Reconstructing a random scenery in polynomial time

Heinrich Matzinger* Silke W.W. Rolles†

March 1, 2002

Abstract

Benjamini asked whether the scenery reconstruction problem can be solved in polynomial time. In this article, we answer his question in the affirmative for an i.i.d. uniformly colored scenery on $\mathbb{Z}$ observed along a random walk path with bounded jumps. We assume the random walk is recurrent, can reach every integer with positive probability, and the number of possible single steps for the random walk exceeds the number of colors. We prove that a finite piece of scenery of length l around the origin can be reconstructed up to reflection and a small translation from the first $p(l)$ observations with high probability; here p is a polynomial and the probability that the reconstruction succeeds converges to 1 as $l \to \infty$. [1]

1 Introduction and Result

We call a coloring of $\mathbb{Z}$ with colors from the finite set $\mathcal{C} := \{1, 2, \ldots, C\}$ a *scenery*. The *scenery reconstruction problem* can be described as follows: Let ξ be a scenery and let $S := (S_k; k \in \mathbb{N}_0)$ be a recurrent random walk on $\mathbb{Z}$. If we are given the color record $\chi := (\xi(S_k); k \in \mathbb{N}_0)$ observed along the random walk path, can we almost surely reconstruct ξ from these observations (of course without knowing S)?

It is not hard to see that in general the reconstruction works only up to a reflection and translation. Recently, Lindenstrauss [7] proved the existence of uncountably many sceneries which cannot be reconstructed. Nevertheless it turns out that in many situations "typical" sceneries can be reconstructed if the scenery ξ is randomly colored and the random walk S is independent of ξ.

Work on the scenery reconstruction problem started with the question how much information can be extracted from the color record χ. This question was

*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

[1] Keywords: Scenery reconstruction, random walk, polynomial time

addressed in the 80s by Benjamini and independently by den Hollander and Keane (see [6]). Kesten [5] proved that a single defect in a 4-color random scenery can be detected if the scenery is i.i.d. uniformly colored.

In his Ph.D. thesis [11] (see also [12] and [13]) Matzinger studied the scenery reconstruction problem for i.i.d. uniformly colored 2-color sceneries. He showed that almost all sceneries can be almost surely reconstructed up to equivalence if they are observed along a simple random walk path (with holding); here we call two sceneries ξ and ξ' equivalent, $\xi \approx \xi'$, if ξ can be obtained from ξ' by a reflection and/or a translation. After Kesten had noticed that Matzinger's proof relies heavily on the skip-freeness of the random walk, Löwe, Matzinger, and Merkl [10] proved that scenery reconstruction still works if the random walk has i.i.d. increments with finite support. More precisely they made the following assumptions: The scenery on $\mathbb{Z}$ is i.i.d. uniformly colored with colors from the set $\mathcal{C}$. The random walk is independent of the scenery, has i.i.d. increments with finite support, is recurrent, and can reach every integer with positive probability. They require that there is at least one color more than possible jumps for the random walk. Under these assumptions it is proved in [10] that almost all sceneries can be almost surely reconstructed up to equivalence.

In this article, we refine the reconstruction result of Löwe, Matzinger, and Merkl. We make the same assumptions on scenery and random walk as in [10] with the only exception that we assume that the maximal jump lengths to the left and to the right of the random walk are equal. This additional assumption is made to keep the exposition as easy as possible. We believe that our result is true without this assumption. Below we prove that for n large a finite piece of scenery of length $l(n) = 10 \cdot 2^n + 1$ around the origin can be reconstructed with high probability from the first $2n^7 + 2 \cdot 2^{12\alpha n}$ observations with a constant $\alpha > 0$; thus the number of observations needed is polynomial in $l(n)$.

In order to state our main result we need some notation. All intervals are taken over the integers, e.g. $[a,b] := \{x \in \mathbb{Z} : a \le x \le b\}$. We write $f|D$ for the restriction of a function f to a set D. For two words $w \in \mathcal{C}^{[0,m]}$ and $w' \in \mathcal{C}^{[0,m']}$ with $m \le m'$ we write $w \preccurlyeq w'$ if there exists an interval $I' \subseteq [0,m']$ such that the restriction of w' to I' can be obtained from w by a reflection and/or a translation. We set for $n_0 \in \mathbb{N}$

$$n_1 := 2^{\lfloor \sqrt{n_0} \rfloor}, \quad n_2 := 2^{\lfloor \sqrt{n_1} \rfloor}; \tag{1}$$

here $\lfloor x \rfloor$ denotes the largest integer $\le x$. The dependence of n_2 on n_0 is always suppressed; we always write n_2 instead of using the more precise notation $n_2(n_0)$. Formally, our result reads as follows:

