Detecting a local perturbation in a continuous scenery

Heinrich Matzinger[1] Serguei Popov[*,2]

December 13, 2005

[1]School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332–0160, USA
e-mail: matzi@math.gatech.edu
[2]Instituto de Matemática e Estatística, Universidade de São Paulo, rua do Matão 1010, CEP 05508–090, São Paulo SP, Brasil
e-mail: popov@ime.usp.br

Abstract

A continuous one-dimensional scenery is a double-infinite sequence of points (thought of as locations of *bells*) in $\mathbb{R}$. Assume that a scenery X is observed along the path of a Brownian motion in the following way: when the Brownian motion encounters a bell different from the last one visited, we hear a ring. The trajectory of the Brownian motion is unknown, whilst the scenery X is known except in some finite interval. We prove that given only the sequence of times of rings, we can a.s. reconstruct the scenery X entirely. For this we take the scenery X to be a local perturbation of a Poisson scenery X'. We present an explicit reconstruction algorithm. This problem is the continuous analog of the "detection of a defect in a discrete scenery" as studied by Kesten [13] and Howard [9, 10]. Many of the essential techniques used with discrete sceneries do not work with continuous sceneries.

Keywords: Brownian motion, Poisson process, localization test, detecting defects in sceneries seen along random walks

[*]This article was written with funding from SFB 701 A3 and CNPq (302981/02–0)

1 Introduction and main result

Suppose that countably many bells are placed on $\mathbb{R}$. Start a Brownian motion from 0; each time it hits a bell *different* from the last one visited, we hear a ring. During this process all the bells remain in the same position. The set of locations of the bells in $\mathbb{R}$ is referred to as the *scenery.* Suppose now that we cannot observe the trajectory of the Brownian motion, and that the scenery is not completely known either. On the other hand, let the sequence of time occurrences of the rings be known to us.

The *detection of a local perturbation* problem can be formulated as follows: is it possible to recover the exact scenery a.s. given only the sequence of rings and the scenery up to a local perturbation?

In this paper, we answer this question affirmatively provided that the scenery is a local perturbation of a random realization of a one-dimensional Poisson process with bounded rate. The realization of the one-dimensional Poisson process is known to us but we do not know in which way and where it was perturbed.

This problem is the continuous analog of the problem of detecting a defect in a scenery seen along the path of a random walk. For periodic sceneries this problem was studied by Howard [9, 10], whilst for random sceneries with at least four colors the detection problem was solved by Harry Kesten [13]. In the discrete case (not in this paper) one considers a discrete scenery $\xi : \mathbb{Z} \to \{0, 1, \dots, C-1\}$ and a random walk $\{S_t\}_{t\in\mathbb{N}}$. The discrete scenery is a coloring of the integers with C colors. One observes the discrete scenery seen along the path of the random walk, i.e. the sequence $\chi_0, \chi_1, \dots$, where $\chi_i := \xi(S_i)$. From this one tries to infer about ξ.

It is worth noticing that in the case of the present paper, i.e. in the case of a continuous scenery, there are no "colors": all the bells ring in the same way. Hence, we have to use the time length between successive rings to estimate where the random walk is. It turn out that bells close to each other tend to confer a lot of information. In discrete scenery reconstruction it is usually the opposite: long blocks of only one color are the essential "markers".

The continuous case considered here contains one of the major difficulties still open in discrete scenery reconstruction. Roughly speaking, in any part of the scenery one can obtain any finite set of observations in the continuous case. Some finite set of observations might be untypical but are never impossible. In all the discrete cases, where scenery reconstruction has been proven possible, there exist patterns which can appear in the observations only when the random walk dwells in some specific regions of the scenery. This is one more reason which makes it worthwhile studying the continuous case.

Also, we should mention that one of the main techniques used in discrete reconstruction does not work here. This is the "going in a straight path from x to y" as is used in a majority of discrete reconstruction papers. Instead we use an estimate of the probability to hear a ring a certain amount of time after being at a marker.

There exists one other related continuous problem solved by Burdzy [3]. He takes an iterated Brownian motion and shows that the path of the outer one can be a.s. reconstructed. This is the continuous analog of reconstructing a random walk path given an iterated random walk path. Matzinger [25] proved that the reconstruction of a 3-color scenery seen along a simple random walk is equivalent to this problem.

Let us present more on the history of discrete scenery reconstruction now. 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. Benjamini and Kesten [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 [10] showed that any pair of periodic, non-equivalent sceneries are distinguishable, as well as periodic sceneries with a single defect [9].

The problem of distinguishing two sceneries which differ at only one point is called detecting a single defect in a scenery. Kesten [13] 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 [24, 25, 26] 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 and independent of ξ and S. The variables ν_t are used to indicate at which times there are errors in the observations. More precisely, if $\nu_t = 1$ then there is an error in the observation at time t. Let $\tilde{\chi}$ denote the sequence of observations χ corrupted by the errors $\{\nu_t\}_{t\in\mathbb{N}}$. Thus, $\tilde{\chi}_t = \chi_t$ when $\nu_t = 0$ and $\tilde{\chi}_t \neq \chi_t$ otherwise. Knowing ξ^a and ξ^b, can we decide a.s. whether $\xi = \xi^a$ or $\xi = \xi^b$ 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 the present article is closely related to a random coin tossing problem which was first investigated by Harris and Keane in [8] and later by Levin, Pemantle and Peres in [23]. They take the error-probability to be equal to 1/2. The *coin tossing problem of Harris and Keane* can be described as follows: Let $X_1, X_2, \dots$ 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 $\tau_1, \tau_2, \dots$ 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 [8] and later by Levin, Pemantle and Peres in [23] 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 observed sequence 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 [23], set the scenery $\xi^a \equiv 0$ and ξ^b to be zero everywhere except at a finite number of points. They take the error probability to be 1/2, except when a "one" is observed. Hence, in their case, $P[\tilde{\chi}_t = 0 \mid \chi_t = 0] = 1/2$, but $P[\tilde{\chi}_t = 0 \mid \chi_t = 1] \neq 1/2$. In the case when the scenery is i.i.d. and has many colors, but is observed with a small error probability, the reconstruction problem was solved by Rolles and Matzinger in [27].

There is an excellent overview of scenery reconstruction and scenery distinguishing by Kesten [14]. 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 [11]. Several important contributions about the properties of the observations were made later. These include Keane and den Hollander [12], den Hollander [4], den Hollander and Steif [5], Heicklen, Hoffman and Rudolph [7], and Levin and Peres [22]. Interest in the scenery distinguishing problem was sparked when Keane and den Hollander, as well as Benjamini, asked if all non-equivalent sceneries could be distinguished. Lindenstrauss was able to prove that there exist pairs of sceneries which can not be distinguished [16]. 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 [26, 25]), Kesten noticed that Matzinger's 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, Löwe and Matzinger [18] were able to prove that scenery reconstruction is also possible on two-dimensional sceneries with many colors. Later, Löwe, Matzinger and Merkl [20] 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 [19]. Den Hollander asked if it would be possible to do reconstruction if the jumps made by the random walk are not bounded. Lenstra and Matzinger [17] were able to answer this question. Finally, following a question of den Hollander, Löwe and Matzinger [21] 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 [28, 29, 30].

