

Report 2000-003
**Reconstructing a 2-color Scenery
by Observing it along
a Simple Random Walk Path**
Heinrich Matzinger
ISSN 1389-2355

RECONSTRUCTING A 2-COLOR SCENERY BY OBSERVING IT ALONG A SIMPLE RANDOM WALK PATH

HEINRICH MATZINGER
Eurandom, P.O. Box 513,
5600 MB, Eindhoven, The Netherlands.
E-mail: matzinger@eurandom.tue.nl

January 21, 2000

Abstract

Let $\{\xi(n)\}_{n \in \mathbb{Z}}$ be a 2-color random scenery, that is a random coloration of \mathbb{Z} in two colors, such that the $\xi(i)$'s are i.i.d. Bernoulli variables with parameter $\frac{1}{2}$. Let $\{S(n)\}_{n \in \mathbb{N}}$ be a symmetric random walk starting at 0. Our main result shows that a.s., $\xi \circ S$ (the composition of ξ and S) determines ξ up to translation and reflection. In other words, by observing the scenery ξ along the random walk path S , we can a.s. reconstruct ξ up to translation and reflection. This result allows us to give a positive answer to the question of H. Kesten of whether one can a.s. detect a single defect in almost any 2-color random scenery by observing it only along a random walk path.

1 Introduction

A scenery will be defined to be a function from \mathbb{Z} to $\{0,1\}$. Let ξ and $\bar{\xi}$ be two sceneries. We say that ξ and $\bar{\xi}$ are equivalent iff there exists $a \in \mathbb{Z}$ and $b \in \{-1,1\}$ such that for all $x \in \mathbb{Z}$ we have that $\xi(x) = \bar{\xi}(a + bx)$. In this case we write $\xi \approx \bar{\xi}$. In other words, two sceneries are equivalent iff they can be obtained from each other by shift and/or reflection around the origin. In everything that follows $\{S(k)\}_{k \geq 0}$ will be a simple random walk on \mathbb{Z} starting at the origin. The question we are interested in is the following: given a scenery ξ which is unknown to us, can we "reconstruct" ξ if we are only given the scenery ξ seen along one path-realization of $\{S(k)\}_{k \geq 0}$. Thus, does one path realization of the process $\{\xi(S(k))\}_{k \geq 0}$ uniquely determine ξ ? The answer to the above question in those general terms is no. First, if ξ and $\bar{\xi}$ are equivalent, we can in general not know whether the observations come from ξ or from $\bar{\xi}$. Second, it is

clear that the reconstruction will in the best case work only almost surely. As a matter of fact, if the random walk $\{S(k)\}_{k \geq 0}$ would decide to walk only to the left (which it could do with probability zero), then we would have no information about the right side of the scenery ξ and thus not be able to reconstruct the scenery ξ . So the best we can hope for is a reconstruction algorithm which works almost surely. Eventually, Lindenstrauss in [12] has been able to exhibit sceneries which one can not reconstruct. However, it is possible to prove that a lot of typical sceneries can be reconstructed up to equivalence and almost surely. For this we take the scenery ξ to be itself the outcome of a random process which is independent of $\{S(k)\}_{k \geq 0}$ in such a way that the $\xi(k)$'s are i.i.d. Bernoulli with parameter $\frac{1}{2}$. Our main result states that, up to equivalence, almost every scenery ξ can be reconstructed a.s. (provided we are given the observation of ξ seen along a path of $\{S(k)\}_{k \geq 0}$ to do the reconstruction. By almost every scenery we mean almost every scenery with respect to the measure which makes the $\xi(k)$'s i.i.d. Bernoulli with parameter $\frac{1}{2}$). Let us now state our main theorem:

Theorem 1 *Let $\{S(k)\}_{k \geq 0}$ and $\{\xi(k)\}_{k \in \mathbb{Z}}$ be two processes independent of each other such that $\{S(k)\}_{k \geq 0}$ is a simple random walk starting at the origin and such that the $\xi(k)$'s are i.i.d. Bernoulli variables with parameter $1/2$. Then, one path realization of the process $\{\xi(S(k))\}_{k \geq 0}$ a.s. determines ξ up to equivalence. In other words, there exists a measurable function $A : \{0, 1\}^{\mathbb{N}} \rightarrow \{0, 1\}^{\mathbb{Z}}$ such that $P(A(\xi \circ S) \approx \xi) = 1$. (Here $\xi \circ S$, designates the path of the process $\{\xi(S(k))\}_{k \geq 0}$ that is, the scenery ξ observed along a path of $\{S(k)\}_{k \geq 0}$. By measurable, we mean, measurable with respect to the σ -algebras induced by the canonical coordinates on $\{0, 1\}^{\mathbb{N}}$ and on $\{0, 1\}^{\mathbb{Z}}$).*

This paper was motivated by Kesten's question to me of whether one can a.s. distinguish a single defect in almost any two color scenery. Let us explain what the scenery distinguishing problem is. Let $\xi, \eta : \mathbb{Z} \rightarrow \{0, 1\}$ and let $\{S(k)\}_{k \in \mathbb{N}}$ be a symmetric random walk on \mathbb{Z} . Let the process $\{\chi(k)\}_{k \in \mathbb{N}}$ be equal to either $\{\xi(S(k))\}_{k \in \mathbb{N}}$ or $\{\eta(S(k))\}_{k \in \mathbb{N}}$. Is it possible by observing only one path realization of $\{\chi(k)\}_{k \in \mathbb{N}}$ to say to which one of the two $\{\xi(S(k))\}_{k \in \mathbb{N}}$ or $\{\eta(S(k))\}_{k \in \mathbb{N}}$, $\{\chi(k)\}_{k \in \mathbb{N}}$ is equal to? (We assume that we know ξ and η .) If yes, we say that it is possible to distinguish between the sceneries ξ and η by observing them along a path of $\{S(k)\}_{k \in \mathbb{N}}$. Otherwise, when it is not possible to figure out almost surely by observing $\{\chi(k)\}_{k \in \mathbb{N}}$ alone whether $\{\chi(k)\}_{k \in \mathbb{N}}$ is generated on ξ or on η , we say that ξ and η are indistinguishable. The problem of distinguishing two sceneries was raised independently by I. Benjamini and by den Hollander and Keane. The motivation came from problems in ergodic theory, such as the T, T^{-1} problem (see Kalikow [7]) and from the study of various aspects of $\{\xi(S(k))\}_{n \in \mathbb{N}}$, where $\{\xi(k)\}_{k \in \mathbb{Z}}$ is random. (See Kesten and Spitzer in [9], Keane and den Hollander in [8], den Hollander in [3]). Benjamini and Kesten showed in [1] that one can distinguish almost any two random sceneries even when the random walk is in \mathbb{Z}^2 . (They assumed the sceneries to be random themselves, so that the $\xi(k)$'s and the $\eta(n)$'s are i.i.d. Bernoulli.) Kesten

in [10] proved that when the random sceneries are i.i.d. and have four colors, i.e., ξ and $\eta : \mathbb{Z} \rightarrow \{0, 1, 2, 3\}$, and differ only in one point, they can be a.s. distinguished. He asked whether this result might still hold with fewer colors. The main result of this paper directly implies that one can distinguish single defects in almost any scenery. In [14], we proved for the three color case that one can a.s. reconstruct almost every three color scenery. We also established that this implies, that one can distinguish single defects for almost all three color sceneries. In the two color case, i.e. in the case we consider in this paper, the same thing is true. This means that our result for scenery reconstruction implies that one can distinguish single defects in almost all sceneries. We state the following corollary to our main result without giving a proof. (The proof that our main result implies the following corollary is very similar to the one given in [14] for the three color case.)

Corollary 2 *Let \mathbb{B} designate the set of all two color sceneries. $\mathbb{B} = \{\xi : \mathbb{Z} \rightarrow \{0, 1\}\} = \{0, 1\}^{\mathbb{Z}}$. Let $(\mathbb{B}, \sigma(\mathbb{B}))$ denote the measurable space, where $\sigma(\mathbb{B})$ is the σ -algebra induced by the canonical coordinates on \mathbb{B} . Let P denote the probability measure on $(\mathbb{B}, \sigma(\mathbb{B}))$ obtained by assuming that the $\xi(i)$'s are i.i.d. Bernoulli variables with parameter $\frac{1}{2}$. Then there exist a $\sigma(\mathbb{B})$ -measurable set S , such that $P(S) = 1$ and such that for ever scenery $\xi \in S$ and every scenery η which is equal to ξ everywhere except in one point, we have that ξ and η are distinguishable.*

The above corollary says that there are many sceneries which one can distinguish or, in other words, that sceneries which are typical in a certain sense can be distinguished. However the above result becomes false if one tries to extend it to all pairs of sceneries which are not equivalent. Recently, Lindenstrauss [12] exhibited a non denumerable set of pairs of non-equivalent sceneries on \mathbb{Z} which he proved to be indistinguishable. Before that, Howard proved in [4], [5] and [6] that any two periodical sceneries of \mathbb{Z} which are not equivalent modulo translation and reflection are distinguishable and that one can a.s. distinguish single defects in periodical sceneries. Kesten asked in [11] whether this result would still hold when the random walk would be allowed to jump. He also asked what would happen in the two dimensional case. Loewe and Matzinger in [13] have been able to prove that one can a.s. reconstruct almost every scenery up to equivalence in two dimensions, provided the scenery has a lot of colors. However the problem of the reconstruction of two color sceneries in \mathbb{Z} seen along the random walk path of a recurrent random walk which is allowed to jump remains open. In our opinion, this is a central open problem at present. Eventually we should also mention that the two color scenery reconstruction problem for a scenery which is i.i.d. is equivalent to the following problem: let $\{R(k)\}_{k \in \mathbb{Z}}$ and $\{S(k)\}_{k \geq 0}$ be two independent simple random walks on \mathbb{Z} both starting at the origin and living on the same probability space. (Here we mean that $\{R(k)\}_{k \geq 0}$ and $\{R(-k)\}_{k \geq 0}$ are two independent simple random walks both starting at the origin.) Does one path realization of the iterated random walk $\{R(S(k))\}_{k \geq 0}$ uniquely determines the path of $\{R(k)\}_{k \in \mathbb{Z}}$ up to shift and reflection around the

origin? If one takes the representation of the scenery ξ as a nearest neighbor walk (which we will define later) for $\{R(k)\}_{k \in \mathbb{Z}}$ then it becomes immediately clear that the two problems are equivalent. We leave it to the reader to check the details. So the main result of this paper is equivalent to the following result for iterated nearest neighbor walks: one path realization of the iterated random walk $\{R(S(k))\}_{k \geq 0}$ a.s. uniquely determines the path of $\{R(k)\}_{k \in \mathbb{Z}}$ up to shift and reflection around the origin. This is a discrete analogous of the result of Burdzy [2] concerning the path of iterated Brownian motion.

2 The reconstruction

We are going to prove our main theorem by explicitly describing a reconstruction algorithm. This means that the measurable function $\mathcal{A} : \{0,1\}^{\mathbb{N}} \rightarrow \{0,1\}^{\mathbb{Z}}$ from our main theorem will be described in terms of an algorithm, the so-called reconstruction algorithm. Thus, $\mathcal{A}(\xi \circ S)$ will denote the outcome of our reconstruction algorithm if we give it the observations $\xi \circ S$. Of course the input, $\xi \circ S$ is infinite. So one may ask how an algorithm can process an infinite amount of data. We will see that our algorithm processes at each step only a finite amount of the data $\xi \circ S$, but as a limit when we let the algorithm work an infinite amount of time we get as output the infinite scenery $\mathcal{A}(\xi \circ S)$.

Our reconstruction algorithm works in two phases. First it reconstructs finite pieces of the scenery ξ . (Here ξ designates the path of the process $\{\xi(k)\}_{k \in \mathbb{Z}}$, i.e. $\xi : \mathbb{N} \rightarrow \{0,1\}$ is the scenery which we try to reconstruct). For each $n \in \mathbb{N}$ the reconstruction algorithm reconstructs a finite piece of scenery. The "partial reconstruction algorithm at level n " will take care of reconstructing that finite piece of scenery. Once we have constructed this sequence of finite pieces of sceneries, the second phase of our reconstruction algorithm begins: we assemble our pieces. We use an assemblage rule to do this. We will show, that when we assemble all our pieces following our assemblage rule, we get almost surely a scenery which is equivalent to ξ . Let us now first start by describing the second phase of our algorithm, that is the assemblage phase.

