Reconstructing a piece of 2-color scenery

Jüri Lember and Heinrich Matzinger

EURANDOM
P.O. Box 513 - 5600 MB Eindhoven, The Netherlands

Contents

1 Introduction and Result

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.

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 $\phi \in \{0,1\}^I$, and a subset $I' \subseteq I$, the piece of scenery $(\phi(i))_{i\in I'}$ is denoted by $\phi|I'$.
Two pieces of scenery $\phi \in \{0,1\}^I$ and $\phi' \in \{0,1\}^{I'}$ are **equivalent**, $\phi \approx \phi'$, 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 $\phi(i) = \phi'(ai+b)$, $\forall i \in I$. If ϕ is obtained from ϕ' by translation, i.e. $\phi(i) = \phi'(b+i)$, then ϕ and ϕ' are called **strongly equivalent**, we denote this $\phi \equiv \phi'$. If ϕ is obtained from ϕ' by reflection i.e. $\phi(i) = \phi'(-i)$, $\forall i \in I$, we write $\phi = \phi'^{-}$. By definition, $\phi \sqsubseteq \phi'$ means that $\phi \approx \phi'|J$ for some $J \subseteq I'$. If, in addition, the equivalence is strong, we write $\phi \sqsubseteq \phi'$. In this case ϕ is equal to $\phi'|J$ up to the translation, only.
For a piece of scenery $\phi|[x,y]$, where $[x,y] = (x,\dots,y) \subset \mathbb{Z}$ is an integer interval, we often write ϕ_x^y. If $x = 0$, then it is skipped, i.e. $\phi|[x,y]$ 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, i.e. a scenery ψ is of random element 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** :

$$\chi := \xi(S(0)), \xi(S(1)), \xi(S(2)), \dots$$

and we interpret χ as a random piece of scenery $\{0,1\}^{\mathbb{N}}$, so that $\chi(k) := \xi(S(k))$ 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 $[x,y] \subset \mathbb{N}$ 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 $w \sqsubseteq w'$. We shall also use the reflected words w^-. Hence, for a word $w = (w_1,\dots,w_N)$, $w^- = (w_N,\dots,w_1)$.

Let $I=[x,y]$. The piece of scenery $\phi|I$ (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 $(1,\dots,N)$). Hence, we can state that "the piece ϕ_x^y is the word w", meaning that the image of $\phi|I$ is w or, equivalently, $\psi_x^y\equiv w$. Depending on ϕ, we shall call w as the observation- or scenery-word.

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 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 $\tau=(\tau(1),\dots,\tau(\exp(\alpha l_1)))$ as an additional input. The stopping times are assumed to satisfy the conditions:

$$\tau(k)-\tau(k-1)\geq 2\exp(2l_1),\quad k=2,3,\dots,\exp(\alpha l_1)+1,\qquad \text{where } \tau(\exp(\alpha l_1)+1):=\exp[12\alpha l_1]. \quad (1.1)$$

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

$$E^1_{\text{stop}}(\tau):=\{|S(\tau(k))|\leq\exp(l_1),\quad k=1,\dots,\exp(\alpha l_1)\}.$$

The condition (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 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

$$E^1_{\text{alg works}}(\tau,I^o):=\Big\{\xi|[-\exp(l_1),\exp(l_1)]\sqsubseteq\mathcal{A}^1\big(\chi^{\exp(12\alpha l_1)},\tau,\xi|I^o\big)\sqsubseteq\xi|[-3\exp(l_1),3\exp(l_1)]\Big\}. \quad (1.2)$$

Obviously the event (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

$$E^1_{\text{alg works}}(\tau) := \bigcap_{I^o \subset [-\exp(l_1), \exp(l_1)]} E^1_{\text{alg works}}(\tau, I^o).$$

The description and formal definition of $\mathcal{A}^1$ is given in Subsection 3.3. The main result of the paper, Theorem 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 1.1 *There exists a constant $k > 0$ not depending on l_1 such that, for l_1 big enough*

$$P\Big(E^1_{\text{stop}}(\tau) \cap \big(E^1_{\text{alg works}}(\tau)\big)^c\Big) \leq e^{-kl_1}. \tag{1.3}$$

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.

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 $y = x + c_1 l_1 L$, where c_1 is a fixed constant, specifies in Section **??**. 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

$$\Big(\psi(x), \psi(x+L), \ldots, \psi(x + (c_1 l_1 - 1)L)), \psi(y)\Big). \tag{1.4}$$

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 (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 (1.4). The formal definition of ladder words is given in Section 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 3.1). Clealry, 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.

1.4.1 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 (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 $y = x + c_1 l_1 L$. 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 (1.4) is rather straightforward. Indeed, for each $t \geq 0$ define the observation words

$$w^1(t) := \chi|[t - lm, t], \quad w^2(t) := \chi|[t, t + c_1 l_1], \quad w^3(t) := \chi|[t + c_1 l_1, t + c_1 l_1 + lm^2] \tag{1.5}$$

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

$$S(t) \leq x \quad \text{and} \quad S(t + c_1 l_1) \geq y. \tag{1.6}$$

On the other hand, during $c_1 l_1$ steps, the random walk S cannot move more than $c_1 l_1 L$. But this is exactly the distance between x and y. Hence, the only possibility for (1.6) to hold is that both inequalities are equalities. In this case, $w^2(t)$ equals the ladder word (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

$$P\Big(\hat{g}(\chi|[t,t+lm^2]) = g_y(\psi), \quad S(t) \in [-\exp(3l_1), y]\Big) = 0 \tag{1.7}$$
$$P\Big(\hat{g}^*(\chi|[t-lm^2,t]) = g^*_x(\psi), \quad S(t) \in [x, \exp(3l_1)]\Big) = 0. \tag{1.8}$$

Since the above-described selection rule now works only on I_1, we have to modify the construction of (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

$$T^1(j) := \tau(j) + \exp(2l_1) + lm^2, \quad T^3(j) := T^1(k) + c_1l_1, \quad j = 1,\dots,\exp(\alpha l_1). \tag{1.9}$$

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

$$w^1(j) := \chi|[T^1(j) - lm^2, T^1(j)]$$
$$w^2(j) := \chi|[T^1(j), T^3(j)]$$
$$w^3(j) := \chi|[T^3(j), T^3(j) + lm^2]$$

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 (1.7) and (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 (1.4) and, therefore, the same. Our modified rule is the following:

Simplified selection: The word w is taken as (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

$$\hat{g}^*(w^1(j)) = g^*_1, \quad \hat{g}(w^3(j)) = g_3; \tag{1.10}$$

b) for every j satisfying (1.10), it holds $w^2(j) = w$.

1.4.2 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 $|z - z'| < Lc_1 l_1$. Consider the possible observation-words that S generates by going from z to z' in $c_1 l_1$ steps. If c_1 is big enough, then the probability that all these words are the same, is small (Proposition 3.1). In Section 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 $c_1 l_1$ steps. Any path of S that consists of $c_1 l_1$ steps has the probability at least $(p_{min})^{c_1 l_1}$.
Suppose w passes the selection rule. Hence, there exists a set $J \subseteq \{1, \ldots, \exp(\alpha l_1)\}$ such that at least $|J| \geq \exp(\gamma l_1)$ and for each $j \in J$ the following holds : $|S(T^3(j)) - S(T^1(j))| < Ll_1 c_1$ and $w^2(j) = w$. Let $Y_k := 1 - I_{w^2(j_1)}(w^2(j_k))$, where $j_1, j_2, \ldots$ 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 $(p_{min})^{c_1 l_1}$. 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 (3.32))

$$P\Big(\sum_{k=2}^{\exp(\gamma l_1)} Y_k = 0 \Big| E^1_{\text{stop}} \cap B^1_{\text{recon straight}}\Big) \leq \exp[-2\exp((\gamma + 2c_1 \ln p_{min}) l_1)].$$

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 3.42). 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 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 1.4.6.

1.4.3 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 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 $\psi_n \in B_{\text{cell\_OK}}(n)$:*

$$P\left(\hat{g}(\chi_0^{m^2}) \sqsubseteq g(\psi_0^m) \Big| S(m^2) = m, \xi = \psi_n\right) > 3/4.$$

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 3.6. Theorem 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,

$$P\Big(\hat{g}(\chi_t^{t+m^2}) \sqsubseteq g(\xi_y^{y+m})\Big) = \Big(\hat{g}(\chi_t^{t+m^2}) \sqsubseteq g(\xi_0^m)\Big) \leq \Big(\frac{1}{2}\Big)^{n^2-1}.$$

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 (1.7) holds

$$P\Big(\hat{g}(\chi|[t, t+m^2]) \sqsubseteq g(\psi|[y, y+m]), \quad S(t) \in [-\exp(3l_1), y - Lm^2)\Big) = \Big(\frac{1}{2}\Big)^{n^2-1}. \tag{1.11}$$

So, taking n big enough, we can make the right side of (1.11) as small as we want.
Unfortunately, for several reasons, the functions from Theorem 1.2 is not good enough. Recall that we want the mistake (1.3) to be exponentially small in l_1. The right side of (1.11) does not depend on l_1. To handle this, we apply Theorem 1.2 repeatedly. This procedure is called *iteration* and it is the subject of Section 2. Let us briefly introduce the main ideas behind the iteration.

From now on, we define

$$l := l_1 \cdot l_2, \quad \text{where } l_2 \text{ is fixed positive integer, specified in Section 3.6.}$$

We shall apply the functions g and $\hat{g}$ from Theorem 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 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 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 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 $\mathcal{I}(\psi|[y, y+lm]) =: \mathcal{I}_y(\psi) \subset \{1, \dots, l\}$ 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

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 2.1 and the estimation (3.21)).
For not OK cells, the statement **2)** of Theorem 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 $g(\chi|[t,t+m^2l])$ reproduces $g_y(\psi)$, if

$$g(\chi|[t,t+m^2l]) \sqsubseteq_{\mathcal{I}(\psi)} g_y(\psi).$$

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.

1.4.4 Getting selected

Let us now give some insight, how do we show that the probability for a ladder word (1.4) to pass the selection is sufficiently high. What follows, is a simplified version of Proposition 3.2. Let

$$E_j(x,y) := \left\{ \begin{array}{c} S(T^1(j)-lm^2) = x-lm \\ S(T^1(j)) = x,\ S(T^3(j)) = y, \\ \hat{g}^*(w^1(j)) \sqsubseteq_{\mathcal{I}^*_x(\xi)} g^*_x(\xi), \\ \hat{g}(w^3(j)) \sqsubseteq_{\mathcal{I}_y(\xi)} g_y(\xi) \end{array} \right\}, \quad Y_j := I_{E_j}, \quad j=1,\dots,e^{\alpha l_1}.$$

Clearly (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 $E_{\text{stop}}(\tau)$, $|S(\tau(j))| \leq \exp(l_1)$. Now, the local central limit theorem (LCLT) can be used to see that for l_1 big enough

$$P_\psi\big(S(\tau(j)+e^{2l_1}) = x-lm\big|E_{\text{stop}}(\tau)\big) \geq \exp(-1.5l_1).$$

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 symmetricity 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,

$$\begin{aligned}
&P_\psi\Big(S(T^1(j)) = x,\ \hat{g}^*(w^1(j)) \sqsubseteq_{\mathcal{I}^*_x(\psi)} g^*_x(\psi)\Big|S(T^1(j) - lm^2) = x - lm\Big) = \\
&P_{\psi^-}\Big(S(lm^2) = -x + lm,\ \hat{g}\big(\chi_{-x}|[0,lm^2]\big) \sqsubseteq_{\mathcal{I}_{-x}(\psi^-)} g_{-x}(\psi^-)\Big) = \\
&P_{\psi^-}\Big(S(T^3(j) + lm^2) = -x + lm,\ \hat{g}\big(\chi|[T^3(j), T^3(j) + lm^2]\big) \sqsubseteq_{\mathcal{I}_{-x}(\psi^-)} g_{-x}(\psi^-)\Big|S(T^3(j) = -x)\Big) = \\
&P_{\psi^-}\Big(S(T^3(j) + lm^2) = -x + lm,\ \hat{g}\big(w^3(j)\big) \sqsubseteq_{\mathcal{I}_{-x}(\psi^-)} g_{-x}(\psi^-)\Big|S(T^3(j) = -x)\Big).
\end{aligned}$$

Suppose $\psi \in B^1_{\text{intervals OK}}$. Then the probability in the previous display has the lower bound

$$\inf_{\psi:\psi|[y,y+lm] \text{ is OK}} P_\psi\Big(S(T^3(j) + lm^2) = y + lm,\ \hat{g}(w^3(j)) \sqsubseteq_{\mathcal{I}_y(\psi)} g_y(\psi)\Big|S(T^3(j)) = y\Big). \tag{1.12}$$

Indeed, (1.12) does not depend on y any more. It is not very hard to see now that by **2)** of Theorem 1.2, (1.12) can be bounded below by

$$\inf_{\psi:\psi|[y,y+lm] \text{ is OK}} \prod_{i \in I(\psi)} P_\psi\Big(\hat{g}(\chi^{im^2}_{(i-1)m^2}) \sqsubseteq g(\psi^{im}_{(i-1)m})\Big|S(im^2) = S((i-1)m^2) + m\Big) \geq \Big(\frac{3}{4}\Big)^l.$$

Finally, for every ψ,

$$P_\psi\Big(S(T^3(j)) = y\Big|S(T^1(j)) = x\Big) = (p_L)^{c_1 l_1}.$$

Hence, if $\psi \in B^1_{\text{intervals OK}}$, we have

$$P_\psi\big(Y_j = 1|E_{\text{stop}}(\tau)\big) \geq \exp(-1.5 l_1)\Big(\frac{3}{4}\Big)^l (p_L)^{c_1 l_1}\Big(\frac{3}{4}\Big)^l = \exp[-(1.5 - 2\ln(\frac{3}{4}) l_2 - c_1 \ln(p_L)) l_1]. \tag{1.13}$$

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 (3.21), we obtain that the the probability of selecting (1.4) is sufficiently big.

1.4.5 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 (1.13) (which also goes to zero as l_1 grows). To explain the meaning of the additional modification, we consider the events

$$E_{z,I} := \big\{\forall i \in I \text{ we have that } S_z(m(i-1)) < m(i-1) - Lm^2\big\}, \quad I \subseteq \{1, \ldots, l\}. \tag{1.14}$$

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 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 $P(\hat{g}(\chi_z^{lm^2}) \sqsubseteq_I g_y(\xi)) \leq (0.5)^{(n^2-1)|I|}$ and, for l big enough the latter yields

$$P\left(B^1_{\text{intervals OK}} \cap \{\hat{g}(\chi_z^{lm^2}) \sqsubseteq_{\mathcal{I}_y(\xi)} g_y(\xi)\} \cap E_{z,\mathcal{I}(\xi)}\right) \leq \exp[-(0.3n)l]. \tag{1.15}$$

(Corollary 2.1). Recall that on $B^1_{\text{intervals OK}}$. Since n can be chosen very big, the right side of (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 $S(T^3(j)) \geq y$, 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 $z \geq y$, 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 $z \geq y$, the additional requirement obviously reduces the probability (1.13); however, if the q-reproduction has a relatively big probability, then the lower bound like (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 (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

$$\psi(y-L) = \cdots = \psi(y-1) \neq \psi(y) = \cdots = \psi(y+m^3L) \neq \psi(y+m^3L+1) = \cdots = \psi(y+m^3L+L)$$
$$\psi(x+L) = \cdots = \psi(x+1) \neq \psi(x) = \cdots = \psi(x-m^3L) \neq \psi(x-m^3L-1) = \cdots = \psi(x-m^3L-L).$$

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 3.1 states that we have sufficiently many barriers. In Proposition 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 $S_z((i_1-1)m^2) \leq m^3L$. The latter obviously holds $S_z(t) \leq m^3L$ $\forall t \leq (i_1-1)m^2$, 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 $S_z(t) \leq m^3L$ $\forall t \leq (i_1-1)m^2$, 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 $\hat{g}(\chi|[(i_1-1)m^2, i_1 m^2]) \sqsubseteq g(\psi|[(i_1-1)m, i_1 m])$. 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 $w(0) \neq w(1) = \cdots = w(u-1) \neq w(u)$. 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 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 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 $i = 1, 2, \ldots, l$ 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: $q(w) = (q_1(w), \ldots, q_l(w))$, $\hat{q}(v) = (\hat{q}_1(v), \ldots, \hat{q}_l(v))$ where

$$\begin{aligned} q_i(w) &:= \text{ number of big clusters contained in sub-vector } (w(0), \ldots, w(im)) \\ \hat{q}_i(v) &:= \text{ number of big blocks contained in sub-vector } (v(0), \ldots, v(im^2)). \end{aligned}$$

As usually we define $q^*(w) := q(w^-)$ and $\hat{q}^*(v) = \hat{q}^*(v^-)$.
We denote

$$\hat{q}(v) \leq q(w) \quad \big(\hat{q}^*(v) \leq q^*(w)\big) \quad \text{if and only if} \quad \hat{q}_i(v) \leq q_i(w) \quad \big(\hat{q}_i^*(v) \leq q_i^*(w)\big) \text{ for all } i.$$

Hence, if $\hat{q}(\chi_z|[0, lm^2]) \leq q(\psi|[0, ml]) =: q(\psi)$, 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 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 (1.13) (Theorem 2.3).