Theorem 1 *There exists $\alpha > 0$ such that for infinitely many $n_0 \in \mathbb{N}$ there exists a map $\mathcal{A}^{n_0}_{\text{initial}} : \mathcal{C}^{[0, n_0^{20} + n_2^7 + 2 \cdot 2^{12\alpha n_2}[} \to \mathcal{C}^{[-5 \cdot 2^{n_2}, 5 \cdot 2^{n_2}]}$ which is measurable with respect to the σ-algebras generated by the canonical projections such that for the events*

$$E^{n_0}_{\text{ini works}} := \left\{ \begin{array}{l} \xi| \left[-2^{n_2-1}, 2^{n_2-1}\right] \preccurlyeq \mathcal{A}^{n_0}_{\text{initial}} \left(\chi|[0, n_0^{20} + n_2^7 + 2 \cdot 2^{12\alpha n_2}[\right) \preccurlyeq \\ \xi| \left[-10 \cdot 2^{n_2}, 10 \cdot 2^{n_2}\right] \end{array} \right\}$$

the following holds:

$$\lim_{n_0\to\infty} P\left(E^{n_0}_{\text{ini works}}\right) = 1.$$

The algorithm $\mathcal{A}^{n_0}_{\text{initial}}$ gets as input the first $n_0^{20}+n_2^7+2\cdot 2^{12\alpha n_2}$ observations of the random walker and produces an output of length $10\cdot 2^{n_2}+1$. If the reconstruction is successful in the sense that the event $E^{n_0}_{\text{ini works}}$ holds, then the output of $\mathcal{A}^{n_0}_{\text{initial}}$ is a piece of the scenery ξ which is typically not centered around the origin. The output is a piece of scenery which is up to equivalence contained in $[-10\cdot 2^{n_2}, 10\cdot 2^{n_2}]$. Note that it is essential to reconstruct a finite piece of scenery close to the origin because the scenery is i.i.d. uniformly colored and therefore any finite color pattern can be almost surely found somewhere in the scenery.

In [14], the authors study the scenery reconstruction problem under the same assumptions as in this article and prove that almost all sceneries can be almost surely reconstructed up to equivalence if there are some errors in the observations. Scenery reconstruction results in different settings have e.g. been obtained by Löwe and Matzinger in [9] and [8]. Related work on the scenery distinguishing problem has been done by Benjamini and Kesten [1] and Howard in [4] and [3].

The remainder of this article is organized as follows: In Section 2, we formally define our setting. In Section 3, we review a result from [10]. Section 4 contains the definition of the reconstruction algorithm $\mathcal{A}^{n_0}_{\text{initial}}$. Section 5 contains some lemmas. In Section 6, we prove Theorem 1.

2 Setting

Let $C\geq 2$. We assume that the scenery $\xi := (\xi_k; k\in\mathbb{Z})$ is i.i.d. uniformly distributed over $\mathcal{C} := \{1,2,\dots,C\}$.

Let μ be a probability measure over $\mathbb{Z}$ with finite support $\mathcal{M}$ and let $S := (S_k; k\in\mathbb{N}_0)$ be a random walk on $\mathbb{Z}$ with μ-distributed increments, independent of ξ. We assume $|\mathcal{M}| < |\mathcal{C}|$, i.e. there is at least one color more than possible steps for the random walk. Furthermore we assume that $\mathcal{M}$ has greatest common divisor 1 and $\sum_{k\in\mathcal{M}} k\mu(k) = 0$; thus S can reach every integer with positive probability and is recurrent. To keep the notation simple we require $\max\mathcal{M} = |\min\mathcal{M}|$, i.e. the maximal jump lengths to the left and to the right agree. We denote the maximal jump length by L.