3 Assembling pieces of sceneries

We first need some definitions. We define a piece of scenery to be a function from an integer interval to \mathbb{Z} . By integer interval we mean the intersection between a real interval and \mathbb{Z} . We say that two pieces of sceneries $\varphi : D \rightarrow \{0,1\}$ and $\bar{\varphi} : \bar{D} \rightarrow \{0,1\}$ are equivalent iff there exists $a \in \mathbb{Z}$ and $b \in \{-1,1\}$ such that $a+bD = \bar{D}$ and for all $k \in D$, we have that $\varphi(k) = \bar{\varphi}(a+bk)$. In other words, two pieces of sceneries are equivalent iff they can be obtained from one another by shift and/or reflection. For the assemblage rule, we suppose that we are given a sequence of pieces of sceneries which we denote by $\xi^1 : D^1 \rightarrow \{0,1\}, \xi^2 : D^2 \rightarrow \{0,1\}, \dots, \xi^n : D^n \rightarrow \{0,1\}, \dots$. Our assemblage will produce another sequence of sceneries $\tilde{\xi}^1, \tilde{\xi}^2, \dots, \tilde{\xi}^n : \bar{D}^n \rightarrow \{0,1\}, \dots$ obtained by shifting and

reflecting the pieces of sceneries ξ^n . Thus for each n , $\bar{\xi}^n$ is obtained from ξ^n by shift and reflection, which means that ξ^n and $\bar{\xi}^n$ are equivalent. The end product of our assemblage is the scenery obtained by taking the pointwise limit of the pieces of sceneries $\bar{\xi}^n$. We will denote this limit if it exists by $\bar{\xi}$. (By pointwise limit of the $\bar{\xi}^n$'s, we mean $\bar{\xi}(k) := \lim_{n \rightarrow \infty} \bar{\xi}^n(k)$ for all $k \in \mathbb{Z}$ if this limit is well defined. For this limit to be well defined, we require that for each given $k \in \mathbb{Z}$ we have that for all but a finite number of n 's, $k \in \bar{D}^n$ and $\lim_{n \rightarrow \infty} \bar{\xi}^n(k)$ exists when we take the sequence $n \rightarrow \bar{\xi}^n(k)$ for those n 's for which $\bar{\xi}^n(k)$ is well defined.) If the pointwise limit is not well defined, then we let the assemblage algorithm break down. In the case of a break down, it is not important how $\bar{\xi}$ gets defined. The rule for the assemblage goes as follows: proceed inductively on n . Once $\bar{\xi}^n$ is well defined, choose any piece of scenery which coincides on an interval of length n with $\bar{\xi}^n$ and which is equivalent to ξ^{n+1} . Call that piece of scenery $\bar{\xi}^{n+1}$. If this is not possible, that is if there exists no piece of scenery which is equivalent to ξ^{n+1} and coincides with $\bar{\xi}^n$ on an interval of length n , then define $\bar{\xi}^{n+1}$ to be equal to $\bar{\xi}^n$. Let us at this stage describe this assemblage algorithm in a more formal way:

Algorithm 3 Proceed inductively on n . Put $\bar{\xi}^1 = \bar{\xi}^1$. Once $\bar{\xi}^n : \bar{D}^n \rightarrow \{0,1\}$ has been obtained, choose any integer interval $J^n \subset D^{n+1}$ of length n , for which there exists $a \in \mathbb{Z}$ and $b \in \{-1,1\}$ such that $a + bJ^n \subset \bar{D}^n$ and such that for all $k \in J^n$, we have $\bar{\xi}^n(a + bk) = \xi^{n+1}(k)$. Put $\bar{D}^{n+1} = a + bD^{n+1}$ and put for all $k \in \bar{D}^{n+1}$, $\bar{\xi}^{n+1}(k) = \bar{\xi}^n(b(k - a))$. If there is no interval J^n satisfying all the above conditions, then put $\bar{\xi}^{n+1} = \bar{\xi}^n$. For all $k \in \mathbb{Z}$, define $\bar{\xi}(k) = \lim_{n \rightarrow \infty} \bar{\xi}^n(k)$ if this limit is well defined. Otherwise let the assemblage algorithm break down.

The next step is to show that our assemblage rule works when it is given the right input. For this purpose let us assume that the sequence of pieces of sceneries $\xi^1, \xi^2, \dots, \xi^n, \dots$ which we will give to the assemblage algorithm as input, is such that all the pieces of sceneries ξ^n live on the same probability space as ξ (and thus these pieces of sceneries are random pieces of sceneries.) Let E^n be the event that there exists an interval I^n (may be random) such that ξ^n is equivalent to $\xi|I^n$ and such that $[-n, n] \subset I^n \subset [-n^3, n^3]$. (Here, $\xi|I^n$ designates the restriction of ξ to I^n . This means that $\xi|I^n$ designates a piece of scenery. From now on, we will write $f|A$ for the restriction of the function f to the set A . Also $[a, b]$ will denote an integer interval, i.e. the intersection between the real interval $[a, b]$ and \mathbb{Z} .) By saying that the assemblage algorithm gets right input, we mean that E^n holds for all but a finite number of n 's for the input of the assemblage algorithm. Let E_0^n designate the event that within the integer interval $[-(n+1)^3, (n+1)^3]$, there are no two different pieces of ξ of length n which are equivalent to each other. More precisely E_0^n is defined as follows: $E_0^n = \{\text{if } i_1, i_2, i_3, i_4 \in [-(n+1)^3, (n+1)^3] \text{ are such that } |i_1 - i_2|, |i_3 - i_4| = n \text{ and such that for all } k \in 0, 1, \dots, n, \text{ we have that } \xi(i_1 + k(i_2 - i_1))/|i_2 - i_1| = \xi(i_3 + k(i_4 - i_3))/|i_4 - i_3|, \text{ then } i_1 = i_3 \text{ and } i_2 = i_4\}$. It is clear, that if for all n but a finite number, E^n and E_0^n both hold,

then the assemblage rule will produce a scenery equivalent to ξ (in case it is given $\xi^1, \xi^2, \dots, \xi^n, \dots$ as input.). As a matter of fact since $\bar{\xi}$ is defined as a limit, it does not depend on a finite number of $\bar{\xi}^n$'s. Furthermore for those n 's for which both E^n and E_0^n hold, our assemblage rule puts $\bar{\xi}^n$ and $\bar{\xi}^{n+1}$ in the "right relative position to each other". We leave it to the reader to check this more in detail. Next we are going to show that almost surely E_0^n holds for all n but a finite number of n 's. This proves then, that our assemblage algorithm produces a.s. a piece of scenery equivalent to ξ , as long as it is given correct input, i.e. as long as E^n holds for all but a finite number of n 's.

Lemma 4 *The probability of the complement of E_0^n is finitely summable over n . In other words: if E_0^{nc} designates the complement of E_0^n , then $\sum_{n=1}^{\infty} P(E_0^{nc}) < \infty$. Thus, a.s., E_0^n holds for all but a finite number of n 's.*

Proof. Let $i_1, i_2, i_3, i_4 \in [-(n+1)^3, (n+1)^3]$ be non-random integer numbers such that $|i_1 - i_2| = n$ and $|i_3 - i_4| = n$ and such that $(i_1, i_2) \neq (i_3, i_4)$. Then there exists a set $K \subset \{0, 1, \dots, n\}$ of cardinality $\lfloor n/2 \rfloor$ such that $(i_1 + K(i_2 - i_1)/|i_2 - i_1|) \cap (i_3 + K(i_4 - i_3)/|i_4 - i_3|) = \emptyset$. Thus, $P(\xi(i_1 + k(i_2 - i_1)/|i_2 - i_1|) = \xi(i_3 + k(i_4 - i_3)/|i_4 - i_3|) \text{ for all } k \in K) = 0.5^{\lfloor n/2 \rfloor}$. Thus, $P(\xi(i_1 + k(i_2 - i_1)/|i_2 - i_1|) = \xi(i_3 + k(i_4 - i_3)/|i_4 - i_3|) \text{ for all } k \in \{0, 1, \dots, n\}) \leq 0.5^{\lfloor n/2 \rfloor}$. However, there are at most $(2n+3)^{12}$ quadruples (i_1, i_2, i_3, i_4) such that $i_1, i_2, i_3, i_4 \in [-(n+1)^3, (n+1)^3]$. This implies that $P(E_0^{nc}) \leq (2n+3)^{12} \times 0.5^{\lfloor n/2 \rfloor}$. The expression on the right side of the previous inequality is finitely summable over n and thus $\sum_{n=1}^{\infty} P(E_0^{nc}) < \infty$. ■

Since, E_0^n holds a.s. for all but a finite number of n 's, we have reduced the problem of reconstructing ξ to the problem of constructing a sequence of pieces of sceneries $\xi^1, \xi^2, \dots, \xi^n, \dots$ for which E^n holds for all n but a finite number. In the next section we are going to define for each n , the so called, partial reconstruction algorithm at level n . The partial reconstruction algorithm at level n produces as output the piece of scenery ξ^n . With the partial reconstruction algorithms at the different levels we will construct the sequence of pieces of sceneries $\xi^1, \xi^2, \dots, \xi^n, \dots$ which we will then assemble in order to get a scenery equivalent to ξ . Thus, ξ^n will denote the piece of scenery, which is the outcome of the partial reconstruction algorithm at level n . The scenery which we obtain as a limit by applying the assemblage algorithm to the sequence of pieces of sceneries ξ^n will be denoted by $\bar{\xi}$. To prove that one can reconstruct ξ a.s. and up to equivalence, that to prove our main theorem it remains to define the partial reconstruction algorithm at level n and to prove that for all but a finite number of n 's, there exists an integer interval (maybe random) I^n such that ξ^n is equivalent to $\xi|I^n$ where $[-n, n] \subset I^n \subset [-n^3, n^3]$.

4 Partial reconstruction algorithm at level n

We are now going to explain the partial reconstruction algorithm at level n . As mentioned already, the outcome of the partial reconstruction algorithm at level

n is a piece of scenery denoted by ξ^n . The goal of the reconstruction algorithm at level n is, by using only the observations $\xi \circ S : k \mapsto \xi(S(k)) \mathbb{N} \rightarrow \{0, 1\}$ as input, to construct a piece of scenery which, with high probability, is going to be equivalent to the restriction of ξ to an interval of length order n around the origin. More precisely, let E^n be the event that there exists an interval I^n (may be random) such that ξ^n is equivalent to $\xi|_{I^n}$ and such that $[-n, n] \subset I^n \subset [-n^3, n^3]$. We are going to prove that the probability of the event of E^n is close to 1, in the sense that the probability of the complements of E^n is finitely summable over n . Let us first explain in an informal way the main ideas behind the partial reconstruction algorithm at level n . Assume for a moment that the scenery ξ , instead of being a two color scenery, would be a four color scenery, i.e. $\xi : \mathbb{Z} \rightarrow \{0, 1, 2, 3\}$. Let us imagine furthermore, that for two integers x, y we have $\xi(x) = 2$ and $\xi(y) = 3$, but outside x and y the scenery has everywhere color 0 or 1. Then, we could reconstruct the portion of the scenery ξ lying between x and y . As a matter of fact, since the random walk $\{S(k)\}_{k \geq 0}$ is recurrent it would a.s. go at least once, (and hence infinitely often,) in the shortest possible way from the point x to the point y . Since we are given the "infinite observations" $\xi \circ S$ we can know what the distance between x and y is. As a matter of fact the distance between x and y is the shortest time laps that a 3 will ever appear in the observations $\xi \circ S$ after a 2. Thus, we can observe all these shortest passages by $\{S(k)\}_{k \geq 0}$ from x to y because they correspond exactly to the times in the observations where the color 3 appears shortest after color 2. But when the random walk $\{S(k)\}_{k \geq 0}$ goes in a shortest possible way from x to y it goes in a straight way which means that between the time it is at x and until it reaches color y it only moves in one direction (when it performs a straight crossing). During that time, the random walk $\{S(k)\}_{k \geq 0}$ reveals the portion of ξ lying between x and y . More precisely, if the couple of integers t_1, t_2 , where $\xi(S(t_1)) = 2$ and $\xi(S(t_2)) = 3$ with $t_2 > t_1$, minimizes $|t_2 - t_1|$ under the condition that $\xi(S(t_1)) = 2$ and $\xi(S(t_2)) = 3$, then the piece of scenery obtained by restricting the scenery ξ to the interval $[\min\{x, y\}, \max\{x, y\}]$ is equivalent to the observations $\xi \circ S$ restricted to the integer interval $[t_1, t_2]$. (Recall that we observe $\xi \circ S : k \mapsto \xi(S(k)) \mathbb{N} \rightarrow \{0, 1\}$).

The main difficulty we have to deal with is that the scenery ξ does not have two extra colors but has only the colors 0 and 1. So instead of the two colors 2 and 3, we will have to use certain patterns in the observations which are likely to appear when the random walk $\{S(k)\}_{k \geq 0}$ is in one specific place in the scenery. We will then look for a shortest time for a pattern A to appear after a pattern B in the observations. During such a shortest time interval, we will assume that the observations $\xi \circ S$ are a copy of a piece of the scenery ξ between the point where the random walk reads the pattern A and the place where it reads the pattern B. There will mainly be four difficulties:

a) if x and y designate the approximate locations in the scenery which generate the patterns A and B, then we want x and y to be on opposite sides of 0. As a matter of fact, for E^n to hold, we need to reconstruct the scenery ξ on an interval containing $[-n, n]$. (Note, that since the scenery is assumed to be i.i.d., saying that we reconstruct a finite piece of ξ without mentioning where

that finite piece is located doesn't make sense: every finite piece occurs a.s. up to translation infinitely often in the scenery ξ . Thus, it is essential to have some control over where that reconstructed piece is located.)

b) $|x|$ and $|y|$ must both be between n and n^3 .

c) Because the $\xi(k)$'s are i.i.d., every finite piece of ξ will appear infinitely often up to translation in ξ . Thus, if at a point x , ξ is likely to generate a certain pattern in the observations (when $\{S(k)\}_{k \geq 0}$ is at x), then there will be infinitely many places in the scenery likely to produce that same finite pattern. Thus, we need the other places in the scenery which are likely to generate the patterns A and B, to be really far away, so that we first get a straight crossing between x and y before $\{S(k)\}_{k \geq 0}$ goes to an other spot likely to generate the same pattern in the observations. This will in general not succeed, so we will use a third pattern C, and only look at when the patterns A and B occur not too long after we observed C. C will be taken such that the next time a spot in the scenery occurs likely to generate the pattern C in the observations is really far away from the points x and y . Furthermore in a certain vicinity of C there will only be the places x and y which are likely to generate the patterns A and B.