1.4.6 Final selection

We are now ready to define the final version of the selection rule.

***** Note, for every $u \in \{0,1\}^{lm+1}$, $q(u) = (q_1, \ldots, q_l)$ is vector, such that $q_i = \{0, 1, \ldots, l\}$, $q_1 = 0$ and $q_i \leq q_{i+1} \leq q_i + 1$. Any such vector is called a q**-vector**. Hence, for every u, $q(u)$ and $q^*(u)$ are q-vectors.

Recall that, for any $u \in \{0,1\}^{lm+1}$, $g(u) = (g_1, \ldots, q_l)$, 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 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 $|\mathcal{I}_y(\xi)|, |\mathcal{I}^*_x(\xi)| \geq (1-3\epsilon)l$ for each pair $x, y \in I_1$.

***** We call (I^*, I, q^*, q, g^*, g) **a set of attributes**, if $I^*, I \subset \{1, \ldots, l\}$, $|I^*|, |I| \geq l(1 - 3\epsilon(n))$, q, q^* are q-vectors and g^*, g are g-words.

Recall the definition if observation words $w^1(j), w^2(j), w^3(j)$, $j = 1, \ldots, \exp(\alpha l_1)$. For each set of attributes (I^*, I, q^*, q, g^*, g) we define the set $J(I^*, I, q^*, q, g^*, g) \subset [1, \exp(\alpha l_1)]$ as follows:

$j \in J(I^*, I, q^*, q, g^*, g)$ if and only if j satisfies

$$\hat{q}^*(w^1(j)) \leq q^*, \quad \hat{g}^*(w^1(j)) \sqsubseteq_{I^*} g^*, \quad \hat{q}(w^3(j)) \leq q, \quad \hat{g}(w^3(j)) \sqsubseteq_I g. \tag{1.16}$$

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 1.4.1.

With g- and q-reproduction, the getting selected part (**a)**) means that (with high probability) for each $x, y \in I_1$, $y - x = Lc_1 l_1$ 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 $j \in J(x,y)$, (1.16) 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 $\exists J' \subset J(I^*, I, q^*, q, g^*, g)$ such that $|J'| \geq \exp(\gamma l_1)$ 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 (1.16), 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 1.3 *We define the set $\mathcal{W} = \mathcal{W}(\chi^{12\alpha l_1}, \tau)$ as follows. A word $w \in \{0,1\}^{c_1 l_1 + 1}$ 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) $|J(I^*, I, q^*, q, g^*, g)| \geq \exp(\gamma l_1)$

b) *if $j \in J(I^*, I, q^*, q, g^*, g)$, then $w^2(j) = w$.*

2 Iteration

In this Section, we formalize g- and q-reproduction, described in Subsection 1.4.3. 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 2.1). In Subsection 2.2, we prove the inequality (1.15) (Theorem 2.2). In Subsection 2.3, we formalize q-reproduction and we found a suitable lower bound for (1.12) (Theorem 2.3). This is the main ingredient for obtaining the lower bound (1.13). Finally, in Subsection 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.

2.1 OK cells

In Theorem 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 $w_1^N := (w(1), \dots, w(N))$, $w(j) \in \{0,1\}$ be a binary 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.

* Define integer intervals

$$D_i :=: [d_{i-1}, d_i] := (d_{i-1}, \dots, d_i), \quad \text{where} \quad d_i := im, \quad i = 1, 2 \dots .$$

Clealy 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:

$$w_i \in \{0,1\}^{m+1}, w_i := w_{d_{i-1}}^{d_i} = (w(d_{i-1}), \dots, w(d_i)), \quad i = 1, \dots, l. \tag{2.1}$$

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 (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 $N \geq m^{1.1}$ 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 (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.

* Let w be as in (2.1). Define

$$\mathcal{I}_{II}^1(w) := \{i \in [2Lm^2, l] : w_i \text{ is isolated}\}$$
$$\mathcal{I}_{II}^2(w) := \{i \in [2Lm^2, l] : w_i \text{ has empty neighborhood}\}$$
$$\mathcal{I}_{II}(w) := \mathcal{I}_{II}^1(w) \cap \mathcal{I}_{II}^2(w), \quad \mathcal{I}(w) := \mathcal{I}_I(w) \cap \mathcal{I}_{II}(w).$$

* 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,

$$E_{\text{OK}} = E_{\text{OK}a} \cap E_{\text{OK}b} \tag{2.2}$$

and on E_{OK}

$$|\mathcal{I}(\xi^{ml})| \geq l(1-3\epsilon(n)), \tag{2.3}$$

provided n is big enough.

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 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*

$$P(E_{\text{OK}}) \geq 1 - e^{-al}.$$

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 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 (2.1) we extend the definition of g as follows

$$g : \{0,1\}^{lm+1} \mapsto \{0,1\}^{l(n^2+1)}, \quad g(w) := (g(w_1), g(w_2), \dots, g(w_l)). \tag{2.4}$$

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

$$T_i =: [t_{i-1}, t_i] := (t_{i-1}, \dots, t_i), \quad \text{where} \quad t_i := im^2, \quad , i = 1, 2 \dots.$$

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 $T_1 \cup \dots \cup T_l = [0, lm^2]$.

* Consider binary words $v = (v(1), \dots, v(lm)) \in \{0,1\}^{lm^2+1}$. For each such a word we define l sub-words, $v_1, \dots, v_l$ as follows:

$$v_i \in \{0,1\}^{m+1}, v_i := v^{t_i}_{t_{i-1}} = (v(t_{i-1}), \dots, v(t_i)), \quad i = 1, \dots, l. \tag{2.5}$$

Regarding v as a mapping, we equivalently define $v_i = v|T_i$.
Using the sub-words (2.5) we define

$$\hat{g} : \{0,1\}^{lm+1} \mapsto \{0,1\}^{ln^2}, \quad \hat{g}(v) := (\hat{g}(v_1), \hat{g}(v_2), \dots, \hat{g}(v_l)).$$

* Let $A = (a'_1, \ldots, a'_l)$, $B = (b'_1, \ldots, b'_l)$ be lp and lr dimensional vectors, respectively. Let $I \subseteq \{1, 2, \ldots, l\}$. We define the following notation:

$$A \sqsubseteq_I B \quad \text{iff for each } i \in I \text{ we have that} \quad a'_i \sqsubseteq b'_i.$$

Recall the definition of $E_{z,I}$ in (1.14). The event $E_{z,I}$ says that for each $i \in I$ we have that 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 (1.15).

Theorem 2.2 *There exists $\alpha_I(n) > 0$ not depending on l, such that for all $z < 0$ the following holds:*

$$P\begin{pmatrix} \exists I \subset \{1,2,...,l\} \text{ with } |I| = l(1-3\varepsilon(n)) \text{ such that} \\ E_{z,I} \text{ holds and } \hat{g}(\chi_z^{lm^2}) \sqsubseteq_I g\left(\xi^{ml}\right) \end{pmatrix} \leq e^{-\alpha_I l}, \tag{2.6}$$

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

$$X_i = 1 \quad \text{iff} \quad \hat{g}(\chi_{z,i}) \sqsubseteq g(\xi_i)$$
$$Y_i = 1 \quad \text{iff} \quad S_z(t_{i-1}) < d_i - Lm^2.$$

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:

$$P(w \sqsubseteq g(\xi_i)) = (0.5)^{n^2-1} \tag{2.7}$$

Note, when $\{Y_i = 1\}$ holds, then $g(\xi_i)$ is independent of $\hat{g}(\chi_{z,i})$. By (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 $I \subset \{1, \ldots, l\}$. 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

$$P\Big(E_{z,I} \cap \{\hat{g}(\chi_z^{lm^2}) \sqsubseteq_I g(\xi^{ml})\}\Big) = E(\prod_{i\in I} X_i Y_i) = P(\prod_{i \in I} X_i Y_i = 1) =$$
$$P(X_i = 1, i \in I|Y_i = 1, i \in I)P(Y_i = 1, i \in I) \leq (0.5)^{(n^2-1)|I|} \tag{2.8}$$

Using(2.8), the probability in (2.6) can bound by

$$\sum_{\substack{I \subset \{1,2,\ldots,l\}, \\ |I| = l(1-3\epsilon(n))}} P\Big(E_{z,I} \cap \big\{\hat{g}(\chi_z^{lm^2}) \sqsubseteq_I g(\xi^{ml})\big\}\Big) \leq \begin{pmatrix} l \\ 3l\varepsilon(n) \end{pmatrix} \left(\frac{1}{2}\right)^{(n^2-1)l(1-3\epsilon(n))}. \tag{2.9}$$

Using Stirling's approximation, one can show that for l big enough

$$\begin{pmatrix} l \\ 3l\varepsilon(n) \end{pmatrix} \leq \exp[-l\big((3\varepsilon(n)\ln(3\varepsilon(n)) + (1-3\varepsilon(n))\ln(1-3\varepsilon(n))\big)] = \exp(-l\epsilon_2(n)),$$

where $\epsilon_2(n) := 3\varepsilon(n)\ln(3\epsilon(n)) + (1-3\epsilon(n))\ln(1-3\epsilon(n)) \to 0$, as n grows. Hence, if n is big enough, then the sum in (2.9) can bounded by

$$\exp(-l\epsilon_2(n))\Big((0.5)^{(n^2-1)l(1-3\epsilon(n))}\Big) \le \exp(-ln^2\frac{\ln 2}{2}) = \exp(-l\alpha_I(n)),$$

where $\alpha_I(n) = n^2\frac{\ln 2}{2}$. ■

2.3 Counting blocks

We now give the formal definition of block.

* Let $w = (w(u), \dots, w(v))$ be a 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 $v-u$. 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**. For two blocks, $A = [a_1, a_2], B = [b_1, b_2]$ we denote $A < B$ if $a_1 < b_1$.

Let $\phi : D \longrightarrow \{0, 1\}$ be a piece of scenery. Let $T = [t_1, t_2] \subset D$ be an integer interval of length at least 3. Since $\phi|T$ can be considered as a word, the definition of block applies to $\phi|T$ as well.
For given ϕ, we also call a location interval $T = [t_1, t_2]$ a block of ϕ, if $\phi|T$ is a block (as word). So, in the following, a block can be a certain pattern (word) or a certain location (T), where ϕ has a block. In the latter case, speaking about blocks, the piece ϕ must be specified. We call a block **big** if its length is at least m^3.

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 $B = [b_1, b_2] \subset \mathbb{Z}$ be a block of scenery ψ. Define

$$l(B) := \min\left\{ n > 1 \middle| \begin{array}{c} \exists R \in \mathcal{R}(n) \quad \text{such that} \quad \psi \circ R = \psi(R(0)), \dots, \psi(R(n)) \\ \text{is a block}, R(0) \le b_1, R(n) \ge b_2 \end{array} \right\}. \tag{2.10}$$

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 (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 ψ 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 $(w(1), w(2), \ldots, w(11)) = 0011111111000$. 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 $w(1), \ldots, w(18) = 011101111111110111$. This word contains a block of length 10 $B = (w(5), \ldots, w(15))$. 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 $P(S(1)-S(0) = i) > 0$, $i = -4, -3, \ldots, 3, 4$. 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 $\psi \circ R(0) = \psi \circ R(1) = \cdots = \psi \circ R(n)$.
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, $i = 1, \ldots, h$;
- $B_1, \ldots, B_h$ 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

$$l(C) := \min\{n > 1 : \exists R \in \mathcal{R}(n) \text{ such that } \quad \psi(R(0)), \ldots, \psi(R(n)) \text{ is a block}, R(0) \leq c, R(n) \geq d\},$$

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, $l(C) \geq m^3$. The reading-path of C is any path that attains the minimum above.

* Let us fix $\psi \in E_{\text{OK}}$. Denote $\mathcal{I} = \mathcal{I}(\psi^{m^2 l})$, $\mathcal{I}_I = \mathcal{I}_I(\psi^{m^2 l})$, $\mathcal{I}_{II} = \mathcal{I}_{II}(\psi^{m^2 l})$.
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.

$$\mathcal{I}_{II}^c = [l_1, l_2] \cup [l_3, l_4] \cup \cdots \cup [l_{2k-1}, l_{2k}], \tag{2.11}$$

where $l_1 = 1, l_2, l_3, \ldots \in [2Lm^2, l]$, $l_i \leq l_{i+1}$.
The set of cell-indexes $[l_{2i-1}, l_{2i}]$ corresponds to the location-interval (cells) $[(l_{2i-1}-1)m, l_{2i}m]$ or $[d_{l_{2i-1}-1}, d_{l_{2i}}]$.
Let us denote

$$r_i := (l_{2i-1} - 1)m, \quad s_i = l_{2i}m, \quad i = 1, \ldots, k. \tag{2.12}$$

By definition, S visits every point in $\mathbb{Z}$ i.o.. This means, there exists an integer $k \geq 1$ 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 $R(0), \ldots, R(\bar{v})$ such that $R(0) = 0$ and $R(\bar{v}) = 1$. Similarly, between points $a < b$ there exists an admissible path $R(0), \ldots, R((b-a)\bar{v})$ such that $R(0) = a$, $R(\bar{v}) = a+1$, $R(2\bar{v}) = a+2, \ldots R((b-a)\bar{v}) = b$. 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 1.4.5 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

$$F_{min}(1) := \big\{\hat{q}(\chi|[0, s_1 m]) \leq q(\psi|[0, s_1]),\ \chi|[s_1 - m\bar{v}, s_1] \text{ contains both colors}, \quad S(s_1 m) = s_1\big\}.$$
$$F_{min}(j) := \big\{\hat{q}(\chi_{r_j}|[0, (s_j - r_j)m]) \leq q(\psi|[r_j, s_j]),\ \chi_{r_j}|[0, m\bar{v}] \text{ and } \chi_{r_j}|[(s_j - r_j)m - m\bar{v}, (s_j - r_j)m] \text{ contain both colors } S_{r_j}((s_j - r_j)m) = s_j\big\},\ j = 2, \ldots, k-1.$$

For the last interval in (2.12) we define $F(k)$ as $F(j)$, $j > 1$, if $s_k < l$. If $r_k = l$, we define

$$F_{min}(k) := \big\{\hat{q}(\chi_{r_k}|[0, (l - r_k)m]) \leq q(\psi|[r_k, l]),\ \chi_{r_k}|[0, m\bar{v}] \text{ contains both colors},\ S_{r_k}((l - r_k)m) = l\big\}.$$

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 $R_j \subset \mathcal{R}((s_j - r_j)m)$ such that

$$R_j(0) = r_j,\ \ R_j((s_j - r_j)m) = s_j, \tag{2.13}$$
$$\hat{q}(\psi \circ R_j) \leq q(\psi[s_j, r_j]), \tag{2.14}$$
$$\text{if } r_j \neq 0 \text{ and } s_j \neq l$$
$$\text{then } (\psi \circ R_j)|[0, m\bar{v}], \quad \text{and} \quad (\psi \circ R_i)|[(s_j - r_j) - m\bar{v}, (s_j - r_j)] \text{ have both colors}, \tag{2.15}$$

then $F_{min}(i) \neq \emptyset$ and $P_\psi\big(F_{min}(i)\big) \geq (p_{min})^{(s_i - r_i)m}$. The following proposition, proved in Appendix, shows that for each i, at least one such admissible path exists.

