Reconstructing a random scenery observed with random errors along a random walk path

Heinrich Matzinger[*] Silke W.W. Rolles[†]

January 31, 2002

Abstract

We show that an i.i.d. uniformly colored scenery on $\mathbb{Z}$ observed along a random walk path with bounded jumps can still be reconstructed if there are some errors in the observations. We assume the random walk is recurrent and can reach every point with positive probability. At time k, the random walker observes the color at her present location with probability $1-\delta$ and an error Y_k with probability δ. The errors Y_k, $k \geq 0$, are assumed to be stationary and ergodic and independent of scenery and random walk. If the number of colors is strictly larger than the number of possible jumps for the random walk and δ is sufficiently small, then almost all sceneries can be almost surely reconstructed up to translations and reflections.

1 Introduction and result

We call a coloring of the integers $\mathbb{Z}$ with colors from the set $\mathcal{C} := \{1, 2, \ldots, C\}$ a *scenery*. Let $(S_k; k \in \mathbb{N}_0)$ be a recurrent random walk on $\mathbb{Z}$. At time k the random walker observes the color $\xi(S_k)$ at her current location. Given the color record $\chi := (\xi(S_k); k \in \mathbb{N}_0)$, can we almost surely reconstruct the scenery ξ without knowing the random walk path? This problem is called *scenery reconstruction problem*. In general, one can only hope to reconstruct the scenery up to *equivalence*, where we call two sceneries ξ and ξ' *equivalent* and write $\xi \approx \xi'$ if ξ is obtained from ξ' by a translation and/or reflection.

Early work on the scenery reconstruction problem was done by Kesten in [14]. He proved that a single defect in a 4-color random scenery can be detected if the scenery is i.i.d. uniformly colored. Reconstruction of typical 2-color sceneries was proved by Matzinger in his Ph.D. thesis [22] (see also [24] and [23]): Almost all i.i.d. uniformly colored sceneries observed along a simple random walk path

[*]Eurandom, PO Box 513, 5600 MB Eindhoven, The Netherlands. Matzinger@eurandom.tue.nl

[†]Eurandom, PO Box 513, 5600 MB Eindhoven, The Netherlands. Rolles@eurandom.tue.nl, Fax: 0031-40-247-8190

(with holding) can be almost surely reconstructed. In [15], Kesten noticed that the proof in [22] heavily relies on the skip-freeness of the random walk. In [21], Löwe, Matzinger and Merkl showed that scenery reconstruction is possible for random walks with bounded jumps if there are sufficiently many colors.

In this article, we prove that scenery reconstruction still works if the observations are seen with certain random errors. We make the same assumptions on scenery and random walk as in [21]: The random walk can reach every integer with positive probability and is recurrent with bounded jumps, and there are strictly more colors than possible single steps for the random walk. To keep the exposition as easy as possible, we assume in addition that for the random walk maximal jump length to the left and maximal jump length to the right are equal; we believe that the results of this paper remain true without this assumption. At time k the random walker observes color $\xi(S_k)$ with probability $1-\delta$, whereas she observes an error Y_k with probability δ. If the errors are independent of scenery and random walk, the occurences of errors are i.i.d. Bernoulli with parameter δ and Y_k, $k \geq 0$, is stationary and ergodic, then for all δ sufficiently small, almost all sceneries can be almost surely reconstructed up to translations and reflections.

More precisely, we consider the following setup: Let $\delta \in]0,1[$. Let μ be a probability measure over $\mathbb{Z}$ with finite support $\mathcal{M}$. With respect to a probability measure P_δ, let $S = (S_k; k \in \mathbb{N}_0)$ be a random walk starting at the origin with independent μ-distributed increments. We assume that $E[S_1] = 0$ and $\mathcal{M}$ has greatest common divisor 1; hence S is recurrent and can reach every $z \in \mathbb{Z}$ with positive probability. Let $\xi = (\xi_k; k \in \mathbb{Z})$ be a family of i.i.d. random variables, uniformly distributed over $\mathcal{C}$. Let $X := (X_k; k \in \mathbb{N}_0)$ be a sequence of i.i.d. random variables taking values in $\{0,1\}$, Bernoulli distributed with parameter δ, and let $Y := (Y_k; k \in \mathbb{N}_0)$ be as sequence of random variables taking values in $\mathcal{C}$ which is stationary and ergodic under P_δ. We assume that (ξ, S, X, Y) are independent. The scenery observed with errors along the random walk path is the process $\tilde{\chi} := (\tilde{\chi}_k; k \in \mathbb{N}_0)$ defined by $\tilde{\chi}_k := \chi_k = \xi(S_k)$ if $X_k = 0$ and $\tilde{\chi}_k := Y_k$ if $X_k = 1$. Our main theorem reads as follows:

Theorem 1.1 *If $|\mathcal{C}| > |\mathcal{M}|$, then there exists $\delta_1 > 0$ and a map $\mathcal{A} : \mathcal{C}^{\mathbb{N}_0} \longrightarrow \mathcal{C}^{\mathbb{Z}}$ which is measurable with respect to the canonical sigma algebras, such that $P_\delta\left(\mathcal{A}(\tilde{\chi}) \approx \xi\right) = 1$ for all $\delta \in]0, \delta_1[$.*

If $\delta = 0$, there are no errors in the observations. In this case, the assertion of Theorem 1.1 was proved by Löwe, Matzinger, and Merkl in [21].

Closely related coin tossing problems have been investigated by Harris and Keane [7] and Levin, Pemantle, and Peres [17]. The present paper has to a large extend been motivated by their work and a question of Peres who asked for generalizations of the existing random coin tossing results for the case of many biased coins.

Let $\chi' := (\chi'_k; k \in \mathbb{N}_0)$ be a coin tossing record, obtained in one of the following ways: a) a (two-sided) fair coin is tossed i.i.d., or b) at renewal times of a renewal process a coin with bias θ is tossed and at all other times a fair coin. Can we almost surely determine from χ' whether we are in case a) or b)?

Let u_n denote the probability of a renewal at time n. Harris and Keane in [7] showed that if $\sum_{n=1}^{\infty} u_n^2 = \infty$ then we can almost surely determine how χ' was produced, whereas this is not possible if $\sum_{n=1}^{\infty} u_n^2 < \infty$ and θ is small enough. Levin, Pemantle, and Peres in [17] showed that to distinguish between a) and b) not only the square-summability of (u_n) but also θ is relevant. They proved that for some renewal sequence (u_n) there is a phase transition: There exists a critical parameter θ_c such that for $|\theta| > \theta_c$ we can almost surely distinguish between a) and b), whereas for $|\theta| < \theta_c$ this is not possible.

The problem we address in this paper can be seen as a generalization of the following coin tossing problem: We have C different coins $\gamma_1, \gamma_2, \ldots, \gamma_C$ each one with C different faces $1, 2, \ldots, C$. Coin γ_i has distribution μ_i which gives probability $1 - \delta + \delta/C$ to face i and probability δ/C to each remaining face. For all $z \in \mathbb{Z}$ we choose i.i.d. uniformly among $\gamma_1, \gamma_2, \ldots, \gamma_C$ a coin $\zeta(z)$. Let $(S_k; k \in \mathbb{N}_0)$ be a random walk on $\mathbb{Z}$ fulfilling the conditions described above, independent of ζ. We generate a coin tossing record $\chi' := (\chi'_k; k \in \mathbb{N}_0)$ by tossing the coin $\zeta(S_k)$ at location S_k at time k. Then χ' has the same distribution as $\tilde{\chi}$ defined above, if we choose Y_k i.i.d. uniformly distributed over $\mathcal{C}$. Theorem 1.1 implies that we can almost surely determine ζ up to equivalence from the coin tossing record χ', as long as δ is small enough.

Research on random sceneries started by work by Keane and den Hollander ([13] and [5]) who studied ergodic properties of a color record seen along a random walk. Their questions were motivated among others by the work of Kalikow [12] in ergodic theory. More recently, den Hollander, Steif [4], and Heicklen, Hoffman, Rudolph [8] contributed to this area.

A preform of the scenery reconstruction problem is the scenery distinguishing problem (for a description of the problem see [15]) which started with the question whether any two non-equivalent sceneries can be distinguished. This question was asked by Benjamini and independently by den Hollander and Keane. The problem has been investigated by Benjamini and Kesten in [2] and [14]. Howard in [11], [10], [9] also contributed to this area. Recently, Lindenstrauss [18] showed the existence of uncountably many sceneries which cannot be reconstructed.

Löwe and Matzinger [20] proved that two-dimensional sceneries can be reconstructed if there are enough colors. In the case of a 2-color scenery and simple random walk with holding, Matzinger [25] showed that the reconstruction can be done in polynomial time. By a result of Löwe and Matzinger [19], reconstruction is possible in many cases even if the scenery is not i.i.d., but has some correlations. In [16], Lenstra and Matzinger showed that scenery reconstruction is still possible if the random walk might jump more than distance 1 with very small probability and the tail of the jump distribution decays sufficiently fast.

The exposition is organized as follows. In Section 2, we introduce some notation and we formally describe our setup. Section 3 describes the structure of the proof of Theorem 1.1: By an ergodicity argument, it suffices to find a partial reconstruction algorithm $\mathcal{A}'$ which reconstructs correctly with probability $>1/2$. To construct $\mathcal{A}'$, we build partial reconstruction algorithms $\mathcal{A}^m$, $m \geq 1$, which reconstruct bigger and bigger pieces of scenery around the origin.

Section 4 contains the proofs of the theorems from Section 3. The core of the reconstruction is an algorithm Alg^n which reconstructs a finite piece of scenery around the origin given as input finitely many observations, stopping times and a small piece of scenery which has been reconstructed earlier. Section 5 contains the definition of Alg^n. In Section 6, we show that Alg^n fulfills its task with high probability.

2 Notation and setup

In this section, we collect frequently used notation.

Sets and functions: The cardinality of a set D is denoted by $|D|$. We write $f|D$ for the restriction of a function f to a set D. For a sequence $\mathcal{S} = (s_i; i \in I)$ we write $|\mathcal{S}| := |I|$ for the number of components of $\mathcal{S}$. If s_i is an entry of $\mathcal{S}$, we write $s_i \in \mathcal{S}$; sometimes we write $s(i)$ instead of s_i. For events B_k, $k \geq 1$, we write $\liminf_{k\to\infty} B_k := \cup_{n=1}^{\infty} \cap_{k=n}^{\infty} B_k$ for the event that all but finitely many B_k's occur.

Integers and integer intervals: $\mathbb{N}$ denotes the set of natural numbers; by definition, $0 \notin \mathbb{N}$. We set $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. If $x \in \mathbb{R}$, we denote by $\lfloor x \rfloor$ the largest integer $\leq x$. Unless explicitly stated otherwise, intervals are taken over the integers, e.g. $[a,b] = \{n \in \mathbb{Z} : a \leq n \leq b\}$, $[a,b[= \{n \in \mathbb{Z} : a \leq n < b\}$.

Sceneries: We fix $C \geq 2$, and denote by $\mathcal{C} := \{1, ..., C\}$ the set of colors. A *scenery* is an element of $\mathcal{C}^{\mathbb{Z}}$. *A piece of scenery* is an element of $\mathcal{C}^I$ for a subset I of $\mathbb{Z}$; here I need not be an integer interval. The cardinality of the set I is called the length of the piece of scenery. We denote by $(1)_I$ the piece of scenery in $\mathcal{C}^I$ which is identically equal to 1. For $I = \{i_1, i_2, \ldots, i_k\} \subseteq \mathbb{Z}$ with $i_1 < i_2 < \ldots < i_k$ and a piece of scenery $\xi \in \mathcal{C}^I$ we define $\xi_{\rightarrow}$ to be the piece of scenery ξ read from left to right and $\xi_{\leftarrow}$ to be ξ read from right to left: $\xi_{\rightarrow} := (\xi(i_j); j \in [1,k])$ and $\xi_{\leftarrow} := (\xi(i_{k-j+1}); j \in [1,k])$.

Equivalence of sceneries: Let $\psi \in \mathcal{C}^I$ and $\psi' \in \mathcal{C}^{I'}$ be two pieces of sceneries. We say that ψ and ψ' are *equivalent* and write $\psi \approx \psi'$ iff I and I' have the same length and there exists $a \in \mathbb{Z}$ and $b \in \{-1, 1\}$ such that for all $k \in I$ we have that $a + bk \in I'$ and $\psi_k = \psi'_{a+bk}$. We call ψ and ψ' *strongly equivalent* and write $\psi \equiv \psi'$ if $I' = a + I$ for some $a \in \mathbb{Z}$ and $\psi_k = \psi'_{a+k}$ for all $k \in I$. We say ψ *occurs in* ψ' and write $\psi \sqsubseteq \psi'$ if $\psi \equiv \psi'|J$ for some $J \subseteq I'$. We write $\psi \preceq \psi'$ if $\psi \approx \psi'|J$ for some $J \subseteq I'$. If the subset J is unique, we write $\psi \preceq_1 \psi'$.

Random walks, random sceneries, and random errors: Let μ be a probability measure on $\mathbb{Z}$ with finite support $\mathcal{M}$. We assume that $|\mathcal{M}| < |\mathcal{C}|$, i.e. the number of colors is strictly larger than the number of possible jumps of the random walk. We assume $\max \mathcal{M} = |\min \mathcal{M}|$, and we write $L := \max \mathcal{M}$ for the maximal jump length of the random walk. Let $\Omega_2 \subseteq \mathbb{Z}^{\mathbb{N}_0}$ denote the set of all paths with jump sizes $S_{k+1} - S_k \in \mathcal{M}$ for all $k \in \mathbb{N}_0$. We denote by Q_x the distribution on $(\Omega_2)^{\mathbb{N}_0}$ of a random walk $(S_k; k \in \mathbb{N}_0)$ starting at x with i.i.d. increments distributed according to μ. We assume that $\sum_{k\in\mathcal{M}} \mu(k) = 0$ and $\mathcal{M}$

has greatest common divisor 1, consequently the random walk is recurrent and can reach every integer with positive probability.

The scenery $\xi := (\xi_k; k \in \mathbb{Z})$ is i.i.d. with ξ_k uniformly distributed on $\mathcal{C}$. Let $X := (X_k; k \in \mathbb{N}_0)$ be a sequence of i.i.d. Bernoulli random variables with values in $\{0,1\}$. If $X_k = 0$, then at time k the random walk observes color $\xi(S_k)$, whereas if $X_k = 1$ an error occurs in the observations at time k: the random walker observes Y_k, where $Y := (Y_k; k \in \mathbb{N}_0)$ is a sequence of random variables taking values in $\mathcal{C}$. We assume that (ξ, S, X, Y) are independent and realized as canonical projections on $\Omega := \left(\mathcal{C}^{\mathbb{Z}}, \Omega_2, \{0,1\}^{\mathbb{N}_0}, \mathcal{C}^{\mathbb{N}_0}\right)$ with the product σ-algebra generated by the canonical projections and probability measures $P_{\delta,x} := \nu^{\otimes\mathbb{Z}} \otimes Q_x \otimes B_\delta^{\otimes\mathbb{N}_0} \otimes \lambda$, $\delta \in [0,1]$, $x \in \mathbb{Z}$; here ν denotes the uniform distribution on $\mathcal{C}$, B_δ the Bernoulli distribution with parameter δ on $\{0,1\}$ and λ a probability measure on $\mathcal{C}^{\mathbb{N}_0}$ such that the left-shift is measure-preserving and ergodic with respect to λ. We abbreviate $P_\delta := P_{\delta,0}$ and $P := P_0$. .

We call $\chi := (\chi_k := \xi(S_k); k \in \mathbb{N}_0)$ the *scenery observed along the random walk path*; sometimes we write $\xi \circ S$ instead of χ. We define $\tilde{\chi} := (\tilde{\chi}_k; k \in \mathbb{N}_0)$, the *scenery observed with errors along the random walk path*, by

$$\tilde{\chi}_k := \begin{cases} \chi_k & \text{if } X_k = 0, \\ Y_k & \text{if } X_k = 1. \end{cases}$$

For a fixed scenery $\xi \in \mathcal{C}^{\mathbb{Z}}$ we set $P_\delta^\xi := \delta_\xi \otimes Q_0 \otimes B_\delta^{\otimes\mathbb{N}_0} \otimes \lambda$, where δ_ξ denotes the Dirac measure at ξ. Thus P_δ^ξ is the canonical version of the conditional probability $P_\delta(\cdot|\xi)$. We use P_δ^ξ and $P_\delta(\cdot|\xi)$ as synonyms; i.e. we never work with a different version of the conditional probability $P_\delta(\cdot|\xi)$.

Admissible paths: Let $I = [i_1, i_2]$ be an integer interval. We call a path $R \in \mathbb{Z}^I$ *admissible* if $R_{i+1} - R_i \in \mathcal{M}$ for all $i \in [i_1, i_2 - 1]$. We call $R(i_1)$ the starting point, $R(i_2)$ the endpoint, and $|I|$ the length of R.

Words: We call the elements of $\mathcal{C}^* := \cup_{n\in\mathbb{N}_0}\mathcal{C}^n$ *words.* If $w \in \mathcal{C}^n$, we say that *w has length n* and write $|w| = n$.

Ladder intervals, ladder paths, and ladder words: A *ladder interval* is a set of the form $I \cap (a + L\mathbb{Z})$ with a bounded interval I and a modulo class $a + L\mathbb{Z} \in \mathbb{Z}/L\mathbb{Z}$. Let I be a ladder interval. We call a path R of length $|I|$ which traverses I from left to right or from right to left a *ladder path* or a *straight crossing of I*. The *ladder words* of a scenery ξ over I are $(\xi|I)_\rightarrow$ and $(\xi|I)_\leftarrow$.

Filtration and shift: We define a filtration over Ω: $\mathcal{G} := (\mathcal{G}_n; n \in \mathbb{N}_0)$ with $\mathcal{G}_n := \sigma(\tilde{\chi}_k; k \in [0,n])$ is the natural filtration of the observations with errors. We define the shift $\theta : \mathcal{C}^{\mathbb{N}_0} \to \mathcal{C}^{\mathbb{N}_0}$, $\eta \mapsto \eta(\cdot + 1)$.

2.1 Conventions about constants

All constants keep their meaning throughout the whole article. Unless otherwise stated, they depend only on C and μ. Constants α, γ, ε, $\bar{\varepsilon}$, c_1, c_2 and n_1 play a special role in the constructions below; we state here how they are chosen. All other constants are denoted by c_i, $i \geq 3$, δ_i, ε_i, $i \geq 1$.

- We choose $\gamma > 0$.
- We choose $c_2 \in]1, \frac{C}{C-1}[$ and $\bar{\varepsilon} \in]0, \bar{\varepsilon}^{\max}[$ with

$$\bar{\varepsilon}^{\max} := \min\{1/30, \varepsilon_1/90, [\ln C - \ln c_2 - \ln(C-1)]/(90 \ln C)\},$$

  where ε_1 is as in Lemma 6.7.
- We choose $c_1 \in \mathbb{N}$ to be a multiple of 36 with $c_1 \geq 27/[\ln C - \ln c_2 - \ln(C-1) - 90\bar{\varepsilon} \ln C]$.
- We set $\varepsilon := c_1 \bar{\varepsilon}$.
- We choose $\alpha > \max\{\gamma, 1 + \gamma - [3c_1 \ln \mu_{\min}]/\ln 2\}$, where we abbreviate $\mu_{\min} := \min\{\mu(i) : i \in \mathcal{M}\}$.
- Finally we choose $n_1 \in \mathbb{N}$, $n_1 \geq \min\{25, c_3\}$, large enough that $2^n \geq c_1 L 2^{\lfloor \sqrt{n} \rfloor}$ for all $n \geq n_1$ and $\varepsilon_2(n_1) + (2\varepsilon_3(n_1))^{1/2} + \sum_{m=2}^{\infty} c_4 e^{-c_5 n_m} < 1/2$ holds, where c_3 is defined in Theorem 3.5, $\varepsilon_2(n_1)$ in Lemma 4.3, $\varepsilon_3(n_1)$ in Theorem 3.3 and c_4 and c_5 in Lemma 4.4.

3 The structure of the reconstruction

In order to prove Theorem 1.1, we reduce the problem of reconstructing the scenery successively to simpler problems. Theorems 3.1 and 3.2 below show that it suffices to find algorithms which do only partial reconstructions. Proofs are postponed to later sections: Sections 5 and 6 are dedicated to the proof of Theorem 3.5, all other statements of this section are proved in Section 4. Our first theorem states that it suffices to find a reconstruction algorithm $\mathcal{A}'$ which reconstructs correctly with probability $> 1/2$:

Theorem 3.1 *If there exist $\delta_1 > 0$ and a measurable map $\mathcal{A}' : \mathcal{C}^{\mathbb{N}_0} \longrightarrow \mathcal{C}^{\mathbb{Z}}$ such that $P_\delta\left(\mathcal{A}'(\tilde{\chi}) \approx \xi\right) > 1/2$ for all $\delta \in]0, \delta_1[$, then there exists a measurable map $\mathcal{A} : \mathcal{C}^{\mathbb{N}_0} \longrightarrow \mathcal{C}^{\mathbb{Z}}$ such that $P_\delta\left(\mathcal{A}(\tilde{\chi}) \approx \xi\right) = 1$ for all $\delta \in]0, \delta_1[$.*

The idea is to apply the reconstruction algorithm $\mathcal{A}'$ to all the shifted observations $\theta^i(\tilde{\chi})$, $i \geq 0$. By the hypothesis and an ergodicity argument, as k tends to infinity the proportion of sceneries $\mathcal{A}'(\theta^i(\tilde{\chi}))$ for $i \in [0, k[$ which are equivalent to ξ is strictly bigger than the proportion of sceneries which are not equivalent to ξ. Therefore we are able to reconstruct the scenery.

