Fractal Percolation and Set-valued Substitutions

Peter van der Wal*

February 9, 2004

1 Introduction

Mandelbrot (1974) introduced fractal percolation as a model for turbulence. In dimension two, the model can be described as follows. Choose an integer $M \geq 2$ and a parameter $0 \leq p \leq 1$ and define random sets $(K_n)_{n\in\mathbb{N}}$ in the unit square as follows. Let K_0 be the unit square itself. In the obvious way, divide K_0 into M^2 equal sub-squares and, for each of these sub-squares independently, color it black with probability p and white with probability $1-p$. The set K_1 will be the set of points that have been colored black. Similarly, define the set K_{n+1} by dividing all black squares from K_n into M^2 sub-squares, each of which is colored black with probability p and white with probability $1-p$, independently of all other colorings. The sequence $(K_n)_{n\in\mathbb{N}}$ is decreasing and we denote the limit set $\bigcap_{n\in\mathbb{N}} K_n$ by K. See figure 1 for a realization of the sets K_n for $M = 3$, $p = 0.7$ and $n = 1, 2, 3, 4$. We say that $K \subseteq [0,1]^2$ percolates if there is a connected component of K that intersects the left and the right side of the unit square and we denote the percolation function $\mathbb{P}_p(K \text{ percolates})$ by $\theta_M(p)$. It is not difficult to show that $\theta_M(p)$ is an increasing, right-continuous function in p. Chayes, Chayes and Durrett (1988) showed that the percolation function exhibits a phase transition in p, i.e., there is a non-trivial critical value $p_c(M)$ such that $\theta_M(p) = 0$ for $p < p_c(M)$ and $\theta_M(p) > 0$ for $p > p_c(M)$. Moreover, Dekking and Meester (1990) proved that the percolation function is discontinuous at the critical value.

For all $M \geq 2$, the value of $p_c(M)$ is unknown, but through the years various bounds have been given. A rather trivial lower bound for $p_c(M)$ is obtained by observing that the sequence $(Z_n)_{n\in\mathbb{N}}$, where Z_n denotes the number of black squares in K_n, is an ordinary branching process. Since this branching process dies out whenever $pM^2 \leq 1$, we obtain $p_c(M) \geq \frac{1}{M^2}$. A more ingenious branching process argument enabled Chayes, Chayes and Durrett (1988) to prove that $p_c(M) \geq \frac{1}{\sqrt{M}}$ and recently White (2001) gave a computer aided proof that $p_c(2) \geq 0.810$.

Chayes, Chayes and Durrett (1988) were the first to give a rigorous upper bound for $p_c(M)$. They showed that $p_c(M) \leq p^*(M)$ for $M \geq 3$, where $p^*(M)$ is the infimum over p for which $x = (px)^{M^2} + (M^2-1)(px)^{M^2-1}(1-px)$ has a root in the half open interval $(0,1]$. The proof of this result was made more transparent by Dekking and Meester (1990) by translating it in terms of set-valued substitutions (Dekking and Meester use the term *multi-valued substitutions*). Set valued substitutions are substitutions on sets of words, rather than on single words. We postpone a formal definition to Section 2 of this paper. For $M = 3$, Dekking and Meester (1990) sharpened the upper bound to $p_c(3) \leq 0.991$.

In this paper, we generalize the set-valued substitution approach. For each M, we construct sequences $(\Phi_{k,M})_{k\in\mathbb{N}}$ and $(\Psi_{k,M})_{k\in\mathbb{N}}$ of set-valued substitutions and to each $\Phi_{k,M}$ and $\Psi_{k,M}$ we associate a critical value $p_c(\Phi_{k,M})$, respectively, $p_c(\Psi_{k,M})$. We prove that the $p_c(\Phi_{k,M})$ monotonically increase to $p_c(M)$ and that the $p_c(\Psi_{k,M})$ monotonically decrease as $k \to \infty$. Unfortunately, we were not able to

*Address: Eurandom, P.O. Box 513, 5600 MB Eindhoven, The Netherlands. Email: vanderwal@eurandom.tue.nl

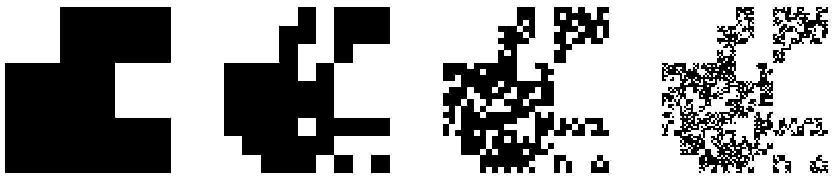


Figure 1: Sets K_1, K_2, K_3 and K_4 of fractal percolation with $M = 3$ and $p = 0.7$.

prove that the $p_c(\Psi_{k,M})$ decrease to $p_c(M)$. We give a computer aided proof that $p_c(\Phi_{k=0,M=2}) \geq 0.784$, $p_c(\Phi_{k=1,M=2}) \geq 0.858$, $p_c(\Phi_{k=0,M=3}) \geq 0.715$ and $p_c(\Psi_{k=0,M=3}) \leq 0.958$.

This paper is organized as follows. In Section 2, we define set-valued substitutions and Bernoulli random substitutions, of which fractal percolation is a specific example. In Section 3, we state our main results. The construction of the sequences $(\Phi_{k,M})_{k\in\mathbb{N}}$ and $(\Psi_{k,M})_{k\in\mathbb{N}}$ is covered in Section 4. The proofs of our main results are postponed to Sections 5, 6 and 7.

2 Definitions

To be able to state our main results, we need to introduce (Bernoulli) random substitutions and (increasing) set-valued substitutions. For ease of exposition, we will give the definitions for dimension 1. At the end of this section we will give some indications on how to generalize these concepts to dimension 2.

2.1 Bernoulli Random Substitutions and Fractal Percolation

Let A be a finite set called the alphabet and let A^* denote the free semi-group generated by A, i.e., the set of all finite words in A. A substitutions is nothing but a homomorphism on A. The random analogue of a substitution is given in the following definition.

Definition 2.1. Let $(\sigma_k)_{k\in\mathbb{N}}$ be a sequence of independent identically distributed random maps from A to A^*. Define a random map σ on A^* by

$$\sigma(u) = \sigma_0(u_0)\ldots\sigma_k(u_k)$$

for $u = u_0 \ldots u_k \in A^*$. The random map σ on A^* is called a *random substitution.* We define the n-fold iterate σ^n to be the composition of n independent copies of the substitution σ.

We say that σ has *fixed base* if there is a positive integer M such that $\sigma(i)$ is a word of length M for all $i \in A$, where M is called the base of the substitution.

Definition 2.2. Let P be a Markov matrix indexed by $A \times A$. We say that a base M random substitution σ is a *Bernoulli random substitution* with transition matrix P, if for all $i \in A$, the letters

$$(\sigma(i))_0, \ldots, (\sigma(i))_{M-1}$$

are independent with $\mathbb{P}((\sigma(i))_k = j) = P_{ij}$ for $j \in A$, $0 \leq k \leq M-1$.

All random substitutions in this paper will be Bernoulli random substitutions with fixed base $M \geq 2$.

Fractal percolation can be viewed as a Bernoulli random substitution. Let A be a finite alphabet and let $\prec$ be a partial ordering of A, such that A has a unique maximum $\bar{a}$ and a unique minimum $\underline{a}$.