Proposition 2.1 *For each $i = 1, \ldots k$ the following holds:*

$$P_\psi\big(F_{min}(i)\big) \geq (p_{min})^{(s_i - r_i)m} = (p_{min})^{(l_{2i} - l_{2i-1} + 1)m^2}. \tag{2.16}$$

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 2.3 *There exist constant $\alpha_{II}(n) > 0$ not depending on l, such that for all $\psi \in E_{\mathrm{OK}}$ the following holds:*

$$P_\psi\Big(\hat{g}(\chi^{lm^2}) \sqsubseteq_{\mathcal{I}} g(\xi^{ml}), \quad \hat{q}(\chi^{m^2 l}) \leq q(\xi^{ml}), \quad S(m^2 l) = ml\Big) \geq e^{-l\alpha_{II}}. \tag{2.17}$$

Proof. For each subset $i \in [1, l]$ and $I \subset [1, l]$ we define

$$E_S(i) := \{S(t_{i-1}) - S(t_i) = m\}, \quad E_S(I) := \cap_{i \in I} E_S(i)$$
$$E_{\sqsubseteq}(i) := \{\hat{g}(\chi_z|T_i) \sqsubseteq g(\psi|D_i)\}, \quad E_{\sqsubseteq}(I) := \cap_{i \in I} E_{\sqsubseteq}(i);$$
$$E_{no-block}(i) := \{ \text{ the sequence } \chi_z|T_i \text{ contains both colors } \}, \quad E_{no-block}(I) := \cap_{i \in I} E_{no-block}(i).$$

Use $[r_j, s_j]$, $j = 1, \ldots, k$ as in (2.12) to define

$$E_{min}(1) := \{\hat{q}(\chi|[0, s_1 m]) \leq q(\psi|[0, s_1]), \quad S(s_1 m) = s_1, \ \chi|[s_1 - m\bar{v}, s_1] \text{ contain both colors}\}$$
$$E_{min}(j) := \{\hat{q}(\chi|[r_j m, s_j m]) \leq q(\psi|[r_j, s_j]), \quad S(s_j m) = s_j, \ \chi|[r_j, r_j + m\bar{v}] \text{ contain both colors},\ \chi|[s_j - m\bar{v}, s_j] \text{ contain both colors}\},$$
$$j = 2, \ldots, k-1 \quad \text{and}$$
$$E_{min} := \cap_{j=1}^k E_{min}(j).$$

If $s_k = l$, then the requirement $\{\chi|[s_k - m\bar{v}, s_k]$ contain both colors$\}$ is dropped for the definition of $E_{min}(k)$.

Consider the event $E_{min} \cap E_S(\mathcal{I}_{II})$. Use the relation $E_{min} \cap E_S(\mathcal{I}_{II}) \subset E_S([1,l])$ to deduce that

$$P_\psi(E_{min} \cap E_S(\mathcal{I}_{II})) = \prod_{j=1}^{k} P_\psi(E_{min}(j)|S(mr_j) = r_j)P_\psi(E_S(\mathcal{I}_{II}))$$
$$= \prod_{j=1}^{k} P_\psi(F_{min}(j))P_\psi(E_S(\mathcal{I}_{II})) \geq (p_{min})^{|\mathcal{I}_{II}^c|m^2} P_\psi(E_S(\mathcal{I}_{II})). \quad (2.18)$$

Let $i \in \mathcal{I}_{II}$. If $i \neq l_{2j-1}$ for each $j = 1,\ldots,k$, then $E^c_{no-block}(i)$ does not depend on $F_{min}(j)$. If $i = l_{2j-1}$ for a $j = 1,\ldots,k$, then $P_\psi(E^c_{no-block}(i)|F_{min}(j)) = P_\psi(E^c_{no-block}(i)|S(t_{i-1}) = d_{i-1})$ and $E^c_{no-block}(i)$ is independent of $F_{min}(j')$, $j' \neq j$. Hence

$$P\big(E^c_{no-block}(i) \cap E_{min} \cap E_S(\mathcal{I}_{II})\big) = \prod_{i=1}^{k} P_\psi(F_{min}(i))P\big(E^c_{no-block}(i)|S(t_{i-1}) = d_{i-1}, E_S(i)\big)P_\psi(E_S(\mathcal{I}_{II})),$$

implying that, for each $i \in \mathcal{I}_{II}$

$$P_\psi\big(E^c_{no-block}(i)|E_{min} \cap E_S(\mathcal{I}_{II})\big) = \quad (2.19)$$
$$P_\psi\big(E^c_{no-block}(i)|S(t_{i-1}) = d_{i-1}, S(t_i) = d_i\big) = P_\psi\big(E^c_{no-block}(i)|E_S([1,l])\big).$$

Let us estimate (2.19). If $S(t_{i-1}) = d_{i-1}$ and $S(t_i) = d_i$, then during T_i 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

$$P(E_S([1,l])) = (p_m)^l. \quad (2.20)$$

So, for each $i \in \mathcal{I}_{II}$, it holds

$$P_\psi\Big(E^c_{no-block}(i)\Big|E_{min} \cap E_S(\mathcal{I}_{II})\Big) = P_\psi\Big(E^c_{no-block}(i)\Big|E_S([1,l])\Big) \leq \frac{\exp(-am^{0.2})}{(p_m)^l}.$$

Now, by local central limit theorem, p_m is of order $\frac{1}{m}$. Thus, when m is big enough

$$P_\psi(E_{no-block}(i)|E_{min} \cap E_S(\mathcal{I}_{II})) > 0.75, \quad P_\psi(E_{no-block}(\mathcal{I}_{II} \backslash \mathcal{I})|E_{min} \cap E_S(\mathcal{I}_{II})) > (0.75)^{|\mathcal{I}_{II}|-|\mathcal{I}|}. \quad (2.21)$$

The second inequality follows because conditional on $E_{min} \cap E_S(\mathcal{I}_{II})$, everything that happens during the time-interval T_i, is independent of events happening during the time-interval T_j, $j,i \in \mathcal{I}_{II}$. Hence, for $j,i \in \mathcal{I}_{II}$ the events $E_{no-block}(i)$ and $E_{no-block}(j)$ are conditionally independent.

Suppose now $i \in \mathcal{I} \subset \mathcal{I}_{II}$. Then $\psi|D_i$ is weak-OK. By **2)** of Theorem 1.2 we now get that

$$P_\psi(E^c_{\sqsubseteq}(i)|E_{min} \cap E_S(\mathcal{I}_{II})) = P_\psi(E^c_{\sqsubseteq}(i)|S_{d_{i-1}}(m^2) = d_i) \leq 0.25.$$

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

$$P_\psi\Big(E_{no-block}(\mathcal{I})\cap E_{\sqsubseteq}(\mathcal{I})\Big|E_{min}\cap E_S(\mathcal{I}_{II})\Big)>(0.5)^{|\mathcal{I}|}. \qquad (2.22)$$

Finally, by the same independence-argument, (2.22) and (2.21),

$$\begin{aligned}&P_\psi\Big(E_{no-block}(\mathcal{I}_{II})\cap E_{\sqsubseteq}(\mathcal{I})\Big|E_{min}\cap E_S(\mathcal{I}_{II})\Big)=\\ &P_\psi\Big((E_{no-block}(\mathcal{I})\cap E_{\sqsubseteq}(\mathcal{I}))\cap E_{no-block}(\mathcal{I}_{II}\backslash\mathcal{I})\Big|E_{min}\cap E_S(\mathcal{I}_{II})\Big)>(0.5)^l\end{aligned} \qquad (2.23)$$

Consider $[r_j,s_j]$, $j=1,\dots,k$ as in (2.12). By the definition of $\mathcal{I}_{II}$, $[s_i-Lm^2,r_{i+1}+Lm^2]$ is empty, for each $i=1,\dots,k-1$ 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 $[s_i-Lm^2-m,s_i-Lm^2]$ as well as $[r_{i+1}+Lm^2,r_{i+1}+Lm^2+m]$ $(i=1,\dots,k-1)$ and $[s_k-Lm^2-m,s_i-Lm^2-m]$, if $s_k<l$, contain a fence. This means that a interval $[s_i-Lm^2,r_{i+1}+Lm^2]$ $(i=1,\dots,k-1)$ as well as $[s_k-Lm^2-m,s_i-Lm^2-m]$ (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 $[s_i,r_j]$ imply that the cluster-counting vector $q(\psi^{m^2l})$ is constant on $\mathcal{I}_{II}$.
The event $E_{no-block}(\mathcal{I}_{II})\cap E_{min}$ ensures that the word $\chi|[s_i-m\bar{v},r_{i+1}+m\bar{v}]$, $i=1,\dots,k-1$ 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 $\hat{q}(\chi^{m^2l})$ is constant on $\mathcal{I}_{II}$. Thus, $\hat{q}(\chi^{t_l})\le q(\psi^{t_l})$ if $\hat{q}_i(\chi^{t_l})\le q_i(\psi^{t_l})$ for each $i\in\mathcal{I}^c_{II}$. The latter holds if and only if $\hat{q}(\chi|[r_jm,s_jm])\le q(\psi|[r_j,s_j])$ for each $j=1,\dots,k$. Hence

$$E_{min}\cap E_{no-block}(\mathcal{I}_{II})\subset\{\hat{q}(\chi^{m^2l})\le q(\xi^{ml})\}.$$

This means

$$E_{min}\cap E_{no-block}(\mathcal{I}_{II})\cap E_{\sqsubseteq}(\mathcal{I})\cap E_S(\mathcal{I}_{II})\subset\Big\{\hat{g}(\chi^{lm^2})\sqsubseteq_{\mathcal{I}}g(\xi^{ml}),\quad \hat{q}(\chi^{m^2l})\le q(\xi^{ml}),\quad S(m^2l)=ml\Big\}. \qquad (2.24)$$

From (2.23), (2.20) and (2.18) it follows

$$\begin{aligned}&P_\psi\Big(E_{min}\cap E_{no-block}(\mathcal{I}_{II})\cap E_{\sqsubseteq}(\mathcal{I})\cap E_S(\mathcal{I}_{II})\Big)=\\ &P_\psi\Big(E_{\sqsubseteq}(\mathcal{I})\cap E_{no-block}(\mathcal{I}_{II})\cap\Big|E_{min}\cap E_S(\mathcal{I}_{II})\Big)P_\psi\Big(E_{min}\cap E_S(\mathcal{I}_{II})\Big)>\\ &(0.5)^lP_\psi\Big(E_{min}\cap E_S(\mathcal{I}_{II})\Big)\ge(0.5)^l(p_{min})^{|\mathcal{I}^c_{II}|m^2}P_\psi(E_S(\mathcal{I}_{II}))\ge(0.5)^l(p_{min})^{|\mathcal{I}^c_{II}|m^2}(p_m)^l.\end{aligned} \qquad (2.25)$$

Hence (2.24), (2.25) and the inequality $|\mathcal{I}^c_{II}|\le l\exp(-m^{0.7})$ imply

$$\begin{aligned}&P_\psi\Big(\hat{g}(\chi^{lm^2})\sqsubseteq_{\mathcal{I}}g(\xi^{ml}),\quad\hat{q}(\chi^{m^2l})\le q(\xi^{ml}),\quad S(m^2l)=ml\Big)\ge(0.5)^l(p_{min})^{|\mathcal{I}^c_{II}|m^2}(p_s)^l\ge\\ &[0.5p_m(p_{min})^{m^2\exp(-m^{0.7})}]^l=\exp[l(\ln(0.5p_m)+m^2\exp(-m^{0.7})\ln(p_{min}))]=\exp[-l\alpha_{II}(m)].\end{aligned}$$

■

Let us show that, for n big enough,

$$8\alpha_{II}(n)=-8\ln(0.5p_s)-m(n)^2\exp(-m(n)^{0.7})\ln(p_{min})<n^2\frac{\ln 2}{2}=\alpha_I(n) \qquad (2.26)$$

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 $-\ln(0.5p_s)$. 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 (2.26) clearly holds.

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 (1.14). The next theorem formalizes the argument explained in Subsection 1.4.5.

Theorem 2.4 *If $z < 0$ then*

$$E_{origin} \cap \{\hat{q}(\chi_z^{t_l}) \leq q(\xi^{ml})\} \subset E_{z,\mathcal{I}(\xi^{ml})}. \tag{2.27}$$

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

$$q_i(\psi^{m^2l}) = q_{i-1-Lm}(\psi^{m^2l}). \tag{2.28}$$

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 (2.28), this number equals $q_i(\psi^{m^2l})$. Hence, if $S_z(t_i) \geq d_{i-1-Lm}$, then $\hat{q}_i(\chi_z^{m^2l}) > q_i(\psi^{m^2l})$. Consequently, $E_{\text{origin}} \cap E^c_{z,\mathcal{I}} \subset E_{\text{origin}} \cap \{\hat{q}(\chi_z^{t_l}) \leq q(\xi^{ml})\}^c$. This proves the statement. ■

Define

$$E_{\text{mistake}}(z) := \left\{\hat{q}(\chi_z^{m^2l}) \leq q(\xi^{ml})\right\} \cap \left\{\hat{g}(\chi_z^{m^2l}) \sqsubseteq_{\mathcal{I}(\xi^{ml})} g(\xi^{ml})\right\} \cap E_{\text{origin}}.$$

Corollary 2.1 *If $z < 0$, then for n and l big enough*

$$P\big(E_{mistake}(z) \cap E_{OK}\big) \leq \exp(-\alpha_I l). \tag{2.29}$$

Proof. By (2.27) we have

$$E_{\text{mistake}}(z) \subset E_{z,\mathcal{I}(\xi^{ml})} \cap \left\{\hat{g}(\chi_z^{m^2l}) \sqsubseteq_{\mathcal{I}(\xi^{ml})} g(\xi^{ml})\right\}.$$

Thus

$$E_{\text{mistake}}(z) \cap E_{OK} \subset E_{z,\mathcal{I}(\xi^{ml})} \cap \left\{\hat{g}(\chi_z^{m^2l}) \sqsubseteq_{\mathcal{I}(\xi^{ml})} g(\xi^{ml})\right\} \cap E_{OK}. \tag{2.30}$$

Consider the right side of (2.30). By E^*_{OK} and (2.3), $|\mathcal{I}(\xi^{ml})| \geq l(1-3\epsilon(n))|$. Thus, if the right side of (2.30) holds, then there exists a subset $I \subset \mathcal{I}(\xi^{ml})$ such that $|I| = |l(1-3\epsilon(n))|$, $\{\hat{g}(\chi_z^{m^2l}) \sqsubseteq_I g(\xi^{ml})\Big\}$ and $E_{z,I}$ holds. By Theorem 2.2, this event has probability not bigger than $\exp(-l\alpha_I)$. ■

3 Reconstruction at level l_1

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 and $\phi \in \{0,1\}^I$. The mapping ϕ is called a **ladder piece**. If $\phi \in \{0,1\}^D$, $I \subset D$ is a ladder interval, we sometimes say that $\phi|I$ is a ladder piece of ϕ (or $\phi|D$).

Hence, a ladder piece of non-random scenery ψ is any vector $(\psi(a), \psi(a+L), \dots, \psi(a+nL))$, $a \in \mathbb{Z}$, $n \in \mathbb{N}$.

Recall: two pieces of scenery $\phi :\in \{0,1\}^D$ and $\phi' \in \{0,1\}^{D'}$ are equivalent, $\phi \approx \phi'$, if there is a mapping $T : \mathbb{Z} \mapsto \mathbb{Z}$, $T(z) = az + b$, $a \in \{+1, -1\}$, $b \in \mathbb{Z}$ such that $T : D \mapsto D'$ is a bijection and $\phi(z) = \phi'(T(z))$ $\forall z \in D$. We also write $T\phi = \phi'$.
Every word $w \in \{0,1\}^{n+1}$ is a mapping $w \in \{0,1\}^{[0,n]}$. Hence, if $I = [a, a+n]$ for some n, and $\phi \in \{0,1\}^I$, then the equivalence $\phi \approx w$ is well defined. By definition, it means that $\phi(a) = w(1), \dots, \phi(a+n) = w(n+1)$ or $\phi(a) = w(n+1), \dots, \phi(a+n) = w(1)$.
Let $I = (a, a+L, \dots, a+nL)$ be a ladder-interval and let $\phi \in \{0,1\}^I$ be a ladder piece. We write $\phi \approx_l w$, if $\phi(a) = w(1), \dots, \phi(a+Ln) = w(n+1)$ or $\phi(a) = w(n+1), \dots, \phi(a+Ln) = w(1)$. Hence, if $L = 1$, then the relation " $\approx_l$ " is the same as the equivalence " $\approx$ ".

Given a ladder piece $\phi \in \{0,1\}^I$, $I \in \mathcal{L}(n)$, we say that $w \in \{0,1\}^{n+1}$ is a **ladder word** of ϕ, if $\phi \approx_l w$. 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 [L,M,M], w is a ladder word of ϕ, if $w \in \{(\phi)_\rightarrow, (\phi)_\leftarrow\}$.)