Let $\Omega_2\subseteq\mathbb{Z}^{\mathbb{N}_0}$ be the set of all paths with jump sizes $S_{k+1}-S_k\in\mathcal{M}$ for all $k\in\mathbb{N}_0$. We realize (ξ, S) as canonical projections on the probability space

$$\Omega := \left(\mathcal{C}^{\mathbb{Z}}, \Omega_2\right), \quad P := \nu^{\otimes\mathbb{Z}}\otimes Q_0$$

with the product σ-algebra generated by the canonical projections; here ν denotes the uniform distribution on the set of colors $\mathcal{C}$ and Q_0 the distribution of the random walk starting at the origin. We write $\chi := (\chi_k := \xi(S_k); k\in\mathbb{N}_0)$ for the observations of the scenery along the random walk path.

We define the shift $\Theta : \Omega \to \Omega$ by

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

Intuitively, Θ shifts the origin to the position of the random walker at time 1.

All constants keep there meaning throughout the article.

3 Review of a Result of Löwe, Matzinger, and Merkl

Löwe, Matzinger, and Merkl [10] showed the existence of measurable maps $\mathcal{A}^{n_2}$ which do "partial reconstructions". In order to define the reconstruction algorithm $\mathcal{A}^{n_0}_{\text{initial}}$, we use these partial algorithms. The maps $\mathcal{A}^{n_2}$ reconstruct with high probability a large piece of scenery around the origin if the observations χ start with a sufficiently large block of ones.

More formally: We define the event that $\mathcal{A}^{n_2}$ reconstructs correctly a piece of scenery around the origin:

$$E^{n_0}_{\text{recon}} := \left\{ \xi|[-2^{n_2}, 2^{n_2}] \preccurlyeq \mathcal{A}^{n_2}\left(\chi|[0, 2 \cdot 2^{12\alpha n_2}[\right) \preccurlyeq \xi|[-9 \cdot 2^{n_2}, 9 \cdot 2^{n_2}] \right\};$$

here $\alpha = \alpha(|\mathcal{C}|, \mu)$ is a sufficiently large positive constant independent of n_2. For $n \in \mathbb{N}$, we denote by $E_B(n)$ the event that the first $n+1$ observations are all equal to 1:

$$E_B(n) := \{\chi_k = 1 \text{ for all } k \in [0, n]\}.$$

For an interval $J \subseteq \mathbb{Z}$, we write $(1)_J$ for the piece of scenery in $\mathcal{C}^J$ which is constantly equal to 1. We define the event that there is "a long block of ones close to the origin":

$$\text{BigBlock} := \left\{ \begin{matrix} \text{There exists an integer interval } J_0 \subseteq \left[-2Ln_0^{20}, 2Ln_0^{20}\right] \\ \text{with } |J_0| \geq n_0^4 \text{ such that } \xi|J_0 = (1)_{J_0} \end{matrix} \right\}.$$

We denote by P_B the image of the conditional distribution $P(\cdot|E_B(n_0^{20}))$ under the shift $\Theta^{n_0^{20}}$, and we abbreviate $\tilde{P} := P_B(\cdot|\text{BigBlock})$. We set

$$\varepsilon_1(n_0) := P_B\left([E^{n_0}_{\text{recon}}]^c\right), \tag{2}$$

i.e. $\varepsilon_1(n_0)$ is the probability that $\mathcal{A}^{n_2}$ does not fulfill its task, and we observe

$$\begin{aligned} P_B\left([E^{n_0}_{\text{recon}}]^c\right) &\leq P_B\left([E^{n_0}_{\text{recon}}]^c \cap \text{BigBlock}\right) + P_B\left([\text{BigBlock}]^c\right) \\ &\leq \tilde{P}\left([E^{n_0}_{\text{recon}}]^c\right) + P_B\left([\text{BigBlock}]^c\right). \end{aligned} \tag{3}$$

In [10], Löwe, Matzinger, and Merkl prove $\lim_{n_0 \to \infty} \tilde{P}\left([E^{n_0}_{\text{recon}}]^c\right) = 0$. The second term on the right-hand side of (3) converges to 0 as $n_0 \to \infty$ by Lemma 3.3 in [10]. Hence it follows from (2) and (3) that

$$\lim_{n_0 \to \infty} \varepsilon_1(n_0) = 0. \tag{4}$$

4 The Algorithm $\mathcal{A}^{n_0}_{\text{initial}}$

In general it will not take "too long" until a long block of ones is observed in the observations; we then apply $\mathcal{A}^{n_2}$ to the observations collected right after this long block of ones.

