k \in \{-1, m\}\}$, i.e. T_m denotes the first hitting time of the set $\{-1, m\}$ by the random walk. In this section, we prove Lemma 9.3.1 and collect some properties of T_m that are needed for the constructions below.

Proof of Lemma 9.3.1. The distribution of T_m is well known, see e.g. [6] for the following identities which hold for all $k, m \in \mathbb{N}$ with $m \geq 2$:

$$P[T_{m-1} = k, S_{T_{m-1}} = -1] = \frac{2p}{m} \sum_{j=1}^{m-1} \sin^2\left[\frac{\pi j}{m}\right] \left[q + 2p\cos\left[\frac{\pi j}{m}\right]\right]^{k-1}, \tag{9.4.1}$$

$$P[T_{m-1} = k, S_{T_{m-1}} = m-1] = \frac{2p}{m} \sum_{j=1}^{m-1} \sin\left[\frac{\pi j(m-1)}{m}\right] \sin\left[\frac{\pi j}{m}\right] \left[q + 2p\cos\left[\frac{\pi j}{m}\right]\right]^{k-1} \tag{9.4.2}$$

It turns out that the main contribution in (9.4.1) and (9.4.2) comes from the sumand for $j = 1$. We write

$$P(T_{m-1} = k, S_{T_{m-1}} = -1) = a_m[\rho_m]^{k-1} + \varepsilon_{k,m} \tag{9.4.3}$$

$$\text{with } a_m := \frac{2p}{m}\sin^2\left[\frac{\pi}{m}\right], \quad \rho_m := q + 2p\cos\left[\frac{\pi}{m}\right] \tag{9.4.4}$$

and error terms $\varepsilon_{k,2} = 0$,

$$\varepsilon_{k,m} := \frac{2p}{m} \sum_{j=2}^{m-1} \sin^2\left[\frac{\pi j}{m}\right] \left[q + 2p\cos\left[\frac{\pi j}{m}\right]\right]^{k-1} \text{ for } m \geq 3.$$

We will show that the $\varepsilon_{k,m}$'s are small so that the distribution of T_{m-1} conditioned on $S_{T_{m-1}} = -1$ is approximately geometrical with parameter ρ_m.

We calculate

$$\rho_1^{(m)} := \sum_{k=i_n}^{\infty} a_m[\rho_m]^{k-1} = \frac{a_m[\rho_m]^{i_n-1}}{1-\rho_m}, \tag{9.4.5}$$

$$\rho_2^{(m)} := \sum_{k=i_n}^{\infty} a_m k[\rho_m]^{k-1} = a_m[\rho_m]^{i_n-1} \sum_{k=1}^{\infty}(k + i_n - 1)[\rho_m]^{k-1} \tag{9.4.6}$$

$$= a_m[\rho_m]^{i_n-1}\left[\frac{i_n}{1-\rho_m} + \frac{\rho_m}{(1-\rho_m)^2}\right], \tag{9.4.7}$$

$$\frac{\rho_2^{(m)}}{\rho_1^{(m)}} = i_n + \frac{\rho_m}{1-\rho_m} \leq i_n + c_{17}m^2 \tag{9.4.8}$$

with a constant $c_{17} > 0$; recall the definition of ρ_m (9.4.4). In order to obtain an upper bound for $\varepsilon_{k,m}$, we will need the following estimate, valid for all $m \geq 3$:

$$\frac{q + 2p\cos\left[\frac{2\pi}{m}\right]}{q + 2p\cos\left[\frac{\pi}{m}\right]} \leq 1 - \frac{8p}{m^2} =: r_m. \tag{9.4.9}$$

In order to prove (9.4.9) we set $y := \frac{\pi}{m}$ and use $\cos y \in [0,1]$, $q+2p=1$, and $\cos(2y) = 2\cos^2 y - 1$ to obtain

$$\begin{aligned} 1 - \frac{q+2p\cos\left[\frac{2\pi}{m}\right]}{q+2p\cos\left[\frac{\pi}{m}\right]} &= 2p\frac{\cos y - \cos(2y)}{q+2p\cos y} \geq 2p[\cos y - \cos(2y)] \\ &= 2p[\cos y - 2\cos^2 y + 1] \geq 2p[1-\cos^2 y] = 2p\sin^2 y \geq \frac{8py^2}{\pi^2}; \end{aligned}$$

for the last inequality we used the estimate $\sin y \geq \frac{2y}{\pi}$ for all $y \in [0, \pi/2]$.

Using (9.4.9), we can bound the error term $\varepsilon_{k,m}$ for all k, m:

$$|\varepsilon_{k,m}| \leq [\rho_m r_m]^{k-1}. \tag{9.4.10}$$

We abbreviate $\varepsilon_1^{(m)} := \sum_{k=i_n}^{\infty} \varepsilon_{k,m}$, $\varepsilon_2^{(m)} := \sum_{k=i_n}^{\infty} k\varepsilon_{k,m}$ and note that by (9.4.10)

$$|\varepsilon_1^{(m)}|, |\varepsilon_2^{(m)}| \leq \frac{[\rho_m r_m]^{i_n-1}(i_n+1)}{(1-\rho_m r_m)^2} \tag{9.4.11}$$

for all $m \geq 3$ (compare (9.4.5) and (9.4.7) with ρ_m replaced by $r_m\rho_m$). In the following, we assume $m \in [3, 2n]$. Then

$$r_m^{i_n-1} = \left(1 - \frac{8p}{m^2}\right)^{(8/n)n^2\ln n - 1} \leq \left(1 - \frac{2p}{n^2}\right)^{(8/p)n^2\ln n - 1} \leq c_{18} n^{-16} \tag{9.4.12}$$

with a constant $c_{18} > 0$. Since $r_m \in]0,1[$, it follows from (9.4.11) that

$$|\varepsilon_1^{(m)}|, |\varepsilon_2^{(m)}| \leq c_{18} n^{-16} \cdot \frac{[\rho_m]^{i_n-1}(i_n+1)}{(1-\rho_m)^2}. \tag{9.4.13}$$

For $|\varepsilon_1^{(m)}|$ we have the following sharper estimate

$$|\varepsilon_1^{(m)}| \leq c_{18} n^{-16} \cdot \frac{[\rho_m]^{i_n-1}}{1-\rho_m} \tag{9.4.14}$$

which follows from (9.4.12) and the definition of $|\varepsilon_1^{(m)}|$. Together with (9.4.5), the estimate (9.4.14) implies

$$|\rho_1^{(m)} + \varepsilon_1^{(m)}| \geq \rho_1^{(m)} - |\varepsilon_1^{(m)}| \geq \frac{[\rho_m]^{i_n-1}}{1-\rho_m} \cdot [a_m - c_{18} n^{-16}].$$

Since $a_m \geq 8p/m^3$, we conclude for all n sufficiently large

$$|\rho_1^{(m)} + \varepsilon_1^{(m)}| \geq \frac{a_m[\rho_m]^{i_n-1}}{2(1-\rho_m)} \geq \frac{4p[\rho_m]^{i_n-1}}{m^3(1-\rho_m)}. \tag{9.4.15}$$

Hence, using (9.4.3),

$$E(T_{m-1} | T_{m-1} \geq i_n, S_{T_{m-1}} = -1) = \frac{\rho_2^{(m)} + \varepsilon_2^{(m)}}{\rho_1^{(m)} + \varepsilon_1^{(m)}} = \frac{\rho_2^{(m)}}{\rho_1^{(m)}} + \varepsilon_3^{(m)} \tag{9.4.16}$$

with an error term $\varepsilon_3^{(2)} = 0$ and

$$|\varepsilon_3^{(m)}| \le \left| \frac{\varepsilon_2^{(m)}\rho_1^{(m)} - \varepsilon_1^{(m)}\rho_2^{(m)}}{[\rho_1^{(m)} + \varepsilon_1^{(m)}]\rho_1^{(m)}} \right| \le \frac{\rho_2^{(m)}}{\rho_1^{(m)}} \cdot \frac{|\varepsilon_1^{(m)}| + |\varepsilon_2^{(m)}|}{|\rho_1^{(m)} + \varepsilon_1^{(m)}|} \le \frac{\rho_2^{(m)}}{\rho_1^{(m)}} \cdot \frac{c_{18} n^{-16} m^3}{4p} \cdot \frac{i_n + 1}{1 - \rho_m}$$

for all $m \ge 3$ and for all n sufficiently large; for the last inequality, we used (9.4.13) and (9.4.15). Combining this with (9.4.8) and the definition of i_n, we obtain for all $m \in [3, 2n]$ with a constant c_{19}

$$|\varepsilon_3^{(m)}| \quad \le \quad (i_n + c_{17} m^2) \cdot \frac{c_{18} n^{-16} m^3}{4p} \cdot \frac{i_n + 1}{1 - \rho_m} \le c_{19} n^{-5} \tag{9.4.17}$$

for all n sufficiently large. (9.4.16) and (9.4.8) yield

$$Z_m := E(T_{m-1} | T_{m-1} \ge i_n, S_{T_{m-1}} = -1) = i_n + \frac{\rho_m}{1 - \rho_m} + \varepsilon_3^{(m)}. \tag{9.4.18}$$

By (9.4.17), we have

$$Z_n - Z_{n-1} \ge \frac{\rho_n}{1 - \rho_n} - \frac{\rho_{n-1}}{1 - \rho_{n-1}} - 2c_{19}(n-1)^{-5}.$$

Since $\frac{\rho_m}{1-\rho_m} \sim \frac{pm^2}{\pi^2}$ as $m \to \infty$ in the sense that the quotient of both sides converges to 1 as $m \to \infty$, there exists c_{20} (large) such that for all $n \ge c_{20}$, $Z_n - Z_{n-1} \ge 2$. We choose d_n such that $d_n - 1 \le Z_{n-1}$ and $Z_n \ge d_n + 1$. It follows from (9.4.18) and (9.4.17) that $Z_m \le Z_{n-1}$ for all $m \in [1, n-1]$ and $Z_m \ge Z_n$ for all $m \in [n, 2n]$. This proves part 1 of the claim. Part 2 follows by a similar argument. The estimates for $E(T_m | T_m \ge i_n, S_{T_m} = m)$ are done analogously: the only difference in the formulas (9.4.1) and (9.4.2) are the terms $\sin\left[\frac{\pi j(m-1)}{m}\right]$ instead of $\sin\left[\frac{\pi j}{m}\right]$. All we used about these terms in the above proof was that their absolute value is ≤ 1; a_m remains the same. □

Lemma 9.4.1. *Recall that $i_n = (8/p)n^2 \ln n$, and let c_3 be as in Lemma 9.3.1. There exist constants $c_{19}, c_{28} > 0$ such that for all $n \ge c_3$ the following hold:*

1. *For all $m \in [1, 2n]$, we have $E(T_m - T_m \wedge n^4 | T_m \ge i_n, S_{T_m}) \le 2c_{19} n^{-1}$.*
2. *For all $m \in [1, 2n]$, we have $P(T_m \ge n^4 | T_m \ge i_n, S_{T_m}) \le e^{-n^{1.5}}$.*
3. *$P(T_n \ge i_n) \ge n^{-c_{28}}$.*

Proof. We continue in the notation of the proof of Lemma 9.3.1. Using (9.4.13) and (9.4.15) and the fact that $\frac{|\varepsilon_2^{(m)}|}{|\rho_1^{(m)} + \varepsilon_1^{(m)}|} \le |\varepsilon_3^{(m)}| \le c_{19} n^{-5}$, we obtain

$$\left|E(T_{m-1} - T_{m-1} \wedge n^4 | T_{m-1} \ge i_n, S_{T_{m-1}} = -1)\right| \le \frac{\sum_{k=n^4}^{\infty} a_m k [\rho_m]^{k-1} + |\varepsilon_2^{(m)}|}{|\rho_1^{(m)} + \varepsilon_1^{(m)}|}$$

$$\le \frac{a_m [\rho_m]^{n^4 - 1}(n^4 + 1)}{(1 - \rho_m)^2 |\rho_1^{(m)} + \varepsilon_1^{(m)}|} + c_{19} n^{-5} \le c_{22} [\rho_m]^{n^4 - i_n} n^6 + c_{19} n^{-5}$$

for all n sufficiently large with a constant $c_{22} > 0$. Recall that $\rho_m = q + 2p \cos\left[\frac{\pi}{m}\right]$. Consequently, $c_{22}[\rho_m]^{n^4 - i_n} n^6 \le c_{19} n^{-1}$ for all $m \in [2, 2n]$ and n sufficiently large. This proves Part 1.

Using (9.4.10) and (9.4.15), we get for all n sufficiently large

$$\begin{aligned} P(T_m \geq n^4 | T_m \geq i_n, S_{T_m} = -1) &\leq \Big[\frac{a_m \rho_m^{n^4-1}}{1-\rho_m} + \frac{[\rho_m r_m]^{n^4-1}}{1-\rho_m r_m}\Big] \cdot |\rho_1^{(m)} + \varepsilon_1^{(m)}|^{-1} \\ &\leq 4\rho_m^{n^4-i_n}. \end{aligned}$$

For ρ_m, we obtain the following estimate:

$$\rho_m = q + 2p\cos\big[\frac{\pi}{m}\big] = 1 - 2p\big[1 - \cos\big[\frac{\pi}{m}\big]\big] \leq 1 - \frac{c_{23}}{n^2}$$

for all n sufficiently large with a constant $c_{23} > 0$. Hence, by the definition of i_n, $4\rho_m^{n^4-i_n} \leq e^{-n^{1.5}}$ for all large n, and Part 2 follows.

Finally we prove Part 3. Using (9.4.15) and the estimate $1 - \rho_{n+1} \leq 1$, we obtain

$$\begin{aligned} P(T_n \geq i_n) &\geq P(T_n \geq i_n, S_{T_n} = -1) = \rho_1^{(n+1)} + \varepsilon_1^{(n+1)} \\ &\geq \frac{4p[\rho_{n+1}]^{i_n-1}}{(n+1)^3(1-\rho_{n+1})} \geq c_{24} n^{-3} [\rho_{n+1}]^{i_n-1} \end{aligned}$$

with a constant $c_{24} > 0$ for all n sufficiently large. We estimate

$$\begin{aligned} \rho_{n+1} &= q + 2p\cos\Big[\frac{\pi}{n+1}\Big] \geq q + 2p\cos^2\Big[\frac{\pi}{n+1}\Big] \\ &= 1 - 2p\sin^2\Big[\frac{\pi}{n+1}\Big] \geq 1 - \frac{2p\pi^2}{(n+1)^2}. \end{aligned}$$

Thus,

$$(\rho_{n+1})^{i_n-1} \geq \left(1 - \frac{2p\pi^2}{(n+1)^2}\right)^{(8/p)n^2 \ln n} \geq n^{-1-4/(p^2\pi^2)}$$

for all n sufficiently large, and the claim follows. □

9.5 Finding block^n

In this section, we prove Theorem 9.3.1 and Proposition 9.3.1.

Let $n \in \mathbb{N}$ be fixed. Recall the definition of block^n: Let block$^{n+} := \xi|[b_l^{n+}, b_r^{n+}]$ designate the leftmost block of ξ of length $\geq n$ with $b_l^{n+} \geq 0$, and let block$^{n-} := \xi|[b_l^{n-}, b_r^{n-}]$ denote the rightmost block of ξ of length $\geq n$ with $b_r^{n-} \leq 0$. Finally, let block$^n = \xi|[b_l^n, b_r^n] \in \{\text{block}^{n+}, \text{block}^{n-}\}$ denote the block which is visited first by S.

9.5.1 Definitions of events

We define in alphabetical order events, which will be needed below.

Definition 9.5.1. *We define*

$$B^n_{\text{average}\geq} := \left\{ \begin{array}{l} \textit{The average lengths of the first } 3^{\lfloor n^{0.2} \rfloor} \textit{ blocks of length} \geq i_n \\ \textit{produced by } S \textit{ on block}^n \textit{ is} \geq d_n \end{array} \right\},$$

$$B^n_{\text{average}<} := \bigcap_{k \in [3^{\lfloor n^{0.2} \rfloor}, 2 \cdot 3^{3n}]} \left\{ \begin{array}{l} \textit{If all blocks } \chi|[\beta_l^n(i), \beta_r^n(i)],\ i \in]k - 3^{\lfloor n^{0.2} \rfloor}, k], \\ \textit{are generated on blocks of } \xi \textit{ of length} < n, \textit{ then} \\ 3^{-\lfloor n^{0.2} \rfloor} \sum_{i \in]k-3^{\lfloor n^{0.2} \rfloor}, k]} [\beta_r^n(i) - \beta_l^n(i) - 1] < d_n \end{array} \right\}.$$

Definition 9.5.2. *Let $s, t \geq 0$. We define*

$$\begin{aligned} r_n^s(1) &:= \min\{k \geq s : (S_k, S_{k+1}) \in \{(b_l^n, b_l^n + 1), (b_r^n, b_r^n - 1)\}\}, \\ r_n^s(i+1) &:= \min\{k > r_n^s(i) : (S_k, S_{k+1}) \in \{(b_l^n, b_l^n + 1), (b_r^n, b_r^n - 1)\}\}, i \geq 1, \\ B^{n,s,t}_{\text{block enough}} &:= \{r_n^s(\lfloor t^{3/10} \rfloor) < s + t/2\}. \end{aligned}$$

$r_n^s(i)$ is the ith entrance time of the random walk into block^n. If $B^{n,s,t}_{\text{block enough}}$ holds, $\chi|[s, s + t/2[$ contains at least $\lfloor t^{3/10} \rfloor - 1$ blocks generated on block^n.

Definition 9.5.3. *We define* $B^n_{\text{b small}} := \left\{\partial \mathsf{block}^{n+} \cup \partial \mathsf{block}^{n-} \subseteq [-3^n/3, 3^n/3]\right\}$.

Definition 9.5.4. *Let γ_n denote the first hitting time of the set $\partial\mathsf{block}^n$ by the random walk: $\gamma_n := \min\{k \geq 0 : S_k \in \partial \mathsf{block}^n\}$. We define* $B^n_{\text{enough blocks}} :=$

$$\{S|[\gamma_n, \gamma_n + 3^{\lfloor n^{0.3} \rfloor}[\text{ produces } > 3^{\lfloor n^{0.2} \rfloor} \text{ blocks of length } \geq i_n \text{ on } \mathsf{block}^n\}.$$

Definition 9.5.5. *We define* $B^n_{\text{only block}} := B^{n+}_{\text{only}} \cap B^{n-}_{\text{only}}$ *with*

$$B^{ni}_{\text{only}} := \left\{ \begin{matrix} \text{There exists no block of length} \geq \lfloor n^{0.4} \rfloor \text{ with a distance} \\ \leq 3^{\lfloor n^{0.3} \rfloor} \text{ from } \partial \mathsf{block}^{ni} \end{matrix} \right\}, i \in \{+, -\}.$$

Definition 9.5.6. *We define* $B^n_{\text{rw hits}} := \{\exists k \in [0, 3^{3n}[\text{ such that } S_k \in \{-3^n/3, 3^n/3\}\}$.

Definition 9.5.7. *We define*

$$B^n_{\text{short block}} := \left\{ \begin{matrix} \text{There exists no block of length} \geq n^2 \text{ in } \chi|[0, 3 \cdot 3^{3n}[\text{ which is} \\ \text{generated by } S \text{ on a block of } \xi \text{ of length} \leq \lfloor n^{0.4} \rfloor \end{matrix} \right\}.$$

Definition 9.5.8. *We define* $B^n_{\text{size block}} := \{|b_l^{ni} - b_r^{ni}| < 2n \text{ for } i \in \{+, -\}\}$.

9.5.2 Proof of Theorem 9.3.1

Recall the definition of the event $E^n_{\nu(0)\text{ ok}}$ from Theorem 9.3.1.

Lemma 9.5.1. *For all $n \geq 2$ the following inclusion holds P-almost surely:*

$$\begin{aligned} B^n_{\text{average}\geq} \cap B^n_{\text{average}<} \cap B^n_{\text{b small}} \cap B^n_{\text{enough blocks}} \cap B^n_{\text{only block}} & \\ \cap B^n_{\text{rw hits}} \cap B^n_{\text{short block}} \quad &\subseteq \quad E^n_{\nu(0)\text{ ok}}. \end{aligned}$$

Proof. Let $n \in \mathbb{N}$ and suppose $B^n_{\text{average}\geq}$, $B^n_{\text{average}<}$, $B^n_{\text{b small}}$, $B^n_{\text{enough blocks}}$, $B^n_{\text{only block}}$, $B^n_{\text{rw hits}}$, and $B^n_{\text{short block}}$ hold. Suppose further $S_0 = 0$. (This holds P-almost surely.)

Since S is a nearest-neighbor random walk and $B^n_{\text{b small}}$ and $B^n_{\text{rw hits}}$ hold, there exists $k \in [0, 3^{3n}[$ such that $S_k \in \partial\mathsf{block}^n$. Thus $\gamma_n \leq 3^{3n}$. Using that $B^n_{\text{enough blocks}}$ and $B^n_{\text{average}\geq}$ hold, we see that $\nu^n(0) \leq \gamma_n + 3^{\lfloor n^{0.3} \rfloor} \leq 2 \cdot 3^{3n}$.

Since $B^n_{\text{average}<}$ holds, we know that $\nu^n(0)$ equals $\beta_r^n(k)$ for some $k \geq 3^{\lfloor n^{0.2} \rfloor}$ with the property that at least one of the blocks $\chi|[\beta_l^n(i), \beta_r^n(i)]$, $i \in]k - 3^{\lfloor n^{0.2} \rfloor}, k]$ is generated on a block of ξ of length $\geq n$. Recall that $i_n = (8/p)n^2 \ln n \geq n^2$ for all $n \geq 2$. Using $\gamma_n + 3^{\lfloor n^{0.3} \rfloor} \leq 2 \cdot 3^{3n}$ and $B^n_{\text{short block}}$, we conclude that all blocks of length $\geq i_n$ generated by $S|[\gamma_n, \gamma_n + 3^{\lfloor n^{0.3} \rfloor}[$ are generated on blocks of ξ of length $> \lfloor n^{0.4} \rfloor$. The only block of ξ qualifying for this is block^n because $B^n_{\text{only block}}$ holds. Hence by the definition of $\nu^n(0)$, it follows $S_{\nu^n(0)} \in \partial\mathsf{block}^n$. □

Proof of Theorem 9.3.1. By Lemma 9.5.1, we have for all $n \geq 2$

$$\begin{aligned} P\left(\left[E^n_{\nu(0)\text{ ok}}\right]^c\right) \;\leq\; & P([B^n_{\text{average}\geq}]^c) + P([B^n_{\text{average}<}]^c) + P([B^n_{\text{b small}}]^c) + P([B^n_{\text{enough blocks}}]^c) \\ & + P([B^n_{\text{only block}}]^c) + P([B^n_{\text{rw hits}}]^c) + P([B^n_{\text{short block}}]^c). \end{aligned}$$

The claim follows from Lemmas 9.5.3, 9.5.4, 9.5.6, 9.5.7, 9.5.8, 9.5.9, and 9.5.10 below. □

9.5.3 Probabilistic estimates

In this subsection, we prove that the complements of the events defined in Subsection 9.5.1 have a probability which is exponentially small in a power of n. Before we treat the events in alphabetical order, we prove a lemma which will be needed below.

Lemma 9.5.2. *There exist constants $c_{25}, c_{26} > 0$ such that for all $t \in \mathbb{N}$ and for any bounded integer interval I the following holds:*

$$P(S_i \in I \text{ for all } i \in [0,t[) \leq c_{25} \exp\Big(-c_{26} \cdot \frac{t}{|I|^2}\Big).$$

Proof. We abbreviate $l := \lfloor t/|I|^2 \rfloor$. By the Markov property of the random walk,

$$\begin{aligned} P(S_i \in I \text{ for all } i \in [0,t[) \;&\leq\; P\Big(\bigcap_{k\in[0,l[} \{\forall i \in [k|I|^2, (k+1)|I|^2[\colon S_i \in I\}\Big) \\ &\leq\; P\Big(\bigcap_{k\in[0,l[} \{\forall i \in [k|I|^2, (k+1)|I|^2[\colon |S_i - S_{k|I|^2}| \leq |I|\}\Big) \\ &\leq\; P(\forall i \in [0,|I|^2[\colon |S_i| \leq |I|)^l \\ &\leq\; P(|S_{|I|^2-1}| \leq |I|)^l. \end{aligned}$$

By the central limit theorem, there exists $c_{27} \in]0,1[$ such that $P(|S_{k^2-1}| \leq k) \leq c_{27}$ for all $k \in \mathbb{N}$. The claim follows with $c_{25} := (c_{27})^{-1}$ and $c_{26} := -\ln c_{27}$. □

Lemma 9.5.3. *There exists c_{28} such that for all $n \geq c_{28}$*

$$P([B^n_{\text{average}\geq}]^c) \leq 2e^{-n}.$$

Proof. Clearly,

$$P([B^n_{\text{average}\geq}]^c) \leq P(\{|b^n_l - b^n_r| < 2n\} \setminus B^n_{\text{average}\geq}) + P(|b^n_l - b^n_r| \geq 2n). \tag{9.5.1}$$

By Lemma 9.5.11,

$$P(|b^n_l - b^n_r| \geq 2n) \leq 4 \cdot 2^{-2n}. \tag{9.5.2}$$

We denote by $\tilde\beta^n_l(k)$ and $\tilde\beta^n_r(k)$ the left and right end of the kth block of length $\geq i_n$ produced by S on block^n. We set $Y^n_k := \tilde\beta^n_r(k) - \tilde\beta^n_l(k) - 1$. Conditioned on $(\tilde\beta^n_l(k))_{k\geq 1}$, the random variables Y^n_k are independent with distribution $P(T_m \in \cdot | T_m \geq i_n)$ with $m = |\mathsf{block}^n|$. We have

$$[B^n_{\text{average}\geq}]^c = \Big\{3^{-\lfloor n^{0.2}\rfloor} \sum_{k\in[1,3^{\lfloor n^{0.2}\rfloor}]} Y^n_k < d_n\Big\} \subseteq \Big\{3^{-\lfloor n^{0.2}\rfloor} \sum_{k\in[1,3^{\lfloor n^{0.2}\rfloor}]} [Y^n_k \wedge n^4] < d_n\Big\}.$$

By part 1 of Lemma 9.4.1 and part 1 of Lemma 9.3.1, we have

$$E[Y_k^n \wedge n^4 |\ |b_l^n - b_r^n| < 2n] \geq d_n + 1 - 2c_{19}n^{-1}$$

for all n sufficiently large. Hence $P\{3^{-\lfloor n^{0.2}\rfloor}\sum_{k\in[1,3^{\lfloor n^{0.2}\rfloor}]}[Y_k^n \wedge n^4] < d_n\big|\ |b_l^n - b_r^n| < 2n, (\tilde{\beta}_l^n(k))_{k\geq 1}\}$ is a large deviation probability for the sequence $(Y_k^n \wedge n^4)_{k\geq 1}$ of bounded, independent random variables. Thus,

$$\begin{aligned}
& P(\{|b_l^n - b_r^n| < 2n\} \setminus B^n_{\text{average}\geq}) \\
\leq\ & P\Big(|b_l^n - b_r^n| < 2n, 3^{-\lfloor n^{0.2}\rfloor}\sum_{k\in[1,3^{\lfloor n^{0.2}\rfloor}]}[Y_k^n \wedge n^4] < d_n\Big) \\
\leq\ & P\Big(\Big\{3^{-\lfloor n^{0.2}\rfloor}\sum_{k\in[1,3^{\lfloor n^{0.2}\rfloor}]}[Y_k^n \wedge n^4] < d_n\Big\}\Big||b_l^n - b_r^n| < 2n\Big) \\
\leq\ & \exp\Big[-c_{29}\frac{3^{\lfloor n^{0.2}\rfloor}}{4n^8}\Big]
\end{aligned}$$

which is $\leq e^{-n}$ for all n sufficiently large; here $c_{29} > 0$ is a constant. Combining this with (9.5.1) and (9.5.2), the claim follows. □

Lemma 9.5.4. *There exists c_{30} such that for all $n \geq c_{30}$*

$$P([B^n_{\text{average}<}]^c) \leq e^{-n}.$$

Proof. For $k \geq 1$, let $\hat{\beta}_l^n(k)$ and $\hat{\beta}_r^n(k)$ be the left and right end of the kth block of length $\geq i_n$ produced by S on a block of ξ of length $< n$. We set $Z_k^n := \hat{\beta}_r^n(k) - \hat{\beta}_l^n(k) - 1$. Conditioned on $(\hat{\beta}_l^n(k))_{k\geq 1}$, the random variables $(Z_k^n)_{k\geq 1}$ are independent with distribution $P(T_m \in \cdot | T_m \geq i_n)$ with $m \in [1, n[$ equal to the length of the underlying block of ξ. We have

$$[B^n_{\text{average}<}]^c \subseteq \bigcup_{k\in[3^{\lfloor n^{0.2}\rfloor}, 2\cdot 3^{3n}]}\Big\{3^{-\lfloor n^{0.2}\rfloor}\sum_{i\in]k-3^{\lfloor n^{0.2}\rfloor},k]} Z_i^n \geq d_n\Big\}. \tag{9.5.3}$$

We abbreviate $B^{n,k}_{Z\text{ small}} := \{\forall i \in]k - 3^{\lfloor n^{0.2}\rfloor}, k] : Z_i^n \leq n^4\}$. Clearly,

$$\begin{aligned}
& P\Big(3^{-\lfloor n^{0.2}\rfloor}\sum_{i\in]k-3^{\lfloor n^{0.2}\rfloor},k]} Z_i^n \geq d_n\Big) \\
\leq\ & P([B^{n,k}_{Z\text{ small}}]^c) + P\Big(3^{-\lfloor n^{0.2}\rfloor}\sum_{i\in]k-3^{\lfloor n^{0.2}\rfloor},k]} Z_i^n \wedge n^4 \geq d_n\Big).
\end{aligned} \tag{9.5.4}$$

Using Lemma 9.4.1, Part 2, yields

$$P([B^{n,k}_{Z\text{ small}}]^c) \leq 3^{\lfloor n^{0.2}\rfloor}e^{-n^{1.5}} \leq e^{-n^{1.4}}$$

for all n sufficiently large. In order to get an upper estimate for the last term in (9.5.4), we use a large deviation estimate for independent, bounded random variables:

$$\begin{aligned}
& P\Big(3^{-\lfloor n^{0.2}\rfloor}\sum_{i\in]k-3^{\lfloor n^{0.2}\rfloor},k]} Z_i^n \wedge n^4 \geq d_n\Big) \\
=\ & E\Big[P\Big(3^{-\lfloor n^{0.2}\rfloor}\sum_{i\in]k-3^{\lfloor n^{0.2}\rfloor},k]} Z_i^n \wedge n^4 \geq d_n\Big|(\hat{\beta}_l^n(k))_{k\geq 1}\Big)\Big] \leq \exp\Big[-c_{31}\frac{3^{\lfloor n^{0.2}\rfloor}}{4n^8}\Big]
\end{aligned}$$

which is $\le e^{-n^2}$ for all n sufficiently large; here $c_{31} > 0$ is a constant. It follows from (9.5.3) and (9.5.4) that

$$P([B^n_{\text{average }<}]^c) \le 2 \cdot 3^{3n} \left[e^{-n^{1.4}} + e^{-n^2} \right]$$

which is $\le e^{-n}$ for all n sufficiently large. □

We define the filtration $\mathcal{H} := (\mathcal{H}_k)_{k \ge 0}$ by $\mathcal{H}_k := \sigma(S_i, \xi_z; i \in [0,k], z \in \mathbb{Z})$. The following lemma gives an estimate for $P([B^{n,s,t}_{\text{block enough}}]^c)$ for certain stopping times s; this estimate will be needed in the proof of Lemma 9.5.7.

Lemma 9.5.5. *There exists c_{32} such that for all $n, t \in \mathbb{N}$ and for all $\mathcal{H}$-adapted stopping times γ with $S_\gamma \in \partial\mathsf{block}^n$*

$$P([B^{n,\gamma,t}_{\text{block enough}}]^c) \le c_{32} t^{-1/30}.$$

Proof. For $i \ge 1$, we set $X_i := r^\gamma_n(i+1) - r^\gamma_n(i)$. Since the random walk is recurrent, X_i is well defined. Note that $\partial\mathsf{block}^n$ depends only on ξ and the random walk up to time γ. Thus, by the strong Markov property of S and the symmetry of the distribution of the random walk jumps $S_{k+1} - S_k$, the random variables X_i, $i \ge 1$, are i.i.d. conditioned on b^n_l and b^n_r. In particular, all X_i have the same moments.

The following inclusion holds:

$$[B^{n,\gamma,t}_{\text{block enough}}]^c \subseteq \Big\{ \sum_{i=1}^{\lfloor t^{3/10} \rfloor} X_i \ge t/2 \Big\} \subseteq \Big\{ \Big[\sum_{i=1}^{\lfloor t^{3/10} \rfloor} X_i^{1/3} \Big]^3 \ge t/2 \Big\}.$$

By Chebyshev's inequality,

$$\begin{aligned} P([B^{n,\gamma,t}_{\text{block enough}}]^c) &\le P\Big(\sum_{i=1}^{\lfloor t^{3/10} \rfloor} X_i^{1/3} \ge (t/2)^{1/3} \Big) \le (t/2)^{-1/3} E\Big(\sum_{i=1}^{\lfloor t^{3/10} \rfloor} X_i^{1/3} \Big) \\ &\le 2t^{-1/3} t^{3/10} E[X_1^{1/3}] = 2t^{-1/30} E[X_1^{1/3}]. \end{aligned} \tag{9.5.5}$$

Conditioned on $\partial\mathsf{block}^n$, X_1 is stochastically dominated by $r_2 - r_1$, where

$$\begin{aligned} r_1 &:= \min\{k \ge 0 : (S_k, S_{k+1}) = (0,1)\}, \\ r_2 &:= \min\{k > r_1 : (S_k, S_{k+1}) = (0,1)\}. \end{aligned}$$

Let T denote the first return time to the origin of the random walk S. Let T_i, $i \ge 1$, be i.i.d. with the same distribution as T. Since the random walk starting at 0 can make k holdings at 0 and m excursions to the left before hitting 1, we obtain

$$\begin{aligned} E[(r_2 - r_1)^{1/3}] &= \sum_{k=0}^\infty \sum_{m=0}^\infty q^k p^{2m+1} \binom{m+1+k}{k} E[(k+2m+1+\sum_{i=1}^m T_i)^{1/3}] \\ &\le c_{33} + c_{34} E[T^{1/3}] \end{aligned}$$

with constants $c_{33}, c_{34} > 0$. By P3, page 381, of [24], $\lim_{n\to\infty} \sqrt{n} P(T > n) = c_{35}$ for some $c_{35} > 0$. Thus with a constant $c_{36} > 0$, $E[T^{1/3}] \le 1 + \sum_{n=1}^\infty P(T^{1/3} > n) \le 1 + c_{36} \sum_{n=1}^\infty n^{-3/2} < \infty$. Hence $E[X_1^{1/3}] < \infty$, and the claim follows. □

Lemma 9.5.6. *There exists c_{37} such that for all $n \geq c_{37}$*

$$P([B^n_{\text{b small}}]^c) \leq e^{-n}.$$

Proof. Starting with $-3^n/3$, we partition the set $[-3^n/3, 3^n/3]$ into N disjoint intervals $I_1, I_2, \ldots, I_N$ of length $n+2$, $N := \lfloor (2 \cdot 3^{n-1}+1)/(n+2) \rfloor$. Let X_k be the Bernoulli random variable defined as follows: $X_k := 1$ if $\xi|I_k$ is a block of length n and $X_k := 0$ otherwise. Then X_k, $k \in [1, N]$ are i.i.d. with $P(X_k = 1) = 2^{-n-2}$. Moreover, $[B^n_{\text{b small}}]^c \subseteq \{\sum_{k=1}^N X_k = 0\}$. Hence

$$P([B^n_{\text{b small}}]^c) \leq P\Big[\sum_{k=1}^N X_k = 0\Big] = \big[1 - 2^{-n-2}\big]^N = \exp\big[N \ln[1 - 2^{-n-2}]\big] \leq e^{-2^{-n-2}N};$$

for the last inequality we used the estimate $\ln(1+x) \leq x$ for $|x| < 1$. Since $2^{-n-2}N \geq n$ for all n sufficiently large, the claim follows. □

Lemma 9.5.7. *There exist $c_{38}, c_{39}, c_{40} > 0$ such that for all $n \geq c_{38}$*

$$P([B^n_{\text{enough blocks}}]^c) \leq c_{39} e^{-c_{40} n^{0.3}}.$$

Proof. Recall the definition of $r^s_n(k)$ from Definition 9.5.2. Let Y^n_k be the Bernoulli random variable defined by $Y^n_k := 1$ if $r^{\gamma_n}_n(k)$ is the left endpoint of a block of χ of length $\geq i_n$ and $Y^n_k := 0$ otherwise. Note that if $Y^n_k = 1$, then the block of χ starting at $r^{\gamma_n}_n(k)$ is generated by the random walk on block^n.

Let $t := 3^{\lfloor n^{0.3} \rfloor}$. If the event $B^{n,\gamma_n,t}_{\text{block enough}}$ holds, then $r^{\gamma_n}_n(k) < \infty$ for all $k \in [1, \lfloor t^{3/10} \rfloor[= [1, \lfloor 3^{\lfloor n^{0.3} \rfloor/10} \rfloor[$. Thus

$$B^{n,\gamma_n,t}_{\text{block enough}} \cap \Big\{ \sum_{k=1}^{\lfloor t^{3/10} \rfloor} Y^n_k > 3^{\lfloor n^{0.2} \rfloor} \Big\} \subseteq B^n_{\text{enough blocks}}. \tag{9.5.6}$$

Because of Lemma 9.5.5 and $t = 3^{\lfloor n^{0.3} \rfloor}$,

$$P([B^{n,\gamma_n,t}_{\text{block enough}}]^c) \leq c_{32} t^{-1/30} = c_{32} 3^{-\lfloor n^{0.3} \rfloor/30}. \tag{9.5.7}$$

It remains to estimate $P(B^{n,\gamma_n,t}_{\text{block enough}} \cap \{\sum_{k=1}^{\lfloor t^{3/10} \rfloor} Y^n_k \leq 3^{\lfloor n^{0.2} \rfloor}\})$. Recall that T_m denotes the first hitting time of the set $\{-1, m\}$ by the random walk. Since block^n has length $\geq n$, it follows from Lemma 9.4.1 that

$$P(Y^n_k = 1 | r^{\gamma_n}_n(k) < \infty) \geq P(T_n \geq i_n) \geq n^{-c_{28}}.$$

Let $\tilde{Y}^n_k$, $k \geq 1$, be i.i.d. Bernoulli random variables with parameter $n^{-c_{28}}$ (on a possibly enlarged probability space). Then

$$\begin{aligned} P\left(B^{n,\gamma_n,t}_{\text{block enough}} \cap \Big\{ \sum_{k=1}^{\lfloor t^{3/10} \rfloor} Y^n_k \leq 3^{\lfloor n^{0.2} \rfloor} \Big\} \right) &\leq P\left(\sum_{k=1}^{\lfloor t^{3/10} \rfloor} Y^n_k \leq 3^{\lfloor n^{0.2} \rfloor} \,\middle|\, B^{n,\gamma_n,t}_{\text{block enough}} \right) \\ &\leq P\left(\sum_{k=1}^{\lfloor t^{3/10} \rfloor} \tilde{Y}^n_k \leq 3^{\lfloor n^{0.2} \rfloor} \right). \end{aligned}$$

For $k \geq 1$, let

$$Z_i^n := \sum_{j \in](k-1)\lfloor n^{c_{28}} \rfloor, k \lfloor n^{c_{28}} \rfloor]} \tilde{Y}_j^n.$$

By the Poisson convergence theorem (see e.g.[5], page 137, theorem (6.1)), Z_i^n converges weakly as $n \to \infty$ to a Poisson(1)-distributed random variable. Thus,

$$\tilde{Z}_k^n := \sum_{i \in]3(k-1)3^{\lfloor n^{0.2} \rfloor}, 3k3^{\lfloor n^{0.2} \rfloor}]} Z_i^n,$$

$k \geq 1$, satisfy $3^{-\lfloor n^{0.2} \rfloor} \tilde{Z}_k^n \to 3$ as $n \to \infty$ in probability. Consequently, there exists $c_{41} \in]0,1[$ such that for all n

$$P(\tilde{Z}_k^n \leq 3^{\lfloor n^{0.2} \rfloor}) \leq c_{41}.$$

Note that each $\tilde{Z}_k^n$ is the sum of $3 \cdot 3^{\lfloor n^{0.2} \rfloor} \lfloor n^{c_{28}} \rfloor$ random variables $\tilde{Y}_j^n$. We set $T := \lfloor t^{3/10} \rfloor / (3 \cdot 3^{\lfloor n^{0.2} \rfloor} \lfloor n^{c_{28}} \rfloor) = \lfloor 3^{3 \lfloor n^{0.3} \rfloor / 10} \rfloor 3^{-\lfloor n^{0.2} \rfloor - 1} \lfloor n^{c_{28}} \rfloor^{-1} \geq 3^{n^{0.3}/10}$ for all n sufficiently large. Then

$$\begin{aligned} P(\sum_{k=1}^{\lfloor t^{3/10} \rfloor} \tilde{Y}_k^n \leq 3^{\lfloor n^{0.2} \rfloor}) &\leq P(\sum_{k=1}^{\lfloor T \rfloor} \tilde{Z}_k^n \leq 3^{\lfloor n^{0.2} \rfloor}) \leq P(\bigcap_{k=1}^{\lfloor T \rfloor} \{\tilde{Z}_k^n \leq 3^{\lfloor n^{0.2} \rfloor}\}) \\ &= P(\tilde{Z}_1^n \leq 3^{\lfloor n^{0.2} \rfloor})^{\lfloor T \rfloor} \leq (c_{41})^{3^{\lfloor n^{0.3} \rfloor / 10}} \leq e^{-n} \end{aligned}$$

for all n sufficiently large. Combining this with (9.5.6) and (9.5.7), the claim follows. □

Lemma 9.5.8. *There exists a constant c_{42} such that for all $n \geq c_{42}$*

$$P([B^n_{\text{only block}}]^c) \leq 2^{-n^{0.4}/2}.$$

Proof. By definition, $[B^n_{\text{only block}}]^c = [B^{n+}_{\text{only}}]^c \cup [B^{n-}_{\text{only}}]^c$ with

$$[B^{ni}_{\text{only}}]^c = \{\exists \text{block of length } \geq \lfloor n^{0.4} \rfloor \text{ with a distance } \leq 3^{\lfloor n^{0.3} \rfloor} \text{ from } \partial \text{block}^{ni}\}$$

for $i \in \{+,-\}$. For $x \in \mathbb{Z}$, the probability that there is a block of length $\geq \lfloor n^{0.4} \rfloor$ in $\xi|[x, x+3^{\lfloor n^{0.3} \rfloor}]$ is bounded by $3^{\lfloor n^{0.3} \rfloor} \cdot 2 \cdot 2^{-\lfloor n^{0.4} \rfloor} \leq 2^{-n^{0.4}/2}/4$ for all $n \geq c_{42}$ with a constant c_{42}. This is because the probability that a piece of scenery of length $\lfloor n^{0.4} \rfloor$ is colored with the same color equals $2 \cdot 2^{-\lfloor n^{0.4} \rfloor}$ and in $[x, x+3^{\lfloor n^{0.3} \rfloor}]$ there are at most $3^{\lfloor n^{0.3} \rfloor}$ possible left ends for such a constantly colored piece of scenery. Thus, conditioning on the endpoints of block^{ni} yields the claim. □

Lemma 9.5.9. *There exists c_{43} such that for all $n \geq c_{43}$*

$$P([B^n_{\text{rw hits}}]^c) \leq e^{-n}.$$

Proof. By the definition of $B^n_{\text{rw hits}}$ and Lemma 9.5.2, we have

$$\begin{aligned} P([B^n_{\text{rw hits}}]^c) &= P(\forall k \in [0, 3^{3n}[: S_k \in]-3^n/3, 3^n/3[) \\ &\leq c_{25} \exp\Big(- c_{26} \frac{3^{3n}}{(2 \cdot 3^{n-1}+1)^2} \Big) \leq c_{25} \exp(-c_{26} 3^n). \end{aligned}$$

The last expression is $\leq e^{-n}$ for all n sufficiently large. □

Lemma 9.5.10. *There exists c_{44} such that for all $n \geq c_{44}$*

$$P([B^n_{\text{short block}}]^c) \leq e^{-n}.$$

Proof. If we set

$$B^{n,k}_{\text{short}} := \left\{ \begin{array}{l} \text{If the } k\text{th block of } \chi \text{ was generated on a block of } \xi \text{ of length} \\ \leq \lfloor n^{0.4} \rfloor, \text{ then it has length} \geq n^2 \end{array} \right\},$$

then $[B^n_{\text{short block}}]^c \subseteq \bigcup_{k=1}^{3\cdot 3^{3n}} B^{n,k}_{\text{short}}$. Let $k \geq 1$. We use the strong Markov property of the random walk at the time when it enters the block of ξ underlying the kth block of χ and Lemma 9.5.2 to obtain

$$\begin{aligned} P(B^{n,k}_{\text{short}}) &\leq P(\forall k \in [0, n^2[\colon S_k \in [0, \lfloor n^{0.4} \rfloor[) \\ &\leq c_{25} \exp\Big[- c_{26} \frac{n^2}{\lfloor n^{0.4} \rfloor^2} \Big] \leq c_{25} e^{-c_{26} n^{1.2}}. \end{aligned}$$

Thus $P([B^n_{\text{short block}}]^c) \leq 3 c_{25} 3^{3n} e^{-c_{26} n^{1.2}} \leq e^{-n}$ for all n sufficiently large. □

Lemma 9.5.11. *For all $n \in \mathbb{N}$,*

$$P([B^n_{\text{size block}}]^c) \leq 4 \cdot 2^{-2n}$$

Proof. By definition, $[B^n_{\text{size block}}]^c = \{\exists i \in \{-,+\} \text{ such that } |b^{ni}_l - b^{ni}_r| \geq 2n\}$. If $|b^{ni}_l - b^{ni}_r| \geq 2n$, then the block starting at b^{ni}_l has length $\geq 2n$. Thus conditioning on b^{ni}_l and using that the length of a block starting at a point x is geometrically distributed, we obtain $P(|b^{ni}_l - b^{ni}_r| \geq 2n) = \sum_{k \geq 2n} 2^{-k} = 2^{1-2n}$ and the claim follows. □

9.5.4 The estimate of block^n

In this subsection, we prove Proposition 9.3.1. Let $T \in \{H^n_l, H^n_r\}$; recall the definitions of H^n_l and H^n_r from (9.3.3). We abbreviate

$$\chi^{n,T} := \chi|[T, T + 3^{\lfloor n^{0.3} \rfloor}[. \tag{9.5.8}$$

We have $\{\mathsf{estimated\text{-}block}^n(\chi^{n,T}) = \mathsf{block}^n\} = B^{n,T}_{\text{color ok}} \cap B^{n,T}_{\text{length ok}}$ with

$$\begin{aligned} B^{n,T}_{\text{color ok}} &:= \{\mathsf{estimated\text{-}block}^n(\chi^{n,T}) \text{ and } \mathsf{block}^n \text{ have the same color}\} \text{ and} \\ B^{n,T}_{\text{length ok}} &:= \{\mathsf{estimated\text{-}block}^n(\chi^{n,T}) \text{ and } \mathsf{block}^n \text{ have the same length}\}. \end{aligned}$$

Proof of Proposition 9.3.1. First, we show that the event $B^{n,T}_{\text{length ok}}$ has high probability. We define the events

$$\begin{aligned} \tilde{B}^{n,T}_{\text{average} \geq} &:= \left\{ \begin{array}{l} \text{The average lengths of the first } 3^{\lfloor n^{0.2} \rfloor} \text{ blocks of length} \geq i_n \\ \text{produced by } S|[T,\infty[\text{ on } \mathsf{block}^n \text{ is} \geq d_n \end{array} \right\}, \\ \tilde{B}^{n,T}_{\text{average} <} &:= \bigcap_{m \in [n,2n]} \left\{ \begin{array}{l} \text{If all blocks } \chi|\big[\bar\beta^n_l(i,\chi^{n,T}), \bar\beta^n_r(i,\chi^{n,T})\big],\ i \in [1, 3^{\lfloor n^{0.2} \rfloor}], \\ \text{are generated on blocks of } \xi \text{ of length} < m, \text{ then} \\ 3^{-\lfloor n^{0.2} \rfloor} \sum_{i \in [1, 3^{\lfloor n^{0.2} \rfloor}]} \big[\bar\beta^n_r(i,\chi^{n,T}) - \bar\beta^n_l(i,\chi^{n,T}) - 1\big] < d^n_m \end{array} \right\}, \\ \tilde{B}^{n,T}_{\text{enough blocks}} &:= \left\{ \begin{array}{l} S|[T, T + 3^{\lfloor n^{0.3} \rfloor}[\text{ produces} > 3^{\lfloor n^{0.2} \rfloor} \text{ blocks of length} \geq i_n \\ \text{on } \mathsf{block}^n \end{array} \right\}, \end{aligned}$$

and claim that the following inclusion holds for all $n \in \mathbb{N}$:

$$\tilde{B}^{n,T}_{\text{average} \geq} \cap \tilde{B}^{n,T}_{\text{average} <} \cap \tilde{B}^{n,T}_{\text{enough blocks}} \cap B^n_{\text{only block}} \cap \{|b^n_l - b^n_r| \leq 2n\} \subseteq B^{n,T}_{\text{length ok}}. \quad (9.5.9)$$

Suppose the events $\tilde{B}^{n,T}_{\text{average} \geq}$, $\tilde{B}^{n,T}_{\text{average} <}$, $\tilde{B}^{n,T}_{\text{enough blocks}}$, $B^n_{\text{only block}}$, and $\{|b^n_l - b^n_r| \leq 2n\}$ hold. Because of $\tilde{B}^{n,T}_{\text{enough blocks}} \cap \tilde{B}^{n,T}_{\text{average} \geq}$, we have

$$3^{-\lfloor n^{0.2} \rfloor} \sum_{i \in [1, 3^{\lfloor n^{0.2} \rfloor}]} \left[\bar{\beta}^n_r(i, \chi^{n,T}) - \bar{\beta}^n_l(i, \chi^{n,T}) - 1\right] \geq d_n = d^n_n.$$

Since $\tilde{B}^{n,T}_{\text{average} <}$ holds, at least one of the blocks $\chi|\left[\bar{\beta}^n_l(i, \chi^{n,T}), \bar{\beta}^n_r(i, \chi^{n,T})\right]$, $i \in [1, 3^{\lfloor n^{0.2} \rfloor}]$, is generated on a block of ξ of length $\geq n$. By $B^n_{\text{only block}}$, this block of ξ must be block^n. We assumed $|b^n_l - b^n_r| \leq 2n$. Hence, block^n has length $m \in [n, 2n-1]$. Because of $\tilde{B}^{n,T}_{\text{average} <}$, the length of block^n is the unique $m \in [n, 2n-1]$ such that (9.3.2) holds. By definition, $\mathsf{estimated\text{-}block}^n(\chi^{n,T})$ has length m as well. This proves inclusion (9.5.9).