* In this chapter, l_1, c_1 stand for positive integers, they will be specified later. We denote

$$I_1 := [-\exp(3l_1), \exp(3l_1)].$$

* The following event, $B^1_{\text{unique fit}}$, states that any ladder piece of $\xi|I_1$ of length $\frac{l_1 c_1}{4}$ has unique ladder word up to equivalence. Formally,

$$B^1_{\text{unique fit}} := \Big\{ \text{if } I, J \in \mathcal{L}(l_1 c_1/4), I, J \subset I_1 \text{ and } I \neq J \text{ then } \xi|I \not\approx \xi|J \Big\}.$$

* Suppose $x, y \in \mathbb{Z}$, $y = x + (l_1 c_1)L$. In this case there is only one admissible path of length $c_1 l_1$ from x to y, i.e. there exist unique $R \in \mathcal{R}(l_1 c_1)$ such that $R(c_1 l_1) - R(0) = (l_1 c_1)L$. Obviously, this path consists of maximum jumps, only, i.e. $R(i+1) - R(i) = L$, $i = 0, 1, \dots, l_1 c_1 - 1$.
Suppose now that $x, y \in \mathbb{Z}$, $x < y$ are such that $y < x + (l_1 c_1)L$. In this case, this might happen that there is no admissible path going from x to y with exactly $l_1 c_1$ 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,

$$B^1_{\text{recon straight}} := \left\{ \begin{array}{c} \text{if } R \in \mathcal{R}(l_1 c_1) \text{ such that } R(0), R(l_1 c_1) \in I_1 \text{ and } R(l_1 c_1) - R(0) < (l_1 c_1)L, \text{ then} \\ \exists R' \in \mathcal{R}(l_1 c_1) \text{ such that } R(0) = R'(0), R(c_1 l_1) = R'(c_1 l_1) \text{ and } \xi \circ R \neq \xi \circ R' \end{array} \right\}.$$

* Let ψ be a non-random scenery. We say that x is a **left-barrier point** of ψ, if

$$\psi(x+L) = \dots = \psi(x+1) \neq \psi(x) = \dots = \psi(x - m^3 L) \neq \psi(x - m^3 L - 1) = \dots = \psi(x - m^3 L - L).$$

We say y is a **right-barrier point** of ψ, if

$$\psi(y-L) = \dots = \psi(y-1) \neq \psi(y) = \dots = \psi(y + m^3 L) \neq \psi(y + m^3 L + 1) = \dots = \psi(y + m^3 L + L).$$

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 $\psi_y := (\psi(i+y))_{i\in\mathbb{Z}}$ belongs to the event E_{origin}. Similarly, x is a left-barrier point, if the translated and reflected scenery $\psi_x^- := (\psi(x-i))_{i\in\mathbb{Z}}$ belongs to the event E_{origin}.

We consider the barriers of ξ, (x, y) such that $y - x = (c_1 l_1)L$. In order to carry on the reconstruction in level l_1, every interval $[z, z + (c_1 l_1/4)L]$, $z \in I_1$ should contain enough left-barrier points of such barriers. This is the meaning of the event $B^1_{\text{enough barriers}}$. More precisely,

$$B^1_{\text{enough barriers}} := \left\{ \begin{array}{c} \text{for any } j = 0, \ldots, L-1 \text{ and for any } z \in I_1, \\ \text{there exists } x \in [z, z + (c_1 l_1/4)L] \text{ such that:} \\ x \bmod L = j \text{ and } (x, x + (c_1 l_1)L) \text{ is a barrier of } \xi \end{array} \right\}.$$

* We now define the left-side counterparts of $g, \hat{g}, q$ and $\hat{q}$. For a word $u = (u_1, \ldots, u_n)$ denote by u^- its reflection, i.e. $u^- := (u_n, \ldots, u_1)$. Now let

$$q^* : \{0,1\}^{lm+1} \mapsto \mathbb{N}^l, \quad \hat{q}^* : \{0,1\}^{lm^2+1} \mapsto \mathbb{N}^l, \quad g^* : \{0,1\}^{lm+1} \mapsto \{0,1\}^{ln^2+1}, \quad \hat{g}^* : \{0,1\}^{lm^2+1} \mapsto \{0,1\}^{ln^2}$$

be as follows

$$q^*(w) = q(w^-), \quad g^*(w) = g(w^-), \quad w \in \{0,1\}^{lm+1} \tag{3.1}$$

$$\hat{q}^*(v) = \hat{q}(v^-), \quad \hat{g}^*(v) = \hat{g}(v^-), \quad v \in \{0,1\}^{lm^2+1}. \tag{3.2}$$

* Finally, we put

$$l = l_1 \cdot l_2.$$

Hence, the requirement "l big enough" in all statement of previous chapter is equivalent to the requirement "l_1 big enough".

3.2 Stopping-time events

3.2.1 Right side

* Let $\tau(1), \tau(2), \ldots$ be a sequence of $\mathcal{F}$-adapted stopping times satisfying

$$\tau(k) - \tau(k-1) \geq 2\exp(2l_1), \quad k = 2, 3, \ldots. \tag{3.3}$$

Let $z \in \mathbb{Z}$. Define

$$\kappa^3(z, 1) := \min\{j : S(\tau(j) + \exp(2l_1) + lm^2 + c_1 l_1) = z\}$$

and, inductively,

$$\kappa^3(z, k) := \min\{j > \kappa^3(z, k-1) : S(\tau(j) + \exp(2l_1) + lm^2 + c_1 l_1) = z\}.$$

Thus, $\kappa^3(z, k)$ is the index of k-th stopping time $\tau(j)$, for which $S(\tau(j) + \exp(2l_1) + lm^2 + c_1 l_1) = z$. Hence, at time

$$T_z^3(k) := \tau\big(\kappa^3(z, k)\big) + \exp(2l_1) + lm^2 + c_1 l_1$$

the random walk is at position z. Let $w_z^3(k)$ denote the observation-word of length lm^2 starting at time $T_z^3(k)$, i.e.

$$w_z^3(k) := \chi|[T_z^3(k), T_z^3(k) + lm^2].$$

* The words $w_z^3(k)$ provide us some information about (unknown) z. This information is captured

in the values of $\hat{q}(w_k^3)$ and $\hat{g}(w_k^3)$ (note that the length of w_k^3 is lm^2). Having sufficiently many k-s, the values $\hat{q}(w_k^3)$ and $\hat{g}(w_k^3)$ give us some information about $g(\xi|[z,z+ml])$ and $q(\xi|[z,z+ml])$. Indeed, by Theorem 2.3, there is a proportion of words w_k^3 such that $\hat{q}(w_k^3) \leq q(\xi|[z,z+ml])$ and $\hat{g}(w_k^3) \sqsubseteq_{\mathcal{I}(\xi|[z,z+ml])} g(\xi|[z,z+ml])$. On the other hand, if $y \in \mathbb{Z}$ is a right-barrier point bigger than z, then, by Corollary 2.1, the probability for the relations above to hold is rather small. Hence we expect that for such y the relations

$$\hat{q}(w_k^3) \leq q(\xi|[y,y+ml]), \quad \hat{g}(w_k^3) \sqsubseteq_{\mathcal{I}(\xi|[y,y+ml])} g(\xi|[y,y+ml]) \tag{3.4}$$

do not occur.
To make these ideas precise, for each $y \in \mathbb{Z}$ we define

$$g_y(\xi) := g(\xi|[y,y+ml]), \quad q_y(\xi) := q(\xi[y,y+ml]).$$

The following event is a counterpart of $E_{\text{mistake}}(z)$. It states that although y is a right-barrier point and $z<y$, the mistake (3.4) still holds.

$$E^1_{\text{mistake-r}}(z,y,k) := \Big\{\hat{q}(w_z^3(k)) \leq q_y(\xi)\Big\} \cap \Big\{\hat{g}(w_z^3(k)) \sqsubseteq_{\mathcal{I}(\xi|[y,y+ml])} g_y(\xi)\Big\} \cap \Big\{y \text{ is a right-barrier point}\Big\}.$$

Finally, let

$$E^1_{\text{mistake-r}} := \bigcup E^1_{\text{mistake-r}}(z,y,k),$$

where the union is taken over all z,y,k such that $z<y$, $z,y \in I_1$ and $k \leq \exp(\alpha l_1)$.

3.2.2 Left side

We now introduce the left-side counterparts of defined notions. At first, let

$$\kappa^1(z,1) := \min\{j : S(\tau(j)+\exp(2l_1)+lm^2) = z\}$$

and, inductively

$$\kappa^1(z,k) := \min\{j > \kappa^1(z,k-1) : S(\tau(j)+\exp(2l_1)+lm^2) = z\}.$$

Thus, $\kappa^l(z,k)$ is the index of k-th stopping time $\tau(j)$, for which $S(\tau(j)+\exp(2l_1)+lm^2) = z$. Hence, at time

$$T_z^1(k) := \tau\big(\kappa^1(z,k)\big) + \exp(2l_1) + lm^2$$

the random walk is at position z. Let $w_z^1(k)$ denote the observation-word of length lm^2 ending at time $T_z^1(k)$, i.e.

$$w_z^1(k) := \chi|[T_z^1(k)-lm^2, T_z^1(k)].$$

As previously, we consider the characteristics $\hat{q}^*(w_z^1(k))$, $\hat{g}^*(w_z^1(k))$ and we compare them with the corresponding functions $q^*(\xi|[x-ml,x])$ and $g^*(\xi|[x-ml,x])$, where $x<z$ is a left-barrier point. For this define

$$q_x^*(\xi) := q^*(\xi|[x-ml,x]), \quad g_x^*(\xi) := g^*(\xi|[x-ml,x]), \quad \mathcal{I}^*(\xi|[x-ml,x]) := \mathcal{I}(\xi|[x-ml,x])^-).$$

The counterpart of $E^1_{\text{mistake-r}}$ is as follows

$$E^1_{\text{mistake-l}}(z,x,k) := \Big\{\hat{q}^*(w_z^1(k)) \leq q_x^*(\xi)\Big\} \cap \Big\{\hat{g}^*(w_z^1(k)) \sqsubseteq_{\mathcal{I}^*(\xi|[x-ml,x])} g_x^*(\xi)\Big\} \cap \Big\{x \text{ is a left-barrier point}\Big\}.$$

Finally, let

$$E^1_{\text{mistake-l}} := \bigcup E^1_{\text{mistake-l}}(z,x,k),$$

where the union is taken over all z,x,k such that $x<z$, $z,x \in I_1$ and $k \leq \exp(\alpha l_1)$.

Finally, let

$$E^1_{\text{no mistake}} := \Big(E^1_{\text{mistake-l}} \cap E^1_{\text{mistake-r}}\Big)^c.$$

3.2.3 Attributes

* Define

$$T^1(j) := \tau(j) + \exp(2l_1) + lm^2, \quad T^3(j) := \tau(j) + \exp(2l_1) + lm^2 + c_1 l_1 = T^1(j) + c_1 l_1.$$

Hence $T^1(j)$ (or $T^3(j)$) is defined as $T^1_z(k)$ (or $T^1_z(k)$) by dropping the requirement that the random walk is at position z. We now define the counterparts of $w^1_z(k)$ and $w^3_z(k)$.

Let, for each $j = 1, 2, \ldots$

$$\begin{aligned} w^1(j) &:= \chi|[T^1(j) - lm^2, T^1(j)] \\ w^2(j) &:= \chi|[T^1(j), T^3(j)] \\ w^3(j) &:= \chi|[T^3(j), T^3(j) + lm^2]. \end{aligned}$$

Let $x, y \in I$ be such that $y - x = Lc_1 l_1$. We consider stopping times $\tau(1), \tau(2), \ldots, \tau(\exp(\alpha l_1))$. The following event states that among these stopping times there is at least $\exp(\gamma l_1)$ stopping times, $\tau(j)$ such that: $S(T^1(j)) = x,\ S(T^3(j)) = y$ and

$$\hat{q}^*(w^1(j)) \leq q^*_x(\xi),\ \hat{g}^*(w^1(j)) \sqsubseteq_{\mathcal{I}^*} g^*_x(\xi) \tag{3.5}$$

$$\hat{q}(w^3(j)) \leq q_y(\xi),\ \hat{g}(w^3(j)) \sqsubseteq_{\mathcal{I}} g_y(\xi), \tag{3.6}$$

where

$$\mathcal{I}^* := \mathcal{I}^*(\xi|[x - lm, x]), \quad \mathcal{I} := \mathcal{I}(\xi|[y, y + lm]).$$

More precisely,

$$E^1_{\text{enough times}}(x, y) := \left\{ \begin{array}{c} \text{there exists a set} \quad J(x,y) \subset [1, \exp(\alpha l_1)] \quad \text{such that} \\ |J(x,y)| \geq \exp(\gamma l_1) \quad \text{and for every} \quad j \in |J(x,y)| \\ S(T^1(j)) = x,\ S(T^3(j)) = y, \\ \hat{q}^*(w^1(j)) \leq q^*_x(\xi),\ \hat{g}^*(w^1(j)) \sqsubseteq_{\mathcal{I}^*} g^*_x(\xi), \\ \hat{q}(w^3(j)) \leq q_y(\xi),\ \hat{g}(w^3(j)) \sqsubseteq_{\mathcal{I}} g_y(\xi) \end{array} \right\}.$$

Finally, let

$$E^1_{\text{enough times}} := \bigcap_{x,y \in I_1,\ x-y = Lc_1 l_1} E^1_{\text{enough times}}(x, y).$$

* Note, for every $u \in \{0,1\}^{lm+1}$, $q(u) = (q_1, \ldots, q_l)$ is vector, such that $q_i = \{0, 1, \ldots, l\}$, $q_1 = 0$ and $q_i \leq q_{i+1} \leq q_i + 1$. Any such vector is called a **q-vector**. Hence, for every u, $q(u)$ and $q^*(u)$ are q-vectors.

Recall that, for any $u \in \{0,1\}^{lm+1}$, $g(u) = (g_1, \ldots, g_l)$, 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.

We say that a word $u \in \{0,1\}^{lm+1}$ is OK*, if u^- is OK. Clearly, if $\xi|[x - lm, x]$ is OK*, then $|\mathcal{I}^*| \geq l(1 - 3\epsilon(n))$. Similarly, if $\xi|[y, y + lm]$ is OK, then $|\mathcal{I}| \geq l(1 - 3\epsilon(n))$.

Note, if (3.5) and (3.6) hold and $\xi|[x - lm, x]$ together with $\xi|[y, y + lm]$ are OK* and OK, respectively, then there exists subsets $I^*, I \subset \{1, \ldots, l\}$, such that $|I^*|, |I| \geq l(1 - 3\epsilon(n))$, the q-vectors q, q^* and g-words g, g^* such that

$$\hat{q}^*(w^1(j)) \leq q^*, \quad \hat{g}^*(w^1(j)) \sqsubseteq_{I^*} g^*, \quad \hat{q}(w^3(j)) \leq q, \quad \hat{g}(w^3(j)) \sqsubseteq_I g. \tag{3.7}$$

We call (I^*, I, q^*, q, g^*, g) **a set of attributes**, if $I^*, I \subset \{1, \ldots, l\}$, $|I^*|, |I| \geq l(1 - 3\epsilon(n))$, q, q^* are q-vectors and g^*, g are g-words.

For every set of attributes (I^*, I, q^*, q, g^*, g) we define the index-set $J(I^*, I, q^*, q, g^*, g) \subset [1, \exp(\alpha l_1)]$ as follows: $j \in J(I^*, I, q^*, q, g^*, g)$ if and only if j satisfies (3.7).

Hence, if $E^1_{\text{enough times}}(x, y)$ holds and $\xi|[x - lm, x]$ together with $\xi|[y, y + lm]$ are OK, then there exists a set of attributes (I^*, I, q^*, q, g^*, g) such that

$$|J(I^*, I, q^*, q, g^*, g)| \geq \exp(\gamma l_1). \tag{3.8}$$

Then also $S(T^1(j)) = x$ and $S(T^3(j)) = y = x + Lc_1 l_1$, i.e. $S(T^3(j)) - S(T^1(j)) = Lc_1 l_1$.

* The following event implies: if $\exists J' \subset J(I^*, I, q^*, q, g^*, g)$ such that $|J'| \geq \exp(\gamma l_1)$ and $\forall j \in J'$ it holds $S(T^3(j)) - S(T^1(j)) < Lc_1 l_1$, then there exists at least two indexes $j^{'}, j^{''} \in J'$ such that $w^2(j^{'}) \neq w^2(j^{''})$.