Then fractal percolation on A^* with parameters M and p is the Bernoulli random substitution with base M and transition matrix P, given by

$$P_{ij} = \begin{cases} 1 & \text{if } i \neq \bar{a}, j = \underline{a} \\ p & \text{if } i = \bar{a}, j = \bar{a} \\ 1-p & \text{if } i = \bar{a}, j = \underline{a} \\ 0 & \text{otherwise.} \end{cases}$$

If we identify the letters $\bar{a}$ in $\sigma^n(\bar{a})$ with black squares and the letters $\underline{a}$ with white squares, then the analogy with the set representation of fractal percolation is obvious.

2.2 Set-valued Substitutions

Turning to the definition of set-valued substitutions, let $\mathcal{A}^*$ denote the set of all finite subsets of A^* and consider two binary operations on $\mathcal{A}^*$:

$$\begin{aligned} V \cup W &= \{u : u \in V \text{ or } u \in W\} \qquad \text{(union)} \\ VW &= \{vw : v \in V \text{ and } w \in W\} \qquad \text{(concatenation).} \end{aligned}$$

Definition 2.3. A *set-valued* substitution Φ is a homomorphism on $\mathcal{A}^*$ respecting unions and concatenations, i.e.,

$$\Phi(V \cup W) = \Phi(V) \cup \Phi(W) \quad \text{and} \quad \Phi(VW) = \Phi(V)\Phi(W),$$

for all $V, W \in \mathcal{A}^*$.

Remark 1.

1. Since $\mathcal{A}^*$ is generated by the singletons, i.e., the sets containing one letter, a set-valued substitution Φ is completely determined by the images $\Phi(i) = \Phi(\{i\})$, $i \in A$, of the singletons.
2. If the sets $\Phi(i)$ are disjoint for all $i \in A$, then the sets $\Phi(V)$ and $\Phi(W)$ are disjoint for all disjoint sets $V, W \in \mathcal{A}^*$.

We say that Φ has *fixed base* if there is a positive integer M such that $\Phi(i) \subset A^M$ for all $i \in A$, where M is called the base of the set-valued substitution. In this paper, we only consider set-valued substitutions Φ with fixed base $M \geq 2$, for which $\{\Phi(i)\}_{i \in A}$ is a partition of A^M. This special class of set-valued substitutions appeared under various names in literature, for example as *multi-valued substitutions* in Dekking and Meester (1990) and as $0L$*-systems* in Rozenberg and Salomaa (1976). It easily follows from the second remark above that $\{\Phi^n(i)\}_{i \in A}$ partitions A^{M^n} for all n, where Φ^n is the n-fold iterate of Φ.

Example 1. Let $A = \{0, 1\}$, $M = 2$, $\Phi(0) = \{00, 01, 10\}$ and $\Phi(1) = \{11\}$. Then

$$\begin{aligned} \Phi(\{10, 101\}) &= \Phi(\{10\} \cup \{101\}) \\ &= \Phi(10) \cup \Phi(101) \\ &= \Phi(1)\Phi(0) \cup \Phi(1)\Phi(0)\Phi(1) \\ &= \{1100, 1101, 1110, 110011, 110111, 111011\}. \end{aligned}$$

2.3 Increasing Set-valued substitutions

Consider a partial ordering $\prec$ of A. For words $v, w \in A^m$ we write $v \preceq w$ if $v_i \preceq w_i$ for all $0 \leq i \leq m-1$. A set $V \in \mathcal{A}^*$ is said to be *increasing* (w.r.t. $\prec$), if $v \in V$ and $v \preceq w$ imply $w \in V$.

Definition 2.4. We say that a set-valued substitution Φ is increasing if $\Phi(V)$ is increasing for all increasing sets $V \in \mathcal{A}^*$.

Remark 2. In fact, Φ is increasing, whenever $\Phi(J)$ is increasing for all increasing $J \subseteq A$. To see this, write

$$V = \bigcup_{v \in V} \{w : v \preceq w\}$$

and observe that for $v \in A^m$

$$\Phi(\{w : v \preceq w\}) = \Phi(\{a : v_0 \preceq a\}) \cdots \Phi(\{a : v_{m-1} \preceq a\}).$$

Since $\{a : v_k \preceq a\}$ is increasing, the claim follows.

2.4 Two Dimensions

In dimension 2, words are blocks of letters. Some sets of such blocks can be glued together to form a larger block, but for other sets this may be impossible. The set of all finite 2-dimensional blocks A^* does not possess the nice semi-group structure as it did in dimension 1, so in order to define a two-dimensional substitution, we have to say what we mean by a homomorphism. By a homomorphism on A^*, we mean a map that respects all *valid* concatenations, i.e., concatenations of blocks in such a way that the result is a block again. With this in mind, the generalization of ordinary, random and set-valued substitutions is straightforward.

3 Main Results

3.1 Bounds on the Critical Value of Fractal Percolation

In Section 4, we construct for each integer $k \geq 0$ and $M \geq 2$:

1. a partially ordered finite alphabet $A_{k,M}$ with unique maximum $\bar{a}_{k,M}$ and unique minimum $\underline{a}_{k,M}$,
2. two-dimensional increasing set-valued substitutions $\Phi_{k,M}$ and $\Psi_{k,M}$ on $\mathcal{A}^*_{k,M}$ with base M,
3. an increasing set $J_{k,M} \subset A_{k,M}$.

Define critical values $p_c(\Phi_{k,M})$, $k \geq 0$, $M \geq 2$, by

$$p_c(\Phi_{k,M}) = \inf\{0 \leq p \leq 1 : \mathbb{P}_p\big(\sigma^n(\bar{a}_{k,M}) \in \Phi^n_{k,M}(J_{k,M}) \text{ for all } n\big) > 0\},$$

where σ denotes 2-dimensional fractal percolation on $A^*_{k,M}$ with parameters M and p. We define $p_c(\Psi_{k,M})$ analogously. These critical values provide upper and lower bounds on the critical value of fractal percolation.

Theorem 3.1. *Let $p_c(M)$ be the critical value of 2-dimensional fractal percolation with base M. Then*

1. $p_c(\Phi_{k,M}) \leq p_c(M) \leq p_c(\Psi_{k,M})$ *for all* $k \in \mathbb{N}$,
2. $(p_c(\Phi_{k,M}))_{k\in\mathbb{N}}$ *increases monotonically and* $(p_c(\Psi_{k,M}))_{k\in\mathbb{N}}$ *decreases monotonically,*
3. $\lim_{k\to\infty} p_c(\Phi_{k,M}) = p_c(M)$.

We conjecture (but are not able to prove) that also $\lim_{k\to\infty} p_c(\Psi_{k,M}) = p_c(M)$, whenever $M \geq 3$.

3.2 Bernoulli Random Substitutions and Set-valued Substitutions

Let A be a finite partially ordered alphabet and $M \geq 2$ an integer. Consider a base M Bernoulli random substitution σ with transition matrix P. Motivated by Theorem 3.1, in this section we study probabilities $\mathbb{P}_P\big(\sigma^n(i) \in \Phi^n(J) \text{ for all } n\big)$ for $i \in A$, increasing $J \subseteq A$ and increasing set-valued substitutions Φ. Actually, it turns out that in the context of Theorem 3.1 it suffices to look at $\mathbb{P}_P\big(\sigma^n(i) \in \Phi^n(J)\big)$.

Lemma 3.1. *Let $\prec$ be a partial ordering of A and Φ an increasing (w.r.t. $\prec$) set-valued substitution. If*

$$\mathbb{P}_P\big(\sigma(i) \in \Phi(J)\big) = 0 \qquad \text{for all increasing } J \subseteq A \text{ and } i \notin J,$$

then $(\mathbb{P}_P(\sigma^n(i) \in \Phi^n(J)))_{n\in\mathbb{N}}$ decreases monotonically to $\mathbb{P}_P(\sigma^n(i) \in \Phi^n(J)$ for all $n)$ for all $i \in A$ and increasing $J \subseteq A$.

