

Higher order corrections for anisotropic bootstrap percolation

Hugo Duminil-Copin^{1,2} · Aernout C. D. van Enter³ ·
Tim Hulshof⁴ 

Received: 11 November 2016 / Revised: 2 October 2017 / Published online: 9 November 2017
© The Author(s) 2017. This article is an open access publication

Abstract We study the critical probability for the metastable phase transition of the two-dimensional anisotropic bootstrap percolation model with $(1, 2)$ -neighbourhood and threshold $r = 3$. The first order asymptotics for the critical probability were recently determined by the first and second authors. Here we determine the following sharp second and third order asymptotics:

$$p_c([L]^2, \mathcal{N}_{(1,2)}, 3) = \frac{(\log \log L)^2}{12 \log L} - \frac{\log \log L \log \log \log L}{3 \log L} + \frac{(\log \frac{9}{2} + 1 \pm o(1)) \log \log L}{6 \log L}.$$

We note that the second and third order terms are so large that the first order asymptotics fail to approximate p_c even for lattices of size well beyond 10^{1000} .

✉ Tim Hulshof
w.j.t.hulshof@tue.nl
Hugo Duminil-Copin
duminil@ihes.fr
Aernout C. D. van Enter
avanenter@gmail.com

- ¹ Institut des Hautes Études Scientifiques, Le Bois-Marie 35, route de Chartres, 91440 Bures-sur-Yvette, France
- ² Département de mathématiques, Université de Genève, 2-4 Rue du Lièvre, 64 1211 Genève 4, Switzerland
- ³ Johann Bernoulli Instituut voor Wiskunde en Informatica, University of Groningen, PO Box 407, 9700 AK Groningen, The Netherlands
- ⁴ Department of Mathematics and Computer Science, Eindhoven University of Technology, PO Box 513, 5600 MB Eindhoven, The Netherlands

Keywords Bootstrap percolation · Finite-size effects · Metastability · Sharp threshold

Mathematics Subject Classification 60K35 · 82B43 · 82C43

1 Introduction

1.1 Motivation and statement of the main result

Bootstrap percolation is a general name for the dynamics of monotone, two-state cellular automata on a graph G . Bootstrap percolation models with different rules and on different graphs have since their invention by Chalupa et al. [20] been applied in various contexts and the mathematical properties of bootstrap percolation are an active area of research at the intersection between probability theory and combinatorics. See for instance [1, 2, 4, 5, 8, 23, 33] and the references therein.

Motivated by applications to statistical (solid-state) physics such as the Glauber dynamics of the Ising model [26, 36] and kinetically constrained spin models [17], the underlying graph is often taken to be a d -dimensional lattice, and the initial state is usually chosen randomly.

Although some progress has recently been made in the study of very general cellular automata on lattices [12, 14, 25], attention so far has mainly focused on obtaining a very precise understanding of the metastable transition for specific simple models [4, 8, 9, 18, 23, 30, 33].

In this paper we will provide the most detailed description so far for such a model; namely, the so-called *anisotropic bootstrap percolation model*, defined as follows: First, given a finite set $\mathcal{N} \subset \mathbb{Z}^d \setminus \{0\}$ (the *neighbourhood*) and an integer r (the *threshold*), define the bootstrap operator

$$\mathcal{B}(\mathcal{S}) := \mathcal{S} \cup \{v \in \mathbb{Z}^d : |(v + \mathcal{N}) \cap \mathcal{S}| \geq r\} \quad (1.1)$$

for every set $\mathcal{S} \subset \mathbb{Z}^d$. That is, viewing \mathcal{S} as the set of “infected” sites, every site v that has at least r infected “neighbours” in $v + \mathcal{N}$ becomes infected by the application of \mathcal{B} . For $t \in \mathbb{N}$ let $\mathcal{B}^{(t)}(\mathcal{S}) = \mathcal{B}(\mathcal{B}^{(t-1)}(\mathcal{S}))$, where $\mathcal{B}^{(0)}(\mathcal{S}) = \mathcal{S}$, and let $\langle \mathcal{S} \rangle = \lim_{t \rightarrow \infty} \mathcal{B}^{(t)}(\mathcal{S})$ denote the set of eventually infected sites. For each $p \in (0, 1)$, let \mathbb{P}_p denote the probability measure under which the elements of the initial set $\mathcal{S} \subset \mathbb{Z}^d$ are chosen independently at random with probability p , and for each set $\Lambda \subset \mathbb{Z}^d$, define the *critical probability* on Λ to be

$$p_c(\Lambda, \mathcal{N}, r) := \inf \{p > 0 : \mathbb{P}_p(\Lambda \subset \langle \mathcal{S} \cap \Lambda \rangle) \geq 1/2\}. \quad (1.2)$$

If $\Lambda \subset \langle \mathcal{S} \cap \Lambda \rangle$ then we say that \mathcal{S} *percolates* on Λ . We remark that since we will usually expect the probability of percolation to undergo a sharp transition around p_c , the choice of the constant $1/2$ in the definition (1.2) is not significant.

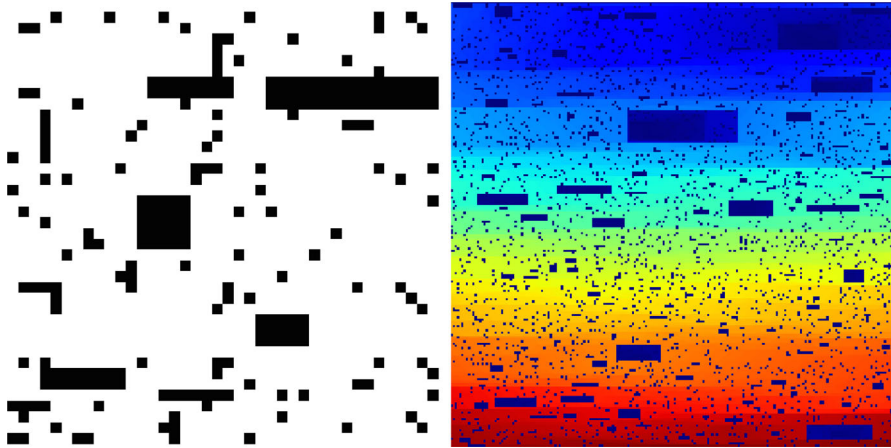


Fig. 1 On the left: a final configuration of the anisotropic model on $[40]^2$. Note that not all stable shapes are rectangles. On the right: a final configuration on $[200]^2$ with $p = 0.085$, where the color of each site represents the time it became infected. Blue sites became infected first, red sites last (color figure online)