It follows from Lemma 9.5.3 that

$$P([\tilde{B}^{n,T}_{\text{average} \geq}]^c) \leq 2e^{-n}$$

for all n sufficiently large. The proof of Lemma 9.5.4 can easily be adapted to show that

$$P([\tilde{B}^{n,T}_{\text{average} <}]^c) \leq e^{-n}$$

for all n sufficiently large. It follows from Lemma 9.5.7 that

$$P([\tilde{B}^{n,T}_{\text{enough blocks}}]^c) \leq c_{39} e^{-c_{40} n^{0.3}}$$

for all n sufficiently large. Combining the preceding estimates with inclusion (9.5.9), Lemma 9.5.8, and Lemma 9.5.11 yields

$$P([B^{n,T}_{\text{length ok}}]^c) \leq c_{45} e^{-c_{46} n^{0.3}} \quad (9.5.10)$$

for all $n \geq c_{47}$ with constants $c_{47}, c_{45}, c_{46} > 0$.

Next, we prove that the event $B^{n,T}_{\text{color ok}}$ has high probability. We define the event $B^{n,T}_{\text{large block}} := \{S | [T, T + 3^{\lfloor n^{0.3} \rfloor}[\text{ produces at least one block of length } \geq n^2\}$ and claim

$$B^{n,T}_{\text{large block}} \cap B^n_{\text{only block}} \cap \Theta^{-T}\left[B^n_{\text{short block}}\right] \subseteq B^{n,T}_{\text{color ok}}. \quad (9.5.11)$$

Suppose the events $B^{n,T}_{\text{large block}}$, $B^n_{\text{only block}}$, and $\Theta^{-T}\left[B^n_{\text{short block}}\right]$ hold, and recall that the color of $\mathsf{estimated\text{-}block}^n(\chi^{n,T})$ is defined to be the color of the first block of length $\geq n^2$ in $\chi^{n,T}$. By $\Theta^{-T}\left[B^n_{\text{short block}}\right]$, the color of $\mathsf{estimated\text{-}block}^n(\chi^{n,T})$ is the color of a block which was generated by S on a block of ξ of length $> \lfloor n^{0.4} \rfloor$. Since $B^n_{\text{only block}}$ holds, this block must be generated on block^n. Hence $B^{n,T}_{\text{color ok}}$ holds.

It follows from (9.5.11) that

$$P([B^{n,T}_{\text{color ok}}]^c) \leq P([B^{n,T}_{\text{large block}}]^c) + P([B^n_{\text{only block}}]^c) + P(\left[\Theta^{-T}[B^n_{\text{short block}}]\right]^c).$$

Using $i_n \geq (8/p) n^2 \ln n \geq n^2$ for all $n \geq 2$ and Lemma 9.5.7, we obtain $P([B^{n,T}_{\text{large block}}]^c) \leq c_{39} e^{-c_{40} n^{0.3}}$. By Lemma 4.1 of [19], the shift Θ preserves the measure P. Hence, we conclude from Lemmas 9.5.8 and 9.5.10

$$P([B^{n,T}_{\text{color ok}}]^c) \leq e^{-c_{48} n^{0.3}} \quad (9.5.12)$$

for all n sufficiently large with a constant $c_{48} > 0$. The claim follows. □

9.6 The stopping times $\nu_b^{n,T}$

In this section, we prove Proposition 9.3.2. Let $T \in \{H_l^n, H_r^n\}$; recall the definitions of H_l^n and H_r^n from (9.3.3).

9.6.1 Definitions of events

We collect in alphabetical order definitions of events, which will be needed in the sequel.

Definition 9.6.1. *Let $b_c^n := \lfloor (b_l^n + b_r^n)/2 \rfloor$. Thus, b_c^n is the center of the interval $[b_l^n, b_r^n]$ or it differs in absolute value from the center by $1/2$. We define*

$$B_{b_c \text{ often}}^{n,T} := \left\{ S | [T, T + 3^{10\alpha \lfloor n^{0.2} \rfloor}/2[\text{ visits } b_c^n \text{ at least } 3^{3\alpha \lfloor n^{0.2} \rfloor} \text{ times} \right\}.$$

Definition 9.6.2. *Let*

$$\mathbb{S}'_{n,T}(\xi, \chi) \;:=\; \{t \in [T, \infty[: S_t = b_c^n, |S_t - S_k| \le n/4 \text{ for all } k \in [t, t + n^2/16[\}.$$

For $t \in \mathbb{S}'_{n,T}$, let $\hat{t}$ be the right end of the block of χ which contains t as an inner point. We set

$$\mathbb{S}_{n,T} := \{\hat{t} : t \in \mathbb{S}'_{n,T}\}.$$

For $k \ge 1$, we denote by $\zeta_{n,T}(k)$ the kth hitting time of b_c^n which is $\ge T$. We define

$$\tilde{B}_{\text{when back recog}}^{n,T} \;:=\; \{3^{-2\alpha \lfloor n^{0.2} \rfloor} | \{k \in [1, 3^{2\alpha \lfloor n^{0.2} \rfloor}] : \zeta_{n,T}(2k \cdot 3^{3 \lfloor n^{0.2} \rfloor}) \in \mathbb{S}'_{n,T}(\xi, \chi)| > q/2\}.$$

Definition 9.6.3. *Recall the definition of the $\nu_b^{n,T}(k)$'s from Section 9.3. We define*

$$\tilde{E}_{\text{no error } \nu_b}^{n,T} := \{\forall k \ge 1 : \text{if } \nu_b^{n,T}(k) < T + 3^{10\alpha \lfloor n^{0.2} \rfloor}, \text{ then } S_{\nu_b^{n,T}(k)} \in \partial \mathsf{block}^n\}.$$

9.6.2 Proof of Proposition 9.3.2

We start this section with two combinatorial lemmas. We set

$$\mathbb{T}^{n,T} := \{\tilde{\nu}_b^{n,T}(k) : k \ge 1\}.$$

Lemma 9.6.1. *For all $n \in \mathbb{N}$ the following inclusion holds:*

$$\tilde{E}_{\text{no error } \nu_b}^{n,T} \cap \{\mathbb{S}_{n,T} \cap [T, T + 3^{10\alpha \lfloor n^{0.2} \rfloor}[\subseteq \mathbb{T}^{n,T}\} \cap B_{b_c \text{ often}}^{n,T} \cap \tilde{B}_{\text{when back recog}}^{n,T} \subseteq E_{\nu_b \text{ ok}}^{n,T}.$$

Proof. The proof is very similar to the proof of Lemma 9.8.1 below. □

Lemma 9.6.2. *The inclusion $\mathbb{S}_{n,T} \cap [T, T + 3^{10\alpha \lfloor n^{0.2} \rfloor}[\subseteq \mathbb{T}^{n,T}$ holds P-almost surely.*

Proof. Let $t \in \mathbb{S}_{n,T} \cap [T, T + 3^{10\alpha \lfloor n^{0.2} \rfloor}[$. Then, t is the right end of a block of $\chi | [T, T + 3^{10\alpha \lfloor n^{0.2} \rfloor}[$. Let $s \in \mathbb{S}'_{n,T}$ be such that $t = \hat{s}$. Since at time s, the random walk is in the center of block^n and $b_r^n - b_l^n \ge n$, we have $S | [s, s + n^2/16[\subseteq [b_l^n, b_r^n]$, and consequently, t is the right end of a block of length $\ge n^2/16$. Hence $t \in \mathbb{T}^{n,T}$. □

Proof of Proposition 9.3.2. By Lemmas 9.6.1 and 9.6.2, we have

$$P([E^{n,T}_{\nu_b \text{ ok}}]^c) \leq P([\tilde{E}^{n,T}_{\text{no error } \nu_b}]^c) + P([B^{n,T}_{b_c \text{ often}}]^c) + P([\tilde{B}^{n,T}_{\text{when back recog}}]^c).$$

The claim follows from Lemmas 9.6.5, 9.6.3, and 9.6.4 below. □

Lemma 9.6.3. *There exist constants $c_{49}, c_{50} > 0$ such that for all $n \geq c_{49}$ and all $T \in \{H_l^n, H_r^n\}$*

$$P([B^{n,T}_{b_c \text{ often}}]^c) \leq e^{-c_{50} n^{0.2}}.$$

Proof. Using arguments which are very similar to the proof of Lemma 9.8.4 below, one can prove that

$$P(\{|b_l^{ni} - b_r^{ni}| < 2n\} \setminus B^n_{b_c \text{ often}}) \leq c_{51} 3^{-\alpha \lfloor n^{0.2} \rfloor / 3}$$

with a constant c_{51}. By Lemma 9.5.11, $P(|b_l^{ni} - b_r^{ni}| \geq 2n) \leq 4 \cdot 2^{-2n}$. The claim follows. □

Lemma 9.6.4. *There exists c_{52} such that for all $n \geq c_{52}$ and all $T \in \{H_l^n, H_r^n\}$*

$$P([\tilde{B}^{n,T}_{\text{when back recog}}]^c) \leq e^{-n}.$$

Proof. By the Markov property of the random walk, the random variables $Y_k := 1\{\zeta_{n,T}(2k \cdot 3^{3\lfloor n^{0.2} \rfloor}) \in \mathbb{S}'_{n,T}(\xi, \chi)\}$, $k \geq 1$, are i.i.d. Furthermore, we have for all $k \geq 1$,

$$\begin{aligned} P(\zeta_{n,T}(k) \in \mathbb{S}'_{n,T}) &= P(|S_k| \leq n/4 \text{ for all } k \in [0, n^2/16[) \\ &\geq 1 - 16 n^{-2} \mathrm{Var}(S_{n^2/16}) = 1 - 2p = q; \end{aligned}$$

for the last inequality, we used Doob's submartingale inequality (see e.g. [5], page 250, example 4.1) and $\mathrm{Var}(S_1) = 2p$. The event $[\tilde{B}^{n,T}_{\text{when back recog}}]^c$ is a large deviation event for the i.i.d. Bernoulli random variables Y_k. Thus, with a constant $c_{53} > 0$, $P([\tilde{B}^{n,T}_{\text{when back recog}}]^c) \leq \exp(-3^{2\alpha \lfloor n^{0.2} \rfloor} c_{53})$ which is $\leq e^{-n}$ for all n sufficiently large. □

Lemma 9.6.5. *There exist constants $c_{54}, c_{55} > 0$ such that for all $n \geq c_{54}$ and all $T \in \{H_l^n, H_r^n\}$*

$$P([\tilde{E}^{n,T}_{\text{no error } \nu_b}]^c) \leq e^{-c_{55} n^{0.4}}.$$

Proof. Clearly,

$$[\tilde{E}^{n,T}_{\text{no error } \nu_b}]^c \subseteq \bigcup_{k \in [1, 3^{10\alpha \lfloor n^{0.2} \rfloor}]} \{\nu_b^{n,T}(k) < T + 3^{10\alpha \lfloor n^{0.2} \rfloor} \text{ and } S_{\nu_b^{n,T}(k)} \notin \partial \mathsf{block}^n\}.$$

Recall the definition of $B^n_{\text{only block}}$ (Definition 9.5.5) and recall that the event $B^n_{\text{only block}}$ is $\sigma(\xi)$-measurable. If $B^n_{\text{only block}}$ holds, $\nu_b^{n,T}(k) < T + 3^{10\alpha \lfloor n^{0.2} \rfloor}$, and $S_{\nu_b^{n,T}(k)} \notin \partial \mathsf{block}^n$, then $\nu_b^{n,T}(k)$ is the right end of a block of ξ of length $< \lfloor n^{0.4} \rfloor$. Hence, using the strong Markov property of S at the time when the random walker enters the block on which the block of χ ending at $\nu_b^{n,T}(k)$ is generated, we obtain

$$\begin{aligned} & P_\xi(B^n_{\text{only block}} \cap \{\nu_b^{n,T}(k) < T + 3^{10\alpha \lfloor n^{0.2} \rfloor} \text{ and } S_{\nu_b^{n,T}(k)} \notin \partial \mathsf{block}^n\}) \\ \leq\ & 1_{B^n_{\text{only block}}} P_\xi(\forall k \in [0, n^2/16[: |S_k| < \lfloor n^{0.4} \rfloor) \\ \leq\ & c_{25} \exp\Big(- c_{26} \cdot \frac{n^2}{16 n^{0.8}} \Big) = c_{25} e^{-c_{26} n^{1.2}/16}; \end{aligned}$$

for the last inequality we used Lemma 9.5.2. Thus,

$$P(B^n_{\text{only block}} \setminus \tilde{E}^{n,T}_{\text{no error }\nu_b}) \le c_{25} 3^{10\alpha\lfloor n^{0.2}\rfloor} e^{-c_{26} n^{1.2}/16}$$

which is $\le e^{-n}$ for all n sufficiently large. Combining this with Lemma 9.5.8, completes the proof. □

9.7 The algorithm SmallAlg^n

Let $n \in \mathbb{N}$ be fixed, but large.

In this section, we define a map SmallAlg^n which fulfills the claim of Theorem 9.3.2. Given $3^{\lfloor n^{0.3}\rfloor}$ observations collected by the random walker starting in the set $\partial\mathsf{block}^n$, SmallAlg^n reconstructs with high probability a piece of scenery around block^n of length of the order $3^{n^{0.2}}$.

The definition of SmallAlg^n has some similarities with the definition of $\text{BigAlg}^{\lfloor n^{0.2}\rfloor}$ in Section 5 of [22]. Alas, we cannot directly use the map $\text{BigAlg}^{\lfloor n^{0.2}\rfloor}$ which reconstructs a piece of scenery given observations, a "typical" piece of scenery close to the origin, and a sequence of stopping times. Here, we would like to reconstruct a piece of the scenery around block^n, but block^n is not typical for the scenery close to the origin. We have to take this into account in the definition of SmallAlg^n.

9.7.1 Finding words

In order to reconstruct a piece of ξ of length of the order $3^{\lfloor n^{0.2}\rfloor}$, we look in the observations χ for words of length $c_1\lfloor n^{0.2}\rfloor$ occurring in the scenery; here $c_1 > 0$ is a constant chosen as in Section 9.2. The idea is to find enough of these words such that we can assemble them to obtain a bigger piece of the scenery. Below we review a criterion from [22] to find such words in the scenery. We abbreviate

$$m = \lfloor n^{0.2}\rfloor \quad \text{and} \quad \nu^n_b(k) := \nu^{n,0}_b(k) \text{ for all } k \ge 1.$$

Let $\eta \in \bigcup_{k\ge 3^{3m}} \mathcal{C}^{[0,k[}$. We consider $\tilde{\mathrm{O}}^m_1 \mathrm{O}^m_2 \tilde{\mathrm{O}}^m_3$ with the following properties: $\tilde{\mathrm{O}}^m_1$ consists of the first $c_1 m$ blocks of η after time 3^{2m}, O^m_2 equals the following $c_1 m/2$ observations in η extended until the next block starts, and $\tilde{\mathrm{O}}^m_3$ consists of the following $c_1 m$ blocks of η. We do the same thing with $\theta^{\nu^n_b(k)}(\eta)$ for all $k \in [1, 3^{\alpha m}]$, i.e. we collect observations after each stopping time. The words $\tilde{\mathrm{O}}^m_1$ and $\tilde{\mathrm{O}}^m_3$ are used to find those O^m_2 which occur in ξ close to block^n. In fact, we consider instead of $\tilde{\mathrm{O}}^m_1$ the sequence $\mathrm{O}^m_1 \in \{1,2,3,4,5\}^{c_1 m}$ where the jth component equals the minimum of 5 and the length of the jth block of $\tilde{\mathrm{O}}^m_1$. The same is done with $\tilde{\mathrm{O}}^m_3$. More formally:

Definition 9.7.1. *Let $\eta \in \bigcup_{k\ge 3^{3m}} \mathcal{C}^{[0,k[}$, and let $\eta^m := \eta|[3^{2m}, 3^{3m}[$. We denote by $B_k(\eta)$ the kth block of η if η possesses at least k blocks; otherwise $B_k(\eta) := 101 \in \mathcal{C}^{[3^{3m}, 3^{3m}+3[}$. Let $o^m_l(\eta)$ be the right end of $B_{c_1 m}(\eta^m)$, the $c_1 m$th block of η^m. Furthermore let $\tilde{o}^m_r(\eta)$ be the left end of the first block of $\eta^m|[o^m_l(\eta) + c_1 m/2 - 2, 3^{3m}[$ and set $o^m_r(\eta) := \tilde{o}^m_r(\eta) + 1$. If $\eta^m|[o^m_l(\eta) + c_1 m/2 - 2, 3^{3m}[$ does not contain a block, then we set $o^m_r(\eta) := o^m_l(\eta)$. We define $\mathrm{O}^m := (\mathrm{O}^m_1, \mathrm{O}^m_2, \mathrm{O}^m_3)$ by*

$$\begin{aligned}
\mathrm{O}^m_1(\eta) &:= (|B_k(\eta^m)| \wedge 5)_{k\in[1,c_1 m]},\\
\mathrm{O}^m_2(\eta) &:= \eta|[o^m_l(\eta), o^m_r(\eta)],\\
\mathrm{O}^m_3(\eta) &:= (|B_k(\theta^{\tilde{o}^m_r(\eta)}(\eta))| \wedge 5)_{k\in[1,c_1 m]}.
\end{aligned}$$

The following picture illustrates the definitions for $c_1 n = 6$:

$$\eta = \underbrace{1110\ldots01110010}_{\eta|[0,3^{2m}[}\underbrace{0111010000011000\boxed{1}000}_{\tilde{\mathrm{O}}_1^m(\eta)}\underbrace{0\boxed{1}1100101111000100111110\ldots}_{\tilde{\mathrm{O}}_3^m(\eta)}$$

$\eta_{o_l^m}$ and $\eta_{o_r^m}$ are marked with boxes. In this example, we have $\mathrm{O}_1^m(\eta) = (3,1,1,5,2,3)$, $\mathrm{O}_2^m(\eta) = 100001$, $\mathrm{O}_3^m(\eta) = (3,2,1,1,4,3)$.

Definition 9.7.2. *For $\eta \in \bigcup_{k\geq 3^{3m}} \mathcal{C}^{[0,k[}$ we define $\mathrm{L}^m := (\mathrm{L}_1^m, \mathrm{L}_2^m, \mathrm{L}_3^m)$ by*

$$\mathrm{L}_1^m(\eta) := \mathrm{O}_1^m(\eta), \quad \mathrm{L}_2^m(\eta) := \eta|[o_l^m(\eta), o_r^m(\eta) - 2], \quad \mathrm{L}_3^m(\eta) := |B_1(\theta^{\tilde{o}_r^m(\eta)}(\eta))|.$$

Furthermore, we define $\mathrm{R}^m := (\mathrm{R}_1^m, \mathrm{R}_2^m, \mathrm{R}_3^m)$ by

$$\mathrm{R}_1^m(\eta) := |B_{c_1 m}(\eta^m)|, \quad \mathrm{R}_2^m(\eta) := \eta|[o_l^m(\eta)+1, o_r^m(\eta)], \quad \mathrm{R}_3^m(\eta) := \mathrm{O}_3^m(\eta).$$

Thus $\mathrm{L}_3^m(\eta)$ is the *non-truncated* length of the first block of $\eta[o_l^m(\eta) + c_1 m/2 - 2, 3^{3m}[$ and $\mathrm{R}_1^m(\eta)$ is the *non-truncated* length of the $(c_1 m)$th block of η^m.

Definition 9.7.3. *For $A \in \{O, L, R\}$ and $\eta \in \mathcal{C}^{[0,2\cdot 3^{10\alpha m}[}$, we define the empirical distribution of O^m observed after the times $\nu_b^n(k)$, $k \in [1, 3^{\alpha m}]$:*

$$\hat{\mu}_\eta^{A,n} := 3^{-\alpha m} \sum_{k\in[1,3^{\alpha m}]} \delta_{\mathrm{A}^m\left(\theta^{\nu_b^n(k)}\eta\right)}.$$

For $\eta \in \mathcal{C}^{\mathbb{N}_0}$ and $T \geq 0$, we set $\hat{\mu}_\eta^{A,n,T} := \hat{\mu}^{A,n}_{\eta|[T,T+2\cdot 3^{10\alpha m}[}$.

Note that block^{n-} and block^{n+} are $\sigma(\xi_z; z \in \mathbb{Z})$-measurable. For $\xi \in \mathcal{C}^{\mathbb{Z}}$ and a (possibly finite) admissible path π, $\mathsf{block}^n(\xi, \pi)$ equals the block in the set $\{\mathsf{block}^{n-}(\xi), \mathsf{block}^{n+}(\xi)\}$ which is visited first by the path π; if π visits neither $\mathsf{block}^{n-}(\xi)$ nor $\mathsf{block}^{n+}(\xi)$, then we set $\mathsf{block}^n(\xi, \pi) := 010 \in \mathcal{C}^{[1,3]}$. The endpoints of $\mathsf{block}^n(\xi, \pi)$ are $b_l^n(\xi, \pi)$ and $b_r^n(\xi, \pi)$. Let $a_{\xi,\pi}^{n,l}$ and $a_{\xi,\pi}^{n,r}$ be the proportion of $k \in [1, 3^{\alpha m}]$ with $\pi_{\nu_b^n(k)} = b_l^n(\xi, \pi)$ and $\pi_{\nu_b^n(k)} = b_r^n(\xi, \pi)$, respectively.

Definition 9.7.4. *For an admissible path $\pi \in \mathbb{Z}^{[0,2\cdot 3^{10\alpha m}[}$, we define*

$$\begin{aligned} \mu_{\xi,\pi}^{A,n} &:= a_{\xi,\pi}^{n,l} P_{b_l^n(\xi,\pi),\xi}\left[\mathrm{A}^n(\chi)\right]^{-1} + a_{\xi,\pi}^{n,r} P_{b_r^n(\xi,\pi),\xi}\left[\mathrm{A}^n(\chi)\right]^{-1}, \\ \varepsilon_{\xi,\pi}^{A,n} &:= \hat{\mu}_{\xi\circ\pi}^{A,n} - \mu_{\xi,\pi}^{A,n}. \end{aligned}$$

For an admissible path $\pi \in \mathbb{Z}^{\mathbb{N}_0}$ and $T \geq 0$, we set $\mu_{\xi,\pi}^{A,n,T} := \mu^{A,n}_{\xi,\pi|[T,T+2\cdot 3^{10\alpha n}[}$.

$\varepsilon_{\xi,\pi}^{A,n}$ measures the difference between the empirical distribution $\hat{\mu}_{\xi\circ\pi}^{A,n}$ and the distribution $\mu_{\xi,\pi}^{A,n}$. By Lemma 5.8 of [22], $\varepsilon_{\xi,S}^{O,n}$ is small with high P-probability provided the stopping times ν_b^n stop correctly. This is used to reconstruct the scenery: $\hat{\mu}_\chi^{A,n}$ can be computed from $\chi|[0, 2\cdot 3^{10\alpha m}[$ and ν_b^n. It is close to $\mu_{\xi,S}^{A,n}$, from which we will extract information about the scenery ξ.

By definition, $\mu_{\xi,\pi}^{A,n}$ and $\hat{\mu}_{\eta}^{A,n}$ are measures on sets of the form $\text{obs} := B \times \text{obs}_2 \times C$ with $\text{obs}_2 := \{w \in \mathcal{C}^k : k \geq c_1 m/2, w_{k-1} \neq w_k, w_j = w_{k-1} \text{ for all } j \in [c_1 m/2 - 1, k-1]\}$ and $B, C \in \{[1,5]^{c_1 m}, \mathbb{N}\}$. We denote by $\Pi_2 : \text{obs} \to \text{obs}_2$ the canonical projections. Furthermore, we introduce the event that an observation $\mathrm{A} \in \text{obs}$ has $\Pi_2(\mathrm{A})$ of length $d \geq c_1 m/2$:

$$E_{\text{block}}^{m,d} := \{\mathrm{A} \in \text{obs} : [\Pi_2(\mathrm{A})]_{d-1} \neq [\Pi_2(\mathrm{A})]_d\}.$$

We order the 2^d elements of $\mathcal{C}^d$ lexicographically and denote them by $v^1, v^2, \cdots, v^{2^d}$. Let $e_{v^k} := (e_{v^k}(i))_{i \in [1,2^d]}$ be defined by $e_{v^k}(i) := \delta_k(i)$; i.e. $\{e_{v^k}; k \in [1, 2^d]\}$ is the canonical basis in $\mathbb{R}^{2^d}$. Let $\{1_{v^k}; k \in [1, 2^d]\}$ be the dual basis, i.e. $1_{v^k}(e_{v^j}) = \delta_k(j)$ for all $j, k \in [1, 2^d]$.

Sometimes it will be convenient to identify a measure λ which is supported on a finite ordered set $\{s_1, s_2, \ldots, s_l\}$ with the vector $(\lambda(\{s_1\}), \lambda(\{s_2\}), \ldots, \lambda(\{s_l\}))$. Similarly, we sometimes identify measures supported on $\mathbb{N}_0$ by one-sided infinite vectors.

Let $w \in \mathcal{C}^d$. For any probability measure λ on $\mathcal{C}^d$ we have $1_w(\lambda) = \lambda(w)$. In particular, if λ gives mass one to w, then $1_w(\lambda) = 1$. We denote by 1 the linear functional defined by $1(\lambda) = \sum_{i=1}^d \lambda_i$. For $j \in \mathbb{N}$, let $e_j^* : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ be defined by $e_j^*((x_i)_{i \in \mathbb{N}}) = x_j$. We define $1_{\geq n^2} := \sum_{j=n^2}^{n^4} e_j^*$. If g_1 and g_2 are two linear functionals we denote by $g_1 \otimes g_2$ their tensor product.

Recall that T_m denotes the first hitting time of $\{-1, m\}$ by the random walk S. For $m \in \mathbb{N}$ we abbreviate

$$\lambda_l^m(\cdot) := P(\{T_m \in \cdot\} \cap \{S_{T_m} = -1\}), \quad \lambda_r^m(\cdot) := P(\{T_m \in \cdot\} \cap \{S_{T_m} = m\}).$$

We define $h : \mathbb{N}_0 \to [1,5]$, $x \mapsto x \wedge 5$. Then

$$\lambda_l^m h^{-1}(\cdot) = P(\{T_m \wedge 5 \in \cdot\} \cap \{S_{T_m} = -1\}).$$

The measures $\lambda_l^m h^{-1}, \lambda_r^m h^{-1}$ are supported on the set $\{1,2,3,4,5\}$. Hence we can identify them with vectors in $\mathbb{R}_+^5$.

Definition 9.7.5. *We define vectors in $\mathbb{R}_+^5$:*

$$\begin{aligned}
\vec{x}_1 &:= (p, pq, pq^2, pq^3, pq^4),\\
\vec{x}_2 &:= \lambda_r^2 h^{-1} = (0, p^2, 2p^2q, p^4 + 3p^2q^2, \lambda_r^2([5,\infty[),\\
\vec{x}_3 &:= (0, 0, p^3, 3p^3q, p^5 + 6p^3q^2),\\
\vec{x}_4 &:= \lambda_r^4 h^{-1} = (0, 0, 0, p^4, \lambda_r^4([5,\infty[),\\
\vec{x}_5 &:= (0, 0, 0, 0, 1).
\end{aligned}$$

Clearly, $\{\vec{x}_i\}_{i \in [1,5]}$ is a basis of $\mathbb{R}^5$. We denote by $\{\vec{x}_i^*\}_{i \in [1,5]}$ the corresponding dual basis.

Definition 9.7.6. *We call a function $f : (\mathbb{R}^5)^{\otimes m} \to \mathbb{R}$* positive *if $f(\otimes_{k=1}^m \vec{x}_{n_k}) \geq 0$ for all $n_1, n_2, \ldots, n_m \in \{1,2,3,4,5\}$.*

The following theorem gives sufficient conditions for a word to be contained in the scenery ξ around block^n. Part (1a) is a criterion to find words in $\xi|[b_l^n - 3^{3m}, b_r^n + 3^{3m}]$. Parts (1b) and (1c) will allow us to reconstruct the words immediately to the left and to the right of block^n. Recall the definitions of the events $B_{\text{only block}}^n$ and $B_{\text{short block}}^n$ from Section 9.5.1.

Theorem 9.7.1. *There exists $c_{11} > 0$ such that for all $n \geq c_{11}$, $d \in [c_1 m/2, c_1 m]$, $T \in \{H_l^n, H_r^n\}$ and $w \in \mathcal{C}^d$ with $w_{d-1} \neq w_d$ the following holds whenever the event $E^{n,T}_{\nu_b \text{ ok}} \cap B^n_{\text{only block}} \cap \Theta^{-T}[B^n_{\text{short block}}]$ holds:*

1. *Case $q \neq 0$:*

   (a) *If there exist positive linear functionals g_1 and g_3 on $(\mathbb{R}^5)^{\otimes c_1 m}$ such that*

   $$\begin{aligned}(g_1 \otimes 1_w \otimes g_3)(\hat{\mu}^{O,n,T}_{\xi \circ S}[\cdot \cap E^{m,d}_{\text{block}}]) &> 1, && (9.7.1)\\ (g_1 \otimes 1 \otimes g_3)(\hat{\mu}^{O,n,T}_{\xi \circ S}[\cdot \cap E^{m,d-1}_{\text{block}}]) &\leq 1/(5m^2), \text{and} && (9.7.2)\\ \|g_1 \otimes g_3\|_2 \cdot \|\varepsilon^{O,n,T}_{\xi,S}\|_1^{1/2} &\leq 1/(2m^2), && (9.7.3)\end{aligned}$$

   then $w \preceq \xi|\,[b_l^n - 3^{3m}, b_r^n + 3^{3m}]$.

   (b) *If there exists a positive linear functional g_1 on $(\mathbb{R}^5)^{\otimes c_1 m}$ such that*

   $$\begin{aligned}(g_1 \otimes 1_w \otimes 1_{\geq n^2})(\hat{\mu}^{L,n,T}_{\xi \circ S}[\cdot \cap E^{m,d}_{\text{block}}]) &> 1, && (9.7.4)\\ (g_1 \otimes 1 \otimes 1_{\geq n^2})(\hat{\mu}^{L,n,T}_{\xi \circ S}[\cdot \cap E^{m,d-1}_{\text{block}}]) &\leq 1/(5m^2), \text{and} && (9.7.5)\\ \|g_1 \otimes 1_{\geq n^2}\|_2 \cdot \|\varepsilon^{L,n,T}_{\xi,S}\|_1^{1/2} &\leq 1/(2m^2), && (9.7.6)\end{aligned}$$

   then $w\mathsf{block}^n \preceq \xi|\,[b_l^n - 3^{3m}, b_r^n + 3^{3m}]$; here $w\mathsf{block}^n$ denotes the concatenation of w and block^n.

   (c) *If there exists a positive linear functional g_3 on $(\mathbb{R}^5)^{\otimes c_1 m}$ such that*

   $$\begin{aligned}(1_{\geq n^2} \otimes 1_w \otimes g_3)(\hat{\mu}^{R,n,T}_{\xi \circ S}[\cdot \cap E^{m,d}_{\text{block}}]) &> 1, && (9.7.7)\\ (1_{\geq n^2} \otimes 1 \otimes g_3)(\hat{\mu}^{R,n,T}_{\xi \circ S}[\cdot \cap E^{m,d-1}_{\text{block}}]) &\leq 1/(5m^2), \text{and} && (9.7.8)\\ \|1_{\geq n^2} \otimes g_3\|_2 \cdot \|\varepsilon^{R,n,T}_{\xi,S}\|_1^{1/2} &\leq 1/(2m^2), && (9.7.9)\end{aligned}$$

   then $\mathsf{block}^n w \preceq \xi|\,[b_l^n - 3^{3m}, b_r^n + 3^{3m}]$.

2. *Case $q = 0$: Replace $d-1$ by $d-2$ in (9.7.2), (9.7.5), and (9.7.8).*

Proof. Part (1a) is an immediate consequence of Theorem 4.1 of [22]. Note that on the event $E^{n,T}_{\nu_b \text{ ok}}$ we have $S_{\nu_b^{n,T}(k)} \in \partial\mathsf{block}^n$, whereas in Theorem 4.1 of [22] the stopping times τ_k are only assumed to satisfy the weaker condition $|S_{\tau_k}| \leq 3^n$.

Parts (1b) and (1c) can be proved using essentially the same arguments as in the proof of Theorem 4.1 of [22]; one replaces g_1 or g_3 in that proof by the functional $1_{\geq n^2}$. Note that on the event $E^{n,T}_{\nu_b \text{ ok}} \cap B^n_{\text{only block}} \cap \Theta^{-T}[B^n_{\text{short block}}]$ blocks of length $\geq n^2$ in $\chi|[T, T + 3^{\lfloor n^{0.3} \rfloor}[$ must be generated on block^n. □

9.7.2 Definition of SmallAlg^n

Theorem 9.7.1 is used to define sets of words which will be used to assemble a larger piece of ξ. Recall $m = \lfloor n^{0.2} \rfloor$.

Definition 9.7.7. *Let $c_7 > 0$ be chosen as in Section 9.2.*

1. *Case $q \neq 0$:*

(a) *We define* $\mathsf{OutsideWords}^n(\eta)$ *to be the set of all* $w \in \mathcal{C}^d$, $d \in [c_1 m/2, c_1 m]$ *such that there exist positive linear functionals* g_1 *and* g_3 *on* $(\mathbb{R}^5)^{\otimes c_1 m}$ *with*

$$\begin{aligned}(g_1 \otimes 1_w \otimes g_3)\,(\hat{\mu}_\eta^{O,n}[\cdot \cap E_{\text{block}}^{m,d}]) &> 1, && (9.7.10)\\ (g_1 \otimes 1 \otimes g_3)(\hat{\mu}_\eta^{O,n}[\cdot \cap E_{\text{block}}^{m,d-1}]) &\le 1/(5m^2), \text{ and} && (9.7.11)\\ \|g_1 \otimes g_3\|_2 &\le e^{c_7 m}. && (9.7.12)\end{aligned}$$

(b) *We define* $\mathsf{LeftWords}^n(\eta)$ *to be the set of all* $w \in \mathcal{C}^d$, $d \in [c_1 m/2, c_1 m]$ *such that there exists a positive linear functional* g_1 *on* $(\mathbb{R}^5)^{\otimes c_1 m}$ *with*

$$\begin{aligned}(g_1 \otimes 1_w \otimes 1_{\ge n^2})(\hat{\mu}_\eta^{L,n}[\cdot \cap E_{\text{block}}^{m,d}]) &> 1, && (9.7.13)\\ (g_1 \otimes 1 \otimes 1_{\ge n^2})(\hat{\mu}_\eta^{L,n}[\cdot \cap E_{\text{block}}^{m,d-1}]) &\le 1/(5m^2), \text{ and} && (9.7.14)\\ \|g_1 \otimes 1_{\ge n^2}\|_2 &\le e^{c_7 m}. && (9.7.15)\end{aligned}$$

(c) *We define* $\mathsf{RightWords}^n(\eta)$ *to be the set of all* $w \in \mathcal{C}^d$, $d \in [c_1 m/2, c_1 m]$ *such that there exists a positive linear functional* g_3 *on* $(\mathbb{R}^5)^{\otimes c_1 m}$ *with*

$$\begin{aligned}(1_{\ge n^2} \otimes 1_w \otimes g_3)(\hat{\mu}_\eta^{R,n}[\cdot \cap E_{\text{block}}^{m,d}]) &> 1, && (9.7.16)\\ (1_{\ge n^2} \otimes 1 \otimes g_3)(\hat{\mu}_\eta^{R,n}[\cdot \cap E_{\text{block}}^{m,d-1}]) &\le 1/(5m^2), \text{ and} && (9.7.17)\\ \|1_{\ge n^2} \otimes g_3\|_2 &\le e^{c_7 m}. && (9.7.18)\end{aligned}$$

2. *Case* $q = 0$*: Replace* $d-1$ *by* $d-2$ *in (9.7.11), (9.7.14), and (9.7.17)*

Finally, we define $\mathsf{Words}^n(\eta) := \mathsf{OutsideWords}^n(\eta) \cup \mathsf{LeftWords}^n(\eta) \cup \mathsf{RightWords}^n(\eta)$.

The algorithm $\mathrm{SmallAlg}^n$ takes as argument $\eta \in \mathcal{C}^{[0,3\lfloor n^{0.3}\rfloor[}$. $\mathrm{SmallAlg}^n(\eta)$ should contain $\mathsf{estimated\text{-}block}^n(\eta)$ in the middle and all subpieces of length $c_1 m/2$ which do not intersect $\mathsf{estimated\text{-}block}^n(\eta)$ should occur in an element of $\mathsf{Words}^n(\eta)$.

Definition 9.7.8. *We define* $\mathsf{Output}^n(\eta) :=$

$$\left\{\begin{array}{l} w \in \mathcal{C}^{[-3\cdot 3^m, 3\cdot 3^m]} : \text{There exists a unique interval } J \subseteq [-3\cdot 3^m, 3\cdot 3^m] \text{ such that}\\ w|J = \mathsf{estimated\text{-}block}^n(\eta) \text{ and for all intervals } I \subseteq [-3\cdot 3^m, 3\cdot 3^m] \setminus J \text{ with}\\ |I| = c_1 m/2 \text{ there exists } w' \in \mathsf{Words}^n(\eta) \text{ such that } w|I \sqsubseteq w' \end{array}\right\}.$$

Definition 9.7.9. *We define* $\mathrm{SmallAlg}^n : \mathcal{C}^{[0,3\lfloor n^{0.3}\rfloor[} \to \mathcal{C}^{[-3\cdot 3^m, 3\cdot 3^m]}$ *as follows:*

If $\mathsf{Output}^n(\eta) \neq \varnothing$*, then we define* $\mathrm{SmallAlg}^n(\eta)$ *to be its lexicographically smallest element. Otherwise we set* $\mathrm{SmallAlg}^n(\eta) := (1)_{[-3\cdot 3^m, 3\cdot 3^m]}$.

9.7.3 Proof of Theorem 9.3.2

We need some definitions. For a word $w \in \mathcal{C}^{[1,d]}$ we define $w^{\leftrightarrow} := (w_k^{\leftrightarrow})_{k\in[1,d]}$ by $w_k^{\leftrightarrow} := w_{d-k+1}$, $k \in [1,d]$, i.e. $w^{\leftrightarrow}$ is obtained by reading w from right to left.

Definition 9.7.10. *We define*

$$\mathsf{InitialPiece}^n(\eta) := \left\{\begin{array}{l} w_l \mathsf{estimated\text{-}block}^n(\eta) w_r \ : \ w_l \in \mathsf{LeftWords}^n(\eta),\\ w_r \in \mathsf{RightWords}^n(\eta), \text{ and } w_l \neq (w_r)^{\leftrightarrow} \end{array}\right\}.$$

The set $\mathsf{InitialPiece}^n(\eta)$ contains concatenations of $\mathsf{estimated\text{-}block}^n(\eta)$ with a word from $\mathsf{LeftWords}^n(\eta)$ to its left and a word from $\mathsf{RightWords}^n(\eta)$ to its right.

Recall the definitions of H_l and H_r from (9.3.3). In the following, let $T \in \{H_l^n, H_r^n\}$. Recall $\chi^{n,T} = \chi|[T, T + 3^{\lfloor n^{0.3} \rfloor}[$.

Definition 9.7.11. *We define*

$$\begin{aligned} E^{n,T}_{\text{ini ok}} \;:=\; & \{\mathsf{estimated\text{-}block}^n(\chi^{n,T}) = \mathsf{block}^n\} \cap \{\mathsf{block}^n \sqsubseteq \xi|[-3^n/3, 3^n/3]\} \cap \\ & \left\{\exists \psi \in \mathsf{InitialPiece}^n(\chi^{n,T}) \text{ with } \psi \sqsubseteq \xi|[b_l^n - c_1 m, b_r^n + c_1 m]\right\}. \end{aligned}$$

Definition 9.7.12. *We define* $E^{n,T}_{\text{Words ok}} := E^{n,T}_{\text{only xi}} \cap E^{n,T}_{\text{all words}}$ *with*

$$\begin{aligned} E^{n,T}_{\text{only xi}} \;:=\; & \left\{\text{If } w \in \mathsf{Words}^n(\chi^{n,T}), \text{ then } w \preceq \xi|\left[b_l^n - 3^{3m}, b_r^n + 3^{3m}\right]\right\}, \\ E^{n,T}_{\text{all words}} \;:=\; & \left\{\begin{array}{l}\text{There exists a unique interval } J \subseteq [-5\cdot 3^m, 5\cdot 3^m] \text{ such} \\ \text{that } \mathsf{block}^n = \xi|J \text{ and for all } w \preceq \xi|\,[[-5\cdot 3^m, 5\cdot 3^m] \setminus J] \\ \text{and } |w| = c_1 m/2, \text{ then } \exists w' \in \mathsf{Words}^n(\chi^{n,T}) \text{ with } w \sqsubseteq w'\end{array}\right\}. \end{aligned}$$

Definition 9.7.13. *For $z \in \mathbb{Z}$ and $k \in \mathbb{N}$ we define $w_{z,k,\rightarrow} := \xi|[z, z+k[$ to be the word of length k starting at z, and we denote by $w_{z,k,\leftarrow}$ the word obtained by reading $w_{z,k,\rightarrow}$ from right to left. We define*

$$B^n_{\text{unique fit}} := \left\{\begin{array}{l}\forall z_1, z_2 \in [b_l^n - 3^{3m}, b_r^n + 3^{3m}] \setminus [b_l^n, b_r^n] \text{ and } \forall i_1, i_2 \in \{\leftarrow, \rightarrow\} \\ \text{with } (z_1, i_1) \neq (z_2, i_2) \text{ we have } w_{z_1,i_1,c_1 m/4} \neq w_{z_2,i_2,c_1 m/4}\end{array}\right\}.$$

Lemma 9.7.1. *There exists c_{57} such that for all $n \geq c_{57}$ and all $T \in \{H_l^n, H_r^n\}$ the following inclusion holds:*

$$E^{n,T}_{\nu_b \text{ ok}} \cap E^{n,T}_{\text{ini ok}} \cap E^{n,T}_{\text{Words ok}} \cap B^n_{\text{unique fit}} \subseteq \Theta^{-T} E^n_{\text{recon Small}}.$$

Proof. The proof is very similar to the proof of Lemma 5.2 of [22]. Roughly speaking, one argues as follows: Suppose the events $E^{n,T}_{\nu_b \text{ ok}}$, $E^{n,T}_{\text{ini ok}}$, $E^{n,T}_{\text{Words ok}}$, and $B^n_{\text{unique fit}}$ hold. Because of $E^{n,T}_{\text{ini ok}}$, $\mathsf{InitialPiece}^n(\chi^{n,T}) \neq \varnothing$. Let $\psi \in \mathsf{InitialPiece}^n(\chi^{n,T})$. Then, there exist $w_l \in \mathsf{LeftWords}^n(\chi^{n,T})$ and $w_r \in \mathsf{RightWords}^n(\chi^{n,T})$ such that

$$\psi = w_l \mathsf{estimated\text{-}block}^n(\chi^{n,T}) w_r = w_l \mathsf{block}^n w_r \sqsubseteq \xi|[b_l^n - c_1 m, b_r^n + c_1 m] \sqsubseteq \xi|[-3^n, 3^n]$$

for all n sufficiently large. We use ψ as initial piece for our reconstruction. Since $E^{n,T}_{\text{Words ok}}$ holds, there exists a word $w \in \mathsf{Words}^n(\chi^{n,T})$ with the property that the right-most $c_1 m/2 - 1$ letters of w agree with the left-most $c_1 m/2 - 1$ letters of ψ. This way we can extend ψ by at least one letter and this extension is a piece of the scenery ξ around block^n. Proceeding like this, we find that $\mathsf{Output}^n(\chi^{n,T}) \neq \varnothing$ and also that every $w \in \mathsf{Output}^n(\chi^{n,T})$ satisfies $w \preceq \xi|[j - 3\cdot 3^m, j + 3\cdot 3^m]$ with $j = b_l^n$ or $j = b_r^n$ depending on whether $T = H_l^n$ or $T = H_r^n$. Consequently, the event $\Theta^{-T}[E^n_{\text{recon Small}}]$ holds. For details, we refer the reader to the similar proof of Lemma 5.2 in [22]. □

Proof of Theorem 9.3.2. By Lemma 9.7.1,

$$\begin{aligned} P([\Theta^{-T} E^n_{\text{recon Small}}]^c) \;\le\; & P([E^{n,T}_{\nu_b \text{ ok}}]^c) + P(E^{n,T}_{\nu_b \text{ ok}} \setminus E^{n,T}_{\text{ini ok}}) \\ & + P(E^{n,T}_{\nu_b \text{ ok}} \setminus E^{n,T}_{\text{Words ok}}) + P([B^n_{\text{unique fit}}]^c). \end{aligned}$$

By Proposition 9.3.2, $P([E^{n,T}_{\nu_b \text{ ok}}]^c) \le e^{-c_{11} n^{0.2}}$ for all $n \ge c_{10}$. Using the definition of the event $E^{n,T}_{\text{ini ok}}$, we see that

$$P(E^{n,T}_{\nu_b \text{ ok}} \setminus E^{n,T}_{\text{ini ok}}) \le P(\mathsf{estimated\text{-}block}^n(\chi^{n,T}) \neq \mathsf{block}^n) + P(\mathsf{block}^n \not\sqsubseteq \xi | [-3^n, 3^n]) +$$
$$P\left[E^{n,T}_{\nu_b \text{ ok}} \setminus \begin{Bmatrix} \exists w_l \in \mathsf{LeftWords}^n(\chi^{n,T}) \; \exists w_r \in \mathsf{RightWords}^n(\chi^{n,T}) \text{ with } w_l \neq (w_r)^{\leftrightarrow} \\ \text{such that } w_l \mathsf{block}^n w_r \sqsubseteq \xi | [b^n_l - c_1 m, b^n_r + c_1 m] \end{Bmatrix}\right].$$

Proposition 9.3.1 states that $P(\mathsf{estimated\text{-}block}^n(\chi^{n,T}) \neq \mathsf{block}^n) \le c_8 e^{-c_9 n^{0.3}}$ for all $n \ge c_7$. By Theorem 9.3.1, $P(\mathsf{block}^n \not\sqsubseteq \xi | [-3^n/3, 3^n/3]) \le c_{12} e^{-c_{13} n^{0.3}}$ for all $n \ge c_{11}$. Using similar arguments as in the proof of Lemma 5.1 of [22] and Section 5.4 of [22], we get that

$$P\left[E^{n,T}_{\nu_b \text{ ok}} \setminus \begin{Bmatrix} \exists w_l \in \mathsf{LeftWords}^n(\chi^{n,T}) \; \exists w_r \in \mathsf{RightWords}^n(\chi^{n,T}) \text{ with} \\ w_l \neq (w_r)^{\leftrightarrow} \text{ such that } w_l \mathsf{block}^n w_r \sqsubseteq \xi | [b^n_l - c_1 m, b^n_r + c_1 m] \end{Bmatrix}\right] \le e^{-c_{58} n^{0.2}}$$

and

$$P(E^{n,T}_{\nu_b \text{ ok}} \setminus E^{n,T}_{\text{Words ok}}) \le e^{-c_{58} n^{0.2}}$$

for all $n \ge c_{59}$ with constants $c_{58}, c_{59} > 0$. The only difference in the two proofs is that in our situation, SmallAlg^n reconstructs around $\mathsf{estimated\text{-}block}^n$, whereas in [22], the reconstruction is done around a typical piece of scenery close to the origin.

Essentially the same arguments as in the proof of Lemma 5.13 of [22] show that

$$P([B^n_{\text{unique fit}}]^c) \le e^{-c_{60} n^{0.2}}$$

for all n sufficiently large. Combining all these estimates, we conclude

$$P([\Theta^{-T} E^n_{\text{recon Small}}]^c) \le e^{-2 c_{61} n^{0.2}}$$

for all $n \ge c_{62}$ with constants $c_{61}, c_{62} > 0$. In order to make statements about the conditional probability $P_\xi([\Theta^{-T} E^n_{\text{recon Small}}]^c)$, we need the following elementary lemma:

Lemma 9.7.2 (see e.g. [19], Lemma 4.6). *Let E be an event and let $r \ge 0$. If $P(E) \le r^2$, then $P[P(E|\xi) > r] \le r$.*

Applying this lemma, we obtain

$$P\big[\xi \in \mathcal{C}^{\mathbb{Z}} : P_\xi([\Theta^{-T} E^n_{\text{recon Small}}]^c) > e^{-c_{61} n^{0.2}}\big] \le e^{-c_{61} n^{0.2}}.$$

This completes the proof of Theorem 9.3.2. □

9.8 More stopping

In this section, we prove Proposition 9.3.3.

9.8.1 Definitions of events

We collect in alphabetical order definitions of events, which will be needed below.

