Scenery reconstruction with branching random walk

Serguei Popov* and Angelica Pachon†

June 14, 2011

Abstract

We study the problem of scenery reconstruction in arbitrary dimension, using observations registered in boxes of size k (for k fixed), seen along a branching random walk. We prove that, using a large enough k for almost all the realizations of the branching random walk, almost all sceneries can be reconstructed up to equivalence.

1 Introduction and the main result

The basic scenery reconstruction problem can be described as follows: suppose that to each $z \in \mathbb{Z}$ a color is assigned so that we have a scenery of colors on $\mathbb{Z}$. Then, discrete time simple random walk starts to move on these colored integers registering the color it sees at each time t, thus producing a new sequence of colors. The question is: Can the original coloring be reconstructed (possibly up to shift and/or reflection) with that sequence produced by the random walk?

In many situations the answer to the above question is positive, see e.g. [3, 7, 8] and references therein. It is interesting to note that, as was proved by Löwe and Matzinger in [5], even in dimension two with (sufficiently) many colors the reconstruction of sceneries

*Instituto de Matemática, Estatística e Computação Científica, rua Sérgio Buarque de Holanda 651, Cidade Universitária, CEP 13083–859, Campinas SP, Brasil.

Email: popov@ime.unicamp.br

†University of Bielefeld Postfach 10 01 31, 33501 Bielefeld, Germany.

Email: apachon@math.uni-bielefeld.de

is possible. It is important to say that solving this reconstruction problem is impossible for dimensions larger than 2 if we use the simple random walk (or any spatially homogeneous random walk), because when the dimension is at least 3, the random walk is transient, so infinitely many sites will not be even visited by the walker.

In this paper we study the question if the reconstruction is possible in arbitrary dimension d using the observations of the scenery by a *branching* random walk. We consider a scenery ξ with two colors, $\{0,1\}$, placed independently in each site of $\mathbb{Z}^d$ with probabilities $(1-p)$ and p respectively (so that $\xi(x)$ stands for the color of site x). As usual for this type of problem, the case of two colors is the most difficult to deal with; here it also means that the method of this paper works for arbitrary number of possible colors (and i.i.d. scenery).

The branching random walk in $\mathbb{Z}^d$ is described in the following way. The transition probabilities are those of the simple random walk, i.e., any particle jumps to one of its nearest neighbors chosen with equal probabilities, independently of everything else. At each step, any particle is substituted by two particles with probability b and is left intact with probability $1-b$, for some fixed $b \in (0,1)$. The process starts with only one particle at the origin. We denote by N_n the total number of particles at the moment n; clearly $(N_n, n = 0, 1, 2, \ldots)$ is a Galton-Watson process with the branching probabilities $\tilde{p}_1 = 1-b$, $\tilde{p}_2 = b$; this process is supercritical, so it is clear that $N_n \to \infty$ a.s.

We suppose also that each particle not only observes the color of the point $z \in \mathbb{Z}^d$ where it is located at the moment, but in fact observes the scenery in the box $z + [-k,k]^d$ (for some fixed k). In this way the branching random walk produces a tree with colored boxes (windows), so at time n we are going to see N_n colored boxes.

An example of the process η and the tree with colored windows can be seen on Figure 1 and Figure 2.

Now, we introduce some notations and give the formal definition of the observed process.

Let $\eta_n(z)$ be the number of particles in z at time n, $z \in \mathbb{Z}^d$ and $n \geq 0$, with $\eta_0(z) = \mathbf{1}\{z = 0\}$. We denote by $\eta_n = (\eta_n(z),\ z \in \mathbb{Z}^d)$ the configuration at time n of the branching random walk on $\mathbb{Z}^d$ starting at the origin with branching probability b. Let $\Omega_1 = \{(\eta_n)_{n\in\mathbb{N}}\}$ be the space of all possible "evolutions" of the branching random walk, and let $\Omega_2 = \{0,1\}^{\mathbb{Z}^d}$ be the space of all possibles sceneries ξ. Let $G = \bigcup_{n=1}^{\infty} G_n$ be the genealogical tree of the