The anisotropic bootstrap percolation model is a specific two-dimensional process in the family described above. To be precise, set $d = 2$ and

$$\mathcal{N}_{(1,2)} := \{(-2, 0), (-1, 0), (0, -1), (0, 1), (1, 0), (2, 0)\}$$

or graphically,

$$\mathcal{N}_{(1,2)} = \begin{matrix} & & \bullet & & \\ & \bullet & \bullet & 0 & \bullet & \bullet \\ & & & \bullet & & \bullet \end{matrix}$$

and set $r = 3$. $\mathcal{N}_{(1,2)}$ is sometimes called the “(1, 2)-neighbourhood” of the origin. See Fig. 1 for an illustration of the behaviour of the anisotropic model.

The main result of this paper is the following theorem:¹

Theorem 1.1 *The critical probability of the anisotropic bootstrap percolation model satisfies*

$$p_c([L]^2, \mathcal{N}_{(1,2)}, 3) = \frac{\log \log L}{12 \log L} \left(\log \log L - 4 \log \log \log L + 2 \log \frac{9e}{2} \pm o(1) \right). \tag{1.3}$$

¹ Throughout this paper we will use the standard Landau order notation: either for all x sufficiently large or sufficiently small, depending on the context,

- $f(x) = O(g(x))$ if there exists $C > 0$ such that $|f(x)| \leq Cg(x)$,
- $f(x) = \Omega(g(x))$ if there exists $c > 0$ such that $|f(x)| \geq cg(x)$,
- $f(x) = \Theta(g(x))$ if $f(x) = O(g(x))$ and $f(x) = \Omega(g(x))$,
- $f(x) = o(g(x))$ if $|f(x)|/|g(x)| \rightarrow 0$.

To put this theorem in context, let us recall some of the previous results obtained for bootstrap processes in two dimensions. The archetypal example of a bootstrap percolation model is the “two-neighbour model”, that is, the process with neighbourhood

$$\mathcal{N}_{(1,1)} := \{(-1, 0), (0, -1), (0, 1), (1, 0)\}$$

and $r = 2$. The strongest known bounds are due to Gravner, Holroyd, and Morris [28,30,37], who, building on work of Aizenman and Lebowitz [4] and Holroyd [33], proved that

$$p_c([L]^2, \mathcal{N}_{(1,1)}, 2) = \frac{\pi^2}{18 \log L} - \Theta\left(\frac{1}{(\log L)^{3/2}}\right). \tag{1.4}$$

The anisotropic model was first studied by Gravner and Griffeath [27] in 1996. In 2007, the second and third authors [41] determined the correct order of magnitude of p_c . More recently, the first and second authors [23] proved that the anisotropic model exhibits a sharp threshold by determining the first term in (1.3).

The “Duarte model” is another anisotropic model that has been studied extensively [13,22,38]. The Duarte model has neighbourhood

$$\mathcal{N}_{\text{Duarte}} = \{(-1, 0), (0, -1), (0, 1)\}$$

and $r = 2$. The sharpest known bounds here are due to the Bollobás et al. [13]:

$$p_c(\mathbb{Z}_L^2, \mathcal{N}_{\text{Duarte}}, 2) = \frac{(\log \log L)^2}{8 \log L} (1 \pm o(1)).$$

Although the Duarte model has the same first order asymptotics for p_c as the anisotropic model (up to the constant), the behaviour is very different. In particular, the Duarte model has a “drift” to the right: clusters grow only vertically and to the right. This asymmetry has severe consequences for the analysis of the model (especially for the shape of critical droplets).

The “ r -neighbour model” in d dimensions generalises the standard (two-neighbour) model described above. In this model, a vertex of \mathbb{Z}^d is infected by the process as soon as it acquires at least r already-infected nearest neighbours. Building on work of Aizenman and Lebowitz [4], Schonmann [40], Cerf and Cirillo [18], Cerf and Manzo [19], Holroyd [33] and Balogh et al. [9, 10], the following sharp threshold result for all non-trivial pairs (d, r) was obtained by Balogh et al. [8]: for every $d \geq r \geq 2$, there exists an (explicit) constant $\lambda(d, r) > 0$ such that

$$p_c([L]^d, \mathcal{N}_{(1,\dots,1)}, r) = \left(\frac{\lambda(d, r) \pm o(1)}{\log_{(r-1)} L}\right)^{d-r+1}.$$

(Here, and throughout the paper, $\log_{(k)}$ denotes a k -times iterated logarithm.)

Finally, we remark that much weaker bounds (differing by a large constant factor) have recently been obtained for an extremely general class of two-dimensional models

by Bollobás et al. [12], see Sect. 1.3, below. Moreover, stronger bounds (differing by a factor of $1 + o(1)$) were proved for a certain subclass of these models (including the two-neighbour model, but not the anisotropic model) by the Duminil-Copin and Holroyd [25].

Although various other specific models have been studied (see e.g. [15, 16, 34]), in each case the authors fell very far short of determining the second term.

1.2 The bootstrap percolation paradox

In [33] Holroyd for the first time determined sharp first order bounds on p_c for the standard model, and observed that they were very far removed from numerical estimates: $\pi^2/18 \approx 0.55$, while the same constant was numerically determined to be 0.245 ± 0.015 on the basis of simulations of lattices up to $L = 28800$ [3]. This phenomenon became known in the literature as the *bootstrap percolation paradox*, see e.g. [2, 21, 28, 31].

An attempt to explain this phenomenon goes as follows: if the convergence of p_c to its first-order asymptotic value is extremely slow, while for any fixed L the transition around p_c is very sharp, then it may appear that p_c converges to a fixed value long before it actually does.