Remark 3. If we take P to be the transition matrix associated with fractal percolation, then the condition in the lemma above is met by any increasing set-valued substitution Φ.

Write

$$\Pi^n_{ij}(P) = \mathbb{P}_P(\sigma^n(i) \in \Phi^n(j)) \qquad i, j \in A,\ n \in \mathbb{N},$$

so that $\mathbb{P}_P\big(\sigma^n(i) \in \Phi^n(J)\big) = \sum_{j\in J} \Pi^n_{ij}(P)$. Theorem 3.1 would not be very useful, if the probabilities $\Pi^n_{ij}(P)$ were just as intractable as the percolation probability. Fortunately, the matrices $\Pi^n(P) = (\Pi^n_{ij}(P)_{ij})_{i,j\in A}$ satisfy a nice recursion formula.

Lemma 3.2. *For all $n \in \mathbb{N}$,*

$$\Pi^{n+1}(P) = \Pi^1(P\Pi^n(P)).$$

Remark 4. This recursion was obtained by Dekking and Meester (1990, Proposition 3.3) for the case that $A = \{0, 1\}$ and $P = P(p)$ is the transition matrix associated with fractal percolation.

If we denote the recursion function $\Pi^1(PX)$ by $F_P(X) = F_{P,\Phi}(X)$, then $F^n_P(I) = \Pi^n(P)$, where F^n_P denotes the n-fold iterate of F_P and I the identity matrix indexed by $A \times A$. Whenever we write $F_p(X)$ for $0 \leq p \leq 1$, we mean $F_{P(p)}(X)$, where $P(p)$ denotes the transition matrix associated with fractal percolation.

Consider a partial ordering $\prec$ of A. For Markov matrices $X = (X_{ij})_{i,j\in A}$ and $Y = (Y_{ij})_{i,j\in A}$ we write $X \preceq Y$ if

$$\sum_{j\in J} X_{ij} \preceq \sum_{j\in J} Y_{ij},$$

for all $i \in A$ and for all increasing sets $J \subseteq A$.

Lemma 3.3. *Let X and Y be Markov matrices indexed by $A \times A$. If Φ is increasing w.r.t. a partial ordering $\prec$ and $X \preceq Y$, then $F_P(X) \preceq F_P(Y)$.*

3.3 Upper Bounds

An almost direct consequence of Lemmas 3.1 and 3.3 is the following theorem.

Theorem 3.2. *Assume that Φ is increasing w.r.t. a partial ordering Φ and that $\Pi^1(P) \preceq I$. Let $i \in A$ and $J \subseteq A$ be an increasing set. If we can find a Markov matrix $Y = (Y_{ij})_{i,j\in A}$ such that*

1. $Y \preceq I$*,*

2. $Y \preceq F_P(Y)$*,*

3. $\sum_{j \in J} Y_{ij} > 0$*,*

then $\mathbb{P}_P(\sigma^n(i) \in \Phi^n(J)$ *for all* $n) > 0$*.*

Remark 5. The assumption in the above theorem that $\Pi^1(P) \preceq I$ is equivalent to

$$\mathbb{P}_P\big(\sigma(i) \in \Phi(J)\big) = 0 \qquad \text{for all increasing } J \subseteq A \text{ and } i \notin J,$$

a condition that appeared in Lemma 3.1.

The theorem above gives us a recipe to find an upper bound for the value of $p_c(\Psi_{k,M})$, and hence an upper bound for $p_c(M)$. We fix $0 \leq p \leq 1$ and search among matrices Y with $Y \preceq I$ and $\sum_{j \in J_{k,M}} Y_{\bar{a},j} > 0$ for one that satisfies $Y \preceq F_{p,\Psi_{k,M}}(Y)$. As soon as we have found such a matrix, we may conclude that $p_c(\Psi_{k,M}) < p$.

3.4 Lower Bounds

The following theorem is, at least in spirit, similar to Proposition 2 of White (2001).

Theorem 3.3. *Let* $k \geq 0$*,* $M \geq 2$ *be integers, let* $0 < p < 1$*, write* $\bar{a} = \bar{a}_{k,M}$*,* $J = J_{k,M}$*,* $\Phi = \Phi_{k,M}$ *and let* $F_p = F_{p,\Phi}$ *denote the recursion function associated with fractal percolation. If*

$$\sum_{j \in J} \left(F_p^n(I)\right)_{\bar{a},j} < \left(\frac{2M^k}{p}\left(4M - 3 + \frac{8M-12}{1-p}\right)\right)^{-2M} \qquad \textit{for some } n \geq 0,$$

then $\mathbb{P}_p(\sigma^n(\bar{a}) \in \Phi^n(J)$ *for all* $n) = 0$*.*

From this theorem we obtain lower bounds for $p_c(\Phi_{k,M})$ and hence for $p_c(M)$ as follows. We fix $0 \leq p \leq 1$ and compute $F_p^n(I)$ for, say, $n = 100$. If this value happens to be smaller than $\left(\frac{2M^k}{p}\left(4M - 3 + \frac{8M-12}{1-p}\right)\right)^{-2M}$, then we may conclude that $p_c(\Phi_{k,M}) > p$.

3.5 Numerical Results

The following numerical bounds were obtained by calculating the recursion functions $F_{p,\Phi_{k,M}}$ and $F_{p,\Psi_{k,M}}$ and applying Theorems 3.2 and 3.3. Since even for small k and M, the entries of $F_{p,\Phi_{k,M}}(X)$ and $F_{p,\Psi_{k,M}}(X)$ are huge polynomials in the entries of X, the work was done by a computer.

Theorem 3.4.

$$\begin{array}{lllllll} p_c(\Phi_{k=0,M=2}) & \geq & 0.784 & \qquad & p_c(\Phi_{k=1,M=2}) & \geq & 0.858 \\ p_c(\Phi_{k=0,M=3}) & \geq & 0.715 & \qquad & p_c(\Psi_{k=0,M=3}) & \leq & 0.958 \end{array}$$

Remark 6. From the definition of $\Psi_{k,M}$ it will be clear that $p_c(\Psi_{k,M=2}) = 1$ for all $k \geq 0$.

4 Constructions

In this section, we will construct a partially ordered alphabet $A_{k,M}$, a set $J_{k,M} \subset A_{k,M}$ and two-dimensional substitutions $\Phi_{k,M}$ and $\Psi_{k,M}$, for all integers $k \geq 0$ and $M \geq 2$. We will fix integers $k \geq 0$ and $M \geq 2$, and we will suppress dependence on these parameters in our notation. So A will denote $A_{k,M}$ for example.

4.1 The Partially Ordered Alphabet A

Let

$$L_0 = \{(1, \tfrac{1}{2}), (\tfrac{1}{2}, 1), (0, \tfrac{1}{2}), (\tfrac{1}{2}, 0)\}$$

be the set of midpoints of the top, bottom, left and right side of the unit square. Fix $M \geq 2$ and $k \geq 0$ and let

$$L = \partial[0,1]^2 \cap M^{-k}(\mathbb{Z}^2 + L_0),$$

be the set of $4M^k$ points equally distributed over the boundary of the unit square $\partial[0,1]^2$. We define the alphabet A to be the set of all equivalence relations on L. Let $\prec$ be the natural partial ordering on A, i.e., $a \prec b$ if and only if $a \subset b$ $(a, b \in A)$. The maximum element $\bar{a}$ w.r.t. $\prec$ is the relation $A \times A$ and the minimum $\underline{a}$ is $\{(a, a) : a \in A\}$.