Definition 9.8.1. *We define*

$$B^n_{\text{block often}} := \{S|[\nu^n(0), 3^{10\alpha n}/3[\text{ visits the set } \partial\mathsf{block}^n \text{ at least } 3^{3\alpha n-1} \text{ times}\}.$$

Definition 9.8.2. *Let*

$$\mathbb{S}_n(\xi,\chi) := \{t \in \mathbb{N}_0 : \text{SmallAlg}^n(\chi|[t, t+3^{\lfloor n^{0.3}\rfloor}[) \preceq \xi|[b^n_l - 3\cdot 3^{\lfloor n^{0.2}\rfloor}, b^n_r + 3\cdot 3^{\lfloor n^{0.2}\rfloor}]\}.$$

For $k \geq 1$, we denote by $\gamma_n(k-1)$ the kth hitting time which is $\geq \nu^n(0)$ of the set $\partial\mathsf{block}^n$ by the random walk. We define

$$B^n_{\text{when back recog}} := \{3^{-2\alpha n}|\{k \in [1, 3^{2\alpha n}] : \gamma_n(2k\cdot 3^{3n}) \in \mathbb{S}_n(\xi,\chi)| > 1/4\}.$$

Definition 9.8.3. *We define* $E^n_{\text{no error }\nu} := \{\forall k \geq 1 : \text{if } \nu^n(k) < 3^{10\alpha n}, \text{ then } |S_{\nu^n(k)}| \leq 3^n\}$.

9.8.2 Proof of Proposition 9.3.3

Lemma 9.8.1. *Recall the definition of $\mathbb{T}_n$ (9.3.5). For all $n \in \mathbb{N}$ the following inclusion holds:*

$$E^n_{\text{no error }\nu} \cap \{\mathbb{S}_n \cap [\nu^n(0), 3^{10\alpha n} - 3^{\lfloor n^{0.3}\rfloor}[\subseteq \mathbb{T}_n\} \cap B^n_{\text{block often}} \cap B^n_{\text{when back recog}} \subseteq E^{n,\nu}_{\text{stop}}.$$

Proof. Suppose the events $E^n_{\text{no error }\nu}$, $\{\mathbb{S}_n \cap [\nu^n(0), 3^{10\alpha n} - 3^{\lfloor n^{0.3}\rfloor}[\subseteq \mathbb{T}_n\}$, $B^n_{\text{block often}}$, and $B^n_{\text{when back recog}}$ hold. Because of $B^n_{\text{block often}}$, $S|[\nu^n(0), 3^{10\alpha n}/3[$ visits $\partial\mathsf{block}^n$ at least $3^{3\alpha n}$ times. Thus, $\gamma_n(2k\cdot 2^{3n}) \leq 3^{10\alpha n}/3 \leq 3^{10\alpha n} - 3^{\lfloor n^{0.3}\rfloor}$ for all $k \in [1, 3^{2\alpha n}]$; recall the definition of $\gamma_n(\cdot)$ from Definition 9.8.2 and note that $2(k+1)3^{3n} \leq 3^{3\alpha n}$ because $\alpha \geq 5$. Since $B^n_{\text{when back recog}}$ holds, at least $3^{2\alpha n}/4 \geq 2\cdot 3^{3n+\alpha n}$ of the times $\gamma_n(2k\cdot 3^{3n})$, $k \in [1, 3^{2\alpha n}]$, belong to $\mathbb{S}_n$. Using that $\mathbb{S}_n \cap [\nu^n(0), 3^{10\alpha n} - 3^{\lfloor n^{0.3}\rfloor}[\subseteq \mathbb{T}_n$, we conclude that $|\mathbb{T}_n| \geq 2\cdot 3^{3n+\alpha n}$. Thus $\nu^n(k) < 3^{10\alpha n}$ for all $k \in [1, 3^{\alpha n}]$. Because of $E^n_{\text{no error }\nu}$, we have $|S_{\nu^n(k)}| \leq 3^n$ for all $k \in [1, 3^{\alpha n}]$. Furthermore, $|\nu^n(k) - \nu^n(j)| \geq 2\cdot 3^{3n}$ for $k, j \in [1, 3^{\alpha n}]$, $k \neq j$, by the definition of the $\nu^n(k)$'s, and we have shown that $E^{n,\nu}_{\text{stop}}$ holds. □

Lemma 9.8.2. *Recall the definition of $B^n_{\text{size block}}$ from Definition 9.5.8. There exists c_{63} such that for all $n \geq c_{63}$ the following inclusion holds:*

$$E^n_{\nu(0)\text{ ok}} \cap \Theta^{-\nu^n(0)}[E^n_{\text{recon Small}}] \cap B^n_{\text{size block}} \subseteq \{\mathbb{S}_n \cap [\nu^n(0), 3^{10\alpha n} - 3^{\lfloor n^{0.3}\rfloor}[\subseteq \mathbb{T}_n\}.$$

Proof. Suppose $E^n_{\nu(0)\text{ ok}}$, $\Theta^{-\nu^n(0)}[E^n_{\text{recon Small}}]$, and $B^n_{\text{size block}}$ hold. For $t \in \mathbb{N}_0$, we abbreviate $\psi_{n,t} := \text{SmallAlg}^n(\chi|[t, t+3^{\lfloor n^{0.3}\rfloor}[)$. Let $t \in \mathbb{S}_n \cap [\nu^n(0), 3^{10\alpha n} - 3^{\lfloor n^{0.3}\rfloor}[$. Then, by the definition of $\mathbb{S}_n$, $\psi_{n,t} \preceq \xi|[b^n_l - 3\cdot 3^{\lfloor n^{0.2}\rfloor}, b^n_r + 3\cdot 3^{\lfloor n^{0.2}\rfloor}]$ and $|\psi_{n,t}| = 6\cdot 3^{\lfloor n^{0.2}\rfloor} + 1$. Since $E^n_{\nu(0)\text{ ok}} \cap \Theta^{-\nu^n(0)}[E^n_{\text{recon Small}}]$ holds, $\psi_{n,\nu^n(0)} \preceq \xi|[s-3\cdot 3^{\lfloor n^{0.2}\rfloor}, s+3\cdot 3^{\lfloor n^{0.2}\rfloor}]$ with $s \in \partial\mathsf{block}^n$. Because of $B^n_{\text{size block}}$, we have $|b^n_i - s| < 2n$ for $i \in \{l, r\}$. Consequently, $\psi_{n,t}$ and $\psi_{n,\nu^n(0)}$ agree on a subpiece of length $\geq 6\cdot 3^{\lfloor n^{0.2}\rfloor} + 1 - 2n \geq 2\cdot 3^{\lfloor n^{0.2}\rfloor} + 1$ for all n sufficiently large. Consequently, by the definition of $\mathbb{T}_n$ (9.3.5), we have $t \in \mathbb{T}_n$. □

Proof of Proposition 9.3.3. Combining Lemmas 9.8.1 and 9.8.2, we obtain for all $n \geq c_{63}$

$$\begin{aligned} & \big[E^n_{\nu(0)\text{ ok}} \cap \Theta^{-\nu^n(0)}[E^n_{\text{recon Small}}] \cap \{\xi \in \Xi^n\}\big] \setminus E^{n,\nu}_{\text{stop}} \\ \subseteq\ & \Big[\big[E^n_{\nu(0)\text{ ok}} \cap \Theta^{-\nu^n(0)}[E^n_{\text{recon Small}}] \cap B^n_{\text{size block}}\big] \setminus E^n_{\text{no error }\nu}\Big] \cup [B^n_{\text{block often}}]^c \\ & \cup [B^n_{\text{size block}}]^c \cup [\{\xi \in \Xi^n\} \setminus B^n_{\text{when back recog}}]. \end{aligned}$$

The claim follows from Lemmas 9.8.3, 9.8.4, 9.5.11, and 9.8.5. □

9.8.3 Probabilistic estimates

Lemma 9.8.3. *There exists c_{64} such that for all $n \geq c_{64}$*

$$P\big(\big[E^n_{\nu(0)\text{ ok}} \cap \Theta^{-\nu^n(0)}[E^n_{\text{recon Small}}] \cap B^n_{\text{size block}}\big] \setminus E^n_{\text{no error }\nu}\big) \leq e^{-n}.$$

Proof. For $i \geq 1$, we denote by v_i the ith time the random walker visits a point in the set $\mathbb{Z} \setminus [-3^n + 3^{\lfloor n^{0.3}\rfloor}, 3^n - 3^{\lfloor n^{0.3}\rfloor}]$, and we set

$$B^{n,i}_{\text{wrong}} := \left\{ \begin{matrix} \exists w \in \mathcal{C}^{[-3^{\lfloor n^{0.2}\rfloor}, 3^{\lfloor n^{0.2}\rfloor}]} \text{ such that } w \preceq \xi[b^n_l - 3\cdot 3^{\lfloor n^{0.2}\rfloor}, b^n_r + 3\cdot 3^{\lfloor n^{0.2}\rfloor}] \\ \text{and } w \preceq \text{SmallAlg}^n(\chi[v_i, v_i + 3^{\lfloor n^{0.3}\rfloor}[) \end{matrix} \right\}.$$

Suppose the event $\big[E^n_{\nu(0)\text{ ok}} \cap \Theta^{-\nu^n(0)}[E^n_{\text{recon Small}}] \cap B^n_{\text{size size}}\big] \setminus E^n_{\text{no error }\nu}$ holds. We claim that for some $i \in [1, 3^{10\alpha n}]$, the event $B^{n,i}_{\text{wrong}}$ holds as well. Because of $[E^n_{\text{no error }\nu}]^c$, there exists $k \geq 1$ such that $\nu^n(k) < 3^{10\alpha n}$ and $|S_{\nu^n(k)}| > 3^n$. Since the random walk jumps in each step a distance of 0 or 1, $|S_{\nu^n(k)-3^{\lfloor n^{0.3}\rfloor}}| > 3^n - 3^{\lfloor n^{0.3}\rfloor}$. Thus, $\nu^n(k) - 3^{\lfloor n^{0.3}\rfloor} = v_i$ for some $i \leq 3^{10\alpha n}$. Using the definition of $\nu^n(k)$, we see that $v_i \in \mathbb{T}_n$. Consequently, there exists $w \in \mathcal{C}^{[-3^{\lfloor n^{0.2}\rfloor}, 3^{\lfloor n^{0.2}\rfloor}]}$ such that $w \preceq \text{SmallAlg}^n(\chi[v_i, v_i + 3^{\lfloor n^{0.3}\rfloor}[)$ and $w \preceq \psi_n$ with $\psi_n := \text{SmallAlg}^n(\chi[\nu^n(0), \nu^n(0) + 3^{\lfloor n^{0.3}\rfloor}[)$. Since the event $E^n_{\nu(0)\text{ ok}} \cap \Theta^{-\nu^n(0)}[E^n_{\text{recon Small}}]$ holds, we have $\psi_n \preceq \xi[s - 3\cdot 3^{\lfloor n^{0.2}\rfloor}, s + 3\cdot 3^{\lfloor n^{0.2}\rfloor}]$ with $s \in \partial\mathsf{block}^n$. Consequently, $\psi_n \preceq \xi[b^n_l - 3\cdot 3^{\lfloor n^{0.2}\rfloor}, b^n_r + 3\cdot 3^{\lfloor n^{0.2}\rfloor}]$. Hence the event $B^{n,i}_{\text{wrong}}$ holds, and we have shown the following inclusion:

$$\big[E^n_{\nu(0)\text{ ok}} \cap \Theta^{-\nu^n(0)}[E^n_{\text{recon Small}}] \cap B^n_{\text{size block}}\big] \setminus E^n_{\text{no error }\nu} \subseteq \bigcup_{i\in[1,3^{10\alpha n}]} \big[E^n_{\nu(0)\text{ ok}} \cap B^{n,i}_{\text{wrong}}\big] \quad (9.8.1)$$

Let $i \in [1, 3^{10\alpha n}]$. By the definition of v_i, we have $S([v_i, v_i + 3^{\lfloor n^{0.3}\rfloor}[) \subseteq \mathbb{Z} \setminus [-3^n + 2\cdot 3^{\lfloor n^{0.3}\rfloor}, 3^n - 2\cdot 3^{\lfloor n^{0.3}\rfloor}]$. Because of $E^n_{\nu(0)\text{ ok}}$, $|b^n_l|, |b^n_r| \leq 3^n/3$. Hence $[b^n_l - 3\cdot 3^{\lfloor n^{0.2}\rfloor}, b^n_r + 3\cdot 3^{\lfloor n^{0.2}\rfloor}] \cap S([v_i, v_i + 3^{\lfloor n^{0.3}\rfloor}[) = \varnothing$ for all n sufficiently large. Consequently, because the scenery is independently colored, $\xi[b^n_l - 3\cdot 3^{\lfloor n^{0.2}\rfloor}, b^n_r + 3\cdot 3^{\lfloor n^{0.2}\rfloor}]$ and $\text{SmallAlg}^n(\chi[v_i, v_i + 3^{\lfloor n^{0.3}\rfloor}[)$ are independent. The probability that two independent, uniformly $\{0,1\}$-colored words in $\mathcal{C}^{[-3^{\lfloor n^{0.2}\rfloor}, 3^{\lfloor n^{0.2}\rfloor}]}$ agree equals $2^{-2\cdot 3^{\lfloor n^{0.2}\rfloor}-1}$. In $\xi[b^n_l - 3\cdot 3^{\lfloor n^{0.2}\rfloor}, b^n_r + 3\cdot 3^{\lfloor n^{0.2}\rfloor}]$ and $\text{SmallAlg}^n(\chi[v_i, v_i + 3^{\lfloor n^{0.3}\rfloor}[)$, there are at most $2\cdot 8\cdot 3^{\lfloor n^{0.2}\rfloor}$ words of length $2\cdot 3^{\lfloor n^{0.2}\rfloor} + 1$ if we count also backward words. Hence there are $\leq 16^2 \cdot 3^{2\lfloor n^{0.2}\rfloor}$ pairs of such words. We conclude

$$P\Big(\bigcup_{i\in[1,3^{10\alpha n}]} \big[E^n_{\nu(0)\text{ ok}} \cap B^{n,i}_{\text{wrong}}\big]\Big) \leq 3^{10\alpha n} \cdot 16^2 \cdot 3^{2\lfloor n^{0.2}\rfloor} \cdot 2^{-2\cdot 3^{\lfloor n^{0.2}\rfloor}-1}$$

which is $\leq e^{-n}$ for all n sufficiently large because $2^{-2\cdot 3^{\lfloor n^{0.2}\rfloor}-1}$ is the leading order term. The claim follows from (9.8.1). □

Lemma 9.8.4. *There exist constants $c_{65}, c_{66}, c_{67} > 0$ such that for all $n \geq c_{65}$*

$$P([B^n_{\text{block often}}]^c) \leq c_{66} e^{-c_{67} n^{0.3}}.$$

Proof. Recall the definition of $B^{n,s,t}_{\text{block enough}}$ from Definition 9.5.2. Let $t := 2 \cdot 3^{10\alpha n-1} - 4 \cdot 3^{3n}$. Clearly, $\lfloor t^{3/10} \rfloor \geq (3^{10\alpha n-1})^{3/10} \geq 3^{3\alpha n-1}$. If $E^n_{\nu(0)\text{ ok}}$ holds, then $\nu^n(0) \leq 2 \cdot 3^{3n}$. Consequently, $\nu^n(0) + t/2 \leq 3^{10\alpha n}$, and we obtain

$$\begin{aligned} & E^n_{\nu(0)\text{ ok}} \cap B^{n,\nu^n(0),t}_{\text{block enough}} \\ \subseteq\ & E^n_{\nu(0)\text{ ok}} \cap \{S | [\nu^n(0), \nu^n(0) + t/2[\text{ visits } \partial\mathsf{block}^n \geq \lfloor t^{3/10} \rfloor \text{ times}\} \\ \subseteq\ & \{S | [\nu^n(0), 3^{10\alpha n}[\text{ visits } \partial\mathsf{block}^n \geq 3^{3\alpha n} \text{ times}\} = B^n_{\text{block often}}. \end{aligned}$$

Hence, by Theorem 9.3.1 and Lemma 9.5.5,

$$\begin{aligned} P([B^n_{\text{block often}}]^c) &\leq P([E^n_{\nu(0)\text{ ok}}]^c) + P([B^{n,\nu^n(0),t}_{\text{block enough}}]^c) \\ &\leq c_{12} e^{-c_{13} n^{0.3}} + c_{32} (3^{10\alpha n-1})^{-1/30}, \end{aligned}$$

and the claim follows. □

Lemma 9.8.5. *There exists c_{68} such that for all $n \geq c_{68}$*

$$P(\{\xi \in \Xi^n\} \setminus B^n_{\text{when back recog}}) \leq [0.9]^{3^{2\alpha n}}.$$

Proof. For $k \geq 1$, we set $Y_k := 1$ if $\gamma_n(2k \cdot 3^{3n}) \in \mathbb{S}_n$ and $Y_k := 0$ otherwise. By the definition of the $\gamma_n(j)$'s, we have $\gamma_n(2(k+1)3^{3n}) - \gamma_n(2k \cdot 3^{3n}) \geq 2 \cdot 3^{3n} > 3^{\lfloor n^{0.3} \rfloor}$. Note that $B^n_{\text{when back recog}} = \{3^{-2\alpha n} \sum_{k=1}^{3^{2\alpha n}} Y_k > 1/4\}$. By the strong Markov property of the random walk, Y_k, $k \geq 1$, are independent, conditioned on ξ.

Suppose $\xi \in \Xi^n$. Since $S_{\gamma_n(k)} \in \{b^n_l, b^n_r\}$, it follows from the strong Markov property of the random walk that $P_\xi(\Theta^{-\gamma_n(2k \cdot 3^{3n})}[E^n_{\text{recon Small}}]^c) \leq e^{-c_{18} n^{0.2}} \leq 1/2$ for all n sufficiently large. If the event $\Theta^{-\gamma_n(2k \cdot 2^{3n})}[E^n_{\text{recon Small}}]$ holds, then $\gamma_n(2k \cdot 3^{3n}) \in \mathbb{S}_n$ and consequently, $Y_k = 1$. Thus, for all n sufficiently large,

$$P_\xi(Y_k = 1) \geq P_\xi(\Theta^{-\gamma_n(2k \cdot 3^{3n})}[E^n_{\text{recon Small}}]) \geq \frac{1}{2}.$$

By the exponential Chebyshev inequality applied to the random variable $\sum_{k=1}^{3^{2\alpha n}} Y_k$, we obtain for $\xi \in \Xi^n$

$$\begin{aligned} P_\xi([B^n_{\text{when back recog}}]^c) &\leq P_\xi(3^{-2\alpha n} \sum_{k=1}^{3^{2\alpha n}} Y_k \leq 1/4) \leq \prod_{k=1}^{3^{2\alpha n}} E_\xi[e^{1/4 - Y_k}] \\ &\leq \Big[\frac{e^{1/4} + e^{-3/4}}{2}\Big]^{3^{2\alpha n}} \leq [0.9]^{3^{2\alpha n}}. \end{aligned}$$

Consequently,

$$P(\{\xi \in \Xi^n\} \setminus B^n_{\text{when back recog}}) = \int_{\{\xi \in \Xi^n\}} P_\xi([B^n_{\text{when back recog}}]^c) dP \leq [0.9]^{3^{2\alpha n}}.$$

□

References

[1] I. Benjamini and H. Kesten. Distinguishing sceneries by observing the scenery along a random walk path. *J. Anal. Math.*, 69:97–135, 1996.

[2] K. Burdzy. Some path properties of iterated Brownian motion. In *Seminar on Stochastic Processes, 1992 (Seattle, WA, 1992)*, volume 33 of *Progr. Probab.*, pages 67–87. Birkhäuser Boston, Boston, MA, 1993.

[3] F. den Hollander and J. E. Steif. Mixing properties of the generalized T, T^{-1}-process. *J. Anal. Math.*, 72:165–202, 1997.

[4] W. Th. F. den Hollander. Mixing properties for random walk in random scenery. *Ann. Probab.*, 16(4):1788–1802, 1988.

[5] R. Durrett. *Probability: Theory and Examples*. Duxbury Press, Second edition, 1996.

[6] William Feller. *An introduction to probability theory and its applications. Vol. I.* John Wiley & Sons Inc., New York, third edition, 1968.

[7] D. Heicklen, C. Hoffman, and D. J. Rudolph. Entropy and dyadic equivalence of random walks on a random scenery. *Adv. Math.*, 156(2):157–179, 2000.

[8] C. D. Howard. Detecting defects in periodic scenery by random walks on $\mathbb{Z}$. *Random Structures Algorithms*, 8(1):59–74, 1996.

[9] C. D. Howard. Orthogonality of measures induced by random walks with scenery. *Combin. Probab. Comput.*, 5(3):247–256, 1996.

[10] C. D. Howard. Distinguishing certain random sceneries on $\mathbb{Z}$ via random walks. *Statist. Probab. Lett.*, 34(2):123–132, 1997.

[11] M. Keane and W. Th. F. den Hollander. Ergodic properties of color records. *Phys. A*, 138(1-2):183–193, 1986.

[12] H. Kesten. Detecting a single defect in a scenery by observing the scenery along a random walk path. In *Itô's stochastic calculus and probability theory*, pages 171–183. Springer, Tokyo, 1996.

[13] H. Kesten. Distinguishing and reconstructing sceneries from observations along random walk paths. In *Microsurveys in discrete probability (Princeton, NJ, 1997)*, pages 75–83. Amer. Math. Soc., Providence, RI, 1998.

[14] D. Levin and Y. Peres. Random walks in stochastic scenery on $\mathbb{Z}$. Preprint, 2002.

[15] D. A. Levin, R. Pemantle, and Y. Peres. A phase transition in random coin tossing. *Ann. Probab.*, 29(4):1637–1669, 2001.

[16] E. Lindenstrauss. Indistinguishable sceneries. *Random Structures Algorithms*, 14(1):71–86, 1999.

[17] M. Löwe and H. Matzinger. Reconstruction of sceneries with correlated colors. Eurandom Report 99-032, accepted by Stochastic Processes and Their Applications, 1999.

[18] M. Löwe and H. Matzinger. Scenery reconstruction in two dimensions with many colors. *Ann. Appl. Probab.*, 12(4):1322–1347, 2002.

[19] M. Löwe, H. Matzinger, and F. Merkl. Reconstructing a multicolor random scenery seen along a random walk path with bounded jumps. Eurandom Report 2001-030. Submitted., 2001.

[20] H. Matzinger. Reconstructing a three-color scenery by observing it along a simple random walk path. *Random Structures Algorithms*, 15(2):196–207, 1999.

[21] H. Matzinger. Reconstructing a 2-color scenery by observing it along a simple random walk path. Eurandom Report 2000-003, 2000.

[22] H. Matzinger and S. W. W. Rolles. Reconstructing a 2-color scenery in polynomial time by observing it along a simple random walk path with holding. Preprint.

[23] H. Matzinger and S. W. W. Rolles. Reconstructing a random scenery observed with random errors along a random walk path. EURANDOM Report 2002-009. Submitted.

[24] Frank Spitzer. *Principles of random walks*. Springer-Verlag, New York, second edition, 1976. Graduate Texts in Mathematics, Vol. 34.

Chapter 10

Markers for error-corrupted observations

Submitted

By Andrew Hart, Heinrich Matzinger

Russel Lyons and Yuval Peres have both posed the question of whether two-color sceneries can be reconstructed when the observations are corrupted by random errors. It has been proved that it is possible to do reconstruction in the case that the observations are contaminated with errors and the scenery has several colors, provided the error probability is small enough. However, the reconstruction problem is more difficult with fewer colors. Although the scenery reconstruction problem for two- color sceneries from error- free observations has been solved, the reconstruction of two-color sceneries from error-corrupted observations remains an open problem. In this paper, we solve one of the two remaining problems needed in order to reconstruct 2-color sceneries when the observations are corrupted with random errors.

Keywords: Scenery reconstruction, scenery distinguishing, large deviations.

2000 MSC: 60K37, 60G50.

10.1 Introduction

We call ξ a scenery if it is a coloring of the integers $\xi : \mathbb{Z} \mapsto \{1, 2, \ldots, C\}$ where C is the number of colors. Let $\{S_t\}_{t\geq 0}$ be a recurrent random walk on the integers. We call $\chi_t := \xi(S_t)$ the observation made of the scenery by the random walk at time t. A realization of the process $\chi := \{\chi_t\}_{t\geq 0}$ is called the "observation".

The **scenery reconstruction problem** can be formulated as follows: If we do not know the scenery ξ but are only given one path realization of χ, can we almost surely recover ξ? In other words, does one path realization of the process $\{\chi_t\}_{t\geq 0}$ determine ξ a.s.? We should point out that it is only possible to reconstruct sceneries up to shift and reflection in general. Thus the scenery reconstruction problem is the problem of trying to reconstruct ξ up to shift and reflection given only one realization of χ. A result of

Lindenstrauss [14] implies that there exists an uncountable number of sceneries which cannot be reconstructed. Fortunately, these unreconstructable sceneries are in a certain sense "untypical". So, in general we take the scenery to be generated by a random process which is independent of the random walk and then show that almost every scenery can be reconstructed a.s. up to shift and reflection from a single realization of χ.

Let $\{\nu_t\}_{t\geq 0}$ be an i.i.d. sequence of Bernoulli random variables which is independent of ξ and S. Here, ν_t represents a possible error in the observation at time t. Let $\tilde{\chi}$ denote the observations χ corrupted by the errors ν. We assume that $\nu_t = 0$ if and only if $\chi_t = \tilde{\chi}_t$. The scenery reconstruction problem with errors can now be formulated as follows: Try to reconstruct ξ a.s. up to shift and reflection when you are only given one realization of $\tilde{\chi}$. In the case that the scenery has many colors and the error probability is small, the problem was solved by Rolles and Matzinger in [25]. In this article, we show how we can construct markers and stopping times telling us when the random walk is back at the markers, despite the errors. The method we use is different from that used to deal with the non-error corrupted case. This solves one of the two remaining problems for scenery reconstruction with 2 colors and errors in the observations.

Scenery reconstruction is closely related to the scenery distinguishing problem. We give a brief account. Let ξ^a and ξ^b be two non-equivalent sceneries which are known to us. Assume that the scenery ξ is either equal to ξ^a or ξ^b but we don't know which. If we are only given one realization of the observation process χ of the scenery ξ by the random walk S, can we almost surely determine if ξ is equal to ξ^a or if it is equal to ξ^b? If so, we say the sceneries ξ^a and ξ^b are distinguishable. Kesten and Benjamini [1] showed that almost every pair of sceneries is distinguishable, even in the two-dimensional case and with only 2 colors. To do this, they took ξ^a to be any non-random scenery and ξ^b to be an i.i.d. scenery with two colors. Earlier, Howard [8] showed that any pair of periodic, non-equivalent sceneries are distinguishable, as well as periodic sceneries with a single defect [7].

The problem of distinguishing two sceneries which differ at only one point is called detecting a single defect in a scenery. Kesten [11] was able to show that one can a.s. detect single defects in the case of four color sceneries. A question Kesten raised concerning the detection of defects in sceneries lead Matzinger [23, 24, 22] to investigate the scenery reconstruction problem.

As with scenery reconstruction, there is a version of the scenery distinguishing problem with observations that are corrupted. Once again, the scenery ξ is either equal to ξ^a or ξ^b, both of which are known to us. However, the observations are now corrupted through an error process $\{\nu_t\}_{t\geq 0}$, which is assumed to be a sequence of i.i.d. Bernoulli random variables with parameter strictly smaller than $1/2$. The observations with errors $\{\tilde{\chi}_t\}_{t\geq 0}$ are such that $\tilde{\chi}_t = \chi_t$ if and only if $\nu_t = 0$. Knowing ξ^a and ξ^b, can we decide a.s. if $\xi = \xi^a$ or if $\xi = \xi^b$ based on one path realization of the process $\tilde{\chi}$? This constitutes the scenery distinguishing problem in the case of error- corrupted observations. The subject of this article is closely related to a random coin tossing problem which was first investigated by Harris and Keane in [6] and later by Levin, Pemantle and Peres in [21]. The **coin tossing problem of Harris and Keane** can be described as follows:

Let $X_1, X_2, \ldots$ denote a sequence of Bernoulli variables where X_k is the result of the k-th coin toss. We consider two ways of doing this.

- The first method is to toss an unbiased coin independently each time. In this case

the variables X_k are a sequence of i.i.d. Bernoulli random variables with parameter 1/2.

- Let $\tau_1, \tau_2, \ldots$ denote a sequence of return times of a random walk to the origin. We toss fair coins independently at all times except at the times τ_k, at which we toss a biased coin with fixed bias ω instead.

The problem investigated by Harris and Keane in [6] and later by Levin, Pemantle and Peres in [21] can now be described as follows: If we are only given one realization of the process $\{X_k\}_{k\geq 0}$, but do not know if it was generated by mechanism 1 or 2, can we determine a.s. from which of the two processes the observation comes? Harris and Keane were able to show that, depending on the finiteness of the moments of the stopping times, we may or may not be able to deduce the method used to generate the observed sequence. Later, Levin, Pemantle and Peres were able to show that there is a phase transition depending on the size of the bias. Furthermore, they were also able to solve the problem in the case where the stopping times halt a random walk at a finite number of points instead of just at the origin.

It is evident that the Harris-Keane coin tossing problem can be viewed as a scenery distinguishing problem with errors. In particular, take ξ^a as the scenery which is everywhere equal to zero, and ξ^b as the scenery which is zero everywhere except at the origin. In the case studied by Levin, Pemantle and Peres [21], set the scenery $\xi^a \equiv 0$ and ξ^b to be zero everywhere except at a finite number of points.

There is an excellent overview of scenery reconstruction and scenery distinguishing by Kesten [12]. Scenery distinguishing and reconstruction belongs to the general area of probability theory which deals with the ergodic properties of observations made by a random process in a random media. An important related problem is the T, T^{-1} problem studied by Kalikow [9]. Several important contributions about the properties of the observations were made later. these include Keane and den Hollander [10], den Hollander [2],den Hollander and Steif [3], Heicklen, Hoffman and Rudolph [5], and Levin and Peres [20]. Interest in the scenery distinguishing problem was sparked when Keane and den Hollander, as well as Benjamin, asked if all non-equivalent sceneries could be distinguished. Indentures was able to prove that there exist pairs of sceneries which can not be distinguished [14]. After Matzinger showed the validity of scenery reconstruction in the simple case of error-free observations made by a one-dimensional random walk without jumps (see [24, 23]), Kesten noticed that Matzingers method was inadequate to solve the reconstruction problem in the 2-dimensional case, as well as in the case when the random walk is allowed to jump. Subsequently, Loewe and Matzinger [16] were able to prove that scenery reconstruction is also possible on two-dimensional sceneries with many colors. Later, Matzinger, Merkl and Loewe [18] proved that with enough colors in one dimension one can do reconstruction even if the random walk is allowed to jump and thus is not a simple random walk. In general, scenery reconstruction becomes more difficult as the number of colors decreases (except in the trivial case when there is only one color). The most difficult case of reconstruction from observations made by a random walk with jumps on two-color sceneries was solved by Lember and Matzinger [17]. Den Hollander asked if it would be possible to do reconstruction if the jumps made by the random walk are not bounded. Lember, Lenstra and Matzinger [15] were able to answer this question. Finally, following a question of den Hollander, Loewe and Matzinger [19] investigated the possibility of reconstructing sceneries that are not i.i.d. but have some correlation. The

possibility to reconstruct finite pieces of sceneries in polynomial time following a question of Benjamini was investigated by Rolles and Matzinger [26] and [25].

In this article, we study one of the crucial techniques for finding markers used in scenery reconstruction and show that one can still construct and use markers when the observations are error- corrupted.

The paper is organized as follows. In Section 10.2, we consider a simplified example without errors. We show how in this simplified case, markers can be constructed and used for scenery reconstruction. Since many scenery reconstruction methods are very complicated, it seems worthwhile to present this simple case. In addition, it also serves as motivation, demonstrating the usefulness of markers. The following sections are concerned with how to define markers in the context of error-corrupted observations and construct stopping times that tell us when the random walk has returned to such a marker. In Section 10.3, we consider the likelihood of a marker being present in the scenery, given that some tell- tale event has occurred in the error-infested observation process. In Section 10.4, we show how to construct a multitude of stopping times which tell us when the random walk has returned to the location of a marker. It is assumed that there is a marker close to the random walk's starting state. Finally, in Section 10.5, we show how to find a Marker for the first time and then construct a series of stopping times which tell us when the random walk returns to that marker.

10.2 An Example of Scenery Reconstruction Using a Single Marker

In this section, we shall illustrate the use of markers in scenery reconstruction. Let us make some special assumptions which will only apply within this section:

- The scenery $\xi : \mathbb{Z} \to \{0, 1, 2\}$ is a 3-color scenery, with colors from the set $\{0, 1, 2\}$.
- The origin is colored with color 2: $\xi(0) = 2$.
- $\xi_1, \xi_{-1}, \xi_2, \xi_{-2}, \xi_3, \xi_{-3}, \ldots$ is a sequence of i.i.d. Bernoulli variables with parameter $1/2$. This means that, excepting the origin, the scenery ξ is a two color-scenery.
- The random walk $\{S_t\}_{t\in\mathbb{N}}$ is a simple random walk starting at the origin.

The only place where there is a 2 in the scenery is the origin. We can use this "2" as a "marker": Every time we see a 2 in the observations, we are at the origin. This implies:

$$\chi_t = 2 \Longrightarrow S_t = 0.$$

Let τ_k be the time of the k-th visit of S to the origin. Note that τ_k is observable since it is also the k-th time we observe a 2 in χ:

$$\tau_k := \min\{t \in \mathbb{N} \mid \chi_t = 2, t > \tau_{k-1}\}, \qquad k \geq 1.$$

By convention, we set $\tau_0 := 0$. Consider the following sequence of binary words:

$$w_1 = 001100, \quad w_2 = 0011001100, \quad w_3 = 00110011001100, \quad \ldots$$

Since the scenery ξ is i.i.d., every finite pattern will occur in ξ infinitely often. Hence all the strings w_k will occur in ξ infinitely often. Let x_k denote the closest place to the origin where w_k occurs in the scenery. (If there be two such places, choose the one to the right of the origin.) Hence x_k is a point z minimizing $|z|$ under the following constraint:

1. If $z > 0$, then
$$\xi_z \xi_{z+1} \dots \xi_{z+4k+1} = w_k.$$
2. If $z < 0$, then
$$\xi_z \xi_{z-1} \dots \xi_{z-4k-1} = w_k.$$

It is easy to see that the only way the string w_k can appear in the observations χ is by walking in a straight line over a portion of the scenery where w_k appears. In other words, we observe the word w_k at time t, that is,

$$\chi_t \chi_{t+1} \cdots \chi_{t+4k+1} = w_k,$$

if and only if, for all $i = 0, \dots, i + 4k + 1$, we have

$$S_{t+i} = S_t + iu$$

and

$$\xi_{S_t} \xi_{S_t+u} \dots \xi_{S_t+4ku+u} = w_k,$$

where $u = \pm 1$.

Almost surely, we have that

$$\lim_{k\to\infty} |x_k| = \infty$$

and on both sides of the origin there are infinitely many points from the sequence $x_k, k \in \mathbb{N}$.

The shortest time after a 2 at which we can observe the word w_k is x_k. It takes the random walk $|x_k|$ steps to go from the origin to x_k in minimal time. When doing so, the random walk must walk in a straight line only taking steps towards x_k. Whenever the random walk travels in a straight line, it produces a copy of the portion of the scenery which it has traversed. This copy is manifest and plain to see in the observations. The random walk goes from the origin to x_k infinitely often in the shortest possible time. This implies that when we observe 2 at time t followed by w_k at time $t + |x_k|$, then we have a copy of $\xi_0 \xi_u \xi_{2u} \dots \xi_{x_k}$ in the observations χ. Here, we take $u = x_k / |x_k|$. Hence we can reconstruct

$$\xi_0 \xi_u \xi_{2u} \dots \xi_{x_k} \tag{10.2.1}$$

using the following algorithm:

Algorithm 10.2.1. *1. Let μ_s denote the first time we observe the word w_k after time τ_s:*
$$\mu_s := \min\{t > \tau_s \mid w_k = \chi_t \chi_{t+1} \cdots \chi_{t+4k+1}\}.$$

2. Let d_k denote the minimum time at which we can observe the finite string w_k after a 2:
$$d_k := \min\{\mu_s - \tau_s \mid s \in \mathbb{N}\}.$$

3. *Let s^* denote any s minimizing $\mu_s - \tau_s$. In other words, s^* is such that*

$$\mu_{s^*} - \tau_{s^*} = d_k.$$

4. *The output of our algorithm is*

$$\chi_{\tau_{s^*}} \chi_{\tau_{s^*}+1} \cdots \chi_{\mu_{s^*}} \tag{10.2.2}$$

For the reasons explained above, the output of the above algorithm is equal to the piece of the scenery located between the origin and x_k inclusive with probability one. This should demonstrate the usefulness of markers in scenery reconstruction.

10.3 Existence of a Marker

In this section, we take the scenery $\xi : \mathbb{Z} \to \{0, 1\}$ to be a two-color i.i.d. scenery. Thus, it is a realization of the process $\{\xi_z\}_{z\in\mathbb{Z}}$ where the ξ_z's are i.i.d. Bernoulli random variables with parameter $1/2$.

As before, the observation of the scenery ξ by the random walk S at time t is denoted by $\chi_t := \xi(S_t)$. We assume that the errors are i.i.d. with the probability $P(\nu_t = 1) = \epsilon$ of an error at time t being strictly smaller than $1/2$. Then, the error-corrupted observation $\tilde{\chi}_t$ at time t is given by

$$\tilde{\chi}_t := (\chi_t + \nu_t) \mod 2.$$

We write $\chi = (\chi_0, \chi_1, \ldots)$ for the error-free observations and $\tilde{\chi} := (\tilde{\chi}_0, \tilde{\chi}_1, \ldots)$ for the error-corrupted observations. The scenery ξ, the random walk S and the error process $\{\nu_t\}_{t\in\mathbb{N}}$ are all assumed to be independent of each other.

For completeness, the following list details all the assumptions we make in this section.

- Let $S = \{S_t\}_{t\geq 0}$ be a recurrent random walk on $\mathbb{Z}$ starting at the origin which can visit any point $z \in \mathbb{Z}$ with positive probability. This means that for every $z \in \mathbb{Z}$ there exists $t_z \geq 0$ such that $P(S_{t_z} = z) > 0$.

- The distribution of the increments of the random walk S has bounded support. That is there exists $L > 0$ such that

$$P(|S_{t+1} - S_t| \leq L) = 1.$$

- The process $\xi = \{\xi_z\}_{z\in\mathbb{Z}}$ is such that the ξ_k's are i.i.d. Bernoulli variables with parameter $1/2$.

- The errors ν_t, for $t \geq 0$, are i.i.d. Bernoulli variables with parameter $\epsilon = P(\nu_t = 1)$ strictly smaller than $1/2$. Here ϵ denotes the probability of an error.

- The three processes ξ, S and $\nu = \{\nu_t\}_{t\geq 0}$ are independent of each other.

Next we need to define a few events. Firstly, define

$$A^n := \left\{ \sum_{t=0}^{n^2} \tilde{\chi}_t \leq \epsilon n^2 \right\}.$$

Let B^n be the event that there exists a contiguous block of zeros in ξ of length greater than $n^{0.1}$ in the interval $[-Ln^2, Ln^2]$. More precisely,

$$B^n := \left\{ \begin{array}{l} \exists z \in [-Ln^2, Ln^2 - n^{0.1}] \text{ such that} \\ \xi_z = \xi_{z+1} = \ldots = \xi_{(z+n^{0.1})} = 0 \end{array} \right\}$$

Let C^n be the event that the error-free observation process reveals more than $n^{1.7}$ 1's in the first n^2 observations:

$$C^n := \left\{ \sum_{t=0}^{n^2} \chi_t \geq n^{1.7} \right\}.$$

We shall denote the complement of an event E by $E^\perp$. Next, let us present the main result of this section.

Theorem 10.3.1. *For n large enough,*

$$P(B^{n\perp} \mid A^n) \leq \exp\left(-(1-2\epsilon)^2 n^{1.4}/3\right).$$

Hence, $P(B^n \mid A^n) \to 1$ as $n \to \infty$.

Theorem 10.3.1 says that if in the first n^2 error-corrupted observations we observe a significantly low number of 1's, then with very high probability there is a contiguous block of zero's of length $n^{0.1}$ very close to the origin in the scenery ξ. This unbroken block of zeros will be used in the next section as a marker to tell us when the random walk is back near the origin.

To prove Theorem 10.3.1, we will need a number of lemmas. The first two of these will be used numerous times throughout this and the following section. Let us start with a large deviation result.

Lemma 10.3.1. *Let $\Delta > 0$. Let $X_1, X_2, \ldots$ be a sequence of zero-mean random variables such that $\{S_t\}_{t\geq 0}$ is a Martingale, where $S_0 = 0$ and $S_t = \sum_{i=1}^t X_i$ for $t \geq 1$. Assume furthermore that the random variables all have bounded range, that is, for some $a > 0$, $|X_i| \leq a$ for all $i = 1, 2, \ldots$. Then, for all $k \geq 1$,*

$$P\left(\frac{\sum_{i=1}^k X_i}{k} \geq \Delta\right) \leq \exp\left(-\frac{k\Delta^2}{2a^2}\right). \tag{10.3.1}$$

Proof. From Chernov's inequality, we have

$$\begin{aligned} & P\left(\frac{\sum_{i=1}^k X_i}{k} \geq \Delta\right) \\ =& B_{\text{cell\_OK}}(n)[\mathbf{1}_{S_k \geq k\Delta}] \\ \leq& B_{\text{cell\_OK}}(n)[\exp(\beta S_k - \beta k\Delta)], \qquad \text{for all } \beta > 0, \\ =& \exp(-\beta k\Delta) B_{\text{cell\_OK}}(n)[\exp(\beta S_k)]. \end{aligned}$$

Then, an application of the Azuma-Hoeffding lemma yields

$$P\left(\frac{\sum_{i=1}^k X_i}{k} \geq \Delta\right) \leq \qquad \exp(-\beta k\Delta)\exp(k\beta^2 a^2/2) = \exp(-\beta k\Delta + k\beta^2 a^2/2).$$

The right-hand side of this final expression obtains its optimal (minimum) value at $\beta = \Delta/a^2$. Substituting $\beta = \Delta/a^2$ into the equation yields the desired result. □

Lemma 10.3.2. *Let $S := \{S_t\}_{t\geq 0}$ be a random walk with bounded jumps which starts at the origin. Then:*

1. *There exists a constant $C' > 0$ such that*

$$P(\max_{0\leq t\leq n^2} |S_t| \leq n) \geq C'$$

for all $n \geq 0$.

2. *As $n \to \infty$,*

$$P(\max_{0\leq t\leq n^{2-\gamma}} |S_t| \leq n) \to 1,$$

for any $0 < \gamma < 2$).

Proof. 1. Define $Z_k := \{Z_k(s)\}_{s\geq 0}$ where $Z_k(s) := \frac{1}{k} S_{sk^2}$ and let $W := \{W_t\}_{t\geq 0}$ denote a Brownian motion. Then, by the invariance principle, $Z_n \xrightarrow{\mathcal{D}} W$ as $n \to \infty$. In particular, $Z_n(s) \xrightarrow{\mathcal{D}} W_s$ and so

$$\max_{0\leq t\leq n^2} \left|\frac{S_t}{n}\right| = \max_{s=0,1/n^2,2/n^2,\ldots,1} |Z_n(s)| \xrightarrow{\mathcal{D}} \max_{0\leq s\leq 1} |W_s|$$

as $n \to \infty$. Thus,

$$P(\max_{0\leq t\leq n^2} |S_t| \leq n) \to P(\max_{0\leq s\leq 1} |W_s| \leq 1) = P(\psi_{[-1,1]} > 1) > 0,$$

where $\psi_{[-1,1]}$ is the first exit time of the Brownian motion W from the interval $[-1, 1]$. The positivity of $P(\psi_{[-1,1]} > 1)$ may be deduced from the analytic expression

$$P(\psi_{[-1,1]} \in dt) = \frac{2}{\sqrt{2\pi t^3}} \sum_{n=-\infty}^{\infty} (4n+1) e^{-\frac{(4n+1)^2}{2t}}\, dt,$$

which is a special case of an expression derived in [**?**].

Thus, since $P(\max_{0\leq t\leq n^2} |S_t| \leq n) > 0$ for all n, it follows that there exists $C' > 0$ such that $P(\max_{0\leq t\leq n^2} |S_t| \leq n) \geq C'$ for all $n \geq 0$.

2. As we are assuming that S has i.i.d. increments, let σ^2 denote the variance of an increment. Then, applying the Kolmogorov inequality (for example, see Chapter 14.6 of [27]), we have

$$\begin{aligned} &P(\max_{0\leq t\leq n^{2-\gamma}} |S_t| \leq n) \\ =&1 - P(\max_{0\leq t\leq n^{2-\gamma}} |S_t| \geq n+1) \\ \geq&1 - \frac{n^{2-\gamma}\sigma^2}{(n+1)^2} \geq 1 - \sigma^2 n^{-\gamma} \to 1, \end{aligned}$$

as $n \to \infty$.

□

Lemma 10.3.3. *there exists a constant $c > 0$ not depending on n such that, for all $n \geq 0$,*

$$P(A^n) \geq c \left(\frac{1}{4}\right)^n. \tag{10.3.2}$$

Proof. Let D^n and E^n be events defined as follows:

$$\begin{aligned} D^n :=& \{\forall z \in [-n, n],\ \xi_z = 0\} \text{ and} \\ E^n :=& \{\forall t \in [0, n^2],\ S_t \in [-n, n]\}. \end{aligned}$$

By Part 1 of Lemma 10.3.2, we know that there exists a constant $c' > 0$, not depending on n, such that $P(E^n) \geq c'$. Furthermore, $P(D^n) = (1/2)^{2n+1}$. Since, conditional on $D^n \cap E^n$, $\sum_{t=1}^{n^2} \tilde{\chi}_t = \sum_{t=1}^{n^2} \nu_t \sim \text{Bin}(n^2, \epsilon)$, we see that $P(A^n \mid D^n \cap E^n) > 0$ for all $n \geq 0$. Furthermore, by the central limit theorem, $P(A^n \mid D^n \cap E^n) \longrightarrow \frac{1}{2}$ as $n \to \infty$. Thus, there exists $c'' > 0$ such that $P(A^n \mid D^n \cap E^n) \geq c''$ for all n. Consequently,

$$P(A^n) \geq c'' P(D^n \cap E^n) = c'' P(D^n) P(E^n) \geq c \left(\frac{1}{2}\right)^{2n},$$

where $c = c'c''/2$. □

Lemma 10.3.4.

$$P(A^n \mid C^n) \leq \exp\left(-\frac{(1-2\epsilon)^2 n^{1.4}}{2}\right) \tag{10.3.3}$$

for all $n \geq 1$.

Proof. Let Z and $\tilde{Z}$ denote the sums

$$Z := \sum_{t=0}^{n^2} \chi_t$$

and

$$\tilde{Z} := \sum_{t=0}^{n^2} \tilde{\chi}_t.$$

Conditional on $\chi_t = 1$, $\tilde{\chi}_t$ has expectation $1 - \epsilon$ whilst conditional on $\chi_t = 0$, $\tilde{\chi}_t$ has expectation ϵ. Thus, $\tilde{Z}$ conditional on Z, has the same distribution as the sum of n^2 independent Bernoulli variables where Z of them have expectation $1 - \epsilon$ and the other $n^2 - Z$ have expectation ϵ. It follows that the conditional expectation of $\tilde{Z}$ given Z is $B_{\text{cell\_OK}}(n)[\tilde{Z} \mid Z] = n^2\epsilon + (1 - 2\epsilon)Z$. Now,

$$P(A^n \mid Z) = P(\tilde{Z} \leq \epsilon n^2 \mid Z) = P\left(\frac{\tilde{Z} - (\epsilon n^2 + (1-2\epsilon)Z)}{n^2} \leq -\frac{(1-2\epsilon)Z}{n^2} \,\middle|\, Z\right). \tag{10.3.4}$$

Since, conditional on Z, $\tilde{Z}$ is distributed like a sum of n^2 independent Bernoulli variables, it follows that we can apply Lemma 12.2.1. Taking $k = n^2$, $a = 1$ and $\Delta = (1 - 2\epsilon)Z/n^2$, we find that the expression on the right-hand side of (10.3.4) is bounded by

$$\exp\left(-\frac{(1-2\epsilon)^2 Z^2}{2n^2}\right).$$

Hence, when $Z \geq n^{1.7}$ is assumed given, we obtain

$$P(A^n \mid C^n) = P(A^n \mid Z \geq n^{1.7}) \leq \exp\left(-\frac{(1-2\epsilon)^2 n^{1.4}}{2}\right)$$

and the proof is complete. □

Next, we define $q^n_{x,y}$ to be the probability that the random walk S visits the point x or y before time $n^{0.21}$:

$$q^n_{x,y} := P\left(\exists t \leq n^{0.21}, S_t \in \{x, y\}\right).$$

Let q^n denote the minimum

$$q^n := \min_{(x,y)\in G_n} q^n_{x,y},$$

where $G_n := \{(x, y) \in [-n^{0.1}, n^{0.1}]^2 \,\big|\, x < 0 < y\}$. The following lemma will be needed to prove Lemma 10.3.6.

Lemma 10.3.5. $\lim_{n\to\infty} q^n = 1$.

Proof. Let n be large and choose two points $x, y \in [-n^{0.1}, n^{0.1}]$ such that $x < 0 < y$. Also, let I_x and I_y denote the intervals $I_x := [x - L, x + L]$ and $I_y := [y - L, y + L]$ respectively. Then, we define τ_{xy} to be the time of the first visit by S to $I_x \cup I_y$ and use E_{xy} to denote the event that S visits x or y before time $n^{0.21}$:

$$E_{xy} := \left\{\exists t \leq n^{0.21}, S_t \in \{x, y\}\right\}.$$

Further, let $E^n_{a,xy}$ denote the event that, within time $n^{0.2}$ of the stopping time τ_{xy}, the random walk visits all the points in a neighborhood of radius L of the point $S_{\tau_{xy}}$. Hence, $E^n_{a,xy}$ denotes the event that for all z satisfying $\left|z - S_{\tau_{xy}}\right| \leq L$, there exists $t \in [\tau_{xy}, \tau_{xy} + n^{0.2}]$ such that $S_t = z$.