We build the algorithm $\mathcal{A}'$ required by Theorem 3.1 by putting together a hierarchy of partial reconstruction algorithms $\mathcal{A}^m$, $m \geq 1$. The algorithm $\mathcal{A}^m$ tries to reconstruct a piece of scenery around the origin of length of order 2^{n_m} with $(n_m; m \in \mathbb{N})$ recursively defined as follows: We choose n_1 as in Section 2.1, and we set for $m \geq 1$

$$n_{m+1} := 2^{\lfloor \sqrt{n_m} \rfloor}. \tag{3.1}$$

Definition 3.1 *For $m \geq 1$ and a measurable map $f : \mathcal{C}^{\mathbb{N}_0} \to \mathcal{C}^{[-3\cdot 2^{n_m}, 3\cdot 2^{n_m}]}$ we define*

$$E^m_{\mathrm{reconst,f}} := \{\xi|[-2^{n_m}, 2^{n_m}] \preceq f(\tilde{\chi}) \preceq \xi|[-4\cdot 2^{n_m}, 4\cdot 2^{n_m}]\}\,. \tag{3.2}$$

$E^m_{\mathrm{reconst,f}}$ is the event that the reconstruction procedure f reconstructs correctly a piece of scenery of length of order 2^{n_m} around the origin. Note that any finite piece of scenery occurs somewhere with probability 1 because the scenery is i.i.d. uniformly colored. Therefore it is crucial to reconstruct a piece of scenery around the origin.

Theorem 3.2 *Suppose there exist $\delta_1 > 0$ and a sequence of measurable maps $\mathcal{A}^m : \mathcal{C}^{\mathbb{N}_0} \to \mathcal{C}^{[-3\cdot 2^{n_m}, 3\cdot 2^{n_m}]}$, $m \geq 1$, such that for all $\delta \in]0, \delta_1[$*

$$\liminf_{m\to\infty} E^m_{\mathrm{reconst},\mathcal{A}^m} = \liminf_{m\to\infty}\left(E^m_{\mathrm{reconst},\mathcal{A}^m} \cap E^{m+1}_{\mathrm{center}}\right)\ \ P_\delta - a.s., \tag{3.3}$$

where $E^{m+1}_{\mathrm{center}} := \left\{\mathcal{A}^{m+1}(\tilde{\chi})|[-3\cdot 2^{n_m}, 3\cdot 2^{n_m}] = \mathcal{A}^m(\tilde{\chi})\right\}$. Suppose further that

$$P_\delta\left(\bigcup_{m=1}^{\infty}\left(E^m_{\mathrm{reconst},\mathcal{A}^m}\right)^c\right) < 1/2 \quad \text{for all } \delta \in]0, \delta_1[. \tag{3.4}$$

Then there exists a measurable map $\mathcal{A}' : \mathcal{C}^{\mathbb{N}_0} \longrightarrow \mathcal{C}^{\mathbb{Z}}$ such that $P_\delta\left(\mathcal{A}'(\tilde{\chi}) \approx \xi\right) > 1/2$ for all $\delta \in]0, \delta_1[$.

In the following, we explain how we construct maps $\mathcal{A}^m$ satisfying the assumptions of Theorem 3.2. The task of $\mathcal{A}^1$ is to reconstruct a piece of scenery of length of order 2^{n_1} around the origin with high probability. It is shown by Löwe, Matzinger, and Merkl in [21] that the whole scenery can be reconstructed with probability one in case there are no errors in the observations. They only prove existence of a reconstruction procedure, but do not explicitly construct an algorithm. In [26] we construct an algorithm which even works in polynomial time: A finite piece of scenery around the origin can be reconstructed with high probability from finitely many error-free observations; the number of observations needed is polynomial in the length of the piece of scenery which is reconstructed. We prove:

Theorem 3.3 *For infinitely many $n \in \mathbb{N}$ there exists a measurable map $\mathcal{A}^n_{\mathrm{initial}} : \mathcal{C}^{[0, 2\cdot 2^{12\alpha n}[} \to \mathcal{C}^{[-3\cdot 2^n, 3\cdot 2^n]}$ such that*

$$\varepsilon_3(n) := P\left(\left\{\xi|[-2^n, 2^n] \preceq \mathcal{A}^n_{\mathrm{initial}}\left(\chi|[0, 2\cdot 2^{12\alpha n}[\right) \preceq \xi|[-4\cdot 2^n, 4\cdot 2^n]\right\}^c\right)$$

satisfies $\lim_{n\to\infty} \varepsilon_3(n) = 0$.

As an immediate consequence of Theorem 3.3 a piece of scenery around the origin can be reconstructed with high probability even if there are errors in the observations. As long as the probability δ to see an error at a particular time is sufficiently small, the probability to see no errors in the first $2\cdot 2^{12\alpha n}$ observations is close to 1. The following corollary makes this precise:

Corollary 3.1 *Let $\mathcal{A}^n_{\text{initial}}$ and $\varepsilon_3(n)$ be as in Theorem 3.3. There exist $\delta_2(n) > 0$ such that for all $\delta \in]0, \delta_2(n)[$*

$$P_\delta \left(\left\{ \xi|[-2^n, 2^n] \preceq \mathcal{A}^n_{\text{initial}} \left(\tilde{\chi}|[0, 2 \cdot 2^{12\alpha n}[\right) \preceq \xi|[-4 \cdot 2^n, 4 \cdot 2^n] \right\}^c \right) \le 2\varepsilon_3(n).$$

We will choose $\mathcal{A}^1 := \mathcal{A}^{n_1}_{\text{initial}}$. The maps $\mathcal{A}^m$, $m \ge 2$, will be defined inductively. Given a partial reconstruction algorithm $\mathcal{A}^m$ we define stopping times which tell us when the random walker is in some sense "close" to the origin: We compare $\mathcal{A}^m(\tilde{\chi})$ with $\mathcal{A}^m(\theta^t(\tilde{\chi}))$, i.e. we compare the output of $\mathcal{A}^m$ if the input consists of the observations collected by the random walker starting at the origin and the observations starting at time t. If both outputs agree up to equivalence on a sufficiently large subpiece, then with a high chance, the random walker is - on an appropriate scale - close to the origin.

The stopping times constructed from $\mathcal{A}^m$ are used to reconstruct a piece of scenery around the origin of length of order $2^{n_{m+1}}$ which is much larger than the piece of scenery reconstructed by $\mathcal{A}^m$; recall our choice of n_m (3.1). Whenever the stopping times indicate that the random walk is "close" to the origin, we collect significant parts of the observations of length $c_1 n_m$. If we have sufficiently many stopping times, the random walk will walk over the same piece of scenery over and over again. This allows us to filter out the errors in the observations. Once this is done, the obtained words are put together like in a puzzle game. The words are used to extend the piece of scenery of length of order 2^{n_m} which has been reconstructed by $\mathcal{A}^m$.

Formally we define stopping times in the following way:

Definition 3.2 *For $m \in \mathbb{N}$ and a measurable map $f : \mathcal{C}^{\mathbb{N}_0} \to \mathcal{C}^{[-3\cdot 2^{n_m}, 3\cdot 2^{n_m}]}$ with the property that $f(\tilde{\chi})$ depends only on $\tilde{\chi}|[0, 2 \cdot 2^{12\alpha n_m}[$, we define*

$$\mathbb{T}^{m+1}_f(\tilde{\chi}) := \begin{Bmatrix} t \in \left[0, 2^{12\alpha n_{m+1}} - 2 \cdot 2^{12\alpha n_m}\right[: \exists w \in \mathcal{C}^{[-2^{n_m}, 2^{n_m}]} \text{ such that} \\ w \preceq f(\tilde{\chi}) \text{ and } w \preceq f(\theta^t(\tilde{\chi})) \end{Bmatrix}.$$

Let $t(1) < t(2) < \cdots$ be the elements of $\mathbb{T}^{m+1}_f(\tilde{\chi})$ arranged in increasing order. We define the sequence $T^{m+1}_f(\tilde{\chi}) := \left(T^{m+1}_{f,k}(\tilde{\chi}); k \ge 1 \right)$ by

$$T^{m+1}_{f,k}(\tilde{\chi}) := \begin{cases} t(2 \cdot 2^{2n_{m+1}} k) + 2 \cdot 2^{12\alpha n_m} & \text{if } 2 \cdot 2^{2n_{m+1}} k \le \left| \mathbb{T}^{m+1}_f(\tilde{\chi}) \right|, \\ 2^{12\alpha n_{m+1}} & \text{otherwise.} \end{cases}$$

$T^{m+1}_f(\tilde{\chi})$ is a sequence of $\mathcal{G}$-adapted stopping times with values in $\left[0, 2^{12\alpha n_{m+1}}\right]$; the stopping times depend only on $\tilde{\chi}|\left[0, 2^{12\alpha n_{m+1}}\right[$. We define the event that a sequence of stopping times fulfils the task of stopping the random walk "close" to the origin (on a rather rough scale).

Definition 3.3 *For $n \in \mathbb{N}$ and a sequence $\tau = (\tau_k; k \ge 1)$ of $\mathcal{G}$-adapted stopping times we define the event $E^{n,\tau}_{\text{stop}} :=$*

$$\bigcap_{k=1}^{2^{\alpha n}} \left\{ \tau_k(\tilde{\chi}) < 2^{12\alpha n}, |S(\tau_k(\tilde{\chi}))| \le 2^n, \tau_j(\tilde{\chi}) + 2 \cdot 2^{2n} \le \tau_k(\tilde{\chi}) \text{ for } j < k \right\}.$$

The next theorem states that given an appropriate partial reconstruction algorithm f, the stopping times T_f^{m+1} fulfil their task with a high probability. By the definition of T_f^{m+1}, we stop at time $t + 2\cdot 2^{12\alpha n_m}$ iff $f(\tilde{\chi})$ and $f(\theta^t(\tilde{\chi}))$ agree on a large enough subpiece. Therefore, for the stopping times to stop the random walk close to the origin, it is necessary that $f(\tilde{\chi})$ is a correctly reconstructed piece of scenery around the origin. Since we apply f often to obtain enough stopping times, we need that given a scenery ξ, there is a high enough chance for the random walk on ξ to be stopped correctly, i.e. f must reconstruct correctly with high enough probability conditional on ξ. This is why we need the event $\left\{P_\delta\left[E^m_{\mathrm{reconst,f}}\,|\,\xi\right]\geq \frac{1}{2}\right\}$ in the following theorem.

Theorem 3.4 *Let $m \geq 1$, and let $f : \mathcal{C}^{\mathbb{N}_0} \to \mathcal{C}^{[-3\cdot 2^{n_m}, 3\cdot 2^{n_m}]}$ be a measurable map with the property that $f(\tilde{\chi})$ depends only on $\tilde{\chi}|[0, 2\cdot 2^{12\alpha n_m}[$. We have for all $\delta \in]0,1[$*

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

The next theorem shows that there exist partial reconstruction algorithms Alg^n (the reader should think of $n = n_m$) with the following properties: Given stopping times which stop the random walk close to the origin, finitely many observations with errors and a small piece of scenery ψ close to the origin, Alg^n reconstructs with high probability a piece of scenery around the origin of length of order 2^n. If the reconstruction is succesful, the output of Alg^n contains ψ in the middle. The reader should think of ψ as a piece of scenery that has been reconstructed before.

Theorem 3.5 *For all $n \in \mathbb{N}$ there exists a measurable map*

$$Alg^n : [0, 2^{12\alpha n}]^{\mathbb{N}} \times \mathcal{C}^{2\cdot 2^{12\alpha n}} \times \bigcup_{k\geq c_1 L} \mathcal{C}^{[-kn,kn]} \to \mathcal{C}^{[-3\cdot 2^n, 3\cdot 2^n]}$$

with the following property: There exist constants $c_3, \delta_3, c_6, c_7 > 0$ such that for all $n \geq c_3$, $\delta \in]0, \delta_3[$ and for any sequence $\tau = (\tau_k; k \geq 1)$ of $\mathcal{G}$-adapted stopping times with values in $[0, 2^{12\alpha n}]$

$$P_\delta\left(E^{n,\tau}_{\mathrm{stop}} \setminus E^{n,\tau}_{\mathrm{reconstruct}}\right) \leq c_6 e^{-c_7 n},$$

where $E^{n,\tau}_{\mathrm{reconstruct}} :=$

$$\left\{\begin{matrix}\text{For all } \psi \in \mathcal{C}^{[-kn,kn]} \text{ with } k \geq c_1 L \text{ and } \psi \preceq \xi|[-2^n, 2^n] \text{ we have} \\ \xi|[-2^n,2^n] \preceq Alg^n(\tau, \tilde{\chi}|\left[0, 2\cdot 2^{12\alpha n}\right[, \psi) \preceq \xi|[-4\cdot 2^n, 4\cdot 2^n].\end{matrix}\right\}.$$

Furthermore if $\xi|[-2^n,2^n] \preceq Alg^n(\tau, \tilde{\chi}|\left[0, 2\cdot 2^{12\alpha n}\right[, \psi) \preceq \xi|[-4\cdot 2^n, 4\cdot 2^n]$ holds, $\psi \in \mathcal{C}^{[-kn,kn]}$ with $k \geq c_1 L$, $\psi \preceq \xi|[-2^n, 2^n]$ and $\xi|[-2^n,2^n] \neq (1)_{[-2^n,2^n]}$, then we conclude that $Alg^n(\tau, \tilde{\chi}|\left[0, 2\cdot 2^{12\alpha n}\right[, \psi)|[-kn, kn] = \psi$.

To motivate the allowed range for the abstract arguments τ in this theorem, recall that the $T^m_{f,k}(\tilde{\chi})$'s in Definition 3.2 take their values in $[0, 2^{12\alpha n_m}]$. We are now able to define $\mathcal{A}^m$, $m \geq 1$, which fulfill the requirements of Theorem 3.2.

Definition 3.4 *We define* $\mathcal{A}^m : \mathcal{C}^{\mathbb{N}_0} \to \mathcal{C}^{[-3\cdot 2^{n_m}, 3\cdot 2^{n_m}]}$ *and sequences* $T^{m+1} = \left(T^{m+1}_k; k \geq 1\right)$ *recursively for* $m \geq 1$ *in the following way:*

- $\mathcal{A}^1(\tilde{\chi}) := \mathcal{A}^{n_1}_{\mathrm{initial}}\left(\tilde{\chi}|[0, 2\cdot 2^{12\alpha n_1}[\right)$ *with* n_1 *as in Section 2.1 and* $\mathcal{A}^{n_1}_{\mathrm{initial}}$ *as in Theorem 3.3,*
- $T^{m+1}(\tilde{\chi}) := T^{m+1}_{\mathcal{A}^m}(\tilde{\chi})$ *with* $T^{m+1}_{\mathcal{A}^m}$ *as in Definition 3.2,*
- $\mathcal{A}^{m+1}(\tilde{\chi}) := Alg^{n_{m+1}}\left(T^{m+1}(\tilde{\chi}), \tilde{\chi}|[0, 2\cdot 2^{12\alpha n_{m+1}}[, \mathcal{A}^m(\tilde{\chi})\right)$ *with* $Alg^{n_{m+1}}$ *as in Theorem 3.5.*

Theorem 3.6 *There exists* $\delta_1 > 0$ *such that the sequence* $(\mathcal{A}^m; m \in \mathbb{N})$ *defined in Definition 3.4 fulfils (3.3) and (3.4) for all* $\delta \in]0, \delta_1[$.

All theorems of this section together yield the proof of our main theorem:
Proof of Theorem 1.1. By Theorem 3.6, the assumptions of Theorem 3.2 are satisfied. Hence the assumptions of Theorem 3.1 are satisfied and Theorem 1.1 follows. ■

4 Proofs

In this section, we prove the statements from Section 3 with the exception of Theorem 3.5 which will be proved in Sections 5 and 6.

Lemma 4.1 *The shift* $\Theta : \Omega \to \Omega$,

$$(\xi, S, X, Y) \mapsto (\xi(\cdot + S(1)), S(\cdot + 1) - S(1), X(\cdot + 1), Y(\cdot + 1))$$

is measure-preserving and ergodic with respect to P_δ *for all* $\delta \in]0, 1[$.

Proof. Let $\delta \in]0, 1[$. By assumption, Y_k, $k \geq 0$, is stationary and ergodic under P_δ. X_k, $k \geq 0$, is i.i.d., hence stationary and ergodic under P_δ. By Lemma 4.1 of [21], $(\xi, S) \mapsto (\xi(\cdot + S(1)), S(\cdot + 1) - S(1))$ is measure-preserving and ergodic with respect to P. The claim follows from these three observations and the fact that (ξ, S, X, Y) are independent. ■

Proof of Theorem 3.1. Let δ_1 and $\mathcal{A}' : \mathcal{C}^{\mathbb{N}_0} \to \mathcal{C}^{\mathbb{Z}}$ be as in the hypothesis of the theorem, and let $\delta \in]0, \delta_1[$. We define for $k \in \mathbb{N}$ measurable maps $\mathcal{A}'_k : \mathcal{C}^{\mathbb{N}_0} \to \mathcal{C}^{\mathbb{Z}}$ as follows: If there exists $j \in [0, k[$ such that

$$\left|\left\{i \in [0, k[: \mathcal{A}'(\theta^i(\tilde{\chi})) \approx \mathcal{A}'(\theta^j(\tilde{\chi}))\right\}\right| > \left|\left\{i \in [0, k[: \mathcal{A}'(\theta^i(\tilde{\chi})) \not\approx \mathcal{A}'(\theta^j(\tilde{\chi}))\right\}\right|,$$

then let j_0 be the smallest j with this property, and define $\mathcal{A}'_k(\tilde{\chi}) := \mathcal{A}'(\theta^{j_0}(\tilde{\chi}))$. Otherwise define $\mathcal{A}'_k(\tilde{\chi})$ to be the constant scenery $(1)_{j\in\mathbb{Z}}$. Finally we define $\mathcal{A} : \mathcal{C}^{\mathbb{N}_0} \to \mathcal{C}^{\mathbb{Z}}$ by

$$\mathcal{A}(\tilde{\chi}) := \begin{cases} \lim_{k\to\infty} \mathcal{A}'_k(\tilde{\chi}) & \text{if this limit exists pointwise,} \\ (1)_{j\in\mathbb{Z}} & \text{else.} \end{cases}$$

As a limit of measurable maps, $\mathcal{A}$ is measurable. For $k \in \mathbb{N}$ we define

$$Z_k := \frac{1}{k} \sum_{i=0}^{k-1} 1\{\mathcal{A}'(\theta^i(\tilde{\chi})) \approx \xi\};$$

here $1B$ denotes the indicator function of the event B. It follows from Lemma 4.1 that the sequence $1\{\mathcal{A}'(\theta^k(\tilde{\chi})) \approx \xi\}$, $k \geq 0$, is stationary and ergodic because it can be written as a measurable function of the sequence $\Theta^k(\xi, S, X, Y)$, $k \geq 0$; note that $\xi \approx \xi(\cdot + S_k)$. Hence we can use the ergodic theorem and our assumption to obtain P_δ-almost surely:

$$\lim_{k\to\infty} Z_k = P_\delta\left(\mathcal{A}'(\tilde{\chi}) \approx \xi\right) > 1/2. \tag{4.1}$$

Note that if $Z_k > 1/2$, then $\mathcal{A}'_k(\tilde{\chi}) \approx \xi$. By (4.1) there exists a.s. a (random) k_0 such that $Z_k > 1/2$ for all $k \geq k_0$, and hence $\mathcal{A}'_k(\tilde{\chi}) = \mathcal{A}'_{k_0}(\tilde{\chi}) \approx \xi$; recall that we chose the smallest possible j_0 in the definition of $\mathcal{A}'_k$. Thus a.s. $\mathcal{A}(\tilde{\chi}) \approx \xi$. ■

Proof of Theorem 3.2. We say a sequence $(\zeta^m; m \in \mathbb{N})$ of pieces of sceneries converges pointwise to a scenery ζ if $\liminf_{m\to\infty} \mathrm{domain}(\zeta^m) = \mathbb{Z}$, and for every $z \in \mathbb{Z}$ there is $m_z > 0$ such that $\zeta^m(z) = \zeta(z)$ for all $m \geq m_z$.