This indeed appears to be the case. The first rigorisation of the “extremely slow convergence” part of this argument appears in [28], for a model related to bootstrap percolation. Theorem 1.1 gives another unambiguous illustration of extremely slow convergence for a bootstrap percolation model: the second term in (1.3) is actually larger than the first while

$$4 \log \log \log L > \log \log L,$$

which holds for all L in the range $66 < L < 10^{2390}$. Moreover, the second term does not become negligible (smaller than 1% of the first term, say) until $L > 10^{10^{1403}}$. On relatively small lattices, even the third term makes a significant contribution to p_c : it is larger than the first term when $L < 10^{60}$ and larger than the second term when $L < 10^{13}$.

The “sharp transition” part of the argument has also been made rigorous: for the standard model, an application of the Friedgut–Kalai sharp-threshold theorem [7] tells us that the “ ε -window of the transition”² is

$$p_{1-\varepsilon}([L]^2, \mathcal{N}_{(1,1)}, 2) - p_\varepsilon([L]^2, \mathcal{N}_{(1,1)}, 2) = O\left(\frac{\log \log L}{\log^2 L}\right).$$

So the ε -window is much smaller than the second order asymptotics in (1.4).

² The ε -window denotes the difference between the value of p_ε where $[L]^2$ is internally filled with probability ε , and the value $p_{1-\varepsilon}$ where this probability equals $1 - \varepsilon$. In other words, the ε -window tells us how sharp the metastable transition is.

For the anisotropic model a similar analysis [11] yields that the ε -window satisfies

$$p_{1-\varepsilon}([L]^2, \mathcal{N}_{(1,2)}, 3) - p_\varepsilon([L]^2, \mathcal{N}_{(1,2)}, 3) = O\left(\frac{\log^3 \log L}{\log^2 L}\right),$$

which is again much smaller than the second and third order asymptotics in Theorem 1.1. So our analysis supports the above explanation of the bootstrap percolation paradox.

1.3 Universality

Recently, a very general family of bootstrap-type processes was introduced and studied by Bollobás et al. [14]. To define this family, let $\mathcal{U} = \{X_1, \dots, X_m\}$ be a finite collection of finite subsets of $\mathbb{Z}^d \setminus \{0\}$, and define the corresponding bootstrap operator by setting

$$\mathcal{B}_\mathcal{U}(\mathcal{S}) = \mathcal{S} \cup \{v \in \mathbb{Z}^d : v + X \subset \mathcal{S} \text{ for some } X \in \mathcal{U}\}$$

for every set $\mathcal{S} \subset \mathbb{Z}^d$. It is not hard to see that all of the bootstrap processes described above can be encoded by such an ‘update family’ \mathcal{U} , and in fact this definition is substantially more general. The key discovery of [14] was that in two dimensions the class of such monotone cellular automata can be elegantly partitioned³ into three classes, each with completely different behaviour. More precisely, for every two-dimensional update family \mathcal{U} , one of the following holds:

- \mathcal{U} is “supercritical” and has polynomial critical probability.
- \mathcal{U} is “critical” and has poly-logarithmic critical probability.
- \mathcal{U} is “subcritical” and has critical probability bounded away from zero.

We emphasise that the first two statements were proved in [14], but the third was proved slightly later, by Balister et al. [6]. Note that the critical class includes the two-neighbour, anisotropic and Duarte models (as well as many others, of course). For this class a much more precise result was recently obtained by Bollobás et al. [12]. In order to state this result, let us first (informally) define a two-dimensional update family to be “balanced” if its growth is asymptotically two-dimensional⁴ (like that of the two-neighbour model), and “unbalanced” if its growth is asymptotically one-dimensional (like that of the anisotropic and Duarte models). The following theorem was proved in [12].

Theorem 1.2 *Let \mathcal{U} be a critical two-dimensional bootstrap percolation update family. There exists $\alpha = \alpha(\mathcal{U}) \in \mathbb{N}$ such that the following holds:*

³ This is the partitioning: We say a direction $u \in S^1$ is *stable* if \mathbb{H}_u , the discrete half-plane that is orthogonal to u , satisfies $\langle \mathbb{H}_u \rangle = \mathbb{H}_u$. A family is *supercritical* if there exists an open semicircle in S^1 containing no stable direction, and it is *subcritical* if every open semicircle contains an infinite number of stable directions. It is *critical* otherwise.

⁴ In other words, in a balanced model the critical droplet is a polygon, all of whose sides have the same length up to a constant factor. For the precise definition, which is somewhat more technical, see [12].

(a) If \mathcal{U} is balanced, then⁵

$$p_c(\mathbb{Z}_L^2, \mathcal{U}) = \Theta\left(\frac{1}{(\log L)^{1/\alpha}}\right).$$

(b) If \mathcal{U} is unbalanced, then

$$p_c(\mathbb{Z}_L^2, \mathcal{U}) = \Theta\left(\frac{(\log \log L)^2}{(\log L)^{1/\alpha}}\right).$$

Theorem 1.2 thus justifies our view of the anisotropic model as a canonical example of an unbalanced model.

1.4 Internally filling a critical droplet

As usual in (critical) bootstrap percolation, the key step in the proof of Theorem 1.1 will be to obtain very precise bounds on the probability that a “critical droplet” R is *internally filled*⁶ (IF), i.e., that $R \subset \langle \mathcal{S} \cap R \rangle$. We will prove the following bounds:

Theorem 1.3 *Let $p > 0$ and $x, y \in \mathbb{N}$ be such that $1/p^2 \leq x \leq 1/p^5$ and $\frac{1}{3p} \log \frac{1}{p} \leq y \leq \frac{1}{p} \log \frac{1}{p}$, and let R be an $x \times y$ rectangle. Then*

$$\mathbb{P}_p(R \text{ is internally filled}) = \exp\left(-\frac{1}{6p} \left(\log \frac{1}{p}\right)^2 + \left(\frac{1}{3} \log \frac{8}{3e} \pm o(1)\right) \frac{1}{p} \log \frac{1}{p}\right).$$

The alert reader may have noticed the following surprising fact: we obtain the first three terms of $p_c([L]^2, \mathcal{N}_{(1,2)}, 3)$ in Theorem 1.1, despite only determining the first two terms of $\log \mathbb{P}_p(R \text{ is IF})$ in Theorem 1.3. We will show how to formally deduce Theorem 1.1 from Theorem 1.3 in Sect. 7, but let us begin by giving a brief outline of the argument.