Lastly, define E^n_b to be the event that the random walk S is outside the interval $[-n^{0.1}, n^{0.1}]$ at time $t = n^{0.205}$:

$$E^n_b := \left\{S_{n^{0.205}} \notin [-n^{0.1}, n^{0.1}]\right\}.$$

Since the random walk S starts at the origin, it must cross (but not necessarily hit) either x or y before leaving the interval $[-n^{0.1}, n^{0.1}]$. Since the step lengths of S are bounded by L, the random walk must visit either I_x or I_y in order to exit the interval $[-n^{0.1}, n^{0.1}]$. Hence, when E^n_b holds, we have

$$\tau_{xy} \leq n^{0.205}. \tag{10.3.5}$$

Now, whenever (10.3.5) and $E^n_{a,xy}$ hold, the set $\{x, y\}$ will be visited before time $n^{0.205} + n^{0.2}$. For n large enough, $n^{0.205} + n^{0.2} < n^{0.21}$. Hence,

$$E^n_{a,xy} \cap E^n_b \subseteq E^n_{xy},$$

for any $(x, y) \in G_n$. This implies that

$$P(E^{n\perp}_{xy}) \leq P(E^{n\perp}_{a,xy}) + P(E^{n\perp}_b).$$

Next, let E_a^n denote the event that the random walk visits all the points in $[-L, L]$ before time $n^{0.2}$. By the strong Markov property of S, we see that $P(E_{a,xy}^{n\perp}) = P(E_a^{n\perp})$ and hence we obtain

$$P(E_{xy}^{n\perp}) \leq P(E_a^{n\perp}) + P(E_b^{n\perp}). \tag{10.3.6}$$

Note that the bound on the right side does not depend on either x or y and that (10.3.6) holds for all $(x, y) \in G_n$. Therefore,

$$q^n = \min_{x,y} P(E_{xy}^n) \geq 1 - P(E_a^{n\perp}) + P(E_b^{n\perp}). \tag{10.3.7}$$

Now, by the central limit theorem, we have

$$\lim_{n\to\infty} P(E_b^{n\perp}) = 0. \tag{10.3.8}$$

Also, by the assumption that S is recurrent and hence has positive probability of visiting all points in $\mathbb{Z}$, we find that $P(E_a^\infty) = 1$. Then, by continuity of probability,

$$\lim_{n\to\infty} P(E_a^n) = P(E_a^\infty) = 1. \tag{10.3.9}$$

Then, by applying (10.3.8) and (10.3.9) to (10.3.7), we conclude that

$$\lim_{n\to\infty} q^n = 1.$$

□

Lemma 10.3.6. *For sufficiently large n,*

$$P(C^{n\perp} \mid B^{n\perp}) \leq \exp(-n^{1.79}/8). \tag{10.3.10}$$

Proof. We begin by defining Bernoulli variables $\{Y_k\}_{k\geq 1}$ in the following way:

$$Y_k = \mathbf{1}_{\sum_{t=(k-1)n^{0.21}}^{kn^{0.21}} \chi_t \geq 1} = \mathbf{1}_{\exists t \in [(k-1)n^{0.21}, kn^{0.21}] \text{ such that } \chi_t = 1}.$$

Clearly, $Y_k \leq \sum_{t=(k-1)n^{0.21}}^{kn^{0.21}} \chi_t$ and

$$\sum_{k=1}^{n^{1.79}} Y_k \leq \sum_{t=0}^{n^2} \chi_t.$$

Thus,

$$C^{n\perp} = \{\sum_{t=1}^{n^2} \chi_t < n^{1.7}\} \subseteq \{\sum_{k=1}^{n^{1.79}} Y_k < n^{1.7}\}$$

and

$$P(C^{n\perp} \mid B^{n\perp}) \leq P(\sum_{k=1}^{n^{1.79}} Y_k < n^{1.7} \mid B^{n\perp}). \tag{10.3.11}$$

Let $\mathcal{F} := \bigcup_{k=1} \infty \mathcal{F}_k$ be the σ-algebra defined by the filtration $\{\mathcal{F}_k\}_{k\geq 1}$, where

$$\mathcal{F}_k := \sigma(S_t, \xi_z \mid t \leq kn^{0.21},\ z \in \mathbb{Z}).$$

The sequence $\{Y_k\}_{k\in\mathbb{N}}$ is $\mathcal{F}$-adapted. Furthermore, $M_k = \sum_{i=1}^{k}(Y_i - B_{\text{cell\_OK}}(n)[Y_i \mid \mathcal{F}_{i-1}])$ is a Martingale with respect to $\{\mathcal{F}_k\}_{k\geq 1}$.

Starting from the origin, the random walk S takes steps with lengths bounded by L. This implies that S stays in the set $[-Ln^{0.21}, Ln^{0.21}]$ during the time interval $[0, n^{0.21}]$. When the event $B^{n\perp}$ holds, there exists , for every point $z \in [-Ln^{0.21}, Ln^{0.21}]$, two random points x^* and y^* such that $z - n^{0.1} < x^* < z < y^* < z + n^{0.1}$ with $\xi_{x^*} = \xi_{y^*} = 1$. By the strong Markov property, given that the random walk is at z at time t, the probability of visiting x or y during the time interval $(t, t + n^{0.21}]$ is equal to $q^n_{x^*-z,y^*-z}$. Hence this probability is larger than q^n. In this case the conditional probability that we observe at least one 1 in χ during the time interval $[t, t + n^{0.21}]$ is larger than or equal to q^n. (Conditional on $B^{n\perp}$ and S_t, where $S_t \in [-Ln^{0.21}, Ln^{0.21}]$.) This means that, when the event $B^{n\perp}$ holds, then

$$P(Y_k = 1 \mid \mathcal{F}_{k-1}) = B_{\text{cell\_OK}}(n)[Y_k \mid \mathcal{F}_{k-1}] \geq q^n,$$

for all $1 \leq k \leq n^{1.79}$. Since

$$\lim_{n\to\infty} q^n = 1$$

by Lemma 10.3.5, we can assume that n is large enough so that $q^n > 3/4$. Thus,

$$B_{\text{cell\_OK}}(n)[Y_k \mid \mathcal{F}_{k-1}] \geq \frac{3}{4} \tag{10.3.12}$$

for n large enough when $B^{n\perp}$ holds and $k \leq n^{1.79}$. Because of (10.3.12) and since $B^{n\perp}$ is $\mathcal{F}_0$-measurable, we find

$$P\left(\sum_{k=1}^{n^{1.79}} Y_k < n^{1.7} \mid B^{n\perp}\right) \leq P\left(\frac{\sum_{k=1}^{n^{1.79}}(Y_k - B_{\text{cell\_OK}}(n)[Y_k \mid \mathcal{F}_{k-1}])}{n^{1.79}} < \frac{n^{1.7}}{n^{1.79}} - \frac{3}{4}\middle| B^{n\perp}\right) \tag{10.3.13}$$

for large n. Since $\{M_k\}_{k\geq 0}$ constitutes a Martingale with respect to the filtration $\{\mathcal{F}_k\}_{k\geq 0}$ and since $B^{n\perp}$ is $\mathcal{F}_0$- measurable, M_k remains a Martingale when we condition on $B^{n\perp}$. Therefore, we can apply Lemma 12.2.1 to bound the probability on the right side of (10.3.13). For this we take $a = 1$ and $k = n^{1.79}$. For n large enough we have that $n^{-0.09} - \frac{3}{4} < -\frac{1}{2}$, which allows us to take the value $\frac{1}{2}$ for the Δ of lemma 12.2.1. In this way we find that the right side of (10.3.13) is smaller than $\exp(-n^{1.79}/8)$. Combining inequalities (10.3.11) and (10.3.13) with this bound completes the proof. □

Lemma 10.3.7. *For n sufficiently large,*

$$P(A^n \mid B^{n\perp}) \leq 3\exp\left(-\frac{(1-2\epsilon)^2 n^{1.4}}{2}\right). \tag{10.3.14}$$

Proof. We have

$$\begin{aligned} P(A^n \mid B^{n\perp}) &= P\left((A^n \cap C^n) \cup (A^n \cap C^{n\perp}) \mid B^{n\perp}\right) \\ &= P(A^n \cap C^n \mid B^{n\perp}) + P(A^n \cap C^{n\perp} \mid B^{n\perp}) \\ &\leq P(A^n \mid C^n)P(B^{n\perp})^{-1} + P(C^{n\perp} \mid B^{n\perp}). \end{aligned}$$

Note that for n large, $P(B^{n\perp})$ is close to one. Thus, let us assume that $P(B^{n\perp}) > 1/2$. With this assumption we obtain

$$P(A^n \mid B^{n\perp}) \leq 2P(A^n \mid C^n) + P(C^{n\perp} \mid B^{n\perp}).$$

We can now apply the bounds from inequalities (10.3.3) and (10.3.10) to the right-hand side of this last inequality. Note that the first of these two bounds is much larger than the second. We therefore find that $P(A^n \mid B^{n\perp})$ is smaller than 3 times the larger bound, provided that n is large enough. In other words,

$$P(A^n \mid B^{n\perp}) \leq 3\exp\left(-\frac{(1-2\epsilon)^2 n^{1.4}}{2}\right)$$

for large n. This completes the proof. □

We can now prove Theorem 10.3.1.

Proof of Theorem 10.3.1. We have

$$P(B^{n\perp} \mid A^n) = P(A^n \mid B^{n\perp}) \cdot \frac{P(B^{n\perp})}{P(A^n)} \leq \frac{P(A^n \mid B^{n\perp})}{P(A^n)}. \tag{10.3.15}$$

Applying inequalities (10.3.2) and (10.3.14) to this expression, we obtain

$$P(B^{n\perp} \mid A^n) \leq \frac{3\exp\left(-(1-2\epsilon)^2 n^{1.4}/2\right)}{c(1/4)^n} = \exp\left(-(1-2\epsilon)^2 n^{1.4}/2 + n\ln 4 + \ln(3/c)\right)$$

for n sufficiently large. In the expression $-(1-2\epsilon)^2 n^{1.4}/2 + n\ln 4 + \ln(3/c)$, the dominating term is the first. This implies that for n large enough, $-(1-2\epsilon)^2 n^{1.4}/2 + n\ln 4 + \ln(3/c)$ is smaller than $-(1-2\epsilon)^2 n^{1.4}/3$. This in turn implies that

$$P(B^{n\perp} \mid A^n) \leq \exp\left(-(1-2\epsilon)^2 n^{1.4}/3\right)$$

for large n and this yields the desired result. ■

We conclude this section with a lemma that will be useful in the next section.

Lemma 10.3.8. *For n large,*

$$P(A^n) \leq 2Ln^2 \left(\frac{1}{2}\right)^{n^{0.1}}.$$

Proof. Let B_z^n denote the event that there is a contiguous block of zeros in the scenery between z and $z + n^{0.1}$ inclusive:

$$B_z^n := \{\xi_z = \xi_{z+1} = \ldots = \xi_{z+n^{0.1}} = 0\}.$$

Since the scenery is generated by i.i.d. Bernoulli random variables, $P(B_z^n) = (1/2)^{n^{0.1}+1}$. Furthermore, with this definition, $B^n = \bigcup_z B_z^n$, where the union is taken over z in $[-Ln^2, Ln^2 - n^{0.1}]$. The length of this interval is smaller than $2Ln^2$. Thus we see that

$$P(B^n) \leq \sum_z P(B_z^n) \leq 2Ln^2 \left(\frac{1}{2}\right)^{n^{0.1}+1}. \tag{10.3.16}$$

Now,

$$P(A^n) = P(A^n \mid B^{n\perp})P(B^{n\perp}) + P(A^n \mid B^n)P(B^n) \leq P(A^n \mid B^{n\perp}) + P(B^n) \quad (10.3.17)$$

We can bound $P(A^n|B^{n\perp})$ using the inequality (10.3.14) and $P(B^n)$ with the aid of (10.3.16). The bound given on the right-hand side of (10.3.16) is asymptotically much larger than that given in (10.3.14). Thus, for large enough n, we can bound (10.3.17) by twice the larger of the two bounds and obtain

$$P(A^n) \leq 2Ln^2 \left(\frac{1}{2}\right)^{n^{0.1}}.$$

□

10.4 Returning to a Marker

The main result of the last section states that, given that we observe a significantly low number of 1's in the first n^2 error-corrupted observations (the event A^n), there is a high probability that the scenery ξ has a contiguous block of $n^{0.1}$ or more zeros in the interval $[-Ln^2, Ln^2]$. In the context of sceneries observed with errors, we shall call such a block a marker.

In this section we shall prove that, by only looking at the observations $\tilde{\chi}$, we can tell $\exp(n^{0.001})$ times with high probability when the random walk is back at the marker. More precisely, we shall show that we can construct $\exp(n^{0.001})$ stopping times, which are observable, that is, $\sigma(\tilde{\chi})$ measurable, and will stop the random walk close to the marker in the interval $[-2Ln^2, 2Ln^2]$. Of course, we need to make the assumption that there is such a marker in the interval $[-Ln^2, Ln^2]$. In order to do this, we will assume that the probability distribution governing our whole world of scenery, random walk and errors has properties similar to the measure we obtain by taking the distribution used in the previous section conditional on the event B^n. To simplify notation, we do not use $P(\cdot \mid B^n)$, but a measure $P_2(\cdot)$ having very similar properties to $P(\cdot \mid B^n)$ instead. Through out this section, $P_2(\cdot)$ denotes a measure which satisfies the following conditions:

- The random walk S and the scenery ξ are independent of each other.
- The random walk S has the same distribution under $P_2(\cdot)$ as it had in the previous section. Moreover, it starts at the origin.
- The scenery outside the interval $[-Ln^2, Ln^2]$ is i.i.d. Bernoulli with parameter 1/2.
- The portion of the scenery inside the interval $[-Ln^2, Ln^2]$ is independent of the remainder outside the interval.
- The scenery ξ P_2-almost surely contains a contiguous block of zeros longer than $n^{0.1}$ in $[-Ln^2, Ln^2]$. We require that

$$P_2\left(\exists z \in [-Ln^2, Ln^2 - n^{0.1}] \text{ such that } \xi_z = \xi_{z+1} = \ldots = \xi_{z+n^{0.1}} = 0\right) = 1.$$

- The errors under $P_2(\cdot)$ are distributed as before and are independent of the random walk and the scenery. In other words, the process $\{\nu_t\}_{t\geq 0}$ is P_2- independent of $\{S_t\}_{t\geq 0}$ and $\{\xi_z\}_{z\in\mathbb{Z}}$. Also, $P_2(\nu_t = 1) = \epsilon$. Once again, for all $t \in \mathbb{N}$,

$$\chi_t := \xi(S_t) \text{ and } \tilde{\chi}_t := \chi_t + \nu_t \mod 2.$$

Next, we define an increasing set of stopping times that are supposed to tell us when the random walk S is back close to the origin.

Definition 10.4.1. *Let T denote the random integer set*

$$T := \left\{ t \geq 0 \ : \ \sum_{s=t}^{t+n^{0.1}} \tilde{\chi}_s \leq \epsilon n^{0.1} \right\}.$$

For $k > 0$, let τ_k denote the k-th element (under the usual ordering on $\mathbb{N}$) of the set T.

We can now state the principal result of this section. It says that, with high P_2-probability, all of the first $\exp(n^{0.001})$ stopping times τ_k stop the random walk in the interval $[-2Ln^2, 2Ln^2]$ near a contiguous block of more than $n^{0.1}$ zeros. Furthermore, it also says that these stopping times all occur prior to time $\exp(n^{0.003})$ with high P_2-probability.

Theorem 10.4.1. *For large n,*

$$\begin{aligned} &P_2\left(\forall k \leq \exp(n^{0.001}), S_{\tau_k} \in [-2Ln^2, 2Ln^2] \text{ and } \tau_{\exp(n^{0.001})} \leq \exp(n^{0.003})\right) \\ \geq &1 - 3\exp(-n^{0.003}/4). \end{aligned}$$

Before continuing, we shall define a few useful intervals and a number of events that we shall need in the sequel.

$$\begin{aligned} I_1^n &:= [-Ln^2, Ln^2], & I_2^n &:= [-1.5Ln^2, 1.5Ln^2], \\ I_3^n &:= [-2Ln^2, 2Ln^2], & I_4^n &:= [-Ln^2, Ln^2 - n^{0.1}], \\ I_5^n &:= [-Ln^2 - n^{0.005}, Ln^2], & I_6^n &:= [-Ln^2 + n^{0.1}/2, Ln^2 - n^{0.1}/2], \\ I_7^n &:= [-L\exp(n^{0.003}), L\exp(n^{0.003})]. \end{aligned}$$

The first event $E^n_{\text{no}-\text{error}}$ says that we never see a significantly low average of 1's in the observations up to time $t = \exp(n^{0.003})$ when we are outside I_2^n.

$$E^n_{\text{no}-\text{error}} := \left\{ \sum_{s=t}^{t+n^{0.1}} \tilde{\chi}_s > \epsilon n^{0.1}, \ \ \forall t \leq \exp(n^{0.003}) \text{ such that } S_t \notin I_2^n \right\}.$$

We know that under $P_2(\cdot)$ there is a block of color zero having length $n^{0.1}$ in I_1^n with probability one. Let z_c denote the center of such a block. Thus, $z_c \in I_6^n$ P_2-almost surely and

$$P_2(\xi_z = 0, \ \forall z \in [z_c - n^{0.1}/2, z_c + n^{0.1}/2]) = 1.$$

Note that, by assumption, z_c is P_2-independent of $\{S_t\}_{t\geq 0}$ and $\{\nu_t\}_{t\geq 0}$. Let κ_l^* denote the l-th visit by S to the point z_c. Let κ_k denote the $l = kn^{0.1}$-th stopping time κ_l^*. More precisely,

$$\kappa_k := \kappa^*_{kn^{0.1}}, \qquad k \in \mathbb{N}.$$

We define the stopping times κ_k in this way to ensure they are separated by time periods of length at least $n^{0.1}$.

Let E^n_{visits} denote the event that there are more than $\exp(n^{0.002})$ visits to z_c before time $\exp(n^{0.003})$:

$$E^n_{\text{visits}} := \left\{\kappa_{\exp(n^{0.002})} \leq \exp(n^{0.003})\right\}.$$

Let Y_k denote the Bernoulli variable which is equal to one if and only if

$$\sum_{s=\kappa_k}^{\kappa_k+n^{0.1}} \tilde{\chi}_s \leq \epsilon n^{0.1}.$$

Let $E^n_{\text{marker-works}}$ denote the event that we observe a significantly low number of ones more than 1/3 of the time after a stopping time $\kappa_k, k \leq \exp(n^{0.002})$:

$$E^n_{\text{marker-works}} := \left\{ \sum_{k=1}^{\exp(n^{0.002})} Y_k \geq \frac{\exp(n^{0.002})}{3} \right\}.$$

The final event we shall need is E^n_{OK} which is the event that our stopping times work the way we want, that is,

$$E^n_{\text{OK}} := \left\{\forall k \leq \exp(n^{0.001}),\ S_{\tau_k} \in I^n_3 \text{ and } \tau_{\exp(n^{0.001})} \leq \exp(n^{0.003})\right\}.$$

With these definitions, we are ready to formulate the four intermediate results which we will need in order to prove Theorem 10.4.1. The first lemma is of a combinatorial nature.

Lemma 10.4.1. *For n sufficiently large, $E^n_{\text{no-error}} \cap E^n_{\text{visits}} \cap E^n_{\text{marker-works}} \subseteq E^n_{\text{OK}}$.*

Proof. When it occurs, the event $E^n_{\text{no-error}}$ guaranties that all the stopping times in T up to time $\exp(n^{0.003})$ stop the random walk inside the interval I^n_2. Since $I^n_2 \subseteq i^n_3$, $E^n_{\text{no-error}}$ implies that

$$S_{\tau_k} \in I^n_3 \text{ for all } \tau_k \leq \exp(n^{0.003}).$$

Next, if E^n_{visits} and $E^n_{\text{marker-works}}$ both hold, then there are at least $\exp(n^{0.002})/3$ stopping times in T which occur prior to time $\exp(n^{0.003})$. In other words,

$$\tau_{\exp(n^{0.002})/3} \leq \exp(n^{0.003}).$$

Now, when n is sufficiently large, $n^{0.001} \leq n^{0.002}/3$ and so

$$\tau_{\exp(n^{0.001})} \leq \exp(n^{0.003}).$$

Consequently, the simultaneous occurrence of both E^n_{visits} and $E^n_{\text{marker-works}}$ implies that $\tau_k \leq \exp(n^{0.003})$ for all $k \leq \exp(n^{0.001})$ when n is large.

Finally, if $E^n_{\text{no-error}}$ holds in addition to E^n_{visits} and $E^n_{\text{marker-works}}$, then we also see that $S_{\tau_k} \in I^n_3$ for all $k \leq \exp(n^{0.001})$. Thus, when all three events

$$E^n_{\text{no-error}}, \quad E^n_{\text{visits}} \text{ and } E^n_{\text{marker-works}}$$

occur simultaneously, then E^n_{OK} must also occur. □

The next three results yield lower bounds on the quantities $P_2(E^n_{\text{no-error}})$, $P_2(E^n_{\text{visits}})$ and $P_2(E^n_{\text{marker-works}})$.

Lemma 10.4.2. *For n large,*

$$P_2(E^n_{\text{no-error}}) \geq 1 - (0.6)^{n^{0.005}}. \tag{10.4.1}$$

Proof. Let $\kappa_{z,l}$ denote the time of the l-th visit by the random walk S to the point z. Let $E^n_{\text{no-error,z,l}}$ denote the event that there is no significantly low number of ones immediately following the stopping time $\kappa_{z,l}$, that is,

$$E^n_{\text{no-error,z,l}} := \left\{ \sum_{s=\kappa_{z,l}}^{\kappa_{z,l}+n^{0.1}} \tilde{\chi}_s > \epsilon n^{0.1} \right\}.$$

Up to time $t = \exp(n^{0.003})$, the random walk can not visit points z further away from the origin than $L\exp(n^{0.003})$ nor can it visit a point more than $\exp(n^{0.003})$ times. Thus, all the times which appear in the definition of the event $E^n_{\text{no-error}}$, that is all the times t for which $t \leq \exp(n^{0.003})$ and $S_t \notin I^n_2$ include the set of times $\kappa_{z,l}$ for which $z \in (I^n_7 \setminus I^n_2)$ and $l \leq \exp(n^{0.003})$. This implies that

$$\bigcap_{z,l} E^n_{\text{no-error,z,l}} \subseteq E^n_{\text{no-error}}, \tag{10.4.2}$$

where the intersection is taken over all $z \in (I^n_7 - I^n_2)$ and $l \leq \exp(n^{0.003})$.

If the random walk S is outside the interval I^n_2 at time t, then it is impossible for the random walk to reach the interval I^n_1 within time $n^{0.1}$. Thus if $S_t \notin I^n_2$ then S_s cannot be in I^n_1 for all times $s \in [t, t + n^{0.1}]$. However, outside the interval I^n_1, the scenery ξ has the same distribution under $P(\cdot)$ as it does under $P_2(\cdot)$. Thus, for $z \notin I^n_2$,

$$P_2(E^n_{\text{no-error,z,l}}) = P(E^n_{\text{no-error,z,l}}).$$

Furthermore, since the distribution of the scenery under $P(\cdot)$ is both time and spatially homogeneous, an application of the strong Markov property yields

$$P(E^n_{\text{no-error,z,l}}) = P(E^n_{\text{no-error,0,0}}) = P\left(\sum_{s=0}^{n^{0.1}} \tilde{\chi}_s > \epsilon n^{0.1} \right),$$

for all $z \notin I^n_2$. However the event $\left\{ \sum_{s=0}^{n^{0.1}} \tilde{\chi}_s > \epsilon n^{0.1} \right\}$ is just the event $A^{m\perp}$ from Section 10.3 with $m = n^{0.05}$. Hence, from Lemma 10.3.8 we obtain

$$P(E^{n\perp}_{\text{no-error,z,l}}) = P(A^m) \leq 2Lm^2 \left(\frac{1}{2}\right)^{m^{0.1}} = 2Ln^{0.1} \left(\frac{1}{2}\right)^{n^{0.005}} \tag{10.4.3}$$

for all $z \notin I^n_2$. By Combining this with (10.4.2), we arrive at

$$P(E^{n\perp}_{\text{no-error}}) \leq \sum_{z \in I^n_7 \setminus I^n_2, 0 \leq l \leq \exp(n^{0.003})} P(E^{n\perp}_{\text{no-error,z,l}}) \leq \; 2L\exp(2n^{0.003}) \cdot 2Ln^{0.1} \left(\frac{1}{2}\right)^{n^{0.005}}.$$

The final inequality comes about by recognizing that there are fewer than $2L\exp(2n^{0.003})$ pairs (z,l) with $z \in I_7^n \setminus I_2^n$ and $0 \le l \le \exp(n^{0.003})$.

Now, the dominating term in the bound on the right-hand side of this inequality is $(1/2)^{n^{0.005}}$. Thus, for n big enough, the expression on the right- hand side of the last inequality is smaller than $(0.6)^{n^{0.005}}$. The result follows by applying this bound to $E^n_{\text{no-error}}$. □

Lemma 10.4.3. *For large n,*

$$P_2(E^n_{\text{marker-works}}) \ge 1 - \exp(-0.225\exp(n^{0.002})). \tag{10.4.4}$$

Proof. Let R be a random walk with increments identical to those of the random walk S but starting at the random point z_c. Thus, $R_t := S_t + z_c$. Let χ_t^R denote the observation made by the random walk R at time t of the scenery ξ, that is, $\chi_t^R := \xi(R_t)$. We shall use $\tilde{\chi}_t^R$ to denote that same observation made with an error:

$$\tilde{\chi}_t^R := \chi_t^R + \nu_t \mod 2.$$

Let E_R^n denote the event that R does not stray from z_c by a distance greater than $n^{0.1}/2$ before time $n^{0.1}$:

$$E_R^n := \left\{\forall t \le n^{0.1}, |R_t - z_c| \le n^{0.1}/2\right\}.$$

Note that when E_R^n occurs, the random walk R stays within the contiguous block of zeros in ξ having z_c at it's center during its first $n^{0.1}$ steps. Consequently, if E_R^n holds, we have

$$\sum_{t=0}^{n^{0.1}} \chi_t^R = 0.$$

It follows, conditional on E_R^n, that $\sum_{t=0}^{n^{0.1}} \tilde{\chi}_t^R \sim \text{Bin}(n^{0.1}, \epsilon)$. Then, by the central limit theorem, as n tends to infinity,

$$P_2\left(\sum_{t=0}^{n^{0.1}} \tilde{\chi}_t^R \le \epsilon n^{0.1} \mid E_R^n\right)$$

converges to 1/2. Now,

$$P_2\left(\sum_{t=0}^{n^{0.1}} \tilde{\chi}_t^R \le \epsilon n^{0.1}\right) = P_2\left(\sum_{t=0}^{n^{0.1}} \tilde{\chi}_t^R \le \epsilon n^{0.1} \mid E_R^n\right) P_2(E_R^n) + P_2\left(\left\{\sum_{t=0}^{n^{0.1}} \tilde{\chi}_t^R \le \epsilon n^{0.1}\right\} \cap E_R^{n\perp}\right). \tag{10.4.5}$$

By Part 2 of Lemma 10.3.2, $P_2(E_R^n)$ converges to one as n converges to infinity. It also follows that $P_2(\{\sum_{t=0}^{n^{0.1}} \tilde{\chi}_t^R \le \epsilon n^{0.1}\} \cap E_R^{n\perp})$ converges to zero as n tends to infinity. Hence,

$$P_2\left(\sum_{t=0}^{n^{0.1}} \tilde{\chi}_t^R \le \epsilon n^{0.1}\right) \longrightarrow \frac{1}{2}$$

as $n \to \infty$.

Next, let us assume that n is large enough so that

$$P_2\left(\sum_{t=0}^{n^{0.1}} \tilde{\chi}_t^R \le \epsilon n^{0.1}\right) \ge 0.49. \tag{10.4.6}$$

Define $\mathcal{G}_k$ to be the σ-algebra

$$\mathcal{G}_k := \sigma(z_c, \xi_z; S_0, S_1, \ldots, S_{\kappa_k} + n^{0.1} \mid z \in \mathbb{Z})$$

and let $\mathcal{G}$ denote the filtration $\mathcal{G} := \bigcup_k \mathcal{G}_k$. It can be seen that the sequence of random variables $Y_1, Y_2, \ldots$ is $\mathcal{G}$-adapted. Furthermore, by definition, the stopping times κ_k are at least $n^{0.1}$ time steps apart from each other. It follows that κ_{k+1} happens no earlier than time $\kappa_k + n^{0.1}$. By the strong Markov property of the random walk S, when we stop the process at a point, it then continues on as though it were a new random walk which was started at that point, independent of what happened beforehand. Putting it another way, conditional on $\mathcal{G}_k$, S is distributed after time κ_{k+1} like R. So,

$$P_2(Y_{k+1} = 1 \mid \mathcal{G}_k) = P_2\left(\sum_{s=\kappa_{k+1}}^{\kappa_{k+1}+n^{0.1}} \tilde{\chi}_s \leq \epsilon n^{0.1} \,\middle|\, \mathcal{G}_k\right) = P_2\left(\sum_{t=0}^{n^{0.1}} \tilde{\chi}_t^R \leq \epsilon n^{0.1}\right) \; P_2\text{-a.s.}$$

According to (10.4.6), the final expression in the equality above is greater than 0.49 for n sufficiently large and, hence, $B_{\text{cell\_OK}}(n)[Y_{k+1}] \geq 0.49$. We can therefore use Lemma 12.2.1. Setting $k = \exp(n^{0.002})$, $a = 1/\sqrt{2}$ and $\Delta = 0.15$, we obtain

$$\begin{aligned}
&P_2(E^{n\perp}_{\text{marker}-\text{works}}) \\
=&P_2\left(\sum_{k=1}^{\exp(n^{0.002})} Y_k < \exp(n^{0.002})/3\right) \\
\leq&P_2\left(\frac{\sum_{k=1}^{\exp(n^{0.002})}(Y_k - B_{\text{cell_OK}}(n)[Y_k])}{\exp(n^{0.002})} < 1/3 - 0.49\right) \\
\leq&P_2\left(\frac{\sum_{k=1}^{\exp(n^{0.002})}(Y_k - B_{\text{cell_OK}}(n)[Y_k])}{\exp(n^{0.002})} \leq -0.15\right) \\
\leq&\exp(-0.225\exp(n^{0.002})).
\end{aligned}$$

Thus, $P_2(E^n_{\text{marker}-\text{works}}) \geq 1 - \exp(-0.225\exp(n^{0.002}))$ asymptotically. □

Lemma 10.4.4. *For large n,*

$$P_2(E_{\text{visits}})^n \geq 1 - \exp(-n^{0.003}/4). \tag{10.4.7}$$

Proof. Let $s := n^{0.1}\exp(n^{0.002})$ and observe that

$$E^n_{\text{visits}} = \left\{\kappa_{\exp(n^{0.002})} \leq \exp(n^{0.003})\right\} = \left\{\kappa^*_s \leq \exp(n^{0.003})\right\}.$$

Without loss of generality, assume that $z_c = 0$. If z_c is not zero, the proof is virtually the same since z_c is at most a distance polynomial in n away from the origin, which has negligible influence on the event, since we are considering exponentially long times in n. When $z_c = 0$, the event E^n_{visits} is simply the event that the random walk S visits the origin no less than s times before time $\exp(n^{0.003})$. Let Z_k denote the k-th interarrival time between consecutive visits by S to the origin. Hence, $\sum_{l=1}^k Z_l$ is the time of the k-th visit by S to the origin. Note that the random variables $Z_k, k \in \mathbb{N}$, are i.i.d. Define n_3 to

be the number $n_3 := \exp(n^{0.003})$. Under the assumption that $z_c = 0$ (which changes the ultimate bound we shall find in only a minute way), we have that

$$P_2(E^{n\perp}_{\text{visits}}) = P_2\left(\sum_{k=1}^{s} Z_k > n_3\right).$$

Now,

$$P_2\left(\sum_{k=1}^{s} Z_k > n_3\right) = P_2\left((\sum_{k=1}^{s} Z_k)^{1/3} > n_4\right),$$

where $n_4 := (n_3)^{1/3}$. For any set of positive numbers $\{a_l\}_1^j$, it is always true that $(\sum_{l=1}^{j} a_l)^3 \geq \sum_{l=1}^{j}(a_l)^3$. Hence, $\sum_{k=1}^{s}(Z_k)^{1/3} \geq (\sum_{k=1}^{s} Z_k)^{1/3}$ and so

$$P_2(E^{n\perp}_{\text{visits}}) \leq P_2(\sum_{k=1}^{s}(Z_k)^{1/3} > n_4).$$

By the Markov inequality,

$$P_2(E^{n\perp}_{\text{visits}}) \leq \frac{sB_{\text{cell\_OK}}(n)_2[(Z_1)^{1/3}]}{n_4} = \frac{n^{0.1}\exp(n^{0.002})B_{\text{cell\_OK}}(n)_2[(Z_1)^{1/3}]}{\exp(n^{0.003}/3)}. \tag{10.4.8}$$

It is known that $B_{\text{cell\_OK}}(n)_2[(Z_k)^{1/3}]$ is finite (see for example Durrett [4]) and thus is a constant not depending on n. Furthermore, the dominating factor in the bound given in (10.4.8) is $\exp(-n^{0.003}/3)$. It follows that, for n large enough, the right-hand side of (10.4.8) is smaller than $\exp(-n^{0.003}/4)$. □

Proof of Theorem 10.4.1. Lemma 10.4.1 yields

$$P_2(E^{n\perp}_{\text{OK}}) \leq P_2(E^{n\perp}_{\text{no-error}}) + P_2(E^{n\perp}_{\text{visits}}) + P_2(E^{n\perp}_{\text{marker-works}}). \tag{10.4.9}$$

For the three quantities $P_2(E^{n\perp}_{\text{no-error}})$, $P_2(E^{n\perp}_{\text{visits}})$ and $P_2(E^{n\perp}_{\text{marker-works}})$, we have the bounds (10.4.1), (10.4.7) and (10.4.4) respectively. The largest of these bounds is given by (10.4.7). Since $P_2(E^{n\perp}_{\text{OK}})$ is asymptotically smaller than 3 times this bound, we can write $P_2(E^{n\perp}_{\text{OK}}) \leq 3\exp(-n^{0.003}/4)$ for n large. ■

10.5 Recognizing Markers in Error-Corrupted Observations

In the preceding section, we investigated the case where we condition on the event B^n. Unfortunately, B^n is not an observable event. So instead, we need to condition on an event we are able to observe. We shall therefore choose to condition on A^n, which is observable. From Theorem 10.3.1, we know that, whenever A^n is observed, there is a block of zeros of length greater than $n^{0.1}$ close to the origin with high probability. (Here, close to the origin means belonging to $[-Ln^2, Ln^2]$.) We can then use this abnormally long block of zeros as a marker. This enables us to construct a total of $\exp(n^{0.001})$ stopping times τ_k and, with high probability, these stopping times all stop the random walk S in the interval

$[-2Ln^2, 2Ln^2]$. This is a situation similar to the one described in Section 10.2, where we had a 2 at the origin. When we previously conditioned on the event B^n, we "forced" the scenery to have a marker close to the origin. We did this in order to simplify notation in the preceding argument. In reality, we have to search for a Marker first. We shall now show how this can be done.

Let τ^* denote the first time t at which we see a string of lenght n^2 with less than ϵn^2 ones in the error-corrupted observations:

$$\tau^* := \min\left\{ t > 0 \ : \ \sum_{s=0}^{n^2} \tilde{\chi}_{t+s} \leq \epsilon n^2 \right\}.$$

Since ξ, S and ν are mutually independent and S is a recurrent random walk, the stopping time τ^* must be almost surely finite, that is, $P(\tau^* < \infty) = 1$. The neighborhood of S_{τ^*} is very similar to the origin under the conditional probability measure $P_2(\cdot)$. Due to the spatial homogeneity of the scenery, the theory which we developed in the last section holds for the point $z = S_{\tau^*}$ instead of the origin. Hence, with high probability, there is a block of more than $n^{0.1}$ contiguous zeros in the interval

$$I_{\tau^*} := [S_{\tau^*} - 2Ln^2, S_{\tau^*} + 2Ln^2].$$

Using this block of zeros as a marker, we can then construct a total of $\exp(n^{0.001})$ stopping times which, with high probability, all stop the random walk S in I_{τ^*}. We shall denote this sequence of stopping times by $\{\bar{\tau}_k\}_{k>0}$. They are defined as follows:

Definition 10.5.1. *For $k > 0$, let $\bar{\tau}_k$ denote the k-th element (under the usual ordering on $\mathbb{N}$) of the set $T \cap [\tau^*, \infty)$. Note that $\bar{\tau}_1 = \tau^*$.*

The result is that with high probability the first $\exp(n^{0.001})$ stopping times $\bar{\tau}_k$ stop S in I_ν.

Theorem 10.5.1. *The probability*

$$P\left(\forall k \leq \exp(n^{0.001}), S_{\bar{\tau}_k} \in I_{\tau^*} \text{ and } (\bar{\tau}_{\exp(n^{0.001})} - \tau^*) \leq \exp(n^{0.003})\right)$$

tends to one as $n \to \infty$.

Proof. The proof is analogous to that of Theorem 10.4.1. □

These stopping times can be used to reconstruct a little piece of the scenery ξ in the neighborhood of the point S_{τ^*}. The methods which can be used for this are similar to what was described in Section 10.2.

In [17], Lember and Matzinger show how beeing able to reconstruct a small amount of information contained in the neighborhood of markers implies that the whole scenery ξ can be reconstructed almost surely. Their proof, however, only pertains to the case of observations made without errors. The question as to whether or not it is possible to perform scenery reconstruction from error-corrupted observations of a two-color scenery remains open.

References

[1] Benjamini, I. and Kesten, H. (1996). Distinguishing sceneries by observing the scenery along a random walk path. *J. Anal. Math.* **69** 97–135.

[2] den Hollander, W. Th. F. (1988). Mixing properties for random walk in random scenery. *Ann. Probab.* **16** 1788–1802.

[3] den Hollander, F. and Steif, J.E. (1997). Mixing properties of the generalized T, T^{-1}-process. *J. Anal. Math.* **72** 165–202.

[4] Durrett, R. (1996). *Probability: theory and examples.* Duxbury Press, Belmont, CA, second edition.

[5] Heicklen, D., Hoffman, C. and Rudolph, D. J. (2000). Entropy and dyadic equivalence of random walks on a random scenery. *Adv. Math.* **156** 157–179.

[6] Harris, M. and Keane, M. (1997). Random coin tossing. *Probab. Theory Related Fields* **109** 27–37.

[7] Howard, C. D. (1996). Orthogonality of measures induced by random walks with scenery. *Combin. Probab. Comput.* **5** 247–256.

[8] Howard, C. D. (1997). Distinguishing certain random sceneries on $\mathbb{Z}$ via random walks. *Statist. Probab. Lett.* **34** 123–132.

[9] Kalikow, S. A. (1982). T, T^{-1} transformation is not loosely Bernoulli. *Ann. Math. (2)* **115** 393–409.

[10] Keane, M. and den Hollander, W. Th. F. (1986). Ergodic properties of color records. *Phys. A* **138** 183–193.

[11] Kesten, H. (1996). Detecting a single defect in a scenery by observing the scenery al a random walk path. *Itô's Stochastic Calculus and Probability Theory* 171–183.Springer, Tokyo.

[12] Kesten, H. (1998). Distinguishing and reconstructing sceneries from observations alon random walk paths. *Microsurveys in Discrete Probability (Princeton, NJ, 1997)* 75–83. Amer. Math. Soc., Providence, RI.

[13] Karatzas, I. and Shreve, S. E. (1991). *Brownian Motion and Stochastic Calculus* **113**. Springer-Verlag, New York.

[14] Lindenstrauss, E. (1999). Indistinguishable sceneries. *Random Structures Algorithms* **14** 71–86.

[15] Lenstra, A. and Matzinger, H. (2001). Reconstructing a 4-color scenery by observing it along a recurre random walk path with unbounded jumps. Eurandom. In preparation.

[16] Löwe, M. and Matzinger, H. (2002). Scenery reconstruction in two dimensions with many colors. *Ann. Appl. Probab.* **12** 1322–1347.

[17] Lember, J. and Matzinger, H. (2003). Reconstructing a 2-color scenery by observing it along a recurre random walk path with bounded jumps. Eurandom. In preparation.

[18] Löwe, M., Matzinger, H. and Merkl, F. (2001). Reconstructing a multicolor random scenery seen along a random wal path with bounded jumps. Eurandom Report 2001-030. Submitted.

[19] Löwe, M. and Matzinger, H. (2003). Reconstruction of sceneries with correlated colors. *Stochastic Process. Appl.* **105** 175–210.

[20] Levin, D. and Peres, Y. (2002). Random walks in stochastic scenery on $\mathbb{Z}$. Preprint

[21] Levin, D. A., Pemantle, R. and Peres, Y. (2001). A phase transition in random coin tossing. *Ann. Probab.* **29** 1637–1669.

[22] Matzinger, H (1999). *Reconstructing a 2-color scenery by observing it along a simp random walk path with holding.* Ph.D. thesis, Cornell University.

[23] Matzinger, H. (1999b). Reconstructing a three-color scenery by observing it along a simpl random walk path. *Random Structures Algorithms* **15** 196–207.

[24] H. Matzinger. Reconstructing a 2-color scenery by observing it along a simple random walk path. Submitted, 2003.

[25] Matzinger, H. and Rolles, S. W. W. (2003). Reconstructing a random scenery observed with random errors along a random walk path, *Probability Theory and Related Fields* **125** 539–577.

[26] Matzinger, H. and Rolles, S. W. W. (2002). , Reconstructing a piece of scenery with polynomially many observations *Stochastic Processes and their Applications* **107** 289–300.

[27] Williams, D. (1991). *Probability with Martingales.* Cambridge University Press, Cambridge.

Chapter 11

Large deviation based upper bounds for the LCS-problem

(submitted)

By Raphael Hauser, Servet Martinez and Heinrich Matzinger

Let $X := (X_1, \ldots, X_n)$ and $Y := (Y_1, \ldots, Y_n)$ be two finite sequences. Let L_n designate the length of the longest sequence which occurs as a subsequence of X as well as of Y. We analyze and apply a large deviation and Montecarlo simulation based method for the computation of improved upper bounds on the Chvàtal-Sankoff constant γ, which is defined by the limit $\gamma = \lim_{n\to\infty} \mathbb{E}[L_n]/n$ when X and Y are random sequences with i.i.d. entries. Our theoretical results show that this method converges to the exact value of γ when a control parameter m converges to infinity. We also give upper bounds on the complexity for numerically computing γ to any given precision via this method. Our numerical experiments confirm the theory and allow us to give new upper bounds that are correct to two digits.

11.1 Introduction

The investigation of longest common subsequences (LCS) of two finite words is one of the main problems in the theory of pattern matching and plays a role in DNA- and Protein-alignments, file-comparison, speech-recognition and so forth.

Let $X := (X_1, \ldots, X_n)$ and $Y := (Y_1, \ldots, Y_n)$ be two independent randomly generated sequences with uniform i.i.d. entries from a finite alphabet $A = \{1, 2, \ldots, C\}$. In the simplest case the entries of X and Y are just i.i.d. Bernoulli variables with parameter $1/2$. Let L_n designate the length of a longest common subsequence of X and Y, that is, a sequence which occurs as a subsequence of both X and Y and which is of maximal length among all sequences with this property. The thus defined random variable L_n and several of its variants have been studied intensively by probabilists, computer-scientists and mathematical biologists; for applications of LCS-algorithms in biology see Waterman [26]. The books of Sankoff-Kruskal [23, 20], Capocelli [12, 13] and Apostolico-Crochemore-Galil-Manbar

[3] present further applications.

Using a subadditivity argument, Chvàtal-Sankoff [14] prove that the limit

$$\gamma := \lim_{n\to\infty} \mathbb{E}[L_n]/n$$

exists. The exact value of γ remains however unknown. Chvàtal-Sankoff [14] derive upper and lower bounds for γ, and similar upper bounds were found by Baeza-Yates, Gavalda, Navarro and Scheihing [10] using an entropy argument. These bounds have been improved by Deken [17], and subsequently by Dancik-Paterson [15, 22]. In this paper we present a Monte Carlo and large deviation based method which allows to further improve the upper bounds on γ. Our approach can be seen as a generalization of the method of Dancik-Paterson.

The most widely used method for the comparison of genetic data is a generalization of the LCS-method. For an excellent overview of this subject see Waterman-Vingron [28]. In this generalization a maximal score is sought over the set of all possible alignments of the two sequences, where gaps are penalized with a fixed parameter $\delta > 0$ and mismatches are penalized by a fixed amount $\mu > 0$: consider for example the two words "brot" and "bat". One possible alignment $\mathbb{A}$ of these words is

b	r	o	t
b	a	$-$	t

The score of this alignment is $1 - \mu - \delta + 1 = S(\mathbb{A})$. The matching pairs of letters "b" and "t" are each valued with a weight of 1. The gap $-$ in "bat" after the "a" costs $-\delta$. Furthermore, the mismatch between "r" and "a" is penalized by adding $-\mu$ to the total score. If $M_{\mu,\delta}(X,Y)$ denotes the maximal score amongst all possible alignments of two words X and Y, and if $M_n(\mu,\delta)$ is the random variable defined by $M_n(\mu,\delta) = M_{\mu,\delta}(X,Y)$, where X and Y are two i.i.d. random sequences of length n, then the LCS-problem is a special case of the investigation of $M_n(\mu,\delta)$, because $L_n = M_n(\infty,0)$. Generalizing the arguments from the LCS-problem, one can prove that the limit

$$a(\mu,\delta) = \lim_{n\to\infty} \frac{\mathbb{E}[M_n]}{n}$$

exists. Arratia-Waterman [8] showed that there is a phase transition phenomenon defined by critical values of μ and δ. In one phase M_n is of linear order in n, whereas in the other it is logarithmically small in n. Waterman [27] conjectures that the deviation of M_n from its mean behaves like $\sqrt{n}$.

As mentioned earlier, the approach we use in this paper to derive upper bounds on γ is inspired by the method of Dancik-Paterson [15, 22]. However, in contrast to the latter, our method can be used in principle to derive upper bounds on $a(\mu,\delta)$ for values of μ and δ that correspond to the linear phase. This is a subject we plan to pursue in future research.

Let us mention a few further details on the history of these problems and the state of knowledge about them: Waterman-Arratia [8] derive a law of large deviation for L_n for fluctuations on scales larger than $\sqrt{n}$. The order of magnitude of the deviation from the

mean of L_n is unknown, and in fact it is not even known if these deviations are larger than a power of n. However, using first passage percolation methods, Alexander [2] proves that $\mathbb{E}[L_n]/n$ converges at a rate of order $\sqrt{\log n/n}$.

Waterman [27] studies the statistical significance of the results produced by sequence alignment methods. An important problem that was open for decades concerns the longest increasing subsequence (LIS) of random permutations and appears to be related to the LCS-problem. However, it is an open question to know if solutions of the LIS-problem can be used to study the LCS problem, see Johansson [11] and Aldous-Diaconis [1].

Another problem related to the LCS-problem is that of comparing sequences X and Y by looking for longest common words that appear both in X and Y, and generalizations of this problem where the word does not need to appear in exactly the same form in the two sequences. The distributions that appear in this context have been studied by Arratia-Gordon-Goldstein-Waterman [4] and Neuhauser [21]. A crucial role is played by the Chen-Stein Method for the Poisson-Approximation. Arratia-Gordon-Waterman [5, 6] shed some light on the relation between the Erdös-Rényi law for random coin tossing and the above mentioned problem. In [7] the same authors also developed an extreme value theory for this problem.

11.2 Overview

As mentioned above, Dancik-Paterson [15, 22] derived the best deterministic bounds on the Chvàtal-Sankoff constant γ, that is, the numbers they derive are analytically proven to be lower and upper bounds on γ respectively.

The results presented here are fundamentally different: we will derive a randomized algorithm that produces an upper bound $\hat{q}$ on γ at a given confidence level. For example, on the 95% level this means that $\mathbb{P}[\hat{q} > \gamma] \geq 0.95$. Thus, $\hat{q}$ is a random variable and a bound that is *not deterministic* but *probabilistic*. Moreover, $\hat{q}$ depends on the number l_0 of simulations and on a control parameter m whose role is further described below. For now it suffices to know that in each of the l_0 simulations we need to evaluate the length of the LCS of two random sequences of length $O(m)$ via the Wagner-Fischer algorithm [24] and collect certain information that is obtained "for free" from intermediate results during the computation. In our theoretical analysis we then show that $\hat{q}$ is a consistent estimator of γ, that is, $\lim_{m,l_0\to\infty} \hat{q} = \gamma$ almost surely. In fact, we show that asymptotically $\mathbb{P}[\gamma < \hat{q} < \gamma + \Xi] \geq \Lambda$ where $\Xi = O(m^{-\frac{\alpha}{2}})$ and $\Lambda = 1 - O(l_0^{-1})$, where $\alpha \in (0,1)$ is a constant.

Ours are not the first results on simulated bounds that are consistent estimators of γ: Alexander [2] described a method that turns Montecarlo estimates $\bar{L}_n/n$ of $\mathbb{E}[L_n]/n$ into consistent upper and lower bounds of γ. Again, these bounds depend on the number l_0 of simulations and on the control parameter n, and it is the case that $\lim_{n,l_0\to\infty} \hat{q} = \gamma$ almost surely. Moreover, the midpoint $\hat{\gamma}_n$ between the upper and lower bounds determined by this method satisfies $\mathbb{P}[|\hat{\gamma}_n - \gamma| < \Xi] \geq \Lambda$ where $\Xi = O(n^{-1/2}) + O(l_0^{-1/2})$ and $\Lambda = 1 - \exp\big(-(O(n) + O(l_0))\big)$.