4.2 The set of letters J

Let $f, g \geq 1$ be integers, let $B^n(f, g) = \{0, \ldots, fM^n - 1\} \times \{0, \ldots, gM^n - 1\}$ be the $fM^n \times gM^n$ box in $\mathbb{Z}^2$ and let $W^n(f, g) = A^{B^n(f,g)}$ denote the set of $fM^n \times gM^n$ blocks of letters in A. To each word $v \in W^n(f, g)$ we associate a graph $G(v)$ as follows. The set of vertices of the graph will be

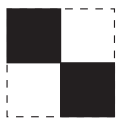
$$L^n(f, g) = L + B^n(f, g).$$

Two points $x, y \in L^n(f, g)$ are connected by an edge if there is $z \in B^n(f, g)$ and $s, t \in L$ such that $x = z + s$, $y = z + t$ and $(s, t) \in v_z$. We say that the graph $G(v)$ percolates if there is a point x on the left side of $L^n(f, g)$ and a point y on the right side, such that x and y are connected in $G(v)$. Define

$$J^n(f, g) = \{w \in W^n(f, g) : G(w) \text{ percolates}\},$$

and define the set J to be $J^0(1, 1)$. In general, if we suppress n and (f, g) in our notation, we assume that $n = 0$ and $(f, g) = (1, 1)$.

Remark 7. There is a slight discrepancy between our definitions of percolation for sets and for words, caused by diagonal connections. For example, the set



percolates, but the graph of the corresponding word

$$\begin{array}{cc} \bar{a} & \underline{a} \\ \underline{a} & \bar{a} \end{array}$$

does not. Consequently, $\mathbb{P}_p(\sigma^n(\bar{a}) \in J^n) < \mathbb{P}_p(K_n \text{ percolates})$ for all $n \geq 1$ and $0 < p < 1$. However, it is not difficult to see that $\mathbb{P}_p(\sigma^n(\bar{a}) \in J^n \text{ for all } n) = \mathbb{P}_p(K_n \text{ percolates for all } n)$.

Example 2. For $k = 0$ and $M = 2$, label the points of L by $x_0 = (1, \frac{1}{2})$, $x_1 = (\frac{1}{2}, 1)$, $x_2 = (0, \frac{1}{2})$ and $x_3 = (\frac{1}{2}, 0)$. Let $v \in W^1$ be defined by

$$\begin{aligned} v_{00} &= \min\{a \in A : (x_0, x_1) \in a\} \\ v_{01} &= \min\{a \in A : (x_2, x_3) \in a\} \\ v_{10} &= \min\{a \in A : (x_0, x_2) \in a\} \\ v_{11} &= \min\{a \in A\}, \end{aligned}$$

where the minima are taken with respect to $\prec$. Then the graph $G(v)$ of v is given in Figure 2. Note that $G(v)$ percolates, so $v \in J^1$.

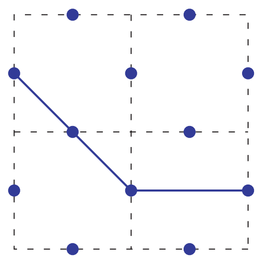


Figure 2: The graph $G(v)$ associated with the word v given in Example 2. The dashed lines represent (part of) the $\mathbb{Z}^2$ lattice.

4.3 The Set-valued Substitutions Φ and Ψ

Consider a graph $G=(V,E)$ with vertex set V and edge set E and let $\mathcal{R}$ be a set of non-empty disjoint subsets of V. We say that $R_1,R_2\in\mathcal{R}$ are connected, if x_1 and x_2 are connected in G for some $x_1\in R_1$ and $x_2\in R_2$. We say that $R_1,R_2\in\mathcal{R}$ are *weakly connected*, if there are $R^1,\dots,R^m\in\mathcal{R}$ such that $R^1=R_1$, $R^m=R_2$ and R^i is connected with R^{i+1} for all $1\le i\le m-1$. Finally, we say that $R_1,R_2\in\mathcal{R}$ are *strongly connected*, if $R_1=R_2$ or there is a connected component C of G such that $|C\cap R_1|\ge\lceil\frac{|R_1|+1}{2}\rceil$ and $|C\cap R_2|\ge\lceil\frac{|R_2|+1}{2}\rceil$. Observe that unlike connectedness, weak and strong connectedness are equivalence relations on $\mathcal{R}$.

For $x\in\mathbb{Z}^2+L$, let $\ell(x)$ be the line segment between two neighboring points in the $M^{-k}\mathbb{Z}^2$ lattice that contains x. Define $R(x)=R_k(x)$ to be the set of M points in $\mathbb{Z}^2+L$ that are contained in the line segment $M\ell(x)$. By default, if G is a graph on L^{n+1}, then weak and strong connectedness will be relations on the set $\mathcal{R}=\{R(x):x\in L^n\}$.

Let the set-valued substitutions Φ and Ψ be defined by

$$\begin{aligned}\Phi(a) &= \{v\in W^1 : R(x)\text{ and }R(y)\text{ are weakly connected in }G(v),\\ &\quad\ \text{if and only if }(x,y)\in a,\ \text{ for all }x,y\in L\},\\ \Psi(a) &= \{v\in W^1 : R(x)\text{ and }R(y)\text{ are strongly connected in }G(v),\\ &\quad\ \text{if and only if }(x,y)\in a,\ \text{ for all }x,y\in L\},\end{aligned}$$

for all $a\in A$.

Example 3. Let $k=0$, $M=3$ and write $L=\{x_0,x_1,x_2,x_3\}$ as in Example 2. At the bottom left of Figure 3, we see the graph $G(v)$ of a word v in which $R(x_0)$, $R(x_1)$ and $R(x_2)$ are weakly connected. Hence $v\in\Phi(a)$, where $a=\inf\{b\in A:(x_0,x_1),(x_1,x_2)\in b\}$ and the infimum is taken with respect to $\prec$. At the bottom right of Figure 3, the graph $G(w)$ of a word w is depicted, in which $R(x_0)$, $R(x_1)$ and $R(x_2)$ are strongly connected. Hence $w\in\Psi(a)$.

5 Proof of Theorem 3.1

5.1 Part 1

Fix integers $k\ge 0$ and $M\ge 2$. Part 1 of Theorem 3.1 follows immediately from the following lemma.

Lemma 5.1. *For all $n\ge 0$,*

$$\Psi^n(J)\subseteq J^n\subseteq\Phi^n(J).$$

An important step in the proof of the above lemma is provided by the following.

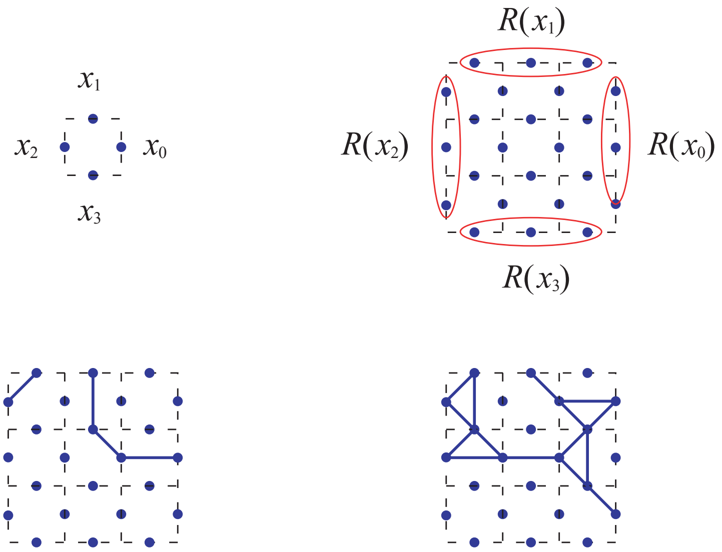


Figure 3: Let $k = 0$ and $M = 3$. Top left: the set $L = \{x_0, x_1, x_2, x_3\}$. Top right: the sets $R(x_i)$, $i = 0, 1, 2, 3$. Bottom left: a graph in which $R(x_0)$, $R(x_1)$ and $R(x_2)$ are weakly connected. Bottom right: a graph in which $R(x_0)$, $R(x_1)$ and $R(x_2)$ are strongly connected.