To slightly simplify the calculations, let us write

$$C_1 := \frac{1}{12} \quad \text{and} \quad C_2 := \frac{1}{6} \log \frac{8}{3e}.$$

We claim (and will later prove) that $p_c = p_c([L]^2, \mathcal{N}_{(1,2)}, 3)$ is essentially equal to the value of p for which the expected number of internally filled critical droplets in $[L]^2$ is equal to 1 (the idea being that a critical droplet with size as in Theorem 1.3 will keep growing indefinitely with probability very close to one). We therefore have

⁵ Here \mathbb{Z}_L^2 denotes the discrete two-dimensional $L \times L$ torus, and $p_c(\mathbb{Z}_L^2, \mathcal{U})$ is defined as in (1.2). We consider the torus since in general undesirable complications may arise due to boundary effects or strongly asymmetrical growth.

⁶ This notion is often referred to as “internally spanned” (especially in the older literature).

$$L^2 \approx \exp\left(\frac{2C_1}{p_c} \left(\log \frac{1}{p_c}\right)^2 - \frac{2C_2}{p_c} \log \frac{1}{p_c}\right),$$

and hence

$$p_c \approx \frac{C_1}{\log L} \left(\log \frac{1}{p_c}\right)^2 - \frac{C_2}{\log L} \log \frac{1}{p_c}.$$

Iterating the right-hand side gives

$$p_c \approx \frac{C_1}{\log L} \left(\log \log L - 2 \log \log \frac{1}{p_c} - \log C_1\right)^2 - \frac{C_2}{\log L} \left(\log \log L - 2 \log \log \frac{1}{p_c}\right).$$

Upon using the approximation $\log \log(1/p_c) \approx \log \log \log L$ and multiplying out, this reduces to

$$p_c \approx \frac{C_1(\log \log L)^2}{\log L} - \frac{4C_1 \log \log L \log \log \log L}{\log L} - (C_2 + 2C_1 \log C_1) \frac{\log \log L}{\log L},$$

which is what we hope to prove. Thus we obtain three terms for the price of two.

1.5 A generalisation of the anisotropic model

One natural way to generalise the anisotropic model is to consider, for each $b > a \geq 1$, the neighbourhood

$$\mathcal{N}_{(a,b)} = \left(\{(0, y) \in \mathbb{Z}^2 : -a \leq y \leq a\} \cup \{(x, 0) \in \mathbb{Z}^2 : -b \leq x \leq b\}\right) \setminus \{(0, 0)\}.$$

It follows from Theorem 1.2 that

$$p_c(\mathbb{Z}_L^2, \mathcal{N}_{(a,b)}, r) = \Theta\left(\frac{(\log \log L)^2}{(\log L)^{1/\alpha}}\right),$$

where $\alpha = r - b$, for each $b + 1 \leq r \leq a + b$.⁷ The arguments developed in [23] can be applied to prove that the leading order behaviour of p_c for the $(1, b)$ -model is⁸

$$p_c([L]^2, \mathcal{N}_{(1,b)}, b + 1) = \left(\frac{(b - 1)^2}{4(b + 1)} \pm o(1)\right) \frac{(\log \log L)^2}{\log L}.$$

⁷ The value of α follows from [12, Definition 1.2]. Furthermore, if $r \leq b$ then the model is supercritical, so $p_c(\mathbb{Z}_L^2, \mathcal{N}_{(a,b)}, r) \leq L^{-c}$ for some $c > 0$, and if $r > a + b$ then the model is subcritical, so $p_c(\mathbb{Z}_L^2, \mathcal{N}_{(a,b)}, r) > c'$ for some $c' > 0$.

⁸ This is contrary to the claim in [23, Sect. 1]. See also [24], the erratum to [23].

Combining the techniques of [23] with those introduced in this paper, it is possible to prove the following stronger bounds:

Theorem 1.4 *Given $b \geq 2$, set*

$$C(b) = \frac{2}{b-1} \log \left(\frac{\binom{2b}{b} - \frac{2b-1}{b+1} \binom{2b-2}{b} - 1 + \binom{2b}{b-1} - \binom{2b-2}{b-3}}{(b+1)e} \right) + 2 \log \left(\frac{(b-1)^2}{4(b+1)} \right).$$

Then

$$p_c([L]^2, \mathcal{N}_{(1,b)}, b+1) = \frac{(b-1)^2}{4(b+1)} \frac{\log \log L}{\log L} \left(\log \log L - 4 \log \log \log L - C(b) \pm o(1) \right). \tag{1.5}$$

Note that in the case $b = 2$ this reduces to Theorem 1.1. We remark that Theorem 1.4 follows from a corresponding generalisation of Theorem 1.3, with the constants $\frac{1}{6}$ and $\frac{1}{3} \log \frac{8}{3e}$ replaced by

$$\frac{(b-1)^2}{2(b+1)} \quad \text{and} \quad \frac{b-1}{b+1} \log \left(\frac{\binom{2b}{b} - \frac{2b-1}{b+1} \binom{2b-2}{b} - 1 + \binom{2b}{b-1} - \binom{2b-2}{b-3}}{(b+1)e} \right),$$

respectively.

We will not prove Theorem 1.4, since the proof is conceptually the same as that in the case $b = 2$, but requires several straightforward but lengthy calculations that might obscure the key ideas of the proof. It is, however, not too hard to see where the numerical factors come from:

A droplet grows horizontally in the $(1, b)$ -model as long as it does not occur that the $b + 1$ consecutive columns to its left and/or right do not contain an infected site. And it grows vertically as long as there are b sites in a ‘‘growth configuration’’ somewhere above and/or below. There are

$$\binom{2b}{b} - \frac{2b-1}{b+1} \binom{2b-2}{b} - 1 + \binom{2b}{b-1} - \binom{2b-2}{b-3} \tag{1.6}$$

such configurations. Indeed, there are $\binom{2b}{b}$ different ways of finding b infected sites inside $\mathcal{N}_{(1,b)} \setminus \{(0, -1), (0, 1)\}$. Of these, $\sum_{i=2}^b \binom{2b-i}{b} + 1 = \frac{2b-1}{b+1} \binom{2b-2}{b} + 1$ are right-shifts of another configuration (e.g. for $\mathcal{N}_{(1,2)}$ the choices $\bullet \circ \bullet \circ$ and $\circ \bullet \circ \bullet$ count as a single growth configuration), so their contribution must be subtracted. If $(0, 1)$ is occupied, there are $\binom{2b}{b-1}$ ways of placing the other $b - 1$ sites in $\mathcal{N}_{(1,b)} \setminus \{(0, -1), (0, 1)\}$. None of these are shift invariant, but some of them cannot grow to fill the entire row. Indeed, when $b \geq 3$, configurations where $(-b, 0)$, $(0, 1)$, and $(b, 0)$ are infected do not cause the row to fill up. Therefore, we must subtract $\binom{2b-2}{b-3}$. This explains (1.6). See Fig. 2 for growth configurations of the case $b = 2$. Finally, it takes $b - 1$ more infected sites for a rectangle to grow a row than it does to grow a column, which explains the remaining factors $b - 1$ in (1.5).