¿From a big-picture viewpoint the two methods thus appear to have similar properties. Note however that the above-mentioned convergence rates are asymptotic worst-case bounds obtained by analytic means and do not necessarily accurately describe the practical convergence behavior. There are therefore at least two strong motivations for analyzing the new approach:

(i) The new method is conceptually very different from Alexander's approach. This opens up a new class of algorithms with possible extensions to other related problems, in particular those appearing in connection with scoring functions in bioinformatics.

(ii) Practical versions of our algorithm converge orders of magnitude faster than the theoretical analysis predicts: with $m = 1000$ our method finds substantially tighter upper bounds on γ than Alexander's approach yields with $n = 50000$. Since the dominant work per simulation is due to an application of the Wagner-Fischer algorithm, the per-simulation complexity of our algorithm is $O(m^2)$ whereas that of Alexander's method is $O(n^2)$. Thus, from a practical point of view our method constitutes a considerable improvement.

11.3 Some Useful Notation and a Key Inequality

Let A be a finite alphabet and $A^* = \bigcup_{n\in\mathbb{N}} A^n$ be the set of finite words. We denote by $|A|$ the cardinal number of A, that is the number of symbols of the alphabet. For $a \in A^*$ denote by $|a|$ its length, that is, the number of letters in a. Trivially, $|ab| = |a| + |b|$ for every pair $(a, b) \in A^* \times A^*$, where ab denotes the concatenation of a and b, that is, the string consisting of the letters of a followed by those of b.

Let Π^n be the class of increasing sequences of $\{1, \ldots, n\}$. We denote the cardinality of any $\pi \in \Pi^n$ by $|\pi|$, and its consecutive components by $\pi(i)$ $(i = 0, \ldots, |\pi|$. For $a \in A^*$ and $\pi \in \Pi^{|a|}$ we use the notation $a_\pi := (a_{\pi(i)} : i = 1, \ldots, |\pi|)$. The main object of study in this paper is the quantity

$$L(a, b) = \max\{k : \exists \pi \in \Pi^{|a|}, \sigma \in \Pi^{|b|}, |\pi| = k = |\sigma|, a_\pi = b_\sigma\},$$

that is, $L(a, b)$ is the length of a longest common subsequence of a and b.

For the analyis it is convenient to use the set of elementary events $\Omega = A^{\mathbb{N}} \times A^{\mathbb{N}}$ endowed with the canonical product σ-algebra. We will also sometimes identify Ω with $(A \times A)^{\mathbb{N}}$, and we denote the points of Ω by $\omega = (x, y)$, where $x = (x_n : n \in \mathbb{N})$ and $y = (y_n : n \in \mathbb{N})$. We use the following notation for the canonical projections defined on Ω: $X(\omega) = x$, $X_i(\omega) = x_i$, $Y(\omega) = y$ and $Y_j(\omega) = y_j$.

We endow Ω with a probability measure $\mathbb{P} = \mathbf{P} \times \mathbf{P}$, where $\mathbf{P}$ is a *Bernoulli measure* on $A^{\mathbb{N}}$, that is, $\mathbf{P} = \xi^{\mathbb{N}}$ where ξ is a probability distribution on the finite alphabet A with $\xi(a) > 0$ for all $a \in A$. In other words yet, all entries in X and Y are i.i.d. random variables with values in A and distribution ξ.

Remark 11.3.1. *It is interesting to note that some of the results presented in this paper extend to the situation where $\mathbb{P}$ is a ergodic shift-invariant measure on Ω. For example, the proof of relation* (11.3.1) *below goes through unchanged, and the argument we will present in* (11.3.3) *extends to the more general probability model if Birkhoff's Ergodic Theorem is invoked. However, since most of the results we present in this paper rely on $\mathbb{P}$ being a Bernoulli measure, and since this is the model of interest in the vast majority of applications, we decided keep to this slightly more restrictive framework.*

Let $x[i,j] = (x_k : i \leq j)$ be the word formed by the letters between i-th and j-th coordinate on x. We use the same notation for words in y and for random vectors, that is, we write for example $X[i,j] = (X_k : i \leq k \leq j)$. Any pair of words $(a,b) \in A^* \times A^*$ defines a measurable set as follows,

$$[[a,b]] = \{(x,y) \in \Omega : x[1,|a|] = a, y[1,|b|] = b\}.$$

Extending this notation, we write $[[S]] = \cup_{(a,b)\in S}[[a,b]]$ for all $S \subseteq A^* \times A^*$.

Let $\left\{L_j^i : \Omega \to \mathbb{N} | i,j \in \mathbb{N}\right\}$ be the family of random variables

$$L_j^i = \begin{cases} L(X[i,j], Y[i,j]) & \text{if } i \leq j, \\ 0 & \text{otherwise.} \end{cases}$$

For ease of notation we will write L_j for L_j^1. Then $\{L_j^i\}$ satisfies the hypotheses of Kingman's subadditive ergodic theorem, which implies that

$$\inf_{n\geq 1} \frac{L_n}{n} = \lim_{n\to\infty} \frac{L_n}{n} = \lim_{n\to\infty} \frac{\mathbb{E}[L_n]}{n} := \gamma \tag{11.3.1}$$

holds $\mathbb{P}$ almost everywhere on Ω for some real number γ, see e.g. [19]. The limit γ, trivially seen to be lying in the interval $(0,1)$, is called the *Chvàtal-Sankoff constant* associated with the law $\mathbb{P}$. It follows from (11.3.1) that for any $q < \gamma$ it is true that $\lim_{n\to\infty} \mathbb{P}\{L_n \geq qn\} = 1$. Therefore, for all $q \in (0,1)$,

$$\lim_{n\to\infty} \mathbb{P}\{L_n \geq qn\} < 1 \Rightarrow q \geq \gamma. \tag{11.3.2}$$

We write $S_1^n(q) := \{(a,b) \in A^n \times A^n : L(a,b) \geq qn\}$. Note that

$$\{L_n \geq qn\} = [[S_1^n(q)]] = \cup_{(a,b)\in S_1^n(q)}[[a,b]].$$

This notation will be useful in the proof of Lemma 11.3.1.

It will sometimes be necessary to have a lower bound for γ. An elementary relation is obtained as follows,

$$\gamma \overset{\text{a.s.}}{=} \lim_{n\to\infty} \frac{L_n}{n} \geq \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \mathbf{1}_{X_k = Y_k} \overset{\text{a.s.}}{=} \sum_{a\in A} \mathbf{P}([[a]])^2 \geq |A|^{-1}. \tag{11.3.3}$$

The following definition introduces one of the key concepts upon which our methods rely:

Definition 11.3.1. *For any word $a \in A^*$ of length $|a| \geq 1$ let $a^- := (a_1, .., a_{|a|-1})$ be the word obtained by removing the last letter from a. For $m \in \mathbb{N}$ we say that a pair $(a, b) \in A^* \times A^*$ is a m-match if*

$$\begin{aligned} L(a, b) &= m, \\ L(a^-, b) &= m - 1 \\ L(a, b^-) &= m - 1. \end{aligned}$$

We write $\mathcal{M}^m$ for the set of m-matches in $A^* \times A^*$.

It follows immediately from Definition 11.3.1 that

$$(a, b) \in \mathcal{M}^m \Rightarrow \min\{|a|, |b|\} \geq m, \tag{11.3.4}$$

$$(X[1, i], Y[1, j]) \in \mathcal{M}^m, k \neq j \Rightarrow (X[1, i], Y[1, k]) \notin \mathcal{M}^m. \tag{11.3.5}$$

The last relation holds point-wise on Ω and says that, for a given $i \in \mathbb{N}$ there is at most one index j such that $(X[1, i], Y[1, j])$ is a m-match.

The following family of random variables will play an important role throughout all parts of this paper:

$$\begin{aligned} L_{i,j} &= L(X[1, i], Y[1, j]), \\ Z_{i,j}^{[m]} &= Z_{i,j} = \mathbf{1}_{\mathcal{M}^m}(X[1, i], Y[1, j]), \\ Z_k^{[m]} &= Z_k = \sum_{(i,j):i+j=k} Z_{i,j}. \end{aligned}$$

We will often use the simplified notation $Z_{i,j}, Z_k$ in contexts where we treat m as a fixed parameter. It follows immediately from (11.3.4) and (11.3.5) that $0 \leq Z_k \leq (k - 2m)_+$, that is, $Z_k = 0$ everywhere on Ω for $k < 2m$. Associated with the variables Z_k is the following measure on $\mathbb{N}$ which will play a key role throughout our analysis: we set

$$\nu^{[m]}(k) = \nu(k) = \mathbb{E}[Z_k] \tag{11.3.6}$$

for all $k \in \mathbb{N}$, and ν is then extended to $\mathbb{N}$ by σ-additivity. Note that the definitions of $Z_{i,j}, Z_k$ and ν all depend on the choice of the parameter m. In order to avoid index cluttering we chose not to account for this dependence explicitly in the notation. This should not lead to confusion, but the reader should bear the dependence on m in mind. Let us mention that, although we cannot exclude at this point that ν be an infinite measure, we will later prove that it is finite because $\nu(\mathbb{N}) \leq |A|m$, see Lemma 11.4.3. However, ν is of course generally not a probability measure. A trivial identity which is sometimes useful is the following,

$$\nu(k) = \sum_{i+j=k} \mathbb{P}\{L_{i,j} = m\}. \tag{11.3.7}$$

We are ready to prove one of a key inequality that drives our approach:

Lemma 11.3.1. *Let $m \in \mathbb{N}$, $q \in [0,1]$, and let $\nu^{*\lfloor qn/m \rfloor}$ be the measure ν, defined in* (11.3.6)*, convoluted $\lfloor qn/m \rfloor$ times with itself. Then*

$$\mathbb{P}\{L_{n-1} \geq qn\} \leq \sum_{l_1+\cdots+l_{\lfloor qn/m \rfloor} \leq 2n} \nu(l_1) \cdot \cdots \cdot \nu(l_{\lfloor qn/m \rfloor}) = \nu^{*\lfloor qn/m \rfloor}\big([0,2n]\big). \tag{11.3.8}$$

Proof. Let us consider the class of words

$$S_2^n(q) = \cup_{(i,j):i+j=2n}\{(a,b) \in A^i \times A^j : L(a,b) \geq qn\}.$$

It is clearly the case that $S_1^n(q) \subseteq S_2^n(q)$. Let

$$S_3^{n,m}(q) := \Big\{\big(a^1 \dots a^{\lfloor qn/m \rfloor} c^1, b^1 \dots b^{\lfloor qn/m \rfloor} c^2\big) : (a^k, b^k) \in \mathcal{M}^m (k = 1, \dots, \lfloor qn/m \rfloor),$$
$$c^1, c^2 \in A^*, \ \sum_{k=1}^{\lfloor qn/m \rfloor} |a^k b^k| + |c^1 c^2| = 2n\Big\}.$$

We claim that $S_2^n(q) \subseteq S_3^{n,m}(q)$. In fact, for any pair $(a,b) \in S_2^n(q)$, there exist two strictly increasing maps $\pi : [1, \lceil qn \rceil] \to [1, |a|]$ and $\sigma : [1, \lceil qn \rceil] \to [1, |b|]$ such that $a_\pi = b_\sigma$, and it is possible to choose π and σ minimal in the sense that for each pair $(\hat{\pi}, \hat{\sigma}) \in \Pi^{|a|} \times \Pi^{|b|}$ that satisfies

$$\begin{aligned} |\pi| = |\sigma| &= \lceil qn \rceil, \\ a_{\hat{\pi}} &= b_{\hat{\sigma}}, \\ \hat{\pi}(k) \leq \pi(k) &\quad (k = 1, \dots, \lceil qn \rceil), \\ \hat{\sigma}(k) \leq \sigma(k) &\quad (k = 1, \dots, \lceil qn \rceil), \end{aligned}$$

we have $\hat{\pi} = \pi$ and $\hat{\sigma} = \sigma$. It is easy to see that when π and σ are minimal in this sense, then

$$(a^k, b^k) := \big(a_{\pi(m(k-1)+1)} \dots a_{\pi(mk)}, b_{\sigma(m(k-1)+1)} \dots b_{\sigma(mk)}\big) \in \mathcal{M}^m$$

for $k = 1, \dots, \lfloor qn/m \rfloor$. Therefore, $\big(a^1 \dots a^{\lfloor qn/m \rfloor} c^1, b^1 \dots b^{\lfloor qn/m \rfloor} c^2\big) \in S_3^n(q)$, where $c^1 := a_{\pi(\lfloor qn/m \rfloor + 1)} \dots a_{|a|}$ and $c^2 := b_{\sigma(\lfloor qn/m \rfloor + 1)} \dots b_{\sigma(|b|)}$. This shows that $S_2^n(q) \subseteq S_3^{n,m}(q)$, as claimed.

It is now useful to introduce the index set

$$\mathcal{I}(q,n,m) = \Big\{\vec{l} := (l_1, \dots, l_{\lfloor qn/m \rfloor}) \in \mathbb{N}^{\lfloor qn/m \rfloor} : \sum_{k=1}^{\lfloor qn/m \rfloor} l_k \leq 2n\Big\}. \tag{11.3.9}$$

With any element $\vec{l} \in \mathcal{I}(q,n,m)$ we associate the set

$$S_3^{n,m}(q,\vec{l}) := \Big\{ \big(a^1 \dots a^{\lfloor qn/m \rfloor} c^1, b^1 \dots b^{\lfloor qn/m \rfloor} c^2\big) \in S_3^{n,m}(q) : \\ |a^k b^k| = l_k, (k = 1, \dots, \lfloor qn/m \rfloor) \Big\}.$$

It is then clearly the case that

$$S_3^{n,m}(q) = \bigcup_{\vec{l} \in \mathcal{I}(q,n,m)} S_3^{n,m}(q,\vec{l}),$$

and hence that

$$\mathbb{P}([[S_3^{n,m}(q)]]) \leq \sum_{\vec{l} \in \mathcal{I}(q,n,m)} \mathbb{P}([[S_3^{n,m}(q,\vec{l})]])$$

which in turn implies

$$\begin{aligned} \mathbb{P}([[S_3^{n,m}(q)]]) &\leq \sum_{\vec{l} \in \mathcal{I}(q,n,m)} \sum_{S_3^n(q,\vec{l})} \mathbb{P}([[(a^1 \dots a^{\lfloor qn/m \rfloor} c^1, b^1 \dots b^{\lfloor qn/m \rfloor}, c^2)]]) \\ &\leq \sum_{\vec{l} \in \mathcal{I}(q,n,m)} \sum_{(a^k,b^k) \in \mathcal{M}^m : |a^k b^k| = l_k, (k=1,\dots,\lfloor qn/m \rfloor)} \prod_{k=1}^{\lfloor qn/m \rfloor} \mathbb{P}([[a^k, b^k]]), \end{aligned}$$

where the last inequality follows from the assumption that $\mathbb{P}$ is a Bernoulli measure and from the trivial inequality $\mathbb{P}([[c^1, c^2]]) \leq 1$. Now, since

$$\nu^{*\lfloor qn/m \rfloor}([0,2n]) = \sum_{\vec{l} \in \mathcal{I}(q,n,m)} \prod_{k=1}^{\lfloor qn/m \rfloor} \nu(l_k), \tag{11.3.10}$$

and

$$\prod_{k=1}^{\lfloor qn/m \rfloor} \nu(l_k) = \sum_{(a^k,b^k) \in \mathcal{M}^m : |a^k b^k| = l_k, (k=1,\dots,\lfloor qn/m \rfloor)} \prod_{k=1}^{\lfloor qn/m \rfloor} \mathbb{P}([[a^k, b^k]]),$$

we can conclude that

$$\mathbb{P}([[S_3^{n,m}(q)]]) \leq \nu^{*\lfloor qn/m \rfloor}([0,2n]). \tag{11.3.11}$$

Finally, since $\{L_n \geq qn\} = [[S_1^n(q)]] \subseteq [[S_2^n(q)]] \subseteq [[S_3^{n,m}(q)]]$, the proof is complete. □

11.4 A Large Deviation Based Upper Bound on γ

In this section we will apply large deviation techniques to find the exponential rate of the bound on the right hand side of (11.3.11). Since ν is not a probability measure in general, we will derive the relevant inequalities without using the classical results stated for probability distributions. Using the usual measure theoretic notation, we have

$$\Big(\int_{\mathbb{N}} e^{t\left(\frac{2m}{q} - x\right)} d\nu(x) \Big)^{\lfloor qn/m \rfloor} = \sum_{(l_1,\dots,l_{\lfloor qn/m \rfloor}) \in \mathbb{N}^{\lfloor qn/m \rfloor}} e^{t \sum_{k=1}^{\lfloor qn/m \rfloor} \left(\frac{2m}{q} - l_k\right)} \prod_{k=1}^{\lfloor qn/m \rfloor} \nu(l_k).$$

Since every $(l_1, \ldots, l_{\lfloor qn/m \rfloor}) \in \mathcal{I}(q,n,m)$ satisfies $\sum_{k=1}^{\lfloor qn/m \rfloor}(2m/q - l_k) \geq -2m/q$, (11.3.10) implies

$$\nu^{*\lfloor qn/m \rfloor}\big([0,2n]\big) \leq \Big(\int_{\mathbb{N}} e^{t\left(\frac{2m}{q}-x\right)} d\nu(x)\Big)^{\lfloor qn/m \rfloor} e^{\frac{2mt}{q}}. \tag{11.4.1}$$

This leads to the following theorem, providing the main tool for the construction of our upper bounds on γ:

Theorem 11.4.1. *Let $t > 0$ and $q \in [0,1]$. If*

$$\sum_{k\in\mathbb{N}} e^{t\left(\frac{2m}{q}-k\right)}\nu(k) < 1 \tag{11.4.2}$$

then $\gamma < q$.

Proof. If (11.4.2) holds then for all n large enough the right hand side of (11.4.1) is < 1. The result then follows from (11.3.2) and (11.3.8). □

Let us now define

$$q_1(m) := \inf\Big\{q \in [0,1] :\ \exists t > 0 \text{ s.t. } \sum_{k\in\mathbb{N}} e^{t\left(\frac{2m}{q}-k\right)}\nu(k) < 1\Big\}. \tag{11.4.3}$$

By Theorem (11.4.1) we it is then true that $\gamma \leq q_1(m)$ for all $m \in \mathbb{N}$. In the remainder of this section, culminating in Theorem 11.4.3 below, we will show that $\lim_{m\to\infty} q_1(m) = \gamma$. The analysis that leads to this result also sets the stage for understanding the practical Montecarlo methods to compute $q_1(m)$ devised in Section 11.5. We start by recalling the following large-deviation inequality:

Lemma 11.4.1 (Azuma-Hoeffding). *Let $t \in \mathbb{N}$, $\mathcal{F} = \cup_{s\in\mathbb{N}_0}\mathcal{F}_s$ a filtration and $V_0, V_1, \ldots, V_t$ a $\mathcal{F}$-adapted martingale such that $V_0 = 0$. Let $a > 0$ and $\Delta > 0$, and let us assume that for all $s \in [0, t-1]$ it is the case that $|V_t - V_{t+1}| \leq a$ a.s. Then the following inequality holds true,*

$$\mathbb{P}\big\{V_t \geq \Delta t\big\} \leq e^{-\frac{t\Delta^2}{2a^2}}$$

Proof. This result is due to Azuma [9] and Hoeffding [18]. A modern proof can be found for example in [25], Section 11.1.4. □

We will now use Lemma 11.4.1 to show that $L_{i,j}$ decays exponentially:

Lemma 11.4.2. *For all $\Delta \geq 0$ it is true that $\mathbb{P}\{L_{i,j} \geq \frac{i+j}{2}(\gamma+\Delta)\} \leq e^{-(i+j)\Delta^2/8}$.*

Proof. We have $L_{i+j,j+i} \geq L_{i,j} + L_{j,i} \circ (\sigma_X^i, \sigma_Y^j)$, where σ_X and σ_Y denote the left-shift operators on the X and Y components of (X,Y) respectively. Since $\mathbb{P}$ is a Bernoulli measure, $L_{i,j}$ and $L_{j,i} \circ (\sigma_X^i, \sigma_Y^j)$ are identically distributed, so that $\mathbb{E}[L_{i+j,j+i}] \geq 2\mathbb{E}[L_{i,j}]$. It follows from subadditivity that $\mathbb{E}[L_{i+j,j+i}] \leq \gamma(i+j)$, implying $\mathbb{E}[L_{i,j}] \leq \gamma(i+j)/2$ and hence,

$$\mathbb{P}\Big\{L_{i,j} \geq \frac{i+j}{2}(\gamma+\Delta)\Big\} \leq \mathbb{P}\Big\{L_{i,j} \geq \mathbb{E}[L_{i,j}] + \frac{(i+j)}{2}\Delta\Big\}. \tag{11.4.4}$$

Let us next consider a fixed path $\Gamma : \{0,\ldots,i+j\} \to \mathbb{Z}^2$ that leads from $\Gamma(0) = (0,0)$ to $\Gamma(i+j) = (i,j)$ by moving one unit in the positive direction of either coordinate in each step. Let $r(k)$ and $s(k)$ be defined by $G(k) = \big(r(k), s(k)\big)$, let $\mathcal{F}_0 = \{\mathbb{R}, \varnothing\}$ be the trivial σ-algebra on $\mathbb{R}$, and let

$$\mathcal{F}_k = \sigma\left(X_u, Y_v : u = 1,\ldots,r(k); v = 1,..,s(k)\right), \quad (k = 1,\ldots,i+j).$$

(Here and elsewhere the notation extends in a natural way to the case where an index set is empty. For example, if $r(k) = 0$ then $\mathcal{F}_k = \sigma\big(Y_1,\ldots,Y_{s(k)}\big)$.) For $k \in \{0,\ldots,i+j\}$ let us define $V_k := \mathbb{E}\big[L_{i,j} - \mathbb{E}[L_{i,j}]\big\|\mathcal{F}_k\big]$.

The sequence $V_0, V_1, \ldots, V_{i+j}$ is then a martingale that satisfies the conditions of Lemma 11.4.1 with $a = 1$. Applying the lemma, we obtain the inequality

$$\mathbb{P}\left(L_{i,j} - \mathbb{E}[L_{i,j}] \geq \frac{(i+j)}{2}\Delta\right) \leq e^{-(i+j)\Delta^2/8}.$$

Combined with (11.4.4) this yields the result. □

Remark 11.4.1. *Applying the Azuma-Hoeffding Lemma to the martingale $(-V_0,\ldots,-V_{i+j})$, where V_k is as in the proof of Lemma 11.4.2, one finds the inequality*

$$\mathbb{P}\left(L_{i,j} - \mathbb{E}[L_{i,j}] \leq -\frac{(i+j)}{2}\Delta\right) \leq e^{-(i+j)\Delta^2/8}. \tag{11.4.5}$$

As a consequence of Lemma 11.4.2 we can now bound $\nu(k)$ for small k:

Corollary 11.4.1. *Let $k \leq 2m/\gamma$ and $\Delta'_k = (2m/k) - \gamma$. Then*

$$\nu(k) \leq 2m|A|e^{-(\Delta'_k)^2 k/8}.$$

Proof. Let us consider a pair (i,j) such that $k := i + j \leq 2m/\gamma$. Then $\Delta'_k \geq 0$, and we can apply Lemma 11.4.2 to find that

$$\mathbb{P}(L_{i,j} = m) \leq \mathbb{P}(L_{i,j} \geq m) \leq e^{-(\Delta'_k)^2 k/8}.$$

Together with (11.3.7) and (11.3.3) this proves the claim. □

As promised in Section 11.3, we will next prove that ν is a finite measure. Recall again that the definitions of $Z_{i,j}$, Z_k and ν depend on the value of the control parameter m.

Lemma 11.4.3. *For every $m \in \mathbb{N}$, it is true that*

$$\sum_{k\geq 1} \nu(k) = \mathbb{E}\left[\sum_{i,j>0} Z_{i,j}\right] \leq |A|m.$$

Proof. The sequence $(\mathcal{Z}_m : m \in \mathbb{N})$ of random variables

$$\mathcal{Z}_m := \min\{k \geq 0 : Z^{[m]}_{m,k} = 1\}$$

is strictly increasing in m. Moreover, we have $\mathcal{Z}_1 = \min\{k \geq 1 : Y_k = X_1\}$. Hence,

$$\mathbb{P}\{\mathcal{Z}_1 = k\} = \sum_{a\in A} \xi(a)(1-\xi(a))^{k-1}\xi(a), \tag{11.4.6}$$

and we find that $\mathbb{E}[\mathcal{Z}_1] = |A|$.

Next, let us set $\mathcal{Y}_0 = 0$ and $\mathcal{Y}_k = \min\{l > \mathcal{Y}_{k-1} : Y_l = X_k\}$ for $k \geq 1$. Then $\mathcal{Y}_1 = \mathcal{Z}_1$ and $\mathcal{Y}_k \geq \mathcal{Z}_k$ holds true for all $k \in \mathbb{N}$. Because $\mathbb{P}$ is a Bernoulli measure, $\mathcal{Y}_{k+1} - \mathcal{Y}_k$ is independent of $\big(\mathcal{Y}_l : l < k\big)$ and is identically distributed as $\mathcal{Y}_1$. Therefore, we have

$$\mathbb{E}[\mathcal{Z}_m] \leq \mathbb{E}[\mathcal{Y}_m] \leq m\mathbb{E}[\mathcal{Z}_1] = m|A|. \tag{11.4.7}$$

Let us now consider the random index set

$$\mathbf{M}^m = \{(i,j) \in \mathbb{N} : (X[1,i], Y[1,j]) \in \mathcal{M}^m\}$$

corresponding to the m-matches occurring in X and Y. Since $(m, \mathcal{Z}_m) \in \mathbf{M}^m$, it follows from (11.3.4), (11.3.5) and the definition of an m-match that

$$|\mathbf{M}^m| = \sum_{i,j>0} Z_{i,j} \leq \mathcal{Z}_m - m < \sum_{k=1}^{m} \mathcal{W}_k, \tag{11.4.8}$$

where $\mathcal{W}_k = \mathcal{Y}_k - \mathcal{Y}_{k-1}$ $(k = 1, \ldots, m)$ are i.i.d. random variables distributed according to (11.4.6). Therefore, by virtue of (11.4.7) we obtain

$$\mathbb{E}\big[\sum_{i,j>0} Z_{i,j}\big] \leq \mathbb{E}[\mathcal{Z}_m - m] \leq |A|(m-1),$$

proving the claim. □

Corollary 11.4.2. *For all $a \in A$ let $\eta(a) = (1-\xi(a))^{1/m}$. Then for all $k \in \mathbb{N}$ the following holds true,*

$$\nu(k) \le \big(\max_{a\in A}\eta(a)\big)^{k-2} \sum_{a\in A} m\xi(a)\frac{k-(k-1)\eta(a)}{\big(1-\eta(a)\big)^2}.$$

Proof. We use the notation and facts derived in the proof of Lemma 11.4.3. Note that $Z_k^{[m]} > 0$ implies that $k \le \mathcal{Z}_m$. Therefore,

$$\begin{aligned}
\nu(k) = \nu^{[m]}(k) &= \mathbb{E}\big[Z_k^{[m]}\big] = \sum_{r=1}^{\infty}\mathbb{P}\big(Z_k^{[m]} \ge r\big)\\
&= \sum_{s=m}^{\infty}\sum_{r=1}^{s}\mathbb{P}\big(Z_k^{[m]} \ge r \| \mathcal{Z}_m = s\big)\cdot\mathbb{P}\big(\mathcal{Z}_m = s\big)\\
&\le \sum_{s=k}^{\infty} s\mathbb{P}\big(\sum_{l=1}^{\infty}\mathcal{W}_l \ge s\big) \le \sum_{s=k}^{\infty} s\sum_{l=1}^{m}\mathbb{P}\Big(\mathcal{W}_l > \frac{s-1}{m}\Big)\\
&= \sum_{s=k}^{\infty} sm\sum_{a\in A}\xi(a)\big(1-\xi(a)\big)^{\lfloor\frac{s-1}{m}\rfloor} \le \sum_{a\in A}\frac{m\xi(a)}{\eta(a)}\sum_{s=k}^{\infty} s\eta(a)^{s-1}\\
&= \sum_{a\in A}\eta(a)^{k-2}\cdot m\xi(a)\cdot\frac{k-(k-1)\eta(a)}{\big(1-\eta(a)\big)^2}.
\end{aligned}$$

□

Our next result is instrumental in proving the consistency of the estimator $q_1(m)$:

Theorem 11.4.2. *Let $\Delta > 0$ be such that $q = \Delta + \gamma \le 1$, and let $0 < t \le \frac{\Delta}{8|A|^2}$. Then*

$$\sum_{k=1}^{\infty} e^{t(2m/q-k)}\nu(k) \le \big(m|A| + 4m^2|A|^2\big)\, e^{-t\Delta m} \tag{11.4.9}$$

Proof. It follows from the hypotheses that $1/\gamma = \Delta/(\gamma q) + 1/q$. Thus, $1/\gamma \ge \Delta + 1/q$ and

$$a := \frac{2m}{q} + m\Delta < \frac{2m}{\gamma} < 2m|A|, \tag{11.4.10}$$

where the last inequality follows from (11.3.3). We split the left hand side of (11.4.9) as follows,

$$\sum_{k=1}^{\infty} e^{t(2m/q-k)}\nu(k) = \sum_{k<a} e^{t(2m/q-k)}\nu(k) + \sum_{k\ge a} e^{t(2m/q-k)}\nu(k), \tag{11.4.11}$$

and we derive bounds on both right-hand terms separately.

To bound the second term, note that for $k \geq a$ we have $2m/q - k \leq 2m/q - a = -\Delta m$. Therefore,

$$\sum_{k \geq a} e^{t(2m/q-k)} \nu(k) \leq e^{-t\Delta m} \sum_{k \geq a} \nu(k) \leq m|A| e^{-t\Delta m}, \tag{11.4.12}$$

where the second inequality follows from the fact that Lemma 11.4.3 implies that $\sum_{k \geq a} \nu(k) \leq \sum_{k \geq 1} \nu(k) \leq m|A|$.

To bound the first term in (11.4.11), note that (11.3.4) implies $\nu(k) = 0$ for $k < 2m$. Using this in conjunction with (11.4.10) and Corollary 11.4.1 we find

$$\begin{aligned} \sum_{k < a} e^{t(2m/q-k)} \nu(k) &\leq 2m|A| \sum_{k=2m}^{a-1} e^{t(2m/q-k)} e^{-(\Delta'_k)^2 k/8} \\ &\leq 2m|A| \sum_{k=2m}^{a-1} e^{t(2m/q-k)} e^{-(\Delta'_k)^2 m/4}, \end{aligned} \tag{11.4.13}$$

where $\Delta'_k := 2m/k - \gamma$, and where the last inequality holds because $k \geq 2m$. If we now use the change of variables $\bar{k} := a - k$, then

$$\sum_{k=2m}^{a-1} e^{t(2m/q-k)} e^{-(\Delta'_k)^2 m/4} = e^{-tm\Delta} \sum_{\bar{k}=1}^{a-2m} e^{t\bar{k}} e^{-(\Delta''_{\bar{k}})^2 m/4}, \tag{11.4.14}$$

where

$$\Delta''_{\bar{k}} := \frac{2m}{a - \bar{k}} - \gamma = \frac{2m}{a} \frac{1}{(1 - \bar{k}/a)} - \gamma \geq \frac{2m}{a} - \gamma + \frac{2m\bar{k}}{a^2}.$$

Note that

$$\frac{2m}{a} - \gamma = \frac{1}{\frac{1}{q} + \frac{\Delta}{2}} - \gamma = \frac{q}{1 + \frac{\Delta q}{2}} - \gamma \geq q\Big(1 - \frac{\Delta q}{2}\Big) - \gamma \geq \Delta - \frac{\Delta q^2}{2} \geq \frac{\Delta}{2}.$$

Together with (11.4.10) this yields

$$(\Delta'_{\bar{k}})^2 \geq \left(\frac{\Delta}{2} + \frac{\bar{k}}{2m|A|^2}\right)^2 > \frac{\Delta\bar{k}}{2m|A|^2}. \tag{11.4.15}$$

Substituting (11.4.15) into (11.4.14), we get

$$\begin{aligned} \sum_{k=2m}^{a-1} e^{t(2m/q-k)} e^{-(\Delta'_k)^2 m/4} &\overset{(11.4.15)}{\leq} e^{-tm\Delta} \sum_{\hat{k}=1}^{a-2m} e^{t\bar{k} - \frac{\Delta\bar{k}}{8|A|^2}} \leq e^{-tm\Delta}(a - 2m) \\ &\overset{(11.4.10)}{<} 2m|A| e^{-tm\Delta}, \end{aligned} \tag{11.4.16}$$

where the second inequality is a consequence of the hypothesis on t. The result now follows from (11.4.11), (11.4.12), (11.4.13) and (11.4.16). □

We are finally ready to prove that $q_1(m)$ is consistent in m:

Theorem 11.4.3. *If $q_1(m)$ is defined as in (11.4.3), then*

$$\lim_{m\to\infty} q_1(m) = \gamma.$$

Proof. Because of Theorem 11.4.1 we already know that $q_1(m) \geq \gamma$ for all $m \in \mathbb{N}$. The result will thus be shown if we can prove that

$$\limsup_{m\to\infty} q_1(m) \leq \gamma. \tag{11.4.17}$$

Let $\epsilon > 0$ be fixed, and let us choose Δ and t as a function of m as follows: $\Delta := m^{-1/(2+\epsilon)}$ and $t =: \Delta/(8|A|^2)$. Then for all m large enough, the conditions of Theorem 11.4.2 are satisfied. Moreover, we have

$$e^{-t\Delta m} = e^{-\frac{\Delta^2 m}{8|A|^2}} = e^{-\frac{m^{\frac{\epsilon}{2+\epsilon}}}{8|A|^2}},$$

so that, again for m large enough, $\big(2m+4m^2|A|^2\big)e^{-t\Delta m} < 1$. Theorem 11.4.2 thus implies that there exists $m_0 \in \mathbb{N}$ such that for all $m \geq m_0$,

$$\sum_{k=1}^{\infty} e^{t(2m/q-k)}\nu(k) < 1, \tag{11.4.18}$$

and then $q_1(m) \leq \gamma + \Delta = \gamma + m^{-1/(2+\epsilon)}$ by (11.4.3), showing that (11.4.17) is indeed true. □

11.5 Montecarlo Simulation

Theorem 11.4.1 revealed that whenever $0 \leq q \leq 1$ and $t > 0$ are such that

$$\sum_{k\in\mathbb{N}} e^{t(2m/q-k)}\nu(k) < 1, \tag{11.5.1}$$

then q is an upper bound on the Chvàtal-Sankoff constant γ. When using this theoretical tool in practical calculations one faces the problem that one cannot evaluate (11.5.1) explicitly, because $\nu(k)$ is not known except for very small values of k. A practical way to get around this problem is to evaluate (11.5.1) via Montecarlo simulation.

Recall that we assumed that $\{X_i\}_{\mathbb{N}} \cup \{Y_j\}_{\mathbb{N}}$ is a family of i.i.d. random variables which take values in the finite alphabet A according to a probability law ξ. Recall also that in Section 11.3 we introduced the notation

$$Z_{i,j} = \mathbf{1}_{\mathcal{M}^m}\big(X[1,i], Y[1,j]\big), \quad Z_k = \sum_{i+j=k} Z_{i,j}, \quad \text{and} \quad \nu(k) = \mathbb{E}[Z_k].$$

Let us now define the random variable

$$W = W(t,q) := \sum_{k>0} e^{t(2m/q-k)} Z_k, \tag{11.5.2}$$

so that $\mathbb{E}[W] = \sum_{k>0} e^{t(2m/q-k)} \nu(k)$ is the expression of interest in (11.5.1).

For the purposes of Montecarlo simulation, we consider $(X_i^l : i,l \in \mathbb{N})$ and $(Y_j^l : j,l \in \mathbb{N})$, two independent collections of i.i.d. random variables with distribution ξ on A. Let us define

$$\begin{aligned} Z_{i,j}^l &:= \mathbf{1}_{\mathcal{M}^m}\big(X^l[1,i], Y^l[1,j]\big), \\ Z_k^l &:= \sum_{i+j=k} Z_{i,j}^l, \quad \text{and} \\ W^l = W^l(t,q) &:= \sum_{k>0} e^{t(2m/q-k)} Z_k^l. \end{aligned}$$

Thus, Z_k^l counts the number of m-matches of length k observed in the l-th realization (X^l, Y^l) of the pair of random sequences. Then

$$\hat{\nu}_k := \frac{1}{l_0} \sum_{l=1}^{l_0} Z_k^l$$

is an unbiased estimator of $\nu(k)$ and

$$\frac{1}{l_0} \sum_{l=1}^{l_0} W^l = \sum_{k>0} e^{t(2m/q-k)} \hat{\nu}_k \tag{11.5.3}$$

is an unbiased estimator of the left hand side of (11.5.1).

In Section 11.5.1 we will show how the estimator (11.5.3) can be used in theory to obtain an upper bound on γ to any given precision and at any given confidence level. In Section 11.5.2 we will also derive an upper bound on the number of elementary computer operations necessary to compute such a bound as a function of the required precision and confidence level. The theoretical analysis is based on estimates which are unnecessarily conservative in practice. Practical implementations are therefore based on a slightly different approach, leading to a number of issues that need careful attention. We discuss these in Section 11.5.3.

11.5.1 Montecarlo Simulation in Theory

The main result of this section is the following theorem, which gives a tool to determine the value of the control parameter m and the number l_0 of simulations necessary to obtain an estimator $\hat{q}_1$ of γ to within a specific precision and on a given confidence level:

Theorem 11.5.1. *Let $\alpha, \delta \in (0,1)$ be constants and $l_0 \in \mathbb{N}$. Let us choose Δ and t as functions of m as follows: $\Delta = m^{-\alpha/2}$ and $t_1 = \Delta/(16|A|^2)$. Let us finally consider*

$$\hat{q}_1 = \Delta + \inf\Big\{q > 0 : \sum_{k=1}^{\infty} e^{t_1(2m/q-k)}\hat{\nu}_k < 1\Big\} \tag{11.5.4}$$

as an estimator for q_1. Then there exists a number

$$m_0 = m_0(\alpha,\xi,\delta) < \max\left(O\left(\Big(\frac{2}{1-\gamma}\Big)^{\frac{2}{\alpha}}\right), O\left(\Big(|A|^4\log\frac{2}{\delta}\Big)^{\frac{1+\epsilon}{1-\alpha}}\big(1-\log\min_{a\in A}\xi(a)\big)\right)\right),$$

where ϵ is a small number, such that for all $m \geq m_0$ it is true that

$$\mathbb{P}\big(\gamma \leq \hat{q}_1 \leq \gamma + 2\Delta\big) \geq 1 - e^{-l_0(1-\delta)^2/2} - \frac{8}{l_0}. \tag{11.5.5}$$

Note that the right hand side of (11.5.5) determines the confidence level that the computed estimator $\hat{q}_1$ *is* an upper bound *and* approximates γ to within precision 2Δ. The confidence level increases if the number l_0 of simulations is increased. The precision on the other hand increases with m. α and δ merely play the role of control parameters. These could be fixed at given values, but treating them as parameters reveals how their values affect the complexity estimates of Section 11.5.2. The same pertains to the dependence of m_0 on the distribution ξ on A. Note finally that γ and $|A|$ are functions of ξ, which is why no extra variables are necessary in $m_0(\alpha,\xi,\delta)$.

Before we can prove Theorem 11.5.1 we need three preliminary results. The first lemma shows that when m is large enough, then with high probability there will be a m-match of length not much larger than what γ predicts:

Lemma 11.5.1. *Let α and Δ be as in Theorem 11.5.1 and let us consider the event*

$$B := \left\{\exists i,j \text{ s.t. } Z_{i,j} = 1 \text{ and } i+j \leq 2\left\lceil\frac{m}{\gamma} + \frac{\Delta m}{2}\right\rceil\right\}. \tag{11.5.6}$$

Then there exists a number $m_1 = m_1(\alpha,\xi)$ such that for all $m \geq m_1$,

$$\mathbb{P}(B) \geq 1 - \exp\Big(-\frac{m^{1-\alpha}|A|^{-4}}{256}\Big). \tag{11.5.7}$$

Proof. Alexander [2] showed that there exists a constant $C > 0$,independent of ξ and A, such that for all $n \geq 1$,

$$0 \leq \gamma - \frac{\mathbb{E}[L_n]}{n} \leq C\sqrt{\frac{\log n}{n}}.$$

A more explicitly quantitative version of this result can be obtained as follows: by choosing $\lambda = 2$, $\theta = 3$ in Proposition 2.4 of [2], a relaxation of Equation (2.13) in [2] shows that

$$0 < \gamma - \frac{\mathbb{E}[L_n]}{n} < 7\sqrt{\frac{\log n}{n}} \qquad \forall n \geq 16. \tag{11.5.8}$$

Let $k' = m/\gamma + \Delta m/2$ and $n' = \lceil k' \rceil$. Then $m \geq 16$ implies

$$n' \geq k' > m \geq 16. \tag{11.5.9}$$

Moreover, if

$$m \geq m_2(\alpha, \xi) := \inf\Big\{ y > 0 : \ \frac{x^{1-\alpha}}{2 \cdot 56^2 |A|^3} > \log x, \ \ \forall x \geq y \Big\}$$

then (11.3.3) implies

$$\log m < \frac{m^{1-\alpha}\gamma^3}{2 \cdot 56^2}. \tag{11.5.10}$$

Finally, if

$$m \geq m_3(\alpha, |A|) := \big(2 \cdot 56^2 |A|^3 \log(|A| + 1/2)\big)^{\frac{1}{1-\alpha}}$$

then (11.3.3) implies

$$\frac{m^{1-\alpha}\gamma^3}{2 \cdot 56^2} > \log\Big(\frac{1}{\gamma} + \frac{1}{2}\Big). \tag{11.5.11}$$

Now, for

$$m \geq m_4(\alpha, |A|) := \max\big(m_2(\alpha, |A|), m_3(\alpha, |A|)\big),$$

(11.5.10) and (11.5.11) show that

$$\begin{aligned} \log k' = \log\Big(\frac{m}{\gamma} + \frac{\Delta m}{2}\Big) &< \log m + \log\Big(\frac{1}{\gamma} + \frac{1}{2}\Big) \\ &\overset{(11.5.10),(11.5.11)}{<} \frac{m^{1-\alpha}\gamma^3}{56^2} \\ &< \frac{m^{-\alpha}\gamma^4}{56^2}\Big(\frac{m}{\gamma} + \frac{m^{-\frac{\alpha}{2}} m}{2}\Big) \\ &= \Big(\frac{\Delta\gamma^2}{8}\Big)^2 \frac{k'}{7^2}, \end{aligned} \tag{11.5.12}$$

and then (11.5.8), (11.5.9) and (11.5.12) show that for

$$m \geq m_5(\alpha, \xi) := \max\big(16, m_4\big)$$

the following holds,

$$0 < \gamma - \frac{\mathbb{E}[L_{n'}]}{n'} < 7\sqrt{\frac{\log n'}{n'}} \leq 7\sqrt{\frac{\log k'}{k'}} < \frac{\Delta\gamma^2}{8}. \tag{11.5.13}$$

Using the notation $\gamma_{n'} := \mathbb{E}[L_{n'}]/n'$, (11.5.13) and (11.3.3) imply

$$\gamma - \gamma_{n'} - \frac{\Delta\gamma^2}{4} \leq -\frac{\Delta\gamma^2}{8} \leq -\frac{\Delta |A|^{-2}}{8} \tag{11.5.14}$$

Now note that when the event $D := \big\{L_{n'} \geq m\big\}$ holds, then a m-match of total length less than or equal to n' must have occurred within $\big(X[1, n'], Y[1, n']\big)$. Hence, $D \subseteq B$, and it follows that

$$\mathbb{P}(\Omega \setminus B) \leq \mathbb{P}(\Omega \setminus D). \tag{11.5.15}$$

We have

$$\Omega \setminus D = \{L_{n'} < m\} = \Big\{ \frac{L_{n'}}{n'} - \gamma_{n'} < \frac{m}{n'} - \gamma_{n'} \Big\}. \tag{11.5.16}$$

Moreover, if

$$m \geq m_1(\alpha, \xi) := \max\big(m_5, 2^{\frac{2}{\alpha}}\big),$$

then $\Delta\gamma/2 < m^{-\frac{\alpha}{2}}/2 < 1/4$. Observe that for $x \in [0, 1/4]$ it is true that $1/(1+x) \leq 1 - x/2$. Applying this inequality to $x = \Delta\gamma/2$, we find

$$\frac{m}{n'} - \gamma_{n'} \leq \frac{m}{k'} - \gamma_{n'} = \frac{\gamma}{1 + \frac{\Delta\gamma}{2}} - \gamma_{n'} \leq \gamma - \gamma_{n'} - \frac{\Delta\gamma^2}{4}.$$

Substituting this into (11.5.16) yields

$$\mathbb{P}(\Omega \setminus D) \leq \mathbb{P}\Big(\frac{L_{n'}}{n'} - \gamma_{n'} \leq \gamma - \gamma_{n'} - \frac{\Delta\gamma^2}{4} \Big). \tag{11.5.17}$$

Combining the last inequality with (11.5.14), we find that

$$\mathbb{P}(\Omega \setminus D) \leq \mathbb{P}\Big(\frac{L_{n'}}{n'} - \frac{\mathbb{E}[L_{n'}]}{n'} \leq -\frac{\Delta |A|^{-2}}{8} \Big) \leq e^{-n' \frac{\Delta^2 |A|^{-4}}{256}}, \tag{11.5.18}$$

where the last inequality follows from (11.4.5). Since $n' > m$, we find that the bound on the right hand side of (11.5.18) is smaller than $\exp(-m\Delta^2|A|^{-4}/256)$. Finally, using $\Delta = m^{-\alpha/2}$, (11.5.15) and (11.5.18), we find that for all $m \geq m_1(\alpha, |A|)$,

$$\mathbb{P}(\Omega \setminus B) \leq \exp(-m^{1-\alpha}|A|^{-4}/256),$$

which is of course equivalent to the claimed inequality (11.5.7). □

Our second lemma shows that if m is large enough then the probability of finding an estimator value $\hat{q}_1$ significantly below γ is very small:

Lemma 11.5.2. *Let α, δ and Δ be as in Theorem 11.5.1, and let $\hat{\nu}_k := \sum_{l=1}^{l_0} Z_k^l / l_0$ for all $k \in \mathbb{N}$. Then there exists a number $m_6 = m_6(\alpha, \xi, \delta)$ such that for all $m \geq m_6$, $t \geq \Delta/(16|A|^2)$ and $q \in (0, \gamma - \Delta)$, it is true that*

$$\mathbb{P}\Big(\sum_{k=1}^{\infty} e^{t(2m/q - k)} \hat{\nu}_k < 1 \Big) \leq e^{-l_0 (1-\delta)^2/2}. \tag{11.5.19}$$

Proof. Note that $|\gamma - q| \geq \Delta$ implies $|2m/\gamma - 2m/q| \geq 2\Delta m$. Thus, when $q \leq \gamma - \Delta$, we find that $2m/q - k \geq \Delta m - 1$ for every $k \leq \lceil 2m/\gamma + \Delta m \rceil$. Hence,

$$\sum_{k=1}^{\infty} e^{t(2m/q-k)} Z_k \geq \sum_{k=1}^{\lceil 2m/\gamma + \Delta m \rceil} e^{t(2m/q-k)} Z_k \geq e^{t(\Delta m - 1)} \sum_{k=1}^{\lceil 2m/\gamma + \Delta m \rceil} Z_k \geq e^{t(\Delta m - 1)} \mathbf{1}_B, \tag{11.5.20}$$

where $\mathbf{1}_B$ denotes the indicator function of the event B defined in (11.5.6). By definition,

$$\sum_{k=1}^{\infty} e^{t(2m/q-k)}\hat{\nu}_k = \sum_{k=1}^{\infty} e^{t(2m/q-k)}\Big(\frac{1}{l_0}\sum_{l=1}^{l_0} Z_k^l\Big) = \frac{1}{l_0}\sum_{l=1}^{l_0}\sum_{k=1}^{\infty} e^{t(2m/q-k)} Z_k^l.$$

It follows therefore from (11.5.20) that

$$\mathbb{P}\Big(\sum_{k=1}^{\infty} e^{t(2m/q-k)}\hat{\nu}_k < 1\Big) \le \mathbb{P}\Big(\frac{1}{l_0}\sum_{l=1}^{l_0} e^{t(\Delta m-1)}\mathbf{1}_B^l < 1\Big). \tag{11.5.21}$$

Here, $\big(\mathbf{1}_B^l : l \in \mathbb{N}\big)$ denotes an i.i.d. sequence of copies of the random variable $\mathbf{1}_B$.