Formally we let for $k \in \mathbb{N}_0$, Z_k be the Bernoulli random variable taking values in $\{0,1\}$ which is equal to one if and only if $\chi(kn_2^6 + j) = 1$ for all $j \in [0, n_0^{20}]$. We set

$$\bar{k} := \min\{k \geq 0 : Z_k = 1\}. \tag{5}$$

Definition 2 *We define* $\mathcal{A}^{n_0}_{\text{initial}} : \mathcal{C}^{[0, n_0^{20} + n_2^7 + 2\cdot 2^{12\alpha n_2}[} \to \mathcal{C}^{[-5\cdot 2^{n_2}, 5\cdot 2^{n_2}]}$ *as follows: If* $\bar{k} \leq n_2$, *then we define* $\mathcal{A}^{n_0}_{\text{initial}}\left(\chi|[0, n_0^{20} + n_2^7 + 2\cdot 2^{12\alpha n_2}[\right) :=$

$$\mathcal{A}^{n_2}\left(\chi|[\bar{k}n_2^6 + n_0^{20}, \bar{k}n_2^6 + n_0^{20} + 2\cdot 2^{12\alpha n_2}[\right).$$

Otherwise we define $\mathcal{A}^{n_0}_{\text{initial}}\left(\chi|[0, n_0^{20} + n_2^7 + 2\cdot 2^{12\alpha n_2}[\right) := (1)_{[-5\cdot 2^{n_2}, 5\cdot 2^{n_2}]}$, *i.e. the output equals the piece of scenery which is constantly equal to* 1.

In the remainder of the article we will prove that $\mathcal{A}^{n_0}_{\text{initial}}$ fulfills its task specified by Theorem 1.

5 Some Lemmas

For $k \in \mathbb{N}$ we define the events

$$\begin{aligned} E^{n_0,k}_{\text{no block}} &:= \left\{\sum_{i=0}^{k-1} Z_i = 0\right\}, \\ E^{n_0,k}_{\text{rw apart}} &:= \left\{\left|S\left(in_2^6\right) - S\left(kn_2^6\right)\right| > 2Ln_0^{20} \text{ for all } i < k\right\}. \end{aligned}$$

Note that $E^{n_0,k}_{\text{no block}} = \{\bar{k} \geq k\}$.

Lemma 3 *For* $n \in \mathbb{N}_0$, *let* $\mathcal{F}(n)$ *be the* σ*-algebra generated by the whole scenery and the random walk up to time* n: $\mathcal{F}(n) := \sigma(\xi, S(i); i \in [0,n])$. *There exists a constant* $c_3 > 0$ *such that for all* $n_0 \in \mathbb{N}$ *and all* $k \leq n_2$

$$P\left(\left[E^{n_0,k}_{\text{rw apart}}\right]^c \middle| \mathcal{F}\left(n_2^6(k-1) + n_0^{20}\right)\right) \quad \leq \quad c_3 n_0^{20} n_2^{-2}.$$

Proof. By the local central limit theorem (see e.g. [2], Theorem (5.2), page 132), the probability that the random walk after $n_2^6 - n_0^{20}$ steps hits a set containing at most $\left(4Ln_0^{20} + 1\right) n_2$ points is bounded by $c_4 n_2^{-3}\left(4Ln_0^{20} \cdot n_2\right) = c_3 n_0^{20} n_2^{-2}$; here $c_4 > 0$ is a constant independent of n_0 and $c_3 = 4Lc_4$. The claim follows. ■

The next lemma gives information about the tail probability of $\bar{k}$. We need some notation: For $k \in \mathbb{N}_0$, let $\bar{F}(k)$ denote the tail probability:

$$\bar{F}(k) := P\left(\bar{k} > k\right). \tag{6}$$

We set

$$\varepsilon_2(n_0) \quad := \quad \sqrt{\max\left(\frac{1}{n_0}, \varepsilon_1(n_0)\right)}, \tag{7}$$

$$k_{n_0} \quad := \quad \min\left\{k : \bar{F}(k) \le \varepsilon_2(n_0)\right\}. \tag{8}$$