Formally, we fix a set of attributes (I^*, I, q^*, q, g^*, g), we consider the stopping times $\tau(0), \tau(1), \dots$ and we define the indexes

$$j_1 := \min\{j \geq 0 : \tag{3.9}$$

$$\hat{q}^*(w^1(j)) \leq q^*,\ \hat{g}(w^1(j)) \sqsubseteq_{I^*} g^*,\ \hat{q}(w^3(j)) \leq q,\ \hat{g}(w^3(j)) \sqsubseteq_I g,\ |S(T^3(j)) - S(T^1(j))| < Lc_1 l_1\} \tag{3.10}$$

$$j_k := \min\{j > j_{k-1} : \tag{3.11}$$

$$\hat{q}^*(w^1(j)) \leq q^*,\ \hat{g}(w^1(j)) \sqsubseteq_{I^*} g^*,\ \hat{q}(w^3(j)) \leq q,\ \hat{g}(w^3(j)) \sqsubseteq_I g,\ |S(T^3(j)) - S(T^1(j))| < Lc_1 l_1\}. \tag{3.12}$$

Here the minimum over empty set is defined to be ∞. Let $\kappa := \max\{k : j_k < \infty\}$.

Clearly the subindexes $j_1, j_2, \dots$ depend on chosen attributes (I^*, I, q^*, q, g^*, g).

Recall $B^1_{\text{recon straight}}$. The following events are of similar nature. Let

$$\begin{aligned} E^1_{\text{recon straight}}(I^*, I, q^*, q, g^*, g) &:= \Big\{\kappa \geq \exp(\gamma l_1), \quad \exists k \leq \exp(\gamma l_1) \quad \text{such that} \quad w^2(j_1) \neq w^2(j_k)\Big\} \cup \Big\{\kappa \leq \exp(\gamma l_1)\Big\} \\ E^1_{\text{recon straight}} &:= \bigcap_{I^*, I, q^*, q, g^*, g} E^1_{\text{recon straight}}(I^*, I, q^*, q, g^*, g), \end{aligned}$$

where the intersection is taken over all sets of attributes.

3.3 Algorithm

We are ready to give the precise definition of the algorithm $\mathcal{A}^1$. The input of $\mathcal{A}^1$ consists of three ingredients

- $\exp(12\alpha l_1) + 1$ observations, $\chi|[0, \exp(12\alpha l_1)]$;
- $\mathcal{F}$-adapted stopping times $\tau = (\tau(1), \dots, \tau(\exp(\alpha l_1)) \subset [0, \exp(12\alpha l_1)]$ satisfying (3.3);
- a piece of original scenery $\psi^o = \xi|I^o$, where $|I^o| \geq 2c_1 L l_1$ and $I^o \subset [-\exp(l_1), \exp(l_1)]$.

The output of $\mathcal{A}^1$ is a piece of scenery of length $4\exp(l_1)$. Thus, formally,

$$\mathcal{A}^1 : \{0,1\}^{[0, \exp(12\alpha l_1)]} \times [0, \exp(12\alpha l_1)]^{[1, \exp(\alpha l_1)]} \times \Big(\cup_{k > 2c_1 l_1 L} \{0,1\}^k \Big) \mapsto \{0,1\}^{[-2\exp(l_1), 2\exp(l_1)]}.$$

The aim of $\mathcal{A}^1$ is to produce a piece of original scenery that lies between $\xi|[-\exp(l_1), \exp(l_1)]$ and $\xi|[-4\exp(l_1), 4\exp(l_1)]$. Thus, $\mathcal{A}^1$ does well, if the following event holds

$$E^1_{\text{alg works}}(\tau, I^o) := \Big\{\xi|[-\exp(l_1), \exp(l_1)] \sqsubseteq \mathcal{A}^1\big(\chi^{\exp(12\alpha l_1)}, \tau, \xi|I^o\big) \sqsubseteq \xi|[-3\exp(l_1), 3\exp(l_1)]\Big\}. \quad (3.13)$$

Obviously the event (3.13) 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

$$E^1_{\text{alg works}}(\tau) := \bigcap_{I^o \subset [-\exp(l_1), \exp(l_1)]} E^1_{\text{alg works}}(\tau, I^o).$$

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 $(w^1(j), w^2(j), w^3(j))$, defined by $\tau(j)$, are used. The word $w^2(j)$ will be collected as a ladder word, if it passes certain selection procedure. We shall specify the selection rule below, this is the core of $\mathcal{A}^1$. The set of collected works, i.e. the set of all words, that pass the selection rule, will be 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.

We now give the description and the definition of the selection rule for **Phase I** and the precise definition of assembling rule for **Phase II**. These definitions complete the definition of $\mathcal{A}^1$.

The selection rule is the most crucial part of the whole scenery reconstruction. The selection rule 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

$$E^1_{\text{only ladders}} := \{\forall w \in \mathcal{W}^1 \text{ there exists } I \in \mathcal{L}(c_1 l_1) \text{ such that } I \subset I_1 \text{ and } \xi|I \approx_l w\}.$$

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

$$E^1_{\text{enough ladders}} := \left\{\begin{array}{c} \text{for any } j = 0, \ldots, L-1 \text{ and for any } z \in I_1, \\ \text{there exists}\, x \in [z, z + (c_1 l_1/4)L] \text{ such that:} \\ x \text{mod} L = j \text{ and } (\xi(x), \xi(x+L), \ldots, \xi(x + (c_1 l_1)L)) \in \mathcal{W}^1 \end{array}\right\}.$$

Let us briefly introduce the main ideas behind the selection rule. The construction of the selection rule used for $\mathcal{A}^1$ starts from the fact that, with high probability, the events $E^1_{\text{enough times}}$ and

$$B^1_{\text{intervals OK}} := \{\xi|[z, z+ml] \text{ is OK } \ \forall z \in I_1\} \cap \{\xi|[z, z-ml] \text{ is OK}^* \ \forall z \in I_1\}$$

both hold. This means that for each $x, y \in I_1$, $y - x = Lc_1 l_1$ 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 $j \in J(x, y)$, (3.7) hold and the word $w^2(j)$ is the same, say w. This yields the first requirement of selection rule – check whether there exists (I^*, I, q^*, q, g^*, g) that satisfies the following condition: $\exists J' \subset J(I^*, I, q^*, q, g^*, g)$ such that $|J'| \geq \exp(\gamma l_1)$ 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 is based on the fact that, with high probability the events $B^1_{\text{enough barriers}}$ and $E^1_{\text{no mistake}}$ hold. This means that if (x,y) form a barrier then 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 $\mathcal{I}^*(\xi|[x-lm,x]), \mathcal{I}(\xi[y,y+lm])$ $q^*_x(\xi)$, $q_y(\xi)$, $g_x(\xi)$, $g_y(\xi)$ 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 (3.7), then $w^2(j) = w$. Thus, the set J' must actually be $J(I^*, I, q^*, q, g^*, g)$. This is the second requirement of the selection rule.
From the argument above, it is clear that if $E^1_{\text{enough times}}$, $B^1_{\text{intervals OK}}$, and $E^1_{\text{no mistake}}$ hold, then the selection rule will select all ladder words $(\xi(x), \xi(x+L), \ldots, \xi(x+L(c_1-1)L), \xi(y))$, where (x,y) is barrier of ξ and $y - x = Lc_1l_1$. With $B^1_{\text{enough barriers}}$ the latter yields $E^1_{\text{enough ladders}}$ (see Lemma 3.1 for formal proof). If, in addition $E^1_{\text{recon straight}}$ holds, then, as it is not hard to see, the selection procedure will select only ladder words (Lemma 3.1). Hence, the selection rule consisting of two requirements described above is sufficient for our purposes. We now give the formal definition of the selection rule.

Definition 3.1 *We define the set $\mathcal{W}^1 = \mathcal{W}^1(\chi^{12\alpha l_1}, \tau)$ as follows. A word $w \in \{0,1\}^{c_1l_1+1}$ belongs to $\mathcal{W}^1$ if and only if there exists a complect of attributes (I^*, I, q^*, q, g^*, g) such that the following conditions are satisfied:*

1. $|J(I^*, I, q^*, q, g^*, g)| \geq \exp(\gamma l_1)$

2. *if $j \in J(I^*, I, q^*, q, g^*, g)$, then $w^2(j) = w$.*

Let us now formalize the Phase II.
For a ladder interval I and a set $D \subset \mathbb{Z}$ we write $|I \cap D| \geq r$ if there exists a ladder interval $J \in \mathcal{L}(r)$ such that $J \subset D \cap I$. Recall that two pieces of scenery ϕ and ϕ' are strongly equivalent, $\phi \equiv \phi'$, 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 $I^o \subset [-\exp(l_1), \exp(l_1)]$.

Definition 3.2 *We say that the piece of scenery $\phi \in \{0,1\}^{[-2\exp(l_1), 2\exp(l_1)]}$ is a solution, formally $\phi \in \mathcal{S}(\chi^{12\alpha l_1}, \tau, \psi^o)$, if and only if there exist $\phi_i \in \{0,1\}^{D_i}$, $i = 1, 2, \ldots, n$ such that $D_i \subset [-3\exp(l_1), 3\exp(l_1)]$ and the following conditions are satisfied:*

1. $D_1 = [0,k]$, $\phi_1 \equiv \psi^o$;

2. *for each $i = 2, \ldots, n$ it holds $\phi_i | D_{i-1} = \phi_{i-1}$;*

3. *for each $i = 2, \ldots, n$ there exists $I_i \in \mathcal{L}(c_1 l_1)$ such that*

- **3a)** $D_i = D_{i-1} \cup V_i$;
- **3b)** $|D_i \cap V_i| \geq \frac{c_1 l_1}{4}$;
- **3c)** $\exists w_i \in \mathcal{W}^1(\chi^{12\alpha l_1}, \tau)$ *such that* $\phi_i | V_i \approx_l w_i$;

4. $[-2\exp(l_1), 2\exp(l_1)] \subset D_n, \quad \phi = \phi_n|[-2\exp(l_1), 2\exp(l_1)]$.

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 3.3 *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,* $\mathcal{A}^1(\chi^{12\alpha l_1}, \tau, \psi^o) := (1)_{[-2\exp(l_1), 2\exp(l_1)]}$.

3.4 Combinatorics for main theorem

Recall the stopping times $\tau = (\tau(0), \ldots, \tau(\exp(\alpha l_1)))$. The aim of the stopping times is to stop the random walk S near the origin. It is enough, if $S(\tau(k)) \in [-\exp(l_1), \exp(l_1)]$. Thus, the stopping times do well, if the following event holds

$$E^1_{\text{stop}}(\tau) := \{|S(\tau(j))| \leq \exp(l_1), \quad j = 0, 1, 2, \ldots, \exp(\alpha l_1)\}.$$

Roughly speaking, the main theorem of the paper states that the algorithm $\mathcal{A}^1$ reconstructs correctly with high probability, provided the stopping times τ indeed stop the random walk close to the origin.

Theorem 3.4 *There exists $a(l_2, n, c_1) > 0$ such that, for l_1 big enough*

$$P\Big(E^1_{stop}(\tau) \cap \big(E^1_{alg\ works}(\tau)\big)^c\Big) \leq e^{-al_1}. \tag{3.14}$$

The rest of the paper is the proof of Theorem 1.1. At first we prove some inclusions.

Lemma 3.1 *The following inclusions hold*

$$E^1_{recon\ straight} \cap E^1_{stop}(\tau) \subset E^1_{only\ ladders}; \tag{3.15}$$
$$B^1_{intervals\ OK} \cap E^1_{stop}(\tau) \cap E^1_{no\ mistake} \cap E^1_{enough\ times} \cap E^1_{enough\ barriers} \subset E^1_{enough\ ladders}; \tag{3.16}$$
$$E^1_{only\ ladders} \cap E^1_{enough\ ladders} \cap B^1_{unique\ fit} \subset E^1_{alg\ works}(\tau), \tag{3.17}$$

provided l_1 is big enough.

Proof. At first note: if $E^1_{\text{stop}}(\tau)$ holds, then, for each $j = 1, 2, \ldots, \exp(\alpha l_1)$, it holds

$$|S(T^3(j))| \leq |S(\tau(j)) + L(\exp(2l_1) + lm^2 + c_1 l_1)| \leq \exp(3l_1), \tag{3.18}$$

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, $j = 1, 2, \ldots, \exp(\alpha l_1)$.

Proof of (3.15):
We prove

$$\big(E^1_{\text{only ladders}}\big)^c \cap E^1_{\text{stop}}(\tau) \subset \big(E^1_{\text{recon straight}}\big)^c. \tag{3.19}$$

Suppose $\big(E^1_{\text{only ladders}}\big)^c \cap E^1_{\text{stop}}(\tau)$ holds. Then there exists a $w \in \mathcal{W}^1$ that is not a ladder word of any ladder piece $\xi|I$ of length $l_1 c_1$ such that $I \subset I_1$. 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 3.1 hold. This means, that

$$|S(T^3(j)) - S(T^1(j))| < c_1 l_1, \quad \forall j \in J(I^*, I, q^*, q, g^*, g). \tag{3.20}$$

Indeed, if there were an index $j^* \in J(I^*, I, q^*, q, g^*, g)$ such that (3.20) 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 $I \subset I_1$. This contradicts our assumption on w.
Recall the definition of κ. Since $|J(I^*, I, q^*, q, g^*, g)| \geq \exp[\gamma l_1]$, we have $\kappa \geq \exp[\gamma l_1]$. On the other hand, by **2** of Definition 3.1, for each j_k, $k = 1, 2, \ldots, \exp[\gamma l_1]$, 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 (3.19).

Proof of (3.16):
Let $x, y \in I_1$ and $y - x = c_1 l_1$. Since $B^1_{\text{intervals OK}}$ holds then, by (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 $j \in J(x, y)$, (3.7) holds. Moreover, $|J(x, y)| \geq \exp(\gamma l_1)$ 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 $J' = J(x,y) \subset J(I^*, I, q^*, q, g^*, g)$ such that $|J'| \geq \exp(\gamma l_1)$ and $w^2(j)$ is constant on J'. Assume (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 (3.7), 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 $S(T^1(j^*)) < y$. Let $S(T^1(j^*) = z > x$. The event $E^1_{\text{stop}}(\tau)$ implies (3.18) 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 it remains to show that there are enough barriers in I_1. This follows immediately from $B^1_{\text{enough barriers}}$.