Lemma 5.2. *Let $n \geq 0$, $x, y \in L^n$, and $v \in W^n$.*

1. *If $w \in \Psi(v)$ and x and y are connected in $G(v)$, then $R(x)$ and $R(y)$ are connected in $G(w)$.*
2. *If $w \in \Phi(v)$ and $R(x)$ and $R(y)$ are connected in $G(w)$, then x and y are connected in $G(v)$.*

Proof. 1. Let $w \in \Psi(v)$ and suppose x and y are connected in $G(v)$. Then there are $z_1, \ldots, z_m \in L^n$ such that $z_1 = x$, $z_m = y$ and (z_i, z_{i+1}) is an edge of the graph $G(v)$, for $1 \leq i \leq m-1$. Since $L^n = \{0, \ldots, M^n - 1\}^2 + L$, we can find

$$r_1, \ldots, r_{m-1} \in \mathbb{Z}^2, \quad s_1, \ldots, s_{m-1} \in L, \quad t_1, \ldots, t_{m-1} \in L,$$

such that

$$z_i = r_i + s_i, \quad z_{i+1} = r_i + t_i, \quad (s_i, t_i) \in v_{r_i}, \qquad 1 \leq i \leq m-1.$$

Fix $0 \leq i \leq m-1$. Since $w \in \Psi(v)$, we have that $w_{Mr_i + B^1} \in \Psi(v_{r_i})$, where $B^1 = \{0, \ldots, M-1\}^2$ and $w_{Mr_i+B^1}$ denotes the $M \times M$ sub-word of w, obtained by restricting w to the indices of $Mr_i + B^1$. So in the graph associated to $w_{Mr_i+B^1}$, the sets $R(s_i)$ and $R(t_i)$ are strongly connected. Consequently, in the graph associated to w, the sets $Mr_i + R(s_i) = R(r_i + s_i)$ and $Mr_i + R(t_i) = R(r_i + t_i)$ are strongly connected. Hence $R(z_i)$ and $R(z_{i+1})$ are strongly connected in $G(w)$. Since strongly connectedness is a transitive relation, it follows that $R(x)$ and $R(y)$ are strongly connected in $G(w)$.

2. Let $w \in \Phi(v)$ and suppose $R(x)$ and $R(y)$ are connected in $G(w)$. Then we can find

$$r_1, \ldots, r_{m-1} \in \mathbb{Z}^2, \quad s_1, \ldots, s_{m-1} \in L, \quad t_1, \ldots, t_{m-1} \in L,$$

such that

$$x = r_1 + s_1, \quad y = r_m + t_m, \quad r_i + t_i = r_{i+1} + s_i,$$

and such that $R(s_i)$ and $R(t_i)$ are connected in $G(w_{Mr_i+B^1})$ for $0 \leq i \leq m-1$. Since $w_{Mr_i+B^1} \in \Phi(v_{r_i})$, we have that $(s_i, t_i) \in v_{r_i}$ and hence, that $r_i + s_i$ and $r_i + t_i$ are connected in $G(v)$. Hence, x and y are connected in $G(v)$.

□

Proof of Lemma 5.1. Both inclusions are proved by induction. Concerning the first inclusion, let $w \in \Psi^{n+1}(J)$. Then there is a unique word $v \in \Psi^n(J)$ such that $w \in \Psi(v)$ and by the induction hypothesis, $v \in J^n$. Let $x, y \in L^n$ be connected in $G(v)$, where x is a point on the left side of L^n and y a point on the right side. By Lemma 5.2, $R(x)$ and $R(y)$ are connected in $G(w)$. Since $R(x)$ is contained in the left side of $[0, M^{n+1}]^2$ and $R(y)$ in the right side, the word w is contained in J^{n+1}.

Concerning the second inclusion, let $w \in J^{n+1}$. Then we can find x on the left side of L^n and y on the right side, such that $R(x)$ and $R(y)$ are connected in $G(w)$. Let $v \in W^n$ be the unique word for which $w \in \Phi(v)$. By Lemma 5.2, x and y are connected in $G(v)$ and hence $v \in J^n$. By the induction hypothesis, $v \in \Phi^n(J)$ and so $w \in \Phi^{n+1}(J)$. □

5.2 Part 2

In this section, we will only suppress dependence on M in our notation. So, e.g., Φ_k denotes $\Phi_{k,M}$.

The key ingredient in the proof that the sequences $(p_c(\Phi_k))_{k\in\mathbb{N}}$ and $(p_c(\Psi_k))_{k\in\mathbb{N}}$ are monotone, is the following lemma.

Let a homomorphism ρ from $\mathcal{A}_{k+1}^*$ to $\mathcal{A}_k^*$, respecting unions and concatenations, be defined by

$$\begin{aligned}\rho(a) \quad &= \quad \{w \in W_k^1 : Mx \text{ and } My \text{ are connected in } G(w), \text{ if and only if} \\ &\qquad (x,y) \in a, \text{ for all } x, y \in L_{k+1}\},\end{aligned}$$

for all $a \in A_{k+1}$.

Lemma 5.3. *For increasing $V \subseteq W_{k+1}^n$,*

$$\rho(\Phi_{k+1}(V)) \subseteq \Phi_k(\rho(V)) \quad \text{and} \quad \Psi_k(\rho(V)) \subseteq \rho(\Psi_{k+1}(V)).$$

Proof. For $w \in W_k^2$, consider the graph $G(w)$. Let $r_1(w)$ and $r_2(w)$ denote the weak connectedness relations on, respectively, $\{R_k(Mx) : x \in L_{k+1}\}$ and $\{R_k(x) : x \in L_k^1\}$. Then for $a \in A_{k+1}$,

$$\rho(\Phi_{k+1}(J)) = \{w \in W_k^2 : (x,y) \in a \Leftrightarrow (R_k(Mx), R_k(My)) \in r_1(w) \text{ for all } x, y \in L_{k+1}\},$$

and hence

$$\begin{aligned}\rho(\Phi_{k+1}(J)) \quad &= \quad \{w \in W_k^2 : \text{ there is } a \in J \text{ such that} \\ &\qquad (x,y) \in a \Rightarrow (R_k(Mx), R_k(My)) \in r_1(w) \text{ for all } x, y \in L_{k+1}\},\end{aligned}$$

for all increasing $J \subseteq A_{k+1}$. Similarly,

$$\begin{aligned}\Phi_k(\rho(J)) \quad &= \quad \{w \in W_k^2 : \text{ there is } a \in J \text{ such that} \\ &\qquad (x,y) \in a \Rightarrow (R_k(Mx), R_k(My)) \in r_2(w) \text{ for all } x, y \in L_{k+1}\}.\end{aligned}$$

Since $\{R_k(Mx) : x \in L_{k+1}\} \subseteq \{R_k(x) : x \in L_k^1\}$, it follows that $r_1(w) \subseteq r_2(w)$. Thus, $\rho(\Phi_{k+1}(J)) \subseteq \Phi_k(\rho(J))$ for all increasing $J \subseteq A_{k+1}$. The first inclusion of the lemma now follows from an argument similar to Remark 2. The second inclusion of the lemma is obtained analogously. □