Let us now explain what these patterns are. We first need the following definitions. Let D be an integer interval. Then we call a function T from D to \mathbb{Z} a nearest neighbor walk, iff for each $x, y \in D$ with $|x - y| = 1$, we have that $|T(x) - T(y)| = 1$. In other words, a nearest neighbor walk represents a movement in integer time on the integers, such that each time unit, it crosses an unit interval, i.e. each unit time interval we go one unit to the left or one unit to the right. (For example, the path S of $\{S(k)\}_{k \geq 0}$ is a.s. a nearest neighbor walk.) Let $\varphi : \mathbb{Z} \rightarrow \{0, 1\}$ be one of the four possible 4-periodic sceneries, where the period is 0011. Then, this scenery φ has a very particular property. Each point in the scenery φ has its two neighboring points (to the left and to the right at distance one), such that one of the neighboring points has color 1 and the other has color 0. This has some very important consequences: Let $D = [d_1, d_2]$ be an integer interval, then for each color record $\chi : D \rightarrow \{0, 1\}$, there exists one and only one nearest neighbor walk T generating the sequence χ on φ , i.e. such that $\varphi \circ T = \chi$ and starting in a specific point, i.e. such that $T(d_1)$ is given.. (We also need to make sure that the nearest neighbor walk starts at a point with the right color, i.e. the point $T(d_1)$ where T starts must be chosen such that $\varphi \circ T(d_1) = \chi(d_1)$). Furthermore, very much unlike other sceneries, φ has the property, that once we know the position of a nearest neighbor walk T at one point in time and the observations $\varphi \circ T$ the nearest neighbor walk produces on the scenery φ , we can immediately reconstruct the nearest neighbor walk T . As a matter of fact, we can proceed inductively on the time: If at time t we know where the nearest neighbor walk T is, then at time $t + 1$ it must be either one to the right or one to the left from where it was at time t . Among these two possible points for the position of T at time $t + 1$ one has color 1 and one has color 0. Thus, by looking at the color record produced by T on φ , we see the color observed at time $t + 1$ and thus know on which of the two possible points T is at time $t + 1$. We are now going to define the

representation of the scenery ξ as a nearest neighbor walk R using what we have explained above about the four periodic sceneries with period 0011. Basically the nearest neighbor walk R which represents the scenery ξ is simply the only nearest neighbor walk generating ξ on the periodic scenery φ and starting at the origin. However if R is to start at the origin we need $\varphi(0) = \xi(0)$. We will thus define two periodical sceneries φ^0 and φ^1 having both same period 0011, but one of them being for the case $\xi(0) = 0$ and the other being for the case $\xi(0) = 1$. This is done in the following way: Let $\varphi^0, \varphi^1 : \mathbb{Z} \rightarrow \{0, 1\}$ be two sceneries both with period 0011 and such that $(\varphi^0(0), \varphi^0(1), \varphi^0(2), \varphi^0(3)) = (0, 0, 1, 1)$ and $(\varphi^1(0), \varphi^1(1), \varphi^1(2), \varphi^1(3)) = (1, 1, 0, 0)$. Let φ be the random scenery which is equal to φ^0 when $\xi(0) = 0$ and equal to φ^1 when $\xi(0) = 1$. Note at this stage, that φ is only random to the extent that $\xi(0)$ is. However, when we know the observations $\xi \circ S$ we also know $\xi(0)$. (This is so, because $\xi(0) = \xi(S(0))$, since the random walk $\{S(k)\}_{k \geq 0}$ starts at the origin.) Thus, when we have the observations $\xi \circ S$, we also know whether $\varphi = \varphi^0$ or $\varphi = \varphi^1$. Let $R : \mathbb{Z} \rightarrow \mathbb{Z}$ be defined to be the only nearest neighbor walk starting at the origin and generating ξ on φ . This means that we request that $R(0) = 0$ and that for all $k \in \mathbb{Z}$ we have $\varphi(R(k)) = \xi(k)$.

How can we now use the just defined representation of the scenery ξ as a nearest neighbor walk to reconstruct ξ ? When we are given the observations $\xi \circ S$ we don't know R a priori. However, the iterated nearest random walk $R \circ S : k \mapsto R(S(k)) : \mathbb{N} \rightarrow \mathbb{Z}$ is well known to us if we have the observations $\xi \circ S$. (We leave it to the reader to check that $R \circ S$ is indeed a nearest neighbor walk.) As a matter of fact, since composition of functions is associative, we get $\varphi \circ (R \circ S) = (\varphi \circ R) \circ S = \xi \circ S$. Now, the expression on the left side of the previous equation is the observations produced by the nearest neighbor walk $R \circ S$ on the scenery φ , whilst the expression on the right are the observations we are given as input to reconstruct the scenery ξ and which are thus known to us. Thus the previous equation means that the nearest neighbor walk $R \circ S$ generates the observations $\xi \circ S$ on the scenery φ . Now the nearest neighbor walk $R \circ S$ starts at the origin and we also know whether $\varphi = \varphi^0$ or $\varphi = \varphi^1$. But, we saw that a nearest neighbor walk is uniquely determined once we know where it starts and the observations it produces on the periodic scenery φ^0 or φ^1 . Thus, since we know $\xi \circ S$, we also know $R \circ S$. So we can from now on assume that $R \circ S$ is known to us and use $R \circ S$ for our partial reconstruction algorithms. To explain how we can use $R \circ S$ for our partial reconstruction algorithms we will need a few definitions. Let $T : D \rightarrow \mathbb{Z}$ be a nearest neighbor walk. Let $t_1, t_2 \in D$ be two integer numbers and let $x_1, x_2 \in \mathbb{Z}$ be two integer numbers different from each other. Then, we call (t_1, t_2) a crossing by T of (x_1, x_2) iff $(T(t_1), T(t_2)) = (x_1, x_2)$ and for all integer t strictly between t_1 and t_2 , we have that $T(t)$ is strictly between x_1 and x_2 . If $t_2 > t_1$ we say that the crossing (t_1, t_2) is positive, otherwise we say that it is negative. If $|t_1 - t_2| = |x_1 - x_2|$ we say that the crossing (t_1, t_2) is straight. Note that a crossing corresponds to going from a point x_1 to a point x_2 such that we leave the point x_1 immediately after we were there and then don't go back to x_1 before we haven't arrived in x_2 . The straight crossing corresponds to the idea of going from a point x_1 to a point

x_2 in a shortest possible way, i.e. going from a point x_1 to a point x_2 without going back and forth but walking only in one direction between x_1 and x_2 . Let (t_1, t_2) be a crossing by T of (x_1, x_2) and (t_3, t_4) be a crossing by T of (x_3, x_4) . Then, we say that (t_3, t_4) is the first crossing by T of (x_3, x_4) during (t_1, t_2) iff $t_3, t_4 \in [\min\{t_1, t_2\}, \max\{t_1, t_2\}]$ and (t_3, t_4) is the crossing by T of (x_3, x_4) which lies in $[\min\{t_1, t_2\}, \max\{t_1, t_2\}]$ (i.e. $t_3, t_4 \in [\min\{t_1, t_2\}, \max\{t_1, t_2\}]$) and is closest to t_1 . (Note, that this definition makes sense, because the different crossings (t_{3i}, t_{4i}) by a nearest random walk T of an interval (x_3, x_4) have their intervals $[\min(t_{3i}, t_{4i}), \max(t_{3i}, t_{4i})]$ mutually disjoint. Thus it makes sense to speak of the crossing by T of (x_3, x_4) which lies in $[\min\{t_1, t_2\}, \max\{t_1, t_2\}]$ and is closest to t_1 .) The idea behind the first crossing is simple. When we follow the path of T from t_1 to t_2 and by doing so cross for the first time from x_3 to x_4 , then we are at the so called first crossing by T of (x_3, x_4) during (t_1, t_2) . Next we are going to state the main properties about iterated nearest neighbor walks. We leave the proof to the reader.

Lemma 5 *Let $T_a : D_a \rightarrow D_b$ and $T_b : D_b \rightarrow \mathbb{Z}$ be two nearest neighbor walks. Then, the composition $T_b \circ T_a : D_a \rightarrow \mathbb{Z}$ $k \mapsto T_b(T_a(k))$ is itself a nearest neighbor walk. Furthermore, (t_1, t_2) is a crossing by $T_b \circ T_a$ of (y_1, y_2) iff (t_1, t_2) is a crossing by T_a of a crossing by T_b of (y_1, y_2) . In other words (t_1, t_2) is a crossing by $T_b \circ T_a$ of (y_1, y_2) , iff (t_1, t_2) is a crossing by T_a and $(T_a(t_1), T_a(t_2))$ is a crossing by T_b of (y_1, y_2) . Furthermore, (t_1, t_2) is a straight crossing by $T_b \circ T_a$ of (y_1, y_2) iff (t_1, t_2) is a straight crossing by T_a and $(T_a(t_1), T_a(t_2))$ is a straight crossing by T_b of (y_1, y_2) . Eventually, let (y_3, y_4) be a couple of integer numbers, then (t_3, t_4) is the first crossing by $T_b \circ T_a$ of (y_3, y_4) during (t_1, t_2) iff (t_3, t_4) is the first crossing by T_a of $(T_a(t_3), T_a(t_4))$ during (t_1, t_2) and $(T_a(t_3), T_a(t_4))$ is the first crossing by T_b of (y_3, y_4) during $(T_a(t_1), T_a(t_2))$.*

How can we now use this lemma for our reconstruction purposes? Let us imagine that (y_1, y_2) and (y_3, y_4) are both such that R crosses each one of them only once and that y_2 and y_3 lie between y_1 and y_4 , that is $y_2, y_3 \in [\min\{y_1, y_4\}, \max\{y_1, y_4\}]$. (This is a.s. never going to be the case, because R is recurrent. We just use this assumption for pedagogical purposes at this stage.) Let (x_1, x_2) designate the only crossing by R of (y_1, y_2) and let (x_3, x_4) be the only crossing by R of (y_3, y_4) . Then, we could reconstruct almost surely and up to equivalence the piece of scenery obtained by restricting the scenery ξ to the interval between y_1 and y_4 , i.e. we could reconstruct up to equivalence the piece of scenery $\xi|[\min\{y_1, y_4\}, \max\{y_1, y_4\}]$. As a matter of fact, by lemma 5, when we observe a crossing (t_1, t_2) by $R \circ S$ of (y_1, y_2) , we know that this crossing is also a crossing by S of a crossing by R of (y_1, y_2) . However, we assumed that there is only one crossing (x_1, x_2) by R of (y_1, y_2) , and thus the crossing (t_1, t_2) by S must be a crossing by S of the crossing (x_1, x_2) . Thus, in this case, whenever we observe a crossing (t_1, t_2) by $R \circ S$ of (y_1, y_2) , we know where S is: $S(t_1) = x_1$ and $S(t_2) = x_2$. The same holds true of course for (y_3, y_4) . Now, when the random walk S goes from x_1 to x_4 in a shortest possible way, i.e. in a straight way, it will first cross (x_1, x_2) and then (x_3, x_4) . Thus, S produces, by

going from x_1 to x_4 at shortest possible time from each other, a crossing by $R \circ S$ of (y_1, y_2) followed by a crossing by $R \circ S$ of (y_3, y_4) . These crossings appear in the shortest possible time after each other. Thus, in order to reconstruct $\xi[\min\{y_1, y_4\}, \max\{y_1, y_4\}]$, we simply need to take a crossing (t_1, t_2) by $R \circ S$ of (y_1, y_2) and a crossing (t_3, t_4) by $R \circ S$ of (y_3, y_4) minimizing $|t_1 - t_4|$. If $|t_1 - t_4|$ is minimal, then the observations $\xi \circ S$ restricted to the time between t_1 and t_4 are a piece of scenery equivalent to $\xi[\min\{y_1, y_4\}, \max\{y_1, y_4\}]$. (Because then, during time t_1 to t_4 , the random walk goes from x_1 to x_4 in a straight way, thus revealing in the observations during that time, the piece of ξ comprised between the points x_1 and x_4 .) Now, the problem is that $\{R(k)\}_{k \in \mathbb{Z}}$ is recurrent and thus the above described reconstruction method doesn't work. (As a matter of fact, $\{R(k)\}_{k \in \mathbb{Z}}$ is a simple random walk starting at the origin as one can easily check.) Thus, all the pairs of integers (y_1, y_2) will get crossed by $\{R(k)\}_{k \in \mathbb{Z}}$ infinitely often a.s. Thus, if we observe different crossing by $R \circ S$ of (y_1, y_2) we can no longer be sure that as crossing by S , they are on the "same spot". Rather, these crossings by $R \circ S$ of (y_1, y_2) can be crossings by S of different crossings by R of (y_1, y_2) . In order to still be able to find out when " S is at the same spot", we need to be able to recognize when two crossings by $R \circ S$ of the same (y_1, y_2) are also crossings by S of the same crossing by R and when they are not. In order to achieve this goal, we are going to introduce a statistical test. This statistical test will be called "Test for two crossings by $R \circ S$ to be crossings by S of the same place". Let y_1, y_2 be two integers lying on the same side of 0 and such that $|y_1 - y_2| = 3n$. Then, we define for $i = 1, 2, \dots, n$ the i -th three unit interval of (y_1, y_2) to be equal to the ordered couple of integer points $(y_1 + 3(i-1)(y_2 - y_1)/|y_2 - y_1|, y_1 + 3i(y_2 - y_1)/|y_2 - y_1|)$. In other words, we partition (y_1, y_2) in n oriented integer intervals of length 3 having same orientation as (y_1, y_2) . Let (t_1, t_2) be a crossing by a nearest neighbor walk T of (y_1, y_2) . (Again we assume that $|y_1 - y_2| = 3n$.) We define the "associated word of the crossing (t_1, t_2) by T " to be the binary word $w = (w(1), \dots, w(i), \dots, w(n))$ with n bits such that $w(i) = 1$ if the first crossing of the i -th three unit interval of (y_1, y_2) by T during (t_1, t_2) is straight and $w(i) = 0$ otherwise. Thus, the binary word associated with a crossing records which of the three unit intervals (of the couple (y_1, y_2) which gets crossed), gets crossed for the first time during the crossing in a straight way and which don't.

Test for two crossings by $R \circ S$ to be crossings by S of the same place

Let y_1, y_2 be two integers such that $|y_1 - y_2| = 3n$. Let (t_1, t_2) be a crossing by $R \circ S$ of (y_1, y_2) having associated binary word $w = (w(1), \dots, w(i), \dots, w(n))$. Let (\bar{t}_1, \bar{t}_2) be another crossing by $R \circ S$ of (y_1, y_2) having associated binary word $\bar{w} = (\bar{w}(1), \dots, \bar{w}(i), \dots, \bar{w}(n))$. We want to test whether (t_1, t_2) and (\bar{t}_1, \bar{t}_2) are crossings by S of the same place, i.e. if $(S(t_1), S(t_2)) = (S(\bar{t}_1), S(\bar{t}_2))$ holds. Our test statistic will be the number of common straight crossings, i.e. $\sum_{i=1}^n w(i) \times \bar{w}(i)$. When $\sum_{i=1}^n w(i) \times \bar{w}(i) \geq n((\frac{3}{4})^3 + (\frac{3}{4})^4)/2$, we decide that the crossings (t_1, t_2) and (\bar{t}_1, \bar{t}_2) are crossings by S of the same place, i.e. that $(S(t_1), S(t_2)) = (S(\bar{t}_1), S(\bar{t}_2))$, otherwise we reject that hypothesis.