Now, for all $m \ge m_7$, where

$$m_7 = m_7(\alpha,\xi,\delta) = \big(1 + 16|A|^2\log(2/\delta)\big)^{\frac{1}{1-\alpha}},$$

we have

$$t(\Delta m - 1) \ge \frac{m^{1-\alpha} - m^{-\frac{\alpha}{2}}}{16|A|^2} > \frac{m^{1-\alpha}-1}{16|A|^2} \ge \log(2/\delta),$$

and hence, $e^{t(\Delta m-1)} > 2/\delta$. Moreover, it follows from Lemma 11.5.1 that if $m \ge m_8$, where

$$m_8 = m_8(\alpha,\xi,\delta) = \max\Big(m_1, \big(256|A|^4\log(2/\delta)\big)^{\frac{1}{1-\alpha}}\Big),$$

then

$$\mathbb{E}\big[\mathbf{1}_B^l\big] = \mathbb{P}(B) \ge 1 - \exp\Big(-\frac{m^{1-\alpha}|A|^{-4}}{256}\Big) \ge 1 - \frac{\delta}{2}.$$

Therefore, for $m \ge m_6 := \max(m_7, m_8)$ we have

$$\Big(\frac{1}{l_0}\sum_{l=1}^{l_0} e^{t_(\Delta m-1)}\mathbf{1}_B^l < 1\Big) \Rightarrow \Big(\frac{1}{l_0}\sum_{l=1}^{l_0}\mathbf{1}_B^l < \frac{\delta}{2}\Big) \Rightarrow \Big(\frac{1}{l_0}\sum_{l=1}^{l_0}\big(\mathbf{1}_B^l - \mathbb{E}[\mathbf{1}_B^l]\big) \le -(1-\delta)\Big). \tag{11.5.22}$$

Applying Lemma 11.4.1 with $a = 1$ to the martingale defined by

$$V_0 = 0, \qquad \mathcal{F}_0 = \{\varnothing, \mathbb{R}\},$$
$$V_k = \sum_{l=1}^{k}\big(\mathbb{E}[\mathbf{1}_B^l] - \mathbf{1}_B^l\big), \qquad \mathcal{F}_k = \sigma(V_1,\dots,V_k), \quad (k = 1,\dots,l_0),$$

we have

$$\mathbb{P}\Big(\frac{1}{l_0}\sum_{l=1}^{l_0}\big(\mathbb{E}[\mathbf{1}_B^l] - \mathbf{1}_B^l\big) \ge 1-\delta\Big) \le e^{-l_0(1-\delta)^2/2}.$$

Together with (11.5.21) and (11.5.22) this finishes the proof. □

The third lemma allows us to give a bound on $VAR(W)$ that will be needed in the proof of Theorem 11.5.1:

Lemma 11.5.3. *Let α and Δ be as in Theorem 11.5.1, and let $q = q(m) = \gamma + \Delta$. Then for all $m \geq m_9(\alpha, \xi) := (1-\gamma)^{-\frac{2}{\alpha}}$ and for all $t \in (0, \Delta/(16|A|^2)]$ it is true that*

$$\mathbb{E}\left[\Big(\sum_{k=1}^{\infty} e^{t_1(2m/q-k)} Z_k\Big)^2\right] \leq \left(161m^4|A|^4 + \frac{2m|A|}{\min_{a\in A}\xi(a)}\right) e^{-2t\Delta m}. \tag{11.5.23}$$

Proof. Let $a = a(m) := 2m/q + m$. (11.3.3) shows that $q > \gamma \geq |A|^{-1}$, and hence,

$$a < 2m\big(|A|+1\big). \tag{11.5.24}$$

We will use the splitting

$$\mathbb{E}\left[\Big(\sum_{k=1}^{\infty} e^{t_1(2m/q-k)} Z_k\Big)^2\right] \leq 2\mathbb{E}\left[\Big(\sum_{k\leq a} e^{t_1(2m/q-k)} Z_k\Big)^2\right] + 2\mathbb{E}\left[\Big(\sum_{k>a} e^{t_1(2m/q-k)} Z_k\Big)^2\right] \tag{11.5.25}$$

and bound both terms on the right hand side separately.

When $k > a$, we have $2m/q - k < -m < -\Delta m$, and hence,

$$\sum_{k>a} e^{t(2m/q-k)} Z_k \leq e^{-t\Delta m} \sum_{k>0} Z_k \overset{(11.4.8)}{\leq} e^{-t\Delta m} \sum_{k=1}^{m} \mathcal{W}_k,$$

where $\{\mathcal{W}_k : k = 1, \ldots, m\}$ are the i.i.d. random variables defined in the proof of Lemma 11.4.3 and whose distribution has the moment generating function

$$\Phi(s) = \sum_{a\in A} \frac{\xi(a)s}{1 - s\big(1-\xi(a)\big)}.$$

This implies that

$$\begin{aligned}
\mathbb{E}\left[\Big(\sum_{k>a} e^{t(2m/q-k)} Z_k\Big)^2\right] &\leq e^{-2t\Delta m}\mathbb{E}\left[\Big(\sum_{k=1}^{m} \mathcal{W}_k\Big)^2\right] \\
&= e^{-2t\Delta m}\left(m\mathbb{E}[\mathcal{W}_1^2] + 2\tbinom{m}{2}\mathbb{E}[\mathcal{W}_1]^2\right) \\
&= e^{-t\Delta m}\big(m\Phi''(1) - m\Phi'(1) + m(m-1)\Phi'(1)^2\big) \\
&= e^{-t\Delta m}\Big(2m\sum_{a\in A}\frac{1}{\xi(a)} - 3m|A| + m(m-1)|A|^2\Big).
\end{aligned} \tag{11.5.26}$$

Note that $t\Delta m$ is a positive power of m with our choice of t and m.

On the other hand,

$$\begin{aligned}
\Big(\sum_{k\leq a} e^{t(2m/q-k)} Z_k\Big)^2 &\leq \Big(\sum_{k\leq q} Z_k\Big)\Big(\sum_{k\leq a} e^{2t(2m/q-k)} Z_k\Big) \\
&\leq a^2 \sum_{k>0} e^{2t(2m/q-k)} Z_k,
\end{aligned} \tag{11.5.27}$$

where the last inequality follows from $Z_k \le (k-2m)_+ \le a$. Since $2t \in (0, \Delta/8|A|]$ satisfies the conditions of Theorem 11.4.2, and since $q = \gamma + \Delta \le 1$ for all $m \ge m_9$, where

$$m_9 = m_9(\alpha, \xi) := (1-\gamma)^{-\frac{2}{\alpha}},$$

(11.4.9),(11.5.24) and (11.5.27) imply

$$\mathbb{E}\left[\Big(\sum_{k\le a} e^{t(2m/q-k)} Z_k\Big)^2\right] \le 4m^3|A|\big(|A|+1\big)^2\big(1+4|A|m\big)e^{-2t\Delta m}. \tag{11.5.28}$$

Using (11.5.25), (11.5.26) and (11.5.28), the result readily follows. □

We are finally ready to give a proof of Theorem 11.5.1:

Proof. Consider the events

$$E_{m,1} := \Big\{\sum_{k=1}^{\infty} e^{t_1(2m/q-k)}\hat{\nu}_k \ge 1, \quad \forall q \in (0, \gamma-\Delta)\Big\}, \qquad \text{and}$$
$$E_{m,2} := \Big\{\sum_{k=1}^{\infty} e^{t_1(2m/(\gamma+\Delta)-k)}\hat{\nu}_k < 1\Big\}.$$

Then (11.5.4) shows $E_{m,1} \subseteq \{\gamma \le \hat{q}_1\}$ and $E_{m,2} \subseteq \{\hat{q}_1 \le \gamma + 2\Delta\}$, which implies that

$$1 - \mathbb{P}\left(\gamma \le \hat{q}_1 \le \gamma + 2\Delta\right) \le \mathbb{P}\left(\Omega \setminus E_{m,1}\right) + \mathbb{P}\left(\Omega \setminus E_{m,2}\right). \tag{11.5.29}$$

Lemma 11.5.2 provides a bound on the first term on the right hand side of this inequality, because it shows that for $m \ge m_6$,

$$\mathbb{P}\left(\Omega \setminus E_{m,1}\right) \le e^{-l_0(1-\delta)^2/2}. \tag{11.5.30}$$

To bound the second term on the right hand side of (11.5.29), let $W(t,q)$ be defined as in (11.5.2). Then $\mathbb{E}[W(t, \gamma+\Delta] = \sum_{k=1}^{\infty} e^{t_1(2m/(\gamma+\Delta)-k)}\nu(k)$. By Chebychev's inequality,

$$\begin{aligned}&\mathbb{P}\left(\Big|\sum_{k=1}^{\infty} e^{t_1(2m/(\gamma+\Delta)-k)}\hat{\nu}_k - \mathbb{E}[W(t_1, \gamma+\Delta)]\Big| \ge \frac{1}{2}\right)\\ &= \mathbb{P}\left(\Big|\frac{1}{l_0}\sum_{l=1}^{l_0} W^l(t_1, \gamma+\Delta) - \mathbb{E}[W(t_1, \gamma+\Delta)]\Big| \ge \frac{1}{2}\right) \le \frac{4\mathbb{E}[W(t_1, \gamma+\Delta)^2]}{l_0}.\end{aligned} \tag{11.5.31}$$

Note that for all $m \ge m_9$, t_1 and Δ satisfy the conditions of Theorem 11.4.2 which shows that for all $m \ge \max(m_9, m_{10})$ with

$$m_{10} = m_{10}(\alpha, \xi) := \inf\Big\{y > 0 : \big(x|A| + 4x^2|A|^2\big)e^{-\frac{x^{1-\alpha}}{16|A|^2}} \le \frac{1}{2}, \quad \forall x \ge y\Big\}$$

it is true that $\mathbb{E}[W(t_1, \gamma + \Delta)] \leq 1/2$. Likewise, Lemma 11.5.3 shows that for all $m \geq \max(m_9, m_{11})$ with

$$m_{11} = m_{11}(\alpha, \xi) := \inf\left\{ y > 0 : \left(161x^4|A|^4 + \frac{2x|A|}{\min_{a\in A}\xi(a)} \right) e^{-\frac{x^{1-\alpha}}{16|A|^2}} \leq \frac{1}{2}, \quad \forall x \geq y \right\}$$

it is the case that $\mathbb{E}[W(t_1, \gamma + \Delta)^2] \leq 1/2$. But then (11.5.31) yields

$$\mathbb{P}\left(\Omega \setminus E_{m,2}\right) \leq \frac{8}{l_0}. \tag{11.5.32}$$

The inequalities (11.5.29), (11.5.30) and (11.5.32) show that the theorem holds true for $m_0(\alpha, \xi, \delta) = \max\big(m_6, m_9, m_{10}, m_{11}\big)$. The claim on the order of m_0 as a function of α, ξ and δ is easy to check directly. □

11.5.2 Theoretical Complexity of Montecarlo Simulation

In this paragraph we will briefly discuss the computational complexity for simulating an upper bound to a given precision and at a given confidence level. The analysis has already been done in Section 11.5.1, all that remains to do is to extract the information from the results we developed there.

Let $\Lambda \in (0,1)$ be a given confidence level and let $\Xi \in (0,1)$ be a given maximum permissible error. If we wish to simulate an estimate $\hat{q}_1$ that is an upper bound on and an approximation of the Chvàtal-Sankoff constant γ to within precision Ξ at the confidence level Λ, then how large do the control parameters m and l_0 have to be chosen to guarantee such an outcome? Theorem 11.5.1 shows that if

$$m \geq \max\left(\left(\frac{2}{\Xi}\right)^{\frac{2}{\alpha}}, m_0(\alpha, \xi, \delta) \right), \qquad l_0 \geq \frac{16}{1-\Lambda}, \qquad \delta = l_0^{-\frac{1}{4}}, \tag{11.5.33}$$

then

$$\mathbb{P}\left(\gamma \leq \hat{q}_1 \leq \gamma + \Xi\right) \geq \mathbb{P}\left(\gamma \leq \hat{q}_1 \leq \gamma + 2\Delta\right) \geq 1 - e^{-l_0(1-\delta)^2/2} - \frac{8}{l_0} \geq 1 - \frac{16}{l_0} \geq \Lambda,$$

that is to say, (11.5.33) guarantees that $\hat{q}_1$ has the desired properties. Since α, ξ and δ are fixed, it is the first part of the term defining m that becomes dominant for small Ξ. The value of α that minimizes the required size of m depends on Ξ but is bounded away both from 0 and 1. Thus, if one works with the lower bounds on m and l_0 derived in (11.5.33), then

$$m = O\left(\Xi^{-\frac{2}{\alpha}}\right) \quad \text{and} \quad l_0 = O\left(\frac{1}{1-\Lambda}\right) \tag{11.5.34}$$

is required to compute an upper bound estimate $\hat{q}_1$ on the levels of accuracy Ξ and confidence Λ. This confirms what we have mentioned earlier: increasing m leads to better precision whereas increasing l_0 leads to a higher confidence level.

Let us now determine an estimate on the number of elementary computer operations required to perform the simulation of $\hat{q}_1$ with values of m and l_0 as given in (11.5.34): to

construct $\hat{q}_1$, one needs to generate l_0 independent copies $\big(X^l, Y^l\big)$ of $(X,Y) = \big((X_1, X_2, \dots), (Y_1, Y_2, \dots)\big)$. In fact, each pair has to be generated only up to the finite random length that contains the full set of m-matches defined by $\big(X^l, Y^l\big)$. Lemma 11.4.3 and equation (11.4.8) show that we can expect that only about $m|A|$ terms have to be generated for each pair (X^l, Y^l) to account for all m-matches contained in it, and Corollary 11.4.2 implies that it is exponentially rare in m that more than $O(m)$ terms need to be generated. Computing the set of all m-matches contained in a pair (X^l, Y^l) takes therefore $O(m^2)$ computer operations when the Wagner-Fischer algorithm [24] is applied. Thus, generating all the m-matches that occur in the l_0 independent copies of (X,Y) takes $O(l_0 m^2)$ time. Since each pair $\big(X^l, Y^l\big)$ contains at most $3m$ m-matches, computing $\hat{\nu}$ and $\hat{q}_1$ from the generated data takes $O(l_0 m)$ time. The overall complexity for the simulation of $\hat{q}_1$ is therefore $O\big(l_0 m^2\big)$ operations. Because of (11.5.34), the complexity of computing an upper bound on γ at the precision level Ξ and confidence level Λ is therefore

$$O\left(\frac{1}{(1-\Lambda)\cdot \Xi^{\frac{4}{\alpha}}}\right) \tag{11.5.35}$$

operations.

Note that the complexity estimate (11.5.35) is an upper bound that corresponds to the worst case scenario. The practical complexity is lower, as we will see in Section 11.5.3. Note also that we did not specify what we mean by a "computer operation". In fact, our arguments are based on the assumption that a computer can perform operations with real numbers. We do not enter a discussion of round-off and finite-precision issues here, but taking these into account it is not difficult to see that (11.5.35) is also a complexity bound in terms of floating-point operations.

11.5.3 Montecarlo Simulation in Practice

Unfortunately, the complexity bound (11.5.35) is valid only asymptotically for very large values of m, because it is assumed that $m \geq m_0(\alpha, \xi, \delta)$: in fact, if $A = \{0,1\}$ with ξ the standard Bernoulli measure characterized by $\xi(0) = 1/2 = \xi(1)$ (i.e., this is coin flipping), and if $\alpha = 0.1$ and $\delta = 0.1$ for example, then the complexity bound (11.5.35) only holds for $m \geq m m_0 = O(10^6)$, which is beyond reach in practical computations. Thus, while the complexity analysis of this section is interesting on theoretical grounds, practical methods cannot rely on it. In this section we will discuss practical methods that can achieve about two correct digits of accuracy with $m = 1000$ for the coin flipping example mentioned above. We will see that such practical methods pose a new set of challenges that need careful attention in the implementations.

Our theoretical analysis of Section 11.5.1 crucially depended on the fact that $\hat{q} \geq \gamma + \Delta$, where $\Delta \simeq O(m^{-\alpha/2})$. In essence, what we proved is that for m large enough and $\hat{q} \geq \gamma + \Delta$, the expression

$$\inf_{t>0} \sum_{k>0} e^{t(2m/\hat{q}-k)} \hat{\nu}_k \tag{11.5.36}$$

is exponentially small in m and thus very close to zero, and that the probability that $\hat{q}_1$ lies in this range is exponentially small in the number l_0 of Montecarlo simulations.

Unfortunately, for $\alpha = 0.5$ and $m = 1000$ for example, this means that $\hat{q}_1 \geq \gamma + 0.1778$, which leads to an upper bound that is not satisfactorily close to the true value of γ. Therefore, in the practical use of the method, we would like to consider estimator values $\hat{q}$ that are allowed to lie in the interval $(\gamma, \gamma + \Delta)$. In this case the expression (11.5.36) is smaller than 1 only by a small amount.

Setting the Stage for a Practical Algorithm

Since (11.5.36) is a random variable, the basic problem of the practical approach is to decide whether the sample of this variable obtained in a Montecarlo simulation lies significantly below 1 or not. An answer to this question is of course provided by Chebychev's inequality: we would like to design an estimator $\hat{q}$ such that the probability of wrongly deciding that $\hat{q}$ is an upper bound on γ is smaller than $1 - \Lambda$, that is, for the specific value of $t > 0$ used in the decision, we require that

$$\mathbb{P}\Big(\hat{q} < \gamma,\ \sum_{l=1}^{l_0} e^{t(2m/\hat{q}-k)}\hat{\nu}_k < 1\Big) \leq 1 - \Lambda \tag{11.5.37}$$

But Chebychev's inequality implies that

$$\begin{aligned}
&\mathbb{P}\Big(\hat{q} < \gamma,\ \sum_{l=1}^{l_0} e^{t(2m/\hat{q}-k)}\hat{\nu}_k < 1\Big) \leq \mathbb{P}\Big(\mathbb{E}\big[W(t,\hat{q})\big] > 1,\ \frac{1}{l_0}\sum_{l=1}^{l_0} W^l(t,\hat{q}) < 1\Big) \\
&\leq \mathbb{P}\Big(\Big|\frac{1}{l_0}\sum_{l=1}^{l_0} W^l(t,\hat{q}) - \mathbb{E}\big[W(t,\hat{q})\big]\Big| > 1 - \frac{1}{l_0}\sum_{l=1}^{l_0} W^l(t,\hat{q})\Big) \\
&\leq \frac{VAR\big(W(t,\hat{q})\big)}{l_0\Big(1 - \frac{1}{l_0}\sum_{l=1}^{l_0} W^l(t,\hat{q})\Big)^2}
\end{aligned}$$

Therefore, (11.5.37) is satisfied if

$$\frac{1}{l_0}\sum_{l=1}^{l_0} W^l(t,\hat{q}) \leq 1 - \sqrt{\frac{\hat{v}(t,\hat{q})}{(1-\Lambda)l_0}}, \tag{11.5.38}$$

where $\hat{v}(t,\hat{q})$ is an upper bound on $VAR\big(W(t,\hat{q})\big)$.

This leads to the problem of determining such an upper bound $\hat{v}(t,\hat{q})$. Note that (11.5.23) only applies for $q \geq \gamma + \Delta$ and hence is not useful in the practical context. A way out of this dilemma is to choose $\hat{v}(t,\hat{q})$ as a statistical estimator, defined in terms of the data $\{Z_k^l :\ k \geq 1, 1 \leq l \leq l_0\}$, such that

$$\mathbb{P}\left(VAR\big(W(t,\hat{q})\big) \leq \widehat{v}(t,\hat{q})\right) \geq 1 - \eta \cdot (1 - \Lambda) \tag{11.5.39}$$

for some $\eta \in (0,1)$, and to accept $\hat{q}$ as an upper bound on γ if

$$\frac{1}{l_0}\sum_{l=1}^{l_0} W^l(t,\hat{q}) \leq 1 - \sqrt{\frac{\hat{v}(t,\hat{q})}{(1-\eta)(1-\Lambda)l_0}}, \tag{11.5.40}$$

is satisfied. This procedure yields an upper bound on γ at the confidence level at least Λ: the probability of wrongly deciding that $\hat{q}$ is an upper bound on γ is bounded as follows,

$$\begin{aligned}
&\mathbb{P}\Big(\hat{q} < \gamma,\ \sum_{l=1}^{l_0} e^{t(2m/\hat{q}-k)}\hat{\nu}_k < 1\Big) \\
&\quad \le \mathbb{P}\Big(\hat{q} < \gamma,\ \sum_{l=1}^{l_0} e^{t(2m/\hat{q}-k)}\hat{\nu}_k < 1, VAR\big(W^l(t,\hat{q})\big) \le \widehat{v}(t,\hat{q})\Big) + \mathbb{P}\Big(VAR\big(W^l(t,\hat{q})\big) > \widehat{v}(t,\hat{q})\Big) \\
&\quad \le (1-\eta)(1-\Lambda) + \eta(1-\Lambda) = 1-\Lambda.
\end{aligned}$$

Thus, the challenge is to design the estimators $\hat{q}$ and $\hat{v}(t,\hat{q})$ so that (11.5.39) and (11.5.40) are satisfied and $\hat{q}$ is as close to γ as possible, that is, as small as possible. Let us assume for a moment that the choice of $\hat{v}(t,q)$ has been fixed. Then a good choice for $\hat{q}$ is the solution of the following nonlinear optimization problem,

$$\begin{aligned}
\hat{q} &= \min_{(t,q)\in\mathbb{R}^2} q \qquad (11.5.41)\\
\text{subject to } &\frac{1}{l_0}\sum_{l=1}^{l_0} W^l(t,q) \le 1 - \sqrt{\frac{\hat{v}(t,q)}{(1-\eta)(1-\Lambda)l_0}} \\
&q \ge 0, t \ge 0.
\end{aligned}$$

Note that η, m, l_0 and the simulated data $\{Z_k^l :\ k \ge 1, 1 \le l \le l_0\}$ are all parameters that define (11.5.41), but when computing $\hat{q}$ we are interested in the situation where these parameters are fixed. Of course, the resulting value of $\hat{q}$ becomes a function of the parameters.

Let us now discuss how to define the estimator $\hat{v}(t,q)$. It follows from Corollary 11.4.2 that $\nu(k)$ is exponentially small in k for large k. Since moreover $\nu(k) = 0$ for $k < 2m$, this implies that $\mathbb{E}[W(t,q)]$, $VAR(W(t,q))$ and moments of all orders of $W(t,q)$ exist, suggesting the following approach: the empirical variance

$$\widehat{VAR}_{l_0}\big(W(t,q)\big) = \frac{1}{l_0-1}\sum_{k=1}^{l_0}\Big(W^k(t,q) - \frac{1}{l_0}\sum_{l=1}^{l_0} W^l(t,q)\Big)^2, \qquad (11.5.42)$$

is an unbiased estimator of $VAR\big(W(t,q)\big)$. Thus, if reasonable assumptions can be made about the distribution of the empirical variance (11.5.42) or a similar expression, then using a confidence interval argument one can define $\hat{v}$ so that it satisfies (11.5.39).

The Pitfalls of Variance Estimation

Before we put the outlined approach into practice, let us explain the pitfalls that need to be avoided. If $2m/\gamma - k > 0$ or, in other words, if k is small in comparison to the typical total length of a m-match, then Corollary 11.4.1 shows that

$$\nu(k) \le 2m|A|e^{-\frac{1}{8}\left(\frac{2m}{\gamma}-k\right)^2\frac{\gamma^2}{k}}.$$

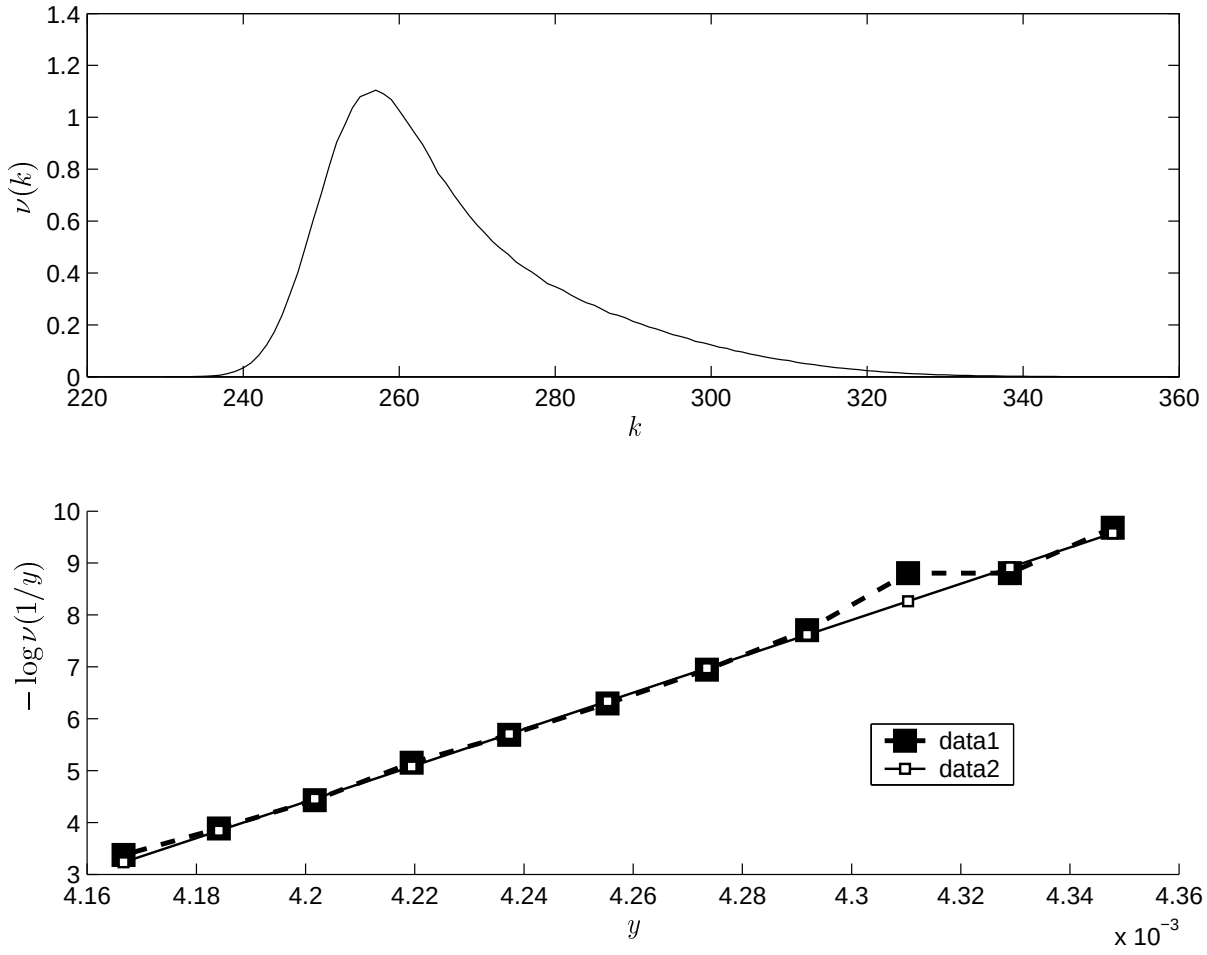


Figure 11.1: $\hat{\nu}$ and its lower tail end for $m = 100$ and $l_0 = 80000$. Data1 represent the function $-log\big(\nu(1/y)\big)$ and data2 the function $35000y - 142.6$.

This behavior is qualitatively correct: the first plot of Figure 11.1 shows the empirical distribution $\hat{\nu}$ obtained in $l_0 = 80000$ simulations for the coin flipping example and $m = 100$. The second plot shows that for k in the lower tail end $230 \leq k \leq 240$ the expression $-log\big(\nu(1/y)\big)$ as a function of $y = k^{-1}$ nearly coincides with the function $35000y - 142.6$. That is, for small k the measure ν nearly behaves like $\nu(k) \simeq \exp(a/k+b)$ with $a = -35000$ and $b = 142.6$. This leads to the following dilemma:

On the one hand, since $\mathbb{P}(Z_k^l > 0) \leq \mathbb{E}[Z_k^l] = \nu(k)$, *the event* $\{Z_k^l > 0\}$ *is exponentially rare in* $1/k$ *for* $2m \leq k < 2m/\gamma$, *and hence this occurs rarely in simulations.*

On the other hand, when $Z_k^l > 0$ *does occur then for* $q \simeq \gamma$ *the term* $\exp(t(2m/q - k))$ *is exponentially large in* $t > 0$ *for the same range of* k, *so that* $\exp(t(2m/q - k))\nu(k)$ *makes a nonnegligible contribution to* $W^l(t, q)$.

To make matters worse, when $Z_k^l > 0$ for some $k \in [2m, 2m/\gamma)$ then generally $Z_k^l > 0$ for other nearby values of k because the random variables $\{Z_k^l : k \in \mathbb{N}\}$ are not independent. Thus, unless t is very small, it could occur that most samples of $W^l(t, q)$ lie around

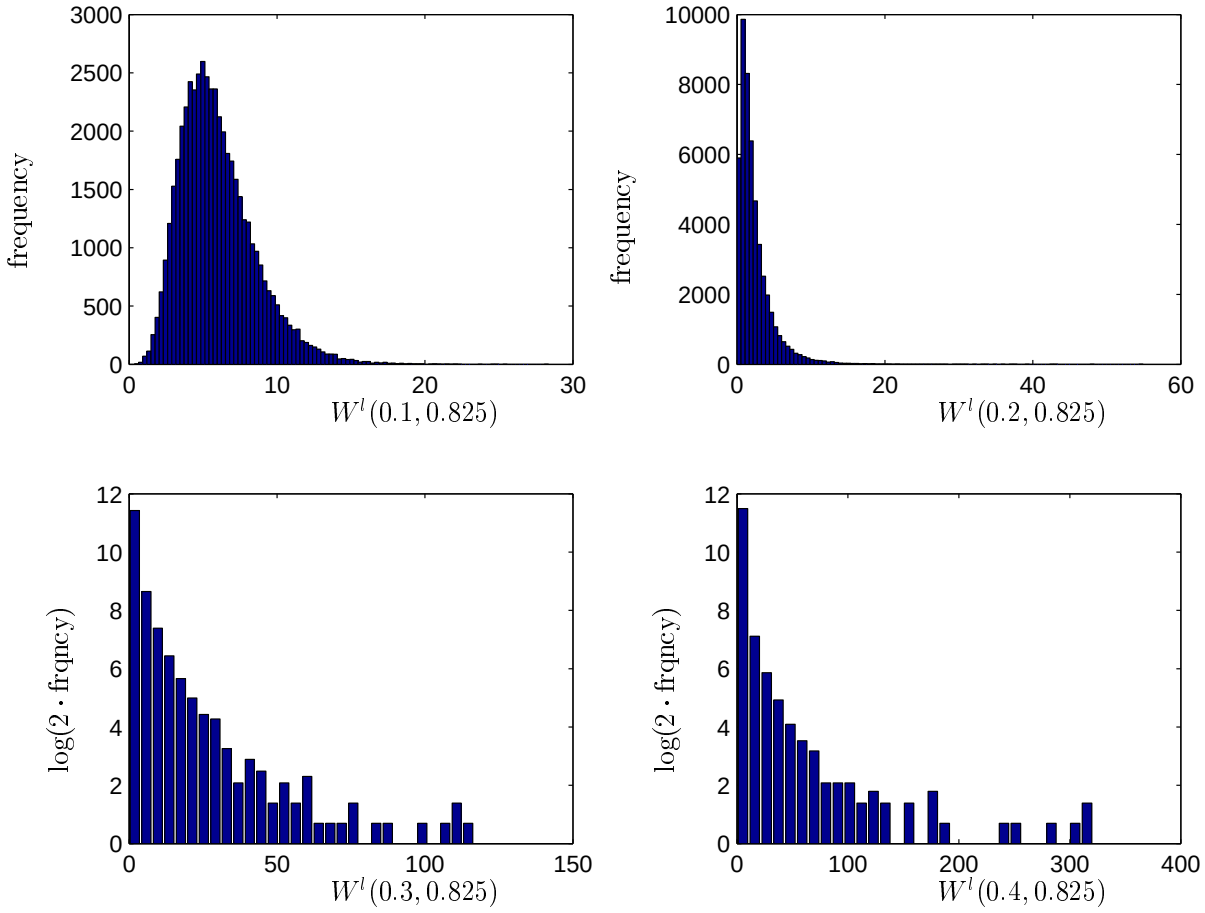


Figure 11.2: Histograms of 50000 samples of $W^l(t,q)$ for the coin flipping example with $m = 100$, $q = 0.825$, and for $t = 0.1, 0.2, 0.3$ and 0.4. In the last two histograms the ordinate is shown on a logarithmic scale

a cluster of small values, while a few outliers take significantly larger values. Such a situation renders accurate estimates of $VAR(W)$ by (11.5.42) extremely costly in terms of the number of simulations required. Figure 11.2 illustrates that this phenomenon occurs not only in theory, but also in practice: the plots show histograms of 50000 samples of $W^l(t,q)$ for the coin flipping example with $m = 100$, $q = 0.825$, and for $t = 0.1, 0.2, 0.3$ and 0.4 respectively. Note that for $t \geq 0.2$ small clusters of large values are observed.

Whether or not the observed outlier points affect the estimation of $VAR(W)$ depends on their probability weights. An investigation of the distribution tails of $W(t,q)$ is therefore revealing. Of course, we already noted that the tails of ν decrease exponentially, and that $\nu(k) = 0$ for $k < 2m$. These two properties assure that for fixed t and q the distribution tails of $W(t,q)$ decay exponentially, that is, $\mathbb{P}\big(W(t,q) > \tau\big) \leq be^{-a\tau}$ for some constants $a, b > 0$ and $\tau \gg 1$. Note however that this exponential decay may become effective only for very large τ, so that from an empirical point of view the decay might be algebraic, that is,

$$\mathbb{P}\big(W(t,q) > \tau\big) \sim e^b \cdot \tau^{-a} \tag{11.5.43}$$

for some constants $a > 0$, $b \in \mathbb{R}$. Figure 11.3 shows that this is indeed the case in the above discussed example. The plots show the function

$$y \mapsto \log\Big(\mathbb{P}\big(W(t,q) > e^y\big)\Big)$$

for $q = 0.825$ and $t = 0.3, 0.4$ respectively. It can be seen from the data that asymptotically the graph behaves like $-ay+b$, where $(a,b) = (2.22, 1.3)$ for $t = 0.3$ and $(a,b) = (1.6, 0.31)$ for $t = 0.4$. The value of a decreases with t.

For a reliable estimate of $VAR(W)$ via (11.5.42) the value of a would have to be substantially larger than 2, since for any distribution whose tail decay is governed by (11.5.43) we have $\mathbb{E}(W(t,q)) < +\infty$ if and only if $a > 1$ and $VAR(W(t,q)) < +\infty$ if and only if $a > 2$. Although for very large τ the tail decay of $W(t,q)$ is exponential, the fact that (11.5.43) holds for intermediate to large values of τ renders the variance estimation via (11.5.42) unreliable.

Therefore, useful information about the distribution of (11.5.42) *is not available, at least for reasonably small values of l_0 and reasonably large values of t.*

To further illustrate this point, Figure 11.3 also shows the histograms of 500 samples of $\widehat{VAR}_{100}\big(W(t,q)\big)$ for $q = 0.825$ and $t = 0.3, 0.4$ respectively, for the coin flipping example with $m = 100$. The ordinate of the second histogram, corresponding to $t = 0.4$, is reported on a logarithmic scale. Heavy tails of the distribution of $\widehat{VAR}_{100}\big(W(t,q)\big)$ are apparent because of the occurrence of massive outliers. The tails become lighter only very slowly with increasing l_0. For example, the tails of $\widehat{VAR}_{1000}(W(t,q))$ are heavier than those of the variable $\widehat{var}_{p,1000,1}$ which is defined below and whose histogram appears in Figure 11.4.

Avoiding the Pitfalls

In the previous paragraph we argued that the evaluating the empirical variance (11.5.42) is unreliable for certain values of (t,q). How can this problem be overcome?

On the one hand, one could impose an upper bound on t, depending on the value of m, so as to guarantee that (11.5.42) does not have major outliers. In fact, one can argue that in the typical range of t where (11.5.41) takes its optimum the distribution of (11.5.42) usually does not have too heavy tails. This observation forms the basis of a practical version of our method and implementations in Visual C++ undertaken in the recent MSc thesis [16]. A drawback of this approach is that certain assumptions about the distribution of (11.5.42) observed for moderately small values of m and l_0 can not be verified experimentally in reasonable time for the typical values of m and l_0 used in actual computations.

We are therefore going to pursue a different approach: recall that we are interested in the empirical variance (11.5.42) only because it is an unbiased estimator of $VAR\big(W(t,q)\big)$. But likewise, so is

$$\frac{1}{r}\sum_{j=1}^{r}\left(\Big(\frac{1}{s}\sum_{k=1}^{s} W^{l_{i,j,k}}(t,q)\Big) - \mathbb{E}[W(t,q)]\right)^2$$

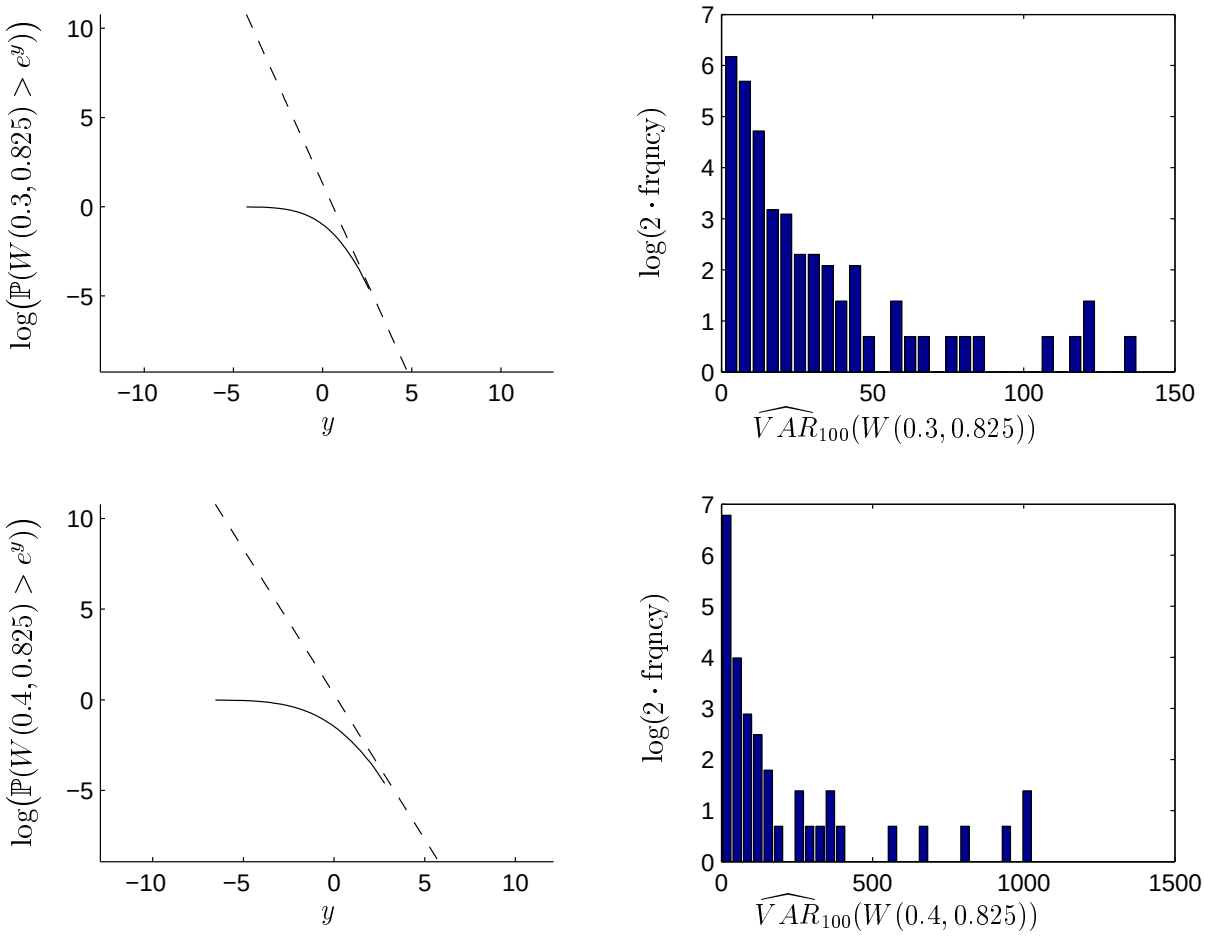


Figure 11.3: Empirical distribution tails of $W(t,q)$ are algebraic. Solid lines represent the function $\log\big(\mathbb{P}(W(t,q) > e^y)\big)$, dashed lines the function $ay + b$. Histograms of 500 samples of $\widehat{VAR}_{100}(W(t,q))$ for $t = 0.3, 0.4$ respectively, plotted on a logarithmic scale on the ordinate.

for each $(i = 1, \ldots, p)$ when $W^{l_{i,j,k}}(t,q)$ $(i = 1, \ldots, p; j = 1, \ldots, r; k = 1, \ldots, s)$ are i.i.d. copies of $W(t,q)$. Choosing $l_0 = prs$ large enough, $\mathbb{E}[W(t,q)] \simeq \frac{1}{prs} \sum_{i,j,k} W^{l_{i,j,k}}(t,q)$, so that

$$\widehat{var}^{\,i}_{p,r,s}(t,q) := \frac{1}{r-1} \sum_{j=1}^{r-1} \left(\Big(\frac{1}{s} \sum_{k=1}^{s} W^{l_{i,j,k}}(t,q) \Big) - \frac{1}{prs} \sum_{i,j,k} W^{l_{i,j,k}}(t,q) \right)^2 \tag{11.5.44}$$

is approximately an unbiased estimator of $VAR\big(W(t,q)\big)$.

Note that (11.5.44) is an improved version of the empirical variance of $\frac{1}{s} \sum_{k=1}^{s} W^{l_{i,j,k}}(t,q)$, $(j = 1, \ldots, r)$ for a fixed i, the only difference being that $\frac{1}{rs} \sum_{j,k} W^{l_{i,j,k}}$ has been replaced by $\frac{1}{prs} \sum_{i,j,k} W^{l_{i,j,k}}(t,q)$ which can be expected to have better converged to $\mathbb{E}[W(t,q)]$. This replacement achieves a slight lightening of the distribution tails, and it also introduces a small bias in the direction of overestimating, which we don't mind, because our aim is to derive an upper bound on $VAR(W(t,q))$.

The really powerful advantage of the new variance estimator is the fact that it is computed on the basis of the averaged data $\frac{1}{s} \sum_{k=1}^{s} W^{l_{i,j,k}}(t,q)$, which still have algebraic empirical decay but with a rate that becomes faster with increasing s. Indeed, Figure 11.4 shows that (11.5.44) has much lighter distribution tails than (11.5.42). In order to make the advantages of averaging apparent, we computed these histograms using the data underlying the first histogram of Figure 11.3. In the first row of Figure 11.4 we chose $(q, r, s) = (50, 1000, 1)$, that is, no averaging was applied. In the second row $(q, r, s) = (50, 200, 5)$, and in the third row $(q, r, s) = (50, 10, 100)$ was chosen. The plots in the left hand column show

$$\log \mathbb{P} \left(\frac{1}{s} \sum_{k=1}^{s} W^{l_{i,j,k}}(0.3, 0.825) > e^y \right)$$

as a function of y for $s = 1, 5, 100$ respectively. Note that the asymptotes of the graphs have decreasing gradient with increasing s, corresponding to a faster algebraic decay of the empirical distribution tails of the averaged data. Not surprisingly, when the decay rate of the averaged data becomes sufficiently fast the distribution of (11.5.44) increasingly resembles a Gaussian: a Lilliefors test applied to the data of the third histogram does not reject the hypothesis that the data is Gaussian with a p-value of 17.9%. The second histogram shows that even averaging of only 5 independent copies of $W(t,q)$ achieves a remarkable decrease in the heaviness of distribution tails and renders the distribution much more symmetric.

A Practical Algorithm

Note that if the distribution of a random variable X is perfectly symmetric, then the probability that 8 out of 10 independent copies $X_1, \ldots, X_{10}$ of X lie below $\mathbb{E}[X]$ equals

$$\mathbb{P}\big(|\{i : X_i \leq \mathbb{E}[X]\}| \geq 8\big) = 0.0107. \tag{11.5.45}$$

We will assume that if a Lilliefors test on the 5% level does not reject the hypothesis that $\widehat{var}_{10,r,s}$ is Gaussian, then the distribution is sufficiently symmetric for (11.5.45) to hold

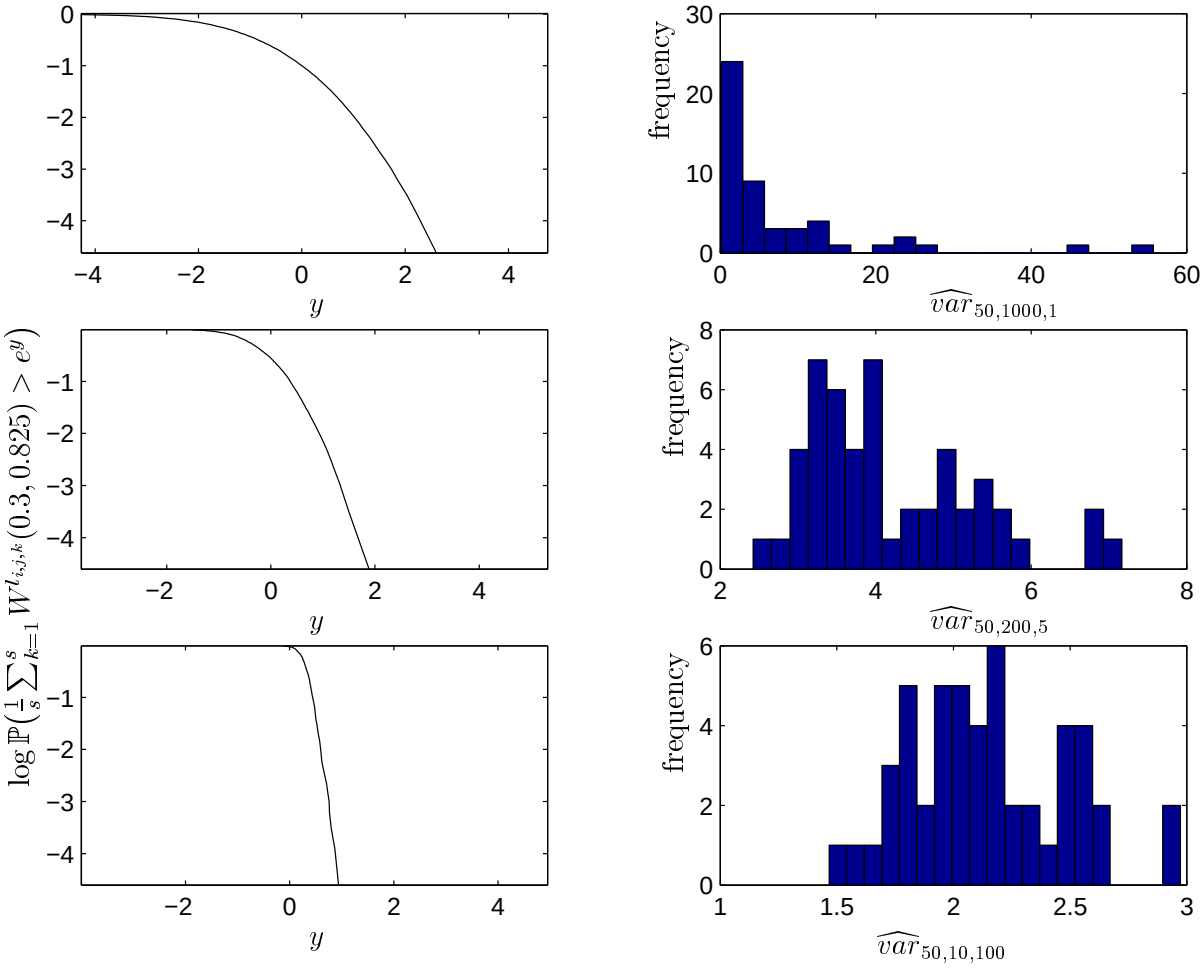


Figure 11.4: Algebraic decay rate of the tails of the empirical distribution tails of $W(t,q)$ becomes sharper with averaging. Variance estimates based on averaged data become increasingly Gaussian.

approximately for $X = \widehat{var}_{10,r,s}$. Since we know moreover that $\mathbb{E}\big[\widehat{var}_{10,r,s}\big] > VAR(W)$, we are confident that in this case

$$\mathbb{P}\Big(VAR\big(W(t,q)\big) > \widehat{var}^{[8]}_{10,r,s}(t,q)\Big) \leq 0.0107 = \eta \cdot (1-\Lambda), \tag{11.5.46}$$

with $\eta = 0.214$ and $\Lambda = 0.95$, and where $\widehat{var}^{[8]}_{10,r,s}$ denotes the 8-th order statistic of $\widehat{var}^{i}_{10,r,s}$ $(i = 1, \ldots, 10)$.

Our practical algorithm for finding an upper bound on the Chvàtal-Sankoff constant γ on the confidence level $\Lambda = 95\%$ is thus as follows:

1. For given input A and ξ, set $\Lambda = 95\%$, $\eta = 0.214$, $p = 10$ and choose m, r and s.
2. Generate the data vectors Z^l $(l = 1 \ldots, l_0 = p \cdot r \cdot s)$ using variant of the Wagner-Fischer algorithm.
3. For $t > 0$, $q \in (0,1)$, evaluate $\hat{v}(t,q) = \widehat{var}^{[8]}_{10,r,s}(t,q)$.
4. Determine $(\hat{t}, \hat{q})$ by solving the optimization problem (11.5.41).
5. If a Lilliefors test on the 5% level rejects the hypothesis that $\widehat{var}^{i}_{10,r,s}(\hat{t},\hat{q})$ $(i = 1, \ldots, 10)$ are Gaussian data, then increase s and/or r and return to Step 2. Otherwise accept $\hat{q}$ as an upper bound on γ on the 95% confidence level.

The last step provides a tool for automatically determining the number of simulations l_0 necessary for the results to be reliable on the 95% confidence level: it guarantees that the assumption on the symmetry of the distribution of $\widehat{var}10, r, s$ holds reasonably well at the optimal values of t and q. Of course, the method can be adapted to other values of p and Λ, but we found that our choice are reasonable values for the limited computing power of a desktop machine.