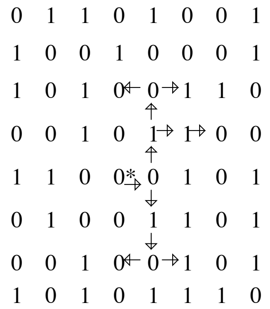

Figure 1: A branching random walk η on a given scenery (0* is the origin)

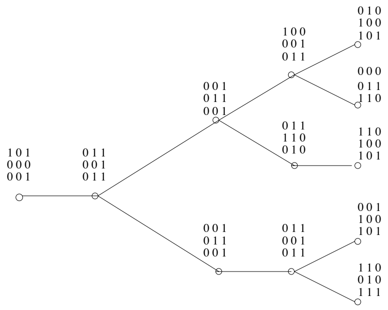

Figure 2: Tree with colored windows from the process at the Figure 1

Galton-Watson process, where $G_n = \{v_1^n, \dots, v_{N_n}^n\}$ are the particles of nth generation.

Let $\Psi : G \to \mathbb{Z}^d$ be the function that shows the position of v_j^i in $\mathbb{Z}^d$ (i.e., $\Psi(v_j^i) = z$ if the particle corresponding to v_j^i is in $z \in \mathbb{Z}^d$ at time i). Observe that, according to our notations,

$$\{z \in \mathbb{Z}^d : \eta_n(z) \geq 1\} = \{\exists j; \Psi(v_j^n) = z, j = 1, \dots, N_n\}.$$

Let $\mathfrak{C}^k(z)$ be the scenery inside the box $z + [-k,k]^d$. So, $\mathfrak{C}^k(z)$ is a $(2k+1) \times (2k+1)$ matrix.

Using these notations, the observed process is represented by a marked tree (in the sense that each mark is a matrix of colors) $\chi = \bigcup_{n \in \mathbb{N}} \chi_n$, where $\chi_n = [\mathfrak{C}^k(\Psi(v_j^i)),\ v_j^i \in G_n]$, and we denote by $\Omega_3 = \{(\chi_n)_{n \in \mathbb{N}}\}$ the space of all possible realizations of the above process.

We assume that $(\eta_n)_{n \in \mathbb{N}}$ and ξ are independent and distributed with the laws P_1 and P_2, respectively.

Two sceneries ξ and ξ' are said to be equivalent (in this case we write $\xi \sim \xi'$), if there exists $a \in \mathbb{Z}^d$ such that $\xi(a + x) = \xi'(x)$ for all x.

Now we are ready to formulate the main result of this paper. In the following, the measure P is the product measure $P_1 \otimes P_2$.

Theorem 1 *Suppose that $p = \frac{1}{2}$. For all $k \geq 5$ independently of the dimension we have the following: for any b there exists a measurable function $\Lambda : \Omega_3 \to \Omega_2$ such that $P(\Lambda(\chi) \sim \xi) = 1$.*

In words, we are going to prove that, if the size of the window is large enough, then for almost all realizations of the branching random walk, almost all sceneries can be reconstructed. As usual, the reconstruction works only up to equivalence.

Remark 1 *First, let us note that the method of this paper works for $p \neq \frac{1}{2}$ as well, however, one may need a larger k (observe that, if p is very close to 0 or 1, then usually particles see the configurations consisting of only 0s or 1s). Then, it is plausible that (especially in higher dimensions) one can refine the method of this paper to obtain that smaller k would be sufficient for the reconstruction. However, the most interesting case $k = 0$ (i.e., a particle only sees the scenery in the site where it is located at the moment) cannot be treated by the approach of this paper. In our opinion, reconstructing many-dimensional sceneries with branching random walk and with $k = 0$ is an interesting and difficult problem (note that the reconstruction may only work up to shift and rotation). We conjecture that it should be possible to do the reconstruction with $k = 0$ for all $p \in (0, 1)$ and $b > 1$ (since by time n one gets an exponential "amount of information" about the scenery in the box of size n), however, it is still not clear to us how the reconstruction algorithm should work in this case.*