What is now the distribution of the test statistic $\sum_{i=1}^n w(i) \times \bar{w}(i)$ under each of the two hypothesis? To be able to answer that question precisely, we are going to introduce a numeration of the crossings. (This is a relatively important point, since if we introduce the wrong numeration we get extremely complicated distributions for the associated words.) Let y_1 and y_2 be two integers different from each other. Let $T : D \rightarrow \mathbb{Z}$ be a nearest neighbor walk. Then, for $i > 0$, we call i -th crossing by T of (y_1, y_2) , the i -th crossing (t_{1i}, t_{2i}) by T of (y_1, y_2) after 0 (if it exists). This means that (t_{1i}, t_{2i}) is the crossing by T of (y_1, y_2) such that $t_{1i}, t_{2i} > 0$ and there are exactly $i - 1$ crossings (t_{1j}, t_{2j}) by T of (y_1, y_2) such that $t_{1j}, t_{2j} \in [0, \min\{t_{1i}, t_{2i}\}]$. For $i < 0$, we call i -th crossing by T of (y_1, y_2) (if it exists), the last $|i|$ -th crossing (t_{1i}, t_{2i}) by T of (y_1, y_2) before zero. This means that (t_{1i}, t_{2i}) is the i -th crossing by T of (y_1, y_2) iff it is a crossing by T of (y_1, y_2) such that $t_{1i}, t_{2i} < 0$ and there are exactly $i - 1$ crossings (t_{1j}, t_{2j}) by T of (y_1, y_2) with $t_{1j}, t_{2j} \in [\max\{t_{1i}, t_{2i}\}, 0]$. If there exists a crossing $\{t_{10}, t_{20}\}$ by T of (y_1, y_2) containing 0 in its interval, i.e. such that $0 \in [\min\{t_{10}, t_{20}\}, \max\{t_{10}, t_{20}\}]$, then we call that crossing $\{t_{10}, t_{20}\}$, the 0-th crossing by T of (y_1, y_2) . (Note, that the different crossings (t_{1i}, t_{2i}) by T of (y_1, y_2) have their intervals $[\min\{t_{1i}, t_{2i}\}, \max\{t_{1i}, t_{2i}\}]$ disjoint from each other. Thus, it is always possible to introduce a strict order on the crossing by T of (y_1, y_2) , in the way defined above and thus the above definitions makes sense.) Now, whenever we have a collection of integer couples (y_{1l}, y_{2l}) with $l \in L$ such that the different intervals $[\min\{y_{1l}, y_{2l}\}, \max\{y_{1l}, y_{2l}\}]$ are mutually disjoint, then all the different crossings (t_{1li}, t_{2li}) by T of the different (y_{1l}, y_{2l}) have their intervals $[\min\{t_{1li}, t_{2li}\}, \max\{t_{1li}, t_{2li}\}]$ disjoint. Thus if $\{T(k)\}_{k \in D}$ is a simple random walk starting at the origin, by the strong Markov property, "what happens with T " during these different crossings is independent. In other words, the random vectors $(T(t_{1li}), T(t_{1li} + 1), T(t_{1li} + 2), \dots, T(t_{1li}))$ for all the different li 's are independent of each other. If none of the intervals $[\min\{y_{1l}, y_{2l}\}, \max\{y_{1l}, y_{2l}\}]$ contains 0, the random vectors $(T(t_{1li}), \dots, T(t_{1li}))$ are also identically distributed for those (y_{1l}, y_{2l}) having same length (i.e. such that $|y_{1l} - y_{2l}|$ is the same). If (y_{1l}, y_{2l}) is such that $0 \notin [\min\{y_{1l}, y_{2l}\}, \max\{y_{1l}, y_{2l}\}]$ and $|y_{1l} - y_{2l}| = 3$, then the probability that the i -th crossing by T of (y_{1l}, y_{2l}) be straight is equal to $\frac{3}{4}$. (Assuming that $\{T(k)\}_{k \in D}$ is a simple random walk starting at the origin.) To see this, note that on an interval (y_{1l}, y_{2l}) of length 3 for each odd number j bigger or equal to three, there is exactly one path of length j starting at y_{1l} at time 0 and arriving at time j at y_{2l} , in such a way that between time 0 and time j it never goes back to y_{1l} . The probability to follow a given path of length j with a simple random walk is $(\frac{1}{2})^j$. Thus the probability for the i -th crossing of (y_{1l}, y_{2l}) where $|y_{1l} - y_{2l}| = 3$ by a simple random walk to be straight, is equal to $(\frac{1}{2})^3 / [(\frac{1}{2})^3 + (\frac{1}{2})^5 + (\frac{1}{2})^7 + \dots] = \frac{3}{4}$. Now $\{R(k)\}_{k \in \mathbb{Z}}$ and $\{S(k)\}_{k \in \mathbb{N}}$ are two independent random walks starting both at the origin. As before, let y_1, y_2 be two integers such that $|y_1 - y_2| = 3n$. Let (t_1, t_2) be a crossing by $R \circ S$ of (y_1, y_2) having associated binary word $w = (w(1), \dots, w(i), \dots, w(n))$. Let (\bar{t}_1, \bar{t}_2) be another crossing by $R \circ S$ of (y_1, y_2) having associated binary word $\bar{w} = (\bar{w}(1), \dots, \bar{w}(i), \dots, \bar{w}(n))$. We can now apply what we explained for

a simple random walk T to S and get:

Case I: when $(S(t_1), S(t_2)) = (S(\bar{t}_1), S(\bar{t}_2))$.

By lemma 5, $w(i)$, $\bar{w}(i) = 1$ iff the first crossing (x_{1i}, x_{2i}) by R during $(S(t_1), S(t_2))$ of the i -th three unit interval of (y_1, y_2) is straight and the first crossings by S of (x_{1i}, x_{2i}) during (t_1, t_2) and during (\bar{t}_1, \bar{t}_2) are both straight. Since, the probability of a crossing by a simple random walk of a three unit interval to be straight is equal to $\frac{3}{4}$ and since, the processes $\{R(k)\}_{k \in \mathbb{Z}}$ and $\{S(k)\}_{k \in \mathbb{N}}$ are both simple random walks independent of each other, we get in the case $(S(t_1), S(t_2)) = (S(\bar{t}_1), S(\bar{t}_2))$ that $P(w(i) \times \bar{w}(i) = 1)$ is equal to $(\frac{3}{4})^3$. Furthermore, because of the strong Markov property of the simple random walk, we get that the different $w(i) \times \bar{w}(i)$'s given that $(S(t_1), S(t_2)) = (S(\bar{t}_1), S(\bar{t}_2))$, are independent of each other. Thus, in case I, our test statistic $\sum_{i=1}^n w(i) \times \bar{w}(i)$ has a binomial distribution with parameters n and $(\frac{3}{4})^3$.

Case II: when $(S(t_1), S(t_2)) \neq (S(\bar{t}_1), S(\bar{t}_2))$.

By our lemma, $w(i)$, $\bar{w}(i) = 1$ iff the first crossings (x_{1i}, x_{2i}) , resp. $(\bar{x}_{1i}, \bar{x}_{2i})$ by R during $(S(t_1), S(t_2))$ resp. during $(S(\bar{t}_1), S(\bar{t}_2))$ of the i -th three unit interval of (y_1, y_2) are both straight and the first crossings by S of (x_{1i}, x_{2i}) , resp. $(\bar{x}_{1i}, \bar{x}_{2i})$ during (t_1, t_2) resp. during (\bar{t}_1, \bar{t}_2) are both straight. Since, the probability for a crossing by a simple random walk of a three unit interval to be straight is equal to $\frac{3}{4}$ and since, the processes $\{R(k)\}_{k \in \mathbb{Z}}$ and $\{S(k)\}_{k \in \mathbb{N}}$ are both simple random walks independent of each other, we get in the case $(S(t_1), S(t_2)) \neq (S(\bar{t}_1), S(\bar{t}_2))$ that $P(w(i) \times \bar{w}(i) = 1)$ is equal to $(\frac{3}{4})^4$. Furthermore, because of the strong Markov property of the simple random walk, we get that the different $w(i) \times \bar{w}(i)$'s given that $(S(t_1), S(t_2)) \neq (S(\bar{t}_1), S(\bar{t}_2))$, are independent of each other. Thus in case II, our test statistic has a binomial distribution with parameters n and $(\frac{3}{4})^4$.

Note that the test defined above has its error of first type and second type both exponentially small in n . (Large deviation principle.) Thus there exists a positive constant c not depending on n , such that in both cases, the probability of an error by our test is smaller than e^{-cn} . We are next going to formulate this in a lemma.

Lemma 6 *Let y_1, y_2 be two integers lying on the same side of zero and such that $|y_1 - y_2| = 3n$. Let l and k be two integers different from zero and let i and j be two natural numbers. Let $w = (w(1), \dots, w(s), \dots, w(n))$ be the characteristic word associated with the crossing by $R \circ S$ of (y_1, y_2) which is equal to the i -th crossing by S of the l -th crossing by R of (y_1, y_2) . Let $\bar{w} = (\bar{w}(1), \dots, \bar{w}(s), \dots, \bar{w}(n))$ be the characteristic word associated with the crossing by $R \circ S$ of (y_1, y_2) which is the j -th crossing by S of the k -th crossing by R of (y_1, y_2) . Then, if $l = k$, we have that $\sum_{s=1}^n w(s) \times \bar{w}(s)$ has a binomial distribution with parameter n and $(\frac{3}{4})^3$. The probability of an error by our test in this case, i.e. $e^{-cn} \geq P(\sum_{s=1}^n w(s) \times \bar{w}(s) < n((\frac{3}{4})^3 - (\frac{3}{4})^4)/2)$, where c is a positive constant not depending on n . If $l \neq k$, we have that $\sum_{s=1}^n w(s) \times \bar{w}(s)$ has a binomial distribution with parameters n and $(\frac{3}{4})^4$. The probability of an error by our test in this case, i.e. $e^{-cn} \geq P(\sum_{s=1}^n w(s) \times \bar{w}(s) \geq n((\frac{3}{4})^3 + (\frac{3}{4})^4)/2)$, where c is a positive constant not depending on n .*

We will now need the following definitions. Let (x_1^{n+}, x_2^{n+}) designate the first crossing after zero by R of $(0, 3n)$ or of $(0, -3n)$ which ever of the two comes first after zero. Thus, if $(x_{1a}^{n+}, x_{2a}^{n+})$ designates the first crossing after zero by R of $(0, 3n)$ whilst $(x_{1b}^{n+}, x_{2b}^{n+})$ designate the first crossing after 0 by R of $(0, -3n)$, then, when $x_{1a}^{n+} < x_{1b}^{n+}$, we have (x_1^{n+}, x_2^{n+}) is equal to $(x_{1a}^{n+}, x_{2a}^{n+})$ whilst otherwise (x_1^{n+}, x_2^{n+}) is equal to $(x_{1b}^{n+}, x_{2b}^{n+})$. This means that x_2^{n+} is the first hitting time by $\{S(k)\}_{k \in \mathbb{Z}}$ on $\{-3n, 3n\}$. In a similar way, let (x_1^{n-}, x_2^{n-}) designate the last crossing before zero by R of $(0, 3n)$ or of $(0, -3n)$ which ever of the two comes last before zero. Because R is a nearest neighbor walk starting at the origin we get that $[-n, n] \subset [x_2^{n-}, x_2^{n+}]$. Furthermore, the probability for the first hitting time by a simple random walk starting at the origin at $\{-3n, 3n\}$ to be bigger than n^3 is exponentially small in n . We already saw that x_2^{n+} is the first hitting time of R on $\{-3n, 3n\}$ after zero. Furthermore, x_2^{n-} is the first hitting time of the reversed random walk $k \mapsto R(-k)$ on $\{-3n, 3n\}$ after zero. Thus with high probability $[x_2^{n-}, x_2^{n+}] \subset [-n^3, n^3]$. Now, recall that we want our partial reconstruction algorithm at level n to reconstruct a piece of the scenery ξ which is the restriction of ξ to an interval containing $[-n, n]$ and contained in $[-n^3, n^3]$. Thus, if we could reconstruct ξ between the point x_2^{n-} and the point x_2^{n+} , we would in most cases meet the requirement which we put on our partial reconstruction algorithm at level n . Now, our partial reconstruction algorithm at level n will try to find a straight crossing by S of (x_2^{n-}, x_2^{n+}) . During such a crossing the observations $\xi \circ S$ are a copy of the piece of scenery $\xi|_{[x_2^{n-}, x_2^{n+}]}$. Thus, if we could construct an algorithm capable of finding with high enough probability a straight crossing by S of (x_2^{n-}, x_2^{n+}) , our problem would be solved (i.e. the problem for the partial reconstruction algorithm at level n .) Let now (t_1^{n+}, t_2^{n+}) , resp. (t_1^{n-}, t_2^{n-}) designate the first crossing by S of (x_1^{n+}, x_2^{n+}) , resp. (x_1^{n-}, x_2^{n-}) . If $t_1^{n+} < t_1^{n-}$ then define (t_1^n, t_2^n) be equal to (t_1^{n+}, t_2^{n+}) and define (t_3^n, t_4^n) to be equal to (t_1^{n-}, t_2^{n-}) . If $t_1^{n+} > t_1^{n-}$ then define (t_1^n, t_2^n) be equal to (t_1^{n-}, t_2^{n-}) and define (t_3^n, t_4^n) to be equal to (t_1^{n+}, t_2^{n+}) . Now we saw that crossings of pairs with disjoint intervals have their interval disjoint too. This means that all the crossings (x_{1i}^n, x_{2i}^n) by R of $(0, 3n)$ or of $(0, -3n)$ which are different from (x_1^{n+}, x_2^{n+}) and from (x_1^{n-}, x_2^{n-}) must be outside $[x_2^{n-}, x_2^{n+}]$, i.e. $x_{1i}^n, x_{2i}^n \notin]x_2^{n-}, x_2^{n+}[$. Now S starts at the origin and thus $S(0) \in [x_1^{n-}, x_1^{n+}]$. So, the first time S crosses a crossing by R of $(0, 3n)$ or of $(0, -3n)$ it must necessarily be a crossing by S of (x_1^{n+}, x_2^{n+}) or of (x_1^{n-}, x_2^{n-}) . Let us call (x_1^n, x_2^n) the one crossing of the two (x_1^{n+}, x_2^{n+}) or of (x_1^{n-}, x_2^{n-}) which gets first crossed by S . As a matter of fact, when S is at $[x_1^{n-}, x_1^{n+}]$ for S to get to another crossing by R of $(0, 3n)$ or of $(0, -3n)$, S must first cross (x_1^{n+}, x_2^{n+}) or (x_1^{n-}, x_2^{n-}) in order to reach another crossing by R of $(0, 3n)$ or of $(0, -3n)$. By lemma 5, the first time S crosses a crossing by R of either $(0, 3n)$ or $(0, -3n)$, this is also the first crossing by $R \circ S$ of either $(0, 3n)$ or $(0, -3n)$. Thus we get the following characterization of the crossing (t_1^n, t_2^n) : (t_1^n, t_2^n) is the first crossing by $R \circ S$ of either $(0, n)$ or $(0, -n)$ which ever of the two $(0, 3n)$ or $(0, -3n)$ gets crossed first by $R \circ S$. (Note that (x_1^n, x_2^n) was defined in such a way that $(S(t_1), S(t_2)) = (x_1^n, x_2^n)$.) Thus we can construct (t_1^n, t_2^n) if we have $R \circ S$. On the other hand, we can use our test for two