Proof of (3.17):
It suffices to show that $E^1_{\text{only ladders}} \cap E^1_{\text{enough ladders}} \cap B^1_{\text{unique fit}}$ ensures that $\mathcal{S}(\chi^{12\alpha l_1}, \tau, \xi|I^o)$ consists of one element that satisfies (3.13).
Consider the "puzzle-playing" algorithm formalized in Definition 3.2. 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 $\exists \phi \in \mathcal{S}$ and let $D_1 \subset D_2 \subset \cdots \subset D_n$ be the sequence of sets ensured by the definition of ϕ. By **1**, $\phi|D_1$ is translated from a piece of $\xi|I_1$ by some b satisfying $|b| \leq \exp(l_1)$, i.e. $\xi|I^o = T[\phi|D_1]$, where $Tz = z+b$ is the translation and $I_o \subset [-e^{l_1}, e^{l_1}] \subset I_1$. We show: if $\phi|D_i$ is translated from a piece of $\xi|I_1$ by b, i.e. $\xi|J_i = T[\phi|D_i]$, for some $J_i \subset I_1$, then the same applies for $\phi|D_{i+1}$. Recall that $\phi|D_{i+1}$ and $\phi|D_i$ differ on V_{i+1}, only. By **3c)** and $E^1_{\text{only ladders}}$, $\phi|V_{i+1} \approx \xi|J(w)$ for some $J(w) \subset I_1$. Thus, there is an affine T' such that $\xi|J(w) = T'[\phi|V_{i+1}]$ and, hence, there is a ladder interval $J' \subset J(w)$ such that $\xi|J' = T'[\phi|(V_{i+1} \cap D_i)]$. So, $\phi|(V_{i+1} \cap D_i)$ is equivalent with some ladder word of $\xi|I_1$ by T'. On the other hand, $\phi|(V_{i+1} \cap D_i)$ 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 $\phi|V_{i+1}$ is translated from a piece of $\xi|I_1$ by b and $\phi|D_{i+1}$ is translated from a piece of $\xi|I_1$ by b as well. The same holds for ϕ, i.e. $\phi \equiv \xi|I(\phi)$ for some interval $I(\phi)$. By **4**, $I(\phi) = [a_o - 2\exp(l_1), a_o + 2\exp(l_1)]$, where $I_o := [a_o, b_o]$. So, ϕ is obtained from a fixed piece of scenery $\xi|I(\phi)$ by a fixed translation, T. Clearly such a ϕ is unique.
Let us show that ϕ satisfies (3.13). Since $|a_o| \leq \exp(l_1)$, we have that

$$[-\exp(l_1), \exp(l_1)] \subset I(\phi) \subset [-3\exp(l_1), 3\exp(l_1)].$$

This means

$$\xi|[-\exp(l_1), \exp(l_1)] \sqsubseteq \phi \sqsubseteq \xi|[-3\exp(l_1), 3\exp(l_1)],$$

i.e. (3.13) holds.
It remains to show that $\mathcal{S}$ is not empty. Consider again the "puzzle playing" algorithm. Let D_i, $i \geq 1$ be the domain of ϕ_i at state i. It suffices to show that there exists $V_{i+1} \in \mathcal{L}(c_1 l_1)$ satisfying all requirements of **3** and such that $|V_{i+1}\backslash D_i| \geq \frac{c_1 l_1}{2}$. Note that $D_i = \cup_{j=0}^{L-1} I(j)$, where $I(j) \subset I_1$ is a ladder interval with length at least $c_1 l_1$. Fix j and let $a_j < b_j$ be the endpoints of $I(j)$. Consider the ladder interval $I^b(j) := \big(b_j - 2(c_1 l_1/4)L, \ldots, b_j - 1(c_1 l_1/4)L\big) \subset I(j)$. By $E^1_{\text{enough ladders}}$ there exists $z \in I^b(j)$ such that a ladder word of $\xi|V(z)$, with $V(z) = \big(z, z+L, \ldots, z+(c_1 l_1)L\big) \in \mathcal{L}(c_1 l_1)$ belongs to $\mathcal{W}^1$. Let this word be $w(z)$. Clearly, $|V(z) \cap D_i| \geq \frac{c_1 l_1}{4}$. By $B^1_{\text{unique fit}}$, $w(z)$ is not a ladder word of any ladder piece $\phi_i|V_j$, $j = 1, \ldots, i$. Hence $w(z)$ and $V(z)$ can be taken as w_{i+1} and V_{i+1}. The same argument applies for $I^a(j) := \big(a_j + 1(c_1 l_1/4)L, \ldots, a_j + 2(c_1 l_1/4)L\big)$, implying that D_i can be efficiently enlarged in other direction as well. ■

3.5 Probabilities for main theorem

3.5.1 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 $P\Big(\big(B^1_{\text{intervals OK}}\big)^c\Big)$

Let

$$E := \{\xi|[z, z+ml] \text{ is OK } \forall z \in I_1\}, \quad E^* := \{\xi|[z, z-ml] \text{ is OK}^* \ \forall z \in I_1\}.$$

Now, by translation invariancy of ξ and Theorem 2.1, it holds that for l_1 big enough

$$P(E^c) \leq \sum_{z\in I_1} P(\xi|[z, z+ml] \text{ is not OK}) \leq 2e^{3l_1} P(E^c_{\text{OK}}) \leq 2\exp[3l_1 - al].$$

Similarly,

$$P(E^{*c}) \leq \sum_{z\in I_1} P(\xi|[z-ml, z] \text{ is not OK}^*) \leq 2e^{3l_1} P(E^c_{\text{OK}}) \leq 2\exp[3l_1 - al].$$

Hence, if l_1 is sufficiently big, then

$$P\Big(\big(B^1_{\text{intervals OK}}\big)^c\Big) \leq 4\exp[(3 - al_2)l_1]. \tag{3.21}$$

The following proposition also specifies the choice of c_1.

Proposition 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:*

$$P\Big(\big(B^1_{\text{unique fit}}\big)^c\Big) \leq \exp[-k_1 l_1] \tag{3.22}$$

$$P\Big(\big(B^1_{\text{recon straight}}\big)^c\Big) \leq \exp[-k_2 l_1] \tag{3.23}$$

$$P\Big(\big(B^1_{\text{enough barriers}}\big)^c\Big) \leq \exp[-k_3 l_1], \tag{3.24}$$

provided l_1 is big enough.

Proof. It follows from Lemma 6.33 in [LMM] that for some constants a_1, a_2 depending on L, only, the bound $P\Big(\big(B^1_{\text{unique fit}}\big)^c\Big)a_1 \leq \exp[-a_2 l_1]$ is valid. Also, there is a fixed constant C_r such that $a_2 > 0$ if $c_1 > C_r$. This implies (3.22) 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 $C_1 \geq 1$) $\mathcal{R}(l_1c_1)(x,y)$ is either empty or has cardinality at least 2. Any admissible path $R \in \mathcal{R}(l_1c_1)(x,y)$ is a sequence $R = (t_1, \ldots, t_{c_1l_1})$ of steps, where $|t_i| \leq L$. Hence, there exists a $R = (t_1, \ldots, t_{c_1l_1}) \in \mathcal{R}(l_1c_1)(x,y)$ such that $t_i \neq t_1$ for a $i = 2, \ldots c_1l_1$ (if no, then $\mathcal{R}(l_1c_1)(x,y)$ would consists of one path, only). Let R be one of such paths. Let $c_1 \geq \ulcorner \frac{100}{2L+1} \urcorner$. The number of possible steps is bounded by $2L+1$. Hence, there is a step t that occurs in R at least $k := 100l_1$ times. Formally, $\exists t \in \{-L, \ldots, L\}$ such that $|\{i = 1, \ldots, c_1l_1 : t_i = t\}| \geq k$. 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 kl_1 steps of size t. Thus $R^1 = \{t^1_1, \ldots, t^1_{c_1l_1}\} \in \mathcal{R}(l_1c_1)(x,y)$ 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, $R^2 = \{t^2_1, \ldots, t^2_{c_1l_1}\} \in \mathcal{R}(l_1c_1)(x,y)$ 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

$$\begin{aligned}
&P(\xi\circ R^1=\xi\circ R^2)\leq P\Big(\big(\xi(x+t),\dots,\xi(x+kt)\big)=\big(\xi(x+u),\xi(x+u+t),\dots,\xi(x+u+kt)\big)\Big)\\
&\leq P\big(\xi(x+t)=\xi(x+u)\big)P\big(\xi(x+2t)=\xi(x+u+t)|\xi(x+t)=\xi(x+u)\big)\times\\
&\times P\big(\xi(x+3t)=\xi(x+u+2t)|\xi(x+t)=\xi(x+u)),\xi(x+2t)=\xi(x+u+t)\big)\cdots\\
&\cdots P\big(\xi(x+kt)=\xi(x+u+(k-1)t)|\xi(x+t)=\xi(x+u)\cdots\xi(x+(k-1)t)=\xi(x+u+(k-2)t)\big)\\
&\leq 2^{-k}=\exp[-100\ln 2l_1].
\end{aligned}$$

Now,

$$\begin{aligned}
E_{\text{recon straight}}&=\bigcup_{x,y\in I_1,|x-y|<l_1c_1}E_{\text{recon straight}}(x,y),\\
P((E_{\text{recon straight}})^c)&\leq\sum_{x,y\in I_1}P(E_{\text{recon straight}}(x,y))\leq 4\exp(6l_1)\exp[-100\ln 2l_1]\leq\exp[-50l_1].
\end{aligned}$$

<u>Estimate $P\big((B^1_{\text{enough barriers}})^c\big)$</u>

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 $x'-x\geq 3m^3L=:r$, 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 $w=\llcorner\frac{c_1l_1}{4r}-\frac{L}{r}\lrcorner>\frac{c_1-4L}{4r}l_1$ and use Höffding's inequality again

$$\begin{aligned}
P\Big(\big(B^1_{\text{enough barriers}}(z,j)\big)^c\Big)&=P\Big(\sum_{x=z+j}^{\frac{z+c_1l_1}{4}}Y_x=0\Big)\leq P\Big(\sum_{k=0}^{\frac{c_1l_1}{4r}}Y_{rk+z+j}=0\Big)\\
&\leq P\Big(\sum_{k=0}^{w}(Y_{rk+z+j}-p)\leq wp\Big)\leq 2\exp[-2wp^2]\\
&\leq 2\exp[-2\frac{c_1-4L}{4r}2^{-6m^3L}l_1]=2\exp[-k_2'l_1],
\end{aligned}$$

for $k_2':=\frac{c_1-4L}{4r}2^{-(6m^3L+1)}$. Obviously, $k_2'>0$, if $c_1>4L$. Thus

$$P\Big(\big(B^1_{\text{enough barriers}}\big)^c\Big)\leq\sum_{z\in I_1,j=\{0,\dots,L\}}P\Big(\big(B^1_{\text{enough barriers}}(z,j)\big)^c\Big)\leq 8\exp[(6-k_2')l_1]\leq\exp[-l_1],$$

if $k_2'\geq 8$. The latter implies $c_1-4L\geq 4\cdot 2^{6m^3+6}$ or $c_1\geq r2^{6m^3+8}+4L=3m^3L2^{6m^3+8}+4L$.

Hence, Proposition 3.1 holds with $C_1(n):=\max\{C_r,\ulcorner\frac{100}{2L+1}\urcorner,3m^3L2^{6m^3+8}+4L\}$. ■

3.5.2 Random-walk depending events

<u>Estimate $P(E^1_{\text{mistake-r}}\cap B^1_{\text{intervals OK}})$</u>.

Fix $y,z\in I_1$, $z<y$ and note

$$E^1_{\text{mistake-r}}(z,y,k)\cap B^1_{\text{intervals OK}}\subset E^1_{\text{mistake-r}}(z,y,k)\cap\{\xi|[y,y+lm] \text{ is OK}\},\quad k=1,2,\dots. \tag{3.25}$$

We now estimate the right side of (3.25). Recall the definitions of $T_z^3(k), w_z^3(k)$ and $g_y(\xi)$. Consider the events

$$E^1_{\text{mistake-r}}(y,z,k)\cap\{\xi|[y,y+lm]\text{is OK}\}=$$
$$\Big\{\hat q(w_z^3(k))\le q_y(\xi),\ \hat g(w_z^3(k))\sqsubseteq_{\mathcal{I}(\xi|[y,y+ml])} g_y(\xi),\ y\text{ is a right barrier point},\ \xi|[y,y+lm]\text{ is OK}\},\tag{3.26}$$

$k=1,2,\ldots$. Because of (3.3), conditionally on ξ the events (3.26) are independent and identically distributed. Hence, the events (3.26) all have the probability equal to

$$P\Big(\hat q(\chi_z^{m^2l})\le q_y(\xi),\quad \hat g(\chi_z^{m^2l})\preccurlyeq_{\mathcal{I}(\xi|[y,y+ml])} g_y(\xi),\quad y\text{ is a right barrier point},\quad \xi|[y,y+lm]\text{ is OK}\Big).\tag{3.27}$$

The event in (3.27) depends on ξ, only. The distribution of ξ is obviously translation invariant. Therefore, by Corollary 2.1, (3.27) can be estimated

$$P\Big(\hat q(\chi_{z-y}^{m^2l})\le q_0(\xi)),\quad \hat g(\chi_{z-y}^{m^2l})\preccurlyeq_{\mathcal{I}(\xi^{ml})} g_0(\xi),\quad 0\text{ is a right barrier point },\quad \xi^{ml}\text{ is OK}\Big)=$$
$$P\Big(\Big\{\hat q(\chi_{z-y}^{m^2l})\le q_0(\xi),\quad \hat g(\chi_{z-y}^{m^2l})\preccurlyeq_{\mathcal{I}(\xi^{ml})} g_0(\xi)\Big\}\cap E_{\text{origin}}\cap E^*_{OK}\Big)=$$
$$P\Big(E_{\text{mistake}}(z-y)\cap E^*_{OK}\Big)\le\exp(-l\alpha_I),$$

provided l_1 is big enough. Therefore,

$$P(E^1_{\text{mistake-r}}\cap B^1_{\text{intervals OK}})\le\sum_{y,z,k}P(E^1_{\text{mistake-r}}(y,z,k)\cap B^1_{\text{intervals OK}})$$
$$\le\sum_{y,z,k}\exp(-l\alpha_I)<4\exp[(6+\alpha)l_1-\alpha_I l].\tag{3.28}$$

The sum here is taken over all $z,y\in I_1$, $z<y$ and $k=1,\ldots,\exp(\alpha l_1)$.

Estimate $P(E^1_{\text{mistake-l}}\cap B^1_{\text{intervals OK}})$.

We need some additional notations. Recall $T_z^1(k)$. Now fix $x'\in I_1$ and define $T_z^1(k_i)$, $i=1,2,\ldots,N(x')$. as the i-th stopping time $T_z^1(k)$, for which $S(T_z^1(k)+\exp(2l_1))=x'$. The indexes k_i depend on chosen x'. Define now

$$E^1_{\text{mistake-l}}(z,x,i,x'):=\Big\{\hat q^*(w_z^1(k_i))\le q_x^*(\xi)\Big\}\cap\Big\{\hat g^*(w_z^1(k_i))\preccurlyeq_{\mathcal{I}^*(\xi|[x-ml,x])} g_x^*(\xi)\Big\}\cap\Big\{x\text{ is a left b. p.}\Big\},$$

$i=1,2,\ldots,N(x')$.
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 (3.25) is

$$E^1_{\text{mistake-l}}(z,x,i,x')\cap B^1_{\text{intervals OK}}\subset E^1_{\text{mistake-r}}(z,x,i,x')\cap\{\xi|[x-lm,x]\text{ is OK*}\}=:E(i,x'),$$

$i=1,2,\ldots,N(x')$.
As previously, we observe that $P(E(i,x'))$ is equal to

$$P\Big(\hat q^*(\chi_{x'}^{m^2l})\le q_x^*(\xi),\ \hat g^*(\chi_{x'}^{m^2l})\preccurlyeq_{\mathcal{I}^*(\xi|[x-ml,x])} g_x^*(\xi),\ S_{x'}(m^2l)=z,\ x\text{ is a left b. p. },\ \xi|[x-lm,x]\text{ is OK*}\Big).\tag{3.29}$$

To calculate (3.29), at first note the following. Let $R(i)$, $i=0,1,\ldots,k$ be an admissible path such that $R(0)=x'$, $R(k)=z$. Thus, for any scenery ψ, the observation $\chi|[0,k]$ equals $\psi(R(i))$, $i=0,\ldots,k$. This means, $\big(\chi|[0,k]\big)^-=\psi(R^-(i))$, where $R^-(i)=-R(k-i)$, $i=0,\ldots,k$. 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 with $P_\psi(\cdot) := P(\cdot|\xi = \psi)$,

$$P_\psi\Big(\big(\chi|[0,k]\big)^- = u,\ S_{x'}(k) = z\Big) = P_\psi\Big(\chi|[0,k] = u,\ S_z(k) = x',\Big)$$

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 left side of last equality equals

$$P_{-\psi}\Big(\chi^k_{-z} = u,\ S_{-z}(k) = -x'\Big).$$

In particular, since $(\psi|[x-lm,x])^- = -\psi|[-x,-x+lm]$

$$P_\psi\Big(\hat q\big((\chi^k_{x'})^-\big) \le q\big((\psi|[x-lm,x])^-\big),\ \hat g\big((\chi^k_{x'})^-\big) \preccurlyeq_{\mathcal{I}((\psi|[x-lm,x])^-)} g\big((\psi|[x-lm,x])^-\big),\ S_{x'}(k) = z\Big) =$$
$$P_{-\psi}\Big(\hat q\big(\chi^k_{-z}\big) \le q(-\psi|[-x,-x+lm]), \hat g\big((\chi^k_{-z})\big) \preccurlyeq_{\mathcal{I}(-\psi|[-x+lm,-x])} g(-\psi|[-x+lm,-x]),\ S_{-z}(k) = -x'\Big).$$

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 $-\psi$ and, by definition, $\psi|[x-lm,x]$ is OK* if and only $(\psi|[x-lm,x])^- = -\psi|[-x,-x+lm]$ is OK. Let

$$A^*(x) := \{x \text{ is a left b. p. of } \psi, \psi|[x-lm,x] \text{ is OK*}\},$$
$$A(x) := \{x \text{ is a right b. p. of } \psi,\ \psi|[x,x+lm] \text{ is OK }\}.$$