Lemma 4 *There exists c_5 such that $k_{n_0} \le n_2$ for all $n_0 \ge c_5$.*

Proof. For $k \in \mathbb{N}_0$, we define the Bernoulli random variable $\bar{Z}_k$ which is equal to 1 if and only if $Z_k = 1$ or $E^{n_0,k}_{\text{rw apart}}$ does not hold. We have that

$$\left\{\sum_{k=0}^{n_2} \bar{Z}_k > 0\right\} \cap \bigcap_{k=0}^{n_2} E^{n_0,k}_{\text{rw apart}} \subset \left\{\bar{k} \le n_2\right\};$$

thus

$$\bar{F}(n_2) = P\left(\bar{k} > n_2\right) \le \sum_{k=0}^{n_2} P\left(\left[E^{n_0,k}_{\text{rw apart}}\right]^c\right) + P\left(\left\{\sum_{k=0}^{n_2} \bar{Z}_k = 0\right\}\right). \tag{9}$$

By Lemma 3,

$$\sum_{k=0}^{n_2} P\left(\left[E^{n_0,k}_{\text{rw apart}}\right]^c\right) \le (n_2+1)c_3 n_0^{20} n_2^{-2}. \tag{10}$$

We define the σ-algebra $\mathcal{G}_k := \sigma\left(Z_i, S(in_2^6) - S(jn_2^6), \xi(S(in_2^6)+z); 0 \le i < j \le k+1, z \in \left[-Ln_0^{20}, Ln_0^{20}\right]\right)$. The sequence $(\bar{Z}_k; k \ge 0)$ is adapted to the filtration $(\mathcal{G}_k; k \ge 0)$. Furthermore, by the definition of $\bar{Z}_k$,

$$\begin{aligned} P(\bar{Z}_k = 1|\mathcal{G}_{k-1}) &= 1E^{n_0,k}_{\text{rw apart}} P(Z_k = 1|\mathcal{G}_{k-1}) + 1\left[E^{n_0,k}_{\text{rw apart}}\right]^c \\ &\ge 1E^{n_0,k}_{\text{rw apart}} P\left(E_B(n_0^{20})\right); \end{aligned}$$

here $1A$ denotes the indicator function of the set A. For the first equality we used $E^{n_0,k}_{\text{rw apart}} \in \mathcal{G}_{k-1}$. For the last inequality we used that Z_k is measurable with respect to $\sigma\left(\xi(S(kn_2^6)+z); z \in \left[-Ln_0^{20}, Ln_0^{20}\right]\right)$ which is independent of $\mathcal{G}_{k-1}$. We abbreviate $p_{n_0} := P\left(E_B(n_0^{20})\right)$. The preceding estimate yields

$$\begin{aligned} P\left(\sum_{k=0}^{n_2} \bar{Z}_k = 0\right) &= P\left(\left\{\sum_{k=0}^{n_2} \bar{Z}_k = 0\right\} \cap \bigcap_{k=0}^{n_2} E^{n_0,k}_{\text{rw apart}}\right) \\ &= E\left[1\left[\left\{\sum_{k=0}^{n_2-1} \bar{Z}_k = 0\right\} \cap \bigcap_{k=0}^{n_2} E^{n_0,k}_{\text{rw apart}}\right] P\left(\bar{Z}_{n_2} = 0 \middle| \mathcal{G}_{n_2-1}\right)\right] \\ &\le (1-p_{n_0}) P\left(\left\{\sum_{k=0}^{n_2-1} \bar{Z}_k = 0\right\} \cap \bigcap_{k=0}^{n_2-1} E^{n_0,k}_{\text{rw apart}}\right); \end{aligned}$$

here E denotes the expectation with respect to P. Using an induction argument, we conclude

$$P\left(\sum_{k=0}^{n_2} \bar{Z}_k = 0\right) \leq (1-p_{n_0})^{n_2}. \tag{11}$$

In order to obtain a lower bound for p_{n_0}, first note $P\left(\xi(z) = 1 \forall z \in \left[-n_0^{11}, n_0^{11}\right]\right) = (1/C)^{2n_0^{11}+1}$. Furthermore, by Doob's inequality (see e.g. [2], page 250),

$$P\left(\max_{i\in[0,n_0^{20}]} |S(i)| > n_0^{11}\right) \leq n_0^{-22}\mathrm{Var}(S(n_0^{20})) = c_6 n_0^{-2}$$

with some constant $c_6 > 0$. Thus, we obtain for all n_0 sufficiently large

$$\begin{aligned} p_{n_0} &= P\left(E_B(n_0^{20})\right) \\ &\geq P\left(\left\{\xi(z) = 1 \forall z \in \left[-n_0^{11}, n_0^{11}\right]\right\} \cap \left\{\max_{i\in[0,n_0^{20}]} |S(i)| \leq n_0^{11}\right\}\right) \\ &\geq C^{-2n_0^{11}-1}(1 - c_6 n_0^{-2}) \geq C^{-2n_0^{11}-2}. \end{aligned} \tag{12}$$