11.6 Implementation and Numerical Results

A straightforward adaptation of the Wagner-Fischer algorithm [24], which is based on dynamic programming, can compute the set of all m-matches contained in a pair of infinite random sequences (X, Y) in $O(m^2)$ time. A careful implementation which avoids computing unnecessary matrix entries achieves a practical complexity which is in effect closer to $O(m \log m)$. Moreover, the method can be implemented in such a way that only $O(m)$ storage of information is needed at any time point during a run of the algorithm. This is important because implementations based on $O(m^2)$ storage quickly spend most of the execution time moving information between different hierarchies of memory. The nontrivial constraint in the optimization problem (11.5.41) was strictly convex in all examples we attempted. Therefore, it is easy to find the global minimizer using standard software tools. We chose the Sequential Quadratic Programming solver of the Matlab Optimization Toolbox which could solve all examples to a precision of 10^{-8} within a few

$\|A\|$	DLB	ALB	BLB	$\hat{\mathbf{q}}$	DUP	AUP	P	s	l_0
2	0.7739	0.8079	0.8118	**0.8182**	0.8376	0.8607	0.0675	400	8000
3	0.6338	–	0.7172	**0.7235**	0.7658	–	>0.2	400	12000
4	0.5528	–	0.6537	**0.6601**	0.7082	–	>0.2	200	8000

Table 11.1: New upper bounds $\hat{q}$ on the 95% confidence level are computed with $m = 1000$. A comparison with BLB shows that $\hat{q}$ approximates the true value of γ to about $5 \cdot 10^{-3}$.

iterations. The Lilliefors test is implemented in standard software packages. We chose to use the Matlab Statistics Toolbox in which the test can be performed using a simple Matlab command.

The method was implemented in Matlab 6.1 and experiments were run on a SunBlade 100 workstation. Our aim was numerical accuracy and reliability rather than speed, and there remains considerable room for optimizing the code from the latter perspective, for example by removing multiple loops and by working with sparse matrix data structures.

In our experiments we considered the LCS problems in which A has $|A| = 2, \ldots, 4$ characters and where ξ is the uniform measure. Each of the experiments reported in Table 11.6 took a few days to complete. The value of $\hat{q}$ did not change significantly after a few hundred simulations, but we continued simulating until the variance data $\widehat{var}^i_{10,r,s}(\hat{t}, \hat{q})$ did not reject the Lilliefors test on the $P = 5\%$ level and hence was sufficiently symmetric. In all four experiments we chose $m = 1000$, $p = 10$ and $\Lambda = 0.95$, that is, $\hat{q}$ is an upper bound on γ at the 95% confidence level. The p-value P of the Lilliefors test and the number s of independent copies of Z used in averaging the raw data are listed in the last two columns. For comparison we also list the best deterministic lower and upper bounds for these examples, denoted by DLB and DUP respectively, which were derived by Dancik-Paterson [15, 22], as well as the best known probabilistic lower and upper bounds at the 95% confidence level, denoted by ALB and AUP respectively, which were derived by Alexander [2] on the basis of two simulations of $E[L_5 0000]$. Finally, we list the best known probabilistic lower bounds BLB *without confidence guarantee* which were obtained by Baeza-Yates, Gavalda, Navarro and Scheihing [10] on the basis of ten simulations of $E[L_{100000}]$.

Although for much larger m roundoff errors might play a significant role, such effects are minimal for $m = 1000$. For example, in the case $|A| = 2$ it is easy to see that the worst-case round-off error for the nontrivial constraint function

$$c(t,q) := \frac{1}{l_0} \sum_{l=1}^{l_0} W^l(t,q) + \sqrt{\frac{\hat{v}(t,q)}{(1-\eta)(1-\Lambda) l_0}} - 1$$

of (11.5.41) can be bounded by 10^{-9}. On the other hand, at the optimal values $(\hat{t}, \hat{q})$ one finds that $\left|\frac{\partial}{\partial q} c(\hat{t}, \hat{q})\right| > 10^{-2}$. Therefore, the backward error $\Delta\hat{q}$ satisfies $10^2 \cdot \Delta\hat{q} < 10^{-9}$, that is $\Delta\hat{q} < 10^{-11}$. However, since $\hat{q}$ approximates γ only to about $5 \cdot 10^{-3}$, we have $\hat{q} - \gamma > 10^{-4} \gg \Delta\hat{q}$. This shows that rounding errors neither wrongly indicate that $\hat{q}$ is an upper bound on γ nor interfere with its approximating quality.

References

[1] David Aldous and Persi Diaconis. Longest increasing subsequences: from patience sorting to the Baik-Deift-Johansson theorem. *Bull. Amer. Math. Soc. (N.S.)*, 36(4):413–432, 1999.

[2] Kenneth S. Alexander. The rate of convergence of the mean length of the longest common subsequence. *Ann. Appl. Probab.*, 4(4):1074–1082, 1994.

[3] A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors. *Combinatorial pattern matching*, volume 684 of *Lecture Notes in Computer Science*, Berlin, 1993. Springer-Verlag.

[4] R. Arratia, L. Goldstein, and L. Gordon. Two moments suffice for Poisson approximations: the Chen-Stein method. *Ann. Probab.*, 17(1):9–25, 1989.

[5] R. Arratia, L. Gordon, and M.S. Waterman. The Erdős-Rényi law in distribution, for coin tossing and sequence matching. *Ann. Statist.*, 18(2):539–570, 1990.

[6] R. Arratia and M.S. Waterman. The Erdős-Rényi strong law for pattern matching with a given proportion of mismatches. *Ann. Probab.*, 17(3):1152–1169, 1989.

[7] Richard Arratia, Louis Gordon, and Michael Waterman. An extreme value theory for sequence matching. *Ann. Statist.*, 14(3):971–993, 1986.

[8] Richard Arratia and Michael S. Waterman. A phase transition for the score in matching random sequences allowing deletions. *Ann. Appl. Probab.*, 4(1):200–225, 1994.

[9] K. Azuma. Weighted sums of certain dependent random variables. *Tohuku Math. J.*, 19:357–367, 1967.

[10] R.A. Baeza-Yates, R. Gavaldà, G. Navarro, and R. Scheihing. Bounding the expected length of longest common subsequences and forests. *Theory Comput. Syst.*, 32(4):435–452, 1999.

[11] Jinho Baik, Percy Deift, and Kurt Johansson. On the distribution of the length of the longest increasing subsequence of random permutations. *J. Amer. Math. Soc.*, 12(4):1119–1178, 1999.

[12] Renato Capocelli, Alfredo De Santis, and Ugo Vaccaro, editors. *Sequences. II.* Springer-Verlag, New York, 1993. Methods in communication, security, and computer science, Papers from the workshop held in Positano, June 17–21, 1991.

[13] Renato M. Capocelli, editor. *Sequences.* Springer-Verlag, New York, 1990. Combinatorics, compression, security, and transmission, Papers from the workshop held in Naples and Positano, June 6–11, 1988.

[14] Václáv Chvatal and David Sankoff. Longest common subsequences of two random sequences. *J. Appl. Probability*, 12:306–315, 1975.

[15] Vlado Dancik and Mike Paterson. Upper bounds for the expected length of a longest common subsequence of two binary sequences. *Random Structures Algorithms*, 6(4):449–458, 1995.

[16] Quentin Decouvelaere. *Upper Bounds for the LCS Problem.* MSc Thesis, Oxford University Computing Laboratory, 2003.

[17] Joseph G. Deken. Some limit results for longest common subsequences. *Discrete Math.*, 26(1):17–31, 1979.

[18] W. Hoeffding. Probability inequalities for sums of bounded random variables. *J. Amer. Statis. Assoc.*, 58:13–30, 1963.

[19] Krengel. *Ergodic Theorems.* W. de Gruyter, Berlin-New York, 1985.

[20] Joseph B. Kruskal. An overview of sequence comparison: time warps, string edits, and macromolecules. *SIAM Rev.*, 25(2):201–237, 1983.

[21] Claudia Neuhauser. A Poisson approximation for sequence comparisons with insertions and deletions. *Ann. Statist.*, 22(3):1603–1629, 1994.

[22] Mike Paterson and Vlado Dancik. Longest common subsequences. In *Mathematical foundations of computer science 1994 (Kosice, 1994)*, volume 841 of *Lecture Notes in Comput. Sci.*, pages 127–142. Springer, Berlin, 1994.

[23] David Sankoff and Joseph B. Kruskal, editors. *Time warps, string edits, and macromolecules: the theory and practice of sequence comparison.* Addison-Wesley Publishing Company Advanced Book Program, Reading, MA, 1983.

[24] Robert A. Wagner and Michael J. Fischer. The sring-to-string correction problem. *J. Ass. Comp. Mach.*, 21:168–173, 1974.

[25] M. Waterman. *Introduction to Computational Biology.* Chapman & Hall, 1995.

[26] Michael S. Waterman. General methods of sequence comparison. *Bull. Math. Biol.*, 46(4):473–500, 1984.

[27] Michael S. Waterman. Estimating statistical significance of sequence alignments. *Phil. Trans. R. Soc. Lond. B*, 344:383–390, 1994.

[28] M.S. Waterman and M. Vingron. Sequence comparison significance and poisson approximation. *Statistical Science*, 9(3):367–381, 1994.

Chapter 12

Deviation from mean in sequence comparison with a periodic sequence

(submitted)

Clement Durringer, Jüri Lember and Heinrich Matzinger

Abstract. Let L_n denote the length of the longest common subsequence of two sequences of length n. We draw one of the sequences i.i.d. whilst the other is non-random and periodic. We study the asymptotic deviation of the mean of L_n when n tends to infinity. We find that $L_n - E[L_n]$ is typically of order $\sqrt{n}$. This confirms the conjecture of Waterman [7] in the special case when one sequence is periodic.

12.1 Introduction

Let $\{X_i\}_{i\in\mathbb{N}}$ and $\{Y_i\}_{i\in\mathbb{N}}$ be two ergodic processes independent of each other. We assume that the variables X_i, and Y_i have a common state space. Let L_n denote the length of the longest common subsequence of the two finite sequences $X_1,\ldots,X_n$ and $Y_1,\ldots,Y_n$. A common subsequence is a subsequence of $X_1,\ldots,X_n$ which is also a subsequence of $Y_1,\ldots,Y_n$. Any common subsequence with maximal possible length is a longest common subsequence (for a formal definition, see Section 12.4).
The investigation of the longest common subsequences (LCS) of two finite words is one of the main problems in the theory of pattern matching. The LCS-problem plays a role for DNA- and Protein-alignments, file-comparison, speech-recognition and so forth. The random variable L_n and several of its variants have been studied intensively by probabilists, computer-scientists and mathematical biologists; for applications of LCS-algorithms in biology see Waterman [6].
One can show by a simple sub-additivity argument, that the limit

$$\lim_{n\to\infty}\frac{E[L_n]}{n}$$

exists (see [5]). An interesting question is: what is the asymptotic behavior of the deviation from its mean of L_n? (Hence, what is the order magnitude of $L_n - E[L_n]$ for large n?) In [7], Waterman conjectures that in many cases the deviation from the mean of L_n is of order $\sqrt{n}$. In [3], Bonetto and Matzinger consider the case where the X_i's and the Y_i's are i.i.d.. They prove that in certain cases the fluctuation is of order $\sqrt{n}$, indeed. However, the case where the variables X_i and Y_i are i.i.d. Bernoulli variables with parameter $1/2$, the order $\sqrt{n}$ seems to be an exception: The simulations of Bonetto and Matzinger [4] suggests that in that case the fluctuation is of order $n^{1/3}$.
In reality, the models like a language or a genetic code are often more complicated than an i.i.d. sequence. Therefore, in order to understand what determines the size of the fluctuations of L_n, it becomes essential to investigate different kind of models. Every model might capture one aspect of a complicated real life system. This is why, through a series of papers, we analyze the order of magnitude of the fluctuations of L_n for different cases.
In this paper, we show that when one of the sequence is non-random and periodic with a short period, then the deviation from the mean of L_n has order $\sqrt{n}$.

Let us mention a little bit more about the history of this field. The most widely used method for the comparison of genetic data is a generalization of the LCS-method. (For an excellent overview of this subject see Waterman-Vingron [8].) In this generalization a maximal score is sought over the set of all possible alignments of the two sequences, where gaps are penalized with a fixed parameter $\delta > 0$ and mismatches are penalized by a fixed amount $\mu > 0$: consider for example the two words "brot" and "bat". One possible alignment $\mathbb{A}$ of these words is

b	r	o	t
b	a	$-$	t

The score of this alignment is $1 - \mu - \delta + 1 = S(\mathbb{A})$. The matching pairs of letters "b" and "t" are each valued with a weight of 1. The gap $-$ in "bat" after the "a" costs $-\delta$. Furthermore, the mismatch between "r" and "a" is penalized by adding $-\mu$ to the total score. If $M_{\mu,\delta}(X, Y)$ denotes the maximal score amongst all possible alignments of the two words X and Y, and if $M_n(\mu, \delta)$ is the random variable defined by $M_n(\mu, \delta) = M_{\mu,\delta}(X, Y)$, where X and Y are two i.i.d. random sequences of length n, then the LCS-problem is a special case of the investigation of $M_n(\mu, \delta)$, because $L_n = M_n(\infty, 0)$. Generalizing the arguments from the LCS-problem, one can prove that the limit

$$a(\mu, \delta) = \lim_{n\to\infty} \frac{\mathbb{E}[M_n]}{n}$$

exists. Arratia-Waterman [2] showed that there is a phase transition phenomenon defined by critical values of μ and δ. In one phase M_n is of linear order in n, whereas in the other it is logarithmically small in n. Waterman [7] conjectures that the deviation of M_n from its mean behaves like $\sqrt{n}$.
Let us mention a few further details on the history of these problems and the state of knowledge about them: Waterman-Arriata [2] derive a law of large deviation for L_n for fluctuations on scales larger than $\sqrt{n}$. Using first passage percolation methods, Alexander [1] proves that $\mathbb{E}[L_n]/n$ converges at a rate of order $\sqrt{\log n/n}$.

12.2 Main result

Let $X_1, X_2, \ldots$ be an i.i.d. sequence of Bernoulli variable with parameter $1/2$. Let $Y_1, Y_2, \ldots$ be a non-random periodic sequence with period p, that is fixed throughout the paper. This means that $p > 1$ is the smallest natural number such that: $Y_{p+n} = Y_n$ for all $n \in \mathbb{N}$. Let L_n be the length of the longest common subsequence of the two finite sequences, $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_n$. A similar argument as in [5] implies that

$$\frac{L_n}{n} \to \gamma_Y, \quad \text{a.s.},$$

where γ_Y is an unknown constant. Of course, γ_Y depends on the periodic scenery Y. In this paper, we study the asymptotic deviation from the mean of the random variable L_n. Let D_n be defined as follows:

$$D_n := \frac{L_n - E[L_n]}{\sqrt{n}} \tag{12.2.1}$$

The main result of this paper is Theorem 12.2.1, which states that $L_n - E[L_n]$ is typically of order $\sqrt{n}$. To prove theorem 12.2.1, we show in Lemma 12.2.2 that the standard deviation of L_n is of order $\sqrt{n}$.
We need the following large deviation result, (which is similar to a result of Arriata and Waterman [2]):

Lemma 12.2.1. *There exists a constant $b > 0$ not depending on n and $\Delta > 0$ such that for all n large enough, we have:*

$$P\left(|L_n - EL_n| \geq n\Delta\right) \leq e^{-bn\Delta^2} \tag{12.2.2}$$

Proof. The inequality (12.2.2) is a straightforward application of the McDiarmid inequality: Let $X_1, \ldots, X_n$ independent A-valued random variables. Let $f : A^n \mapsto \mathbb{R}$ be a function that satisfies

$$\sup_{x_1,\ldots,x_n,x_i' \in A} |f(x_1, \ldots, x_n) - f(x_1, \ldots, x_{i-1}, x_i', x_{i+1}, \ldots, x_n)| \leq c_i, \quad i = 1, \ldots, n.$$

Then for any $\Delta > 0$

$$P\Big(|f(X_1, \ldots, X_n) - Ef(X_1, \ldots, X_n)| \geq \Delta\Big) \leq 2\exp\Big[-\frac{2\Delta^2}{\sum_{i=1}^n c_i^2}\Big]. \tag{12.2.3}$$

Take $f : \{0,1\}^n \to \mathbb{R}$ to be the length of the longest common subsequence between i.i.d. random variables $X_1, \ldots, X_n$ and non-random $Y_1, \ldots, Y_n$. So $L_n = f(X_1, \ldots, X_n)$. Clearly the following holds: by changing an element in a binary sequence $(x_1, \ldots, x_n) \in \{0,1\}^n$, the length of a longest common subsequence of $x_1, \ldots, x_n$ and $Y_1, \ldots, Y_n$ changes at most by one. Thus, the assumptions of McDiarmid inequality are satisfied with $c_i = 1$, $i = 1, \ldots, n$. Hence, the inequality (12.2.3) holds, and (12.2.2) trivially follows. □

Our main Lemma about the variance is:

Lemma 12.2.2. *There exist $0 < k < K < \infty$ not depending on n, such that for all n large enough:*

$$Kn \geq \mathrm{VAR}\,[L_n] \geq kn.$$

The proof of Lemma 12.2.2 is presented at the end of Section 12.3.

Our main theorem studies the sequence $\{D_n\}$ as defined in (12.2.1).

Theorem 12.2.1. *The sequence $\{D_n\}$ is tight. Moreover, the limit of any weakly convergent subsequence of $\{D_n\}$ is not a Dirac measure.*

Proof. For $s > 0$, the inequality (12.2.2) with $\Delta = \frac{s}{\sqrt{n}}$ implies

$$P(|D_n| \geq s) = P(|D_n| \geq \sqrt{n}\frac{s}{\sqrt{n}}) \leq \exp[-cn\frac{s^2}{n}] = \exp[-cs^2].$$

The last inequality implies that for any $r \geq 1$, the sequence $\{D_n\}$ is uniformly bounded in L_r, i.e.

$$\sup_n E|D_n|^r = \sup_n \int_0^\infty P(|D_n|^r \geq s)ds \leq \int_0^\infty \exp[-cs^{\frac{2}{r}}]ds < \infty. \tag{12.2.4}$$

Hence, the sequence $\{D_n\}$ is uniformly integrable and, therefore, tight.
Let $D_{n_i} \Rightarrow Q$ be a weakly converging subsequence of $\{D_n\}$. Suppose $Q = \delta_c$, for a $c \in (-\infty, \infty)$. By the continuous mapping theorem, $D^2_{n_i} \Rightarrow \delta_{c^2}$ or, equivalently, the sequence $D^2_{n_i}$ converges to the constant c^2 in probability. Since $\sup_n E|D_n|^3 < \infty$, the sequence $\{D^2_n\}$ is uniformly integrable, as well. Hence, the weak convergence implies that: $ED^2_{n_i} = \mathrm{VAR}D_{n_i} \to 0$, which contradicts Lemma 12.2.2. □

12.3 Proof of Lemma 12.2.2

This section is dedicated to the proof of Lemma 12.2.2.

12.3.1 Main idea and numerical example

Lemma 12.2.2 states that the variance of L_n is of order n. To prove this, we show that L_n can be written as the sum of two independent parts: $Z_{\vec{T}}$ and $L_n^{\vec{T}}$ (see 12.3.7). The variance of $Z_{\vec{T}}$ is of order n, and so is the variance of L_n.
Let us present a simple numerical example: Let the periodic sequence Y have period 2, such that:

$$Y_1Y_2Y_3Y_4Y_5Y_6 \ldots = 010101 \ldots$$

Let $l \in 12\mathbb{N}$. (Here the number 12 corresponds to $4p^2$). Assume that in the neighborhood of l, the sequence X is equal to the periodic sequence Y (except possibly in l). More precisely, assume that we observe:

$$Y_{l-12}Y_{l-11}Y_{l-10} \ldots Y_{l+9}Y_{l+10}Y_{l+11} = 010101010101a101010101010,$$

where a can be equal to either zero or one. A point l satisfying the last equality above is called a *replica point*. If a coincides with the periodic pattern, we say that the replica point l *matches*. In this numerical example, this would happen if $a = 0$. We call $[l - 4p^2, l + 2p^2]$ the *interval of the replica point l*. The main combinatorial idea in this article is contained in Lemma 12.3.1. It states that for a replica point l, the score L_n is increased by one

when l matches. Furthermore, this not influenced by the sequence X outside the interval of the replica point l. This fact is intuitively clear and it is simple to find a heuristic proof. However, the formal proof of Lemma 12.3.1 is difficult. The whole Section 12.4 is dedicated to it.
The variable $Z_{\vec{T}}$ is defined to be the number of replica points (among the first cn replica points, where $c > 0$ is a constant not depending on n). From Lemma 12.3.1, it follows directly that L_n can be written as a sum of $Z_{\vec{T}}$ and a term which depends only on the sequence X "outside the replica points intervals". This leads directly to the independence $Z_{\vec{T}}$ and $L_n^{\vec{T}}$.

12.3.2 Replica points

We can assume without restriction that $Y_0 = 1$. For $l \in \mathbb{N}$ we define the integer interval:

$$J_l := [l - 4p^2, l + 4p^2 - 1].$$

Let I_l designate J_l minus its center:

$$I_l := J_l - \{l\}.$$

Definition 12.3.1. *Let $l \in \mathbb{N}$, with $l > 4p^2$. We say that l is a* **replica point** *if the following condition holds:*

$$Y_z = X_z, \forall z \in I_l.$$

If l is a replica point and $X_l = Y_l$, then we say that the **replica point l matches**.

We need some more notation. We denote by A_l the event that l is a replica point and denote by Z_l the Bernoulli variable which is equal to one if and only if l is a replica point which matches. Thus, $Z_l = 1$ if A_l and $X_l = Y_l$ both hold, otherwise $Z_l = 0$.
We denote by $X_{|l}$ the finite sequence obtained from $X_1, \ldots, X_n$ by removing X_l, i.e.

$$X_{|l} := (X_1, X_2, \ldots, X_{l-2}, X_{l-1}, X_{l+1}, X_{l+2}, \ldots, X_n).$$

We denote by Σ_l the σ-algebra generated by $X_{|l}$, i.e.

$$\Sigma_l := \sigma(X_i | 1 \leq i \leq n, i \neq l).$$

Let L_n^l designate the length of the longest common subsequence of $X_{|l}$ and $Y_1, \ldots, Y_n$.

The next Lemma is the fundamental combinatorial idea for replica points. It says that when l is a replica point, then the length of the longest common subsequence can be decomposed as $L_n = Z_l + L_n^l$, where Z_l comes from the replica point and L_n^l depends on $X_{|l}$, only. Such a decomposition is useful, because $A_l \in \Sigma_l$, i.e. whether l is a replica point or not does not depend on X_l. The proof of Lemma 12.3.1 is given in Section 12.4.

Lemma 12.3.1. *Let $l \in \mathbb{N}$ so that $4p^2 < l \leq n - 4p^2 - 1$. If A_l holds, then*

$$L_n = Z_l + L_n^l. \tag{12.3.1}$$

12.3.3 Several replica points

In the following, $c > 0$ is a constant not depending on n such that $cn \in \mathbb{N}$. (We choose $c > 0$ to be small enough, so that with high probability there are at least cn replica points in $[0, n]$. By Lemma 12.3.4, it is enough to take c such that: $0 < c < (0.5)^{8p^2-1}$.) Let $K^n \subseteq \mathbb{N}^{cn}$ designate the set of all integer vectors

$$\vec{k} = (k_1, k_2, \ldots, k_{cn})$$

such that $k_i + 8p^2 \leq k_{i+1}, \forall i = 1, \ldots, cn-1$ and $4p^2 < k_1$ and $k_{cn} < n - 4p^2$.

Let $\vec{k} = (k_1, k_2, \ldots, k_{cn}) \in K^n$. We define the σ-algebra:

$$\Sigma_{\vec{k}} := \sigma\left(X_i |\ i \in [0, n] \text{ and } i \neq k_j, \forall j \in [1, cn]\right).$$

We denote by $A_{\vec{k}}$ the event that k_i is a replica point for all $i = 1, \ldots, cn$. Clearly $A_{\vec{k}} \in \Sigma_{\vec{k}}$.

Suppose $A_{\vec{k}}$ holds. Let $Z_{\vec{k}}$ designate the number of replica points among $k_1, k_2, \ldots, k_{cn}$ which are matches. So, if $A_{\vec{k}}$ holds, and $\vec{k} = (k_1, \ldots, k_{cn}) \in K^n$, then

$$Z_{\vec{k}} := \sum_{i=1}^{cn} Z_{k_i}.$$

Let $X_{|\vec{k}}$ designate the finite sequence one obtains by removing from X the bits $X_{k_i}, i = 1, \ldots, cn$. Hence, for $\vec{k} = (k_1, \ldots, k_{cn}) \in K^n$,

$$X_{|\vec{k}} := \{X_i |\ i \in [0, n] \text{ and } i \neq k_j, \forall j \in [1, cn]\}.$$

Finally, let $L_n^{\vec{k}}$ designate the length of the longest common subsequence of $X_{|\vec{k}}$ and Y.

Lemma 12.3.2. *Let $\vec{k} \in K^n$. When $A_{\vec{k}}$ holds, then*

$$L_n = Z_{\vec{k}} + L_n^{\vec{k}}. \tag{12.3.2}$$

Proof. The proof follows from Lemma 12.3.1 by induction.
Let $cn = 2$, i.e. $\vec{k} = (l_1, l_2)$. Let $Z_i = Z_{l_i}$, $i = 1, 2$. Let us show that

$$L_n = L_n^{\vec{k}} + Z_1 + Z_2. \tag{12.3.3}$$

Let L_n^{1+} be length of the longest common subsequence of $X|_{l_1}$ and $Y_1, \ldots, Y_n$ provided that $Z_2 = 1$. Let L_n^{1-} be length of the longest common subsequence of $X|_{l_1}$ and $Y_1, \ldots, Y_n$ provided that $Z_2 = 0$. Finally, let $L_n^1 := L_n^{l_1}$, so L_n^1 is either L_n^{1+} or L_n^{1-}.
At first note,

$$L_n^{1-} + 1 = L_n^{1+}. \tag{12.3.4}$$

Let L_n^+ and L_n^- denote the length of the longest common subsequence of $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ provided that $Z_2 = 1$ and $Z_2 = 0$, respectively. From Lemma 12.3.1 follows that $L_n^+ = L_n^- + 1$ as well as $L_n^{1+} + Z_1 = L_n^+$ and $L_n^{1-} + Z_1 = L_n^-$. Hence, (12.3.4) holds.
Clearly, $L_n^1 \geq L_n^{\vec{k}} \geq L_n^1 - 1$. Hence, $L_n^{\vec{k}}$ is equal to L_n^{1+} or $L_n^{1+} - 1 = L_n^{1-}$. If $L_n^{\vec{k}} = L_n^{1+}$, we

would have that $L_n^{\vec{k}} > L_n^{1-}$, a contradiction. Hence $L_n^{\vec{k}} = L_n^{1+} - 1 = L_n^{1-}$. Suppose $Z_2 = 1$. Then $L_n = L_n^+ = L_n^{1+} + Z_1$, so

$$L_n^{\vec{k}} + Z_1 + Z_2 = L_n^{\vec{k}} + Z_1 + 1 = L_n^{1+} + Z_1 = L_n^+ = L_n.$$

Suppose $Z_2 = 0$. Then $L_n = L_n^- = L_n^{1-} + Z_1$, so

$$L_n^{\vec{k}} + Z_1 + Z_2 = L_n^{\vec{k}} + Z_1 = L_n^{1-} + Z_1 = L_n^- = L_n.$$

Let $cn = m + 1$, i.e. $\vec{k} = (l_1, l_2, \ldots, l_{m+1})$. Let $\vec{m} := (l_1, l_2, \ldots, l_m)$, $Z_m = \sum_{i=1}^m Z_{l_i}$, $Z_{m+1} := Z_{l_{m+1}}$. Suppose (12.3.2) holds for $cn = m$, i.e.

$$L_n = L_n^{\vec{m}} + Z_m. \tag{12.3.5}$$

Let us show that

$$L_n = L_n^{\vec{k}} + Z_m + Z_{m+1}.$$

The argument is similar to the case $m = 2$. Let L_n^{m+} be equal to $L_n^{\vec{m}}$ provided that $Z_{m+1} = 1$. Let L_n^{m-} be equal to $L_n^{\vec{m}}$ provided that $Z_{m+1} = 0$. At the first, we prove that

$$L_n^{m-} + 1 = L_n^{m+}. \tag{12.3.6}$$

Let L_n^+ and L_n^- denote the length of the longest common subsequence of $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ provided that $Z_{m+1} = 1$ and $Z_{m+1} = 0$, respectively. From Lemma 12.3.1 follows that $L_n^+ = L_n^- + 1$. From (12.3.5) follows $L_n^{m+} + Z_m = L_n^+$ and $L_n^{m-} + Z_m = L_n^- = L_n^+ - 1$. Hence, (12.3.6) holds.
Clearly, $L_n^{\vec{m}} \geq L_n^{\vec{k}} \geq L_n^{\vec{m}} - 1$. Hence, $L_n^{\vec{k}}$ is equal to L_n^{m+} or $L_n^{m+} - 1 = L_n^{m-}$. If $L_n^{\vec{k}} = L_n^{m+}$, we would have that $L_n^{\vec{k}} > L_n^{m-}$, a contradiction. Hence $L_n^{\vec{k}} = L_n^{m+} - 1 = L_n^{m-}$.
Suppose $Z_{m+1} = 1$. Then by (12.3.5), $L_n = L_n^+ = L_n^{m+} + Z_m$, so

$$L_n^{\vec{k}} + Z_m + Z_{m+1} = L_n^{\vec{k}} + Z_m + 1 = L_n^{m+} + Z_m = L_n^+ = L_n.$$

Suppose $Z_{m+1} = 0$. Then by (12.3.5), $L_n = L_n^- = L_n^{m-} + Z_m$, so

$$L_n^{\vec{k}} + Z_m + Z_{m+1} = L_n^{\vec{k}} + Z_m = L_n^{m-} + Z_m = L_n^- = L_n.$$

□

12.3.4 Intervals

Let U_i, $i = 1, 2, \ldots$ be the disjoint consecutive intervals with length $8p^2$, i.e. (recall the definition of J_l)

$$U_i := J_{i4p^2+1} = [(i-1)8p^2 + 1, i8p^2], \quad i = 1, 2, \ldots.$$

Let $u_i := i4p^2 + 1$. Whether u_i is a replica point or not, depends on $\{X_z : z \in U_i, z \neq u_i\}$.

Let T_i designate the i-th replica point. Formally, we define T_i by induction on i. For $i = 1$, we put:

$$T_1 := \min\{u_j | u_j \text{ is a replica point}, j > 0\}.$$

Once, T_i is defined, we define T_{i+1} in the following way:

$$T_{i+1} := \min\{u_j > T_i | u_j \text{ is a replica point }, j > 0\}.$$

Let $c > 0$ be a constant not depending on n. We define the event

$$E_n := \{T_{cn} \leq n\}$$

which guarantees that there are at least cn replica points in $[0, n]$.

Let

$$\vec{T} := \begin{cases} (T_1, T_2, \ldots, T_{cn}), & \text{if } E_n \text{ holds}, \\ 0, & \text{otherwise} \end{cases},$$

$$X_{|\vec{T}} := \begin{cases} X_{|\vec{k}}, & \text{if } \vec{T} = \vec{k}, \\ X, & \text{if } \vec{T} = 0. \end{cases} \qquad Z_{\vec{T}} := \begin{cases} Z_{\vec{T}} := Z_{\vec{k}}, & \text{if } \vec{T} = \vec{k}. \\ 0, & \text{if } \vec{T} = 0. \end{cases}$$

In other words, when E_n holds, $X_{|\vec{T}}$ is the sequence obtained by removing the bits $X_{T_1}, X_{T_2}, \ldots, X_{T_{cn}}$ from the sequence X and $Z_{\vec{T}}$ is the number of matching replica points in $\vec{T}$.
With $L_n^0 := L_n$, we obviously have

$$L_n = Z_{\vec{T}} + L_n^{\vec{T}}. \tag{12.3.7}$$

Finally, let

$$\Sigma := \sigma(\vec{T}, X_{|\vec{T}}).$$

Clearly, $L_n^{\vec{T}}$ is Σ-measurable and $E_n \in \Sigma$.

Lemma 12.3.3. *Conditional on Σ and E_n, $Z_{\vec{T}}$ has binomial distribution with parameters $1/2$ and cn:*

$$\mathcal{L}\left(Z_{\vec{T}} \middle| \vec{T} = \vec{k}, X_{|\vec{k}}\right) = B(1/2, cn),$$

for all $\vec{k} \in K^n$.

Proof. By interval construction, it holds that $\{\vec{T} = \vec{k}\} \in \sigma(X_{|\vec{k}})$. The vector $\vec{Z} := (Z_{k_1}, \ldots Z_{k_{cn}})$ is $\sigma(X_{k_1}, \ldots, X_{k_{cn}})$-measurable. Those σ-algebras are independent, hence $\vec{Z}$ is independent of $\sigma(X_{|\vec{k}})$. By interval-construction, $\vec{Z}$ consists of independent components. Since X_i is a Bernoulli $1/2$-random variable, the statement holds. □

The next Lemma shows that we can choose $c > 0$ so that for big n, there are typically at least cn replica points in $[0, n]$.

Lemma 12.3.4. *If $c < (0.5)^{8p^2-1}$, then $\lim_{n\to+\infty} P(E_n) = 1$.*

Proof. Let ξ_i be a Bernoulli random variable that is 1 if and only if u_i is a replica point. Clearly, $P(\xi_i = 1) = (0.5)^{8p^2-1} =: q$ and

$$E_n = \Big\{\sum_{i=1}^{n} \xi_i \geq cn\Big\}.$$

Then, by Hoeffding inequality,

$$P(E_n^c) = P\Big(\sum_{i=1}^{n} \xi_i < cn\Big) = P\Big(\sum_{i=1}^{n} \xi_i - qn < (c-q)n\Big) \leq \exp[-2(c-q)^2 n] \to 0.$$

□

12.3.5 Proof of Lemma 12.2.2

From (12.2.4) it follows: $\exists K < \infty$ such that

$$\sup_n ED_n^2 = \sup_n \frac{\mathrm{VAR}[L_n]}{n} < K.$$

We now prove the existence of $k > 0$.
Clearly

$$\mathrm{VAR}[L_n] = E\left(\mathrm{VAR}[L_n|\Sigma]\right) + \mathrm{VAR}\left(E[L_n|\Sigma]\right) \geq E\left(\mathrm{VAR}[L_n|\Sigma]\right).$$

By (12.3.7), $L_n = Z_{\vec{T}} + L_n^{\vec{T}}$. Since $L_n^{\vec{T}}$ is Σ-measurable, it holds that:

$$\mathrm{VAR}[L_n|\Sigma] = \mathrm{VAR}[Z_{\vec{T}}|\Sigma]. \tag{12.3.8}$$

By Lemma 12.3.3, on $E_n = \{T \neq 0\}$, the conditional distribution of $Z_{\vec{T}}$ is binomial. On E_n^c, $Z_{\vec{T}} = 0$ and hence $E(I_{E_n^c}\mathrm{VAR}[Z_{\vec{T}}|\Sigma]) = 0$. Therefore:

$$E(\mathrm{VAR}[L_n|\Sigma]) = E(\mathrm{VAR}[Z_{\vec{T}}|\Sigma]) = E(I_{E_n}\mathrm{VAR}[Z_{\vec{T}}|\Sigma]) + E(I_{E_n^c}\mathrm{VAR}[Z_{\vec{T}}|\Sigma]) = 0.25cn \cdot P(E_n).$$

By Lemma 12.3.4, for all n large enough we have:

$$0.25cn \cdot P(E_n) \geq kn,$$

for any $k > 0$ not depending on n, such that $k < 0.25c$.

12.4 Combinatorics

The rest of this paper is devoted to the proof of Lemma 12.3.1.

12.4.1 Preliminaries

Blocks

We need to introduce some necessary formalism. In the present Section, we consider the non-random sequences, only. At first, we formalize the common subsequence.

Let $x_1, \ldots, x_n$ and $y_1, \ldots, y_m$ be two fixed finite sequences. A *common subsequence* of $x_1, \ldots, x_n$ and $y_1, \ldots, y_m$ is a strictly increasing mapping

$$v : \{1, \ldots, n\} \hookrightarrow \{1, \ldots, m\}. \tag{12.4.1}$$

Notation (12.4.1) means: There exists $I \subseteq \{1, \ldots, n\}$ and a mapping

$$v : I \to \{1, \ldots, m\}$$

such that

$$y_{v(i)} = x_i, \quad \forall i \in I$$

and v is strictly increasing: $v(i_2) > v(i_1)$, if $i_2 > i_1$.

Let $x_1, \ldots, x_n$ and $y_1, \ldots, y_m$ be two sequences and let v be a common subsequence. Since v is defined as a mapping (12.4.1), in what follows, we would like to distinguish the sequence on which v is defined from the image sequence of v. Therefore, we say: v is a common subsequence between $x_1, \ldots, x_n$ and $y_1, \ldots, y_m$, implying that v is defined as (12.4.1), i.e. from the sequence $x_1, \ldots, x_n$ into $y_1, \ldots, y_m$.

The set I in (12.4.1) shall be denoted by

$$\mathrm{Dom}(v).$$

The length of v, denoted as $|v|$, is $|\mathrm{Dom}(v)|$.

With $J \subseteq \{1, \ldots, n\}$, we denote by $v|_J$ the restriction of v to J. The restriction as a subsequence of the common sequence v is defined even when J is not a subset of $\mathrm{Dom}(v)$.

For $a \in \{1, \ldots, n\}$, we define

$$\underline{v}(a) = v(\max\{i \in \mathrm{Dom}(v) : i < a\}) + 1, \quad \bar{v}(a) = v(\min\{i \in \mathrm{Dom}(v) : i > a\}) - 1.$$

Our analysis is based on the *optimality principle*: If v is a longest common subsequence, then for any $[a, b] \subseteq \{1, \ldots, n\}$, the subsequences:

$$\begin{aligned} v|_{[1,a-1]} &: \{1, \ldots, a-1\} \hookrightarrow \{1, \ldots, \bar{v}(a-1)\} \\ v|_{[a,b]} &: \{a, \ldots, b\} \hookrightarrow \{\underline{v}(a), \ldots, \bar{v}(b)\} \\ v|_{[b+1,a]} &: \{b+1, \ldots, n\} \hookrightarrow \{\underline{v}(b+1), \ldots, m\} \end{aligned}$$

are all with the longest possible length.

Note: $[\underline{v}(a), \bar{v}(b)]$ can also be empty. Moreover, the intervals $[1, \bar{v}(a-1)]$ and $[\underline{v}(a), \bar{v}(b)]$ as well as $[\underline{v}(a), \bar{v}(b)]$ and $[\underline{v}(b+1), m]$ can be overlapping, but the overlapping region does not contain any elements of common subsequence v.

Let v be a common subsequence, i.e. a mapping satisfying (12.4.1). Let $\{A_1, \ldots, A_l\}$ be a partition of $\mathrm{Dom}(v)$ that satisfies:

i) A_i is an integer interval for every i, i.e. $A_i = \{j, j+1, \ldots, j+s\}$ for some $s \geq 0$.

ii) v is linear on A_i, i.e.

$$v(j+1) = v(j) + 1, \quad \text{for every } j \in A_i \text{ such that } j+1 \in A_i.$$

Clearly there exists at least one partition that satisfies **i)** and **ii)**: the partition, where $A_i = \{i\}$ for every $i \in \mathrm{Dom}(v)$. This is the maximal partition. Let $B^*(v) = B^* = \{B_1, \cdots, B_r\}$ be the minimal partition that satisfies **i)** and **ii)**, i.e. every other partition is a sub-partition of B^*. Clearly B^* exists and is unique. We call the elements of B^* the *blocks* of v. By **i)**, every block is an interval, the *length* of B is the number of the elements in B.

Proposition 12.4.1. *Let* $\{B_1, \ldots, B_r\}$ *be the blocks of*

$$v : \{1, \ldots, n\} \hookrightarrow \{1, \ldots, m\}.$$

Then

$$\max\{n, m\} \geq \lfloor \frac{r-1}{2} \rfloor + \sum_i^r |B_i| = \lfloor \frac{r-1}{2} \rfloor + |\mathrm{Dom}(v)|. \tag{12.4.2}$$

Proof. Let $n_j := \max B_j$, $j = 1, 2, \ldots, r$. From the definition of blocks, it follows: $n_2 \geq |B_1| + |B_2| + 1$ or $v(n_2) \geq |B_1| + |B_2| + 1$, i.e by changing the block, v "loses" an element either in the set on which v is defined or in the image set of v. Similarly, $n_4 \geq |B_1| + |B_2| + 2$ or $v(n_4) \geq |B_1| + |B_2| + 2$. Hence, for an even r,

$$\max\{n_r, v(n_r)\} \geq \sum_i^r |B_i| + \frac{r}{2}.$$

Since $\max\{n, m\} \geq \max\{n_r, v(n_r)\}$, (12.4.2) follows. □

The blocks between two subsequences of a periodic sequence

In the following, we investigate common subsequences between finite periodic sequences. We start with a simple but yet useful observation, proved in the Appendix.

Proposition 12.4.2. *Let* $x_1, x_2, \ldots$ *be a periodic sequence with period* p. *If* $k \leq p$ *is a nonnegative integer such that*

$$x_j = x_{k+j}, \quad \forall j = 1, \ldots, p, \tag{12.4.3}$$

then $k = p$.

Assume now that $x_1, \ldots, x_n$ and $x_{m+1}, \ldots, x_{m+n}$ are two subsequences of a periodic sequence $\{x_n\}$ with period p. Let v be a common subsequence of $x_1, \ldots, x_n$ and $y_1, \ldots, y_n = x_{m+1}, \ldots, x_{m+n}$, i.e.

$$v : \{1, \ldots, n\} \hookrightarrow \{1, \ldots, n\}.$$

Let B be a block of v. The difference $v(i) + m - i$, where $i \in B$ is called the *bias* of B.

What is the meaning of the bias? Suppose v is a common subsequence, $B = \{j, \dots, j+s\}$ is a block of v with the bias 2. This means that the common subsequence v includes the elements $x_j, \dots, x_{j+s}$ of $x_1, \dots, x_n$. We also know, how these elements are matched with the elements of $y_1, \dots, y_n$: $x_j = y_{j+2-m}$, $x_{j+1} = y_{j+3-m}, \dots, x_{j+s} = y_{j+s+2-m}$. Since $y_j = x_{j+m}$, we get $x_j = x_{j+2}$, $x_{j+1} = x_{j+3}, \dots, x_{j+s} = x_{j+s+2}$. Moreover, for x_{j-1} (x_{j+m+1}), it holds: x_{j-1} (x_{j+m+1}) either does not belong to the common subsequence or it is matched with an element not equal to x_{j+1} (x_{j+m+3}).
Hence, the bias 0 means that every element of B is matched with itself – the *identity matching*. By periodicity, the bias np means essentially the same. We say that B is *unbiased*, if the bias of B is np for a $n \in \mathbb{N}$. Otherwise B is *biased*. Proposition 12.4.2 can be restated:

Proposition 12.4.3. *Let B be a biased block. Then the length of B is at most $p-1$.*

Example 12.4.1. *Let us give a numerical example. Let*

$$(x_1, \dots, x_{20}) = (00111001110011100111),$$
$$(y_1, \dots, y_{20}) := (x_2, \dots, x_{21}) = (01110011100111001110).$$

So, we consider the subsequences of a periodic sequence with the period $p = 5$. Let

$$v : \{1, \dots, 20\} \hookrightarrow \{1, \dots, 20\},$$

with

$$v(1) = 1, v(3) = 3, v(4) = 4, v(5) = 7, v(6) = 10, v(7) = 11, v(8) = 12$$
$$v(14) = 13, v(15) = 14, v(16) = 15, v(17) = 16, v(18) = 17, v(19) = 18$$

be a common subsequence. Obviously,

$$\mathrm{Dom}(v) = \{1, 3, 4, 5, 6, 7, 8, 14, 15, 16, 17, 18, 19\}$$

and v has 5 blocks:

$$B_1 = \{1\},\ B_2 = \{3, 4\},\ B_3 = \{5\},\ B_4 = \{6, 7, 8\},\ B_5 = \{14, 15, 16, 17, 18, 19\}.$$

Since $m = 1$, the corresponding biases are

$$b(B_1) = 1 - 1 + 1 = 1,\ b(B_2) = 1,\ b(B_3) = 7 - 5 + 1 = 3,\ b(B_4) = 5,\ b(B_5) = 0.$$

Hence, the blocks B_4 and B_5 are unbiased. The lengths of the blocks are, respectively, 1,2,1,3,6. The length of v, is $|v| = |B_1| + |B_2| + |B_3| + |B_4| + |B_5| = 1 + 2 + 1 + 3 + 6 = 13$.

Sometimes we regard v as a subsequence between

$$(x_1, \dots, x_{20}) = (00111001110011100111),$$
$$(x_2, \dots, x_{21}) = \ \ (01110011100111001110),$$

i.e. v is a mapping

$$v : \{1, \dots, 20\} \hookrightarrow \{2, \dots, 21\}.$$

with

$$v(1) = 2, v(3) = 4, v(4) = 5, v(5) = 8, v(6) = 11, v(7) = 12, v(8) = 13$$
$$v(14) = 14, v(15) = 15, v(16) = 16, v(17) = 17, v(18) = 18, v(19) = 19.$$

With this notation, the blocks and their biases remain unchanged, the bias of a block $B = \{i, \dots, j\}$ is just defined as $v(i) - i$.

12.4.2 The structure of a common subsequence between periodic subsequences

The structure of a common subsequence between periodic subsequences with length $8p^2$

In the present Subsection, we consider the subsequences of a periodic sequence with length $8p^2$, i.e. we consider the sequences $x_1, \ldots, x_{8p^2}$ and $x_{m+1}, \ldots, x_{m+8p^2}$. We are interested in the length and the structure of (any) longest common subsequence of these two subsequences. Of course, when m is a multiple of p, then the longest common subsequence is just the identity matching. Hence, we assume that m is not a multiple of p. Without loss of generality, we assume that $0 < m < p$. Moreover, it is easy to see that without loss of generality we can (and we do) assume that

$$0 < m \leq \frac{p}{2}.$$

Obviously, there exists a common subsequence v with length $8p^2 - m$: the identity matching. Such a v has only one block with bias 0.

Proposition 12.4.4. *Let $x_1, \ldots, x_{8p^2}$ and $x_{m+1}, \ldots, x_{m+8p^2}$ be the subsequences of a periodic sequence, $0 \leq m \leq \frac{p}{2}$. Then the length of the longest common subsequence is $8p^2 - m$.*

Proof. Let v be a longest common subsequence, let $\{B_1, \ldots, B_r\}$ be the blocks of v. Note: if v has an unbiased block, then the length of v is at most $8p^2 - m$. Indeed: suppose that the bias of $B_j = \{i_j, i_j + 1, \ldots, i_j + s\}$ $s \geq 0$ is 0. Let $n_{j-1} = \max B_{j-1}$. Since $v(n_{j-1}) \leq v(i_j) - 1 = i_j - 1 - m$, we have that the length of $v|_{B_1 \cup \ldots \cup B_{j-1}}$ is at most $v(n_{j-1}) = i_j - m - 1$. Similarly, the length of $v|_{B_{j+1} \cup \ldots \cup B_r}$ is at most $8p^2 - (i_j + s)$. So the length of v is at most $(i_j - m - 1) + (s + 1) + (8p^2 - (i_j + s)) = 8p^2 - m$.
If the bias of B_j is kp for a $k \in \mathbb{N}, k \neq 0$ the same argument holds.
Hence, if the length of v is bigger than $8p^2 - m$, then all blocks $\{B_1, \ldots, B_r\}$ must be biased. By Proposition 12.4.3, the length of a biased block is at most $p - 1$. Thus, the number of blocks is bounded below $r \geq \frac{8p^2 - m + 1}{p}$ and

$$\lfloor \frac{r-1}{2} \rfloor \geq \lfloor \frac{8p^2 - m + 1 - p}{2p} \rfloor \geq \lfloor 4p - \frac{m-1}{2p} - \frac{1}{2} \rfloor \geq 4p - 1 > m + 1. \qquad (12.4.4)$$

From Proposition 12.4.1, it follows $|\mathrm{Dom}(v)| < 8p^2 - m - 1$ that contradicts the assumption that the length of v is at least $8p^2 - m + 1$. □

Corollary 12.4.1. *Let v be a longest common subsequence, and let $\{B_1, \ldots, B_r\}$ be its blocks. Then there exists one and only one block B_j that is unbiased. Moreover, the bias of B_j is 0 or p, and it can be p only, when $m = \frac{p}{2}$.*

Proof. From (12.4.4) follows that v has at least one unbiased block. Since v is the longest, Proposition 12.4.1 implies that v has only one unbiased block, say B_j. If $m < \frac{p}{2}$, the argument used in the beginning of the proof of Proposition 12.4.4 yields that the bias of B_j is 0. If $m = \frac{p}{2}$, then the bias of B_j can be p as well. □

Corollary 12.4.2. *Let v be a longest common subsequence, let $\{B_1, \ldots, B_r\}$ be its blocks. Let $B_j = \{i_j, \ldots, i_j + s\}$ be its unbiased block. Let $b \in \{0, p\}$ be the bias of B_j. Then the length of $v|_{B_1 \cup \ldots \cup B_{j-1}}$ is $i_j - m - 1 + \frac{b}{2}$ and the length of $v|_{B_{j+1} \cup \ldots \cup B_r}$ is $8p^2 - (i_j + s) - \frac{b}{2}$.*

Proposition 12.4.5. *Let v be a longest common subsequence, let $B_j = \{i_j, \dots, i_j + s\}$ be the unbiased block of v. Let $b \in \{0, p\}$ be the bias of B_j. Then the integer interval $[mp + 1 - \frac{b}{2}, 8p^2 - m(p-1) - \frac{b}{2}] \subseteq B_j$. In particular, $[mp + 1, 8p^2 - mp] \subseteq B_j$.*

Proof. Let us first consider the case $b = 0$. By Corollary 12.4.2, the length of $v|_{B_1 \cup \dots \cup B_{j-1}}$ is $i_j - m - 1$. Since

$$v|_{B_1 \cup \dots \cup B_{j-1}} : \{1, \dots, i_j - 1\} \hookrightarrow \{1, \dots, i_j - m - 1\},$$

it holds that:

$$v|_{B_1 \cup \dots \cup B_{j-1}}(\{1, \dots, i_j - 1\}) = \{1, \dots, i_j - m - 1\}.$$

This means that

$$v(n_{j-1}) = i_j - m - 1 = |B_1| + \dots + |B_{j-1}|, \tag{12.4.5}$$

where $n_{j-1} = \max B_{j-1}$. Hence, by changing the blocks, v loses only the elements on the set where it is defined. Up to the block B_j there are $j - 1$ changes. Hence, v loses at least $j - 1$ elements, so that:

$$i_j > |B_1| + \dots + |B_{j-1}| + j - 1.$$

On the other hand, by (12.4.5):

$$i_j = |B_1| + \dots + |B_{j-1}| + (m + 1),$$

and thus $j - 1 < m + 1$ or $j - 1 \leq m$. Since the blocks $B_1, \dots, B_{j-1}$ are biased, their length is at most $p - 1$. Therefore, $i_j \leq m(p-1) + (m+1) = mp + 1$.
By Corollary 12.4.2, the length of $v|_{B_{j+1} \cup \dots \cup B_r}$ is at most $8p^2 - (i_j + s)$. Since

$$v|_{B_{j+1} \cup \dots \cup B_r} : \{i_j + s + 1, \dots, 8p^2\} \hookrightarrow \{i_j + s - m + 1, \dots, 8p^2\},$$

it holds:

$$\mathrm{Dom}(v|_{B_{j+1} \cup \dots \cup B_r}) = \{i_j + s + 1, \dots, 8p^2\}.$$

The last equality implies that:

$$8p^2 - (i_j + s) = |B_{j+1}| + \dots + |B_r|. \tag{12.4.6}$$

Hence, after B_j, by changing the blocks, v loses the elements on the image set, only. From B_j to B_r there are $r - j$ changes, so that:

$$v(i_j + s) + (r - j) + |B_{j+1}| + \dots + |B_r| \leq 8p^2.$$

Hence, with $v(i_j + s) = i_j + s - m$, we have that:

$$(r - j) \leq 8p^2 - (|B_{j+1}| + \dots + |B_r|) - v(i_j + s) = i_j + s - v(i_j + s) = m.$$

Therefore, (12.4.6) implies $8p^2 - (i_j + s) \leq m(p-1)$, so $i_j + s \geq 8p^2 - m(p-1)$.