2 Main idea: the good sites

Let e_i be the canonical d-dimensional vectors, $i = 1, \ldots, d$. Consider the events

$$A^k_{ab}(z) = \{\exists y \in [-k, k]^d \text{ such that } \xi(z + y + a) \neq \xi(z + y + b) \text{ and } y + a, y + b \in [-k, k]^d\},$$

and

$$A^k(z) = \bigcap_{a,b \in \{\pm e_i, i=1,\ldots,d\}, a \neq b} A^k_{ab}(z).$$

Consider the following

Definition 1 *We say that the site z is k-good (or, equivalently, that $\mathfrak{C}^k(z)$ is a k-good configuration), if the event $A^k(z)$ occurs.*

The key observation here is that, if a particle jumps from a good site, we are seeing two different configurations which determine the direction of the jump, so we are able to deduce the particle's relative position after the jump.

Lemma 1 *For $k' \geq 2$ independently of the dimension,*

$$P_2(\textit{configuration of size } k' \textit{ with center at the origin is good}) \geq p_{cr}(2), \tag{1}$$

where $p_{cr}(2)$ is the critical probability for the site percolation in $\mathbb{Z}^2$.

Proof: We have for $a, b \in \{\pm e_i, i = 1, \dots, d\}$, $a \neq b$

$(A^{k'}_{ab}(z))^c = \{\forall y \in [-k', k']^d$ such that $y + a, y + b \in [-k', k']^d, \xi(z + y + a) = \xi(z + y + b)\}$.

We now want to obtain an upper bound for the probability of the above event. Consider first the case $a \neq -b$; by symmetry, without restriction of generality we may assume that $a = e_1$, $b = e_2$. Then, if $(A^{k'}_{e_1 e_2}(z))^c$ occurs, the configuration in $z + [-k', k']^d$ is uniquely determined by the values of ξ on the set

$$z + \big(\{y \in [-k', k']^d : y \cdot e_1 = k\} \cup \big(\{y \in [-k', k']^d : y \cdot e_2 = -k\}\big)$$

($a \cdot b$ stands for the scalar product of vectors a and b). So, we obtain (by conditioning on ξ-values on this set) that

$$P_2((A^{k'}_{e_1 e_2}(z))^c) = 2^{-(2k')^2(2k'+1)^{d-2}}.$$

Now, suppose that $a = -b$; then we may assume that $a = e_1$, $b = -e_1$. In this case, if the event $(A^{k'}_{e_1,-e_1}(z))^c$ occurs, the configuration in $z + [-k', k']^d$ is uniquely determined by the values of ξ on the set

$$z + \{y \in [-k', k']^d : y \cdot e_1 \leq -(k - 1)\}.$$

So, in the same way we obtain

$$P_2((A^{k'}_{e_1,-e_1}(z))^c) = 2^{-(2k'-1)(2k'+1)^{d-1}}.$$

Thus, in both cases we obtain that $P_2((A^{k'}_{ab}(z))^c) \leq 2^{-m^d}$, where $m = 2k' - 1$. So,

$$P(A^{k'}(z)) \geq 1 - 2^{-m^d}\binom{2d}{2}. \tag{2}$$

In what follows, we show that for $m \geq 3$ we obtain

$$1 - 2^{-m^d}\binom{2d}{2} \geq p_{cr}(2) \quad \text{for all } d \geq 2, \tag{3}$$

we use the value of the rigorous upper bound $p_{cr}(2) \leq 0.679492$, see [9].

To show (3), it is enough to prove that

$$\frac{1}{d}\ln\log_2\big(\frac{d(2d-1)}{1-p_{cr}(2)}\big) \leq \ln 3$$

for all $d \geq 2$. This is clearly true for $d = 2$; also, denoting $f(x) = \frac{1}{x}\ln\log_2\big(\frac{x(2x-1)}{a}\big)$, it is straightforward to obtain that $f'(x) \leq 0$ for $x \geq 2$, where a is any value in $[0.32, 0.5)$. The latter interval comes out when one uses the upper bound for $p_{cr}(2)$ given above, and the lower bound $\frac{1}{2}$ (it is known that the critical probability for site percolation is bigger than the critical probability for bond percolation, and the latter value is exactly $\frac{1}{2}$, see [6]).

Thus, with $m \geq 3$ we obtain $k' \geq 2$, and this concludes the proof of Lemma 1. □

3 Proof of Theorem 1

3.1 Existence of an infinite cluster of good sites

First, we prove that if k is large enough, then the good sites are "connected".