Let δ_1 and $\mathcal{A}^m$ be as in the hypothesis of the theorem, and let $\delta \in]0, \delta_1[$. We set $\mathcal{A}'(\tilde{\chi}) := \lim_{m\to\infty} \mathcal{A}^m(\tilde{\chi})$ if this limit exists pointwise on $\mathbb{Z}$; otherwise we set $\mathcal{A}'(\tilde{\chi}) := (1)_{j\in\mathbb{Z}}$. Being a pointwise limit of measurable maps, $\mathcal{A}' : \mathcal{C}^{\mathbb{N}_0} \to \mathcal{C}^{\mathbb{Z}}$ is measurable. We abbreviate $E^m := E^m_{\mathrm{reconst},\mathcal{A}^m}$, and define the events

$$E^m_{1\mathrm{fit}} := \left\{\xi|[-2^{n_m}, 2^{n_m}] \preccurlyeq_1 \xi|[-4 \cdot 2^{n_{m+1}}, 4 \cdot 2^{n_{m+1}}]\right\}.$$

We claim:

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

Together with the assumption $P_\delta\left[\cup_{m=1}^\infty (E^m)^c\right] < 1/2$ these two statements imply that $P_\delta\left(\mathcal{A}'(\tilde{\chi}) \approx \xi\right) > 1/2$ which yields the claim of the theorem.

Proof of claim 1: We show for any integer intervals $I_1 \neq I_2$ with $|I_1| = |I_2|$

$$P(\xi|I_1 \approx \xi|I_2) \leq 2 \cdot C^{-|I_j|/3}. \tag{4.2}$$

First we define $f_j : [0, |I_j|[\to I_j$ for $j = 1, 2$ to be the unique translation which maps $[0, |I_j|[$ onto I_j. An argument similar to the proof of (6.26) below shows that there exists a subset $J \subseteq [0, |I_j|[$ of cardinality $|J| \geq |I_j|/3$ with $f_1(J) \cap f_2(J) = \emptyset$. Since ξ_k, $k \in \mathbb{Z}$, are i.i.d. with a uniform distribution, we conclude

$$P(\xi|I_1 \equiv \xi|I_2) \leq P(\xi|f_1(J) = \xi|f_2(J)) = C^{-|J|} \leq C^{-|I_j|/3}.$$

Since $\xi|I_1 \approx \xi|I_2$ means $\xi|I_1 \equiv \xi|I_2$ or $\xi|I_1 \equiv (\xi|I_2)^{\leftrightarrow}$ with $(\xi|I_2)^{\leftrightarrow}$ denoting the piece of scenery obtained from $\xi|I_2$ by reflection, estimate (4.2) follows.

We apply (4.2) for $I_1 = [-2^{n_m}, 2^{n_m}]$ and all integer intervals $I_2 \subseteq [-4 \cdot 2^{n_{m+1}}, 4 \cdot 2^{n_{m+1}}]$, $I_1 \neq I_2$, of length $|I_1| = |I_2| = 2 \cdot 2^{n_m} + 1$; there are not more than $8 \cdot 2^{n_{m+1}}$ choices for I_2. We obtain

$$P\left((E_{\mathrm{1fit}}^m)^c\right) \leq 8 \cdot 2^{n_{m+1}} \cdot 2 \cdot C^{-(2 \cdot 2^{n_m}+1)/3} \leq 16 \cdot 2^{2^{\sqrt{n_m}} - 2 \cdot 2^{n_m}/3},$$

which is summable over m; recall $C \geq 2$ and (3.1). Hence by the Borel-Cantelli lemma $(E_{\mathrm{1fit}}^m)^c$ occurs P_δ-a.s. only finitely many times ; this proves claim 1.

Proof of claim 2: By the assumption of this claim, there is a (random) M such that the events E_{1fit}^m and E^m hold for all $m \geq M$. By the assumption of Theorem 3.2, M can be chosen in such a way that $E_{\mathrm{center}}^{m+1}$ holds for all $m \geq M$, too. Consequently, $\mathcal{A}^{m+1}(\tilde{\chi})|\,[-3 \cdot 2^{n_m}, 3 \cdot 2^{n_m}] = \mathcal{A}^m(\tilde{\chi})$ for all $m \geq M$ and it follows that

$$\mathcal{A}'(\tilde{\chi})|[-k,k] = \mathcal{A}^m(\tilde{\chi})|[-k,k] \tag{4.3}$$

for all $k \geq 1$ and all m large enough. In particular, $\lim_{m\to\infty} \mathcal{A}^m(\tilde{\chi})$ exists.

Since E^m and E_{1fit}^m hold, $\mathcal{A}^m(\tilde{\chi}) \preceq_1 \xi|[-4 \cdot 2^{n_m}, 4 \cdot 2^{n_m}]$. Hence there exists a unique map $h^m : \mathbb{Z} \to \mathbb{Z}$ of the form $x \mapsto a_m + b_m x$ with $a_m \in \mathbb{Z}$ and $b_m \in \{-1, 1\}$ that maps $\mathcal{A}^m(\tilde{\chi})$ onto a subpiece of $\xi|[-4 \cdot 2^{n_m}, 4 \cdot 2^{n_m}]$. It follows from (4.3) that h^m is independent of m and maps $\mathcal{A}'(\tilde{\chi})$ to ξ. This finishes the proof of claim 2. ■

Proof of Theorem 3.3. By Theorem 1.1 of [26], we know that there exists $\beta > 0$ and for infinitely many $n \in \mathbb{N}$ there exists a measurable map $\mathcal{A}_{\mathrm{ini}}^n : \mathcal{C}^{[0, 2n^7 + 2 \cdot 2^{12\beta n}[} \to \mathcal{C}^{[-5 \cdot 2^n, 5 \cdot 2^n]}$ such that $\lim_{n\to\infty} P\left([E_{\mathrm{ini}}^n]^c\right) = 0$, where

$$E_{\mathrm{ini}}^n := \left\{\xi|\left[-2^{n-1}, 2^{n-1}\right] \preccurlyeq \mathcal{A}_{\mathrm{ini}}^n\left(\chi|[0, 2n^7 + 2 \cdot 2^{12\beta n}[\right) \preccurlyeq \xi|\left[-10 \cdot 2^n, 10 \cdot 2^n\right]\right\}.$$

Small modifications in the proof of Theorem 1.1 in [26] prove our claim. We remark that alternatively, we could work directly with the maps $\mathcal{A}_{\mathrm{ini}}^n$ from [26] without adjusting the constants; all proofs in the remainder of the article go through, but the notation becomes more cumbersome. ■

Proof of Corollary 3.1. We estimate the probability under consideration by intersecting with the event $B_0 := \left\{X_k = 0 \text{ for all } k \in \left[0, 2 \cdot 2^{12\alpha n}\right[\right\}$ that there are no errors in the first $2 \cdot 2^{12\alpha n}$ observations: For any $\delta > 0$ we have

$$\begin{aligned}
& 1 - P_\delta\left(\xi|[-2^n, 2^n] \preceq \mathcal{A}_{\mathrm{initial}}^n\left(\tilde{\chi}|[0, 2 \cdot 2^{12\alpha n}[\right) \preceq \xi|[-4 \cdot 2^n, 4 \cdot 2^n]\right) \\
\leq\ & 1 - P_\delta\left(\left\{\xi|[-2^n, 2^n] \preceq \mathcal{A}_{\mathrm{initial}}^n\left(\tilde{\chi}|[0, 2 \cdot 2^{12\alpha n}[\right) \preceq \xi|[-2^{n+2}, 2^{n+2}]\right\} \cap B_0\right) \\
=\ & 1 - \delta(n) P\left(\xi|[-2^n, 2^n] \preceq \mathcal{A}_{\mathrm{initial}}^n\left(\chi|[0, 2 \cdot 2^{12\alpha n}[\right) \preceq \xi|[-2^{n+2}, 2^{n+2}]\right) \\
=\ & 1 - \delta(n)(1 - \varepsilon_3(n));
\end{aligned}$$

with $\delta(n) := (1-\delta)^{2 \cdot 2^{12\alpha n}}$ and $\varepsilon_3(n)$ as in Theorem 3.3. We choose $\delta_2(n) > 0$ such that the last expression is bounded above by $2\varepsilon_3(n)$ for all $\delta \in]0, \delta_2(n)[$. ■

Proof of Theorem 3.4. The proof is very similar to the proof of Theorem 3.11 in section 7 of [21] (Our Theorem 3.4 is the analogon of their Theorem 3.11

for our setting). The errors in the observations do not require adaptations of their arguments; note that the errors are independent of scenery and random walk and occurences of errors are i.i.d. Bernoulli. ■

The rest of this section is dedicated to the proof of Theorem 3.6. Throughout we assume, $\mathcal{A}^m$, $m \geq 1$, are as in Definition 3.4, and we set $\delta_1 := \min\{\delta_3, \delta_2(n_1)\}$ with δ_3 as in Theorem 3.5 and $\delta_2(n_1)$ as in Corollary 3.1. We set for $m \geq 1$

$$E^m := E^m_{\text{reconst},\mathcal{A}^m}. \tag{4.4}$$

Definition 4.1 *For $\delta \in]0, \delta_1[$ we define events of sceneries*

$$\begin{aligned}
\Xi_1^\delta &:= \left\{\xi \in \mathcal{C}^{\mathbb{Z}} : P_\delta\left[(E^1)^c \middle| \xi\right] \leq (2\varepsilon_3(n_1))^{1/2}\right\}, \\
\Xi_2^\delta &:= \bigcap_{m=2}^{\infty} \left\{\xi \in \mathcal{C}^{\mathbb{Z}} : P_\delta\left[E^{m-1}|\xi\right] \geq \frac{1}{2} \Rightarrow P_\delta\left[E^{m-1} \setminus E^{n_m,T^m}_{\text{stop}} \middle| \xi\right] \leq e^{-\frac{n_m}{2}}\right\} \\
&= \bigcap_{m=2}^{\infty} \left\{\xi \in \mathcal{C}^{\mathbb{Z}} : P_\delta\left[\left[E^{m-1} \setminus E^{n_m,T^m}_{\text{stop}}\right] \cap \left\{P_\delta\left[E^{m-1}\,|\,\xi\right] \geq \frac{1}{2}\right\} \middle| \xi\right] \leq e^{-\frac{n_m}{2}}\right\}, \\
\Xi_3^\delta &:= \bigcap_{m=2}^{\infty} \left\{\xi \in \mathcal{C}^{\mathbb{Z}} : P_\delta\left[E^{m-1} \cap \left(E^{n_m,T^m}_{\text{stop}} \setminus E^m\right) \,\middle|\, \xi\right] \leq (c_6)^{1/2} e^{-\frac{c_7 n_m}{2}}\right\}, \\
\Xi^\delta &:= \Xi_1^\delta \cap \Xi_2^\delta \cap \Xi_3^\delta,
\end{aligned}$$

where $\varepsilon_3(n_1)$ is as in Theorem 3.3 and c_6 and c_7 are as in Theorem 3.5.

Note the similarity between these events and the bounds in Corollary 3.1, Theorems 3.4 and 3.5. The following lemma provides a link between bounds with and without conditioning on the scenery ξ:

Lemma 4.2 ([21], Lemma 4.6) *Let A be an event, $r \geq 0$, and let Q be a probability measure on Ω. If $Q(A) \leq r^2$, then $Q\left(Q(A|\xi) > r\right) \leq r$.*

Lemma 4.3 *For all $n \in \mathbb{N}$ there exist $\varepsilon_2(n) > 0$ with $\lim_{n\to\infty} \varepsilon_2(n) = 0$ such that $P_\delta\left(\xi \notin \Xi^\delta\right) \leq \varepsilon_2(n_1)$ for all $\delta \in]0, \delta_1[$.*

Proof. Let $\delta \in]0, \delta_1[$. Using Corollary 3.1 and Lemma 4.2 for $Q = P_\delta$, we obtain

$$P_\delta\left(\xi \notin \Xi_1^\delta\right) \leq (2\varepsilon_3(n_1))^{1/2}. \tag{4.5}$$

An application of Theorem 3.4 with $f = \mathcal{A}^m$ yields for $m \geq 2$

$$P_\delta\left(\left(E^{m-1} \setminus E^{n_m,T^m}_{\text{stop}}\right) \cap \left\{P_\delta\left[E^{m-1}\,|\,\xi\right] \geq \frac{1}{2}\right\}\right) \leq e^{-n_m}.$$

An application of Lemma 4.2 with $Q = P_\delta$ yields

$$P_\delta\left(\xi \notin \Xi_2^\delta\right) \leq \sum_{m=2}^{\infty} e^{-n_m/2} \leq e^{-c_8 n_1} \tag{4.6}$$

for some constant $c_8 > 0$, recall our choice of n_m (3.1). Let $m \geq 2$, and recall the definition of the event $E^{n_m,T^m}_{\text{reconstruct}}$ from Theorem 3.5. By Definition 3.4, we have that $\mathcal{A}^m(\tilde{\chi}) = \mathrm{Alg}^{n_m}\left(T^m(\tilde{\chi}), \tilde{\chi}|[0, 2\cdot 2^{12\alpha n_m}[, \psi\right)$ with $\psi := \mathcal{A}^{m-1}(\tilde{\chi})$. By our choice of n_1, $(|\psi|-1)/2 = 3\cdot 2^{n_{m-1}} \geq c_1 n_m L$. If E^{m-1} holds, then $\psi \preceq \xi|\,[-2^{n_m}, 2^{n_m}]$. Hence the inclusion

$$E^{m-1} \cap \left(E^{m,T^m}_{\text{stop}} \setminus E^m\right) \subseteq E^{m,T^m}_{\text{stop}} \setminus E^{n_m,T^m}_{\text{reconstruct}} \tag{4.7}$$

holds. Together with Theorem 3.5 the last inclusion implies

$$P_\delta\left(E^{m-1} \cap \left(E^{n_m,T^m}_{\text{stop}} \setminus E^m\right)\right) \leq P_\delta\left(E^{m,T^m}_{\text{stop}} \setminus E^{n_m,T^m}_{\text{reconstruct}}\right) \leq c_6 e^{-c_7 n_m}.$$

Another application of Lemma 4.2 yields for some constant $c_9 > 0$

$$P_\delta\left(\xi \notin \Xi^\delta_3\right) \leq \sum_{m=2}^{\infty} (c_6)^{1/2} e^{-c_7 n_m/2} \leq e^{-c_9 n_1}. \tag{4.8}$$

The claim of the lemma follows from (4.5), (4.6), and (4.8); recall $\varepsilon_3(n) \to 0$ as $n \to \infty$. ■

Lemma 4.4 *For all $\delta \in]0, \delta_1[$, $\xi \in \Xi^\delta$, and $m \geq 2$ the following holds for some constants $c_4, c_5 > 0$:*

$$P_\delta(E^{m-1} \mid \xi) \geq 1 - (2\varepsilon_3(n_1))^{1/2} - \sum_{k=2}^{m-1} c_4 e^{-c_5 n_k} \geq \frac{1}{2}, \tag{4.9}$$

$$P_\delta(E^{m-1} \setminus E^m \mid \xi) \leq c_4 e^{-c_5 n_m}. \tag{4.10}$$

Proof. Let $\delta \in]0, \delta_1[$ and $\xi \in \Xi^\delta$. We prove (4.9) and (4.10) simultaneously by induction over m: For $m = 2$ it follows from $\xi \in \Xi^\delta_1$

$$P_\delta(E^1 \mid \xi) \;=\; 1 - P_\delta\left[\left(E^1\right)^c \mid \xi\right] \geq 1 - (2\varepsilon_3(n_1))^{1/2} \geq 1/2; \tag{4.11}$$

recall our choice of n_1 from Section 2.1. Thus (4.9) holds for $m = 2$.

Suppose (4.9) holds for some $m \geq 2$. Then we have

$$\begin{aligned} P_\delta[E^{m-1} \setminus E^m|\xi] &\leq P_\delta\left[(E^{m-1} \setminus E^m) \cap E^{m,T^m}_{\text{stop}} \,\middle|\, \xi\right] + P_\delta\left[E^{m-1} \setminus E^{m,T^m}_{\text{stop}} \,\middle|\, \xi\right] \\ &\leq (c_6)^{1/2} e^{-\frac{c_7 n_m}{2}} + e^{-n_m/2} \leq c_4 e^{-c_5 n_m} \end{aligned} \tag{4.12}$$

for some constants $c_4, c_5 > 0$; for the first term we used $\xi \in \Xi^\delta_3$ and for the second term we used $\xi \in \Xi^\delta_2$ and our induction hypothesis (4.9). Using (4.12) and our induction hypothesis (4.9) we obtain

$$\begin{aligned} P_\delta(E^m \mid \xi) &\geq P_\delta(E^{m-1} \mid \xi) - P_\delta(E^{m-1} \setminus E^m \mid \xi) \\ &\geq 1 - (2\varepsilon_3(n_1))^{1/2} - \sum_{k=2}^{m} c_4 e^{-c_5 n_k} \geq \frac{1}{2}; \end{aligned}$$

for the last inequality we used our choice of n_1. This completes the induction step. ■

Proof of Theorem 3.6. Let $\delta \in]0, \delta_1[$; recall our choice $\delta_1 = \min\{\delta_3, \delta_2(n_1)\}$. By Theorem 3.5 we know that whenever the events E^{m-1} and E^m hold and $\xi|[-2^{n_m}, 2^{n_m}] \neq (1)_{[-2^{n_m}, 2^{n_m}]}$, then E^m_{center} holds. Since P_δ-a.s. $\xi \neq (1)_{\mathbb{Z}}$, relation (3.3) holds. Using Lemma 4.3 we have

$$\begin{aligned} P_\delta\left(\bigcup_{m=1}^{\infty}(E^m)^{\mathrm{c}}\right) &\leq P_\delta\left(\xi \notin \Xi^\delta\right) + P_\delta\left(\{\xi \in \Xi^\delta\} \cap \bigcup_{m=1}^{\infty}(E^m)^{\mathrm{c}}\right) \\ &\leq \varepsilon_2(n_1) + \int_{\{\xi \in \Xi^\delta\}} P_\delta\left(\bigcup_{m=1}^{\infty}(E^m)^{\mathrm{c}} \,\middle|\, \xi\right) \mathrm{d}P_\delta. \end{aligned} \tag{4.13}$$

To bound the integrand, we use Lemma 4.4: For all $\xi \in \Xi^\delta$ and $k \geq 1$, we obtain

$$\begin{aligned} P_\delta\left(\bigcup_{m=1}^{k}(E^m)^{\mathrm{c}} \,\middle|\, \xi\right) &\leq P_\delta\left((E^1)^{\mathrm{c}} \mid \xi\right) + \sum_{m=2}^{k+1} P_\delta(E^{m-1} \setminus E^m \mid \xi) \\ &\leq (2\varepsilon_3(n_1))^{1/2} + \sum_{m=2}^{k+1} c_4 e^{-c_5 n_m}, \end{aligned} \tag{4.14}$$

and taking limits as $k \to \infty$, we conclude

$$P_\delta\left(\bigcup_{m=1}^{\infty}(E^m)^{\mathrm{c}} \,\middle|\, \xi\right) \leq (2\varepsilon_3(n_1))^{1/2} + \sum_{m=2}^{\infty} c_4 e^{-c_5 n_m}.$$

Together with (4.13) the last estimate yields (3.4):

$$P_\delta\left(\bigcup_{m=1}^{\infty}(E^m)^{\mathrm{c}}\right) \leq \varepsilon_2(n_1) + (2\varepsilon_3(n_1))^{1/2} + \sum_{m=2}^{\infty} c_4 e^{-c_5 n_m} < \frac{1}{2}; \tag{4.15}$$

for the last inequality we used that n_1 is chosen as in Section 2.1. ■

5 The key algorithm of the reconstruction

In this section, we define algorithms Alg^n for which Theorem 3.5 holds. We fix $n \in \mathbb{N}$.

For two words $w, w' \in \mathcal{C}^*$ of the same length we define their distance

$$d(w, w') := |\{k \in [1, |w|] : w_k \neq w'_k\}|; \tag{5.1}$$

$d(w, w')$ is the number of places where w and w' disagree. Clearly, d is a metric.

When the random walk observes a piece of scenery and δ is small, the observations with errors differ "typically" from the errorfree observations in only a small proportion of the letters because the probability to see an error at a particular time is small under P_δ. Since the random walk observes a given piece of scenery very often, we are able to filter out the errors using a majority rule f^*.

The following notions will be used in this context. For $w = w_1 w_2 \dots w_m \in \mathcal{C}^m$ we define $\mathsf{Cut}(w) := w_2 \dots w_{m-1}$; $\mathsf{Cut}(w)$ is obtained from w by cutting off the first and the last letter.