Let us start with the formal definitions used in this paper. A *scenery* is a double infinite sequence $X = (\ldots, X_{-1}, X_0, X_1, \ldots)$, such that $X_n < X_{n+1}$ for all $n \in \mathbb{Z}$ and $\lim_{n\to-\infty} X_n = -\infty$, $\lim_{n\to+\infty} X_n = +\infty$. The last condition guarantees that the number of points of X in any finite interval is finite.

With some abuse of notation, we denote the set of points in the scenery by the same letter, $X = \{\ldots, X_{-1}, X_0, X_1, \ldots\}$. Let $\mathcal{M}$ be the set of all such sceneries. Let $\xi(n) := X_n - X_{n-1}$ for all $n \in \mathbb{Z}$. The sequence ξ is thus the sequence of distances between the successive bell-locations.

Definition 1.1 *Scenery $\tilde{X}$ is a local perturbation of X if they coincide everywhere except possibly in a finite interval, i.e., there exist $a, b \in \mathbb{R}$ such that $\tilde{X} \setminus [a, b] = X \setminus [a, b]$.*

Let $(W_t, t \geq 0)$ be the standard Brownian motion (starting from 0, unless otherwise indicated). When it is necessary to consider a Brownian motion starting from an arbitrary $x \in \mathbb{R}$, we use the notations $\mathbb{P}^x, \mathbb{E}^x$ for the corresponding probability and expectation. Let $\mathcal{M}^+$ be the set of all infinite sequences $U = (0 = U_0, U_1, U_2 \ldots)$, such that $U_n < U_{n+1}$ for all $n \in \mathbb{Z}_+$, and such that $\lim_{n\to+\infty} U_n = +\infty$. Using the scenery X and the trajectory of the Brownian motion W_t, we define the specific sequence of stopping times $Y = (0 = Y_0, Y_1, Y_2 \ldots) \in \mathcal{M}^+$ that corresponds to the sequence of ringing-times. More precisely (see Figure **??**, the marks on the horisontal line correspond to the bells, the marks on the vertical line correspond to the rings):

$$Y_{n+1} := \inf \big\{ t \geq Y_n : W_t \in X \setminus \{W_{Y_n}\} \big\},$$

$n \geq 0$ (note that the sequence Y always begins with 0, regardless of whether $0 \in X$ or not). From the fact that $X \in \mathcal{M}$ it is elementary to obtain that $Y_{n+1} > Y_n$ for all $n \in \mathbb{Z}_+$, and that $\lim_{n\to+\infty} Y_n = +\infty$ a.s., so indeed $Y \in \mathcal{M}^+$. Denote by $\chi(n)$, $n = 1, 2, 3, \ldots$, the sequence of time lapses between successive rings. Hence, $\chi(n) := Y_n - Y_{n-1}$.

Our main result is the following

Theorem 1.1 *Suppose that (the known scenery) X' is a realization of a one-dimensional inhomogeneous Poisson process with intensity bounded away from 0 and $+\infty$. Suppose also that (the unknown scenery) X is a local perturbation of X', and Y is the sequence of rings defined above. Then, almost surely, Y and X' together determine X. In other words, there exists a measurable function $\Psi_{X'} : \mathcal{M}^+ \mapsto \mathcal{M}$ such that $\mathbb{P}[\Psi_{X'}(Y) = X] = 1$.*

2 Proof of Theorem 1.1

In the proof of this theorem we will suppose for definiteness that X' is a realization of a Poisson process with rate 1, the general case is completely analogous.

The idea of the proof is, roughly speaking, the following: we use couples of bells which are untypically close to each other. The distance to neighbouring bells in the scenery should be much larger. The Brownian motion is likely to produce a long sequence of rings separated by short time intervals when visiting such a couple of bells (one can observe this on Figure **??**). The Brownian motion tends to visit the two bells many times before moving on to another bell in the scenery.

When we hear many rings shortly after one another, then this is likely to be caused by two bells at short distance form each other in the scenery. Hence a sequence many rings in a short time permits to estimate the distance between the underlying two bells. (Provided the sequence was really generated on only two bells close to each other, which is likely.) We discuss this in Section 2.1. Then, for a given (large) n, we define a location ζ_n (with a bell there) and construct a sequence of stopping times $\tau_i^{(n)}$ depending only on Y and X' (i.e., on known information) such that, with overwhelming probability $W_{\tau_i^{(n)}} = \zeta_n$, whenever i is not too large. In other words, with large probability we are able to tell whether we are back to the same place. For this we use the information provided by the estimated distances between couple of bells close to each other. This is done in Section 2.2 (see Lemma 2.5). In Section 2.3, we present an algorithm for reconstructing the local perturbation with a high precision, then we consider a sequence of such algorithms which permits us to reconstruct X exactly; however, this is done supposing that the interval where the perturbation took place is known. In Section 2.4 we explain the reconstruction procedure in the case when the interval of perturbation is unknown.

2.1 The main idea: trills and couples

Fix some $\varepsilon_0, \delta_0, \delta_1 > 0$ such that

$$\varepsilon_0 + \delta_0 + \delta_1 < 1/2, \quad (1)$$

$$6\varepsilon_0 < \delta_0. \quad (2)$$

Let z_0 be such that

$$\int_{z_0}^{+\infty} (2\pi u^3)^{-1/2} \exp\Big(-\frac{1}{2u}\Big) du = \frac{1}{2}. \tag{3}$$

Denote also

$$\mathfrak{A}_n^k = \big(z_0^{-1}\mathsf{median}\{\chi(k+1),\dots,\chi(k+n^{\delta_0/2})\}\big)^{1/2}.$$

The next two definitions play an important role in our construction.

Definition 2.1 *We say that there is a level-n trill at the mth position of the sequence Y, if $\chi(m+k) \le n^{-2+2\varepsilon_0+2\delta_0+\delta_1}$ for all $k = 1,\dots,n^{\delta_0/2}$.*

Definition 2.2 *Suppose that there is a level-n trill at the mth position of the sequence Y. We say that this trill is good, if $\mathfrak{A}(m) \le n^{-1+\varepsilon_0}$.*

The main idea is that if there is a good level-n trill in the kth position of the sequence Y, it is very probable that it was produced by the alternating visits of the Brownian motion to some two neighboring points from X that are roughly $\mathfrak{A}_n(k)$ away from each other (by alternating visits we mean here that the rings in the piece of the sequence Y under consideration were caused by only two bells). Consider the following

Definition 2.3 *A pair of two consecutive points (X_k, X_{k+1}) is called level-n couple if $\xi(k+1) = X_{k+1}-X_k \le n^{-1+\varepsilon_0}(1-z_0^{-1}n^{-\delta_0/6})^{-1}$, and $\min\{\xi(k),\xi(k+2)\} \ge n^{-1+\varepsilon_0+\delta_0+\delta_1}$.*