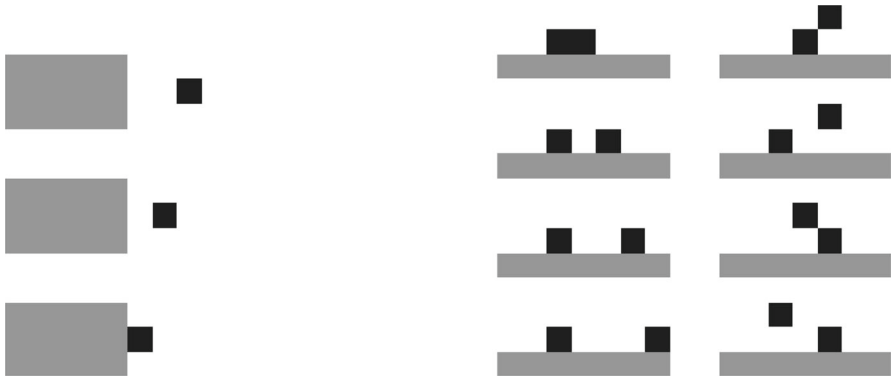


Fig. 2 On the left: the three relative positions of an infected site that can cause horizontal growth of the grey rectangle. On the right: the 8 pairs of sites (up to shifts) that can cause vertical growth of the grey rectangle

1.6 Comparison with simulations

One might be tempted to hope that the third-order approximation of p_c in Theorem 1.1 is reasonably good already for lattices that a computer might be able to handle. Simulations indicate that this is not the case. Indeed, for lattices with $L \leq 10,000$, the third-order approximation is even farther from to the simulated values than the first-order approximation (and recall that the second-order approximation is negative here). We believe that this should not be surprising, because it is not at all obvious that the fourth order term should be significantly smaller: careful inspection of our proof suggests that the $o(\frac{1}{p} \log \frac{1}{p})$ term in Theorem 1.3 is at most $O(\frac{1}{p} \log \log \frac{1}{p})$. Although we do not prove this, we have no reason to believe that a correction term of that order does not exist. Even if we suppose that the third order correction in Theorem 1.3 can be sharply bounded by C_3/p , say, so that we would have the bound

$$\mathbb{P}_p(R \text{ is internally filled}) \stackrel{?}{=} \exp\left(-\frac{2C_1}{p} \left(\log \frac{1}{p}\right)^2 + \frac{2C_2}{p} \log \frac{1}{p} + \frac{C_3 \pm o(1)}{p}\right),$$

for critical droplets instead, then a computation like the one in Sect. 1.4 above suggests that this would yield

$$\begin{aligned} p_c([L]^2, \mathcal{N}_{(1,2)}, 3) &\stackrel{?}{=} \frac{(\log \log L)^2}{12 \log L} - \frac{(\log \log L)(\log \log \log L)}{3 \log L} \\ &+ \frac{(\log \frac{9}{2} + 1) \log \log L}{6 \log L} \\ &+ \frac{(\log \log \log L)^2}{3 \log L} - \frac{(\log \frac{9}{2} + 1) \log \log \log L}{3 \log L} \\ &+ \frac{C_3 + \log \frac{1}{12} (2 + \log \frac{27}{16}) \pm o(1)}{12 \log L}, \end{aligned}$$

so the fourth, fifth, and sixth order terms of p_c would also be comparable to the first for moderately sized lattices. Moreover, because of the extremely slow decay of these correction terms (e.g. $(\log \log 10^{10})^2 \approx 10$), it might be too optimistic to expect that one would be able to determine C_3 by fitting to the simulated values of p_c , if indeed C_3 exists.

1.7 Comparison with the two-neighbor model

Comparing Theorem 1.1 with the analogous result for the two-neighbor model, (1.4), it may seem remarkable how much sharper the former is than the latter. We believe the following heuristic discussion goes a way towards explaining this difference.

Both approximations of p_c are proved using essentially the same critical droplet heuristic described above. Once a critical droplet has formed, the entire lattice will easily fill up. But filling a droplet-sized area is exponentially unlikely: it is essentially a large deviations event. The theory of large deviations tells us that if a rare event occurs, it will occur in the most probable way that it can. For filling a droplet, this means that one should find an optimal “growth trajectory”: a sequence of dimensions from which a very small infected area (a “seed”) steadily grows to fill up the entire droplet. For the anisotropic model, in [23], the first and second authors determined this trajectory to be close to $x = \frac{e^{3py}}{3p}$, where x and y denote the horizontal and vertical dimensions of the seed as it grows. This approximation was enough to yield the first term of p_c . In the current paper we establish tighter bounds of optimal trajectory around $x = \frac{e^{3py}}{3p}$, allowing us to give the sharper estimate for the probability of filling a droplet in Theorem 1.3. As we showed in Sect. 1.4 above, this correction is enough to obtain the first three terms of p_c for the anisotropic model.

For the two-neighbor model, however, finding this optimal growth trajectory is not at all the challenge: by symmetry it is trivially $x = y$. The correction to p_c that Gravner, Holroyd, and Morris determined in [28,30,37], is instead due to the much smaller entropic effect of random fluctuations around this trajectory (see also the introduction of [29] for a more detailed explanation of this effect). We believe that such fluctuations also influence p_c for the anisotropic model, but that their effect will be much smaller than the improvements that can still be made in controlling the precise shape of the optimal growth trajectory.

1.8 About the proofs

The proof of Theorem 1.1 uses a rigorisation of the iterative determination of p_c in Sect. 1.4 above, combined with Theorem 1.3 and the classical argument of Aizenman and Lebowitz [4].