Definition 5.1 *Let $W = (w_j; 1 \le j \le K) \in (\mathcal{C}^{c_1 n})^K$ be a vector consisting of K words of length $c_1 n$. For $i \in [1, c_1 n]$ we define $f_i(W)$, the* favorite letter at position i, *to be the element in $\mathcal{C}$ which most of the first $2^{\gamma n}$ words in W have at position i. If there is no unique letter with this property, then we define the favorite letter to be the smallest one. Formally, we set*

$$f_i(W) = k \quad \textit{iff} \quad |\{j \in [1, 2^{\gamma n}] : w_j(i) = k\}| = \max_{k' \in \mathcal{C}} |\{j \in [1, 2^{\gamma n}] : w_j(i) = k'\}|$$

and k is the smallest element in $\mathcal{C}$ satisfying the last equality; here $w_j(i)$ denotes the i^{th} letter of the word w_j. We set $f(W) := f_1(W) f_2(W) \dots f_{c_1 n}(W)$. Furthermore, we define $f^*(W) :=$

$$\begin{cases} \mathsf{Cut}(f(W)), & \textit{if } K \ge 2^{\gamma n} \textit{ and } \max_{j \in [1, 2^{\gamma n}]} d(\mathsf{Cut}(w_j), \mathsf{Cut}(f(W))) \le \varepsilon n \\ (-1)_{[1, c_1 n - 2]}, & \textit{otherwise.} \end{cases}$$

$f^*(W)$ equals the word $\mathsf{Cut}(f(W))$ which is composed of the favorite letters iff the vector W has sufficiently many components and each of the first $2^{\gamma n}$ words in W differs from $f(W)$ in not more than εn letters. In the proof of Lemma 6.9 below it will be essential that we use $\mathsf{Cut}(f(W))$ and not $f(W)$ in the definition of $f^*(W)$. Note that $-1 \notin \mathcal{C}$ so that $(-1)_{[1, c_1 n - 2]}$ differs from all words $w \in \mathcal{C}^{c_1 n - 2}$.

The algorithm Alg^n which will be defined below takes input data

$$\tau \in \left[0, 2^{12\alpha n}\right]^{\mathbb{N}}, \; \eta \in \mathcal{C}^{2 \cdot 2^{12\alpha n}}, \text{ and } \psi \in \bigcup_{k \ge c_1 L} \mathcal{C}^{[-kn, kn]}. \tag{5.2}$$

First we define the set of all observations of length $3c_1 n$ which are collected within a time horizon of length 2^{2n} after a time $\tau_k, k \in [1, 2^{\alpha n}]$:

Definition 5.2 *We define $\mathsf{Collection}^n(\tau, \eta) :=$*

$$\left\{ (w_1, w_2, w_3) \in (\mathcal{C}^{c_1 n})^3 : \exists k \in [1, 2^{\alpha n}] \textit{ such that } w_1 w_2 w_3 \sqsubseteq \eta | [\tau_k, \tau_k + 2^{2n} [\right\}.$$

The set $\mathsf{PrePuzzle}^n(\tau, \eta)$ contains only $(w_1, w_2, w_3) \in \mathsf{Collection}^n(\tau, \eta)$ with the following property: If $(w_1', w_2', w_3') \in \mathsf{Collection}^n(\tau, \eta)$ and w_1' and w_3' are "not too different" from w_1 and w_3 respectively, then w_2' is "not too different" from w_2. Formally:

Definition 5.3 *We define $\mathsf{PrePuzzle}^n(\tau, \eta) :=$*

$$\left\{ \begin{array}{l} (w_1, w_2, w_3) \in \mathsf{Collection}^n(\tau, \eta) : \textit{If } (w_1', w_2', w_3') \in \mathsf{Collection}^n(\tau, \eta) \textit{ with} \\ d(w_1, w_1') \le 2\varepsilon n \textit{ and } d(w_3, w_3') \le 2\varepsilon n, \textit{ then } d(w_2, w_2') \le 2\varepsilon n. \end{array} \right\}.$$

Definition 5.4 *For an element* $(w_1, w_2, w_3) \in \mathsf{PrePuzzle}^n(\tau, \eta)$ *we denote by* $\mathcal{S}^n_{\tau,\eta}(w_1, w_2, w_3)$ *the sequence of (random) times* $s \in \cup_{k=1}^{2^{\alpha n}} \left[\tau_k, \tau_k + 2^{2n} - 3c_1 n\right]$ *such that* $w_1' w_2' w_3' := \eta|[s, s + 3c_1 n[\in \mathsf{PrePuzzle}^n(\tau, \eta)$, $d(w_1, w_1') \le 2\varepsilon n$, *and* $d(w_3, w_3') \le 2\varepsilon n$; *we assume that the elements of the sequence* $\mathcal{S}^n_{\tau,\eta}(w_1, w_2, w_3)$ *are arranged in increasing order. We define*

$$\mathsf{List}^n_{\tau,\eta}(w_1, w_2, w_3) := \left(\eta|[s + c_1 n, s + 2c_1 n[; s \in \mathcal{S}^n_{\tau,\eta}(w_1, w_2, w_3)\right)$$

to be the sequence with components $\eta|[s + c_1 n, s + 2c_1 n[$ *indexed by the set* $\mathcal{S}^n_{\tau,\eta}(w_1, w_2, w_3)$. *We set*

$$\mathsf{PuzzleLists}^n(\tau, \eta) := \left\{\mathsf{List}^n_{\tau,\eta}(w_1, w_2, w_3) : (w_1, w_2, w_3) \in \mathsf{PrePuzzle}^n(\tau, \eta)\right\}.$$

Clearly, $w_2 \in \mathsf{List}^n_{\tau,\eta}(w_1, w_2, w_3)$. Note that $\mathsf{List}^n_{\tau,\eta}(w_1, w_2, w_3)$ is a sequence, and not a set. If by coincidence observations $\eta|[s + c_1 n, s + 2c_1 n[$ coincide for two different values of s, we want to keep them both. The components of $\mathsf{List}^n_{\tau,\eta}(w_1, w_2, w_3)$ are close to w_2 in d-distance because we assumed $(w_1, w_2, w_3) \in \mathsf{PrePuzzle}^n(\tau, \eta)$.

Definition 5.5 *We define* $\mathsf{Puzzle}^n(\tau, \eta) := \{f^*(W) : W \in \mathsf{PuzzleLists}^n(\tau, \eta)\}$.

$\mathsf{Puzzle}^n(\tau, \eta)$ is the set of all words of length $c_1 n - 2$ which are obtained by the majority rule f^* from the lists in $\mathsf{PuzzleLists}^n(\tau, \eta)$. We use the words in $\mathsf{Puzzle}^n(\tau, \eta)$ like the pieces in a puzzle game to reconstruct a piece of scenery. We want the piece of scenery reconstructed by Alg^n to contain in the middle the piece of scenery ψ from the input data of the algorithm.

Definition 5.6 *For* $\psi \in \mathcal{C}^{[-kn,kn]}$ *we define* $\mathsf{SolutionPiece}^n(\tau, \eta, \psi) :=$

$$\left\{\begin{matrix} w \in \mathcal{C}^{[-3\cdot 2^n, 3\cdot 2^n]} : w|[-kn, kn] = \psi \text{ and for all ladder intervals } I \subseteq \\ [-3\cdot 2^n, 3\cdot 2^n] \text{ with } |I| = c_1 n - 2 \text{ we have } (w|I)_{\to} \in \mathsf{Puzzle}^n(\tau, \eta) \end{matrix}\right\}.$$

We will see in the proof of Lemma 6.4 below that under appropriate conditions, there is precisely one element in $\mathsf{SolutionPiece}^n(\tau, \eta, \psi)$.

Definition 5.7 *We define*

$$Alg^n : [0, 2^{12\alpha n}]^{\mathbb{N}} \times \mathcal{C}^{2\cdot 2^{12\alpha n}} \times \bigcup_{k \ge c_1 L} \mathcal{C}^{[-kn,kn]} \to \mathcal{C}^{[-3\cdot 2^n, 3\cdot 2^n]}$$

as follows: If $\mathsf{SolutionPiece}^n(\tau, \eta, \psi)$ *is not empty, then we define* $Alg^n(\tau, \eta, \psi)$ *to be its lexicographically smallest element. Otherwise we define* $Alg^n(\tau, \eta, \psi)$ *to be the constant scenery* $(1)_{[-3\cdot 2^n, 3\cdot 2^n]}$.

6 The key algorithm reconstructs correctly

In this section, we prove Theorem 3.5. Throughout we fix $n \in \mathbb{N}$. We assume that $\tau \in [0, 2^{12\alpha n}]^{\mathbb{N}}$ is a sequence of $\mathcal{G}$-adapted stopping times. Recall that ε was chosen in Section 2.1.

6.1 Definition of the key events

In this subsection, we collect the definitions of all the "basic" events which we will need to prove the correctness of Alg^n. The event $B^{n,\tau}_{\text{all paths}}$ holds if the random walk traverses all paths of length $3c_1n$ in the region where we want to do the reconstruction. $B^n_{\text{few mistakes}}$ makes sure that there are not too many mistakes in the words in $\mathsf{Collection}^n(\tau,\eta)$. $B^n_{\text{ladder diff}}$ gives a lower bound for the d-distance of two different ladder words in the neighborhood of the origin. $B^{n,\tau}_{\text{majority}}$ garanties that the majority decision f^* is not corrupted by the errors in the observations. If $B^n_{\text{outside out}}$ holds, then we can distinguish ladder words from the region where we want to reconstruct from observations which are read further outside. B^n_{signals} implies that there are "signal words" which can be read only left from a certain point $z\in\mathbb{Z}$ or only right from a certain $z\in\mathbb{Z}$; this event allows use to reconstruct all ladder words in a region around the origin. $B^{n,\tau}_{\text{straight often}}$ guarantees that certain ladder paths are traversed often enough.

We arranged the definitions of the events in alphabetical order so that the reader can easily find them while following the proofs in the next two subsections. We suggest to have a quick look at the definitions, and then to skip ahead to the next subsection and look up definitions when needed.

Definition 6.1 *For $z\in\mathbb{Z}$ and n such that $c_1n\in\mathbb{N}$, we denote by $w_{z,\rightarrow,n}$ the ladder word of length c_1n starting at z read from left to right, and by $w_{z,\leftarrow,n}$ the word $w_{z,\rightarrow,n}$ read from right to left:*

$$w_{z,\rightarrow,n}:=(\xi(z+kL);k\in[0,c_1n[)_{\rightarrow}\quad\textit{and}\quad w_{z,\leftarrow,n}:=(w_{z,\rightarrow,n})_{\leftarrow}.$$

Note that $w_{z-(c_1n-1)L,\rightarrow,n}$ is the ladder word of length c_1n ending at z.

Definition 6.2 *We define*

$$B^{n,\tau}_{\text{all paths}}:=\left\{\begin{array}{l}\textit{For any admissible piece of path } R\in\mathbb{Z}^{[0,3c_1n[}\textit{ with starting}\\ \textit{point in }[-7\cdot2^n,7\cdot2^n]\textit{ there exists }t\in\cup_{k=1}^{2^{\alpha n}}[\tau_k,\tau_k+2^{2n}-\\ 3c_1n]\textit{ such that }R(i)=S(t+i)\textit{ for all }i\in[0,3c_1n[\end{array}\right\}.$$

Definition 6.3 *We define*

$$B^n_{\text{few mistakes}}:=\left\{\sum_{k=t-c_1n+1}^{t}X_k\le\varepsilon n\textit{ for all }t\in\left[c_1n-1,2\cdot2^{12\alpha n}\right[\right\}.$$

Definition 6.4 *We define*

$$B^n_{\text{ladder diff}}:=\left\{\begin{array}{l}\forall z_1,z_2\in[-8\cdot2^n,8\cdot2^n]\textit{ and }\forall i_1,i_2\in\{\leftarrow,\rightarrow\}\textit{ with}\\ (z_1,i_1)\neq(z_2,i_2)\textit{ we have }d(w_{z_1,i_1,n/3},w_{z_2,i_2,n/3})\ge10\varepsilon n\end{array}\right\}.$$

Definition 6.5 *Let $\mathcal{I}_L$ denote the set of ladder intervals $I\subseteq[-7\cdot2^n,7\cdot2^n]$ of length c_1n. For $w_1,w_3\in\mathcal{C}^{c_1n}$ and $I\in\mathcal{I}_L$, we denote by $\mathcal{S}^{I\rightarrow}_{w_1,w_3}:=\left(s^{I\rightarrow}_i;i\ge1\right)$ ($\mathcal{S}^{I\leftarrow}_{w_1,w_3}:=\left(s^{I\leftarrow}_i;i\ge1\right)$) the sequence of all times $s\in\cup_{k=1}^{2^{\alpha n}}\left[\tau_k,\tau_k+2^{2n}-3c_1n\right]$*

such that $S|[s+c_1n, s+2c_1n[$ *is a straight crossing from left to right (right to left) of* I *and* $d(\tilde{\chi}|[s+(i-1)c_1n, s+ic_1n[, w_i) \leq 2\varepsilon n$ *for* $i = 1, 3$. *We assume that the components of* $\mathcal{S}^{I_\to}_{w_1,w_3}$ *and* $\mathcal{S}^{I_\leftarrow}_{w_1,w_3}$ *are arranged in increasing order. We define*

$$B^{n,\tau}_{\text{majority}} := \bigcap_{w_1,w_3 \in \mathcal{C}^{c_1 n}} \bigcap_{I \in \mathcal{I}_L} \left(B^{n,\tau,I_\to}_{\text{maj}}(w_1,w_3) \cap B^{n,\tau,I_\leftarrow}_{\text{maj}}(w_1,w_3) \right) \text{ with}$$

$$B^{n,\tau,I_\to}_{\text{maj}}(w_1,w_3) := \left\{ \begin{array}{l} \text{If } \left|\mathcal{S}^{I_\to}_{w_1,w_3}\right| \geq 2^{\gamma n}, \text{ then } \forall j \in [1, c_1n-1[\text{ the} \\ \text{following holds: } \sum_{i=1}^{2^{\gamma n}} X_{s_i^{I_\to}+c_1n+j} < 2^{\gamma n}/2 \end{array} \right\}$$

and $B^{n,\tau,I_\leftarrow}_{\text{maj}}(w_1,w_3)$ *defined analogously.*

Definition 6.6 *We define* $B^n_{\text{outside out}} :=$

$$\left\{ \begin{array}{l} \forall z \in [-5\cdot 2^n, 5\cdot 2^n], \text{ for any admissible piece of path } R \in ([-2L\cdot 2^{2n}, 2L\cdot \\ 2^{2n}] \setminus [-6\cdot 2^n,\ 6\cdot 2^n])^{[0,c_1n/2[} \text{ and } \forall i \in \{\leftarrow, \rightarrow\} \text{ we have that } d(\xi \circ \\ R, w_{z,i,n/2}) \geq 3\varepsilon n \end{array} \right\}.$$

Definition 6.7 *We define* $B^n_{\text{recogn straight}} :=$

$$\left\{ \begin{array}{l} \text{For any admissible piece of path } R_1 \in [-7\cdot 2^n, 7\cdot 2^n]^{[0,c_1n[} \text{ which is} \\ \text{not a ladder path there exists an admissible piece of path } R_2 \in [-8\cdot \\ 2^n, 8\cdot 2^n]^{[0,c_1n[} \text{ with } R_2(0) = R_1(0),\ R_2(c_1n-1) = R_1(c_1n-1) \text{ and} \\ d(\xi\circ R_1, \xi\circ R_2) \geq 5\varepsilon n \end{array} \right\}.$$

Definition 6.8 *We define*

$$\begin{array}{rcl} B^n_{\text{signals}} & := & B^n_{\text{sign,l},\to} \cap B^n_{\text{sign,r},\to} \cap B^n_{\text{sign,l},\leftarrow} \cap B^n_{\text{sign,r},\leftarrow} \quad \text{with} \\ B^n_{\text{sign,l},\to} & := & \left\{ \begin{array}{l} \forall z \in [-6\cdot 2^n, 6\cdot 2^n] \text{ and for any admissible piece of path} \\ R \in [-2L\cdot 2^{2n}, 2L\cdot 2^{2n}]^{[0,c_1n[} \text{ with } R(c_1n-1) > z \text{ we} \\ \text{have that } d(\xi\circ R, w_{z-(c_1n-1)L,\to,n}) \geq 5\varepsilon n \end{array} \right\}, \\ B^n_{\text{sign,r},\to} & := & \left\{ \begin{array}{l} \forall z \in [-6\cdot 2^n, 6\cdot 2^n] \text{ and for any admissible piece of path} \\ R \in [-2L\cdot 2^{2n}, 2L\cdot 2^{2n}]^{[0,c_1n[} \text{ with } R(0) < z \text{ we have} \\ \text{that } d(\xi\circ R, w_{z,\to,n}) \geq 5\varepsilon n \end{array} \right\}, \\ B^n_{\text{sign,l},\leftarrow} & := & \left\{ \begin{array}{l} \forall z \in [-6\cdot 2^n, 6\cdot 2^n] \text{ and for any admissible piece of path} \\ R \in [-2L\cdot 2^{2n}, 2L\cdot 2^{2n}]^{[0,c_1n[} \text{ with } R(0) > z \text{ we have} \\ \text{that } d(\xi\circ R, w_{z-(c_1n-1)L,\leftarrow,n}) \geq 5\varepsilon n \end{array} \right\}, \\ B^n_{\text{sign,r},\leftarrow} & := & \left\{ \begin{array}{l} \forall z \in [-6\cdot 2^n, 6\cdot 2^n] \text{ and for any admissible piece of path} \\ R \in [-2L\cdot 2^{2n}, 2L\cdot 2^{2n}]^{[0,c_1n[} \text{ with } R(c_1n-1) < z \text{ we} \\ \text{have that } d(\xi\circ R, w_{z,\leftarrow,n}) \geq 5\varepsilon n \end{array} \right\}. \end{array}$$

Definition 6.9 *We denote the collection of ladder intervals* $I \subseteq [-6\cdot 2^n, 6\cdot 2^n]$ *of length* $3c_1n$ *by* $\mathcal{J}_L$. *For* $I \in \mathcal{J}_L$, *we denote by* $\mathcal{S}_\to(I)$ $(\mathcal{S}_\leftarrow(I))$ *the sequence of all times* $s \in \cup_{k=1}^{2^{\alpha n}} [\tau_k, \tau_k + 2^{2n} - 3c_1n]$ *such that* $S|[s, s+3c_1n[$ *is a straight* *crossing from left to right (right to left) of I; we assume that the components of $\mathcal{S}_{\rightarrow}(I)$ and $\mathcal{S}_{\leftarrow}(I)$ are arranged in increasing order. We define*

$$B^{n,\tau}_{\text{straight often}} := \bigcap_{I \in \mathcal{J}_L} \{|\mathcal{S}_{\rightarrow}(I)| \geq 2^{\gamma n} \text{ and } |\mathcal{S}_{\leftarrow}(I)| \geq 2^{\gamma n}\}.$$

6.2 Combinatorics

In this subsection, we prove that Alg^n reconstructs correctly in the sense that the event $E^{n,\tau}_{\text{reconstruct}}$ holds, under the assumption that $E^{n,\tau}_{\text{stop}}$ and all the "basic" events defined in the previous subsection hold. We abbreviate

$$\tilde{\chi}^n := \tilde{\chi}|\left[0, 2 \cdot 2^{12\alpha n}\right[.$$

The task is split in four parts: Lemma 6.1 states a property of the elements in the set $\mathsf{PrePuzzle}^n(\tau, \tilde{\chi}^n)$. Lemma 6.2 shows that all words in $\mathsf{Puzzle}^n(\tau, \tilde{\chi}^n)$ which are observed while the random walk is approximately in the region of the scenery which we want to reconstruct, are ladder words. Lemma 6.3 states that $\mathsf{Puzzle}^n(\tau, \tilde{\chi}^n)$ contains all the ladder words we need. Finally Lemma 6.4 shows that the reconstruction works.

Definition 6.10 *We say $(w_1, w_2, w_3) \in \mathsf{Collection}^n(\tau, \tilde{\chi}^n)$ is* read while the random walk is walking on $J \subseteq \mathbb{Z}$ *if there exists $t \in \cup_{k=1}^{2^{\alpha n}} \left[\tau_k, \tau_k + 2^{2n} - 3c_1 n\right]$ such that $S(t+j) \in J$ for all $j \in [0, 3c_1 n[$ and $w_1 w_2 w_3 = \tilde{\chi}|[t, t+3c_1 n[$. If we know the time t, we say that (w_1, w_2, w_3) is* read during $[t, t+3c_1 n[$.

Definition 6.11 *We define $E^{n,\tau}_{\text{pre ladder}} :=$*

$$\left\{\begin{array}{l}\textit{If } (w_1, w_2, w_3) \in \mathsf{PrePuzzle}^n(\tau, \tilde{\chi}^n) \textit{ and there exists } t \in \cup_{k=1}^{2^{\alpha n}}[\tau_k, \tau_k + \\ 2^{2n} - 3c_1 n] \textit{ such that } (w_1, w_2, w_3) \textit{ is read during } [t, t+3c_1 n[\textit{ while the} \\ \textit{random walk is walking on } [-7 \cdot 2^n, 7 \cdot 2^n]\textit{, then } S|[t+c_1 n, t+2c_1 n[\textit{ is a} \\ \textit{ladder path.}\end{array}\right\}.$$

Lemma 6.1 *For all $n \in \mathbb{N}$ the following holds:*

$$E^{n,\tau}_{\text{pre ladder}} \supseteq B^{n,\tau}_{\text{all paths}} \cap B^n_{\text{few mistakes}} \cap B^n_{\text{recogn straight}}.$$

Proof. Suppose the events $B^{n,\tau}_{\text{all paths}}$, $B^n_{\text{few mistakes}}$, and $B^n_{\text{recogn straight}}$ hold. Let $(w_1, w_2, w_3) \in \mathsf{PrePuzzle}^n(\tau, \tilde{\chi}^n)$, and suppose there exists $t \in \cup_{k=1}^{2^{\alpha n}}[\tau_k, \tau_k + 2^{2n} - 3c_1 n]$ such that the triple (w_1, w_2, w_3) is read during $[t, t+3c_1 n[$ while the random walk is walking on $[-7 \cdot 2^n, 7 \cdot 2^n]$.