crossings by $R \circ S$ to be on the same place, in order to try to find (t_3^n, t_4^n) . As a matter of fact (t_3^n, t_4^n) can be characterized as using the following idea: after a positive crossing by S of $(S(t_2^n), S(t_1^n)) = (x_2^n, x_1^n)$ we have that S is at the point x_1^n . So from there on the first time S will cross again a crossing by R of $(0, 3n)$ or $(0, -3n)$ (that is the first time after that time when S is at x_1^n that we observe a crossing by $R \circ S$ of $(0, 3n)$ or $(0, -3n)$) it will either be that S crosses (x_1^{n+}, x_2^{n+}) or (x_1^{n-}, x_2^{n-}) . This gives us the following characterization of (t_3^n, t_4^n) : the first time that after a positive crossing by S of (x_2^n, x_1^n) the next crossing by S of (x_1^{n+}, x_2^{n+}) or (x_1^{n-}, x_2^{n-}) is not a crossing by S of (x_1^n, x_2^n) , (i.e. it is a crossing by S of the one (x_1^{n+}, x_2^{n+}) or (x_1^{n-}, x_2^{n-}) which is different from (x_1^n, x_2^n)). Because of lemma 5 this characterization is equivalent to the following: (t_3^n, t_4^n) is the first crossing (s, t) $s < t$, by $R \circ S$ of $(0, 3n)$ or $(0, -3n)$ for which the following holds: the last crossing by $R \circ S$ of $(3n, 0)$ or $(-3n, 0)$ before that crossing, i.e. before time s , is a positive crossing by S of (x_2^n, x_1^n) and (s, t) is not a crossing by S of (x_2^n, x_1^n) . (Thus, it is a crossing by S of the one (x_1^{n+}, x_2^{n+}) or (x_1^{n-}, x_2^{n-}) which is different from (x_1^n, x_2^n) .) Now we can use the above characterization of (t_3^n, t_4^n) to determine (t_3^n, t_4^n) using our test for crossings. As a matter of fact, with our test for a crossing by $R \circ S$ to be a crossing by S on the same place, we can with high probability find out whether a crossing by $R \circ S$ of $(0, 3n)$ and of $(0, -3n)$ are also crossings by S of (x_1^n, x_2^n) or not. (This is so because we know that (t_1^n, t_2^n) is a crossing by S of (x_1^n, x_2^n) . Thus we can compare other crossings by $R \circ S$ of $(0, 3n)$ and of $(0, -3n)$ with the crossing (t_1^n, t_2^n) by $R \circ S$ using our test for crossings.) Thus we will proceed as follows in order to get an estimate $(\hat{t}_3^n, \hat{t}_4^n)$ for (t_3^n, t_4^n) . We take the first positive crossing by $R \circ S$ of $(0, 3n)$ or $(0, -3n)$ which our test identifies as not being a crossing by S of (x_1^n, x_2^n) but such that the last crossing by $R \circ S$ of $(0, 3n)$ or $(0, -3n)$ before \hat{t}_3^n is identified by our test for crossings to be a positive crossing by S of (x_2^n, x_1^n) . We explained how our partial reconstruction algorithm finds (t_1^n, t_2^n) and how it can with high probability find (t_3^n, t_4^n) .

Let us go on with our informal discussion of the partial reconstruction algorithm at level n . The next step for our partial reconstruction algorithm at level n will be to try to find a straight crossing by S of $(S(t_2^n), S(t_4^n))$. (Note that $\{S(t_2^n), S(t_4^n)\} = \{x_2^{n+}, x_2^{n-}\}$. We already saw that in most cases, if we are able to determine a straight crossing by S of (x_2^{n+}, x_2^{n-}) then we have solved the problem which we ask the partial reconstruction algorithm at level n to solve.) The idea which comes first to mind now for finding such a straight crossing by S , would be to use our test in order to find a crossing by S of (x_2^n, x_1^n) followed in minimal time by a crossing by S of (x_3^n, x_4^n) . Here $(x_3^n, x_4^n) = (S(t_3^n), S(t_4^n))$. In this simple setting this doesn't quite work. As a matter of fact, the distance between the points $S(t_2^n)$ and $S(t_4^n)$ is typically of order n^2 . The probability for a simple random walk to cross an interval of length n^2 in a straight way is negatively exponentially small in n^2 . Thus we need about $\exp(n^2)$ crossings by S of $(S(t_2^n), S(t_4^n)) = (x_2^n, x_4^n)$ before we get a straight one. However our test to test whether a crossing by $R \circ S$ is also a crossing by S of $(S(t_1^n), S(t_2^n))$ has a negative exponential probability of making a mistake. Thus, if we apply that test e^{n^2} times, many mistake will happen. So before we will be able to

identify a straight crossing by S of (x_2^n, x_4^n) our test will have made a mistake. This problem can be fixed in the following way. On top of the crossings by $R \circ S$ of $(0, 3n)$ and $(0, -3n)$ we will also look at the crossings by $R \circ S$ of $(0, 3n^5)$ and $(0, -3n^5)$. These crossings by $R \circ S$ of $(0, 3n^5)$ and $(0, -3n^5)$ will only be used to be able to find a lot of returns of S close to the origin. Let $(t_1^{n^5}, t_2^{n^5})$ be the first crossing by $R \circ S$ of $(0, 3n^5)$ or of $(0, -3n^5)$ which ever of the two gets crossed first. Thanks to our test we can with high probability find approximately e^{n^5} crossings by S of $(S(t_1^{n^5}), S(t_2^{n^5}))$ before making a mistake. This is more than needed to get with high probability an interval of length n^2 to get crossed by S in a straight way shortly after one of these e^{n^5} crossings. We will thus restrict our attention to the crossings by $R \circ S$ of either $(0, 3n)$ or $(0, -3n)$ which appear within time n^{25} from a crossing which we identified by our test as being a crossing by S of $(S(t_1^{n^5}), S(t_2^{n^5}))$. Let w and \bar{w} be two binary words having same length. Then, we say that w is bigger than \bar{w} , iff for each bit for which $\bar{w}(i) = 1$, we also have that $w(i) = 1$. Let w^n designate the characteristic word of the crossing (t_2^n, t_1^n) by $R \circ S$. Let w_{34}^n designate the associated word of the crossing (t_3^n, t_4^n) by $R \circ S$. When, can a crossing by $R \circ S$ of $(R(S(t_2^n)), R(S(t_1^n)))$ have its associated word strictly bigger than w^n . By lemma 5, this can only happen if it is a crossing by S of a crossing (x_a, x_b) by R of $(R(S(t_2^n)), R(S(t_1^n)))$ and the crossing (x_a, x_b) by R must be such that for each $i \in 1, 2, \dots, n$ for which $w^n(i) = 1$ we also have that the first crossing by R during (x_a, x_b) of the i -th three unite interval of $(R(S(t_2^n)), R(S(t_1^n)))$ is straight. We will prove that with very high probability the only crossings by R of $(R(S(t_2^n)), R(S(t_1^n)))$ which come within n^{26} from the crossing $(S(t_1^{n^5}), S(t_2^{n^5}))$ and which satisfies that property is $(S(t_1^{n^5}), S(t_2^{n^5}))$ itself. Thus, in this case, if within time n^{26} from a crossing by S of $(S(t_1^{n^5}), S(t_2^{n^5}))$ we observe a crossing by $R \circ S$ of $(R(S(t_2^n)), R(S(t_1^n)))$ having associated word bigger than w^n , then this must be a crossing by S of $(S(t_2^n), S(t_1^n))$. The same holds for $(R(S(t_3^n)), R(S(t_4^n)))$ and w_{34}^n . We can now explain how our partial reconstruction algorithm at level n work: We take the set of all the couples of natural numbers (t_a, t_b) where $t_a < t_b$ and such that t_a, t_b lie within time n^{26} from a crossing which we identified as a crossing by S of $(S(t_1^{n^5}), S(t_2^{n^5}))$ and such that $e^{n^4} \geq t_a, t_b$ and such that there exists $t > t_a$ and $s < t_b$ such that (t_a, t) is a crossing by $R \circ S$ of $(R(S(t_2^n)), R(S(t_1^n)))$ having associated word bigger than w^n and such that (s, t_b) is a crossing by $R \circ S$ of $(R(S(t_3^n)), R(S(t_4^n)))$ having associated word bigger than w_{34}^n . Among all the couples (t_a, t_b) satisfying all of the above conditions, we pick one (t_5^n, t_6^n) which minimizes $t_b - t_a$. In other words we pick two crossings at minimum distance from each other (in the sense that the distance between the ends which are furthest away from each other is minimized.) For the two crossings we require that they both be within time n^{26} from a crossing which we identified as a crossing by S of $(S(t_1^{n^5}), S(t_2^{n^5}))$ and that both crossings come before time e^{n^4} . Furthermore we require that the first crossing of the two be a crossing by $R \circ S$ of $(R(x_2^n), R(x_1^n))$ (this means that it is a crossing of either $(3n, 0)$ or $(-3n, 0)$) and that it has associated word bigger than w^n . In a similar way,

we require that the second of the two crossings picked be a crossing by $R \circ S$ of $(R(x_3^n), R(x_4^n))$ which has its associated word bigger than w_{34}^n . The piece of scenery which is going to be the output of the partial reconstruction algorithm at level n will be the restriction of the observations $\xi \circ S$ to the interval comprised between the crossings which we pick. (To be more precise let us say that we take the interval between the ends t_5^n and t_6^n which lie furthest away from each other of the crossings which we picked.) If everything goes correctly (which according to our above explanations should be the case with high probability), the couple (t_5^n, t_6^n) should be a straight crossing by S of (x_2^n, x_4^n) . Our partial reconstruction algorithm at level n has as outcome the piece of scenery obtained by restricting the observations $\xi \circ S$ to the space between t_5^n and t_6^n . (This means that the outcome of the partial reconstruction algorithm at level n is equal to the following piece of scenery: $k \mapsto \xi(S(k)) [t_5^n, t_6^n] \rightarrow \{0, 1\}$.) In the case that the couple (t_5^n, t_6^n) is a straight crossing by S of (x_2^n, x_4^n) , this means that the outcome of the partial reconstruction algorithm at level n is equivalent to $\xi[\min\{x_2^n, x_4^n\}, \max\{x_2^n, x_4^n\}]$ and thus is equivalent to $\xi[x_2^{n-}, x_2^{n+}]$. Let us, at this stage define the partial reconstruction algorithm at level n in a precise way:

Algorithm 7 *The input of the partial reconstruction algorithm at level n are the observations $\xi \circ S$. Step a) Let $\varphi : \mathbb{Z} \rightarrow \{0, 1\}$ be the four periodic scenery such that $(\varphi(0), \varphi(1), \varphi(2), \varphi(3)) = (0, 0, 1, 1)$ when $\xi(S(0)) = 0$ and such that $(\varphi(0), \varphi(1), \varphi(2), \varphi(3)) = (1, 1, 0, 0)$ otherwise. Step b) Construct $R \circ S$. Note for this purpose, that $R \circ S$ is the only nearest neighbor walk starting at the origin and generating the observation $\xi \circ S$ on φ . Thus $R \circ S$ is the only nearest neighbor walk starting at the origin and such that for all integer k , we have $R \circ S(k) = \xi(S(k))$. Step c) Find the first positive crossing by $R \circ S$ of $(0, 3n)$ or $(0, -3n)$ which ever gets crossed first. Call it (t_1^n, t_2^n) . Let w^n designate the word associated with the crossing (t_1^n, t_2^n) by $R \circ S$. Step d) Find the first positive crossing by $R \circ S$ of $(0, 3n)$ or $(0, -3n)$ for which the following holds: our test indicates it not be a crossing by S of $(S(t_1^n), S(t_2^n))$ (by this we mean that when we compare that crossing by $R \circ S$ with the crossing (t_1, t_2) our test indicates that S is on different places for the two crossings) but such that before that, the last positive crossing by $R \circ S$ of $(3n, 0)$ or of $(-3n, 0)$, is such that our test indicates it to be a crossing by S of $(S(t_2^n), S(t_1^n))$. (By this we mean that when we compare that crossing by $R \circ S$ with the crossing (t_1, t_2) our test indicates that S is on the same place.) Call that crossing $(\hat{t}_3^n, \hat{t}_4^n)$. Let w_{34}^n designate the word associated with the crossing $(\hat{t}_3^n, \hat{t}_4^n)$ by $R \circ S$. Step e) Find the first positive crossing by $R \circ S$ of $(0, 3n^5)$ or $(0, -3n^5)$ which ever gets crossed first. Call it $(t_1^{n^5}, t_2^{n^5})$. Step f) Up to time e^{n^4} find all the crossings by $R \circ S$ of $(0, 3n^5)$ or $(0, -3n^5)$ which our test for crossings identifies as being crossings by S of $(S(t_1^{n^5}), S(t_2^{n^5}))$. Let $\Gamma^n \subset [0, e^{n^4}]$ be the integer random set of the points up to e^{n^4} which lie within time n^{25} of a crossing which our test for crossings identifies as being a crossing by S of $(S(t_1^{n^5}), S(t_2^{n^5}))$. Step g) Let $\{(t_{2i}^n, t_{1i}^n) | i \in I\} \subset \Gamma^n \times \Gamma^n$ designate the set of all positive crossings by $R \circ S$ of $(R(S(t_2^n)), R(S(t_1^n)))$ which lie within $\Gamma^n \times \Gamma^n$ and which have there associated word bigger than w^n . Step h) Let $\{(t_{3j}^n, t_{4j}^n) | j \in J\} \subset \Gamma^n \times \Gamma^n$ designate the set*

of all positive crossings by $R \circ S$ of $(R(S(t_3^n)), R(S(t_4^n)))$ which lie within $\Gamma^n \times \Gamma^n$ and which have there associated word bigger then $w^{n^{34}}$. *Step i)* find a crossing (t_{2i}^n, t_{1i}^n) in $\{(t_{2i}^n, t_{1i}^n) | i \in I\}$ and a crossing (t_{3j}^n, t_{4j}^n) in $\{(t_{3j}^n, t_{4j}^n) | j \in J\}$ such that $t_{4j}^n - t_{2i}^n$ is minimal under the constraints $t_{4j}^n - t_{2i}^n > 0$. (Equivalently, find a couple $(i, j) \in I \times J$ minimizing $t_{4j}^n - t_{2i}^n$ under the constraint $t_{4j}^n - t_{2i}^n > 0$.) Put t_5^n to be equal to t_{2i}^n and t_6^n to be equal to t_{4j}^n . *Step j)* The outcome of our partial reconstruction algorithm at level n is going to be the observations $\xi \circ S$ restricted to the integer interval $[t_5^n, t_6^n]$. More precisely the output of the partial reconstruction algorithm at level n is equal to the piece of scenery $k \mapsto \xi(S(k))$ $[t_5^n, t_6^n] \rightarrow \{0, 1\}$.