Let $T_a = \inf\{t \ge 0 : W_t = a\}$ be the hitting time of $a > 0$ by Brownian motion. Then, provided that the Brownian motion starts at 0, the density $f_a(s)$ of T_a is given by (see [2], formula **1**.2.0.2)

$$f_a(s) = a(2\pi s^3)^{-1/2} \exp\Big(-\frac{a^2}{2s}\Big). \tag{4}$$

We recall also the following elementary fact: if $a < b < c$, then (see [2], formula **1**.3.0.4)

$$\mathbb{P}^b[T_a < T_c] = \frac{c-b}{c-a}. \tag{5}$$

Let us consider now a level-n couple (X_k, X_{k+1}). Abbreviate $a := X_k - n^{-1+\varepsilon_0+\delta_0}$, $b := X_k$, $c := X_{k+1}$, $d := X_{k+1} + n^{-1+\varepsilon_0+\delta_0}$. Note that, by Definition 2.3, it holds that $X_{k-1} < a$ and that $X_{k+2} > d$. By (5), there is $C_1 > 0$ such that

$$\min\{\mathbb{P}^b[T_c < T_a], \mathbb{P}^c[T_b < T_d]\} \geq 1 - C_1 n^{-\delta_0},$$

so for any $x \in \{b, c\}(= \{X_k, X_{k+1}\})$

$$\mathbb{P}[W_{Y_{m+s}} \in \{b, c\} \text{ for any } 1 \leq s \leq n^{\delta_0/2} \mid W_{Y_m} = x] \geq 1 - C_1 n^{-\delta_0/2}, \quad (6)$$

i.e., with a large probability the Brownian motion will commute between the points of a level-n couple at least $n^{\delta_0/2}$ times. Now, it is elementary to see that

$$\mathbb{P}^b[\min\{T_a, T_c\} \leq n^{-2+2\varepsilon_0+2\delta_0+\delta_1} \mid T_c < T_a] \geq 1 - \exp(-C_2 n^{\delta_1}) \quad (7)$$

and that the same bound holds if b, a, c are substituted by c, d, b (in this order). Indeed, since the conditional density of $\min\{T_a, T_c\}$ is known (see **1**.3.0.6 of [2]), it is possible to obtain (7) by a direct (although not so simple) computation. In any case, to see that (7) holds, it is sufficient to consider the following intuitive argument: for any starting point within the interval $[a, c]$, the probability that the Brownian motion hits $\{a, c\}$ in time at most $n^{-2+2\varepsilon_0+2\delta_0}$ is bounded away from 0. The time interval $[0, n^{-2+2\varepsilon_0+2\delta_0+\delta_1}]$ contains n^{δ_1} non-intersecting intervals of length $n^{-2+2\varepsilon_0+2\delta_0}$, so we have at least n^{δ_1} tries to enter $\{a, c\}$. Also, it is easy to obtain that the conditional probability in (7) is of the same order as the unconditional one. Thus, using (7), we obtain that

$$\begin{aligned} \mathbb{P}[\chi(m+s) \leq n^{-2+2\varepsilon_0+2\delta_0+\delta_1} & \text{ for any } 1 \leq s \leq n^{\delta_0/2} \\ & \mid Y_m = x, Y_{m+s} \in \{b, c\} \text{ for any } 1 \leq s \leq n^{\delta_0/2}] \\ & \geq \ 1 - n^{\delta_0/2} \exp(-C_2 n^{\delta_1}). \end{aligned} \quad (8)$$

for any $x \in \{b, c\}$. This shows that if the Brownian motion commutes between b and c (without hitting other points of X) at least $n^{\delta_0/2}$ times, then, with overwhelming probability, we obtain a level-n trill. To show that this trill should normally be good, we have to work a bit more.

First, let us recall the Chernoff's bound for the binomial distribution:

Lemma 2.1 *[see e.g. [31], p. 68.] Let $\{\zeta_i, i \geq 1\}$ be i.i.d. random variables with $\mathbb{P}[\zeta_i = 1] = \theta$ and $\mathbb{P}[\zeta_i = 0] = 1 - \theta$. Then for any $0 < \theta < \alpha < 1$ and for any $s \geq 1$ we have*

$$\mathbb{P}\Big[\frac{1}{s}\sum_{i=1}^{s} \zeta_i \geq \alpha\Big] \leq \exp\{-sH(\alpha,\theta)\}, \tag{9}$$

where

$$H(\alpha,\theta) = \alpha \log\frac{\alpha}{\theta} + (1-\alpha)\log\frac{1-\alpha}{1-\theta} > 0.$$

If $0 < \alpha < \theta < 1$, then (9) holds with $\mathbb{P}[s^{-1}\sum_{i=1}^{s}\zeta_i \leq \alpha]$ in the left-hand side.

Now, we define another sequence of stopping times $(Y'_m, m \geq 0)$, constructed in a similar way as the sequence Y, this time supposing, however, that the only bells are in b and c (i.e., in X_k and in X_{k+1}):

$$\begin{aligned} Y'_0 &= 0, \text{ and} \\ Y'_{n+1} &= \inf\big\{t \geq Y'_n : W_t \in \{b,c\} \setminus \{W_{Y'_n}\}\big\}. \end{aligned}$$

Analogously, define $\chi'(i) = Y'_i - Y'_{i-1}$ and

$$\mathfrak{A}'_n = \big(z_0^{-1}\mathsf{median}\{\chi'(1),\dots,\chi'(n^{\delta_0/2})\}\big)^{1/2}.$$

Lemma 2.2 *There is a positive constant γ_1 such that for all n large enough we have*

$$\begin{aligned} \mathbb{P}\big[\beta^2(z_0 - n^{-\delta_0/6}) \leq \mathsf{median}\{\chi'(1),\dots,\chi'(n^{\delta_0/2})\} \leq \beta^2(z_0 + n^{-\delta_0/6})\big] \\ \geq \; 1 - \exp(-\gamma_1 n^{-\delta_0/6}) \end{aligned} \tag{10}$$

and also

$$\mathbb{P}\big[\mathfrak{A}'_n(1 - z_0^{-1}n^{-\delta_0/6}) \leq \beta \leq \mathfrak{A}'_n(1 + z_0^{-1}n^{-\delta_0/6})\big] \geq 1 - \exp(-\gamma_1 n^{-\delta_0/6}), \tag{11}$$

where $\beta := c - b = X_{k+1} - X_k$.

Proof. Denote

$$Z = \mathsf{median}\{\chi'(1),\dots,\chi'(n^{\delta_0/2})\},$$

and for any $y \in (0,1)$ let $\hat{M}_y$ be such that

$$\int_0^{\hat{M}_y} \beta(2\pi s^3)^{-1/2}\exp\Big(-\frac{\beta^2}{2s}\Big) = y. \tag{12}$$

Fix a number $p \in (0,1/2)$ (to be chosen later), and define the random variable $\eta_i = \mathbf{1}\{\chi'(i) \geq \hat{M}_{\frac{1}{2}+p}\}$, so that $\mathbb{P}[\eta_i = 1] = 1 - \mathbb{P}[\eta_i = 0] = \frac{1}{2} - p$. Now, we have

$$\mathbb{P}[Z \geq \hat{M}_{\frac{1}{2}+p}] = \mathbb{P}\Big[n^{-\delta_0/2} \sum_{i=1}^{n^{\delta_0/2}} \geq \frac{1}{2}\Big]. \qquad (13)$$

Let us use Lemma 2.1 with $s = n^{\delta_0/2}$, $\alpha = 1/2$, $\theta = \frac{1}{2} - p$. It holds that

$$\begin{aligned} H(\alpha,\theta) &= \frac{1}{2}\ln\frac{1}{1-2p} + \frac{1}{2}\ln\frac{1}{1+2p} \\ &= \frac{1}{2}\ln\frac{1}{1-4p^2} \\ &\geq p^2 \end{aligned}$$

for all p small enough. So, by (13) and Lemma 2.1 we obtain that

$$\mathbb{P}[Z \geq \hat{M}_{\frac{1}{2}+p}] \leq \exp(-p^2 n^{\delta_0/2}).$$

By symmetry, the same estimate holds for $\mathbb{P}[Z \leq \hat{M}_{\frac{1}{2}-p}]$, so we obtain

$$\mathbb{P}[\hat{M}_{\frac{1}{2}-p} \leq Z \leq \hat{M}_{\frac{1}{2}+p}] \geq 1 - 2\exp(-p^2 n^{\delta_0/2}). \qquad (14)$$