Lemma 2 *For all k large enough there is percolation by good sites, that is:*

$$P(\textit{there exists an infinite cluster of } k\textit{-good sites with respect to } k) = 1.$$

Proof: Taking k' given above we have

$$P(x \text{ is } k'\text{-good}) > p_{cr}(2),$$

where $p_{cr}(2)$ is the critical probability for the site percolation in $\mathbb{Z}^2$. Since $p_{cr}(d) \leq p_{cr}(2)$, where $p_{cr}(d)$ is the critical probability for the site percolation in $\mathbb{Z}^d$, then we also have

$$P(x \text{ is } k'\text{-good}) > p_{cr}(d).$$

Now, we divide $\mathbb{Z}^d$ into disjoint boxes of size k' (here and in the sequel "box of size m" means a translate of $[-m, m]^d$) and construct an independent model of site percolation, where the sites are identified with the boxes of size k' at the moment. Note that k' chosen in this way does not depend on the dimension.

Since the event "x is k'-good" is equivalent to "the configuration (box) of size k' with center at x is good ", we have that

$$P(\mathfrak{C}^{k'}(x) \text{ is good}) = P(x \text{ is } k'\text{-good}) > p_{cr}(d).$$

Hence, we have percolation of good boxes of size k', that is, a.s. there exists an infinite cluster of good configurations of size k'. Precisely, this means that the set $\{z \in \mathbb{Z}^d : z \text{ belongs to some good configuration}\}$ is connected.

However, we want percolation of good sites, therefore we consider now the boxes of size $k = 2k' + 1$. Then for any point y in $\mathbb{Z}^d$ which is in some box of size k' from the partition, for instance in $x + [-k', k']^d$, all the unit translations (i.e., by $\pm e_i$) of $y + [-k, k]^d$ contain the box $x + [-k', k']^d$.

This fact implies that if the configuration $\mathfrak{C}^{k'}(x)$ is good, then all the translations of $\mathfrak{C}^k(y)$ by $\pm e_i$ are different, therefore $\mathfrak{C}^k(y)$ is a good configuration and y is a k-good site.

From this, we have that a site is k-good if it lies in any good box of size k'. Since we proved that there exists percolation of good boxes of size k', then there is also percolation of k-good sites. $\square$

With all this to hand, the idea is to begin the reconstruction on the points of the infinite cluster of good points, since for them we know the particle's jump direction (see Definition 1). However, we must still prove that the restricted process visits all points of this cluster a.s.

3.2 Scenery reconstruction on the infinite cluster of good sites

To describe the reconstruction algorithm, we will need two more results, namely Lemma 3 and Lemma 4.

Lemma 3 *The branching random walk η is recurrent on $\mathbb{Z}^d$, i.e., all the sites are visited an infinite number of times a.s.*

Proof: Since N_n is a Galton-Watson process with the mean offspring $1+b$, we have

$$E(N_n) = (1+b)^n. \tag{4}$$

From (4) and using the local CLT for the simple random walk (see e.g. [4] Theorem 1), we obtain that for some c^*

$$E\eta_n(0) = E(\# \text{ of particles of } \eta_n \text{ at the origin}) \quad \geq \quad (1+b)^n \frac{c^*}{n^{\frac{d}{2}}}, \tag{5}$$

which is greater than 1 for all n large enough. Let t_0 be the smallest of these (of course, t_0 depends on the dimension), so that $E\eta_{t_0}(0) > 1$.

Now, using the process η, we are going to construct a Galton-Watson process with expected number of offspring greater than 1. It is constructed using (5) and can be restricted to a box of size t_0 as follows: we start with one particle at the origin at time 0 (this is the 0th generation) and take the particles that are in the origin at time t_0 as the first generation. Let us denote their number by Z_1 (that is, Z_1 is the size of the first generation). For the second generation of size Z_2 we consider the descendants of each one of those Z_1 particles which are at the origin at time $2t_0$, and so on.