Proof of part 2. If τ is the Bernoulli random substitution on A_k^* with base M and transition matrix equal to the identity matrix, then

$$\mathbb{P}_p(\sigma^n(\bar{a}_{k+1}) \in \Phi_{k+1}^n(J_{k+1})) \quad = \quad \mathbb{P}_p(\tau(\sigma^n(\bar{a}_k)) \in \rho(\Phi_{k+1}^n(J_{k+1}))).$$

Applying Lemma 5.3 n times, we obtain

$$\mathbb{P}_p(\sigma^n(\bar{a}_{k+1}) \in \Phi_{k+1}^n(J_{k+1})) \quad \leq \quad \mathbb{P}_p(\tau(\sigma^n(\bar{a}_k)) \in \Phi_k^n(\rho(J_{k+1}))).$$

By Lemma 5.1, $\rho(J_{k+1}) = J_k^1 \subseteq \Phi_k(J_k)$ and hence

$$\begin{aligned}\mathbb{P}_p(\sigma^n(\bar{a}_{k+1}) \in \Phi_{k+1}^n(J_{k+1})) \quad &\leq \quad \mathbb{P}_p(\tau(\sigma^n(\bar{a}_k)) \in \Phi_k^{n+1}(J_k)) \\ &= \quad \mathbb{P}_p(\sigma^n(\bar{a}_k) \in \Phi_k^n(J_k)).\end{aligned}$$

It follows from Lemma 3.1 that $p_c(\Phi_k) \leq p_c(\Phi_{k+1})$. Monotonicity of $(p_c(\Psi_k))_{k\in\mathbb{N}}$ is proved similarly. $\square$

5.3 Part 3

Lemma 5.4. *For $n \geq 0$ and $k \geq n$,*

$$\mathbb{P}_p(\sigma^n(\bar{a}_k) \in \Phi_k^n(J_k)) \quad = \quad \mathbb{P}_p(\sigma^n(\bar{a}_0) \in J_0^n).$$

Proof. Since $J_k^n \subseteq \Phi_k^n(J_k)$ by Lemma 5.1, we have that

$$\begin{aligned}\mathbb{P}_p(\sigma^n(\bar{a}_0) \in J_0^n) \quad &= \quad \mathbb{P}_p(\sigma^n(\bar{a}_k) \in J_k^n) \\ &\leq \quad \mathbb{P}_p(\sigma^n(\bar{a}_k) \in \Phi_k^n(J_k)).\end{aligned}$$

For the reversed inequality, we use that for $n \geq 1$,

$$\mathbb{P}_p(\sigma^n(\bar{a}_k) \in \Phi_k^n(J_k)) \quad \leq \quad \mathbb{P}_p(\tau(\sigma^n(\bar{a}_{k-1})) \in \Phi_{k-1}^n(\rho(J_k))),$$

an intermediate result obtained in the proof of part 2 of Theorem 3.1. Since $\rho(J_k) = J_{k-1}^1$, we have

$$\mathbb{P}_p(\sigma^n(\bar{a}_k) \in \Phi_k^n(J_k)) \quad \leq \quad \mathbb{P}_p(\sigma^n(\bar{a}_{k-1}) \in \Phi_{k-1}^{n-1}(J_{k-1}^1)).$$

Repeating the previous two steps n times, we obtain

$$\begin{aligned}\mathbb{P}_p(\sigma^n(\bar{a}_k) \in \Phi_k^n(J_k)) \quad &\leq \quad \mathbb{P}_p(\sigma^n(\bar{a}_{k-n}) \in J_{k-n}^n) \\ &= \quad \mathbb{P}_p(\sigma^n(\bar{a}_0) \in J_0^n).\end{aligned}$$

$\square$

Fix integers $k \geq 0$ and $M \geq 2$. For integers $f, g \geq 1$ and $a \in A$, let $a(f,g)$ denote the $f \times g$ block consisting of solely a's. Define

$$p_c(\Phi, 1, 2) = \inf\{0 \leq p \leq 1 : \mathbb{P}_p(\sigma^n(\bar{a}(1,2)) \in \Phi^n(J(1,2)) \text{ for all } n) > 0\}.$$

The following lemma is similar in spirit to Theorem 5.4 of Dekking and Meester (1990).

Lemma 5.5. *For $p \geq p_c(\Phi, 1, 2)$,*

$$\mathbb{P}_p(\sigma^n(\bar{a}(1,2)) \in \Phi^n(J(1,2)) \textit{ for all } n) \geq \left(4M - 3 + \frac{8M-12}{1-p}\right)^{-2}.$$

Proof. Define sets $\boxminus, \text{◫}, \text{◳}, \text{◲}, \text{◱}, \text{◰} \subset A$ as follows. Let L_l, L_r, L_t and L_b denote, respectively, the points on the left, right, top and bottom side of L. Define

$$\begin{aligned} \boxminus &= \{a \in A: \text{ there are } x \in L_l \text{ and } y \in L_r \text{ such that } (x,y) \in a\} \\ \text{◫} &= \{a \in A: \text{ there are } x \in L_t \text{ and } y \in L_b \text{ such that } (x,y) \in a\}. \end{aligned}$$

Let

$$\text{◱} = \{a \in A: \text{ there are } x \in L_l \text{ and } y \in L_b \text{ such that } (x,y) \in a, \text{ but } a \notin \boxminus \text{ or } \text{◫}\}$$

and define ◰, ◳ and ◲ likewise. Finally, define $\text{⌸} \subset W(1,2)$ by

$$\text{⌸} = \{w \in W(1,2): \text{ there are } x \in L_l(1,2) \text{ and } y \in L_r(1,2) \text{ such that } x \text{ and } y \text{ are connected in } G(w)\},$$

where $L_l(1,2)$ and $L_r(1,2)$ denote, respectively, the points on the left and right side of $L(1,2)$.

Consider a word $v \in \Phi(\text{⌸})$. It follows from the definition of Φ that at least one of the following should hold for the letters v_{0j}, $0 \le j \le 2M-1$ in the first column of v:

1. $\genfrac{}{}{0pt}{}{v_{0,j+1}}{v_{0j}} \in \text{⌸}$ for some $0 \le j \le 2M-2$,
2. $v_{0j} \in \text{◫}$ for some $1 \le j \le 2M-2$,
3. $v_{0j} \in \text{◱}$ for $j = 0$ or $j = M$,
4. $v_{0j} \in \text{◰}$ for $j = M-1$ or $j = 2M-1$.

A similar requirement holds for the last column of v. Hence,

$$\begin{aligned} \mathbb{P}_p(\sigma^n(\bar{a}(1,2)) \in \Phi^n(\text{⌸})\ \forall n) \le \Big(&(2M-1)p^2 \mathbb{P}_p(\sigma^n(\bar{a}(1,2)) \in \Phi^n(\text{⌸})\ \forall n) \\ &+ (2M-2)p\mathbb{P}_p(\sigma^n(\bar{a}) \in \Phi^n(\text{◫})\ \forall n) + 4p\mathbb{P}_p(\sigma^n(\bar{a}) \in \Phi^n(\text{◱})\ \forall n)\Big)^2, \end{aligned}$$

where we used that the probabilities $\mathbb{P}_p(\sigma^n(\bar{a}) \in \Phi^n(H)\ \forall n)$ are equal for $H = \text{◱}, \text{◰}, \text{◳}, \text{◲}$. Hence

$$\mathbb{P}_p(\sigma^n(\bar{a}(1,2)) \in \Phi^n(\text{⌸})\ \forall n) \le \big((4M-3)\mathbb{P}_p(\sigma^n(\bar{a}(1,2)) \in \Phi^n(\text{⌸})\ \forall n) + 4p\mathbb{P}_p(\sigma^n(\bar{a}) \in \Phi^n(\text{◱})\ \forall n)\big)^2, \quad (1)$$

where we used that $\mathbb{P}_p(\sigma^n(\bar{a}) \in \Phi^n(\text{◫})\ \forall n) \le \mathbb{P}_p(\sigma^n(\bar{a}(1,2)) \in \Phi^n(\text{⌸})\ \forall n)$.