To proceed, we notice that it is straightforward to obtain from (3) and (4) that $\hat{M}_{1/2} = z_0\beta^2$. Since by (4), $f_\beta(y)$ is of order β^{-2} when y is of order β^2, there exist positive constants C_4, C_5 such that

$$\begin{aligned} \hat{M}_{\frac{1}{2}+p} &\leq z_0\beta^2 + C_4 p\beta^2, \\ \hat{M}_{\frac{1}{2}-p} &\geq z_0\beta^2 - C_5 p\beta^2, \end{aligned}$$

for all p small enough. Now, it remains only to take $p = n^{-\delta_0/6}$ and use (14) to obtain (10) and (11), thus finishing the proof of Lemma 2.2. □

Consider the events

$$R_{n,m} = \{\chi(m+s) \leq n^{-2+2\varepsilon_0+2\delta_0+\delta_1} \text{ for any } 1 \leq s \leq n^{\delta_0/2}\}$$

and

$$D_{n,m} = \{\mathfrak{A}_n^m(1 - z_0^{-1}n^{-\delta_0/6}) \leq \beta \leq \mathfrak{A}_n^m(1 + z_0^{-1}n^{-\delta_0/6})\}, \qquad (15)$$

where, as before, $\beta := c - b = X_{k+1} - X_k$. We are going to estimate the conditional probability $\mathbb{P}[D_{n,m} \mid R_{n,m}]$ from below. To this end, define also the events

$$D'_{n,m} = \{\mathfrak{A}'_n(1 - z_0^{-1} n^{-\delta_0/6}) \le \beta \le \mathfrak{A}'_n(1 + z_0^{-1} n^{-\delta_0/6})\},$$

and

$$E_{n,m} = \big\{Y_{m+s} \in \{b, c\} \text{ for all } 0 \le s \le n^{\delta_0/2}\big\}.$$

Write

$$\begin{aligned}
\mathbb{P}[D_{n,m} \mid R_{n,m}] &\ge \mathbb{P}[D_{n,m} E_{n,m} \mid R_{n,m}] \\
&= \mathbb{P}[D'_{n,m} E_{n,m} \mid R_{n,m}] \\
&\ge 1 - \mathbb{P}[(D'_{n,m})^c \mid R_{n,m}] - \mathbb{P}[E^c_{n,m} \mid R_{n,m}] \\
&\ge 1 - \frac{\mathbb{P}[(D'_{n,m})^c]}{\mathbb{P}[R_{n,m}]} - \mathbb{P}[E^c_{n,m} \mid R_{n,m}]. \qquad (16)
\end{aligned}$$

Recall that $\{b, c\}$ is a level-n couple, so that $\min\{\xi(k), \xi(k+2)\} \ge n^{-1+\varepsilon_0+\delta_0+\delta_1}$. It is elementary to obtain from (4) that for some $C_6 > 0$ and all n it holds that

$$\mathbb{P}[T_{n^{-1+\varepsilon_0+\delta_0+\delta_1}} \le n^{-2+2\varepsilon_0+2\delta_0+\delta_1}] \le \exp(-C_6 n^{\delta_1}),$$

so

$$\mathbb{P}[E^c_{n,m} \mid R_{n,m}] \le n^{\delta_0/2} \exp(-C_6 n^{\delta_1}). \qquad (17)$$

By (8), $\mathbb{P}[R_{n,m}] \ge 1/2$ for all n large enough, and we can bound $\mathbb{P}[(D'_{n,m})^c]$ from above by using Lemma 2.2. So, using (16) and (17), we obtain

$$\mathbb{P}[D_{n,m} \mid R_{n,m}] \ge 1 - 2\exp(-\gamma_1 n^{-\delta_0/6}) - n^{\delta_0/2} \exp(-C_6 n^{\delta_1}). \qquad (18)$$

In words, the above equation shows that if a level-n couple causes a level-n trill, then, with a very high probability, that trill will be good and that one will be able to obtain the distance between the points in the couple with a good precision. Also, by (6) and (reftr1), we obtain that

$$\mathbb{P}[R_{n,m} \mid H^*_m] \ge 1 - C_7 n^{-\delta_0/2}, \qquad (19)$$

where the event H^*_m is defined by

$$H^*_m = \{W_{Y_m} \text{ is a point of some level-}n \text{ couple}\}.$$

Now, we have also to figure out how likely it is to produce good level-n trills elsewhere, not in level-n couples. First, we observe that, since the interval between any two consecutive rings in a level-n trill are at most $n^{-2+2\varepsilon_0+2\delta_0+\delta_1}$, the bells where the rings were produced should not be at distance more than $n^{-1+\varepsilon_0+\delta_0+\delta_1}$ from each other (otherwise the probability of producing such closely placed rings would be stretched-exponentially small). On the other hand, if we have three or more close bells (with distance of order $n^{-1+\varepsilon_0}$ from each other), then such a group of bells is, in principle, capable to produce a good level-n trill as well.

Suppose, however, that we know that we are in some region where there are no triples of close points (bells). More precisely, suppose that there are bells in points $a, b, c, d \in \mathbb{R}$, and $|b-c| < n^{-1+\varepsilon_0+\delta_0+\delta_1}$, while $\min\{|a-b|, |c-d|\} > n^{-1+\varepsilon_0+\delta_0+\delta_1}$; however, b is not close enough to c to form a level-n couple. Then, analogously to the proof of Lemma 2.2, it is straightforward to prove that

$$\mathbb{P}[\text{there is a good level-}n\text{ trill at } m \mid H_m^*(b)] \leq \exp(-C_8 n^{-\delta_0/6}), \qquad (20)$$

where $H_m^*(b) = \{W_{Y_m} = b\}$.

Now, for the sake of convenience we introduce some definitions concerning trills and couples:

Definition 2.4 *A level-n trill is compatible with a level-n couple with the distance β between the points, if (supposing for definiteness that the trill begins at the mth position of the sequence Y) the event $D_{n,m}$, defined in (15), occurs.*

Definition 2.5 *We say that a level-n trill was produced by a level-n couple, if all the rings of the trill occurred in the bells of the couple.*

Definition 2.6 *(i) Two level-n couples with the distances between their points being respectively β_1, β_2 are called n-similar if*

$$\min\{|\beta_1\beta_2^{-1} - 1|, |\beta_1^{-1}\beta_2 - 1|\} \leq 5z_0^{-1}n^{\delta_0/6}.$$

(ii) Two level-n trills (in positions m_1, m_2) are called n-similar if

$$\min\{|\mathfrak{A}_n^{m_1}(\mathfrak{A}_n^{m_2})^{-1} - 1|, |(\mathfrak{A}_n^{m_1})^{-1}\mathfrak{A}_n^{m_2} - 1|\} \leq 5z_0^{-1}n^{\delta_0/6}.$$

Two level-n couples (trills) are called n-different, if they are not n-similar.