Let $R_i(j) := S(t + (i-1)c_1 n + j)$ for $j \in [0, c_1 n[$ and $i = 1, 2, 3$. Then $|R_i(j)| \leq 7 \cdot 2^n$ for all $j \in [0, c_1 n[$ and

$$d(\xi \circ R_i, w_i) \leq \varepsilon n \quad \text{for } i = 1, 2, 3 \tag{6.1}$$

because $B^n_{\text{few mistakes}}$ holds. We have to show that R_2 is a ladder path. Suppose not. Since $B^n_{\text{recogn straight}}$ holds, there exists an admissible piece of path $R_2' \in [-8 \cdot 2^n, 8 \cdot 2^n]^{[0, c_1 n[}$ with the same starting and endpoint as R_2 and

$$d(\xi \circ R_2, \xi \circ R_2') \geq 5\varepsilon n. \tag{6.2}$$

Since $B^{n,\tau}_{\text{all paths}}$ holds and the concatenation $R_1R_2'R_3$ is an admissible piece of path with starting point in $[-7\cdot 2^n, 7\cdot 2^n]$, there exists $t' \in \cup_{k=1}^{2^{\alpha n}}[\tau_k, \tau_k + 2^{2n} - 3c_1 n]$ such that $R_1R_2'R_3(i) = S(t'+i)$ for all $i \in [0, 3c_1n[$. Using the triangle inequality, we obtain

$$\begin{aligned} d(w_2, \tilde{\chi}|[t'+c_1n, t'+2c_1n[) &\geq d(w_2, \chi|[t'+c_1n, t'+2c_1n[) - \varepsilon n \\ &= d(w_2, \xi\circ R_2') - \varepsilon n \\ &\geq d(\xi\circ R_2, \xi\circ R_2') - d(w_2, \xi\circ R_2) - \varepsilon n \\ &\geq 5\varepsilon n - \varepsilon n - \varepsilon n = 3\varepsilon n; \end{aligned} \tag{6.3}$$

for the first inequality we used that $B^n_{\text{few mistakes}}$ holds, and for the last inequality we used (6.2) and (6.1). The fact that $B^n_{\text{few mistakes}}$ holds together with inequality (6.1) yields

$$\begin{aligned} d(w_1, \tilde{\chi}|[t', t'+c_1n[) &\leq d(w_1, \chi|[t', t'+c_1n[) + \varepsilon n \\ &= d(w_1, \xi\circ R_1) + \varepsilon n \leq 2\varepsilon n. \end{aligned}$$

By the same argument, $d(w_3, \tilde{\chi}|[t'+2c_1n, t'+3c_1n[) \leq 2\varepsilon n$. Together with (6.3) this contradicts $(w_1, w_2, w_3) \in \mathsf{PrePuzzle}^n(\tau, \tilde{\chi}^n)$. Hence R_2 is a ladder path. ■

Definition 6.12 *We define*

$$\begin{aligned} \mathit{Puzzle}_1^n(\tau, \tilde{\chi}^n) &:= \left\{ \begin{array}{l} f^*\left(\mathit{List}^n_{\tau,\tilde{\chi}^n}(w_1, w_2, w_3)\right) \in \mathcal{C}^{c_1n-2} : (w_1, w_2, w_3) \in \\ \mathit{PrePuzzle}^n(\tau, \tilde{\chi}^n) \text{ and } \exists (w_1', w_2', w_3') \in \\ \mathit{PrePuzzle}^n(\tau, \tilde{\chi}^n) \text{ such that } d(w_1, w_1') \leq 2\varepsilon n, \\ d(w_3, w_3') \leq 2\varepsilon n \text{ and } (w_1', w_2', w_3') \text{ is read while the} \\ \text{random walk is walking on } \mathbb{Z}\setminus[-6\cdot 2^n, 6\cdot 2^n]. \end{array} \right\}, \\ \mathit{Puzzle}_2^n(\tau, \tilde{\chi}^n) &:= \mathit{Puzzle}^n(\tau, \tilde{\chi}^n) \setminus \left[\mathit{Puzzle}_1^n(\tau, \tilde{\chi}^n) \cup \left\{(-1)_{[1,c_1n-2]}\right\}\right]. \end{aligned}$$

Note that $\mathsf{Puzzle}_i^n(\tau, \tilde{\chi}^n)$, $i = 1, 2$, together with $\left\{(-1)_{[1,c_1n-2]}\right\}$, form a partition of the set $\mathsf{Puzzle}^n(\tau, \tilde{\chi}^n)$. If we are given an element of $\mathsf{Puzzle}^n(\tau, \tilde{\chi}^n)$, we cannot decide to which set of the partition it belongs. Nevertheless the sets $\mathsf{Puzzle}_i^n(\tau, \tilde{\chi}^n)$, $i = 1, 2$, will be useful in the following.

Definition 6.13 *We define*

$$E^{n,\tau}_{\text{only ladder}} := \left\{ \begin{array}{l} \text{If } w_2 \in \mathit{Puzzle}_2^n(\tau, \tilde{\chi}^n), \text{ then } w_2 \preceq \xi|\,[-7\cdot 2^n, 7\cdot 2^n] \\ \text{and } w_2 \text{ is a ladder word} \end{array} \right\}.$$

Let $c_{10} > 0$ be chosen in such a way that for all $n \geq c_{10}$

$$3c_1nL \leq 2^n. \tag{6.4}$$

Lemma 6.2 *For all $n \geq c_{10}$ the following holds:*

$$E^{n,\tau}_{\text{only ladder}} \supseteq E^{n,\tau}_{\text{pre ladder}} \cap B^n_{\text{few mistakes}} \cap B^n_{\text{ladder diff}} \cap B^{n,\tau}_{\text{majority}}.$$

Proof. Let $n \geq c_{10}$, and suppose the events $E^{n,\tau}_{\text{pre ladder}}$, $B^n_{\text{few mistakes}}$, $B^n_{\text{ladder diff}}$ and $B^{n,\tau}_{\text{majority}}$ hold. Let $(w_1, w_2, w_3) \in \mathsf{PrePuzzle}^n(\tau, \tilde{\chi}^n)$ and abbreviate $W := \mathsf{List}^n_{\tau,\tilde{\chi}^n}(w_1, w_2, w_3)$. Suppose $f^*(W) \in \mathsf{Puzzle}^n_2(\tau, \tilde{\chi}^n)$. Let $w'_2 \in W$. Then there exist w'_1, w'_3 such that $(w'_1, w'_2, w'_3) \in \mathsf{PrePuzzle}^n(\tau, \tilde{\chi}^n)$, $d(w_1, w'_1) \leq 2\varepsilon n$, and $d(w_3, w'_3) \leq 2\varepsilon n$. By definition of $\mathsf{Puzzle}^n_2(\tau, \tilde{\chi}^n)$, at least once the random walk is in $[-6 \cdot 2^n, 6 \cdot 2^n]$ while it reads (w'_1, w'_2, w'_3). Since the random walk jumps at most a distance of L in each step, it can move in $3c_1 n$ steps at most a distance of $3c_1 nL \leq 2^n$. Hence (w'_1, w'_2, w'_3) is observed while the random walk is walking on $[-7 \cdot 2^n, 7 \cdot 2^n]$. Using that $E^{n,\tau}_{\text{pre ladder}}$ holds, we obtain that w'_2 is observed while the random walk is walking on a ladder word. Since $B^n_{\text{few mistakes}}$ holds, there exists a ladder word $\widehat{w}_2 \preceq \xi |\, [-7 \cdot 2^n, 7 \cdot 2^n]$ such that

$$d(w'_2, \widehat{w}_2) \leq \varepsilon n. \tag{6.5}$$

Suppose $w''_2 \in W$. Then by the above argument, there exists a ladder word $\bar{w}_2 \preceq \xi |\, [-7 \cdot 2^n, 7 \cdot 2^n]$ such that

$$d(w''_2, \bar{w}_2) \leq \varepsilon n. \tag{6.6}$$

Since $(w_1, w_2, w_3) \in \mathsf{PrePuzzle}^n(\tau, \tilde{\chi}^n)$, we have that $d(w'_2, w_2) \leq 2\varepsilon n$ and $d(w_2, w''_2) \leq 2\varepsilon n$. Hence

$$d(w'_2, w''_2) \leq 4\varepsilon n. \tag{6.7}$$

Using the triangle inequality, (6.5), (6.7) and (6.6) we obtain

$$\begin{aligned} d(\widehat{w}_2, \bar{w}_2) &\leq d(\widehat{w}_2, w'_2) + d(w'_2, w''_2) + d(w''_2, \bar{w}_2) \\ &\leq \varepsilon n + 4\varepsilon n + \varepsilon n = 6\varepsilon n. \end{aligned} \tag{6.8}$$

If $\widehat{w}_2 \neq \bar{w}_2$, then it follows from $B^n_{\text{ladder diff}}$ that $d(\widehat{w}_2, \bar{w}_2) \geq 10\varepsilon n$, which contradicts (6.8). Hence $\widehat{w}_2 = \bar{w}_2$.

We have shown that any $w'_2 \in W$ is observed while the random walk reads the ladder word $\widehat{w}_2$. Hence for $j \in [0, c_1 n[$, $w'_2(j)$ equals $\widehat{w}_2(j)$ or an error in the observations. Since by assumption, $f^*(W) \neq (-1)_{[1, c_1 n - 2]}$, W has at least $2^{\gamma n}$ components; recall the definition of f^* (Definition 5.1). An application of $B^{n,\tau,I}_{\text{maj}}(w_1, w_3)$ with I equal to the ladder interval underlying $\widehat{w}_2$ shows that more than half of the first $2^{\gamma n}$ words in W have j^{th} letter equal to $\widehat{w}_2(j)$. Consequently, $f(W) = \widehat{w}_2$, and since $B^n_{\text{few mistakes}}$ holds, $f^*(W) = \mathsf{Cut}(\widehat{w}_2)$. ■

Definition 6.14 *We define* $E^{n,\tau}_{\text{all ladder}} :=$

$$\left\{ \forall z \in [-5 \cdot 2^n, 5 \cdot 2^n] : \mathit{Cut}(w_{z,\rightarrow,n}), \mathit{Cut}(w_{z,\leftarrow,n}) \in \mathit{Puzzle}^n(\tau, \tilde{\chi}^n) \right\}.$$

Lemma 6.3 *For all $n \geq c_{10}$ the following holds:*

$$\begin{aligned} E^{n,\tau}_{\text{all ladder}} \supseteq\ & B^{n,\tau}_{\text{all paths}} \cap B^n_{\text{few mistakes}} \cap B^{n,\tau}_{\text{majority}} \cap B^n_{\text{signals}} \\ & \cap B^{n,\tau}_{\text{straight often}} \cap E^{n,\tau}_{\text{stop}}. \end{aligned}$$

Proof. Let $n \geq c_{10}$ and $z \in [-5 \cdot 2^n, 5 \cdot 2^n]$. Suppose the events $B^{n,\tau}_{\text{all paths}}$, $B^{n}_{\text{few mistakes}}$, $B^{n,\tau}_{\text{majority}}$, B^{n}_{signals}, $B^{n,\tau}_{\text{straight often}}$, and $E^{n,\tau}_{\text{stop}}$ hold. We will prove $\mathsf{Cut}(w_{z,\rightarrow,n}) \in \mathsf{Puzzle}^n(\tau, \tilde{\chi}^n)$. The proof for $w_{z,\leftarrow,n}$ is similar. We define

$$w_1 := w_{z-c_1nL,\rightarrow,n}, \quad w_2 := w_{z,\rightarrow,n}, \quad w_3 := w_{z+c_1nL,\rightarrow,n}.$$

Clearly, $w_1w_2w_3$ is the ladder word of length $3c_1n$ starting at $z - c_1nL$ and ending at $z + (2c_1n - 1)L$. We define $R : [0, 3c_1n[\rightarrow \mathbb{Z}$ by $R(i) = z - c_1nL + iL$. Then R is a ladder path with starting point $z - c_1nL \geq -6 \cdot 2^n$ and endpoint $z + (2c_1n - 1)L \leq 6 \cdot 2^n$ by our choice of z and n; recall (6.4). Furthermore $\xi \circ R = w_1w_2w_3$. Since $B^{n,\tau}_{\text{all paths}}$ holds, there exists $t \in \cup_{k=1}^{2^{\alpha n}} [\tau_k, \tau_k + 2^{2n} - 3c_1n]$ such that $R = S|[t, t + 3c_1n[$. We set

$$\widehat{w}_{i,t} := \tilde{\chi}|[t + (i-1)c_1n, t + ic_1n[\quad \text{for } i = 1,2,3. \tag{6.9}$$

Since $B^{n,\tau}_{\text{straight often}}$ holds, there are at least $2^{\gamma n}$ different t's with this property. Fix t. Clearly, $(\widehat{w}_{1,t}, \widehat{w}_{2,t}, \widehat{w}_{3,t}) \in \mathsf{Collection}^n(\tau, \tilde{\chi}^n)$. We want to show $(\widehat{w}_{1,t}, \widehat{w}_{2,t}, \widehat{w}_{3,t}) \in \mathsf{PrePuzzle}^n(\tau, \tilde{\chi}^n)$. The word $\widehat{w}_{i,t}$ differs from w_i only by errors in the observations. Since $B^n_{\text{few mistakes}}$ holds,

$$d(w_i, \widehat{w}_{i,t}) \leq \varepsilon n \quad \text{for } i = 1,2,3. \tag{6.10}$$

Suppose $(w'_1, w'_2, w'_3) \in \mathsf{Collection}^n(\tau, \tilde{\chi}^n)$ and $d(w'_i, \widehat{w}_{i,t}) \leq 2\varepsilon n$ for $i = 1, 3$. Then there exists $t' \in \cup_{k=1}^{2^{\alpha n}} [\tau_k, \tau_k + 2^{2n} - 3c_1n]$ such that $w'_1w'_2w'_3 = \tilde{\chi}|[t', t' + 3c_1n[$. Using (6.10) and the triangle inequality, we obtain

$$d(w'_i, w_i) \leq d(w'_i, \widehat{w}_{i,t}) + d(\widehat{w}_{i,t}, w_i) \leq d(w'_i, \widehat{w}_{i,t}) + \varepsilon n \leq 3\varepsilon n \quad \text{for } i = 1, 3.$$

We set $I_1 := [t', t' + c_1n[$, $I_3 := [t' + 2c_1n, t' + 3c_1n[$. Since $B^n_{\text{few mistakes}}$ holds,

$$\begin{aligned} d(\xi \circ S|I_i, w_i) &\leq d(\xi \circ S|I_i, w'_i) + d(w'_i, w_i) \\ &\leq \varepsilon n + d(w'_i, w_i) \leq 4\varepsilon n \quad \text{for } i = 1, 3. \end{aligned} \tag{6.11}$$

Since $E^{n,\tau}_{\text{stop}}$ holds, $|S(\tau_k)| \leq 2^n$, and for all $i \in [0, 2^{2n}[$, $|S(\tau_k + i)| \leq 2^n + L \cdot 2^{2n} \leq 2L \cdot 2^{2n}$ because each jump of the random walk has length $\leq L$. Hence we can use that $B^n_{\text{sign,l},\rightarrow}$ holds for $w_1 = w_{z-c_1nL,\rightarrow,n}$ (note that $|z - L| \leq 6 \cdot 2^n$) and $S|I_1$ to conclude from (6.11) that $S(t' + c_1n - 1) \leq z - L$. Similarly, we can use that $B^n_{\text{sign,r},\rightarrow}$ holds for $w_3 = w_{z+c_1nL,\rightarrow,n}$ (note that $|z + c_1nL| \leq 6 \cdot 2^n$) and $S|I_3$ to conclude that $S(t' + 2c_1n) \geq z + c_1nL$. The only path of length $c_1n + 2$ from $z - L$ to $z + c_1nL$ is the ladder path which visits precisely the points $z + iL$, $0 \leq i \leq c_1n - 1$. Hence w'_2 is observed with errors by the random walk walking on the ladder word w_2. Using the fact that $B^n_{\text{few mistakes}}$ holds and (6.10), we obtain

$$d(w'_2, \widehat{w}_{2,t}) \leq d(w'_2, w_2) + d(w_2, \widehat{w}_{2,t}) \leq \varepsilon n + \varepsilon n = 2\varepsilon n.$$

Consequently, $(\widehat{w}_{1,t}, \widehat{w}_{2,t}, \widehat{w}_{3,t}) \in \mathsf{PrePuzzle}^n(\tau, \tilde{\chi}^n)$. We set

$$W := \mathsf{List}^n_{\tau,\tilde{\chi}^n}(\widehat{w}_{1,t}, \widehat{w}_{2,t}, \widehat{w}_{3,t}).$$

Clearly, $W \in \mathsf{PuzzleLists}^n(\tau, \tilde{\chi}^n)$. Consider $\widehat{w}_{i,s}$ for $s \neq t$. Recall that there are at least $2^{\gamma n} - 1$ different s with this property. By the triangle inequality and (6.10), $d(\widehat{w}_{i,s}, \widehat{w}_{i,t}) \leq d(\widehat{w}_{i,s}, w_i) + d(w_i, \widehat{w}_{i,t}) \leq 2\varepsilon n$ for $i = 1, 2, 3$. Consequently, $(\widehat{w}_{1,s}, \widehat{w}_{2,s}, \widehat{w}_{3,s}) \in W$, and we conclude that W has at least $2^{\gamma n}$ components.

Suppose $w_2' \in W$. Then there exist w_1', w_3' with $d(w_i', \widehat{w}_{i,t}) \leq 2\varepsilon n$ for $i = 1, 3$ and $(w_1', w_2', w_3') \in \mathsf{PrePuzzle}^n(\tau, \tilde{\chi}^n)$. We have shown above (after (6.10)) that under these conditions, w_2' must be observed while the random walk reads the ladder word w_2. In particular, for $j \in [0, c_1 n[$, $w_2'(j) = w_2(j)$ or $w_2'(j)$ is an error in the observations. Since $B_{\mathrm{maj}}^{n,\tau,I}(\widehat{w}_{1,t}, \widehat{w}_{3,t})$ holds for the ladder interval $I = \{z + iL; i \in [0, c_1 n[\}$, in more than half of the words in W the j^{th} letter equals $w_2(j)$. Consequently, the j^{th} letter of $f(W)$ equals $w_2(j)$, and we have proved that $\mathsf{Cut}(w_2) \in \mathsf{Puzzle}^n(\tau, \tilde{\chi}^n)$. ■

Recall the definition of $E_{\mathrm{reconstruct}}^{n,\tau}$ from Theorem 3.5.

Lemma 6.4 *For all $n \geq c_{10}$ with c_{10} as in (6.4) the following holds:*

$$\begin{aligned} E_{\mathrm{reconstruct}}^{n,\tau} \quad \supseteq \quad & E_{\mathrm{only\,ladder}}^{n,\tau} \cap E_{\mathrm{all\,ladder}}^{n,\tau} \cap B_{\mathrm{few\ mistakes}}^{n} \cap B_{\mathrm{ladder\,diff}}^{n} \\ & \cap B_{\mathrm{outside\,out}}^{n} \cap E_{\mathrm{stop}}^{n,\tau}. \end{aligned}$$

Proof. Let $n \geq c_{10}$, and suppose all the events $E_{\mathrm{only\,ladder}}^{n,\tau}$, $E_{\mathrm{all\,ladder}}^{n,\tau}$, $B_{\mathrm{ladder\,diff}}^{n}$, $B_{\mathrm{few\ mistakes}}^{n}$, $B_{\mathrm{outside\,out}}^{n}$, and $E_{\mathrm{stop}}^{n,\tau}$ hold. Let $\psi \in \mathcal{C}^{[-kn,kn]}$ for some $k \geq c_1 n L$, and suppose $\psi \preceq \xi|[-2^n, 2^n]$. There exist $a \in [-2^n, 2^n]$ and $b \in \{-1, 1\}$ such that for all $j \in [-kn, kn]$

$$\psi(j) = \xi(a + bj) \quad \text{and} \quad a + bj \in [-2^n, 2^n]. \tag{6.12}$$

First we show $w := (\xi(a + bj); j \in [-3 \cdot 2^n, 3 \cdot 2^n]) \in \mathsf{SolutionPiece}^n(\tau, \tilde{\chi}^n, \psi)$. By (6.12), $\psi = w|[-kn, kn]$. Let $I \subseteq [-3 \cdot 2^n, 3 \cdot 2^n]$ be a ladder interval of length $c_1 n - 2$. The image of I under the map $j \mapsto a + bj$ is a ladder interval which is contained in $[-4 \cdot 2^n, 4 \cdot 2^n]$ because $|a| \leq 2^n$. Since $E_{\mathrm{all\,ladder}}^{n,\tau}$ holds, $(w|I)_{\rightarrow} \in \mathsf{Puzzle}^n(\tau, \tilde{\chi}^n)$. Consequently, $w \in \mathsf{SolutionPiece}^n(\tau, \tilde{\chi}^n, \psi)$, and in particular, $\mathsf{SolutionPiece}^n(\tau, \tilde{\chi}^n, \psi)$ is not empty.