Finally, let us consider the case $b = p$. This can happen only, when $m = \frac{p}{2}$. Then

$$v|_{B_1 \cup \dots \cup B_{j-1}} : \{1, \dots, i_j - 1\} \hookrightarrow \{1, \dots, i_j + m - 1\},$$
$$v|_{B_{j+1} \cup \dots \cup B_r} : \{i_j + s + 1, \dots, 8p^2\} \hookrightarrow \{i_j + s + m + 1, \dots, 8p^2\}$$

and the arguments used before yield $i_j \leq (p-1)m + 1$ and $8p^2 - (i_j + s) \leq mp$. □

Proposition 12.4.5 states that a certain neighborhood of $(4p^2+1)$ belongs to the unbiased block. This means that, for every longest common subsequence, the elements

$$x_{(4p^2+1)-p^2}, x_{(4p^2+1)-p^2+1}, \ldots, x_{4p^2+1}, \ldots, x_{(4p^2+1)+p^2}$$

are included and directly matched. In particular, the element x_{4p^2+1} belongs to the same block and are directly matched. Similarly, x_{2p^2+1+m} is directly matched. This implies that we can define $x_1, \ldots, x_n = x_{m+1}, \ldots, x_{m+n}$ and $y_1, \ldots, y_n = x_1, \ldots, x_n$. Then, for every longest common subsequence, the element x_{2p^2+1} is directly matched.

The structure of a common subsequence between periodic subsequences with unequal length

In the previous Subsection, we analyzed the longest common subsequences of two periodic subsequences with length $8p^2$ in detail. We now consider the longest common subsequences between two finite periodic subsequence with unequal length. We study the case, when one sequence is still with length $8p^2$ and length of the other sequence differs from $8p^2$ by at most $2(p-1)$. Our aim is still to show that any longest common subsequence contains a unbiased block that is located in the center.
The proofs used in the present Subsection are essentially the same as the ones in the previous Subsection, but for a few additional technicalities. Therefore, we leave the proofs for the Appendix.

Proposition 12.4.6. *Let $x_1, \ldots, x_{8p^2}$ and $x_{l-m_1+1}, \ldots, x_{l+8p^2+m_2}$ be the subsequences of a periodic sequence, with $0 \leq m_1 \leq p-1$, $-m_1 \leq m_2 \leq p-1$ and $l = jp$, for a $j \in \mathbb{Z}$. Let $t_1 = (p-m_1) \mod p$, $t_2 = \max\{-m_2, 0\}$. Then the length of the longest common subsequence is $8p^2 - \min\{t_1, t_2\}$ and any longest common subsequence between $x_1, \ldots, x_{8p^2}$ and $x_{l-m_1+1}, \ldots, x_{l+8p^2-1}$ includes an unbiased block which contains x_{4p^2+1}.*

Proposition 12.4.7. *Let $x_1, \ldots, x_{8p^2}$ and $x_{l+m_1+1}, \ldots, x_{l-m_2+8p^2}$ be the subsequences of a periodic sequence, $0 \leq m_1 \leq p-1$, $-m_1 \leq m_2 \leq p-1$ and $l = jp$, for a $j \in \mathbb{Z}$. Let $t_1 = (p-m_1) \mod p$, $t_2 = \max\{-m_2, 0\}$. Then the length of the longest common subsequence is $8p^2 - m_1 - m_2$, if $m_2 \geq 0$ and $8p^2 - \min\{m_1, p+m_2\}$, else. Moreover, any longest common subsequence between $x_1, \ldots, x_{8p^2}$ and $x_{l+m_1+1}, \ldots, x_{l-m_2+8p^2}$ includes an unbiased block which contains x_{4p^2+1}.*

The structure of a common subsequence between periodic subsequences with mismatch

In the present Subsection, we consider the subsequences of a periodic sequence with the length $8p^2$. The only difference is that sequence $x_1, \ldots, x_{8p^2}$ has a *mismatch :* the element x_{4p^2+1} has been changed. So, formally, we consider the sequences $z_1, \ldots, z_{8p^2}$ and $x_{m+1}, \ldots, x_{m+8p^2}$, where $z_i = x_i$, $i = 1, \ldots, 4p^2, 4p^2+2, \ldots, 8p^2$ and $z_{4p^2+1} \neq x_{4p^2+1}$.

Proposition 12.4.8. *Let $z_1, \ldots, z_{8p^2}$ and $x_{t+1}, \ldots, x_{t+8p^2+h}$ be the subsequences of a periodic sequence with mismatch, $0 \leq t \leq \frac{p}{2}, 0 \leq h \leq p-2t$. Then the length of the longest common subsequence is $8p^2 - t - 1$.*

Proof. Let v be a longest common subsequence of $z_1, \dots, z_{8p^2}$ and $x_{t+1}, \dots, x_{t+8p^2+h}$. The length of v is clearly at least $8p^2 - m - 1$.
Let us show that both subsequences $v|_{[1,4p^2]}$ and $v|_{[4p^2+2,8p^2]}$ have an unbiased block. By (12.4.4), v has at least one unbiased block $B_j = \{i_j, \dots, n_j\}$. Assume $i_j > 4p^2 + 1$. It holds that:

$$v|_{[1,i_j-1]} : \{1, \dots, i_j - 1\} \hookrightarrow \{1, \dots, i_j - 1 - m + b\},$$

where $b \in \{0, 2t\}$ is the bias of B_j. Clearly the length of $v|_{[1,i_j-1]}$ is at least $i_j - 1 - t + \frac{b}{2}$. Let $B_1, \dots, B_{r_1}$ be the blocks of $v|_{[1,i_j-1]}$. Suppose they all are biased. Then, with $u = i_j - (4p^2 + 2)$, we find:

$$\frac{r_1 - 1}{2} \geq \frac{i_j - 1 - (p-1) - t + \frac{b}{2}}{2(p-1)} = \frac{4p(p-1) + 2 + 3p + u - t + \frac{b}{2}}{2(p-1)} > 2p.$$

By Proposition 12.4.1, $i_j - 1 + \frac{b}{2} \geq 2p + |v|_{[1,i_j-1]}|$ or $|v|_{[1,i_j-1]}| \leq i_j - 1 + \frac{b}{2} - 2p$, which is a contradiction. Since the argument holds for any u, the unbiased block is contained in $\{1, \dots, 4p^2\}$.
Hence, $B_1, \dots, B_{r_1}$ contain at least one unbiased block.

Suppose the unbiased block B_j is contained in $\{1, \dots, 4p^2\}$. It holds that:

$$v|_{[n_j+1,8p^2]} : \{n_j + 1, \dots, 8p^2\} \hookrightarrow \{n_j + 1 + b, \dots, t + 8p^2 + h\},$$

where $h = 0$, if $b = 2t$. Then $|v|_{[n_j+1,8p^2]}| \geq 8p^2 - n_j - 1 - \frac{b}{2}$. Let $C_1, \dots, C_{r_2}$ be the blocks of $v|_{[n_j+1,8p^2]}$. Suppose they all are biased, hence, with $u = 4p^2 - n_j$,

$$\frac{r_2 - 1}{2} \geq \frac{4p(p-1) + 3p + u - \frac{b}{2}}{2(p-1)} > 2p.$$

By Proposition 12.4.1,

$$h + t + 8p^2 - n_j - 1 \geq |v|_{[n_j+1,8p^2]}| + 2p \geq 8p^2 - n_j - 1 + 2p - \frac{b}{2},$$

which is a contradiction. Since the argument holds for any u, the unbiased block is contained in $\{4p^2 + 1, \dots, 8p^2\}$.

Let $l > j$ and B_j,B_l be unbiased blocks: $B_i \subseteq \{1, \dots, 4p^2\}$, $B_l \subseteq \{4p^2 + 2, \dots, 8p^2\}$.
If $t < \frac{p}{2}$, then the bias of both blocks is 0. Since v is the longest common subsequence, it follows that $|v| = 8p^2 - t - 1$ and the blocks are consecutive: $l = j + 1$ and

$$B_j = \{i_j, \dots, 4p^2\}, \quad B_l = B_{j+1} = \{4p^2 + 2, \dots, 4p^2 + s\}. \tag{12.4.7}$$

If $\frac{t}{2}, p > 2$, then the bias of both blocks can be p as well. However, the length of v is still $8p^2 - t - 1$ and (12.4.7) holds. In both cases, the element z_{4p^2+1} is not included in v.
Finally, if $t = 1$ and $p = 2$, it might be that the bias of B_j is 0, the bias of B_{j+1} is 2 and the element z_{4p^2+1} is included in v. The length of v is still however equal to $8p^2 - t - 1$ □

Proposition 12.4.9. *Let $z_1, \dots, z_{8p^2}$ and $x_{m+1}, \dots, x_{m+8p^2-h}$ be the subsequences of a periodic sequence with mismatch, $0 \leq 2m \leq p + h, 0 \leq h \leq m$. Then the length of the longest common subsequence is $8p^2 - m - 1$.*

Proof. The proof of Proposition 12.4.8 holds without changes. □

Proposition 12.4.10. *Let $z_1, \ldots, z_{8p^2}$ and $x_{l-m_1+1}, \ldots, x_{l+8p^2+m_2}$ be the subsequences of a periodic sequence with mismatch, where $m_1 \leq p-1$, $-m_1 \leq m_2 \leq p-1$ and $l = jp$, for a $j \in \mathbb{Z}$. Let $t_1 = (p - m_1) \mod p$, $t_2 = \max\{-m_2, 0\}$. Then the length of the longest common subsequence is $8p^2 - \min\{t_1, t_2\} - 1$.*

Proposition 12.4.11. *Let $z_1, \ldots, z_{8p^2}$ and $x_{l+m_1+1}, \ldots, x_{l-m_2+8p^2}$ be the subsequences of a periodic sequence with mismatch, $m_1 \leq p-1$, $-m_1 \leq m_2 \leq p-1$ and $l = jp$, for a $j \in \mathbb{Z}$. Let $t_1 = (p - m_1) \mod p$, $t_2 = \max\{-m_2, 0\}$. Then the length of the longest common subsequence is $8p^2 - m_1 - m_2 - 1$, if $m_2 \geq 0$ and $8p^2 - \min\{m_1, p + m_2\} - 1$, else.*

12.4.3 Sequences with periodic pieces

Sequence with a periodic piece

Let $y_1, \ldots, y_n$ be a periodic sequence. Let $x_1, \ldots, x_n$ be a sequence with property:

$$\exists k \leq n - 8p^2 \text{ such that } x_{k+1} = y_{k+1}, x_{k+2} = y_{k+2}, \ldots, x_{k+8p^2} = y_{k+8p^2}. \tag{12.4.8}$$

So, the sequence $x_1, \ldots, x_n$ contains a periodic piece of length $8p^2$.
Let v be a longest common subsequence between $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$. We consider the integer interval $[\underline{v}(k+1), \bar{v}(k+8p^2)]$, and we show that the length of $[\underline{v}(k+1), \bar{v}(k+8p^2)]$ is about $8p^2$. The proofs of the following two propositions can be found in the Appendix.

Proposition 12.4.12. *Suppose the length of $[\underline{v}(k+1), \bar{v}(k+8p^2)]$ is not smaller than $8p^2$. Then there exist integers l, m_1, m_2 such that*

$$[\underline{v}(k+1), \bar{v}(k+8p^2)] = [l + 1 - m_1, l + 8p^2 + m_2], \tag{12.4.9}$$

where $|k - l| = jp$, for a non-negative $j \in \mathbb{N}$, $0 \leq m_1 \leq p-1$ and $-m_1 \leq m_2 \leq p-1$. In particular, the length of $[\underline{v}(k+1), \bar{v}(k+8p^2)]$ is at most $8p^2 + 2(p-1)$.

Proposition 12.4.13. *Suppose the length of $[\underline{v}(k+1), \bar{v}(k+8p^2)]$ is not bigger than $8p^2$. Then there exist integers l, m_1, m_2 such that*

$$[\underline{v}(k+1), \bar{v}(k+8p^2)] = [l + 1 + m_1, l + 8p^2 - m_2], \tag{12.4.10}$$

where $|k - l| = jp$, for a non-negative $j \in \mathbb{N}$ and $0 \leq m_1 \leq p-1$, $-m_1 \leq m_1 \leq p-1$. In particular, the length of $[\underline{v}(k+1), \bar{v}(k+8p^2)]$ is at least $8p^2 - 2(p-1)$.

Subsequence with a periodic piece and mismatch

Let $y_1, \ldots, y_n$ be a periodic sequence. Let $z_1, \ldots, z_n$ be a sequence with property: $\exists k \leq n - 8p^2$ such that

$$z_{k+1} = y_{k+1}, \ldots, z_{k+4p^2} = y_{k+4p^2}, z_{k+4p^2+1} \neq y_{k+4p^2+1}, z_{k+4p^2+2} = y_{k+4p^2+2}, \ldots, z_{k+8p^2} = y_{k+8p^2}. \tag{12.4.11}$$

Hence, the sequence $z_1, \ldots, z_n$ contains a periodic piece of length $8p^2$ with mismatch. From the proofs of Propositions 12.4.12 and 12.4.13, the following corollaries can be deduced.

Corollary 12.4.3. *There exists a longest common subsequence v between $z_1, \ldots, z_n$ and $y_1, \ldots, y_n$ such that either (12.4.9) or (12.4.10) holds.*

12.4.4 Proof of Lemma 12.3.1

Corollary 12.4.4. *Let $y_1, \dots, y_n$ be a periodic sequence. Let $x_1, \dots, x_n$ be a sequence with property (12.4.8). Then any longest common subsequence between $x_1, \dots, x_n$ and $y_1, \dots, y_n$ has an unbiased block that contains the element x_{k+4p^2+1}.*

Proof. Let v be a longest common subsequence between $x_1, \dots, x_n$ and $y_1, \dots, y_n$. We consider $[\underline{v}(k+1), \bar{v}(k+8p^2)]$. By optimality principle,

$$v|_{[k+1,k+8p^2]} : \{k+1, \dots, k+8p^2\} \hookrightarrow \{\underline{v}(k+1), \dots, \bar{v}(k+8p^2)\}$$

must be the longest common subsequence.
Suppose that the length of $[\underline{v}(k+1), \bar{v}(k+8p^2)]$ is bigger than $8p^2$. Then Proposition 12.4.12 and Proposition 12.4.6 apply.
Suppose that the length of $[\underline{v}(k+1), \bar{v}(k+8p^2)]$ is smaller than $8p^2$. Then Proposition 12.4.13 and Proposition 12.4.7 apply. □

Corollary 12.4.5. *Let L_n be the length of the longest common subsequence of a periodic sequence $y_1, \dots, y_n$ and a sequence $x_1, \dots, x_n$ with the property (12.4.8). Let $z_1, \dots, z_n$ be a sequence with the property (12.4.11). Then the length of the longest common subsequence of $y_1, \dots, y_n$ and $z_1, \dots, z_n$ is $L_n - 1$.*

Proof. Let v be a longest common subsequence between $z_1, \dots, z_n$ and $y_1, \dots, y_n$ that satisfies (12.4.9) ((12.4.10), resp.). By Corollary 12.4.3, such a v exists. Recall that $|L_n - |v|| \geq 1$. The length of v is the sum of the length of restrictions:

$$\begin{aligned} v|_{[1,k]} &: \{1, \dots, k\} \hookrightarrow \{1, \dots, \underline{v}(k+1) - 1\} \\ v|_{[k+1,k+8p^2]} &: \{k+1, \dots, k+8p^2\} \hookrightarrow \{\underline{v}(k+1), \dots, \bar{v}(k+8p^2)\} \\ v|_{[k+8p^2+1,n]} &: \{k+8p^2+1, \dots, n\} \hookrightarrow \{\underline{v}(k+8p^2+1), \dots, \bar{v}(n)\}. \end{aligned}$$

In this case, Proposition 12.4.10 (Proposition 12.4.11 resp.) specifies the length of $v|_{[k+1,k+8p^2]}$. Proposition 12.4.6 (Proposition 12.4.7 resp.) states: if $z_{k+1}, \dots, z_{k+8p^2}$ is replaced with $x_{k+1}, \dots, x_{k+8p^2}$, i.e. the mismatch has been removed, then there exists a common subsequence

$$v' : \{k+1, \dots, k+8p^2\} \hookrightarrow \{\underline{v}(k+1), \dots, \bar{v}(k+8p^2)\}$$

with length $|v|_{[k+1,k+8p^2]}| + 1$. Hence, the sequence v^* between $x_1, \dots, x_n$ and $y_1, \dots, y_n$, defined as

$$v^*|_{[1,k]} = v|_{[1,k]}, \;\; v^*|_{[k+1,k+8p^2]} = v', \;\; v^*|_{[k+8p^2+1,n]} = v|_{[k+8p^2+1,n]}$$

has length $|v| + 1$ and is, therefore, the longest common subsequence of $x_1, \dots, x_n$ and $y_1, \dots, y_n$. This proves the statement. □

Proof of Lemma 12.3.1. Let $x_1, \dots, x_n$ be a realization of $X_1, \dots, X_n$ such that l is a replica point. Denote $y_1, \dots, y_n := Y_1, \dots, Y_n$. Recall that L_n is the length of the longest common subsequence of $x_1, \dots, x_n$ and $y_1, \dots, y_n$, and L_n^l is the length of the longest common subsequence of $x_1, \dots, x_{l-1}, x_{l+1}, \dots x_n$ and $y_1, \dots, y_n$. Recall

$$L_n - 1 \leq L_n^l \leq L_n. \tag{12.4.12}$$

Assume that A_l holds, i.e. l is a replica point. If the replica point matches, then $x_1, \dots, x_n$ is a sequence satisfying (12.4.8) with $x_{k+4p^2+1} = x_l$ being the replica point. Let L_n^+ be the length of the longest common subsequence of $x_1, \dots, x_n$ and $y_1, \dots, y_n$ with matching replica point. Suppose $L_n^+ = L_n^l$. Then any longest common subsequence of s $x_1, \dots, x_{l-1}, x_{l+1}, \dots x_n$ and $y_1, \dots, y_n$ would also be a longest common subsequence of $x_1, \dots, x_n$ and $y_1, \dots, y_n$. This contradicts Corollary 12.4.4 which states that any longest common subsequence of $x_1, \dots, x_n$ and $y_1, \dots, y_n$ contains x_l. Hence, $L_n^+ = L_n^l + 1 = L_n^l + Z_n$.
Suppose that the replica point does not match. Then $x_1, \dots, x_n$ is a sequence as in (12.4.11) with $x_{k+4p^2+1} = x_l$ being the mismatching replica point. Let L_n^- be the length of the longest common subsequence of $x_1, \dots, x_n$ and $y_1, \dots, y_n$ with mismatching replica point. By Corollary 12.4.5, $L_n^- = L_n^+ - 1$. By (12.4.12), $L_n^l \le L_n^- = L_n^+ - 1 \le L_n^l$, i.e. $L_n^- = L_n^l$.

12.5 Appendix

Proof of Proposition 12.4.2. Assume that there exists $k < p$ such that (12.4.3) hold. Then

$$x_{mk+j} = x_j \quad \forall m \ge 1, \quad j = 1, \dots, p. \tag{12.5.1}$$

The latter implies

$$x_{k+n} = x_n \quad \forall n \ge 1.$$

that contradicts the definition of p.
Let us proof (12.5.1). Use induction: For $m = 1$, (12.5.1) is equivalent to (12.4.3).
Suppose that (12.5.1) holds for m. Let $k+j \le p$. Then $x_{(m+1)k+j} = x_{mk+(k+j)} = x_{k+j} = x_j$. If $k + j > p$, then $x_{(m+1)k+j} = x_{mk+(k+j)} = x_{mk+k+j-p} = x_{k+j-p} = x_{k+j} = x_j$, To get the third inequality note that from $j \le p$ follows $k + j - p < p$, and use (12.5.1).

12.5.1 Proofs of Propositions 12.4.6 and 12.4.7

Proposition 12.5.1. *Let $x_1, \dots, x_{8p^2}$ and $x_{t+1}, \dots, x_{t+8p^2+h}$ be the subsequences of a periodic sequence, $0 \le t \le \frac{p}{2}$, $0 \le h \le p - 2t$. Then the length of the longest common subsequence is $8p^2 - t$. Moreover, any longest common subsequence has an unbiased block B_j that contains the integer-interval $[tp + 1, 7p^2] \subseteq B_j$.*

Proof. Since $h \le p - 2t$, we have $p - (t + h) \ge t$, so t is the minimal bias between the two subsequences. In the proof of Proposition 12.4.4, replace the inequalities (12.4.4) with

$$\lfloor \frac{r-1}{2} \rfloor \ge \lfloor \frac{8p^2 - t + 1 - p}{2p} \rfloor \ge \lfloor 4p - \frac{t-1}{2p} - \frac{1}{2} \rfloor \ge 4p - 1 \ge t + h, \tag{12.5.2}$$

where the last inequality holds, because $t \le \frac{p}{2}$ and $h \le p$.
Let assume $b = 0$. Then the first half of the proof of Proposition 12.4.5 holds with any changes. For the second half, replace $8p^2$ by $8p^2 + h$. Then $8p^2 - (i_j + s) \le (t+h)(p-1) \le p(p-1)$ implying $(i_j + s) \le 7p^2 + p$. For $t = \frac{p}{2}$, $h = 0$. □

Proposition 12.5.2. *Let $x_1, \dots, x_{8p^2}$ and $x_{m+1}, \dots, x_{m+8p^2-h}$ be the subsequences of a periodic sequence, $0 \le 2m \le p + h$, $0 \le h \le m$. Then the length of the longest common*

subsequence is $8p^2-m$. *Moreover, any longest common subsequence has an unbiased block* B_j *that contains the integer-interval* $[mp+1, 8p^2-mp] \subseteq B_j$.

Proof. By assumption, $2m \leq p+h \leq m+p$, i.e., $m \leq p$. It holds, $p-m+h \geq m$, i.e. m is the minimal bias between the two subsequences. But it might be that $m > \frac{p}{2}$. The proof of Proposition 12.4.4 holds without any changes. Since $0 \leq h \leq m$, Proposition 12.4.5 holds, the only formal change is

$$v|_{B_{j+1}\cup\cdots\cup B_r} : \{i_j+s+1, \ldots, 8p^2\} \hookrightarrow \{i_j+s-m+1, \ldots, 8p^2-h\}. \tag{12.5.3}$$

□

Proposition 12.5.3. *Let* $x_{m_1+1}, \ldots, x_{m_1+8p^2}$ *and* $x_1, \ldots, x_{m_1+8p^2+m_2}$ *be the subsequences of a periodic sequence,* $0 \leq m_1, m_2 \leq p-1$. *The length of the longest common subsequence is* $8p^2$ *and each such subsequence of* $x_{m_1+1}, \ldots, x_{m_1+8p^2}$ *and* $x_1, \ldots, x_{m_1+8p^2+m_2}$ *includes an unbiased block which contains the interval* $[p^2, 7p^2]$.

Proof. Let $v : [1, 8p^2] \hookrightarrow [1, 8p^2+m_1+m_2]$ be a longest common subsequence, the length of v is clearly $8p^2$. Let $\{B_1, \ldots, B_r\}$ be the blocks of v. Suppose that all blocks are unbiased. Then $r \geq \frac{8p^2}{p-1}$. Since all the elements of the smallest subsequence are included in the longest common subsequence, by changing the blocks, v loses the elements on the bigger subsequence, only. Thus,

$$8p^2 + (r-1) = \sum_i^r |B_i| + (r-1) \leq 8p^2 + m_1 + m_2,$$

implying that $r-1 \leq m_1+m_2 \leq 2(p-1)$. This contradicts the lower bound for r.
Hence, there exists one and only one unbiased block $B_j = \{i_j, \ldots, i_j+s\}$. The bias of B_j can only be 0. Before the unbiased block, there are at most m_1 biased blocks, implying: $i_j \leq m_1(p-1) < p^2$. Similarly, $i_j + s \geq 7p^2$. □

Proposition 12.5.4. *Let* $x_1, \ldots, x_{8p^2}$ *and* $x_{m_1+1}, \ldots, x_{8p^2-m_2}$ *be the subsequences of a periodic sequence,* $0 \leq m_1, m_2 \leq p-1$. *The length of a longest common subsequence is* $8p^2-(m_1+m_2)$ *and every such a subsequence of* $x_1, \ldots, x_{8p^2}$ *and* $x_{m_1+1}, \ldots, x_{8p^2-m_2}$ *includes an unbiased block which contains the interval* $[p^2, 7p^2]$.

Proof. Let v be a longest common subsequence, the length of v is clearly $8p^2-m_1-m_2$. Let $\{B_1, \ldots, B_r\}$ be the blocks of v. Suppose that all blocks are unbiased. Then $r \geq \frac{8p^2-2p}{p-1}$. Since all the elements of the smallest subsequence are included in the longest common subsequence, by changing the blocks, v loses the elements on the bigger subsequence, hence. Thus,

$$8p^2 - (m_1+m_2) + (r-1) = \sum_i^r |B_i| + (r-1) \leq 8p^2,$$

implying that $r-1 \leq m_1+m_2 \leq 2p$. This contradicts with the lower bound of r.
So, there exists one and only one unbiased block $B_j = \{i_j, \ldots, i_j+s\}$. The bias of B_j is 0. Since before the unbiased block, there are at most t_1 biased blocks, we have: $i_j \leq m_1(p-1)+m_1 \leq p^2$. Similarly, $i_j + s + m_2(p-1) + m_2 \geq 8p^2$, so $i_j + s \geq 7p^2$. □

Proof of Proposition 12.4.6. Suppose $t_2 = 0$. Then Proposition 12.5.3 applies. If $t_1 = 0$, then $m_1 = 0$ and $t_2 = 0$, Proposition 12.5.3 applies again.
Suppose $t_1 > 0, t_2 > 0$. Assume $t_1 \leq t_2$. Note that $t_1 \leq \frac{p}{2}$. If not, then $m_1 = p - t_1 \leq \frac{p}{2}$, a contradiction with the assumption $m_1 \geq t_2$.
Since $l - m_1 = (l-1)p + t_1 = l^* + t_1$, we have

$$x_{l-m_1+1}, \ldots, x_{l-m_1+m_1+8p^2+m_2} = x_{l^*+t_1+1}, \ldots, x_{l^*+t_1+8p^2+m_2+m_1}.$$

Let $h = m_2 + m_1$. Clearly, $h = m_2 + m_1 \geq 0$ and $h = m_2 + m_1 = p - t_1 - t_2 \leq p - 2t_1$ since $-t_2 \leq -t_1$. Hence Proposition 12.5.1 applies.
Assume $t_1 \geq t_2$. Then $t_2 > \frac{1}{2}$ would imply that $m_1 > \frac{1}{2}$ and $t_1 \geq \frac{1}{2}$, a contradiction. We reverse the sequences, i.e we define

$$x'_1 = x_{8p^2}, x'_2 = x_{8p^2-1}, \ldots, x'_{8p^2} = x_1.$$

Then $x'_{t_2+1} = x'_{8p^2+m_2}, x'_{t_2+2} = x'_{8p^2+m_2-1}, \ldots, x'_{t_2+8p^2} = x_{m_2+1}, \ldots, x'_{t_2+8p^2+m_1-t_2} = x_{-m_1+1}$.
Take $h = m_1 - t_2 = m_1 + m_2 \geq 0$. It holds: $p - 2t_2 \geq p - t_1 - t_2 = m_1 - t_2 = h$. Now apply Proposition 12.5.1 to the reversed sequences. The reversing does not change the longest common subsequences (except reversing them). The element x_{4p^2+1} in the original sequence is the element x'_{4p^2}. By Proposition 12.5.1, it belongs to the unbiased block of any longest common subsequence.

Proof of Proposition 12.4.7. If $t_1 = 0$ then $m_2 \geq 0$. If $m_2 \geq 0$, then apply Proposition 12.5.4.
Let $0 < m_1 \leq p + m_2$. Define $h = m_1 + m_2 \geq 0$. Since $2m_1 \leq p + m_2 + m_1 = p + h$, Proposition 12.5.2 applies.
Let $0 < m_2 + p \leq m_1$. Then reverse the sequences as in the proof of Proposition 12.4.6 and apply Proposition 12.5.2.

12.5.2 Proofs of Proposition 12.4.10 and 12.4.11

Proposition 12.5.5. *Let $z_{m_1+1}, \ldots, z_{m_1+8p^2}$ and $m_1, \ldots, x_{m_1+8p^2+t_2}$ be the subsequences of a periodic sequence with mismatch, $0 \leq m_1, m_2 \leq p-1$. The length of the longest common subsequence is $8p^2 - 1$.*

Proof. Let v be a longest common subsequence of $z_{m_1+1}, \ldots, z_{m_1+8p^2}$ and $x_1, \ldots, x_{m_1+8p^2+m_2}$. By the argument used in the proof of Proposition 12.5.3, v has at least one unbiased block. The same argument, applied again, yields that the subsequences $v|_{[1,4p^2]}$ and $v|_{[4p^2+1,8p^2]}$ both have an unbiased block. If $p > 2$, then the bias of the unbiased blocks is 0, implying that the length of the longest common subsequence is $8p^2 - 1$.
When $p = 2$, the statement is easy to see. □

Proposition 12.5.6. *Let $z_{m_1+1}, \ldots, z_{m_1+8p^2}$ and $x_1, \ldots, x_{8p^2-m_2}$ be the subsequences of a periodic sequence with mismatch, $0 \leq m_1, m_2 \leq p-1$. The length of the longest common subsequence is $8p^2 - 1 - (m_1 + m_2)$.*

Proof. Let v be a longest common subsequence of $z_{m_1+1}, \ldots, z_{m_1+8p^2}$ and $x_1, \ldots, x_{m_1+8p^2-m_2}$. By the argument used in the proof of Proposition 12.5.4, v has at least one unbiased block, by the same argument, $v|_{[1,4p^2]}$ and $v|_{[4p^2+1,8p^2]}$ both have an unbiased block. If $p > 2$,

then the bias of the unbiased blocks is 0, implying that the length of the longest common subsequence is $8p^2 - (m_1 + m_2) - 1$.
When $p = 2$, the statement is easy to see. □

Proof of Proposition 12.4.10. Suppose $t_2 = 0$, i.e. $m_2 \geq 0$. If $t_1 = 0$, then $m_1 = 0$ and $t_2 = 0$. For $m_2 \geq 0$, Proposition 12.5.5 applies.
Suppose $t_1 > 0, t_2 > 0$. Assume $t_1 \leq t_2$. Then $t_1 \leq \frac{p}{2}$.
Since $l - m_1 = (l-1)p + t_1 = l^* + t_1$, we have

$$x_{l-m_1+1}, \ldots, x_{l-m_1+m_1+8p^2+m_2} = x_{l^*+t_1+1}, \ldots, x_{l^*+t_1+8p^2+m_2+m_1}.$$

Let $h = m_2 + m_1$. Clearly, $h = m_2 + m_1 \geq 0$ and $h = m_2 + m_1 = p - t_1 - t_2 \leq p - 2t_1$ since $-t_2 \leq -t_1$. Hence Proposition 12.4.8 applies.
Assume $t_1 \geq t_2$. Then $t_2 \leq \frac{1}{2}$. Reverse the sequences as in the proof of Proposition 12.4.6, i.e. we define

$$z'_1 = z_{8p^2}, z'_2 = z_{8p^2-1}, \ldots, z'_{8p^2} = z_1.$$

Note that in the reversed sequence, the mismatching element is z'_{4p^2} instead of z'_{4p^2+1}. However, it is easy to see that the proof of Propositions 12.4.8 holds also in this case.

Proof of Proposition 12.4.11. If $t_1 = 0$ then $m_2 \geq 0$. If $m_2 \geq 0$, then apply Proposition 12.5.6.
Let $0 < m_1 \leq p + m_2$. Define $h = m_1 + m_2 \geq 0$. Since $2m_1 \leq p + m_2 + m_1 = p + h$, Proposition 12.4.8 applies.
Let $0 < m_2 + p \leq m_1$. Then reverse the sequences as in the proof of Proposition 12.4.10 and apply Proposition 12.4.8.

12.5.3 Proofs of Propositions 12.4.12 and 12.4.13

Proof of Proposition 12.4.12. If $|[\underline{v}(k+1), \bar{v}(k+8p^2)]| = 8p^2$, the statement clearly holds. Suppose $|[\underline{v}(k+1), \bar{v}(k+8p^2)]| > 8p^2$. Then it holds: either $k+1 > \underline{v}(k+1)$ or $\bar{v}(k+8p^2) > (k+8p^2)$. Without loss of generality assume

$$\underline{v}(k+1) < k+1. \tag{12.5.4}$$

There $\exists l \geq 0$ such that $|k - l| = jp$, for a non-negative $j \in \mathbb{N}$ and

$$\underline{v}(k+1) = l - ip - m_1 + 1, \quad \bar{v}(k+8p^2) = l + 8p^2 + m_2,$$

where $0 \leq m_1 \leq p-1$ and $-m_1 \leq m_2 \leq p-1$, when $i = 0$ and $0 \leq m \leq p-1$, when $i \geq 1$.
The proposition is proven, if we show that $i = 0$. Suppose not. Then $0 \leq m \leq p-1$.
By the optimality principle, the subsequence

$$v|_{[k+1,k+8p^2]} : \{k+1, \ldots, k+8p^2\} \hookrightarrow \{l - ip - m_1 + 1, \ldots, l + 8p^2 + m_2\}$$

is the longest possible and its length is therefore equal to $8p^2$. Let

$$v' : \{k+1, \ldots, k+8p^2\} \hookrightarrow \{l+1, \ldots, l+8p^2\}$$

be a common subsequence that consists of a direct match:

$$v'(k+1) = l+1, \ldots, v'(k+8p^2) = l+8p^2.$$

The length of v' is also $8p^2$.
Let

$$w : \{1, \ldots, n\} \hookrightarrow \{1, \ldots, n\}$$

be a common subsequence of $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ that is defined as follows:

$$\begin{aligned} w|_{[1,k]} &= v|_{[1,k]} \\ w|_{[k+1,k+8p^2]} &= v' \\ w|_{[k+8p^2+1,n]} &= v|_{[k+8p^2+1,n]} \end{aligned}$$

Hence, w is a modification of v obtained by $v|_{[k+1,k+8p^2]}$ replaced by a direct matching v'. Of course, the length of w is the same as the length of v, hence, w is the longest common subsequence.
The subsequence w has the following property: $[1, \bar{w}(k)] = [1, l]$, but

$$\begin{aligned} \underline{w}(k+1) &= w(\max\{i \le k : i \in \mathrm{Dom}(w)\}) + 1 \\ &= v(\max\{i \le k : i \in \mathrm{Dom}(v)\}) + 1 = \underline{v}(k+1) = l - ip - m_1 + 1. \end{aligned}$$

Hence, the interval $[l - ip - m_1 + 1, l]$ does not contain any element of w. This means that the subsequence

$$w|_{[1,k]} : \{1, \ldots, k\} \hookrightarrow \{1, \ldots, l\} \tag{12.5.5}$$

is actually a subsequence

$$w|_{[1,k]} : \{1, \ldots, k\} \hookrightarrow \{1, \ldots, l - ip - m_1\}.$$

We shall show that this property contradicts the optimality principle.
By (12.5.4), $k > l - m_1 - ip$. Let

$$t = \max\{i \le k : i \notin \mathrm{Dom}(v)\}.$$

We have: $w(t+1), \ldots, w(k) \le l - ip - m_1$. Define $w' : \{1, \ldots, k\} \hookrightarrow \{1, \ldots, l\}$,

$$\begin{aligned} w'|_{[1,t]} &= w|_{[1,t]} \\ w'(t+1) &= w(t+1) + p, \ldots, w'(k) = w(k) + p. \end{aligned}$$

Since $w(k) \le l$, the sequence w' is well defined and has the length as (12.5.5). Let s be the last element of w before t, i.e. $s = \max\{i < t : i \in \mathrm{Dom}(w)\}$. By definition of w', $w'(t+1) = w(t+1) + p \ge w'(s) + 1 + p$, so the interval $[w'(s)+1, w'(s)+p]$ does not contain any elements of w'. By periodicity, the interval $[y_{w'(s)+1}, y_{w'(s)+p}]$ contains at least one 0 and at least one 1. On the other hand, the unconnected element x_t is either 0 or 1. Therefore, we can connect the element x_t with an element of $[y_{w'(s)+1}, y_{w'(s)+p}]$. The possibility of such a connection shows that w' is not the longest common subsequence. This, in turn, implies that (12.5.5) can not be the longest common subsequence. By the optimality principle, the latter implies that w and, hence, v cannot be the longest common subsequences as well. This is a contradiction. The reason for the contradiction is

the assumption $i \geq 1$.

Proof of Proposition 12.4.13. If $|[\underline{v}(k+1), \bar{v}(k+8p^2)]| = 8p^2$, the statement clearly holds. Suppose $|[\underline{v}(k+1), \bar{v}(k+8p^2)]| < 8p^2$. Then it holds: either $k+1 < \underline{v}(k+1)$ or $\bar{v}(k+8p^2) < (k+8p^2)$. Without loss of generality assume

$$\bar{v}(k+8p^2) < (k+8p^2). \tag{12.5.6}$$

There $\exists l \geq 0$ such that $|k-l| = jp$, for a non-negative $j \in \mathbb{N}$ and

$$\underline{v}(k+1) = l + m_1 + 1, \quad \bar{v}(k+8p^2) = l - ip + 8p^2 - m_2 =: u_l,$$

where $0 \leq m_1 \leq p-1$ and $-m_1 \leq m_2 \leq p-1$, when $i = 0$, and $0 \leq m \leq p-1$, when $i \geq 1$. Proposition is proved, if we show that $i = 0$. Suppose $i > 0$. Then $0 \leq m_1 \leq p-1$. By the optimality principle, the subsequence

$$v|_{[k+1,k+8p^2]} : \{k+1, \ldots, k+8p^2\} \hookrightarrow \{l+m_1+1, \ldots, u_l\}$$

is the longest possible, the length of it is, therefore, $L := 8p^2 - (m_1 + m_2 + ip)$. Let

$$v' : \{k+1, \ldots, k+8p^2\} \hookrightarrow \{l+m_1+1, \ldots, u_l\}$$

be a common subsequence that consists of a direct match:

$$v'(k+1+m_1) = l+m_1+1, \ldots, v'(k+8p^2-ip-m_2) = l-ip+8p^2-m_2 = u_l.$$

The length of v' is also L.
Let

$$w : \{1, \ldots, n\} \hookrightarrow \{1, \ldots, n\}$$

be a common subsequence of $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ that is defined as follows:

$$\begin{aligned} w|_{[1,k]} &= v|_{[1,k]} \\ w|_{[k+1,k+8p^2]} &= v' \\ w|_{[k+8p^2+1,n]} &= v|_{[k+8p^2+1,n]} \end{aligned}$$

Of course, the length of w is the same as the length of v, hence, w is the longest common subsequence of $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$.
The subsequence w has the following property: $u_k := k + 8p^2 - ip - m_2 \in \mathrm{Dom}(w)$, and the next element in $\mathrm{Dom}(w)$ is not earlier as $k+8p^2+1$: $\min\{i \geq u_k : i \in \mathrm{Dom}(w)\} \geq k+8p^2+1$. In particular, this implies: $\underline{w}(k+8p^2+1) = u_l + 1$ or

$$|w|_{[u_k+1,n]}| = |w|_{[k+8p^2+1,n]}|. \tag{12.5.7}$$

Note:

$$w|_{[k+8p^2+1,n]} : \{k+8p^2+1, \ldots, n\} \hookrightarrow \{u_l+1, \ldots, n\}.$$

By (12.5.6), $k+8p^2+1 < u_l+1$, so there exists at least one element $j \in [u_l+1, n]$ such that y_j does not belong to the subsequence $w|_{[k+8p^2+1,n]}$. Let

$$t = \min\{j \geq u_l + 1 : j \notin w([k+8p^2+1, n])\}. \tag{12.5.8}$$

Suppose $t \in [u_l + 1, l + 8p^2]$. Let r be such that $w(r) = t - 1$, i.e. $r = w^{-1}(t-1)$. Obviously, $r \in [k + 8p^2 + 1, n]$. Define

$$v'' : \{k+1, \ldots, u_k + (t - u_l)\} \hookrightarrow \{l + m_1 + 1, \ldots, t\}$$

be a common subsequence that consists of a direct match:

$$v''(k+1+m_1) = l + m_1 + 1, \ldots, v''(u_k) = u_l, v''(u_k+1) = u_l + 1, \ldots, v''(u_k + (t-u_l)) = t.$$

The definition of v'' is possible, since $t - u_l \leq ip + m_2$ and $(t - u_l) \leq k + 8p^2$.
The length of v'' is $L + (t - u_l)$. Define $w' : \{k+1, \ldots, n\} \hookrightarrow \{l + m_1 + 1, \ldots, n\}$,

$$\begin{aligned} w'|_{[k+1, u_k + (t-u_l)]} &= v'' \\ w'|_{[u_k + (t-u_l)+1, n]} &= v|_{[r+1, n]}. \end{aligned} \tag{12.5.9}$$

By the definition of t and r, $|v|_{[k+8p^2+1,n]}| - |v|_{[r+1,n]}| = |v|_{[k+8p^2+1,r]}| = (t - u_r) - 1$. Hence, the length of w' is $L + (t - u_l) + |v|_{[k+8p^2+1,n]}| - (t - u_r) + 1 = L + |v|_{[k+8p^2+1,n]}| + 1 = |w|_{[k+1,n]}| + 1$ which contradicts the assumption that w is a longest common subsequence.

Suppose $t \in [l + 8p^2 + 1, n]$. Then $t - p \geq u_l + 1$ and by the definition of t, the elements $y_{u_l+1}, \ldots, y_{t-1}$ all belong to the common subsequence w. Let

$$v'' : \{u_k + 1, \ldots, w^{-1}(t-p)\} \hookrightarrow \{u_l + 1, \ldots, t\}$$

be defined as follows:

$$v''(u_k+1) = u_l+1, \ldots, v''(u_k+p) = u_l+p, v''(w^{-1}(u_l+1)) = u_l+1+p, \ldots, v''(w^{-1}(t-p)) = t.$$

The definition of v'' is possible, because $w^{-1}(u_l + 1) \geq k + 8p^2 + 1$. The length of v'' is $|w|_{[k+8p^2+1, w^{-1}(t-p)]}| + p$. Note

$$|w|_{[w^{-1}(t-p)+1, w^{-1}(t-1)]} = |w|_{[k+8p^2+1, w^{-1}(t-1)]}| - |w|_{[k+8p^2+1, w^{-1}(t-p)]}| = p - 1.$$

We define $w' : \{u_k + 1, \ldots, n\} \hookrightarrow \{u_l + 1, \ldots, n\}$, where

$$\begin{aligned} w'|_{[u_k+1, w^{-1}(t-p)]} &= v'', \\ w'|_{[w^{-1}(t-p)+1, n]} &= w|_{[w^{-1}(t-1)+1, n]}. \end{aligned}$$

The definition of w' is correct, because $w(w^{-1}(t-1)+1) > t$. The length of w' is

$$|w|_{[k+8p^2+1, w^{-1}(t-p)]}| + p + |w|_{[w^{-1}(t-1)+1, n]}| = |w|_{[k+8p^2+1, w^{-1}(t-1)]}| + |w|_{[w^{-1}(t-1)+1, n]}| + 1 = \\ = |w|_{[k+8p^2+1, n]}| + 1.$$

By (12.5.7), the length of w' is strictly bigger than that of

$$w|_{[u_k+1, n]} : \{u_k + 1, \ldots, n\} \hookrightarrow \{u_l + 1, \ldots, n\}.$$

This contradicts the assumption that w is a longest common subsequence.

Proof of Corollary 12.4.3 Let v be a longest common subsequence of $z_1, \ldots, z_n$ and

$y_1, \dots, y_n$. Suppose $[\underline{v}(k+1), \bar{v}(k+1)]$ is bigger than $8p^2$ but does not satisfy (12.4.9). Then there exists $0 \le m_1, m_2 \le p-1$, $i > 0$, such that

$$v|_{[k+1,k+8p^2]} : \{k+1, \dots, k+8p^2\} \hookrightarrow \{l - ip - m_1 + 1, \dots, l + 8p^2 + m_2\}. \qquad (12.5.10)$$

Suppose $i \ge 2$. Then, assuming (12.5.4), it holds $\underline{v}(k+1) + 2 < k+1$. The length of $|v|_{[k+1,k+8p^2]}|$ is $8p^2$. Define the common subsequence w as in the proof of Proposition 12.4.12. The length of w is $|v|-1$, but the length of the empty interval $[l-ip-m_1+1, l]$ is at least $2p$. Since there are at least two elements in $[1,k]$, say t_1 and t_2, not included into $\mathrm{Dom}(v)$, by rearranging the elements of $w|_{[1,k]}$ as in the proof of Proposition 12.4.12, both $z_{t_1} = x_{t_1}$ and $z_{t_2} = x_{t_2}$ can be matched with an empty period. So, the length of $w|_{[1,k]}$ can be increased by 2. This contradicts the assumption that v is the longest common subsequence.
This means that in (12.5.10), $i = 1$. Now, again, use the argument of Proposition 12.4.12: Define the common subsequence w and note that the length of w is $|v|-1$. Then rearrange the elements of $w|_{[t,k]}$ by defining $w'(t+1) = w(t+1) + p, \dots, w'(k) = w(k) + p = l - p - m_1 + p = l - m_1$ and connect the element x_t with some element on $[y_{w'(s)+1}, y_{w'(s)+p}]$. Let

$$w^* : \{1, \dots, k\} \hookrightarrow \{1, \dots, l - m_1\},$$

be modification of w' with connected x_t so the length of w^* is $|w|_{[1,k]}| + 1$. Hence, the sequence v^* with

$$\begin{aligned} v^*|_{[1,k]} &= w^* \\ v^*|_{[k+1,k+8p^2]} &= w \\ v^*|_{[k+8p^2+1,n]} &= w|_{[k+8p^2+1,n]} \end{aligned}$$

has length $|w| + 1$, which is the same as the length of v. Since $v^*(k) = w^*(k) = w'(k) = l - m_1$, the sequence v^* satisfies (12.4.9).

Suppose $[\underline{v}(k+1), \bar{v}(k+1)]$ is bigger than $8p^2$ but does not satisfy (12.4.10). The proof is similar: as in the proof of Proposition 12.4.13, define the subsequence w and note that the length of w is $|v| - 1$. Define t as in (12.5.8), and w' as in (12.5.9). With the help of w', construct the common subsequence v^* with $v^*|_{[1,k]} = v|_{[1,k]}$ and $v^*|_{[k+1,n]} = w'$. The length of v^* is the same as the length of v. If v^* does not satisfy (12.4.10), then use the proof of Proposition 12.4.13 to see that w^* satisfies (12.4.10).

References

[1] Kenneth S. Alexander. The rate of convergence of the mean length of the longest common subsequence. *Ann. Appl. Probab.*, 4(4):1074–1082, 1994.

[2] Richard Arratia and Michael S. Waterman. A phase transition for the score in matching random sequences allowing deletions. *Ann. Appl. Probab.*, 4(1):200–225, 1994.

[3] Federico Bonetto and Heinrich Matzinger. Fluctuations of the longest common subsequence in the case of 2- and 3-letter alphabets. *in preparation*, 2004.

[4] Federico Bonetto and Heinrich Matzinger. Simulations for the longest common subsequence problem. *in preparation*, 2004.

[5] Václáv Chvatal and David Sankoff. Longest common subsequences of two random sequences. *J. Appl. Probability*, 12:306–315, 1975.

[6] Michael S. Waterman. General methods of sequence comparison. *Bull. Math. Biol.*, 46(4):473–500, 1984.

[7] Michael S. Waterman. Estimating statistical significance of sequence alignments. *Phil. Trans. R. Soc. Lond. B*, 344:383–390, 1994.

[8] M.S. Waterman and M. Vingron. Sequence comparison significance and poisson approximation. *Statistical Science*, 9(3):367–381, 1994.