5 Proof that the partial reconstruction algorithm works

In this section we are going to formally prove that the partial reconstruction algorithm at level n works with high probability. Recall that we say that the partial reconstruction algorithm at level n works when there exists an integer (random) interval I^n , such that $[-n, n] \subset I^n \subset [-n^3, n^3]$ and such that the output of the partial reconstruction algorithm at level n is a piece of scenery equivalent to the restriction of ξ to I^n . (We denote that restriction by $\xi|_{I^n}$.) Recall also that E^n designates the event that the reconstruction algorithm at level n works. Let E^{nc} be the complement of E^n . This section is devoted to proving that $\sum_{n=1}^{\infty} P(E^{nc}) < \infty$. This together with lemma 4, proves then that the main result of this paper holds, i.e. that one can a.s. reconstruct a scenery up to equivalence if one is only given the observations $\xi \circ S$. To prove that $\sum_{n=1}^{\infty} P(E^{nc}) < \infty$ holds, we are going to proceed as follows: We will define 7 events related to E^n : $E_1^n, E_2^n, \dots, E_7^n$. We will show that when $E_1^n, E_2^n, \dots, E_7^n$ all hold, then E^n also holds. In other words, $E_1^n \cap E_2^n \cap \dots \cap E_7^n \subset E^n$. Then we will show that for each E_i^n with $i \in 1, 2, \dots, 6, 7$ the probability of the complement of E_i^n is finitely summable over n . (E_i^{nc} will designate the complement of E_i^n .) This proves then that $\sum_{n=1}^{\infty} P(E^{nc}) < \infty$, since $P(E_1^{nc}) + P(E_2^{nc}) + \dots + P(E_7^{nc}) \geq P(E^{nc})$. Before starting to prove that $E_1^n \cap E_2^n \cap \dots \cap E_7^n \subset E^n$. Let us define the events $E_1^n, E_2^n, \dots, E_7^n$.

Let E_1^n be the event $x_2^{n+}, x_2^{n-} \subset [-n^3, n^3]$. Recall, that (x_1^{n+}, x_2^{n+}) was defined to be the first crossing after zero by R of $(0, 3n)$ or of $(0, -3n)$ which ever of the two comes first after zero. Furthermore, (x_1^{n-}, x_2^{n-}) was defined to be the last crossing before zero by R of $(0, 3n)$ or of $(0, -3n)$ which ever of the two comes last before zero. Thus, $E_1^n = \{R([0, n^3]) \cap \{-3n, 3n\} \neq \emptyset \text{ and } R([-n^3, 0]) \cap \{-3n, 3n\} \neq \emptyset\}$.

Let E_2^n be the event that S visits both points $-n^3$ and n^3 before time n^{15} , that is $E_2^n = \{-n^3, n^3 \in S([0, n^{15}])\}$.

Let E_3^n be the event that our test for crossings by $R \circ S$ to be crossings by S of the same place, does not fail a single time for all crossings by $R \circ S$ of $(0, 3n)$ or $(0, -3n)$ which have the numbers defining them smaller than n^{15} .

Thus $E_3^n = \{ \text{For any } l, k \in \mathbb{Z} \text{ and } i \text{ and } j \in \mathbb{N} \text{ such that } n^{15} \geq |l|, |k|, i, j \text{ our statistical test for crossings to be on the same place, will determine correctly if } l = k \text{ or not, when comparing the two crossings by } R \circ S \text{ which correspond to the } i\text{-th crossing by } S \text{ of the } l\text{-th crossing by } R \text{ of } (0, 3n) \text{ or } (0, -3n) \text{ and to the } j\text{-th crossing by } S \text{ of the } k\text{-th crossing by } R \text{ of } (0, 3n) \text{ or } (0, -3n). \}$

Recall, that $(t_1^{n^5}, t_2^{n^5})$ designates the first crossing by $R \circ S$ of $(0, 3n^5)$ or $(0, -3n^5)$ which ever of the two gets crossed first by $R \circ S$. Let E_4^n be the event that up to time e^{n^4} our test for crossings by $R \circ S$ to be crossings by S of the same place never makes a mistake when comparing the crossing $(t_1^{n^5}, t_2^{n^5})$ to another crossing by $R \circ S$ of $(0, 3n^5)$ or $(0, -3n^5)$. That is, $E_4^n = \{ \text{for each crossing } (t, s) \text{ by } R \circ S \text{ of } (R(S(t_1^{n^5})), R(S(t_2^{n^5}))) \text{ where } e^{n^4} \geq t, s, \text{ our test recognizes correctly whether } (S(t), S(s)) = (S(t_1^{n^5}), S(t_2^{n^5})) \text{ or not.} \}$

Let E_5^n be the event that the first crossing after 0 by R of $(0, 3n^5)$ or of $(0, -3n^5)$ happens during $[0, n^{11}]$ and the last crossing before 0 by R of $(0, 3n^5)$ or of $(0, -3n^5)$ happens during $[-n^{11}, 0]$. More precisely $E_5^n = \{R([0, n^{11}]) \cap \{-3n^5, 3n^5\} \neq \emptyset \text{ and } R([-n^{11}, 0]) \cap \{-3n^5, 3n^5\} \neq \emptyset\}$.

Let E_6^n be the event that in the interval $[-n^{26}, n^{26}]$ the only crossing by R of $(3n, 0)$ or $(-3n, 0)$ having associated binary word bigger than w^n is (x_2^n, x_1^n) and the only crossing by R of $(0, 3n)$ or $(0, -3n)$ having associated binary word bigger than w_{34}^n is (x_3^n, x_4^n) . In other words, $E_6^n = \{ \text{for any crossing } (x_a, x_b) \text{ by } R \text{ of } (3n, 0) \text{ or } (-3n, 0) \text{ for which } x_a, x_b \in [-n^{26}, n^{26}] \text{ and } (x_a, x_b) \neq (x_2^n, x_1^n) \text{ there exists } i \text{ such that } w^n(i) = 1 \text{ and } n \geq i \text{ but the first crossing by } R \text{ of the } i\text{-th three unit interval of } (R(x_a), R(x_b)) \text{ during } (x_a, x_b) \text{ is not straight} \} \cap \{ \text{for any crossing } (x_a, x_b) \text{ by } R \text{ of } (0, 3n) \text{ or } (0, -3n) \text{ for which } x_a, x_b \in [-n^{26}, n^{26}] \text{ and } (x_a, x_b) \neq (x_3^n, x_4^n) \text{ there exists } i \leq n \text{ such that } w_{34}^n(i) = 1 \text{ but the first crossing by } R \text{ of the } i\text{-th three unit interval of } (R(x_a), R(x_b)) \text{ during } (x_a, x_b) \text{ is not straight.} \}$

Let E_7^n designate the event that before time e^{n^4} , S crosses $(-n^3, n^3)$ in both directions in a straight way within time n^{25} from a crossing by S of $(x_1^{n^5}, x_2^{n^5})$. That is $E_7^n = \{ \text{there exists } (s_a, t_a), (s_b, t_b), (s_c, t_c) \text{ such that } s_a < t_a < s_b < t_b < s_c < t_c \leq e^{n^4}, |t_a - t_c| \leq n^{25} \text{ and such that } (s_a, t_a) \text{ is a crossing by } S \text{ of } (S(t_1^{n^5}), S(t_2^{n^5})) \text{ whilst } (s_b, t_b) \text{ is a straight crossing by } S \text{ of } (-n^3, n^3) \text{ and } (s_c, t_c) \text{ is a straight crossing by } S \text{ of } (n^3, -n^3). \}$

We are now ready to start our proofs:

Proof that $E_1^n \cap E_2^n \cap \dots \cap E_6^n \cap E_7^n \subset E^n$. Let us recall once more, that $(x_1^{n^+}, x_2^{n^+})$ was defined to be the first crossing after zero by R of $(0, 3n)$ or of $(0, -3n)$ which ever of the two gets crossed first after zero. In a similar way, $(x_1^{n^-}, x_2^{n^-})$ was defined to be the last crossing before zero by R of $(0, 3n)$ or of $(0, -3n)$ which ever of the two comes last before zero. (x_1^n, x_2^n) was defined to be the one of the two crossings $(x_1^{n^-}, x_2^{n^-})$ and $(x_1^{n^+}, x_2^{n^+})$ which gets first crossed by S whilst (x_3^n, x_4^n) was defined to be the one which gets crossed second. Thus, $\{(x_1^{n^-}, x_2^{n^-}), (x_1^{n^+}, x_2^{n^+})\} = \{(x_1^n, x_2^n), (x_3^n, x_4^n)\}$. We have already mentioned that all the different crossings (t_{ij1}, t_{ij2}) by a nearest neighbor walk T of different pairs (y_{i1}, y_{i2}) such that the intervals $[\min\{y_{i1}, y_{i2}\}, \max\{y_{i1}, y_{i2}\}]$ are mutually disjoint for different i 's, have their intervals $[\min\{t_{ij1}, t_{ij2}\}, \max\{t_{ij1}, t_{ij2}\}]$ also

mutually disjoint. Thus, all the crossings (x_a, x_b) by R of $(0, 3n)$ and of $(0, -3n)$ which are different from (x_1^n, x_2^n) and from (x_3^n, x_4^n) must lie outside $[x_2^{n-}, x_2^{n+}]$, that is $x_a, x_b \notin [x_2^{n-}, x_2^{n+}]$. Thus, when at a time t , S is between x_1^{n-} and x_1^{n+} , i.e. $S(t) \in [x_1^{n-}, x_1^{n+}]$, then the first time after time t that S will cross a crossing by R of $(0, 3n)$ or of $(0, -3n)$ (and that we will thus observe a crossing by $R \circ S$ of $(0, 3n)$ or of $(0, -3n)$), S will cross either (x_1^{n-}, x_2^{n-}) or (x_1^{n+}, x_2^{n+}) . So, by lemma 5, this means that if at a time t , S is between x_1^{n-} and x_1^{n+} , then the first time after time t , that we observe a crossing by $R \circ S$ of $(0, 3n)$ or of $(0, -3n)$, this must be a crossing by S of either (x_1^{n-}, x_2^{n-}) or (x_1^{n+}, x_2^{n+}) . Recall that (t_3^n, t_4^n) designates the first crossing by S of (x_3^n, x_4^n) . Now, (t_3^n, t_4^n) can be characterized as follows: let (s, t) be the first positive crossing by S of (x_2^n, x_1^n) such that the first crossing by $R \circ S$ of $(0, 3n)$ or of $(0, -3n)$ after t is not a crossing by S of (x_1^n, x_2^n) . Then, (t_3^n, t_4^n) is the first crossing by $R \circ S$ of $(0, 3n)$ or of $(0, -3n)$ after t . When our algorithm searches for (t_3^n, t_4^n) and constructs the estimate $(\hat{t}_3^n, \hat{t}_4^n)$ for (t_3^n, t_4^n) , it uses the above characterization of (t_3^n, t_4^n) in conjunction with our test for crossings by $R \circ S$ to be crossings by S of the same place. Thus, if up to time t_4^n our test for crossings by $R \circ S$ of $(0, 3n)$ and of $(0, -3n)$ to be crossings by S on the same place, never makes a mistake, then our algorithm is able to identify (t_3^n, t_4^n) correctly. In this case $(t_3^n, t_4^n) = (\hat{t}_3^n, \hat{t}_4^n)$. Now, when E_1^n and E_2^n both hold, then $n^{15} \geq t_4^n$. But when E_3^n holds, our test for crossings by $R \circ S$ of $(0, 3n)$ and of $(0, -3n)$ to be crossings by S on the same place makes no mistake up to time n^{15} . (As a matter of fact until time n^{15} , S will never get further then the crossings number $\pm n^{15}$ by R of $(0, 3n)$ and of $(0, -3n)$. Furthermore, during time n^{15} , S will at most cross any crossing n^{15} times.) Thus, when E_1^n, E_2^n and E_3^n all hold, we get that $(t_3^n, t_4^n) = (\hat{t}_3^n, \hat{t}_4^n)$.