From the theory of branching processes we know that if $E(Z_1) > 1$, as is the case, the process does not die with positive probability (see [1] Theorem 1), i.e.,

$$P(Z_n > 0 \text{ for all } n) = p^*, \text{ with } p^* \in (0,1).$$

For any $z_i \in \mathbb{Z}$, we define the following events:

$$B_i \quad = \quad \{\text{the Galton-Watson process as above starting in } z_i \text{ survives}\}$$

$i = 1, \ldots, \infty$. These events have probability p^* and they are independent. This shows that a.s. at least one of the events B_i occurs, so at least one site is visited infinitely often. By the irreducibility of the simple random walk, all sites of $\mathbb{Z}^d$ are visited an infinite number of times a.s. Thus, we proved the recurrence of η. $\square$

Before formulating the next lemma we need the following definition:

Definition 2 *We define the branching random walk η restricted on $U \subset \mathbb{Z}^d$ by removing all particles that step outside U, and we denote such a process by η^U.*

Lemma 4 *For $k \geq 5$, with positive probability, the branching random walk restricted to the infinite cluster of good sites visits all cluster's points infinitely many times.*

Proof: As observed in Lemma 2, the infinite cluster of good sites defined on k with $k = 2k'+1$ and $k' \geq 2$, contains a cluster of good configurations (boxes) of size at least k'.

We will take for this proof the space of the partition of $\mathbb{Z}^d$ into disjoint boxes of size k'. Define for $x \in \mathbb{Z}^d$

$$Y_x = \begin{cases} 1 & \text{if } \mathfrak{C}^{k'}((2k'+1)x) \text{ is good} \\ 0 & \text{otherwise,} \end{cases}$$

and consider the following events:

$$\begin{aligned} U_x &:= \{\mathfrak{C}^{k'}((2k'+1)x) \text{ belongs to the infinite cluster of } k'\text{-good boxes}\}, \\ W_x &:= \{\mathfrak{C}^{k'}((2k'+1)x) \text{ belongs to a box of size } N \text{ of } k'\text{-good boxes.}\} \end{aligned}$$

These events are increasing with respect to the field $(Y_x, x \in \mathbb{Z}^d)$ and they have positive probability. Since the family $(Y_x, x \in \mathbb{Z}^d)$ is i.i.d., using the FKG inequality, (see [2] Theorem 2.4), we have

$$P(U_x \cap W_x) \geq P(U_x)P(W_x) > 0.$$

Let $T^i(w) = w - ie_d$, $f(x) = \mathbf{1}\{U_x \cap W_x \text{ holds, for } i = 0, \ldots, \infty\}$, and let $\gamma_i(w)_{i \geq 0}$ be a sequence of random variables with Bernoulli distribution, defined by $\gamma_i(w) = f\left(T^i(w)\right)$.

Since f is an integrable function and T is an ergodic transformation, then by the ergodic theorem (see e.g. [10] Section 1.2) we have

$$\lim_{n \to \infty} \frac{\gamma_0 + \gamma_1 + \gamma_2 + \cdots + \gamma_{n-1}}{n} = E(f) \quad \text{a.s.,}$$

since $E(f) = P(U_x \cap W_x) > 0$, we have

$$\gamma_0 + \gamma_1 + \gamma_2 + \cdots + \gamma_{n-1} \underset{n \to \infty}{\longrightarrow} \infty.$$

Thus, the event $\{U_x \cap W_x\}$ occurs for infinite many configurations. It means that the super critical cluster of good configurations contains arbitrarily large "boxes" of them. Since all the sites in any $\mathfrak{C}^{k'}(x)$ good are k-good, then, it is possible to find boxes with size t_0 of k-good sites, where t_0 comes from the proof of Lemma 3.

We choose any of those boxes of size t_0 from the cluster of good sites. Now we start a Galton-Watson process restricted to the box of size t_0, identical to that considered in the proof of Lemma 3.

By construction, $E(Z_1) > 1$, so

$$P(\text{survival}) = P(Z_n > 0 \text{ for all } n) > 0,$$

in other words, the probability that the center of the box has at least one particle at each time nt_0, for $n \geq 1$ is greater than zero. Then, the probability that the restricted process visits the center of this box infinitely many times is positive as well.