Using the above definition, we summarize the results of this section in the following

Lemma 2.3 *There is a positive constant γ_2 such that:*

(i) With probability at least $1-\exp(-\gamma_2 n^{-\delta_0/6})$, given that a level-$n$ couple produces a level-n trill, the former will be compatible with the latter.

(ii) With at least the same probability n-different couples produce n-different trills.

(iii) Suppose that $W_{Y_m} = b$, where b is not from a level-n couple, and in the interval $[b - 2n^{-1+\varepsilon_0+\delta_0+\delta_1}, b + 2n^{-1+\varepsilon_0+\delta_0+\delta_1}]$ there are at most two bells (including the one in b). Then, with probability at least $1 - \exp(-\gamma_2 n^{-\delta_0/6})$, there is no level-$n$ trill at the mth position of the sequence Y.

Proof. Items (i) and (iii) follow from (18) and (20) respectively, and (ii) is an immediate consequence of (i). □

2.2 Localization test

The purpose of this section is to construct a test which, with high probability, is able to tell us if the Brownian motion is back to the same place.

Suppose that the local perturbation of the scenery X' was made in the interval $[-\ell, \ell]$, in other words, the "real" scenery X is known precisely in $\mathbb{R} \setminus [-\ell, \ell]$. We construct now a localization test depending on parameters n and ℓ. Define the events

$$\begin{aligned} G_{i,1}^{(n)} &= \Big\{\text{in the interval } [in^{1-\frac{\varepsilon_0}{2}}, (i+1)n^{1-\frac{\varepsilon_0}{2}}) \text{ there are at most } n^{\frac{3\varepsilon_0}{4}} \\ &\qquad \text{pairs } X_k, X_{k+1} \text{ such that } X_{k+1} - X_k \le \frac{n^{-1+\varepsilon_0}}{1 - z_0^{-1} n^{-\delta_0/6}}\Big\}, \\ G_{i,2}^{(n)} &= \Big\{\text{in the interval } [in^{1-\frac{\varepsilon_0}{2}}, (i+1)n^{1-\frac{\varepsilon_0}{2}}) \text{ there are at least } n^{\frac{\varepsilon_0}{4}} \\ &\qquad \text{level-}n \text{ couples which are } n\text{-different} \\ &\qquad \text{from all the level-}n \text{ couples in } [\ell, 5n]\Big\}, \end{aligned}$$

and let $G_i^{(n)} = G_{i,1}^{(n)} \cap G_{i,2}^{(n)}$.

Now, we define the values of n for which the localization test will be constructed.

Definition 2.7 *We say that $n > 2\ell$ is good, if:*

(i) On the interval $[n/2, \pi n]$ there are at least $n^{\varepsilon_0}/3$ level-n couples, and the same holds for the interval $[\pi n, 5n]$.

(ii) All the level-n couples on the interval $[\ell, 5n]$ are n-different.

(iii) Any subinterval of $[\ell, 5n]$ of length $4n^{-1+\varepsilon_0+\delta_0+\delta_1}$ contains at most two bells. Note that this implies that any pair of consecutive bells X_k, X_{k+1} such that $X_{k+1} - X_k \le n^{-1+\varepsilon_0}(1 - z_0^{-1} n^{-\delta_0/6})^{-1}$ and $\{X_k, X_{k+1}\} \subset [\ell, 5n]$ is a level-n couple.

(iv) for any $i \in \mathbb{Z}$ such that $[in^{1-\frac{\varepsilon_0}{2}}, (i+1)n^{1-\frac{\varepsilon_0}{2}}) \cap [\ell, \pi n] = \emptyset$ and that $|i| < \exp(n^{\frac{\varepsilon_0}{8}})$ the event $G_i^{(n)}$ holds.

(v) On any interval of length $n^{1-\frac{\varepsilon_0}{2}}$, which is within $[\ell, 5n]$, there is at least $n^{\frac{\varepsilon_0}{4}}$ level-n couples.

(vi) In the interval $[-n^2, n^2]$, the minimal distance between two neighboring bells is at least n^{-3}.

The following lemma ensures that there is an infinite sequence of good ns:

Lemma 2.4 *There exists $C > 0$ such that $\mathbb{P}[n \text{ is good}] \ge 1 - n^{-C}$.*

Proof. The proof of this lemma is completely elementary, so we shall give only a sketch. First, one can easily see that the probability that there exists a level-n couple on an interval of length 1 is (up to smaller terms) $n^{-1+\varepsilon_0}$. This implies that the probability of the events in (i) and (v) are high enough (even stretched-exponentially high). Similarly, it is elementary to obtain that the probability of the event in (iii) is of order $1 - n^{-1+2\varepsilon_0+\delta_0+\delta_1}$ (and $2\varepsilon_0 + \delta_0 + \delta_1 < 1$, recall (1)), and for the event in (vi), that probability is of order $1 - \frac{const}{n}$. Also, note that the amount of "classes" of n-different level-n couples is of order $n^{\delta_0/6}$; since the total number of level-n couples there is of order n^{ε_0} (recall (2)), this takes care of (ii). To deal with (iv), we note that a "random" level-n couple has a good (bounded away from 0) chance to be different from all those in the interval $[\ell, \pi n]$. With this observation, one can obtain by a straightforward computation that (iv) holds with stretched-exponentially high probability. □

Now, we define the localization test. Suppose that n is good and consider all the level-n couples in the interval $[n/2, \pi n]$. Let $(\zeta'_n, \zeta'_n + \Delta'_n)$ be the smallest level-n couple on that interval, $(\zeta_n, \zeta_n + \Delta_n)$ the largest one, and let $\psi(n)$ be the number of other level-n couples there (note that, by (i) of Definition 2.7, $\psi(n) \geq n^{\varepsilon_0}/3 - 2$).

Define $\tau_0^{(n)} = 0$ and, for $i \geq 1$,

$$\tau_i^{(n)} = \inf\{t \geq \tau_{i-1}^{(n)} + 3n^2 : t \text{ satisfies (A), (B), (C), (D) below}\}, \tag{21}$$

where

(A) there exists $s \in [t - n^2, t)$ and $m_1 \in \mathbb{Z}_+$ such that $Y_{m_1} = s$ and there is a level-n trill in m_1 compatible with the couple $(\zeta'_n, \zeta'_n + \Delta'_n)$;

(B) the number of n-different level-n trills on the time interval $[t - n^2, t)$ is at least $\frac{\psi(n)}{2}$;

(C) for any level-n trill from that interval there exists a level-n couple on $[n/2, \pi n]$ which is compatible to that trill;

(D) (suppose without restricting of generality that $\lfloor n^{\delta_0/2} \rfloor$ is even) for some $m_2 \in \mathbb{Z}_+$ there is a level-n trill in m_2 such that it is compatible with the couple $(\zeta_n, \zeta_n + \Delta_n)$ and $Y_{m_2 + \lfloor n^{\delta_0/2} \rfloor} = t$.

In words, the above (A)–(D) are what we typically observe when the Brownian motion crosses the interval $[n/2, \pi n]$.

The main result of this section is the following

Lemma 2.5 *There exist $\delta_2, \delta_3 > 0$ such that*

$$\mathbb{P}[W_{\tau_i^{(n)}} = \zeta_n \textit{ for all } i = 1, \ldots, \exp(n^{\delta_2})] \geq 1 - \exp(-n^{\delta_3}). \tag{22}$$

Proof. Choose a number $\delta_2 > 0$ such that

$$\delta_2 < \min\left\{\frac{\delta_0}{6}, \delta_1, \frac{\varepsilon_0}{8}\right\} \tag{23}$$

(in fact, due to (2), in the above display $\frac{\delta_0}{6}$ is redundant).