It remains to show that $\xi|[-2^n, 2^n] \preceq w \preceq \xi|[-4 \cdot 2^n, 4 \cdot 2^n]$ for any element $w \in \mathsf{SolutionPiece}^n(\tau, \tilde{\chi}^n, \psi)$. Let $w \in \mathsf{SolutionPiece}^n(\tau, \tilde{\chi}^n, \psi)$. Then $w|[-kn, kn] = \psi$, and it follows from (6.12) that for all $j \in [-kn, kn]$

$$w(j) = \xi(a + bj). \tag{6.13}$$

Suppose we prove (6.13) for all $j \in [-3 \cdot 2^n, 3 \cdot 2^n]$. Then we know there is precisely one element in $\mathsf{SolutionPiece}^n(\tau, \tilde{\chi}^n, \psi)$. Since $\psi \preceq \xi|[-2^n, 2^n]$, there are more than $2 \cdot 2^n$ letters to the left and to the right of ψ in w, and consequently $\xi|[-2^n, 2^n] \preceq w$. On the other hand, in w, there are less than $3 \cdot 2^n$ letters to the left and to the right of ψ. Hence $w \preceq \xi|[-4 \cdot 2^n, 4 \cdot 2^n]$.

Thus, to finish the proof, it suffices to verify (6.13) for all $j \in [-3 \cdot 2^n, 3 \cdot 2^n]$. We have already seen that (6.13) holds for all $j \in [-kn, kn]$. Suppose we know that (6.13) holds for all $j \in [-s, s]$ for some $s \in [kn, 3 \cdot 2^n - 1]$. We set

$$\begin{aligned} w_l &:= (w|I_l)_{\rightarrow} \text{ with } I_l := (-s-1+iL; i \in [0, c_1 n - 2[) \,, \\ w_r &:= (w|I_r)_{\rightarrow} \text{ with } I_r := (s+1+(i-c_1 n+3)L; i \in [0, c_1 n - 2[) \,; \end{aligned}$$

note that I_l denotes the ladder interval of length $c_1 n - 2$ which contains $-s-1$ as leftmost point, and I_r denotes the ladder interval of length $c_1 n - 2$ which contains $s+1$ as rightmost point. The words w_l and w_r are well defined because $c_1 n L \leq |\psi| = 2kn + 1$. Since $w \in \mathsf{SolutionPiece}^n(\tau, \tilde{\chi}^n, \psi)$, we have $w_l, w_r \in \mathsf{Puzzle}^n(\tau, \tilde{\chi}^n)$. Note that w_l and w_r have both precisely $c_1 n - 3$ points in common with $w|[-s, s]$; w_l extends $w|[-s, s]$ one letter to the left, and w_r extends $w|[-s, s]$ one letter to the right.

Suppose $w_l \in \mathsf{Puzzle}_1^n(\tau, \tilde{\chi}^n)$. Then we have $w_l = f^*(W)$ for some $W = \mathsf{List}^n_{\tau, \tilde{\chi}^n}(w_1, w_2, w_3)$ and there exists $(w'_1, w'_2, w'_3) \in \mathsf{PrePuzzle}^n(\tau, \tilde{\chi}^n)$ such that $d(w_i, w'_i) \leq 2\varepsilon n$, for $i = 1, 3$ and (w'_1, w'_2, w'_3) is read while the random walk is walking on $\mathbb{Z} \setminus [-6 \cdot 2^n, 6 \cdot 2^n]$. Thus, there exists $t \in \cup_{k=1}^{2^{\alpha n}} \left[\tau_k, \tau_k + 2^{2n} - 3c_1 n\right]$ such that $|S(t+j)| > 6 \cdot 2^n$ for all $j \in [0, 3c_1 n[$ and $w'_2 = \tilde{\chi}|J$ with $J = [t + c_1 n, t + 2c_1 n[$. Using that $E^{n,\tau}_{\text{stop}}$ holds, we know that $|S(\tau_k)| \leq 2^n$ for all k. Since the random walk jumps a distance $\leq L$ in each step, it follows that $|S(t+j)| \leq 2^n + L \cdot 2^{2n} \leq 2L \cdot 2^{2n}$ for all $j \in [0, 3c_1 n[$. For a word $w = w_1 w_2 \ldots w_m \in \mathcal{C}^m$ of length $m \geq c_1 n/2$, we define $\mathsf{Last}(w) := w_{m - c_1 n/2 + 1} \ldots w_m$ to be the word consisting of the last $c_1 n/2$ letters of w. Let $z \in [-5 \cdot 2^n, 5 \cdot 2^n]$ and $i \in \{\leftarrow, \rightarrow\}$. Since $B^n_{\text{few mistakes}}$ and $B^n_{\text{outside out}}$ hold, we obtain

$$\begin{aligned} d(\mathsf{Last}(\mathsf{Cut}(w'_2)), w_{z,i,n/2}) &= d(\mathsf{Last}(\mathsf{Cut}(\tilde{\chi}|J)), w_{z,i,n/2}) \\ &\geq d(\mathsf{Last}(\mathsf{Cut}(\chi|J)), w_{z,i,n/2}) - \varepsilon n \geq 3\varepsilon n - \varepsilon n = 2\varepsilon n. \end{aligned} \tag{6.14}$$

By definition of $f^*(W)$, $d(\mathsf{Cut}(f(W)), \mathsf{Cut}(w)) \leq \varepsilon n$ for all $w \in W$. Hence

$$d(\mathsf{Last}(w_l), \mathsf{Last}(\mathsf{Cut}(w'_2))) \leq \varepsilon n. \tag{6.15}$$

Combining (6.14) and (6.15), we obtain

$$\begin{aligned} d(\mathsf{Last}(w_l), w_{z,i,n/2}) &\geq d(\mathsf{Last}(\mathsf{Cut}(w'_2)), w_{z,i,n/2}) - d(\mathsf{Last}(w_l), \mathsf{Last}(\mathsf{Cut}(w'_2))) \\ &\geq 2\varepsilon n - \varepsilon n = \varepsilon n. \end{aligned} \tag{6.16}$$

Recall that w_l is a ladder word of w of length $c_1 n - 2$ and the $c_1 n - 3$ right-most letters of w_l overlap with $w|[-s, s]$. Using that (6.13) holds for all $j \in [-s, s]$ together with $|a| \leq 2^n$ and $|s| \leq 3 \cdot 2^n$, yields $\mathsf{Last}(w_l) \preceq \xi|\,[-4 \cdot 2^n, 4 \cdot 2^n]$. This contradicts (6.16), which implies that $\mathsf{Last}(w_l)$ is different from any ladder word of $\xi|[-4 \cdot 2^n, 4 \cdot 2^n]$. We conclude $w_l \in \mathsf{Puzzle}_2^n(\tau, \tilde{\chi}^n)$. Since $E^{n,\tau}_{\text{only ladder}}$ holds, $w_l \preceq \xi|\,[-7 \cdot 2^n, 7 \cdot 2^n]$, and w_l is a ladder word of ξ.

Suppose (6.13) does not hold for $j = -s-1$. Let $I_{l,\xi}$ denote the image of I_l under the map $j \mapsto a + bj$. Then $\xi|I_{l,\xi} \neq w_l$; more precisely, $\xi|I_{l,\xi}$ and w_l disagree in precisely one point, namely the leftmost point $\xi(a + b(-s-1)) \neq w_l(0)$. Thus we found two ladder words of length $c_1 n - 2$ in $\xi|[-7 \cdot 2^n, 7 \cdot 2^n]$ which disagree in precisely one point. Consequently, there exist $z, z' \in [-8 \cdot 2^n, 8 \cdot 2^n]$, $i, i' \in \{\leftarrow$

$,\rightarrow\}$ with $(z,i) \neq (z',i')$ such that $\xi|I_{l,\xi} = \mathsf{Cut}(w_{z,i,n})$ and $w_l = \mathsf{Cut}(w_{z',i',n})$. Consequently, there exist $z_1, z_2 \in [-8\cdot 2^n, 8\cdot 2^n]$, $i_1, i_2 \in \{\leftarrow, \rightarrow\}$ with $(z_1,i_1) \neq (z_2,i_2)$ such that the two ladder words consisting of the last $c_1 n/3$ letters of $\xi|I_{l,\xi}$ and w_l respectively, equal $w_{z_1,i_1,n/3}$, $w_{z_2,i_2,n/3}$, respectively. Since $B^n_{\text{ladder diff}}$ holds, $w_{z_1,i_1,n/3} \neq w_{z_2,i_2,n/3}$ which is a contradiction. We conclude that (6.13) holds for $j = -s-1$.

To see that (6.13) holds for $j = s+1$, one applies the above argument with $\bar{w}$ defined by $\bar{w}(j) := w(-j)$ for $j \in [-3\cdot 2^n, 3\cdot 2^n]$ in place of w. By the induction principle, (6.13) holds for all $j \in [-3\cdot 2^n, 3\cdot 2^n]$. ■

6.3 The basic events have high probabilities

In this subsection, we prove that the events $B^n_{\dots}$ defined in Subsection 6.1 have a probability which is exponentially small in n. For some events $B^n_{\dots}$ this is only true under the assumption that $E^{n,\tau}_{\text{stop}}$ holds, i.e. if the stopping times stop correctly. We treat the events from Subsection 6.1 in alphabetical order.

Recall that unless otherwise stated, constants depend only on the distribution of the random walk increments and the number of colors of the scenery. In particular, the constants c_i in this section do not depend on n.

Lemma 6.5 *There exists a constant $c_{11} > 0$ such that for all $n \geq c_{11}$,*

$$P\left(E^{n,\tau}_{\text{stop}} \setminus B^{n,\tau}_{\text{all paths}}\right) \leq e^{-n}.$$

Proof. We have $P(S_0 = S_2 = 0) > 0$ because the random walk has a positive probability to make first a step of maximal length L to the right and then a step of maximal length L to the left. Hence 2 divides the period of the random walk, and the period must be 1 or 2. Therefore there exists $c_{12} > 0$ such that for all $n \geq c_{12}$ and for all $x, z \in [-7\cdot 2^n, 7\cdot 2^n]$, the random walk starting at x can reach z with positive probability in 2^{2n-1} or $2^{2n-1}+1$ steps:

$$P_x\left(S(2^{2n-1}) = z \text{ or } S(2^{2n-1}+1) = z\right) > 0. \tag{6.17}$$

We denote by $\mathcal{R}$ the set of all admissible pieces of path $R \in \mathbb{Z}^{[0,3c_1 n[}$ with starting point in $[-7\cdot 2^n, 7\cdot 2^n]$. For $R \in \mathcal{R}$ and $t \in \mathbb{N}_0$, we define the event

$$E(t,R) := \{S(t+i) = R(i)\ \forall i \in [0, 3c_1 n[\text{ or } S(t+1+i) = R(i)\ \forall i \in [0,3c_1 n[\}.$$

Let $n \geq \max\{c_{12}, c_{10}\}$ with c_{10} as in (6.4), and let $k \in [1, 2^{\alpha n}]$. We set $t_{k,n} := \tau_k + 2^{2n-1}$ and we define random variables $Y_k(R)$ as follows: If $|S(\tau_k)| \leq 2^n$ and $E(t_{k,n}, R)$ does not hold, then we set $Y_k(R) = 0$. Otherwise we set $Y_k(R) = 1$. Using the definitions of $E^{n,\tau}_{\text{stop}}$ and $B^{n,\tau}_{\text{all paths}}$, we see that

$$E^{n,\tau}_{\text{stop}} \setminus B^{n,\tau}_{\text{all paths}} \subseteq \bigcup_{R\in\mathcal{R}} E^{n,\tau}_{\text{stop}} \cap \left\{\sum_{k=1}^{2^{\alpha n}} Y_k(R) = 0\right\} \subseteq \bigcup_{R\in\mathcal{R}} E_{2^{\alpha n}}(R) \tag{6.18}$$

with

$$E_M(R) := \bigcap_{k=1}^{M} \{|S_{\tau_k}| \le 2^n, \tau_{k-1} + 2\cdot 2^n \le \tau_k, Y_k(R) = 0\}$$

for $M \in [1, 2^{\alpha n}]$. Let $R \in \mathcal{R}$. Since $n \ge c_{10}$, we have $3c_1 n \le 2^n$ by (6.4). Hence $t_{k,n} + 1 + 3c_1 n = \tau_k + 1 + 2^{2n-1} + 3c_1 n \le \tau_k + 2^{2n}$. Consequently, $\{\tau_k + 2\cdot 2^{2n} < \tau_{k+1}\} \cap E(t_{k,n}, R) \in \mathcal{F}_{\tau^n_{k+1}}$; here $\mathcal{F}_k := \sigma(S_i, \tilde{\chi}_i; i \in [0,k])$ denotes the natural filtration of random walk and observations with errors. Using the strong Markov property at time τ_M, we obtain

$$\begin{aligned}
P\left[E_M(R)\right] \le\ & P\left[E_{M-1}(R) \cap \left\{|S_{\tau_M}| \le 2^n, \tau_{M-1} + 2^{n+1} \le \tau_M, Y_M(R) = 0\right\}\right] \\
=& P\left[E_{M-1}(R) \cap \left\{|S(\tau^n_M)| \le 2^n, \tau_{M-1} + 2^{n+1} \le \tau_M\right\} \cap E(t_{M,n}, R)^c\right] \\
\le& P\left[E_{M-1}(R) \cap \{|S(\tau^n_M)| \le 2^n\} P_{S(\tau^n_M)}(E(2^{2n-1}, R)^c)\right] \\
\le& P[E_{M-1}(R)] \max_{x \in [-2^n, 2^n]} P_x[E(2^{2n-1}, R)^c].
\end{aligned}$$

An induction argument yields

$$P\left(E_{2^{\alpha n}}(R)\right) \le \left[\max_{x \in [-2^n, 2^n]} P_x(E(2^{2n-1}, R)^c)\right]^{2^{\alpha n}}. \tag{6.19}$$

To estimate the right-hand side of (6.19), let $b \in \mathbb{N}$ be minimal and let $h \in \mathbb{N}$ be maximal such that $P(S_1 - S_0 \in b + h\mathbb{Z}) = 1$. We set $\sigma^2 := E[(S_1 - S_0)^2]$, and $\mathcal{L}_m := \{(mb + hy)/\sqrt{m} : y \in \mathbb{Z}\}$. By the local central limit theorem ([6], page 132, Theorem (5.2)),

$$\lim_{m \to \infty} \sup_{y \in \mathcal{L}_m} \left| \frac{\sqrt{m}}{h} P\left(\frac{S_m}{\sqrt{m}} = y\right) - \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{y^2}{2\sigma^2}\right)\right| = 0.$$

We apply this with $m \in \{2^{2n-1}, 2^{2n-1}+1\}$, $y := (R_0 - x)/\sqrt{m}$ and R_0 equal to the starting point of R. Note that $|R_0| \le 7\cdot 2^n$ so that $|R_0 - x|/\sqrt{m} \le 16$ for all $x \in [-2^n, 2^n]$. Hence $\min_{x \in [-2^n, 2^n], R \in \mathcal{R}} \exp\left(-\frac{(R_0 - x)^2}{2m\sigma^2}\right) > 0$. We conclude that there exist constants $c_{13} > 0$ and $c_{14} \ge \max\{c_{12}, c_{10}\}$ such that for all $n \ge c_{14}$

$$\begin{aligned}
& \min_{x \in [-2^n, 2^n], R \in \mathcal{R}} P_x\left(S(2^{2n-1}) = R_0 \text{ or } S(2^{2n-1}+1) = R_0\right) \\
=\ & \min_{x \in [-2^n, 2^n], R \in \mathcal{R}} P\left(\frac{S(2^{2n-1})}{\sqrt{2^{2n-1}}} = \frac{R_0 - x}{\sqrt{2^{2n-1}}} \text{ or } \frac{S(2^{2n-1}+1)}{\sqrt{2^{2n-1}+1}} = \frac{R_0 - x}{\sqrt{2^{2n-1}+1}}\right) \\
\ge\ & c_{13} 2^{-n}
\end{aligned} \tag{6.20}$$

We set $\mu_{\min} := \min\{\mu(j) : j \in \mathcal{M}\}$; recall that μ is the distribution of the random walk increments $S_{k+1} - S_k$. The probability that the random walk starting at R_0 follows the path R for the next $3c_1 n - 1$ steps is bounded below by $\mu_{\min}^{3c_1 n - 1}$. Thus, (6.20) yields

$$\min_{x \in [-2^n, 2^n], R \in \mathcal{R}} P_x(E(2^{2n-1}, R)) \ge c_{13} 2^{-n} \mu_{\min}^{3c_1 n - 1} = c_{15} 2^{-n} \mu_{\min}^{3c_1 n}$$

with $c_{15} := c_{13}\mu_{\min}^{-1}$. Combining the last inequality with (6.18) and (6.19), we obtain

$$\begin{aligned} P\left(E_{\text{stop}}^{n,\tau} \setminus B_{\text{all paths}}^{n,\tau}\right) &\leq |\mathcal{R}| \left(1 - c_{15} 2^{-n} \mu_{\min}^{3c_1 n}\right)^{2^{\alpha n}} \\ &\leq (14 \cdot 2^n + 1)|\mathcal{M}|^{3c_1 n - 1} \exp\left(2^{\alpha n} \ln\left(1 - c_{15} 2^{-n} \mu_{\min}^{3c_1 n}\right)\right). \end{aligned} \tag{6.21}$$

Note that choosing a path in $\mathcal{R}$ one has $14 \cdot 2^n + 1$ possible starting points and $|\text{supp}(\mu)| = |\mathcal{M}|$ possibilities for each step of the path. Using the estimate $\ln(1-x) \leq -x$, we obtain

$$(6.21) \leq 2^{n+4}|\mathcal{M}|^{3c_1 n} \exp\left[-c_{15} 2^{(\alpha-1)n} \mu_{\min}^{3c_1 n}\right] = 2^{n+4}|\mathcal{M}|^{3c_1 n} \exp\left[-c_{15} e^{c_{16} n}\right]$$

and the last expression is $\leq e^{-n}$ for all n sufficiently large because $c_{16} = (\alpha - 1)\ln 2 + 3c_1 \ln \mu_{\min} > 0$ by our choice of α. ■

Lemma 6.6 *There exist $\delta_4 > 0$ such that for all $n \in \mathbb{N}$ and $\delta \in]0, \delta_4[$*

$$P_\delta\left(\left(B_{\text{few mistakes}}^n\right)^c\right) \leq e^{-n}.$$

Proof. Using Definition 6.3 and our convention $\varepsilon = c_1 \bar{\varepsilon}$ we obtain

$$\left(B_{\text{few mistakes}}^n\right)^c = \bigcup_{t \in [c_1 n - 1, 2 \cdot 2^{12\alpha n}[} \left\{ \sum_{k=t-c_1 n+1}^{t} X_k > c_1 \bar{\varepsilon} n \right\}. \tag{6.22}$$

Recall that X_k, $k \geq 0$, are i.i.d. Bernoulli random variables with parameter δ under P_δ. Hence $E_\delta\left[\sum_{k=t-c_1 n+1}^{t} X_k\right] = c_1 \delta n$. By the large deviation principle (see e.g. [3]), we have for all $\delta \in]0, \bar{\varepsilon}[$

$$P_\delta\left(\sum_{k=t-c_1 n+1}^{t} X_k > c_1 \bar{\varepsilon} n\right) \leq \exp\left(-I_\delta(\bar{\varepsilon} - \delta) c_1 n\right) \tag{6.23}$$

with rate function

$$I_\delta(x) = (1-x)\log\left(\frac{1-x}{1-\delta}\right) + x \log\left(\frac{x}{\delta}\right), x \in]0,1[. \tag{6.24}$$

Combining (6.22) with (6.23) we obtain for all $\delta \in]0, \bar{\varepsilon}[$

$$P_\delta\left(\left(B_{\text{few mistakes}}^n\right)^c\right) \leq \exp\left([1 + 12\alpha n]\ln 2 - I_\delta(\bar{\varepsilon} - \delta) c_1 n\right).$$

Since

$$\lim_{\delta \to 0} I_\delta(\bar{\varepsilon} - \delta) = \lim_{\delta \to 0} (1 - \bar{\varepsilon} + \delta) \log\left[\frac{1 - \bar{\varepsilon} + \delta}{1 - \delta}\right] + (\bar{\varepsilon} - \delta) \log\left[\frac{\bar{\varepsilon} - \delta}{\delta}\right] = +\infty,$$

there exists $\delta_4 \in]0, \bar{\varepsilon}[$ such that $[1 + 12\alpha]\ln 2 - I_\delta(\bar{\varepsilon} - \delta)c_1 < -1$ for all $\delta \in]0, \delta_4[$. The assertion of the lemma follows. ■

We will need the following lemma in the proofs of Lemmas 6.8, 6.10, and 6.13.