Furthermore, since the chosen box is within the cluster of good points, in this cluster we can find a path between the center of the box and any other site. This means that, with positive probability a particle at the center of the box or any of its descendants walk along this path without dying (since the path is inside the cluster). Therefore, with positive probability any other site of the cluster will be visited infinitely many times. With this we finish the proof of Lemma 4. □

Let us introduce some more notation. We denote by $\mathcal{A}_k$ the infinite cluster of good sites with respect to k and by $\delta(\mathcal{A}_k)$ the internal boundary of $\mathcal{A}_k$.

For the next lemma it is important to note the following: when the process η arrives to a point y_1 which belongs to $\delta(\mathcal{A}_k)$ and we restrict the process to $\mathcal{A}_k$, from the point of view of the tree obtained by the branching random walk, we are eliminating the branches with non-good sites on the subtrees beginning in y_1. Thus, we obtain subtrees with good configurations only.

Lemma 5 *With probability 1 there are infinitely many infinite subtrees of good configurations, which contain all the good points of the cluster $\mathcal{A}_k$.*

Proof: This is an immediate consequence of Lemma 4. □

Therefore, we have that a.s. there exist an infinite number of subtrees containing all the points of the infinite cluster of good sites. With any of those subtrees we can reconstruct the scenery inside the cluster.

The algorithm for the reconstruction is described as follows: Each time η arrives to a good point, even though it is not in $\mathcal{A}_k$, (we do not know this at the moment), we start the reconstruction following the subtree with good configurations only. Then, since we are following all of these subtrees which containing all the good sites of the cluster with good configurations, the scenery inside the cluster will be reconstructed.

Thus, we have the first part of an algorithm that proves Theorem 1, i.e., we have found a measurable function that, applied to the observed process (the tree with windows), allows us to obtain a scenery which is equivalent to the original one, but only in the cluster of good sites.

3.3 Scenery reconstruction out of the infinite cluster of good sites

Let w be a site which belongs to $\delta(\mathcal{A}_k)$ and define by $\rho(w)$ the vertical semi-infinite line starting at w in the direction of $-e_d$ (in dimension two or three, this is the downward direction), i.e., $\rho(w) = \{w, w - e_d, w - 2e_d, w - 3e_d, \dots\}$.

Lemma 6 *For all w in $\delta(\mathcal{A}_k)$, $\rho(w)$ contains infinitely many good sites from $\mathcal{A}_k$ a.s.*

Proof: Let $(\gamma_i(w))_{i\geq 0}$ be a sequence of random variables with Bernoulli distribution as in Lemma 4. Define $T^i(w) = w - ie_d$ for $i \geq 0$, and let $f(x) = \mathbf{1}\{x \text{ is a good site from } \mathcal{A}_k\}$. Since $E(f) = P(x \text{ is a good point}) > p_{cr}(d)$ then by the ergodic theorem (see e.g. [10] Section 1.2) we have

$$\gamma_0 + \gamma_1 + \gamma_2 + \cdots + \gamma_{n-1} \underset{n\to\infty}{\longrightarrow} \infty.$$

Thus, we proved that there are infinitely many sites in $\mathcal{A}_k \cap \rho(w)$. □

This last lemma says that, a.s., for each $w \in \delta(\mathcal{A}_k)$ there exists a site $w' \in \delta(\mathcal{A}_k)$ that is exactly below w.

Lemma 7 *The infinite cluster of good points $\mathcal{A}_k$ is "uniquely determined", that is, one cannot find a translation (i.e., a map $x \to x + a$, with $a \in \mathbb{Z}^d$) that maps the cluster onto itself.*

Proof: To prove this lemma, it is enough to show that $\mathcal{A}_k + a \neq \mathcal{A}_k$ for any $a \in \mathbb{Z}^d$. We are not going to think about the scenery inside the cluster, the cluster's form will be enough.