Consider a word $v \in \Phi(\text{◱})$. Then at least one of the following conditions should hold for the letters in the first column of v:

1. $\genfrac{}{}{0pt}{}{v_{0,j+1}}{v_{0j}} \in \text{⌸}$ for some $0 \le j \le M-2$,
2. $v_{0j} \in \text{◫}$ for some $1 \le j \le M-2$,
3. $v_{00} \in \text{◱}$.

One might wonder whether there is a $v \in \Phi(\text{◱})$ with $v_{0,M-1} \in \text{◰}$ such that none of the conditions above apply. To see that this is not the case, observe that for such v there are $x \in L_r$, $y \in L_t$ and $z \in L_b$ such that $R(x)$, $R(y)$ and $R(z)$ are weakly connected in $G(v)$. But then $v \in \Phi(\text{◫})$ and thus $\text{◫} \cap \text{◱} \ne \emptyset$, which contradicts the definition of ◱.

Hence, we have that

$$
\begin{aligned}
\mathbb{P}_p(\sigma^n(\bar{a} \in \Phi^n(\boxbslash)\ \forall n) &\leq (M-1)p^2\mathbb{P}_p(\sigma^n(\bar{a}(1,2)) \in \Phi^n(\boxminus)\ \forall n) + (M-2)p\mathbb{P}_p(\sigma^n(\bar{a}) \in \Phi^n(\boxbar)\ \forall n) \\
&\quad + p\mathbb{P}_p(\sigma^n(\bar{a}) \in \Phi^n(\boxbslash)\ \forall n) \\
&\leq (2M-3)\mathbb{P}_p(\sigma^n(\bar{a}(1,2)) \in \Phi^n(\boxminus)\ \forall n) + p\mathbb{P}_p(\sigma^n(\bar{a}) \in \Phi^n(\boxbslash)\ \forall n),
\end{aligned}
$$

and so

$$
\mathbb{P}_p(\sigma^n(\bar{a} \in \Phi^n(\boxbslash)\ \forall n) \leq \frac{2M-3}{1-p}\mathbb{P}_p(\sigma^n(\bar{a}(1,2)) \in \Phi^n(\boxminus)\ \forall n). \tag{2}
$$

Combining inequalities 1 and 2, we obtain

$$
\mathbb{P}_p(\sigma^n(\bar{a}(1,2)) \in \Phi^n(\boxminus)\ \forall n) \leq \left(4M - 3 + \frac{8M-12}{1-p}\right)^2 \mathbb{P}_p(\sigma^n(\bar{a}(1,2)) \in \Phi^n(\boxminus)\ \forall n)^2.
$$

Hence,

$$
\mathbb{P}_p(\sigma^n(\bar{a}(1,2)) \in \Phi^n(J(1,2))\ \forall n) \geq \left(4M - 3 + \frac{8M-12}{1-p}\right)^{-2},
$$

whenever $\mathbb{P}_p(\sigma^n(\bar{a}(1,2)) \in \Phi^n(J(1,2))\ \forall n) > 0$. It follows from standard arguments (conform e.g. Chayes et al. (1988)) that $\mathbb{P}_p(\sigma^n(\bar{a}(1,2)) \in \Phi^n(J(1,2))\ \forall n)$ is a non-decreasing right-continuous function of p, and hence we have obtained the statement of the lemma. $\square$

Proof of part 3. By a trivial extension of Lemma 5.4, we have that for $n \geq 0$ and $k \geq n$,

$$
\mathbb{P}_p(\sigma^n(\bar{a}_k(1,2)) \in \Phi_k^n(J_k(1,2))) \quad = \quad \mathbb{P}_p(\sigma^n(\bar{a}_0(1,2)) \in J_0^n(1,2)).
$$

Fix any $0 \leq p < p_c(M,1,2)$, where

$$
p_c(M,1,2) = \inf\{0 \leq p \leq 1 : \mathbb{P}_p(\sigma^n(\bar{a}_0(1,2)) \in J_0^n(1,2) \text{ for all } n) > 0\}.
$$

Since $(\mathbb{P}_p(\sigma^n(\bar{a}_0(1,2))) \in J_0^n(1,2)))_{n\in\mathbb{N}}$ converges to 0, we can choose n such that

$$
\mathbb{P}_p(\sigma^n(\bar{a}_0(1,2))) \in J_0^n(1,2)) < \left(4M - 3 + \frac{8M-12}{1-p}\right)^{-2}.
$$

Choosing $k \geq n$, we obtain

$$
\mathbb{P}_p(\sigma^n(\bar{a}_k(1,2)) \in \Phi_k^n(J_k(1,2))) < \left(4M - 3 + \frac{8M-12}{1-p}\right)^{-2},
$$

and it follows from Lemma 5.5 that $p < p_c(\Phi_k, 1, 2)$. Since obviously $p_c(\Phi_k,1,2) \leq p_c(\Phi_k)$ and since $p_c(\Phi_k) \leq p_c(M)$ by Theorem 3.1, part 1, we have $p < p_c(\Phi_k) \leq p_c(M)$. Using a qualitative analogue to the classical RSW theorem from ordinary percolation, Dekking and Meester (1990, Lemma 5.1) proved that $p_c(M,1,2) = p_c(M)$. Hence, for all $0 \leq p < p_c(M)$, we can find k such that $p < p_c(\Phi_k) \leq p_c(M)$. $\square$

6 Proofs of Lemmas 3.1, 3.2 and 3.3

For notational convenience, we will assume in the Lemmas 3.2, 3.1 and 3.3 that the substitutions are 1-dimensional. Generalizing the proofs to higher dimensions is trivial.

Proof of Lemma 3.1. Since $\mathbb{P}_P\big(\sigma(i) \in \Phi(J)\big) = 0$ for all increasing $J \subseteq A$ and $i \notin J$, we have that $\mathbb{P}_P(\sigma(i) \in \Phi(\{j : j \preceq i\})) = 1$ for all $i \in A$, and hence $\mathbb{P}_P(\sigma(w) \in \Phi(\{v : v \preceq w\})) = 1$ for all $w \in W^n$.

Fix $i \in A$, $J \subseteq A$ increasing and suppose that $w \in W^n$ is not contained in $\Phi^n(J)$. Since $w \in \Phi^n(J^c)$ and $\Phi^n(J^c)$ is a decreasing set, it follows that $\{v : v \preceq w\} \subseteq \Phi^n(J^c)$ and thus $\Phi(\{v : v \preceq w\}) \subseteq \Phi^{n+1}(J^c)$. From this we have that

$$\mathbb{P}_P(\sigma(w) \in \Phi^{n+1}(J^c)) \geq \mathbb{P}_P(\sigma(w) \in \Phi(\{v : v \preceq w\})) = 1,$$

and hence $\mathbb{P}_P(\sigma(w) \in \Phi^{n+1}(J)) = 0$. Taking $w = \sigma^n(i)$, it follows that $\{\sigma^{n+1}(i) \in \Phi^{n+1}(J)\} \subseteq \{\sigma^n(i) \in \Phi^n(J)\}$ a.s., which implies the statement of the lemma. □