It follows from (11) and (12) that

$$P\left(\sum_{k=0}^{n_2} \bar{Z}_k = 0\right) \leq \left(1 - C^{-2n_0^{11}-2}\right)^{n_2} \leq \exp\left(-n_2 C^{-2n_0^{11}-2}\right).$$

Using the definition of n_2 (1), we see that the right-hand side is bounded above by $2^{-1}n_0^{-1/2}$ for all n_0 sufficiently large. Combining this with (9) and (10) yields for all n_0 sufficiently large

$$\bar{F}(n_2) \leq (n_2+1)c_3 n_0^{20} n_2^{-2} + 2^{-1}n_0^{-1/2} \leq n_0^{-1/2} \leq \varepsilon_2(n_0);$$

for the last inequality we used the definition of $\varepsilon_2(n_0)$ (7). The claim follows from the definition of k_{n_0}. ■

We define

$$\begin{aligned} E_{\text{rw apart}}^{n_0} &:= \left\{\left|S\left(in_2^4\right) - S\left(\bar{k}n_2^4\right)\right| > 2Ln_0^{20} \text{ for all } i < \bar{k}\right\}, \\ E_{\text{ok}}^{n_0} &:= \left\{\bar{k} \leq k_{n_0}\right\} \cap E_{\text{rw apart}}^{n_0}. \end{aligned}$$

Lemma 5 *The following holds:*

$$\lim_{n_0\to\infty} P\left([E_{\text{ok}}^{n_0}]^c\right) = 0.$$

Proof. We observe that

$$P\left([E_{\text{ok}}^{n_0}]^c\right) = P\left(\bar{k} > k_{n_0}\right) + P\left(\left\{\bar{k} \leq k_{n_0}\right\} \cap \left[E_{\text{rw apart}}^{n_0}\right]^c\right).$$

By the definition of $\bar{F}$ and k_{n_0}, $P\left(\bar{k} > k_{n_0}\right) = \bar{F}(k_{n_0}) \le \varepsilon_2(n_0)$, which converges to 0 as $n_0 \to \infty$; recall the definition of $\varepsilon_2(n_0)$ and (4). Using Lemmas 3 and 4, we obtain for all n_0 sufficiently large

$$P\left(\left\{\bar{k} \le k_{n_0}\right\} \cap \left[E^{n_0}_{\text{rw apart}}\right]^c\right) \le \sum_{i=0}^{k_{n_0}} P\left(\left[E^{n_0,i}_{\text{rw apart}}\right]^c\right) \le c_3 n_0^{20} n_2^{-2}(n_2+1)$$

which converges to 0 as $n_0 \to \infty$ by the definition of n_2 (1). The claim follows. ■

6 Proof of Theorem 1

Proof of Theorem 1. The following holds:

$$P\left(\left[E^{n_0}_{\text{ini works}}\right]^c\right) \le P\left(\left[E^{n_0}_{\text{ini works}}\right]^c \cap E^{n_0}_{\text{ok}}\right) + P\left(\left[E^{n_0}_{\text{ok}}\right]^c\right).$$

By Lemma 5, the second term on the right-hand side converges to 0 as $n_0 \to \infty$. It suffices to prove that the same is true for the first term. We observe that

$$P\left(\left[E^{n_0}_{\text{ini works}}\right]^c \cap E^{n_0}_{\text{ok}}\right) = \sum_{k=0}^{k_{n_0}} P\left(\left[E^{n_0}_{\text{ini works}}\right]^c \cap E^{n_0}_{\text{ok}} \,|\, \bar{k} = k\right) P\left(\bar{k} = k\right) \tag{13}$$

By the definition of the shift Θ, we have for $m \ge 0$

$$\Theta^m(\xi, S) = \left(\xi\left(\cdot + S(m)\right), S(\cdot + m) - S(m)\right).$$