Thus, for each ψ,

$$P_\psi\Big(\hat q^*(\chi^k_{x'}) \le q^*(\psi|[x-lm,x]),\ \hat g^*(\chi^k_{x'}) \preccurlyeq_{\mathcal{I}^*(\psi|[x-lm,x])} g^*(\psi|[x-lm,x]),\ S_{x'}(k) = z\Big) I_{A^*(x)}(\psi) =$$
$$P_{-\psi}\Big(\hat q\big(\chi^k_{-z}\big) \le q(-\psi|[-x,-x+lm]), \hat g(\chi^k_{-z}) \preccurlyeq_{\mathcal{I}(-\psi|[-x+lm,-x])} g(-\psi|[-x+lm,-x]),\ S_{-z}(k) = -x'\Big)\times$$
$$\times I_{A(-x)}(-\psi).$$

Finally, integrate over ξ and use the fact that ξ and $-\xi$ have the same distribution to get

$$P\Big(\hat q^*(\chi^k_{x'}) \le q^*(\xi|[x-lm,x]),\ \hat g^*(\chi^k_{x'}) \preccurlyeq_{\mathcal{I}^*(\xi|[x-lm,x])} g^*(\xi|[x-lm,x]),\ S_{x'}(k) = z, \xi \in A^*(x)\Big)=$$
$$P\Big(\hat q\big(\chi^k_{-z}\big) \le q(\xi|[-x,-x+lm]), \hat g(\chi^k_{-z}) \preccurlyeq_{\mathcal{I}(\xi|[-x+lm,-x])} g(\xi|[-x+lm,-x]),\ S_{-z}(k) = -x',\ \xi \in A(-x)\Big) \tag{3.30}$$

Now take $k = m^2 l$, $y := -x$, $z := -z$ and sum over x' to obtain that $P(E(i,x'))$ equals

$$P\Big(\hat q\big(\chi^{m^2 l}_z\big) \le q_y(\xi), \hat g\big(\chi^{m^2 l}_z\big) \preccurlyeq_{\mathcal{I}(\xi|[y+lm,y])} g_y(\xi),\ y \ \text{ is a right b. p., } \xi|[y,y+lm] \ \text{ is OK}\Big).$$

Hence, $P(E(i,x')$ equals (3.27) and, hence, it is bounded by $\exp(-l\alpha_I)$. This means

$$\begin{aligned} P(E^1_{\text{mistake-l}} \cap B^1_{\text{intervals OK}}) &\le \sum_{y,z,i,x'} P(E^1_{\text{mistake-r}}(y,z,i,x') \cap B^1_{\text{intervals OK}}) \\ &\le \sum_{y,z,i,x'} \exp(-l\alpha_I) < 8\exp[(9+\alpha)l_1 - \alpha_I l], \end{aligned} \tag{3.31}$$

where the sum is taken over all $z, y, x' \in I_1$, $z < y$ and $i = 1, \ldots, \exp(\alpha l_1)$.

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 $j_1, \ldots, j_\kappa$ as in (3.9) (3.12). 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 $k = 1, \ldots, \exp(\gamma l_1) + 1$, it holds $w^2(j_k) = w^2(j_1)$. Define

$$Y_k := 1 - I_{w^2(j_1)}(w^2(j_k)), \quad k = 2, \ldots, \kappa.$$

Hence $Y_k = 1$ if and only if $w^2(j_k) \neq w^2(j_1)$. Therefore, $E^c = \{\sum_{k=1}^{\exp(\gamma l_1)+1} Y_k = 0\}$. We now consider the following σ-algebra

$$\mathcal{A} := \sigma\Big(\xi(z),\ S(\tau(j)),\ S(T^1(j_k)),\ S(T^3(j_k)),\ z \in \mathbb{Z},\ j = 1, \ldots, \exp(\alpha l_1),\ k = 1, \ldots, \kappa\Big).$$

Given $\mathcal{A}$, the values of κ as well as $S(T^1(j_k)) = x_k$ and $S(T^3(j_k)) = y_k$, $k = 1, \ldots, \kappa$ are known. This means that the random variables $Y_1, \ldots, Y_\kappa$ 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 $Y_1, \ldots, Y_\kappa$ are independent.
Consider now the events $E^1_{\text{stop}}(\tau)$ and $B^1_{\text{recon straight}}$. Obviously they both belong to $\mathcal{A}$. Note that on $E^1_{\text{stop}}(\tau)$, it holds $x_k, y_k \in I_1$, for every $k = 1, \ldots, \kappa$. Hence, if in addition also $B^1_{\text{recon straight}}$ holds, then for each $k = 2, \ldots, \kappa$ 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 $P(Y_k = 1|\mathcal{A}) \geq (p_{min})^{c_1 l_1}$, $k = 2, \ldots, \kappa$. Hence, by Höffding's inequality on $E^1_{\text{stop}}(\tau) \cap B^1_{\text{recon straight}}$ we get for a constant $d > 0$

$$P(E^c|\mathcal{A}) = P\Big(\sum_{k=2}^{\exp(\gamma l_1)+1} Y_k = 0\Big|\mathcal{A}\Big) \leq \exp[-d\exp((\gamma + 2c_1 \ln p_{min}) l_1)]. \tag{3.32}$$

Indeed, for $Y_1, \ldots, Y_{e^b}$ independent Bernoulli random variables with $E(X_i) \geq e^a$, the Höffding's inequality states

$$P\Big(\sum_{i=1}^{e^b} Y_i = 0\Big) = P\Big(\sum_{i=1}^{e^b}(Y_i - EY_i) \leq -\sum_{i=1}^{e^b} EY_i\Big) \leq \exp\Big[-de^{-b}\big(\sum_{i=1}^{e^b} EY_i\big)^2\Big] \leq \exp[-de^{b+2a}]$$

Now take $b = \gamma l_1$, $a = c_1 l_1 \ln(p_{min})$ to obtain (3.32).
Integrate (3.32) over $E^1_{\text{stop}}(\tau) \cap B^1_{\text{recon straight}}$ to obtain

$$P\Big(E^c \cap E^1_{\text{stop}}(\tau) \cap B^1_{\text{recon straight}}\Big) \leq \exp[-d\exp((\gamma + 2c_1 \ln p_{min}) l_1)]. \tag{3.33}$$

Finally, estimate

$$P\Big(\big(E^1_{\text{recon straight}}\big)^c \cap E^1_{\text{stop}}(\tau) \cap B^1_{\text{recon straight}}\Big) \leq \sum_{(I^*, I, q^*, q, g^*, g)} E^1_{\text{recon straight}}(I^*, I, q^*, q, g^*, g),$$

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,

$$\begin{aligned}
&\big(E^1_{\text{stop}}(\tau) \cap B^1_{\text{recon straight}} \cap (E^1_{\text{recon straight}})^c\big) \leq \\
&\leq \exp[l_1(2(n^2 l_2 + l_2)\ln 2 + (4l_2)(\ln l_1 + \ln l_2)) - d\exp((\gamma + 2c_1 \ln p_{min}) l_1)].
\end{aligned} \tag{3.34}$$

Estimate $P\Big(E^1_{\text{stop}}(\tau)\cap(E^1_{enough\ times})^c\cap B^1_{intervals\ OK}\Big)$

Let $p_L := P(S(1)-S(0)=L)$ and define

$$p^* := \exp[-(1.5+2\alpha_{II}l_2+c_1\ln p_L)l_1].$$

Proposition 3.2 *If*

$$\exp(\alpha l_1)p^*\geq 2\exp(\gamma l_1), \tag{3.35}$$

then

$$P\Big(E^1_{\text{stop}}(\tau)\cap(E^1_{enough\ times})^c\cap B^1_{intervals\ OK}\Big)\leq 36\exp[(2-\exp(2\gamma-\alpha))l_1], \tag{3.36}$$

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 $x,y\in I_1$ be such that $y=x+c_1l_1L$ and define

$$E_j(x,y) := \left\{\begin{array}{c} S(T^1(j)-lm^2)=x-lm \\ S(T^1(j))=x,\ S(T^3(j))=y, \\ \hat q^*(w^1(j))\leq q^*_x(\xi),\ \hat g^*(w^1(j))\preccurlyeq_{\mathcal{I}^*} g^*_x(\xi), \\ \hat q(w^3(j))\leq q_y(\xi),\ \hat g(w^3(j))\preccurlyeq_{\mathcal{I}} g_y(\xi)\end{array}\right\},\quad Y_j := I_{E_j},\quad j=1,\dots,e^{\alpha l_1}.$$

Obviously,

$$\Big\{\sum_{j=1}^{e^{\alpha l_1}}Y_j\geq e^{\gamma l_1}\Big\}\subset E^1_{\text{enough times}}(x,y). \tag{3.37}$$

For each j and for every scenery ψ, it holds

$$\begin{aligned}&P_\psi(Y_j=1)=P_\psi(S(T^1(j)-lm^2)=x-lm)\times\\ &P_\psi(S(T^1(j))=x,\hat q^*(w^1(j))\leq q^*_x(\xi),\ \hat g^*(w^1(j))\preccurlyeq_{\mathcal{I}^*} g^*_x(\xi)|S(T^1(j)-lm^2)=x-lm^2)\times\\ &P_\psi(S(T^3(j))=y|S(T^1(j)-lm^2)=x-lm,S(T^1(j))=x,\hat q^*(w^1(j))\leq q^*_x(\xi),\ \hat g^*(w^1(j))\preccurlyeq_{\mathcal{I}^*} g^*_x(\xi))\times\\ &P_\psi(\hat q(w^3(j))\leq q_y(\xi),\ \hat g(w^3(j))\preccurlyeq_{\mathcal{I}} g_y(\xi)|\\ &S(T^1(j)-lm^2)=x-lm^2,\ S(T^1(j))=x,\ S(T^3(j))=y,\hat q^*(w^1(j))\leq q^*_x(\xi),\ \hat g^*(w^1(j))\preccurlyeq_{\mathcal{I}^*} g^*_x(\xi)).\end{aligned}$$

Now, by the Markov property of S

$$\begin{aligned}P_\psi(Y_j=1|E_{\text{stop}}(\tau))=&P_\psi(S(T^1(j)-lm^2)=x-lm|E_{\text{stop}}(\tau)))\\ &\times P_\psi(S(T^1(j))=x,\hat q^*(w^1(j))\leq q^*_x(\xi),\ \hat g^*(w^1(j))\preccurlyeq_{\mathcal{I}^*} g^*_x(\xi)|S(T^1(j)-lm^2)=x-lm^2)\\ &\times P_\psi(S(T^3(j))=y|S(T^1(j))=x)\\ &\times P_\psi(\hat q(w^3(j))\leq q_y(\xi),\ \hat g(w^3(j))\preccurlyeq_{\mathcal{I}} g_y(\xi)|S(T^3(j))=y).\end{aligned}$$

Use local CLT to estimate

$$\begin{aligned}P_\psi\Big(S(T^1(j)-lm^2)=x-lm\Big|E_{\text{stop}}(\tau)\Big)&=P_\psi\Big(S(T^1(j)-lm^2)=x-lm\Big|S(T^1(j)-lm^2-e^{2l_1}\Big)\\ &\geq\exp(-1.5l_1)\end{aligned}$$

By Theorem 2.3 and symmetry of S, it holds

$$\begin{aligned}&P_\psi(S(T^1(j))=x,\hat q^*(w^1(j))\leq q^*_x(\xi),\ \hat g^*(w^1(j))\preccurlyeq_{\mathcal{I}^*} g^*_x(\xi)|S(T^1(j)-lm^2)=x-lm^2)\geq\exp[-l\alpha_{II}]\\ &P_\psi(\hat q(w^3(j))\leq q_y(\xi),\ \hat g(w^3(j))\preccurlyeq_{\mathcal{I}} g_y(\xi)|S(T^3(j))=y)\geq\exp[-l\alpha_{II}],\end{aligned}$$

provided $\psi \in B^1_{\text{intervals OK}}$.
Finally,

$$P_\psi(S(T^3(j)) = y | S(T^1(j)) = x) = (p_L)^{c_1 l_1 L}.$$

This means, for $\psi \in B^1_{\text{intervals OK}}$

$$P_\psi(Y_j = 1 | E_{\text{stop}}(\tau)) \geq \exp[-1.5 l_1] \exp[-2l\alpha_{II}](p_L)^{c_1 l_1 L} = p^* \tag{3.38}$$

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

$$P_\psi\Big(\sum_{j=1}^{e^{\alpha l_1}} Y_j < e^{\gamma l_1} \big| E_{\text{stop}}(\tau)\Big) \leq P\Big(\sum_{j=1}^{e^{\alpha l_1}} Z_j < e^{\gamma l_1}\Big) = P\Big(\sum_{j=1}^{e^{\alpha l_1}} (Z_j - p^*) < e^{\gamma l_1} - e^{\alpha l_1} p^*\Big), \tag{3.39}$$

where Z_i are independent Bernoulli random variables with parameter p^*. By (3.35), the right side of (3.39) is bounded by

$$P\Big(\sum_{j=1}^{e^{\alpha l_1}} (Z_j - p^*) < e^{\gamma l_1} - e^{\alpha l_1} p^*\Big) \leq P\Big(\sum_{j=1}^{e^{\alpha l_1}} (Z_j - p^*) < -e^{\gamma l_1} p^*\Big).$$

Use Höffding's inequality to get

$$P\Big(\sum_{j=1}^{e^{\alpha l_1}} (Z_j - p^*) < -e^{\gamma l_1}\Big) \leq \exp[-d e^{(2\gamma - \alpha) l_1}].$$

Finally, integrate over $E^1_{\text{stop}}(\tau) \cap B^1_{intervals\ OK}$ and use (3.37) to deduce

$$P\Big(E^1_{\text{stop}}(\tau) \cap (E^1_{\text{enough times}}(x, y))^c \cap B^1_{\text{intervals OK}}\Big) \leq \exp[-\exp(2\gamma - \alpha)) l_1].$$

Sum over all pairs $(x, y) \in I_1$ to get (3.36). ■

3.6 Tuning parameters

Recall that for big n, $\alpha_I > 8\alpha_{II}$.

- Choose n so big that statements of Theorem 2.1, Theorem 2.2, relation (2.26) and the statement of Corollary 2.1 hold.
- Then choose $c_1(n) > C_1(n)$, where $C_1(n)$ is specified in Proposition 3.1.
- Then choose $l_2(c_1, n)$ so big that simultaneously

$$\alpha_{II} l_2 > 1.5 + \ln 2 - c_1 \ln p_L \tag{3.40}$$
$$(\alpha_I - 7\alpha_{II}) l_2 > 9 \tag{3.41}$$
$$4\alpha_{II} l_2 > -2c_1 \ln p_{min} \tag{3.42}$$
$$\alpha_{II} l_2 > \ln 2 \tag{3.43}$$
$$a l_2 > 3 \tag{3.44}$$

- Then take $\gamma(n, c_1, l_2) = 4\alpha_{II} l_2$
- Then take $\alpha(n, c_1, l_2) = 7\alpha_{II} l_2$

3.7 Proof of the main theorem

Recall Lemma 3.1. By (3.15), (3.16) and (3.17), for l_1 big enough, it holds

$$P\Big(\big(E^1_{\text{alg works}}(\tau)\big)^c \cap E^1_{\text{stop}}(\tau)\Big) \leq$$
$$P\Big(\big(E^1_{\text{only ladders}}\big)^c \cap E^1_{\text{stop}}(\tau)\Big) + P\Big(\big(E^1_{\text{all ladders}}\big)^c \cap E^1_{\text{stop}}(\tau)\Big) + P\Big(\big(B^1_{\text{unique fit}}\big)^c\Big); \tag{3.45}$$

$$P\Big(\big(E^1_{\text{only ladders}}\big)^c \cap E^1_{\text{stop}}(\tau)\Big) \leq P\Big(\big(E^1_{\text{recon straight}}\big)^c \cap E^1_{\text{stop}}(\tau)\Big) \leq \tag{3.46}$$
$$P\Big(\big(E^1_{\text{recon straight}}\big)^c \cap E^1_{\text{stop}}(\tau) \cap B^1_{\text{recon straight}}\Big) + P\Big(\big(B^1_{\text{recon straight}}\big)^c\Big);$$

$$P\Big(\big(E^1_{\text{all ladders}}\big)^c \cap E^1_{\text{stop}}(\tau)\Big) \leq P\Big(\big(B^1_{\text{enough barriers}}\big)^c\Big)$$
$$+ P\Big(\big(E^1_{\text{no mistake}}\big)^c\Big) + P\Big(\big(B^1_{\text{enough paths}}\big)^c \cap E^1_{\text{stop}}(\tau)\Big); \tag{3.47}$$

$$P\Big(\big(E^1_{\text{no mistake}}\big)^c\Big) \leq P\Big(\big(E^1_{\text{no mistake}}\big)^c \cap B^1_{\text{intervals OK}}\Big) + P\Big(\big(B^1_{\text{intervals OK}}\big)^c\Big); \tag{3.48}$$

$$P\Big(\big(B^1_{\text{enough times}}\big)^c \cap E^1_{\text{stop}}(\tau)\Big) \leq P\Big(\big(B^1_{\text{enough times}}\big)^c \cap E^1_{\text{stop}}(\tau) \cap B^1_{\text{intervals OK}}\Big) + P\Big(\big(B^1_{\text{intervals OK}}\big)^c\Big). \tag{3.49}$$

Recall the definitions of l_2. The condition (3.45) 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,

$$\exp(\alpha l_1)\exp(-1.5l_1 - 2\alpha_{II}l)(p_L)^{c_1 l_1} > 2\exp(\gamma l_1).$$

Recall the definition of p^* and note that the inequality in the previous display is (3.35). Hence, by Proposition 3.2, we have the bound (3.36). By (3.43), $k_4 := \exp(2\gamma - \alpha) = \exp(\alpha_{II}l_2) > 2$, implying that (3.36) is exponentially small in l_1. By (3.44), there exist $k_5 > 0$ such that (3.21) is bounded by $4\exp[-k_5 l_1]$. With (3.36), we obtain that (3.49) is bounded by $40\exp[-(k_4 \wedge k_5)l_1]$.