Lemma 6.7 *There exist $\varepsilon_1, c_{17}(\varepsilon') > 0$ such that for all m with $c_1 m \in \mathbb{N}$, $\varepsilon' \in]0, \varepsilon_1[$, $w \in \mathcal{C}^{[0,c_1 m[}$, and for any admissible piece of path $R \in \mathbb{Z}^{[0,c_1 m[}$ the following holds:*

$$P\left(d(\xi \circ R, w) < c_1 \varepsilon' m\right) \leq c_{17}(\varepsilon')(c_2)^{c_1 m} \max_J P((\xi \circ R)|J = w|J),$$

where the maximum is taken over all subsets $J \subseteq [0, c_1 m[$ with cardinality $|J| = c_1 m - \lfloor c_1 \varepsilon' m \rfloor$ and c_2 is as in Section 2.1.

Proof. Let m be such that $c_1 m \in \mathbb{N}$, let $w \in \mathcal{C}^{[0,c_1 m[}$, and let $R \in \mathbb{Z}^{[0,c_1 m[}$ be an admissible piece of path. If $d(\xi \circ R, w) < c_1 \varepsilon' m$, then $c_1 m - \lfloor c_1 \varepsilon' m \rfloor$ letters of $\xi \circ R$ and w agree. Since there are $\binom{c_1 m}{\lfloor c_1 \varepsilon' m \rfloor}$ possibilities of choosing $c_1 m - \lfloor c_1 \varepsilon' m \rfloor$ out of $c_1 m$ letters, we have

$$P\left(d(\xi \circ R, w) < c_1 \varepsilon' m\right) \leq \binom{c_1 m}{\lfloor c_1 \varepsilon' m \rfloor} \max_J P((\xi \circ R)|J = w|J),$$

where the maximum is taken over all subsets $J \subseteq [0, c_1 m[$ with cardinality $c_1 m - \lfloor c_1 \varepsilon' m \rfloor$. By Stirling's formula ([1], p.24, formula (3.9)) we have for $k \in \mathbb{N}$, $k! = \sqrt{2\pi} k^{k+1/2} e^{-k+\theta(k)}$ with $\theta(k) \in]0, 1[$ and $\lim_{k\to\infty} \theta(k) = 0$. Thus

$$\binom{c_1 m}{\lfloor c_1 \varepsilon' m \rfloor} \quad \leq \quad c_{17}(\varepsilon') \varphi\left(\frac{\lfloor c_1 \varepsilon' m \rfloor}{c_1 m}\right)^{c_1 m}$$

with $\varphi(x) = x^{-x}(1-x)^{-(1-x)}$ and some constant $c_{17}(\varepsilon') > 0$ independent of m. Note that φ is continuous at 0 with $\varphi(0) = 1$, and recall that $c_2 \in]1, C/(C-1)[$. There exists ε_1 such that $\varphi(x) < c_2$ for all $x \in]0, \varepsilon_1[$. Note that $\lfloor c_1 \varepsilon' m \rfloor / (c_1 m) \leq \varepsilon'$. The claim follows. ∎

Lemma 6.8 *There exists a constant $c_{18} > 0$ such that for all $n \in \mathbb{N}$*

$$P\left((B^n_{\text{ladder diff}})^c\right) \leq c_{18} e^{-n}.$$

Proof. Let

$$\mathcal{J} := \left\{(z_1, i_1, z_2, i_2) \in \left([-8 \cdot 2^n, 8 \cdot 2^n] \times \{\leftarrow, \rightarrow\}\right)^2 : (z_1, i_1) \neq (z_2, i_2)\right\}.$$

By Definition 6.4,

$$(B^n_{\text{ladder diff}})^c = \bigcup_{(z_1,i_1,z_2,i_2)\in\mathcal{J}} \{d(w_{z_1,i_1,n/3}, w_{z_2,i_2,n/3}) < 10\varepsilon n\}. \tag{6.25}$$

Let $(z_1, i_1, z_2, i_2) \in \mathcal{J}$. For $k = 1, 2$ we set $o_k := +1$ if $i_k = \rightarrow$, $o_k := -1$ if $i_k = \leftarrow$, and we set $f_k(j) := z_k + o_k j L$ for $j \in [0, c_1 n/3[$. First we prove that there exists a subset $J \subseteq [0, c_1 n/3[$ of cardinality $|J| \geq c_1 n/9$ such that

$$f_1(J) \cap f_2(J) = \emptyset. \tag{6.26}$$

We distinguish two cases. Case $z_1 = z_2$: By assumption, $i_1 \neq i_2$. Hence $o_1 \neq o_2$, and we conclude that (6.26) is satisfied for $J =]0, c_1 n/3[$.

Case $z_1 \neq z_2$: We show by induction over $k \in [1, c_1 n/9]$ that there exists J with $|J| \geq k$ such that (6.26) holds. For $k = 1$ the set $J = \{0\}$ has the required property. Suppose there exists J' with $|J'| = k \in [1, c_1 n/9 - 1]$ such that (6.26) holds. The sets $J'_i := f_i(J')$, $i = 1, 2$, have cardinality $|J'_i| = |J'| \leq c_1 n/9 - 1$. We set

$$\bar{J} := \{j \in [0, c_1 n/3[: f_1(j) \notin J'_1 \cup J'_2, f_2(j) \notin J'_1, \text{ and } f_1(j) \neq f_2(j)\}.$$

Then $|\bar{J}| \geq c_1 n/3 - |J'_1 \cup J'_2| - |J'_1| - 1 = c_1 n/3 - 3(c_1 n/9 - 1) - 1 = 2$; note that there exists at most one j with $f_1(j) \neq f_2(j)$. In particular $\bar{J}$ is not empty. Let $j \in \bar{J}$, and set $J := J' \cup \{j\}$. Since $f_1(j) \notin J'_1$, we have $|J| = |J'| + 1$. It follows from $f_1(j) \notin J'_2 \cup \{f_2(j)\}$ that $f_1(j) \notin f_2(J)$. Similarly, it follows from $f_2(j) \notin J'_1 \cup \{f_1(j)\}$ that $f_2(j) \notin f_1(J)$, and we have proved that (6.26) holds for J. By the induction principle, (6.26) holds for a set $J \subseteq [0, c_1 n/3[$ of cardinality $|J| = c_1 n/9$.

Let $J \subseteq [0, c_1 n/3[$ with $|J| = c_1 n/9$ such that (6.26) holds. Then the words $w_{z_k, i_k, n/3}|f_k(J)$, $k = 1, 2$, are independent. Note that $P(\xi_k = \xi_{k'}) = 1/C$ for $k \neq k'$. We use Lemma 6.7 with $m := n/9$, $\varepsilon' := 90\varepsilon/c_1$ and R equal to the ladder path underlying $w_{z_1, i_1, n/3}$ to obtain

$$\begin{aligned}
& P(d(w_{z_1,i_1,n/3}, w_{z_2,i_2,n/3}) < 10\varepsilon n) \\
\leq \; & P\left(d(w_{z_1,i_1,n/3}|f_1(J), w_{z_2,i_2,n/3}|f_2(J)) < 10\varepsilon n\right) \\
\leq \; & c_{17}(90\varepsilon/c_1)(c_2)^{c_1 n/9} C^{\lfloor 10\varepsilon n \rfloor - c_1 n/9}.
\end{aligned} \tag{6.27}$$

Since the intersection in (6.25) is taken over $4(16 \cdot 2^n + 1)^2$ possible pairs (z_1, i_1), (z_2, i_2), it follows from (6.27) that

$$P(B^n_{\text{ladder diff}}) \leq 4(16 \cdot 2^n + 1)^2 c_{17}(90\varepsilon/c_1)(c_2)^{c_1 n/9} C^{\lfloor 10\varepsilon n \rfloor - c_1 n/9}.$$

Note that $C^{\lfloor 10\varepsilon n \rfloor} \leq \exp(10\varepsilon n \ln C)$. Let $c_{18} > 0$ be chosen in such a way that $4(16 \cdot 2^n + 1)^2 c_{17}(90\varepsilon/c_1) \leq c_{18} 2^{2n}$. Then

$$P(B^n_{\text{ladder diff}}) \leq c_{18} e^{n[2\ln 2 + 10\varepsilon \ln C + (c_1/9)[\ln c_2 - \ln C]]}.$$

Since $2\ln 2 + 10\varepsilon \ln C + (c_1/9)[\ln c_2 - \ln C] < -1$ by our choice of ε and c_1, the claim follows. ■

Lemma 6.9 *There exist constants $c_{19}, \delta_5 > 0$ such that for all $n \geq c_{19}$ and $\delta \in]0, \delta_5[$*

$$P_\delta\left(\left(B^{n,\tau}_{\text{majority}}\right)^c\right) \leq e^{-n}.$$

Proof. Recall the notation from Definition 6.5. Let $w_1, w_3 \in \mathcal{C}^{c_1 n}$, $I \in \mathcal{I}_L$. Let r_i, $i \geq 1$, denote all the times $s \in \cup_{k=1}^{2^{\alpha n}} \left[\tau_k + c_1 n, \tau_k + 2^{2n} - 2c_1 n\right]$ such that $S|[r_i, r_i + c_1 n[$ is a straight crossing of I from left to right. Clearly, the intervals

$[r_i, r_i + c_1 n[$, $i \geq 1$, are pairwise disjoint. Let $\mathcal{H} := \sigma(r_i, \tau_i; i \geq 1)$. Since S and X are independent, we know that conditioned on $\mathcal{H}$, the random variables X_{r_i+j}, $i \geq 1$, $j \in [0, c_1 n[$, are i.i.d. Bernoulli with parameter δ under P_δ.

We obtain the random variables $s_i^{I\rightarrow} + c_1 n$, $i \geq 1$, from r_i, $i \geq 1$, by checking whether $d(\tilde{\chi}|[r_i + (k-2)c_1 n, r_i + (k-1)c_1 n[, w_k) \leq 2\varepsilon n$ for $k = 1, 3$. Since at time $r_i + c_1 n - 1$ the random walk is at the right endpoint of I and at time r_{i+1} at the left endpoint of I, the time interval $[r_i + c_1 n - 1, r_{i+1}]$ has length $\geq c_1 n$. Consequently, the time intervals $[r_i, r_i + c_1 n[$, $[r_{i+1}, r_{i+1} + c_1 n[$ have a distance $\geq c_1 n - 2$ from each other. Since ξ, S, Y are independent of X, we conclude that $\tilde{\chi}|[s_i^{I\rightarrow} + kc_1 n, s_i^{I\rightarrow} + (k+1)c_1 n[$, $k = 0, 2$, $i \geq 1$, is independent of $\sigma(X_{s_i^{I\rightarrow} + c_1 n + j}; j \in [1, c_1 n - 1[, i \geq 1)$. Hence conditioned on $\bar{\mathcal{H}} := \sigma\left(s_i^{I\rightarrow} + c_1 n, \tau_i, \tilde{\chi}|[s_i^{I\rightarrow} + kc_1 n, s_i^{I\rightarrow} + (k+1)c_1 n[; i \geq 1, k = 0, 2\right)$ the random variables $X_{s^{I\rightarrow} + c_1 n + j}$, $j \in [1, c_1 n - 1[$, are i.i.d. Bernoulli with parameter δ under P_δ.

By the large deviation principle (see e.g. [3]), we have for all $\delta \in]0, 1/2[$ and $n \in \mathbb{N}$ P_δ-almost surely on the set $\left\{|\mathcal{S}_{w_1,w_3}^{I\rightarrow}| \geq 2^{\gamma n}\right\}$

$$P_\delta\left(\sum_{i=1}^{2^{\gamma n}} X_{s_i + c_1 n + j} \geq 2^{\gamma n}/2 \,\middle|\, \bar{\mathcal{H}}\right) \leq \exp\left(-I_\delta(1/2 - \delta)2^{\gamma n}\right) \tag{6.28}$$

with rate function I_δ given by (6.24). Since

$$\lim_{\delta \to 0} I_\delta(1/2 - \delta) = \lim_{\delta \to 0} (1/2 + \delta) \log\left[\frac{1/2 + \delta}{1 - \delta}\right] + (1/2 - \delta) \log\left[\frac{1/2 - \delta}{\delta}\right] = +\infty,$$

there exists $\delta_5 > 0$ such that $I_\delta(1/2 - \delta) > 1$ for all $\delta \in]0, \delta_5[$. It follows from (6.28) that for all $\delta \in]0, \delta_5[$ P_δ-almost surely on the set $\left\{|\mathcal{S}_{w_1,w_3}^{I\rightarrow}| \geq 2^{\gamma n}\right\}$

$$P_\delta\left(\sum_{i=1}^{2^{\gamma n}} X_{s_i + c_1 n + j} \geq 2^{\gamma n}/2 \,\middle|\, \bar{\mathcal{H}}\right) \leq \exp\left(-2^{\gamma n}\right). \tag{6.29}$$

Consequently, $P_\delta\left(\sum_{i=1}^{2^{\gamma n}} X_{s_i + c_1 n + j} \geq 2^{\gamma n}/2\right) \leq \exp\left(-2^{\gamma n}\right)$. By Definition 6.5, $B_{\text{majority}}^{\mathrm{n},\tau} = B_{\text{maj},\rightarrow}^{n,\tau} \cap B_{\text{maj},\leftarrow}^{n,\tau}$ with

$$B_{\text{maj},\rightarrow}^{\mathrm{n},\tau} = \left\{|\mathcal{S}_{w_1,w_3}^{I\rightarrow}| < 2^{\gamma n}\right\} \cup \bigcap_{w_1, w_3 \in \mathcal{C}^{c_1 n}} \bigcap_{I \in \mathcal{I}_L} B_{\text{maj}}^{n,\tau,I\rightarrow}(w_1, w_3)$$

and $B_{\text{maj},\leftarrow}^{n,\tau}$ defined analogously. Hence

$$\left[B_{\text{maj},\rightarrow}^{\mathrm{n},\tau}\right]^c = \bigcup_{w_1, w_3 \in \mathcal{C}^{c_1 n}} \bigcup_{I \in \mathcal{I}_L} \bigcup_{j \in [1, c_1 n - 1[} \left\{|\mathcal{S}_{w_1,w_3}^{I\rightarrow}| \geq 2^{\gamma n}, \sum_{i=1}^{2^{\gamma n}} X_{s_i + c_1 n + j} \geq \frac{2^{\gamma n}}{2}\right\}.$$

Since there are less than $14 \cdot 2^n$ ladder intervals in $\mathcal{I}_L$, it follows that

$$P_\delta\left(\left(B_{\text{maj},\rightarrow}^{n,\tau}\right)^c\right) \leq 14 \cdot 2^n c_1 n C^{2c_1 n} \exp\left(-2^{\gamma n}\right).$$

We choose $c_{19} > 0$ large enough that $14 \cdot 2^n c_1 n C^{2c_1 n} \exp\left(-2^{\gamma n}\right) \leq e^{-n}/2$ for all $n \geq c_{19}$. The claim follows. ■

Lemma 6.10 *There exist constants $c_{20}, c_{21} > 0$ such that for all $n \geq c_{10}$ (with c_{10} as in (6.4))*

$$P\left((B^n_{\text{outside out}})^c\right) \leq c_{21} e^{-c_{20} n}.$$

Proof. We set

$$\mathcal{J} := \left\{ \begin{array}{l} (z, i, R) : R \in ([-2L\cdot 2^{2n}, 2L\cdot 2^{2n}] \backslash [-6\cdot 2^n, 6\cdot 2^n])^{[0, c_1 n/2[} \text{ admissible} \\ \text{piece of path, } z \in [-5 \cdot 2^n, 5 \cdot 2^n],\ i \in \{\leftarrow, \rightarrow\} \end{array} \right\}.$$

By Definition 6.6,

$$(B^n_{\text{outside out}})^c = \bigcup_{(z,i,R) \in \mathcal{J}} \left\{ d(\xi \circ R, w_{z,i,n/2}) < 3\varepsilon n \right\},$$

and consequently,

$$P\left((B^n_{\text{outside out}})^c\right) \leq |\mathcal{J}| \max_{(z,i,R) \in \mathcal{J}} P\left(d(\xi \circ R, w_{z,i,n/2}) < 3\varepsilon n\right). \tag{6.30}$$

Let $(z, i, R) \in \mathcal{J}$, and let $n \geq c_{10}$. The piece of scenery $\xi \circ R$ depends only on $\xi | [-2L \cdot 2^{2n}, 2L \cdot 2^{2n}] \setminus [-6 \cdot 2^n, 6 \cdot 2^n]$, whereas $w_{z,i,n/2}$ depends only on $\xi | [-5 \cdot 2^n - c_1 n L/2, 5 \cdot 2^n + c_1 n L/2]$. Since $n \geq c_{10}$, $c_1 n L/2 \leq 2^n$ by (6.4), and therefore $w_{z,i,n/2}$ depends only on $\xi | [-6 \cdot 2^n, 6 \cdot 2^n]$. Since the scenery ξ is i.i.d. uniformly colored, $\xi \circ R$ and $w_{z,i,n/2}$ are independent and $P(\xi_j = \xi_{j'}) = 1/C$ for $j \neq j'$. Thus

$$P\left(\xi(R(j)) = w_{z,i,n/2}(j) \ \ \forall j \in J\right) = C^{\lfloor 3\varepsilon n \rfloor - c_1 n/2}$$

for any subset $J \subseteq [0, c_1 n/2[$ with cardinality $|J| = c_1 n/2 - \lfloor 3\varepsilon n \rfloor$. Applying Lemma 6.7 with $\varepsilon' = 6\varepsilon/c_1$ and $m = n/2$, we obtain

$$P\left(d(\xi \circ R, w_{z,i,n/2}) < 3\varepsilon n\right) \leq c_{17}(6\varepsilon/c_1)(c_2)^{c_1 n/2} C^{\lfloor 3\varepsilon n \rfloor - c_1 n/2}. \tag{6.31}$$

The cardinality of $|\mathcal{J}|$ satisfies

$$|\mathcal{J}| \leq 2(10 \cdot 2^n + 1) 4L \cdot 2^{2n} (C - 1)^{c_1 n/2} \tag{6.32}$$

for the following reason: There are $10 \cdot 2^n + 1$ possible values for z, 2 possible values for i and at most $4L \cdot 2^{2n}$ possible starting points for R. An admissible piece of path has at each step at most $|\mathcal{M}| \leq C - 1$ possible steps; recall that there are strictly more colors than possible steps for the random walk. Hence the number of possible paths R is bounded by $4L \cdot 2^{2n} (C-1)^{c_1 n/2}$.

Clearly, $C^{\lfloor 3\varepsilon n \rfloor} \leq e^{(3\varepsilon n \ln C)}$. We choose $c_{21} > 0$ such that $c_{17}(6\varepsilon/c_1) 2(10 \cdot 2^n + 1) 4L \cdot 2^{2n} \leq c_{21} \cdot 2^{3n}$. Combining (6.30), (6.31), and (6.32), we obtain

$$P\left((B^n_{\text{outside out}})^c\right) \leq c_{21} e^{n(3 \ln 2 + 3\varepsilon \ln C)} \left(\frac{c_2(C-1)}{C}\right)^{c_1 n/2}.$$

Finally, we set $c_{20} := -\left(3\ln 2 + 3\varepsilon \ln C + (c_1/2)\ln\left(\frac{c_2(C-1)}{C}\right)\right)$, and the claim follows because $c_{20} > 0$ by our choice of ε and c_1. ■

We will need the following lemma in the proof of Lemma 6.12.

Lemma 6.11 *There exists c_{22} such that for all $n \geq c_{22}$ and for any admissible piece of path $R \in \mathbb{Z}^{[0,c_1 n[}$ with $R(0) \leq R(c_1 n - 1)$ there exists an admissible piece of path $\bar{R} \in \mathbb{Z}^{[0,c_1 n[}$ such that $\bar{R}(0) = R(0)$, $\bar{R}(c_1 n - 1) = R(c_1 n - 1)$, and the first $c_1 n/3$ steps of $\bar{R}$ are steps of maximal length L to the right.*

Proof. Let $R \in \mathbb{Z}^{[0,c_1 n[}$ be an admissible piece of path. We set $x := R(0)$, $y := R(c_1 n - 1)$; note $x \leq y$.

Suppose R contains at least $c_1 n/3$ steps of maximal length L to the right. Then we define $\bar{R} \in \mathbb{Z}^{[0,c_1 n[}$ to be the admissible piece of path starting at x and ending at y obtained from R by permuting the order of the steps in such a way that all the steps of maximal length L to the right are at the beginning.