When E_5^n holds, then $n^{11} \geq |S(t_1^{n^5})|, |S(t_2^{n^5})|$. Now, $n^{26} \geq n^{11} + n^{25}$, (at least for n big enough which we will assume here). Thus, whenever S at a time t is within time n^{25} from a crossing by S of $(S(t_2^{n^5}), S(t_1^{n^5}))$, then when E_5^n holds, $|S(t)|$ is smaller equal than n^{26} . Now, when E_4^n holds, we have that up to time e^{n^4} , all the crossings by S of $(S(t_2^{n^5}), S(t_1^{n^5}))$ are correctly recognized by our test. Recall that Γ^n was defined to be the random set of the integer points up to time e^{n^4} which come within time n^{25} of a crossing which our test identifies to be a crossing by S of $(S(t_2^{n^5}), S(t_1^{n^5}))$. Thus when both E_4^n and E_5^n hold, then $S(\Gamma^n) \subset [-n^{26}, n^{26}]$. However, when E_6^n holds, it is not possible, during a time when S is in the interval $[-n^{26}, n^{26}]$, to observe a crossing by $R \circ S$ of $(3n, 0)$ or of $(-3n, 0)$, with associated word bigger than w^n and which is not a crossing by S of (x_2^n, x_1^n) . Thus, when E_5^n, E_4^n, E_6^n all hold, then during time Γ^n , all the crossings by $R \circ S$ of $(3n, 0)$ or of $(-3n, 0)$, with associated word bigger than w^n must be crossings by S of (x_2^n, x_1^n) . In a similar way, when E_5^n, E_4^n, E_6^n all hold, then during time Γ^n , all the crossings by $R \circ S$ of $(0, 3n)$ or of $(0, -3n)$, with associated word bigger than w_{34}^n must be crossings by S of (x_3^n, x_4^n) . Thus, whenever E_5^n, E_4^n, E_6^n all hold, the set of crossings $\{(t_{2i}^n, t_{1i}^n) | i \in I\} \subset \Gamma^n \times \Gamma^n$, resp. $\{(t_{3j}^n, t_{4j}^n) | j \in J\} \subset \Gamma^n \times \Gamma^n$ constructed in step g, resp. h of our partial reconstruction algorithm at level n are all crossings by S of (x_2^n, x_1^n) , resp. of

(x_3^n, x_4^n) . Thus in this case, $S(t_{2i}^n) = x_2^n$ for all $i \in I$ and $S(t_{4j}^n) = x_4^n$ for all $j \in J$.

This implies that when E_5^n, E_4^n, E_6^n all hold, then $|t_{4j}^n - t_{2i}^n| \geq |x_4^n - x_2^n|$. When, E_5^n, E_4^n, E_6^n all hold, we would thus have that $|t_{4j}^n - t_{2i}^n| = |x_4^n - x_2^n|$ iff (t_{2j}^n, t_{4i}^n) is a straight crossing by S . Thus, when, E_5^n, E_4^n, E_6^n all hold, if there exists two crossings (t_{2a}^n, t_{1a}^n) and (t_{3b}^n, t_{4b}^n) with $a \in I$ and $b \in J$ and such that (t_{2a}^n, t_{4b}^n) is a straight crossing by S , then the pair (t_5^n, t_6^n) obtained in step i of the partial reconstruction algorithm at level n , would have to be a straight crossing by S of (x_2^n, x_4^n) . But, when E_1^n, E_4^n and E_7^n all hold, then there exists two crossings (t_{2a}^n, t_{1a}^n) and (t_{3b}^n, t_{4b}^n) with $a \in I$ and $b \in J$, such that (t_{2a}^n, t_{4b}^n) is a straight crossing by S . As a matter of fact, when E_1^n holds, then $x_1^n, x_2^n, x_3^n, x_4^n \in [-n^3, n^3]$. When, E_7^n holds, then before time e^{n^4} and within time n^{25} after a crossing by S of $(S(t_1^{n^5}), S(t_2^{n^5}))$, S crosses $(-n^3, n^3)$ in a straight way. Thus, when E_1^n and E_7^n both hold, then, before time e^{n^4} and within time n^{25} after a crossing by S of $(S(t_1^{n^5}), S(t_2^{n^5}))$, S crosses (x_2^n, x_4^n) in a straight way. When E_7^n holds, then, up to time e^{n^4} , our test for crossings identifies correctly all the crossings by S of $(S(t_1^{n^5}), S(t_2^{n^5}))$. Thus, when E_1^n, E_4^n and E_7^n all hold, there exists a straight crossing (s, t) by S of (x_2^n, x_4^n) lying in $\Gamma^n \times \Gamma^n$. (Recall that Γ^n is the random set of integer points constructed in step f of our partial reconstruction algorithm at level n .) When, S crosses (x_2^n, x_4^n) in a straight way, then it first crosses (x_2^n, x_1^n) in a straight way and ends up crossing (x_3^n, x_4^n) in a straight way. Thus, if (s, t) is a straight crossing by S of (x_2^n, x_4^n) , then there exists $s_a, s_b \in [\min\{s, t\}, \max\{s, t\}]$ such that (s, s_a) is a straight crossing by S of (x_2^n, x_1^n) and (s_b, t) is a straight crossing by S of (x_3^n, x_4^n) . It is easy to check that a straight crossing by S of (x_2^n, x_1^n) has, when seen as a crossing by $R \circ S$, its associated word bigger than w^n . In a similar manner, a straight crossing by S of (x_3^n, x_4^n) , has its associated word (here we mean the binary word associated with that crossing seen as a crossing by $R \circ S$) bigger than the binary word w_{34}^n . We already saw that when E_1^n, E_2^n and E_3^n all hold, we get that $(t_3^n, t_4^n) = (t_3^n, t_4^n)$. This means, that when E_1^n, E_2^n and E_3^n all hold then the binary word w_{34}^n , constructed in step d of our partial reconstruction algorithm at level n , is equal to the binary word of the crossing (t_3^n, t_4^n) by $R \circ S$. Thus, when $E_1^n, E_2^n, E_3^n, E_4^n$, and E_7^n all hold then there exist a straight crossing (s, t) by S of (x_2^n, x_4^n) lying in $\Gamma^n \times \Gamma^n$. and such that there exists $a \in I$ and $b \in J$ such that $t_{2a}^n = s$ and $t_{4b}^n = t$ (i.e. the straight crossing (s, t) "begins" with a crossing by $R \circ S$ of (x_2^n, x_1^n) which has its associated word bigger than w^n and "ends" with a crossing by $R \circ S$ of (x_3^n, x_4^n) which has its associated word bigger than w_{34}^n .) This then finishes to prove that when $E_1^n, E_2^n, E_3^n, E_4^n, E_5^n, E_6^n$ and E_7^n all hold, the pair (t_5^n, t_6^n) obtained in step i of the partial reconstruction algorithm at level n , is a straight crossing by S of (x_2^n, x_4^n) . Thus, when $E_1^n, E_2^n, E_3^n, E_4^n, E_5^n, E_6^n$ and E_7^n all hold, the piece of scenery which is the outcome of our partial reconstruction algorithm at level n , is equivalent to the restriction of ξ to $[\min\{x_2^n, x_4^n\}, \max\{x_2^n, x_4^n\}]$. But, when E_1^n holds $[\min\{x_2^n, x_4^n\}, \max\{x_2^n, x_4^n\}] \subset [-n^3, n^3]$. On the other hand by definition, it's always true that $[-n, n] \subset [\min\{x_2^n, x_4^n\}, \max\{x_2^n, x_4^n\}]$. Thus,

when E_1^n holds, the restriction of ξ to $[\min\{x_2^n, x_4^n\}, \max\{x_2^n, x_4^n\}]$ satisfies the demands we put on the outcome of the partial reconstruction algorithm at level n . We just finished to prove that $E_1^n, E_2^n, E_3^n, E_4^n, E_5^n, E_6^n$ and E_7^n jointly imply E_1^n . Thus, $E_1^n \cap E_2^n \cap E_3^n \cap E_4^n \cap E_5^n \cap E_6^n \cap E_7^n \subset E^n$.

Proof that the probability of E_1^{nc} is finitely summable over n . x_2^{n+} is equal to the first passage after zero of R at $\{-3n, 3n\}$. We know that for a simple random walk starting at the origin, the probability that from 0 up to time k , R does not hit on $\{-3n, 3n\}$ is bounded above by $e^{-k_1 k / (n^2)}$, where k_1 is a strictly positive constant not depending on k or n . Thus the probability that $x_2^{n+} \geq n^3$ is smaller, than $e^{-k_1 n}$. In a similar way, one can show that $P(-n^3 \geq x_2^{n-}) \leq e^{-k_1 n}$. Thus, $P(E_1^{nc})$ is smaller than $2e^{-k_1 n}$. Thus, $\sum_{n=1}^{\infty} P(E_1^{nc}) < \infty$.

Proof that the probability of E_2^{nc} is finitely summable over n . Let us introduce a sequence of increasing stopping times in the following way. Let us stop the random walk S when it first visits -1 , then when it first visits -2 , and let us go on like this, until the random walk S first reaches $-n^3$. From there on let us wait until the random walks comes back to $-n^3 + 1$, then to $-n^3 + 2$, and so on until the random walk S reaches $+n^3$. In other words, we define the stopping times $\tau(l)$ for each $l \in 1, 2, \dots, 3n^3$, in the following way: for $l \in 1, 2, \dots, n^3$, let $\tau(l)$ be the first passage of S at $-l$. For $l \in n^3 + 1, n^3 + 2, \dots, 3n^3$, let $\tau(l)$ be the first passage of S at $-n^3 + l$ after time $\tau(n^3)$. For $l = 0$, let $\tau(0) = 0$. Then by symmetry and the strong Markov property of the random walk we get that the times between our stopping times are i.i.d., i.e. the collection $\tau(l) - \tau(l-1)$, where $l \in 1, 2, \dots, 3n^3$ is i.i.d. Obviously, when $n^{15} \geq \tau(3n^3)$, then E_2^n holds. Thus, $P(\tau(3n^3) \geq n^{15}) \geq P(E_2^{nc})$. Let $X(l)$ be the random variable equal to $\tau(l) - \tau(l-1)$. Then, $P(\tau(3n^3) \geq n^{15}) = P(X(1) + X(2) + \dots + X(3n^3) \geq n^{15})$. For any set of positive numbers a, b, c, d, e, \dots we have that $(a + b + c + d + \dots)^3 \geq a^3 + b^3 + c^3 + \dots$. Thus, $X(1)^{1/3} + X(2)^{1/3} + \dots + X(3n^3)^{1/3} \geq (X(1) + X(2) + \dots + X(3n^3))^{1/3}$. This implies, that $P(X(1)^{1/3} + X(2)^{1/3} + \dots + X(3n^3)^{1/3} \geq n^5) \geq P(X(1) + X(2) + \dots + X(3n^3) \geq n^{15})$. By Chebychev, we get $3n^3 E[X(1)^{1/3}] / (n^5) \geq P(X(1)^{1/3} + X(2)^{1/3} + \dots + X(3n^3)^{1/3} \geq n^5)$. Thus, we get that $3E[X(1)^{1/3}] / n^2 \geq P(E_2^{nc})$. It is a well known fact that the expectation of $X(1)^{1/3}$ is finite and thus we get that the expression on the left side of the last inequality is finitely summable over n . Thus $\sum_{n=1}^{\infty} P(E_2^{nc}) < \infty$.

Proof that the probability of E_3^{nc} is finitely summable over n . By lemma 6, the probability of our test to make a mistake, when comparing the i -th crossing by S of the l -th crossing by R of $(0, 3n)$ or $(0, -3n)$ with the j -th crossing by S of the k -th crossing by R of $(0, 3n)$ or $(0, -3n)$ is smaller than e^{-cn} , where c is a positive constant not depending on n . There are at most $4n^{60}$ quadruples (i, j, k, l) such that $l, k \in \mathbb{Z}; i, j \in \mathbb{N}$ and $n^{15} \geq |l|, |k|, i, j$. Thus, $4n^{60} e^{-cn} \geq P(E_3^{nc})$. Obviously, the expression on the right side of the previous equation is finitely summable over n , and thus $\sum_{n=1}^{\infty} P(E_3^{nc}) < \infty$.

Proof that the probability of E_4^{nc} is finitely summable over n . Within time e^{n^4} the random walk S can at most cross e^{n^4} crossings by R of $(0, 3n^5)$ or $(0, -3n^5)$. Furthermore, up to time e^{n^4} the random walk S can

cross any crossing at most e^{n^4} times. Thus, for E_4^{nc} to hold it is enough to ask that our test does not make a mistake whenever it compares the i -th crossing by S of the l -th crossing by R of $(0, 3n^5)$ or $(0, -3n^5)$ with the j -th crossing by S of the k -th crossing by R of $(0, 3n^5)$ or $(0, -3n^5)$, where $l, k \in \mathbb{Z}$; $i, j \in \mathbb{N}$ and $e^{n^4} \geq |l|, |k|, i, j$. Furthermore, by lemma 6, the probability of our test to make a mistake, when comparing the i -th crossing by S of the l -th crossing by R of $(0, 3n^5)$ or $(0, -3n^5)$ with the j -th crossing by S of the k -th crossing by R of $(0, 3n^5)$ or $(0, -3n^5)$ is smaller than e^{-cn^5} . There are at most $4e^{4n^4}$ quadruples (i, j, k, l) such that $l, k \in \mathbb{Z}$; $i, j \in \mathbb{N}$ and $e^{n^4} \geq |l|, |k|, i, j$. Thus, $4e^{4n^4} \times e^{-cn^5} \geq P(E_4^{nc})$. Obviously, the expression on the left side of the previous equation is finitely summable over n , and thus $\sum_{n=1}^{\infty} P(E_4^{nc}) < \infty$.

Proof that the probability of E_5^{nc} is finitely summable over n . We know that for a simple random walk starting at the origin, the probability that from 0 up to time k , R does not hit on $\{-3n^5, 3n^5\}$ is bounded above by $e^{-k_1 k / (n^{10})}$, where k_1 is a strictly positive constant not depending on k or n . Thus, the probability that the random walk R , between time zero and time n^{11} does not hit on $\{-3n^5, 3n^5\}$ is smaller, than $e^{-k_1 n}$. Thus, $e^{-k_1 n} \geq P(R([0, n^{11}]) \cap \{-3n^5, 3n^5\} = \emptyset)$. By symmetry, one can show that $e^{-k_1 n} \geq P(R([-n^{11}, 0]) \cap \{-3n^5, 3n^5\} = \emptyset)$. Thus, $2e^{-k_1 n} \geq P(E_5^{nc})$. From this it follows that $\sum_{n=1}^{\infty} P(E_5^{nc}) < \infty$.