Let us say that a time interval $[t_1, t_2]$ is a crossing of the interval $[a, b]$ by the Brownian motion, if $W_{t_1} = a$, $W_{t_2} = b$, and $W_s \notin \{a, b\}$ for $s \in (t_1, t_2)$. We say that a crossing $[t_1, t_2]$ of the interval $[n/2, \pi n]$ by the Brownian motion is *good*, if $t_2 - t_1 \le n^2$, and there is j_0 such that $\tau_{j_0}^{(n)} \in [t_1, t_2]$ (see (A)–(D) above). Define the events

$$\begin{aligned}
U_1^{(n)} &= \Big\{\text{up to time } \exp(3n^{\delta_2}), \text{ there are at least } \exp(n^{\delta_2}) \\
&\qquad \text{good crossings of the interval } [n/2, \pi n]\Big\}, \\
U_2^{(n)} &= \Big\{\text{up to time } \exp(3n^{\delta_2}), \text{ all the level-}n \text{ trills produced when} \\
&\qquad \text{the Brownian motion was in the interval } [\ell, 5n] \text{ correspond} \\
&\qquad \text{to level-}n \text{ couples compatible with those trills}\Big\}, \\
U_3^{(n)} &= \Big\{\text{up to time } \exp(3n^{\delta_2}), \text{ on any time interval } I \text{ of length at least} \\
&\qquad n^{2-\frac{\varepsilon_0}{2}} \text{ and such that } \{W_s, s \in I\} \cap [n/2, \pi n] = \emptyset, \text{ one finds} \\
&\qquad \text{at least } n^{\frac{\varepsilon_0}{4}} \text{ level-}n \text{ trills and at least } \tfrac{1}{2} n^{\frac{\varepsilon_0}{4}} \text{ of those trills} \\
&\qquad \text{are not compatible with any of the couples from } [n/2, \pi n]\Big\}, \\
U_4^{(n)} &= \Big\{\text{up to time } \exp(3n^{\delta_2}), \text{ on any time interval } I \text{ of length at least} \\
&\qquad n^{2-\frac{\varepsilon_0}{2}} \text{ the range of the Brownian motion is at most } n^{1-\frac{\varepsilon_0}{8}}\Big\},
\end{aligned}$$

where the range of the Brownian motion on a time interval is the difference between the maximum and the minimum of the Brownian motion on that interval.

First, let us show that on $U_1^{(n)} \cap U_2^{(n)} \cap U_3^{(n)} \cap U_4^{(n)}$ the event $\{W_{\tau_i^{(n)}} = \zeta_n$ for all $i = 1, \ldots, \exp(n^{\delta_2})\}$ holds. It is straightforward to see that on $U_1^{(n)}$ we have that $\tau^{(n)}_{\exp(n^{\delta_2})} \le \exp(3n^{\delta_2})$. Now, let us suppose that there exists $i_0 \le \exp(n^{\delta_2})$ such that $a_0 := W_{\tau_{i_0}^{(n)}} \ne \zeta_n$. Consider the two possible cases: $a_0 \in [\ell, 5n]$, or $a_0 \notin [\ell, 5n]$. We know that $\tau_{i_0}^{(n)}$ is at the end of a level-n trill compatible with the level-n couple $(\zeta_n, \zeta_n + \Delta_n)$, so on the event $U_2^{(n)}$ it is impossible to have $a_0 \in [\ell, 5n]$. On the other hand, if $a_0 \notin [\ell, 5n]$, then (since $(5-\pi)n > n^{1-\frac{\varepsilon_0}{2}}$) on the event $U_4^{(n)}$ we have that $W_s \notin [n/2, \pi n]$ for all $s \in [\tau_{i_0}^{(n)} - n^{2-\frac{\varepsilon_0}{2}}, \tau_{i_0}^{(n)}]$. So, on $U_3^{(n)}$ we have that on the time interval $[\tau_{i_0}^{(n)} - n^{2-\frac{\varepsilon_0}{2}}, \tau_{i_0}^{(n)}]$ there will be level-n trills which are not compatible with any of the level-n couples from $[n/2, \pi n]$; clearly, this contradicts (21).

Now let us estimate the probabilities of the events $U_i^{(n)}$, $i = 1, 2, 3, 4$.

First, we deal with $U_2^{(n)}$. Recall that, by Definition 2.7 (vi), the minimal distance between the bells in $[\ell, 5n]$ is at least n^{-3}. So, given that the particle is in some bell there, the time until the next ring will be greater than n^{-7} with probability at least $1-\exp(-C_1 n^{1/2})$ for some $C_1 > 0$. Thus, up to time $\exp(3n^{\delta_2})$ we will have at most $n^7\exp(3n^{\delta_2})$ rings produced by the bells in $[\ell, 5n]$, with probability at least

$$1 - n^7 \exp(-C_1 n^{1/2} + 3n^{\delta_2})$$

(recall that $\delta_2 < 1/2$ by e.g. (1)). Using Lemma 2.3, one obtains

$$\mathbb{P}[U_2^{(n)}] \geq 1 - n^7 \exp(-C_1 n^{1/2} + 3n^{\delta_2}) - n^7 \exp(-\gamma_2 n^{\frac{\delta_0}{6}} + 3n^{\delta_2}). \tag{24}$$

To estimate the probability of $U_1^{(n)}$, we note that by (19) and Lemma 2.3, the probability that a crossing of the interval $[n/2, \pi n]$ is good, is bounded away from 0 by some constant C_2. Also, with probability at least $1 - C_3\exp(-n^{\delta_2})$ up to time $\exp(3n^{\delta_2})$ there will be at least $2C_2^{-1}\exp(n^{\delta_2})$ crossings of that interval. So,

$$\mathbb{P}[U_1^{(n)}] \geq 1 - C_4 \exp(-n^{\delta_2}) \tag{25}$$

for some $C_4 > 0$.

Now, note that the event $U_4^{(n)}$ occurs if on each of the intervals (of length $\frac{1}{2}n^{2-\frac{\varepsilon_0}{2}}$) $[(i-1)n^{2-\frac{\varepsilon_0}{2}}, in^{2-\frac{\varepsilon_0}{2}}]$, $i = 1, \dots, 2n^{-2+\frac{\varepsilon_0}{2}}\exp(3n^{\delta_2})$, the range of the Brownian motion is at most $n^{1-\frac{\varepsilon_0}{8}}$. So, since for each i that happens with probability at least $1-\exp(-n^{\frac{\varepsilon_0}{8}})$, we obtain

$$\mathbb{P}[U_4^{(n)}] \geq 1 - 2n^{-2+\frac{\varepsilon_0}{2}} \exp(-n^{\frac{\varepsilon_0}{8}} + 3n^{\delta_2}). \tag{26}$$

The probability of the event $U_3^{(n)}$ can be bounded from below in the following way. Note that for each time interval of length $n^{2-\frac{\varepsilon_0}{2}}$ the range of the Brownian motion on that interval is greater than $2n^{1-\frac{\varepsilon_0}{2}}$ with probability at least $1 - \exp(-C_5 n^{\frac{\varepsilon_0}{4}})$. Note also that

$$\mathbb{P}\Big[\max_{s\leq \exp(3n^{\delta_2})} |W_s| \leq \exp(n^{\frac{\varepsilon_0}{8}})\Big] \geq 1 - \exp\Big(-n^{\frac{\varepsilon_0}{8}} - \frac{3}{2}n^{\delta_2}\Big).$$

Then we use Definition 2.7 (iv) and (v) and Lemma 2.3 to obtain that

$$\mathbb{P}[U_3^{(n)}] \geq 1 - \exp(-C_6 n^{\frac{\varepsilon_0}{8}}) \tag{27}$$

for some $C_6 > 0$.