Consequently, we obtain for n_0 sufficiently large and $k \le k_{n_0}$

$$\Theta^{-kn_2^6 - n_0^{20}}\left(E^{n_0}_{\text{recon}}\right) \cap \left\{\bar{k} = k\right\} \subseteq E^{n_0}_{\text{ini works}} \cap \left\{\bar{k} = k\right\};$$

here we used that if $\bar{k} = k$, then the first observation which $\mathcal{A}^{n_0}_{\text{initial}}$ uses is $\chi\left(kn_2^6 + n_0^{20}\right)$ and for all n_0 sufficiently large,

$$\left|S\left(kn_2^6 + n_0^{20}\right)\right| \le \left(kn_2^6 + n_0^{20}\right) L \le \left(k_{n_0} n_2^6 + n_0^{20}\right) L \le \left(n_2^7 + n_0^{20}\right) L \le 2^{n_2 - 1}$$

because the random walker starts at the origin and jumps in each step at most a distance of L. Hence the reconstruction starts not at the origin, but at position $S\left(kn_2^6 + n_0^{20}\right)$ which has a distance $\le 2^{n_2-1}$ from the origin. This is why different pieces of scenery of ξ are concerned in the definitions of $E^{n_0}_{\text{recon}}$ and $E^{n_0}_{\text{ini works}}$. Thus (13) yields

$$\begin{aligned} & P\left(\left[E^{n_0}_{\text{ini works}}\right]^c \cap E^{n_0}_{\text{ok}}\right) && (14) \\ \le & \sum_{k=0}^{k_{n_0}} P\left(\left[\Theta^{-kn_2^6 - n_0^{20}}\left(E^{n_0}_{\text{recon}}\right)\right]^c \cap E^{n_0}_{\text{ok}} \,\middle|\, \bar{k} = k\right) P\left(\bar{k} = k\right) \\ \le & \sum_{k=0}^{k_{n_0}} P\left(\left[\Theta^{-kn_2^6 - n_0^{20}}\left(E^{n_0}_{\text{recon}}\right)\right]^c \,\middle|\, \{\bar{k} = k\} \cap E^{n_0}_{\text{ok}}\right) P\left(\bar{k} = k\right). && (15) \end{aligned}$$

Note that $E^{n_0,k}_{\text{no block}} = \{\bar{k} \geq k\}$ and $E^{n_0,k}_{\text{rw apart}} = E^{n_0}_{\text{rw apart}} \cap \{\bar{k} = k\}$. Consequently, for $k \leq k_{n_0}$,

$$\{\bar{k} = k\} \cap E^{n_0}_{\text{ok}} = \Theta^{-kn_2^6}\left(E_B(n_0)\right) \cap E^{n_0,k}_{\text{no block}} \cap E^{n_0,k}_{\text{rw apart}}. \tag{16}$$

By the Markov property, we know that $\left(S\left(t + kn_2^6\right) - S\left(kn_2^6\right); t \geq 0\right)$ is independent of $\left(S(t); t \in [0, kn_2^6]\right)$. The event $\Theta^{-kn_2^6}\left(E_B(n_0)\right)$ depends only on $\left(S\left(t + kn_2^6\right) - S\left(kn_2^6\right); t \geq 0\right)$ and $\left(\xi\left(S(kn_2^6) + z\right); z \in \left[-Ln_0^{20}, Ln_0^{20}\right]\right)$ because the random walker can jump in each step at most a distance of L. On the other hand, we have that $E^{n_0,k}_{\text{no block}} \cap E^{n_0,k}_{\text{rw apart}}$ only depends on $\left(S(t); t \in \left[0, kn_2^6\right]\right)$ and the scenery $\left(\xi\left(S(kn_2^6) + z\right); z \notin \left[-Ln_0^{20}, Ln_0^{20}\right]\right)$. Since the scenery ξ is i.i.d., the event $\Theta^{-kn_2^6}\left(E_B(n_0)\right)$ is independent of $E^{n_0,k}_{\text{no block}} \cap E^{n_0,k}_{\text{rw apart}}$. For any events A, B, C with the property that A and B are independent and $P(B) > 0$ the following inequality holds:

$$P(C|A \cap B) \leq \frac{P(C|A)}{P(B)}.$$

In our case this yields together with (16):