The lower bound in Theorem 1.3 is a refinement of the computation in [23].

Most of the work of this paper goes into the proof of the upper bound of Theorem 1.3. Like many recent entries in the bootstrap percolation literature, our proof centers around the “hierarchies” argument of Holroyd [33]. In particular, we sharpen the argument of [23] by incorporating the idea of “good” and “bad” hierarchies from

[30], and by using very precise bounds on horizontal and vertical growth of infected rectangular regions.

The main new contributions of this paper (besides the iterative determination of p_c) can be found in Sects. 3 and 6.

In Sect. 3, we introduce the notion of *spanning time* (Definition 3.3), which characterises to a large extent the structure of configurations of vertical growth. We show that if the spanning time is 0, then such structures have a simple description in terms of paths of infected sites, whereas if the spanning time is not 0, then this description can still be given in terms of paths, but these paths now also involve more complex arrangements of infected sites. We call such arrangements *infectors* (Definition 3.7), and show that they are sufficiently rare that their contribution does not dominate the probability of vertical growth.

In Sect. 6 we generalise the variational principle of Holroyd [33] to a more general class of growth trajectories. This part of the proof is intended to be more widely applicable than the current anisotropic case, and is set up to allow for precise estimates.

1.9 Notation and definitions

A *rectangle* $[a, b] \times [c, d]$ is the set of sites in \mathbb{Z}^2 contained in the Euclidean rectangle $[a, b] \times [c, d]$. For a finite set $\mathcal{Q} \subset \mathbb{Z}^2$, we denote its *dimensions* by $(\mathbf{x}(\mathcal{Q}), \mathbf{y}(\mathcal{Q}))$, where $\mathbf{x}(\mathcal{Q}) = \max\{a_1 - b_1 + 1 : \{(a_1, a_2), (b_1, b_2)\} \in \mathcal{Q} \times \mathcal{Q}\}$, and similarly, $\mathbf{y}(\mathcal{Q}) = \max\{a_2 - b_2 + 1 : \{(a_1, a_2), (b_1, b_2)\} \in \mathcal{Q} \times \mathcal{Q}\}$. So in particular, a rectangle $R = [a, b] \times [c, d]$ has dimensions $(\mathbf{x}(R), \mathbf{y}(R)) = (|[a, b] \cap \mathbb{Z}|, |[c, d] \cap \mathbb{Z}|)$. Oftentimes, the quantities that we calculate will only depend on the size of R , and be invariant with respect to the position of R . In such cases, when there is no possible confusion, we will write R with $\mathbf{x}(R) = x$ and $\mathbf{y}(R) = y$ as $[x] \times [y]$. A *row* of R is a set $\{(m, n) \in R : n = n_0\}$ for some fixed n_0 . A *column* is similarly defined as a set $\{(m, n) \in R : m = m_0\}$. We sometimes write $[a, b] \times \{c\}$ for the row $\{(m, c) \in \mathbb{Z}^2 : m \in [a, b] \cap \mathbb{Z}\}$, and use similar notation for columns.

We say that a rectangle $R = [a, b] \times [c, d]$ is *horizontally traversable* (hor-trav) by a configuration \mathcal{S} if

$$R \subset \langle (R \cap \mathcal{S}) \cup ([a - 2, a - 1] \times [c, d]) \rangle.$$

That is, R is horizontally traversable if the rectangle becomes infected when the two columns to its left are completely infected. Under \mathbb{P}_p , this event is equiprobable to the event that $R \subset \langle (R \cap \mathcal{S}) \cup ([b + 1, b + 2] \times [c, d]) \rangle$, and more importantly, it is equivalent to the event that R does not contain three or more consecutive columns without any infected sites and the rightmost column contains an infected site.

We say that R is *up-traversable* (up-trav) by \mathcal{S} if

$$R \subset \langle (R \cap \mathcal{S}) \cup ([a, b] \times \{c - 1\}) \rangle.$$

That is, R becomes entirely infected when all sites in the row directly below R are infected. Similarly, we say that R is *down-traversable* by \mathcal{S} if $R \subset \langle (R \cap \mathcal{S}) \cup ([a, b] \times$

$\{d + 1\}$). Again, under \mathbb{P}_p up and down traversability are equiprobable, so we will only discuss up-traversability. If \mathcal{S} is a random site percolation, then we simply say that R is horizontally- or up- or down-traversable.

Given rectangles $R \subset R'$ we write $\{R \Rightarrow R'\}$ for the event that the dynamics restricted to R' eventually infect all sites of R' if all sites in R are infected, i.e., for the event that $R' = \langle (\mathcal{S} \cap R') \cup R \rangle$.

We will frequently make use of two standard correlation inequalities: The first is the *Fortuin–Kasteleyn–Ginibre inequality* (FKG-inequality), which states that for increasing events A and B , $\mathbb{P}_p(A \cap B) \geq \mathbb{P}_p(A)\mathbb{P}_p(B)$. The second is the *van den Berg–Kesten inequality* (BK-inequality), which states that for increasing events A and B , $\mathbb{P}_p(A \circ B) \leq \mathbb{P}_p(A)\mathbb{P}_p(B)$, where $A \circ B$ means that A and B occur disjointly (see [32, Chapter 2] for a more in-depth discussion).

1.10 The structure of this paper

In Sect. 2 we state two key bounds, Lemmas 2.2 and 2.3, giving primarily lower bounds on the probabilities of horizontal and vertical growth of an infected rectangular region, and we use them to prove the lower bound of Theorem 1.3. In Sect. 3 we prove a complementary upper bound on the vertical growth of infected rectangles, Lemma 3.1. In Sect. 4 we prove Lemma 4.1, which combines the upper bounds on horizontal and vertical growth from Lemmas 2.2 and 3.1. This lemma is crucial for the upper bound of Theorem 1.3. We prove the upper bound of Theorem 1.3 in Sect. 5, subject to a variational principle, Lemma 5.9, that we prove in Sect. 6. Finally, in Sect. 7 we use Theorem 1.3 to prove Theorem 1.1.

2 The lower bound of Theorem 1.3

Recall that $C_1 = \frac{1}{12}$ and $C_2 = \frac{1}{6} \log \frac{8}{3e}$.