Using (24)–(27) it is straightforward to obtain (22), thus finishing the proof of Lemma 2.5. □

2.3 Reconstruction algorithm for the case when the interval of perturbation is known

In this section we describe the algorithm that reconstructs the local perturbation using the localization test of Section 2.2. As in the previous section, we assume here that it is known that the perturbation took place on the interval $[-\ell, \ell]$.

Let $k_1 = \min\{k : X_k \in [-\ell, \ell]\}$, $k_2 = \max\{k : X_k \in [-\ell, \ell]\}$. Denote by $m = k_2 - k_1 + 1$ the number of points of the (true) scenery in the interval $[-\ell, \ell]$, and abbreviate by $a_i = X_{k_1+i-1} + \ell$ the distance from the left end of the interval to the ith point of the scenery there, $i = 1, \dots, m$. Moreover, for $i = 1, 2, 3, \dots$ denote $A_i = a_1^i + \dots + a_m^i$. Now, the idea is to reconstruct first the quantity m; then, given m, reconstruct A_1; then, given m and A_1, reconstruct A_2, and so on.

We need the following technical fact:

Lemma 2.6 *Suppose that $\theta = o(n^{-3})$ and $x = O(n)$. Then*

$$\begin{aligned}&\mathbb{P}\big[W_t = x \text{ for some } t \in [n^2, n^2+\theta]\big]\\ &= \frac{1}{\sqrt{2\pi}n}\exp\Big(-\frac{x^2}{2n^2}\Big)\Big(\frac{2\sqrt{2}}{\sqrt{\pi}}\theta^{1/2} + n^{-2}O(\theta^{3/2})\Big).\end{aligned} \tag{28}$$

Proof. By (4) and conditioning on W_{n^2}, the left-hand side of (28) can be written as follows:

$$\begin{aligned}&\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}n}\exp\Big(-\frac{(y-x)^2}{2n^2}\Big)\int_0^\theta \frac{|y|}{\sqrt{2\pi}s^{3/2}}\exp\Big(-\frac{y^2}{2s}\Big)ds\,dy\\ &= \frac{1}{\sqrt{2\pi}n}\exp\Big(-\frac{x^2}{2n^2}\Big)\int_0^\theta \frac{1}{\sqrt{2\pi}s^{3/2}}\int_{-\infty}^{+\infty}|y|\exp\Big(-\frac{y^2}{2n^2}+\frac{xy}{n^2}-\frac{y^2}{2s}\Big)dy\,ds.\end{aligned}$$

Then, in the last integral we change the variables $u := \frac{y^2}{2s}$, and after some elementary calculus we obtain (28). □

Define $\theta_n = \exp(-n^{\delta_2/2})$. Let

$$Z_i^{(n)} = \mathbf{1}\{\text{there is a ring in the interval } [\tau_i^{(n)} + n^2, \tau_i^{(n)} + n^2 + \theta_n]\},$$

and let

$$Z^{(n)} = \exp(-n^{\delta_2}) \sum_{i=1}^{\exp(n^{\delta_2})} Z_i^{(n)}.$$

Let

$$h^{(n)} = \mathbb{P}^{\zeta_n}[\text{there is a ring in the interval } [n^2, n^2 + \theta_n]].$$

By Lemma 2.5 and usual large deviation technique (use e.g. Lemma 2.1), we obtain that

$$\mathbb{P}\Big[|Z^{(n)} - h^{(n)}| > \exp\Big(-\frac{n^{\delta_2}}{2}\Big)\Big] \leq \exp\Big(-\frac{n^{\delta_2}}{4}\Big). \tag{29}$$

Let X'' be the local perturbation of the scenery X' (and X) obtained by removing all the bells from the interval $[-\ell, \ell]$ (so, since we supposed that X' was perturbed on $[-\ell, \ell]$, X'' is completely known to us). Define

$$\mu^{(n)} = \mathbb{P}^{\zeta_n}[\text{there is a ring in the interval } [n^2, n^2 + \theta_n] \mid \text{the scenery is } X''].$$

Let $\hat{Z}^{(n)} = Z^{(n)} - \mu^{(n)}$ and abbreviate also $b_n = (\zeta_n + \ell)/n$. Let

$$B^{(n)} = \frac{2\theta_n^{1/2}}{\pi n} \exp\Big(-\frac{b^2}{2}\Big) \sum_{i=1}^{m} \exp\Big(-\frac{a_i^2}{2n^2} + \frac{ba_i}{n}\Big).$$

Using Definition 2.7 (vi), it is straightforward to obtain that

$$\mathbb{P}^{\zeta_n}[\text{there are at least two rings in the interval } [n^2, n^2 + \theta_n]] \leq e^{-n} \tag{30}$$

for all n large enough. So, by Lemma 2.6 and (29)–(30), we can write

$$\mathbb{P}\Big[\Big|\hat{Z}^{(n)} - B^{(n)}\Big| > 2\exp(-n^{\delta_2/2})\Big] \leq \exp\Big(-\frac{n^{\delta_2}}{4}\Big). \tag{31}$$

Consider now the function $\varphi_b(x) = \exp\big(-\frac{x^2}{2} + bx\big)$ and its Taylor decomposition in $x = 0$:

$$\varphi_b(x) = \exp\Big(-\frac{x^2}{2} + bx\Big) = 1 + \sum_{k=1}^{\infty} M_k(b) x^k.$$

It is easy to see that $M_k(b)$ is a polynomial of kth degree of b, so if b is a transcendental number, $M_k(b) \neq 0$ for all k. By Definition 2.7 (v), we have that $b_n/n \to \pi$, so if n is large enough, then we have $M_i(b_n) \neq 0$ for all $i \leq m$.

Now, we can write

$$B^{(n)} = \frac{2\theta_n^{1/2}}{\pi n} \exp\Big(-\frac{b_n^2}{2}\Big)\Big(m + \frac{M_1(b_n)A_1}{n} + \frac{M_2(b_n)A_2}{n^2} + \cdots\Big). \tag{32}$$

Let us define the estimator for m (the number of points of the scenery in $[-\ell, \ell]$):

$$\hat{m}(n) = \Big[\hat{Z}^{(n)} \exp\Big(\frac{b_n^2}{2}\Big)\frac{\pi n}{2\theta_n^{1/2}}\Big]; \tag{33}$$

here $[y]$ stands for the integer part of $y + \frac{1}{2}$, i.e., $[y]$ is the integer closest to y.