$$\begin{aligned}
& P\left(\left[\Theta^{-kn_2^6 - n_0^{20}}\left(E^{n_0}_{\text{recon}}\right)\right]^c \middle| \{\bar{k} = k\} \cap E^{n_0}_{\text{ok}}\right) \\
= \quad & P\left(\left[\Theta^{-kn_2^6 - n_0^{20}}\left(E^{n_0}_{\text{recon}}\right)\right]^c \middle| \Theta^{-kn_2^6}\left(E_B(n_0)\right) \cap E^{n_0,k}_{\text{no block}} \cap E^{n_0,k}_{\text{rw apart}}\right) \\
\leq \quad & \frac{P\left(\left[\Theta^{-kn_2^6 - n_0^{20}}\left(E^{n_0}_{\text{recon}}\right)\right]^c \middle| \Theta^{-kn_2^6}\left(E_B(n_0)\right)\right)}{P\left(E^{n_0,k}_{\text{no block}} \cap E^{n_0,k}_{\text{rw apart}}\right)} && (17) \\
= \quad & \frac{P\left(\left[\Theta^{-n_0^{20}}\left(E^{n_0}_{\text{recon}}\right)\right]^c \middle| E_B(n_0)\right)}{P\left(E^{n_0,k}_{\text{no block}} \cap E^{n_0,k}_{\text{rw apart}}\right)} = \frac{\varepsilon_1(n_0)}{P\left(E^{n_0,k}_{\text{no block}} \cap E^{n_0,k}_{\text{rw apart}}\right)}; && (18)
\end{aligned}$$

for the second but last inequality we used that the shift Θ preserves the measure P by Lemma 4.1 of [10]; for the last equality we used definition (2). Using the monotonicity of $\bar{F}$ and the definition of k_{n_0}, we obtain for all $k \leq k_{n_0}$

$$P\left(E^{n_0,k}_{\text{no block}}\right) = P\left(\bar{k} \geq k\right) = \bar{F}\left(k - 1\right) \geq \bar{F}\left(k_{n_0} - 1\right) \geq \varepsilon_2\left(n_0\right).$$

Combinig the last inequality with Lemma 3 and the fact $E^{n_0,k}_{\text{no block}} \in \mathcal{F}(t)$ for $t = n_2^6\left(k - 1\right) + n_0^{20}$, we obtain

$$\begin{aligned}
P\left[P\left[E^{n_0,k}_{\text{no block}} \cap E^{n_0,k}_{\text{rw apart}} \middle| \mathcal{F}(t)\right]\right] \quad & \geq \quad \left[1 - c_3 n_0^{20} n_2^{-2}\right] P\left[E^{n_0,k}_{\text{no block}}\right] \\
& \geq \quad \left[1 - c_3 n_0^{20} n_2^{-2}\right] \varepsilon_2\left(n_0\right).
\end{aligned}$$

For n_0 big enough we get that $\left(1 - c_3 n_0^{20} n_2^{-2}\right) > 1/2$. In that case we conclude

$$P\left(E^{n_0,k}_{\text{no block}} \cap E^{n_0,k}_{\text{rw apart}}\right) \geq \varepsilon_2\left(n_0\right)/2.$$

Combining the last estimate with (18), we obtain

$$P\left(\left[\Theta^{-kn_2^6}\left(E_{\text{recon}}^{n_0}\right)\right]^c \middle| \{\bar{k}=k\} \cap E_{\text{ok}}^{n_0}\right) \leq \frac{2\varepsilon_1(n_0)}{\varepsilon_2(n_0)}.$$

By the definition of $\varepsilon_2(n_0)$ (7), we have $\varepsilon_2(n_0) \geq \sqrt{\varepsilon_1(n_0)}$. Thus,

$$P\left(\left[\Theta^{-kn_2^6}\left(E_{\text{recon}}^{n_0}\right)\right]^c \middle| \{\bar{k}=k\} \cap E_{\text{ok}}^{n_0}\right) \leq 2\sqrt{\varepsilon_1(n_0)}.$$

Using (15) we get

$$P\left(\left[E_{\text{ini works}}^{n_0}\right]^c \cap E_{\text{ok}}^{n_0}\right) \leq 2\sqrt{\varepsilon_1(n_0)}.$$

It follows from (4) that $\lim_{n_0\to\infty} P\left(\left[E_{\text{ini works}}^{n_0}\right]^c \cap E_{\text{ok}}^{n_0}\right) = 0$. This completes the proof of Theorem 1. ■

References

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

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

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

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

[5] 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.

[6] 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.

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

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

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

[10] 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.

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

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

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

[14] H. Matzinger and S. W. W. Rolles. Reconstructing a random scenery observed with random errors along a random walk path. Preprint.