Proposition 2.1 *Let $p > 0$ and $\frac{1}{p^2} \leq x \leq \frac{1}{p^5}$ and $\frac{1}{3p} \log \frac{1}{p} \leq y \leq \frac{1}{p^5}$. Then*

$$\mathbb{P}_p([x] \times [y] \text{ is IF}) \geq \exp\left(-\frac{2C_1}{p} \log^2 \frac{1}{p} + (2C_2 - o(1)) \frac{1}{p} \log \frac{1}{p}\right).$$

Note that the upper bound on y is different from the bound in Theorem 1.3.

For the proof it suffices to show that there exists a subset of configurations that has the desired probability. We choose a subset of configurations that follow a typical “growth trajectory”: configurations that contain a small area that is locally densely infected (a *seed*). We bound the probability that such a seed will grow a bit (which is likely), and then a lot more (which is exponentially unlikely), until the infected region reaches a size where the growth is again very likely, because the boundary of the infected region is large and the dynamics depend only on the existence of infected sites on the boundary, not on their number.

To prove this proposition we will need bounds on the probability that a rectangle becomes infected in the presence of a large infected cluster on its boundary. We state two lemmas that achieve this, which are improvements upon [23, Lemmas 2.1 and 2.2].

Lemma 2.2 *For any rectangle $[x] \times [y]$,*

$$e^{-xf(p,y)} \leq \mathbb{P}_p([x] \times [y] \text{ is hor-trav}) \leq e^{-(x-2)f(p,y)},$$

where $f(p, y) := -\log(\alpha(1 - (1 - p)^y))$ and where $\alpha(u)$ is the positive root of the polynomial

$$X^3 - uX^2 - u(1 - u)X - u(1 - u)^2. \tag{2.1}$$

Moreover, $f(p, y)$ satisfies the following bounds:

(a) when $p \rightarrow 0$ and $py \rightarrow \infty$,

$$f(p, y) = e^{-3py} + \Theta(e^{-4py}),$$

(b) when $y \geq \frac{2}{p} \log \log \frac{1}{p}$,

$$f(p, y) = e^{-3py} \left(1 + \Theta\left(\log^{-2}(1/p)\right) \right),$$

(c) when $p \rightarrow 0$, $y \rightarrow \infty$, and $(1 - p)^y \rightarrow 1$,

$$f(p, y) \geq \frac{1}{2}py - 3p^2y^2.$$

Proof From [23, Lemma 2.1]⁹ we know that

$$\alpha(1 - (1 - p)^y)^x \leq \mathbb{P}_p([x] \times [y] \text{ is hor-trav}) \leq \alpha(1 - (1 - p)^y)^{x-2}.$$

When u is close to 1, $X = e^{-(1-u)^3}$ is an approximate solution for the positive root, since

$$e^{-3(1-u)^3} - ue^{-2(1-u)^3} - u(1 - u)e^{-(1-u)^3} - u(1 - u)^2 = \Theta((1 - u)^4).$$

So, as $p \rightarrow 0$ and $py \rightarrow \infty$,

$$-\log \alpha(1 - (1 - p)^y) = (1 - p)^{3y} + \Theta((1 - p)^{4y}) = e^{-3py} + \Theta(e^{-4py}).$$

This establishes (a) and (b) simply follows.

To prove (c), recall Rouché’s Theorem (see e.g. [39, Theorem 10.43]), which states that if two functions $g(z)$ and $h(z)$ are holomorphic on a bounded region $U \subset \mathbb{C}$

⁹ Note that in the proof of [23, Lemma 2.1] there are a number of (unimportant) sign errors.

with continuous boundary ∂U and satisfy $|g(z) - h(z)| < |g(z)|$ for all $z \in \partial U$, then g and h have an equal number of roots on U . Applying Rouché’s Theorem with $h(z) = a_0 + a_1z + a_2z^2 + a_3z^3$ and $g(z) = a_0$, it follows that the moduli of the roots of $h(z)$ are all bounded from below by $|a_0|/(|a_0| + \max\{|a_1|, |a_2|, |a_3|\})$. Applying this bound to (2.1) we find that when $u > 0$ is sufficiently small,

$$\alpha(u) \geq \frac{u(1 - u)^2}{u(1 - u)^2 + 1} \geq u - 3u^2,$$

where the second inequality is due to a series expansion around $u = 0$. (We remark that an explicit computation gives $\alpha(u) \geq u - 3u^2$ for all $u > 0$, but without relying on a computer this may take several pages to verify.) Since we assumed $(1 - p)^y \rightarrow 1$ we thus have

$$f(p, y) \geq (1 - (1 - p)^y) - 3(1 - (1 - p)^y)^2 \geq \frac{1}{2}py - 3p^2y^2,$$

where we used $\frac{1}{2}py \leq 1 - (1 - p)^y \leq py$ for p sufficiently small. □

Lemma 2.3 (a) *If p^2x is sufficiently small, then we have, for any rectangle $[x] \times [y]$,*

$$\mathbb{P}_p([x] \times [y] \text{ is up-trav}) \geq \exp\left(y \log(8p^2x)(1 + O(p^2x + p))\right).$$

(b) *As long as $\frac{8p^2x}{5} \leq 1$ we have*

$$\mathbb{P}_p([x] \times [y] \text{ is up-trav}) \geq \left(\frac{8p^2x}{5e}\right)^y.$$

Proof We say that a rectangle is *North-traversable* (N-trav) if the intersection of every row with R contains a site (a, b) such that $((a, b) + \mathcal{N}_{(1,2)}) \setminus \{(a, b - 1)\}$ contains at least two infected sites. Observe that North-traversability implies up-traversability, so

$$\mathbb{P}_p([x] \times [y] \text{ is up-trav}) \geq \mathbb{P}_p([x] \times [y] \text{ is N-trav}).$$

We can similarly define South-traversability by requiring that the intersection of every row with R contains a site (a, b) such that $((a, b) + \mathcal{N}_{(1,2)}) \setminus \{(a, b + 1)\}$ contains at least two infected sites. South-traversability implies down-traversability. Again, from a probabilistic point of view North- and South-traversability are equivalent, so we will henceforth only discuss North-traversability.