Proof of Lemma 3.2. For $i, j \in A$ and $n \in \mathbb{N}$,

$$\begin{aligned}
\Pi_{ij}^{n+1}(P) &= \mathbb{P}_P\big(\sigma^{n+1}(i) \in \Phi^{n+1}(j)\big) \\
&= \sum_{v \in \Phi(j)} \prod_{m=0}^{M-1} \big(\sum_{k \in A} P_{ik}\, \Pi_{k,v_m}^n(P)\big) \\
&= \sum_{v \in \Phi(j)} \prod_{m=0}^{M-1} \big(P\Pi^n(P)\big)_{i,v_m} \\
&= \mathbb{P}_{P\Pi^n(P)}\big(\sigma(i) \in \Phi(j)\big) \\
&= \big(\Pi^1(P\Pi^n(P))\big)_{ij}.
\end{aligned}$$

□

Proof of Lemma 3.3. Fix $i \in A$. Observe that if $X \preceq Y$, we also have that $PX \preceq PY$. Since Φ is an increasing set-valued substitution, it suffices to prove that $\mu_X = \mathbb{P}_X(\sigma(i) \in \cdot)$ is stochastically dominated by $\mu_Y = \mathbb{P}_Y(\sigma(i) \in \cdot)$, whenever $X \preceq Y$. Fix $X \preceq Y$, $0 \leq k \leq M-1$ and define $\mu_X^k = \mathbb{P}_X((\sigma(i))_k \in \cdot)$. Then by definition,

$$\int \mathbb{1}_J \, \mathrm{d}\mu_X^k \leq \int \mathbb{1}_J \, \mathrm{d}\mu_Y^k \qquad \text{for all increasing } J \subseteq A.$$

In fact, since every increasing function from A to $\mathbb{R}$ can be written as a positive linear combination of indicator functions of increasing sets, we have that μ_X^k is stochastically dominated by μ_Y^k. From this, it easily follows that μ_X is stochastically dominated by μ_Y. □

7 Proofs of Theorems 3.2 and 3.3

Proof of Theorem 3.2. We start by proving that $Y \preceq F_P^n(I)$ for all $n \in \mathbb{N}$. Indeed, for $n = 0$ this follows from assumption 1 of the theorem. Suppose that $Y \preceq F_P^n(I)$. Then $F_P(Y) \preceq F_P^{n+1}(I)$ by Lemma 3.3 and hence $Y \preceq F_P^{n+1}(I)$ by assumption 2, establishing the claim. So we have for all $n \in \mathbb{N}$,

$$\mathbb{P}(\sigma^n(i) \in \Phi^n(J)) = \sum_{j \in J} (F_P^n(I))_{ij} \geq \sum_{j \in J} Y_{ij}.$$

Since $(\mathbb{P}_P(\sigma^n(i) \in \Phi^n(J)))_{n\in\mathbb{N}}$ decreases monotonically to $\mathbb{P}_P(\sigma^n(i) \in \Phi^n(J) \text{ for all } n)$ by Lemma 3.1, it follows that

$$\mathbb{P}_P(\sigma^n(i) \in \Phi^n(J) \text{ for all } n) \geq \sum_{j\in J} Y_{ij} > 0,$$

by assumption 3. □

Proof of Theorem 3.3. Fix $p \geq p_c(\Phi, 1, 2)$. By Lemma 5.5 we have that

$$\mathbb{P}_p(\sigma^n(\bar{a}(1,2)) \in \Phi^n(J(1,2)) \text{ for all } n) \geq \left(4M - 3 + \frac{8M-12}{1-p}\right)^{-2}.$$

For x on the left side of $L(1,2)$ and y on the right side, define

$$J(1,2;x,y) = \{w \in J(1,2) : \ x \text{ and } y \text{ are connected in } G(w)\}.$$

Then of course

$$\begin{aligned}\left(4M - 3 + \frac{8M-12}{1-p}\right)^{-2} &\leq \mathbb{P}_p(\sigma^n(\bar{a}(1,2)) \in \Phi^n(J(1,2)) \text{ for all } n)\\ &\leq \sum_{x,y} \mathbb{P}_p(\sigma^n(\bar{a}(1,2)) \in \Phi^n(J(1,2;x,y)) \text{ for all } n),\end{aligned}$$

where the sum extends over all x on the left side of $L(1,2)$ and all y on the right side. Since both the left and the right side of $L(1,2)$ contain $2M^k$ points, we may fix x and y such that

$$\mathbb{P}_p(\sigma^n(\bar{a}(1,2)) \in \Phi^n(J(1,2;x,y)) \text{ for all } n) \geq \left(2M^k\left(4M - 3 + \frac{8M-12}{1-p}\right)\right)^{-2}.$$

For $z = (z_1, z_2) \in L_k$, let $\hat{z} = (1 - z_1, z_2)$ denote the reflection of z in the vertical line $\{\frac{1}{2}\} \times \mathbb{R}$. By symmetry, we have that

$$\mathbb{P}_p(\sigma^n(\bar{a}(1,2)) \in \Phi^n(J(1,2;x,y)) \text{ for all } n) = \mathbb{P}_p(\sigma^n(\bar{a}(1,2)) \in \Phi^n(J(1,2;\hat{x},\hat{y})) \text{ for all } n).$$

Define a set $V \subset W^1$ by

$$V = \{v \in W^1 : \ \begin{matrix} v_{i1} \\ v_{i0}\end{matrix} \in J(1,2;x,y) \text{ if } i \text{ is even}, \ \begin{matrix} v_{i1} \\ v_{i0}\end{matrix} \in J(1,2;\hat{x},\hat{y}) \text{ if } i \text{ is odd}, 0 \leq i \leq M-1\}.$$

Observe that $V \subseteq J^1$, i.e., every word in V percolates. Hence,

$$\begin{aligned}\mathbb{P}_p(\sigma^n(\bar{a}) \in \Phi^n(J) \text{ for all } n) &\geq \mathbb{P}_p(\sigma^{n+1}(\bar{a}) \in \Phi^n(V) \text{ for all } n)\\ &\geq p^{2M}\mathbb{P}_p(\sigma^n(\bar{a}(1,2)) \in \Phi^n(J(1,2;x,y)) \text{ for all } n)^M\\ &\geq \left(\frac{2M^k}{p}\left(4M - 3 + \frac{8M-12}{1-p}\right)\right)^{-2M}.\end{aligned}$$

This result implies that $p_c(\Phi) = p_c(\Phi, 1, 2)$. Hence for all $p \geq p_c(\Phi)$ and $n \geq 0$ we have

$$\begin{aligned}\sum_{j\in J}\left(F_p^n(I)\right)_{\bar{a},j} &= \mathbb{P}_p(\sigma^n(\bar{a}) \in \Phi^n(J))\\ &\geq \mathbb{P}_p(\sigma^n(\bar{a}) \in \Phi^n(J) \text{ for all } n)\\ &\geq \left(\frac{2M^k}{p}\left(4M - 3 + \frac{8M-12}{1-p}\right)\right)^{-2M}.\end{aligned}$$

□

References

Chayes, J. T., L. Chayes, and R. Durrett. 1988. *Connectivity properties of Mandelbrot's percolation process*, Probab. Theory Related Fields **77**, 307–324.

Dekking, F. M. and R. W. J. Meester. 1990. *On the structure of Mandelbrot's percolation process and other random Cantor sets*, J. Statist. Phys. **58**, 1109–1126.

Mandelbrot, Benoit. 1974. *Intermittent turbulence in self-similar cascades: divergence of high moments and dimension of the carrier*, J. Fluid Mech. **62**, 331–358.

Rozenberg, G. and A. Salomaa. 1976. *The mathematical theory of L systems*, Advances in Information Systems Science, Vol. 6, Plenum Press, New York, pp. 161–206.

White, Damien G. 2001. *On the value of the critical point in fractal percolation*, Random Structures Algorithms **18**, 332–345.