Given m, define the estimator for A_1 (cf. (32)):

$$\hat{A}_1(n; m) = \Big(\hat{Z}^{(n)} \exp\Big(\frac{b^2}{2}\Big)\frac{\pi n}{2\theta_n^{1/2}} - m\Big)\frac{n}{M_1(b_n)},$$

and, for all $i \geq 2$, given m and $A_1, \ldots, A_{i-1}$, define the estimator for A_i:

$$\hat{A}_i(n; m, A_1, \ldots, A_{i-1}) = \Big(\hat{Z}^{(n)} \exp\Big(\frac{b^2}{2}\Big)\frac{\pi n}{2\theta_n^{1/2}} - m - \sum_{j=1}^{i-1}\frac{M_j(b_n)A_j}{n^j}\Big)\frac{n^i}{M_i(b_n)}.$$

Using (31), one can observe that

$$\mathbb{P}[\hat{m}(n) \neq m] \leq \exp\Big(-\frac{n^{\delta_2}}{4}\Big) \tag{34}$$

and

$$\mathbb{P}\big[|\hat{A}_i(n; m, A_1, \ldots, A_{i-1}) - A_i| \geq Cn^i \exp(-n^{\delta_2/2})\big] \leq \exp\Big(-\frac{n^{\delta_2}}{4}\Big). \tag{35}$$

Now, informally, the idea is the following: take a sequence of ns going fast to infinity, reconstruct m (using also Borel-Cantelli), then reconstruct A_1, and so on. Formally, consider the sequence $n_k = 2^k$, $k = 1, 2, 3, \ldots$. Then, by Lemma 2.4, n_k will be good for all but finitely many k. Using (34) and Borel-Cantelli lemma, we obtain that there is k_0 such that

$$\hat{m}(n_k) = m \qquad \text{for all } k \geq k_0. \tag{36}$$

Then, given m, we are able to determine A_1 in the following way: by (35),

$$\lim_{k\to\infty} \hat{A}_1(n_k, m) = A_1 \qquad \text{a.s.} \tag{37}$$

Inductively, given m and $A_1, \dots A_{i-1}$, we determine A_i by

$$\lim_{k\to\infty} \hat{A}_i(n_k, m, A_1, \dots, A_{i-1}) = A_i \qquad \text{a.s.,} \tag{38}$$

for all $i \le m$.

At this point we need the following elementary fact:

Lemma 2.7 *Suppose that $a_1, \dots, a_m$ are positive numbers satisfying the following system of algebraic equations*

$$\begin{cases} a_1 + \cdots + a_m &= d_1 \\ & \cdots \\ a_1^m + \cdots + a_m^m &= d_m \end{cases} \tag{39}$$

Suppose also that $(a_1', \dots, a_m')$ is another solution of the system (2.7). Then $\{a_1, \dots, a_m\} = \{a_1', \dots, a_m'\}$, i.e., $a_1', \dots, a_m'$ is simply a reordering of the collection $a_1, \dots, a_m$.

Proof. This is an easy consequence of Newton's and Vieta's formulas. □

To conclude this section, it remains to note that, by Lemma 2.7, one can uniquely determine $a_1, \dots, a_m$ from $A_1, \dots, A_m$.

2.4 Reconstruction algorithm for the general case

Now, suppose that we do not know about where the perturbation took place, and that we only know it is local in the sense of Definition 1.1. That means that there exists N_0 (which is not known to us) such that the interval of perturbation is fully inside $[-N, N]$ for all $N \ge N_0$. Denote by $\tilde{X}^{(N)}$ the result of application of the reconstruction algorithm of Section 2.3 with $\ell := N$. Note, however, that it is not clear if the algorithm of Section 2.3 produces any result (i.e., (36), (38) hold) when the perturbation is not actually limited to $[-N, N]$. When the algorithm does not produce the result, we formally define $\tilde{X}^{(N)} := \emptyset$.

Then, it is clear that the true scenery X can be obtained as

$$X = \lim_{N\to\infty} \tilde{X}^{(N)},$$

where the limit can be formally defined in any reasonable sense, since actually $\tilde{X}^{(N)} = X$ for all $N \ge N_0$. This concludes the proof of Theorem 1.1. □

References

[1] I. Benjamini, H. Kesten (1996) Distinguishing sceneries by observing the scenery along a random walk path. *J. Anal. Math.* **69**, 97–135.

[2] A.N. Borodin, P. Salminen (2002) *Handbook of Brownian motion — Facts and Formulae.* Birkhäuser Verlag, Basel-Boston-Berlin.

[3] K. Burdzy (1993) Some path properties of iterated Brownian motion. *Seminar on Stochastic Processes, 1992 (Seattle, WA, 1992), Progr. Probab.* **33**, 67–87, Birkhäuser Boston.

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

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

[6] R. Durrett (1996) *Probability: Theory and Examples.* Duxbury Press, Belmont, CA (2nd ed.).

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

[8] M. Harris, M. Keane (1997) Random coin tossing. *Probab. Theory Relat. Fields* **109**, 27–37.

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

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

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

[12] M. Keane, F. den Hollander (1986) Ergodic properties of color records. *Phys. A* **138**, 183–193.

[13] H. Kesten (1996) Detecting a single defect in a scenery by observing the scenery along a random walk path. *Itô's Stochastic Calculus and Probability Theory* 171–183. Springer, Tokyo.

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

[15] I. Karatzas, S.E. Shreve (1991) *Brownian Motion and Stochastic Calculus.* Springer-Verlag, New York.

[16] E. Lindenstrauss (1999) Indistinguishable sceneries. *Random Struct. Algorithms* **14**, 71–86.

[17] A. Lenstra, H. Matzinger (2001) Reconstructing a 4-color scenery by observing it along a recurrent random walk path with unbounded jumps. Eurandom. In preparation.

[18] M. Löwe, H. Matzinger (2002) Scenery reconstruction in two dimensions with many colors. *Ann. Appl. Probab.* **12**, 1322–1347.

[19] J. Lember, H. Matzinger (2003) Reconstructing a 2-color scenery by observing it along a recurrent random walk path with bounded jumps. Eurandom. In preparation.

[20] M. Löwe, H. Matzinger, F. Merkl (2004) Reconstructing a multicolor random scenery seen along a random walk path with bounded jumps. Electron. J. Probab. **9**, 436–507.

[21] M. Löwe, H. Matzinger (2003) Reconstruction of sceneries with correlated colors. *Stochastic Process. Appl.* **105**, 175–210.

[22] D. Levin, Y. Peres (2004) Identifying several biased coins encountered by a hidden random walk. *Random Struct. Algorithms* **25** (1), 91–114.

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

[24] H. Matzinger (1999) *Reconstructing a 2-color scenery by observing it along a simple random walk path with holding.* Ph.D. thesis, Cornell University.

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

[26] H. Matzinger (2005) Reconstructing a 2-color scenery by observing it along a simple random walk path. Ann. Appl. Probab. **15**, 778–819.

[27] H. Matzinger, S.W.W. Rolles (2003) Reconstructing a random scenery observed with random errors along a random walk path. *Probab. Theory Related Fields* **125**, 539–577.

[28] H. Matzinger, S.W.W. Rolles (2002) Reconstructing a piece of scenery with polynomially many observations. *Stochastic Process. Appl.* **107**, 289–300.

[29] H. Matzinger, S.W.W. Rolles (2005) Finding blocks and other patterns in a random coloring of Z. To appear in: *Random Struct. Algorithms.*

[30] H. Matzinger, S.W.W. Rolles (2006) Retrieving random media. To appear in: *Probab. Theory Related Fields.*

[31] A. Shiryaev (1989) *Probability* (2nd ed.). Springer, New York.