Consider the following events:

$$H_n = \{(n-1)a \text{ is in the cluster, } na \text{ is not good}\}.$$

It is easy to show that $P(H_0) > 0$ and, given that the sequence $\{\mathbf{1}_{H_n}\}_{n\geq 0}$ is ergodic, a.s. at least one of the events H_n occurs. This shows that $\mathcal{A}_k + a$ is different from $\mathcal{A}_k$, because $(n-1)a$ is in the cluster but $(n-1)a + a = na$ is not. $\square$

What the previous lemma says is that, for the cluster of good points (clearly, by comparison with the site percolation, there exists only one such cluster), it is not relevant where we begin to observe the cluster since we will always be able to identify the starting point. Thus, the observed sceneries "relate" (for instance, it is possible to identify the point where we are starting to observe the cluster along the first subtree in the second subtree and vice versa). Note, however, that one can easily construct (periodic) deterministic colorings for which all the points are good, but, because of periodicity, it is not possible to identify the points of the cluster generated by the first subtree in the cluster generated by the second one (so, it is important that the cluster was obtained in an i.i.d. way).

We denote by $g_i(y_i)$ the i-th subtree with good configurations only and starting at y_i and by $g = \bigcup_i g_i(y_i)$ the forest composed by the subtrees $g_i(y_i)$, $i = 1, \ldots, \infty$.

Since in the good sites we know the direction of each jump, in the cluster generated by $g_1(y_1)$ we have the relative location of each site with respect to y_1. Furthermore, since the cluster is uniquely determined, we are able to identify y_1 in any cluster generated by $g_i(y_i)$, for $i = 2, \ldots, \infty$.

On the other hand, since every subtree of g contains all the points of $\mathcal{A}_k$, by Lemma 7, we are able to know when we are in w for every subtree, for all w in $\delta(\mathcal{A}_k)$ and all $i = 1, \ldots, \infty$.

Once again, since we know the direction of each jump in the good sites and, since the cluster of good sites is uniquely determined, we know its form. Thus, the distance between w and w' is known as well, where w' is the point of $\delta(\mathcal{A}_k)$ that is exactly below w.

So, the algorithm for the scenery reconstruction outside of the infinite cluster of good sites can be described as follows.

For each subtree $g_i(y_i)$ we locate all the vertices corresponding to w and consider the paths that connect w to the next point in $\delta(\mathcal{A}_k)$ in the branches with sites outside the cluster of good sites, i.e., the sequences of distinct vertices

$$\{v_0^i, v_1^i, \dots, v_l^i;\ v_j^i \in g_i(y_i),\ j = 0, \dots, l \text{ and } l \in \mathbb{N}\}$$

which are not k-good, $w \sim v_0^i$, $v_k^i \sim v_{k+1}^i$ and $v_l^i \sim w'$, but w' is k-good. We know by Lemma 4 that with positive probability we can obtain from this new site in $\delta(\mathcal{A}_k)$ the whole cluster of good sites, thus, we are able to identify if we are in w' or not. Then, with positive probability a particle can go from w to w' by a path which lies outside the cluster of good sites. Since the subtree $g_i(y_i)$ has infinite number of vertices corresponding to w we have a.s. infinite number of paths from w to w' that pass outside the cluster of good sites. We describe now how one can recognize such a path.

Let $d(w, w')$ be the $\mathcal{L}_1$-distance between w and w'. Choosing the path such that the number of steps between hittings of w and w' is equal to $d(w, w')$ we obtain the shortest path between w and w' and it is exactly the vertical one. When any particle walks along this path, we obtain the scenery there.

Doing the same for all the points of $\delta(\mathcal{A}_k)$ we reconstruct the scenery outside of the cluster of good points.

In this way one can reconstruct the whole scenery, thus proving Theorem 1. $\square$

Acknowledgements

We thank the referee for valuable comments and suggestions.

References

[1] K.B. Athreya, P.E. Ney (1972) *Branching Processes.* Springer, New York.

[2] G. Grimmett (1989) Percolation. Springer-Verlag New York.

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

[4] G. Lawler (1991) *Intersections of Random Walks.* Birkhäuser Boston.

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

[6] H. Kesten (1980) The critical probability of bond percolation on the square lattice equals $\frac{1}{2}$. *Comm. Math. Phys.* **74**, 41-59.

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

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

[9] J.C. Wierman (1995) Substitution method critical probability for the square lattice site percolation model. *Combinatorics, Probability and Computing* **4**, 181-188.

[10] K. Ulrich (1985) *Ergodic Theorems.* Walter de Gruyter, Berlin-New York.