If R contains less than $c_1 n/3$ steps of maximal length L to the right, then

$$y - x \leq \left(\frac{c_1 n}{3} - 1\right) L + \frac{2c_1 n}{3}(L-1) \leq c_1 n L - \frac{2c_1 n}{3}. \tag{6.33}$$

In this case, let $R_1 \in \mathbb{Z}^{[0,t_1[}$ denote the path which starts at x and goes with maximum steps to the right until it reaches the interval $]y-L, y]$. In other words, $R_1(0) = x$, $R_1(t_1 - 1) \in]y-L, y]$, and for all $s \in [0, t_1 - 1[$ we have that $R_1(s+1) - R_1(s) = L$. Let $y' := R_1(t_1 - 1)$ be the endpoint of R_1. We have $(t_1 - 1)L \leq y - x$ and using (6.33), we obtain

$$t_1 \leq \frac{y-x}{L} + 1 \leq c_1 n - \frac{2c_1 n}{3L} + 1. \tag{6.34}$$

As we noticed already in the proof of Lemma 6.5, the random walk has period 1 or 2. Thus there exists c_{23} such that for all $z \in]y-L, y]$ there exists an admissible piece of path of length $\leq c_{23}$ starting at z and ending at y. If furthermore the random walk is aperiodic, then c_{23} can be chosen in such a way that for all $z \in]y-L, y]$ there exist admissible pieces of path of even and odd length $\leq c_{23}$ starting at z and ending at y. We choose c_{22} such that $\min\left\{\frac{c_1 n}{3} - 2, \frac{2c_1 n}{3L} - 2\right\} > c_{23}$ for all $n \geq c_{22}$.

Case 1: The random walk is periodic (with period 2). Let $R_3 \in \mathbb{Z}^{[0,t_3[}$ be an admissible piece of path starting at y', ending at y with $t_3 \leq c_{23}$. The concatenation $R_1 R_3$ is an admissible piece of path starting at x, ending at y of length $t_1 + t_3 \leq c_1 n - 1$ by (6.34). By assumption, R also starts at x and ends at y. Thus by periodicity we have that $l := |R| - |R_1 R_3| \geq 0$ is even. Let R_2 be the admissible piece of path starting and ending at y' which makes first $l/2$ steps of length L to the right and then $l/2$ steps of length L to the left. We set $\bar{R} := R_1 R_2 R_3$. We have $|R_1 R_2| \geq c_1 n - c_{23} \geq 2 + 2c_1 n/3$. Since all steps of R_1 and half of the steps of R_2 are maximum steps to the right, $\bar{R}$ contains at least $c_1 n/3$ steps of maximal length L at the beginning. By construction, $\bar{R}$ starts at x and ends at y.

Case 2: The random walk is aperiodic. Let $R_3 \in \mathbb{Z}^{[0,t_3[}$ be an admissible piece of path starting at y', ending at y of length $t_3 \leq c_{23}$. We may assume that t_3 is even iff $c_1 n - t_1$ is even. Then $c_1 n - t_1 - t_3$ is even, and we can define R_2 as before. The same argument as above shows that $\bar{R} := R_1 R_2 R_3$ fulfills the claim. ■

Lemma 6.12 *There exists c_{24} such that for all $n \geq c_{24}$*

$$P\left(\left(B^n_{\text{recogn straight}}\right)^c\right) \leq c_{18} e^{-n};$$

c_{18} is specified in Lemma 6.8.

Proof. Let $c_{24} := \max\{c_{10}, c_{22}\}$ with c_{22} as in Lemma 6.11, and let $n \geq c_{24}$. We will show that the following inclusion holds:

$$B^n_{\text{ladder diff}} \subseteq B^n_{\text{recogn straight}}. \tag{6.35}$$

The claim follows then from Lemma 6.8.

Suppose the event $B^n_{\text{ladder diff}}$ holds. Let $R_1 \in [-7 \cdot 2^n, 7 \cdot 2^n]^{[0,c_1 n[}$ be an admissible piece of path which is not a ladder path. We set $x := R_1(0)$ and $y := R_1(c_1 n - 1)$. We have to show that there exists an admissible piece of path $R_2 \in [-8 \cdot 2^n, 8 \cdot 2^n]^{[0,c_1 n[}$ with starting point x, endpoint y, and $d(\xi \circ R_1, \xi \circ R_2) \geq 5\varepsilon n$. We assume that $x \leq y$. The case $x > y$ is reduced to this case by considering the reversed path $k \mapsto R_1(c_1 n - 1 - k)$. By Lemma 6.11 applied to R_1, there exists an admissible piece of path $R_3 \in \mathbb{Z}^{[0,c_1 n[}$ such that $R_3(0) = x$, $R_3(c_1 n - 1) = y$ and the first $c_1 n/3$ steps of R_3 are steps of maximal length L to the right. Since $y - x \neq (c_1 n - 1)L$, at least one step of R_3 is not a step of maximum length to the right. We construct an admissible piece of path R_4 by permuting the steps of R_3. We set $R_4(0) := x$. The first step of R_4 is the first step of R_3 which is not a step of maximum length to the right. Formally we set $j := \min\{i \in [1, c_1 n[: R_3(i) - R_3(i-1) \neq L\}$, and define

$$R_4(i) := \begin{cases} R_3(i), & \text{if } i \in [0, c_1 n[\setminus [1, j] \\ R_3(i-1) + R_3(j) - R_3(j-1), & \text{if } i \in [1, j]. \end{cases}$$

Clearly, R_4 is an admissible piece of path of length $c_1 n$ with $R_4(0) = x$ and $R_4(c_1 n - 1) = y$. Using that R_4 jumps in each step at most a distance of L, we obtain that $|R_4(i)| \leq |R_4(0)| + c_1 n L = x + c_1 n L \leq 8 \cdot 2^n$ for all $i \in [0, c_1 n[$ because $c_1 n L \leq 2^n$ for $n \geq c_{10}$. The same is true for R_3.

Since R_3 starts with $c_1 n/3$ steps of maximum length L to the right, we have that $\xi \circ R_3|[1, c_1 n/3] = w_{x+L,\rightarrow,n/3}$, and by definition of R_4, we have $\xi \circ R_4|[1, c_1 n/3] = w_{x',\rightarrow,n/3}$ with $x' = x + R_3(j) - R_3(j-1)$. By construction, $R_3(j) - R_3(j-1) \neq L$ so that $x + L \neq x'$. Since R_3 and R_4 take only values in $[-8 \cdot 2^n, 8 \cdot 2^n]$, we have that $x + L, x' \in [-8 \cdot 2^n, 8 \cdot 2^n]$. Using that $B^n_{\text{ladder diff}}$ holds, yields $d(w_{x+L,\rightarrow,n/3}, w_{x',\rightarrow,n/3}) \geq 10\varepsilon n$, and by the triangle inequality, we get that $\xi \circ R_1$ cannot have a distance smaller than $5\varepsilon n$ to both $\xi \circ R_3$ and $\xi \circ R_4$. Hence there exists $i \in \{3, 4\}$ such that $d(\xi \circ R_1, \xi \circ R_i) \geq 5\varepsilon n$. Let $R_2 := R_i$ in the definition of $B^n_{\text{recogn straight}}$. ■

Lemma 6.13 *There exist constants $c_{25}, c_{26} > 0$ such that for all $n \in \mathbb{N}$*

$$P\left(\left(B^n_{\text{signals}}\right)^c\right) \leq c_{25} e^{-c_{26} n}.$$

Proof. We show that there exist $c_{25}, c_{26} > 0$ such that for all n

$$P\left(\left(B^n_{\text{sign,r},\rightarrow}\right)^c\right) \leq \frac{c_{25}}{4} e^{-c_{26} n}. \tag{6.36}$$

Analogously, one proves statements for $B^n_{\text{sign,l},\rightarrow}$, $B^n_{\text{sign,l},\leftarrow}$, and $B^n_{\text{sign,r},\leftarrow}$. The claim follows from these four inequalities and the definition of B^n_{signals}. We set

$$\mathcal{R} := \left\{ \begin{matrix} (z, R) : z \in [-6 \cdot 2^n, 6 \cdot 2^n],\, R \in \left[-2L \cdot 2^{2n}, 2L \cdot 2^{2n}\right]^{[0, c_1 n[} \text{ admis-} \\ \text{sible piece of path with } R(0) < z \end{matrix} \right\}.$$

By Definition 6.8,

$$\left(B^n_{\text{sign,r},\rightarrow}\right)^c = \bigcup_{(z,R) \in \mathcal{R}} \left\{ d(\xi \circ R, w_{z,\rightarrow,n}) < 5\varepsilon n \right\}. \tag{6.37}$$

Let $(z, R) \in \mathcal{R}$. By Definition 6.1, $w_{z,\rightarrow,n}(k) = \xi(z + kL)$. Note that $R(k) < z + kL$ for all $k \in [0, c_1 n[$: For $k = 0$ this is true by assumption. Suppose $R(k) < z + kL$ holds for some $k \in [0, c_1 n - 1[$. Since the maximal jump length of R is L, we obtain $R(k+1) \leq R(k) + L < z + (k+1)L$, and the claim follows by induction.

We prove by induction over the cardinality of J, that

$$P((\xi \circ R)|J = w_{z,\rightarrow,n}|J) = C^{-|J|} \tag{6.38}$$

for any $J \subseteq [0, c_1 n[$: For $J = \{j\}$ we use that $\xi(R(j))$ and $w_{z,\rightarrow,n}(j) = \xi(z + jL)$ are independent because $R(j) < z + jL$. Suppose (6.38) holds for any $J \subseteq [0, c_1 n[$ with $|J| = k$ for some $k \in [1, c_1 n - 1[$. Let $J' \subseteq [0, c_1 n[$ with $|J'| = k + 1$, and let $j := \max J'$. Then $\xi(z + jL)$ is independent of $\xi(z + j'L)$, $j' \in J' \setminus \{j\}$, and of $\xi(R(j'))$, $j' \in J'$, because $R(j') < z + j'L \leq z + jL$. Hence

$$\begin{aligned} P((\xi \circ R)|J' = w_{z,\rightarrow,n}|J') &= C^{-1} P((\xi \circ R)|J' \setminus \{j\} = w_{z,\rightarrow,n}|J' \setminus \{j\}) \\ &= C^{-\left(1 + |J' \setminus \{j\}|\right)} = C^{|J'|}; \end{aligned}$$

for the second but last equality with used the induction hypothesis. We use Lemma 6.7 with $\varepsilon' := 5\varepsilon$ and $m := n$ to obtain

$$P\left(d(\xi \circ R, w_{z,\rightarrow,n}) < 5\varepsilon n\right) \leq c_{17}(5\varepsilon / c_1)(c_2)^{c_1 n} C^{\lfloor 5\varepsilon n \rfloor - c_1 n}. \tag{6.39}$$

It is easy to see that the cardinality of $\mathcal{R}$ is bounded by $(12 \cdot 2^n + 1)(4L \cdot 2^{2n} + 1)(C - 1)^{c_1 n}$ Combining this with (6.37) and (6.39), we obtain

$$P\left(\left(B^n_{\text{sign,r},\rightarrow}\right)^c\right) \leq c_{17}(5\varepsilon / c_1)(12 \cdot 2^n + 1)(4L \cdot 2^{2n} + 1) C^{\lfloor 5\varepsilon n \rfloor} \left(\frac{c_2(C-1)}{C}\right)^{c_1 n}.$$

We choose c_{25} such that $c_{17}(5\varepsilon/c_1)(12\cdot 2^n+1)(4L\cdot 2^{2n}+1)\le c_{25}2^{3n}/4$ for all $n\in\mathbb{N}$. Then

$$P\left(\left(B^n_{\text{sign,r},\to}\right)^c\right)\le \frac{c_{25}}{4}e^{n[3\ln 2+5\varepsilon\ln C]}\left(\frac{c_2(C-1)}{C}\right)^{c_1 n}.$$

We set $c_{26}:=-\left(3\ln 2+5\varepsilon\ln C+c_1\ln\left(\frac{c_2(C-1)}{C}\right)\right)$. Since $c_{26}>0$ by our choice of ε and c_1, the claim follows. ■

Lemma 6.14 *There exists a constant $c_{27}>0$ such that for all $n\ge c_{27}$*

$$P\left(E^{n,\tau}_{\text{stop}}\setminus B^{n,\tau}_{\text{straight often}}\right)\le e^{-n}.$$

Proof. Recall Definition 6.9. We will show for all n sufficiently large,

$$P\left(E^{n,\tau}_{\text{stop}}\setminus\left(\bigcap_{I\in\mathcal{J}_L}\{|\mathcal{S}_\to(I)|\ge 2^{\gamma n}\}\right)\right)\le e^{-n}/2. \tag{6.40}$$

A similar consideration shows that the same estimate is true if we replace $\mathcal{S}_\to(I)$ by $\mathcal{S}_\leftarrow(I)$, and the claim then follows from the definition of $B^{n,\tau}_{\text{straight often}}$. Since the proof is very similar to the proof of Lemma 6.5, we will omitt some of the details.

Let $I\in\mathcal{J}_L$. We denote by R^I the ladderpath in $\mathbb{Z}^{[0,3c_1n[}$ which traverses I from left to right. For $t\in\mathbb{N}_0$ we define the event $E(t,I):=$

$$\left\{S(t+i)=R^I(i)\ \forall i\in[0,3c_1n[\text{ or }S(t+1+i)=R^I(i)\ \forall i\in[0,3c_1n[\right\}.$$

Let $n\ge c_{10}$ with c_{10} as in (6.4), and let $k\in[1,2^{\alpha n}]$. We set $t_{k,n}:=\tau_k+2^{2n-1}$ and we define random variables $Y_k(I)$ as follows: If $|S(\tau_k)|\le 2^n$ and $E(t_{k,n},I)$ does not hold, then we set $Y_k(I)=0$. Otherwise we set $Y_k(I)=1$. By Definition 6.9, we have

$$\begin{aligned}E^{n,\tau}_{\text{stop}}\setminus\left(\bigcap_{I\in\mathcal{J}_L}\{|\mathcal{S}_\to(I)|\ge 2^{\gamma n}\}\right)&\subseteq\bigcup_{I\in\mathcal{J}_L}E^{n,\tau}_{\text{stop}}\cap\left\{\sum_{k=1}^{2^{\alpha n}}Y_k(I)<2^{\gamma n}\right\}\\&\subseteq\bigcup_{I\in\mathcal{J}_L}\bigcup_{j=1}^{2^{\gamma n}}E^{n,\tau}_{\text{stop}}\cap\left\{\sum_{k=(j-1)2^{(\alpha-\gamma)n}+1}^{j\cdot 2^{(\alpha-\gamma)n}}Y_k(I)=0\right\}.\end{aligned} \tag{6.41}$$

Using the strong Markov property and induction (see the proof of Lemma 6.5, in particular (6.19), for a similar argument) we obtain for $n\ge c_{10}$ and $m,M\in[1,2^{\alpha n}]$ with $m\le M$

$$P\left(E^{n,\tau}_{\text{stop}}\cap\left\{\sum_{k=m}^{M}Y_k(I)=0\right\}\right)\le\left[\max_{x\in[-6\cdot 2^n,6\cdot 2^n]}P_x(E(2^{2n-1},I)^c)\right]^{M-m+1}. \tag{6.42}$$

By the local central limit theorem, there exist constants $c_{27}, c_{28} > 0$ such that for all $n \geq c_{27}$

$$\min_{x,z \in [-6\cdot 2^n, 6\cdot 2^n]} P_x\left(S(2^{2n-1}) = z \text{ or } S(2^{2n-1}+1) = z\right) \geq c_{28} 2^{-n}. \tag{6.43}$$

The probability that the random walk starting at x makes $3c_1 n - 1$ consecutive steps of maximum length to the right equals $\mu(L)^{3c_1 n - 1}$. Since all intervals in $\mathcal{J}_L$ are contained in $[-6\cdot 2^n, 6\cdot 2^n]$, we obtain

$$\min_{x\in[-2^n,2^n]} \min_{I\in\mathcal{J}_L} P_x(E(2^{2n-1}, I)) \geq c_{28} 2^{-n} \mu(L)^{3c_1 n - 1} = c_{29} 2^{-n} \mu(L)^{3c_1 n}$$

with $c_{29} := c_{28}\mu(L)^{-1}$. Combining the last inequality with (6.42), we obtain

$$P\left(E^{n,\tau}_{\text{stop}} \cap \left\{\sum_{k=m}^{M} Y_k(I) = 0\right\}\right) \leq \left(1 - c_{29} 2^{-n}\mu(L)^{3c_1 n}\right)^{M-m+1}. \tag{6.44}$$

From (6.41) and (6.44) it follows that

$$\begin{aligned}
&P\left[E^{n,\tau}_{\text{stop}} \setminus \left[\bigcap_{I\in\mathcal{J}_L} \{|\mathcal{S}_{\rightarrow}(I)| \geq 2^{\gamma n}\}\right]\right] \leq 2^{4+[1+\gamma]n}\left[1 - c_{29}2^{-n}\mu(L)^{3c_1 n}\right]^{2^{[\alpha-\gamma]n}} \\
&\leq 2^{4+[1+\gamma]n} \exp\left[2^{(\alpha-\gamma)n} \ln\left[1 - c_{29}2^{-n}\mu(L)^{3c_1 n}\right]\right] \\
&\leq 2^{4+[1+\gamma]n} \exp\left[-c_{29}2^{[\alpha-1-\gamma]n}\mu(L)^{3c_1 n}\right] \leq 2^{4+[1+\gamma]n} \exp\left[-c_{29}e^{c_{30}n}\right] \leq e^{-n}/2
\end{aligned}$$

for all n sufficiently large because $c_{30} = (\alpha - 1 - \gamma)\ln 2 + 3c_1 \ln\mu(L) > 0$ by our choice of α. ■

6.4 Alg^n reconstructs with high probability

Proof of Theorem 3.5. Suppose $\xi|[-2^n, 2^n] \preceq \mathrm{Alg}^n(\tau, \tilde{\chi}|\left[0, 2\cdot 2^{12\alpha n}\right[, \psi) \preceq \xi|[-4\cdot 2^n, 4\cdot 2^n]$. Assume $\psi \in \mathcal{C}^{[-kn,kn]}$ with $k \geq c_1 L$, $\psi \preceq \xi|[-2^n, 2^n]$, and assume $\xi|[-2^n, 2^n] \neq (1)_{[-2^n,2^n]}$. Then $\mathrm{Alg}^n(\tau, \tilde{\chi}|\left[0, 2\cdot 2^{12\alpha n}\right[, \psi)|[-kn, kn] = \psi$ by the definition of Alg^n (Definition 5.7) and the definition of $\mathsf{SolutionPiece}^n$ (Definition 5.6).

In order to show that Alg^n reconstructs with high probability, we combine Lemmas 6.4, 6.3, 6.2, and 6.1 to obtain

$$\begin{aligned}
E^{n,\tau}_{\text{stop}} \setminus E^{n,\tau}_{\text{reconstruct}} \quad \subseteq \quad & \left(E^{n,\tau}_{\text{stop}} \setminus B^{n,\tau}_{\text{all paths}}\right) \cup \left(B^n_{\text{few mistakes}}\right)^c \cup \left(B^n_{\text{ladder diff}}\right)^c \\
& \cup \left(B^{n,\tau}_{\text{majority}}\right)^c \cup \left(B^n_{\text{outside out}}\right)^c \cup \left(B^n_{\text{signals}}\right)^c \\
& \cup \left(B^n_{\text{recogn straight}}\right)^c \cup \left(E^{n,\tau}_{\text{stop}} \setminus B^{n,\tau}_{\text{straight often}}\right).
\end{aligned}$$

The claim follows from Lemmas 6.5, 6.6, 6.8, 6.9, 6.10, 6.12, 6.13, and 6.14. ■

References

[1] E. Artin. *The Gamma Function.* Holt, Rinehart and Winston, 1964.

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

[3] F. den Hollander. *Large deviations.* American Mathematical Society, Providence, RI, 2000.

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

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

[6] R. Durrett. *Probability: Theory and Examples.* Duxbury Press, Second edition, 1996.

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

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

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

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

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

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

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

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

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

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

[17] D. Levin, R. Pemantle, and Y. Peres. A phase transition in random coin tossing. Preprint, 2001.

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

[19] M. Löwe and H. Matzinger. Reconstruction of sceneries with correlated colors. Eurandom Report 99-032, 1999.

[20] M. Löwe and H. Matzinger. Scenery reconstruction in two dimensions with many colors. Eurandom Report 99-018, 1999.

[21] M. Löwe, H. Matzinger, and F. Merkl. Reconstructing a multicolor random scenery seen along a random walk path with bounded jumps. Eurandom Report 2001-030, 2001.

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

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

[24] H. Matzinger. Reconstructing a 2-color scenery by observing it along a simple random walk path. Eurandom Report 2000-003, 2000.

[25] 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, 2000.

[26] H. Matzinger and S. W. W. Rolles. Reconstructing a random scenery in polynomial time. Preprint.