Proof that the probability of E_6^{nc} is finitely summable over n . The bits of w^n and w_{34}^n are not i.i.d. As a matter of fact, S chooses among (x_1^{n+}, x_2^{n+}) and (x_1^{n-}, x_2^{n-}) the one it crosses firsts, to decide which one of the two will be equal to (x_1^n, x_2^n) and which one will be equal to (x_3^n, x_4^n) . So we will have to work with (x_1^{n+}, x_2^{n+}) and (x_1^{n-}, x_2^{n-}) directly. Let w^{n+} be the binary word associated with the crossing (t_1^{n+}, t_2^{n+}) by $R \circ S$. (Here, (t_1^{n+}, t_2^{n+}) denotes the first crossing by S of (x_1^{n+}, x_2^{n+}) .) Let w^{n-} be the binary word associated with the crossing (t_1^{n-}, t_2^{n-}) by $R \circ S$. (Here, (t_1^{n-}, t_2^{n-}) denotes the first crossing by S of (x_1^{n-}, x_2^{n-}) .) We can now find an upper bound for the probability of the event E_6^{nc} by using w^{n+} and w^{n-} . As a matter of fact, E_6^{nc} holds when in the interval $[-n^{26}, n^{26}]$ the only crossing by R of $(0, 3n)$ or $(0, -3n)$ having associated characteristic word bigger than w^{n+} is (x_1^{n+}, x_2^{n+}) and the only crossing by R of $(0, 3n)$ or $(0, -3n)$ having associated characteristic word bigger than w^{n-} is (x_1^{n-}, x_2^{n-}) . Now let (x_a, x_b) be any crossing by R of $(R(x_1^{n+}), R(x_2^{n+}))$ different from (x_1^{n+}, x_2^{n+}) . Then the interval $[\min\{x_1^{n+}, x_2^{n+}\}, \max\{x_1^{n+}, x_2^{n+}\}]$ and $[\min\{x_a, x_b\}, \max\{x_a, x_b\}]$ must be disjoint, since the different crossings of an interval, as well as crossings of disjoint intervals have their own interval disjoint. By the strong Markov property for R and for S , we have that what is happening in disjoint intervals is independent. Thus, for given $i \in 1, 2, \dots, n$, the event $\{w^{n+}(i) = 1\}$ is independent of the event $\{\text{the first crossing by } R \text{ of the } i\text{-th three unit interval of } (R(x_a), R(x_b)) \text{ during } (x_a, x_b) \text{ is straight}\}$. Recall, that for $\{w^{n+}(i) = 1\}$ to hold we need to have that the first crossing by R of the i -th three unit interval of $(R(x_1^{n+}), R(x_2^{n+}))$ during (x_1^{n+}, x_2^{n+}) is straight as well, as the first crossing by S of that first crossing during (t_1^{n+}, t_2^{n+}) . We already saw why the probability for a crossing over a three unit interval to be

straight is equals to $\frac{3}{4}$. Thus, $P(w^{n+}(i) = 1) = (\frac{3}{4})^2$. Using this and the above mentioned independence leads for a fix i to $P(\text{the first crossing by } R \text{ of the } i\text{-th three unit interval of } (R(x_a), R(x_b)) \text{ during } (x_a, x_b) \text{ is straight if } w^{n+}(i) = 1) = (\frac{3}{4})^3 + \frac{7}{16} = \frac{55}{64} < 1$. Thus, the probability that the binary word associated with the crossing (x_a, x_b) by R is bigger than w^{n+} is equal to $(\frac{55}{64})^n$. Now, there are at most n^{26} crossings (x_a, x_b) by R of $(0, 3n)$ or $(0, -3n)$ different from (x_1^{n+}, x_2^{n+}) and such that $x_a, x_b \in [-n^{26}, n^{26}]$. (To see this note that the different crossings considered here have their intervals mutually disjoint.) Thus the probability that there exists a crossing (x_a, x_b) by R of $(0, 3n)$ or $(0, -3n)$ having associated binary word bigger than w^{n+} , but different from (x_1^{n+}, x_2^{n+}) and such that $x_a, x_b \in [-n^{26}, n^{26}]$ is smaller than $n^{26} \times (\frac{55}{64})^n$. The same inequality holds for w^{n-} and (x_1^{n-}, x_2^{n-}) . This implies that $2n^{26} \times (\frac{55}{64})^n \geq P(E_8^{nc})$. Since the expression on the left side of the previous inequality is finitely summable over n , we get that $\sum_{n=1}^{\infty} P(E_8^{nc}) < \infty$.

Proof that the probability of E_7^{nc} is finitely summable over n .

Let x_2 be a non random number, such that $n^{11} \geq |x_2|$. Then the distance between x_2 and the integer point n^3 is smaller than $\frac{1}{2}n^{12}$. (At least for n big enough, which we will assume here.) Thus, the probability that the random walk S hits on the point n^3 between time t and $t + n^{25}$, given that at time t , S is at x_2 , (i.e. $S(t) = x_2$), is bigger than a constant p . (Where p does not depend on n or x_2 as long as $n^{11} \geq |x_2|$.) The probability that the random walk S crosses $(-n^3, n^3)$ in both directions in a straight way within time $4n^3$ right after time t , given that $S(t) = n^3$, is equal to $(\frac{1}{2})^{4n^3}$. Thus the probability, that S crosses $(-n^3, n^3)$ in both directions in a straight way within time n^{25} from a time t , such that $S(t) = x_2$ where $n^{11} \geq |x_2|$ is bigger than $p(\frac{1}{2})^{4n^3}$. Let $\nu_1, \nu_2, \dots, \nu_i, \dots$ be the sequence of stopping times defined inductively on i , in the following way: Let ν_1 be the time that for the first time S ends its first crossing of $(x_1^{n^5}, x_2^{n^5})$. Thus, $\nu_1 = t_2^{n^5}$. For a crossing (s, t) , we call $\max\{s, t\}$ the right end of the crossing (s, t) . Let ν_{i+1} be the right end of the first positive crossing by S of $(x_1^{n^5}, x_2^{n^5})$ after time $\nu_i + n^{25}$. (I.e. we ask that there exists $t < \nu_i$ such that (t, ν_i) is the first positive crossing by S of $(x_1^{n^5}, x_2^{n^5})$ such that $t, \nu_i \in [\nu_i + n^{25}, \infty[$.) Let $Y_1, Y_2, \dots, Y_i, \dots$ designate the collection of Bernoulli variables defined in the following way: Y_i is equal to one iff the random walk S within time n^{25} after ν_i crosses $(-n^3, n^3)$ in both directions in a straight way. When we condition under $x_2^{n^5}$, then the variables Y_i become i.i.d.. If on top of that we have that we condition under $x_2^{n^5} = x_2$, where x_2 is an integer number which is smaller in absolute value then n^{11} , then we get that the Y_i 's have parameter bigger than $p \times (\frac{1}{2})^{4n^3}$. Let now E_{71}^n be the event that for $i=1, 2, \dots$ up to $e^{n^4/10}$ we have that there exists at least one $i \in 1, 2, \dots, e^{n^4/10}$ such that $Y_i = 1$. Let E_{71}^{nc} designate the complement of E_{71}^n . Then, we get that given that $n^{11} \geq |x_2^{n^5}|$, we have that the event E_{71}^{nc} , has probability less then $\exp(\ln((1 - p \cdot \exp(-4n^3 \ln 2)) \cdot \exp(n^4/10)))$ which is smaller than $\exp(-p \cdot \exp(-4n^3 \ln 2) \cdot \exp(n^4/10))$. Roughly speaking, the previous formula behaves like $\exp(-e^{n^4/10})$, and thus is finitely summable over

n. This implies, that $\exp(-p \cdot \exp(-4n^3 \ln 2)) \cdot \exp(n^4/10) + P(x_2^{n^5} > n^{11}) \geq P(E_7^{nc})$. However, $P(E_5^{nc}) \geq P(x_2^{n^5} > n^{11})$. We already proved that $P(E_5^{nc})$ is finitely summable over n, and thus $P(E_7^{nc})$ is finitely summable over n too. Let E_{72}^n designate the event that $\nu_i \leq e^{n^4}$ for all $i \in 1, 2, \dots, e^{n^4/10}$. Then, $E_{71}^n \cap E_{72}^n \subset E_7^n$. Thus, $P(E_7^{nc}) \geq P(E_{71}^{nc}) + P(E_{72}^{nc})$. Since we already proved that $P(E_{71}^{nc})$ is finitely summable over n, we only need to prove that $P(E_{72}^{nc})$ is also finitely summable over n and this will then imply that $P(E_7^{nc})$ is finitely summable over n. Let E_{73}^n be the event that there are more than $n^{25} \cdot e^{n^4/10}$ positive crossing by S of $(x_1^{n^5}, x_2^{n^5})$ up to time e^{n^4} . Obviously, $E_{73}^n \subset E_{72}^n$ and thus $P(E_{72}^{nc}) \geq P(E_{73}^{nc})$. So if we could prove that $P(E_{73}^{nc})$ is finitely summable over n we would be done. When E_5^n holds, we get that $n^{11} \geq |x_1^{n^5}|, |x_2^{n^5}|$. Thus, for the same reasons as in the proof of $\sum_{n=1}^{\infty} P(E_2^{nc}) < \infty$, we get that $P(X(1) + X(2) + \dots + X(5 \cdot n^{11} \cdot n^{25} \cdot e^{n^4/10}) \geq e^{n^4}) \geq P(E_{73}^{nc} | E_5^n)$. Here, as before, the $X(i)$'s are i.i.d. such that $X(1)$ has the same distribution than the first passage time of the random walk $\{S(k)\}_{k \geq 0}$ at 1. (Recall that $\{S(k)\}_{k \geq 0}$ starts at zero.) Using the same trick as in the proof of $\sum_{n=1}^{\infty} P(E_2^{nc}) < \infty$, we get that $P(X(1) + X(2) + \dots + X(5 \cdot n^{11} \cdot n^{25} \cdot e^{n^4/10}) \geq e^{n^4})$ is smaller than $P(X(1)^{1/3} + X(2)^{1/3} + \dots + X(5 \cdot n^{11} \cdot n^{25} \cdot e^{n^4/10})^{1/3} \geq e^{n^4/3})$. Thus, by Tchebycheff, we get $E[X(1)^{1/3}] \cdot 5 \cdot n^{11} \cdot n^{25} \cdot e^{n^4/10} \cdot e^{-n^4} \geq P(E_{73}^{nc} | E_5^n)$. Since, $E[X(1)^{1/3}]$ is finite, the expression on the left side of the last inequality is finitely summable over n. However, $E_{73}^{nc} \subset (E_{73}^{nc} \cap E_5^n) \cup E_5^{nc}$, which implies that $P(E_{73}^{nc} \cap E_5^n) + P(E_5^{nc}) \geq P(E_{73}^{nc})$. But, $P(E_{73}^{nc} | E_5^n) \geq P(E_{73}^{nc} \cap E_5^n)$ and thus $P(E_{73}^{nc} | E_5^n) + P(E_5^{nc}) \geq P(E_{73}^{nc})$. Both of the terms in the sum on the right side of the last inequality have been proven to be finitely summable over n, and thus $P(E_{73}^{nc})$ is also finitely summable over n. This achieves to prove that $\sum_{n=1}^{\infty} P(E_7^{nc}) < \infty$.

Acknowledgement : I want to express my deepest gratitude to my thesis advisor Professor Harry Kesten, for providing such an interesting question and for all the time he spend helping me.

References

- [1] I. Benjamini and H. Kesten, *Distinguishing sceneries by observing the scenery along a random walk path*, J. Anal. Math. **69**, (1996), 97-135.
- [2] K. Burdzy, *Some path properties of iterated Brownian motion*, Seminar on Stochastic processes 1992 K.L., Chung, E. Cinlar and M.J. sharpe, eds., Birkhauser, Boston, 1993, 67-87.

- [3] W.Th.F den Hollander, *Mixing Properties for random walk in random scenery*, Ann. Probab. **16** (1988), 1788-1802.
- [4] C.D Howard, *Detecting defects in periodic scenery by random walks on \mathbb{Z}* , Random Structures Algorithms **8** (1996), no.1,59-74.
- [5] C.D. Howard, *Orthogonality of measures induced by random walks with scenery*, Combin. Probab. Comput **5** (1996), no. 3, 247-256.
- [6] C.D. Howard, *Distinguishing certain random sceneries on \mathbb{Z} via random walks*, Statis. Probab. Lett. **34** (1997), no.2, 123-132.
- [7] S. A. Kalikov, T, T^{-1} transformation is not loosely Bernoulli, Annals of Math, **115** (1982), 393-409.
- [8] M. Keane and W.Th.F. den Hollander, *Ergodic properties of color records*, Physica **138A** (1986), 183-193.
- [9] H. Kesten and F. Spitzer, *A limit theorem related to a new class of self similar processes*, Z. Wahrsch. verw. Geb. **50** (1979), 5-25.
- [10] H. Kesten, *Detecting a single defect in a scenery by observing the scenery along a random walk path*, Ito's Stochastic Calculus and Probability theory, Springer, Tokyo, (1996), 171-183.
- [11] H. Kesten, *Distinguishing and reconstructing sceneries from observations along a random walk path*, DIMACS; Series in Discrete Mathematics and Theoretical Computer Science, Vol. **41** (1998), 75-83, P. aldous and J. Propp, eds.
- [12] E. Lindenstrauss, *Indistinguishable sceneries*, Random Struct. and Alg. **14**, (1999), 71-86.
- [13] M. Loewe and H. Matzinger, *Scenery reconstruction in two dimensions with many colors*, submitted, (1999).
- [14] H. Matzinger, *Reconstructing a 3-color scenery by observing it along a simple random walk path*, accepted in Random Structures Algorithms