Use (3.31) and (3.28) with (3.41) to obtain that $P((E^1_{\text{no mistake}})^c \leq 12\exp[(9+\alpha)l_1 - \alpha_I l] = 12\exp[-k_6 l_1]$ for a $k_6 > 0$. Hence, (3.48) 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 (3.24), we now get that (3.47) is bounded by $40\exp[-(k_4 \wedge k_5)l_1] + 16\exp[-(k_6 \wedge k_5)l_1] + \exp[-k_3 l_1] \leq 56\exp[-k_7 l_1]$ for a $k_7 > 0$.

The requirement (3.42) states that $\gamma + 2c_1 \ln p_{min} > 0$ implying that

$$\exp[l_1(2(n^2 l_2 + l_2)\ln 2 + (4l_2)(\ln l_1 + \ln l_2)) - d\exp((\gamma + 2c_1 \ln p_{min})l_1)] \leq \exp[k_8 l_1]$$

for l_1 big enough. This means (3.46) is bounded by $\exp[-k_9 l_1]$ for l_1 big enough.

Finally, we get that (3.45) is bounded by $\exp[-kl_1]$, if l_1 is big enough. This proves Theorem 1.1.

4 Appendix

4.1 Proof of Theorem 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 $E_{\mathrm{OK}a}$. We can rewrite

$$E_{\mathrm{OK}a} = \Big\{ \sum_{i=2Lm^2}^{l} X_i \leq l2\epsilon(n) \Big\},$$

where X_i is Bernoulli random variable that is one 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, $E_{\mathrm{OK}a}$ is independent on ξ^{Lm^3}. Then also $l - l^* = c - 1 \leq 2Lm^2$. Let us estimate

$$\begin{aligned} E_{\mathrm{OK}_a^c} &= \Big\{ \sum_{i=2Lm^2}^{l} X_i > l2\epsilon(n) \Big\} \subset \Big\{ \sum_{i=2Lm^2}^{l-2Lm^2} X_i > l2\epsilon(n) - 2Lm^2 \Big\} \\ &\subset \bigcup_{j=-c+1}^{c} \Big\{ \sum_{k=k_*}^{k^*} X_{ik2c-j} > \frac{l2\epsilon(n) - 2Lm^2}{2c} \Big\} \\ &\subset \bigcup_{j=-c+1}^{c} \Big\{ \sum_{k=k_*}^{k^*} X_{k2c-j} - (k^* - k_* + 1)\epsilon(n) > \frac{l2\epsilon(n) - 2Lm^2}{2c} - \frac{l\epsilon(n)}{2c} \Big\}. \end{aligned}$$

Here $k_* := \ulcorner \frac{2Lm^2+c}{2c} \urcorner$ and $k^* := \llcorner \frac{l-2Lm^2-c+1}{2c} \lrcorner$. Thus $k^* - k_* \leq \frac{l-4Lm^2+1}{2c} < \frac{l}{2c}$, $k^* - k_* + 1 < l$.
Note, by definition $X_i \in \sigma(\xi_j | j = i - c, i - c + 1, \ldots, i + c - 1)$. Thus, X_k and X_{k2c} are independent. This means, for each j we can apply Höffding's inequality. Thus, for each j

$$\begin{aligned} P\Big(\sum_{k=k_*}^{k^*} (X_{k2c-j} - \epsilon(n)) > \frac{l\epsilon(n) - 2Lm^2}{2c} \Big) &\leq P\Big(\sum_{k=k_*}^{k^*} (X_{k2c-j} - EX_{k2c-j}) > \frac{l\epsilon(n) - 2Lm^2}{2c} \Big) \\ &\leq \exp\Big[-\frac{(l\epsilon(n) - 2Lm^2)^2}{c(k^* - k_*)} \Big] \leq \exp[-\frac{l\epsilon^2(n)}{2c}], \end{aligned}$$

provided l is big enough to satisfy $l\epsilon(n) - 2Lm^2 \geq l\frac{\epsilon(n)}{2}$. Hence,

$$P(E_{\mathrm{OK}_a^c}) \leq 2c \exp\Big(\frac{-\epsilon^2(n) l}{2c} \Big) \leq \exp(-a_1(n) l), \tag{4.1}$$

for some $a_1(n) > 0$, provided l is big enough.

We estimate $P(E_{\mathrm{OK}_b^c})$ by the same argument. Define

$$E^{i*}_{\mathrm{OK}b} := \left\{ \left| \mathcal{I}^i_{II}(\xi^{ml}) \right| \geq l(1 - \exp(-m^{0.8})) \right\}, \quad i = 1, 2.$$

Clearly, for n big enough,

$$E^{1*}_{\mathrm{OK}b} \cap E^{2*}_{\mathrm{OK}b} \subset E_{\mathrm{OK}b} \quad \text{and } \mathrm{P}(E_{\mathrm{OK}}{}^c) \leq P(E^{1*}_{\mathrm{OK}b}) + P(E^{2*}_{\mathrm{OK}b}). \tag{4.2}$$

Let us estimate $P(E^{2*}_{\mathrm{OK}b})$.

Let Y_i be Bernoulli random variable that is 1 iff ξ_i has not empty neighborhood. Let us estimate $P(Y_i = 1)$. If $d_{i-1} - Lm^2 \geq 0$ and $d_i + Lm^2 \leq lm$, then

$$\begin{aligned} P(Y_i = 1) &= \big(\exists j \in [d_{i-1} - Lm^2, d_i + Lm^2] : \xi(j) = \cdots = \xi(j + m^{0.9})\big) \\ &\leq (2Lm^2 + m + 1)(0.5)^{m^{0.9}} \leq \exp(-m^{0.85}), \end{aligned}$$

in m is big. Otherwise, by definition, $Y_i = 1$. Let N be such that the inequality above holds as well as (4.2) if $n > N$. Note that $E^{2*}_{\mathrm{OK}b}$ is independent of ξ^{Lm^3}.
Clearly $Y_i \in \sigma(\xi_{i-Lm}, \ldots, \xi_{i+Lm})$. 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 $i^* \leq \frac{l}{k}$ we get

$$\begin{aligned} {E^{2*}_{\mathrm{OK}b}}^c &= \Big\{ \sum_{i=2Lm^2}^{l} Y_i > l\exp(-m^{0.8}) \Big\} \subset \Big\{ \sum_{i=2Lm^2}^{l-2Lm^2} Y_i > l\exp(-m^{0.8}) - 2Lm^2 \Big\} \\ &\subset \bigcup_{j=0}^{k-1} \Big\{ \sum_{i=0}^{i^*} Y_{2Lm^2+j+ik} > \frac{l\exp(-m^{0.8}) - 2Lm^2}{k} \Big\} \\ &\subset \bigcup_{j=0}^{k-1} \Big\{ \sum_{i=0}^{i^*} Y_{2Lm^2+j+ik} - i^*\exp(-m^{0.85}) > \frac{l(\exp(-m^{0.8}) - \exp(-m^{0.85})) - 2Lm^2}{k} \Big\} \\ &\subset \bigcup_{j=0}^{k-1} \Big\{ \sum_{i=0}^{i^*} (Y_{2Lm^2+j+ik} - EY_{2Lm^2+j+ik}) > \frac{l(\exp(-m^{0.8}) - \exp(-m^{0.85})) - 2Lm^2}{k} \Big\}. \end{aligned}$$

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,

$$P({E^{2*}_{\mathrm{OK}b}}^c) \leq 2(k+1)\exp(-a_2(m)l) \leq 2(m+1)\exp(-a_2(m)l) \leq \exp(-a_3(m)l),$$

for some $a_3(m) > 0$.

Similarly we estimate $P(E^{1*}_{\mathrm{OK}b})$.
Let Z_i be Bernoulli random variable that is 1 iff ξ_i is not isolated. If $i \geq l - Lm$, then, by definition $Z_i = 1$. Thus

$${E^{1*}_{\mathrm{OK}b}}^c = \{ \sum_{i=2Lm^2}^{l} Z_i > l\exp(-m) \Big\} \subset \Big\{ \sum_{i=2Lm^2}^{l-Lm} Z_i > l\exp(-m) - Lm \Big\}.$$

Again, $E^{1*}_{\mathrm{OK}b}$ 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 iff the srandom 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

$$P({E^{1*}_{\mathrm{OK}b}}^c) \leq P\Big(\sum_{j=j^*}^{l} Z'_j > \frac{1}{2}(l\exp(-m) - Lm - 1) \Big).$$

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

$$P\Big(\sum_{j=j^*}^{l} Z'_j \geq \frac{1}{2}(le^{-m^{0.8}} - Lm)\Big) \leq P\Big(\sum_{j=1}^{l} Z'_j - le^{-cm} \geq \frac{1}{2}(le^{-m^{0.8}} - Lm) - le^{-cm}\Big) =$$

$$P\Big(\sum_{j=1}^{l}(Z'_j - EZ'_j) > \frac{1}{2}l(e^{-m^{0.8}} - 2e^{-cm}) - \frac{L}{2}m\Big) \leq \exp\Big[-\frac{(l(e^{-m^{0.8}} - 2e^{-cm}) - Lm)^2}{2l}\Big]. \quad (4.3)$$

The right side of (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,

$$P(E_{\mathrm{OK}}{}_b^c) \leq \exp(-a_3 l) + \exp(-a_4 l) \leq \exp(-a_5 l) \quad (4.4)$$

Now, by (2.2), (4.1), (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.

4.2 Proof of Proposition

By definition,

$$E_{min}(i) \in \sigma\Big(S(t) - S(t-1)\Big| t \in [1, (s_i - r_i)m]\Big).$$

This means, if $E_{min}(i) \neq \emptyset$, then $P_\psi(E_{min}(i)) \geq (p_{min})^{(s_i - r_i)m}$. We shall show that $E_{min}(i) \neq \emptyset$.
Let $i \in \{1, \ldots, k\}$. Let us describe an admissible path $R_i \in \mathcal{R}((s_i - r_i)m)$ such that simultaneously satisfies (2.13), (2.14), (2.15). If such an path exists then and (2.16) holds.

Consider an arbitrary index-interval $[l_{2i-1}, l_{2i}]$, $i > 1$. It corresponds to the location-interval $[r_i, s_i]$. Let $C_1 < \cdots < C_q$ be the big clusters of ψ in $[s_i, r_i]$. Denote by c_j, d_j, $j = 1, \ldots, q$ the beginnings and ends of big clusters, respectively. Hence, $C_j \subset [c_j, d_j]$. The path R_i should read the big clusters as one block, i.e. along the reading-path.
Moreover, let $B_1 < B_2 < \cdots < B_p$ be the blocks of ψ in the set $[s_i, r_i] \backslash (\cup_{j=1}^q [c_j + 2, d_j - 2])$ that are bigger than $m^2/2\bar{v}$. By definition, $l(B_j) < m^3$, $j = 1, \ldots, p$. Indeed, if $l(B_j) \geq m^3$, then B_j would be a (part of) big cluster. We refer to a B_j as a **small block**. The small blocks should be red as shortly as possible, i.e. along the reading path.
Finally let $A_1 < A_2 < \ldots < A_K$, $K = p + q$ be the ordered big clusters and small blocks. Let a_j, b_j denote (an arbitrary) reading-beginning and reading-end of A_j.
Since $i > 1$, it holds $l_{2i-1} \in \mathcal{I}_{II}$. Then D_{2i-1} has empty neighborhood, hence $[r_i, r_i + Lm^2]$ is empty (for ψ) and, therefore, does not contain any small blocks. Also D_{2i-1} is isolated. This implies that there is no point in $[r_i, r_i + Lm^2]$ that is connected with any point in $[r_i + Lm^2 + m, s_i]$. In particular, all objects of interest, $A_1, \ldots, A_K$ are outside of $[r_i, r_i + Lm^2]$ or, formally, $a_1 > r_i + Lm^2$.
If $s_i - r_i \leq 2Lm^2$, then the interval does not contain blocks that are bigger than $m^{0.9}$. In this case the path R_i starts at r_i, i.e. $R(0) = r_i$ and goes to the point s_i with $(l_{2i} - l_{2i-1} + 1)m^2$ step without generating more than $m\bar{v}$ consecutive same colors in observations. This is clearly possible.
If $s_i - r_i > 2Lm^2$, then we define the minimum-blocks path R_i for interval $[r_i, s_i]$ backwards. More precisely, we define or prescribe a path R^* that starts at s_i and goes to r_i with $(s_i - r_i)m^2$ steps. The prescription of R^* is the following: start at s_i, i.e. $R^*(0) = s_i$. Then move stepwise to b_K (recall, this is a reading-end of the last small block or the last big cluster in $[r_i, s_i]$). Recall $s_i = l_{2i}m$ If $s_i \neq l$, then $l_{2i} \in \mathcal{I}_{II}$ and $[s_i - Lm^2, s_i + m]$ is empty and $[s_i - Lm^2 - m, s_i - Lm^2]$ contains a fence. As explained above, this implies that $b_K \leq s_i - Lm^2$. So, by moving stepwise from s_i 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 stepwise until $r_i + Lm^2$. Since $a_1 > r_i + Lm^2$, for such a path less than $((s_i - r_i) - Lm^2)\bar{v}$ steps are needed. This means that the path has more than $(s_i - r_i)(m - \bar{v}) + Lm^2\bar{v}$ steps to cover the interval $[r_i, r_i + Lm^2]$ with length Lm^2 without generating more than $m\bar{v}$ consecutive same colors in observations and satisfying $R^*((s_i - r_i)m) = r_i$. This is obviously possible, because the interval does not contain more than $m^{0.9}$ consecutive same colors. Finally define R_i as R^* backwards, i.e. $R_i(0) = R^*((s_i - r_i)m) = r_i, R_k(1) = R^*((s_i - r_i)m - 1), \ldots, R_i(j) = R^*((s_i - r_i)m - j), \ldots, R_k((s_i - r_i)m) = R^*(0) = s_i$ (recall, S is symmetric).
Such definition of R_i ensures that (2.13) and (2.15) hold. Let us show that (2.14) holds as well.

Note that the number of big blocks in $\psi \circ R_i$ is equal with the number of big clusters in $[r_i, s_i]$. Let this number be M. That means

$$\hat{q}_V(\psi \circ R_i) = q_V([r_i, s_i]) = M,$$

where $V := l_{2i} - l_{2i-1} + 1$. Let

$$T(j) := \inf\{k : q_k(\psi|[r_i, s_i]) = j\}, \quad \hat{T}(j) := \inf\{k : \hat{q}_k(\psi \circ R_i) = j\} \quad j = 1, \ldots, M.$$

Clearly, (2.14) is violated if there exists $j \in \{1, \ldots, M\}$ such that $\hat{T}(j) < T(j)$. Fix a $j \in \{1, \ldots, M\}$. The inequality $\hat{T}(j) < T(j)$ means that after reading the j-th big cluster, R_i has more than $(V - T(j) + 1)m^2$ steps to go to s_i. However, the path R_i is constructed such that after reaching to the b_j we have at most $(V - T(j) + 1)m\bar{v}$ step to go s_i. That proves (2.14).
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 (2.14) still holds.
Hence, $E^*_{min}(i) \neq \emptyset$ for each $i = 1, \ldots, k$.

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.