If $[x] \times [y]$ is North-traversable then for each of the y rows there must exist an infected pair of sites u and v and a site z in the row such that $u, v \in z + \mathcal{N}_{(1,2)}$. By the FKG inequality we thus have the lower bound

$$\mathbb{P}_p([x] \times [y] \text{ is N-trav}) \geq \mathbb{P}_p(\exists \text{ an infected pair for a row of length } x)^y.$$

For the proof of (a) we apply Janson’s inequality [35]. The expected number of infected pairs immediately above an infected rectangle of width x is at least $\mu = (8x -$

$16)p^2$. To see this, consider that up to translations there are 8 possible pairs of infected sites above the rectangle that can infect the whole row, see Fig. 2 above. The variance is¹⁰ $\Delta = O(p^3x) \ll \mu$, so the probability that some pair is infected is at least

$$1 - \exp(-\mu + \Delta/2) \geq \left(8p^2x - O(p^3x + p^4x^2)\right),$$

using the inequality $1 - e^{-x} \geq x - x^2$ for $x \geq 0$.

For the proof of (b) we use a cruder approximation: For $(a, b) \in [x] \times [y]$ let $A_{(a,b)}$ be the event that (a, b) is the leftmost site of an infected pair as in Fig. 2. These pairs all have width at most 5, so the probability that a row of length x does not have an infected pair can be bounded from above by

$$(1 - 8p^2)^{\lfloor x/5 \rfloor} \leq \exp\left(-\frac{8p^2x}{5}\right) \leq 1 - \frac{8p^2x}{5e}$$

when $\frac{8p^2x}{5} \leq 1$. The claim follows. □

Proof of Proposition 2.1 We start by constructing a seed. Let $r := \lfloor \frac{2}{p} \log \log \frac{1}{p} \rfloor$ and infect sites $(1, 2i)$ and $(2, 2i + 1)$ for $2i \leq r$. The probability that a rectangle $[2] \times [r]$ is a seed is p^r . Note that the infected sites internally fill $[2] \times [r]$.

The growth of the seed to a rectangle of arbitrary size can be divided into three stages:

Stage 1. By Lemma 2.2(a) the probability of finding a seed of size r that will grow to size $\lceil e^{3rp}/(3p) \rceil \times [r]$ is about the same as the probability of just finding the seed, i.e.,

$$p^r \cdot \exp\left(-\frac{e^{3rp}}{3p} \cdot \left(e^{-3rp} + O\left(e^{-4rp}\right)\right)\right) \geq p^r e^{-O(1/p)}. \tag{2.2}$$

Stage 2. Next we bound the probability that the infected rectangle grows to size

$$R := \left\lceil \frac{1}{3p^2} \right\rceil \times \left\lceil \frac{1}{3p} \log \frac{1}{p} \right\rceil,$$

that is, we want to bound

$$\mathbb{P}_p \left(\left\lceil \frac{e^{3rp}}{3p} \right\rceil \times [r] \Rightarrow \left\lceil \frac{e^{3mp}}{3p} \right\rceil \times [m] \right), \tag{2.3}$$

where $m := \frac{1}{3p} \log \frac{1}{p}$. This is the bottleneck for the growth dynamics. We bound (2.3) by considering the growth in many small steps. In each such step, the rectangle will

¹⁰ For two positive sequences a_n and b_n we write $a_n \gg b_n$ when $a_n/b_n \rightarrow \infty$ and $a_n \ll b_n$ when $a_n/b_n \rightarrow 0$.

either infect an entire row above or below it, or it will infect an entire row to the left or right of it (with the help of infected sites on the boundary of the rectangle). Because vertical growth is less probable than horizontal growth, we will consider sequences where the rectangle grows by one vertical step, from height ℓ to $\ell + 1$, followed by horizontal growth that infects many columns successively, with the rectangle growing from width x_ℓ to $x_{\ell+1}$ where $x_\ell := \frac{e^{3\ell p}}{3p}$. That this choice is close to optimal can be seen in Sect. 6 below, where a variational principle for the upper bound of Theorem 1.3 is derived.

Having divided the growth into steps, we can bound (2.3) from below using the FKG-inequality:

$$\begin{aligned}
 & \mathbb{P}_p \left(\left[\frac{e^{3rp}}{3p} \right] \times [r] \Rightarrow \left[\frac{e^{3mp}}{3p} \right] \times [m] \right) \\
 & \geq \prod_{\ell=r}^m \mathbb{P}_p ([x_\ell] \times [\ell] \Rightarrow [x_{\ell+1}] \times [\ell]) \\
 & \quad \times \prod_{\ell=r}^{m-r} \mathbb{P}_p ([x_\ell] \times [\ell] \Rightarrow [x_\ell] \times [\ell + 1]) \\
 & \quad \times \prod_{\ell=m-r+1}^m \mathbb{P}_p ([x_\ell] \times [\ell] \Rightarrow [x_\ell] \times [\ell + 1]). \tag{2.4}
 \end{aligned}$$

We bound these three products separately.

It follows from Lemma 2.2(a) that the horizontal growth from width x_ℓ to $x_{\ell+1}$ occurs with probability approximately $1/e$, i.e.,

$$\begin{aligned}
 & \mathbb{P}_p ([x_\ell] \times [\ell] \Rightarrow [x_{\ell+1}] \times [\ell]) \\
 & \geq \exp \left(-\frac{1}{3p} (e^{3(\ell+1)p} - e^{3\ell p}) e^{-3\ell p} (1 + O(\log^{-4/3}(1/p))) \right) \\
 & \geq e^{-1-o(1)}. \tag{2.5}
 \end{aligned}$$

Therefore,

$$\prod_{\ell=r}^m \mathbb{P}_p ([x_\ell] \times [\ell] \Rightarrow [x_{\ell+1}] \times [\ell]) \geq e^{-m(1+o(1))}. \tag{2.6}$$

When $\ell \leq m - r$, then $p^2 x_\ell \leq \log^{-2} \frac{1}{p}$, so we can apply Lemma 2.3(a) to bound

$$\mathbb{P}_p ([x_\ell] \times [\ell] \Rightarrow [x_\ell] \times [\ell + 1]) \geq 8p^2 x_\ell e^{O(\log^{-4/3} \frac{1}{p})}.$$

Therefore we can bound the second product in (2.4) from below by

$$\begin{aligned}
 & \prod_{\ell=r}^{m-r} 8p^2 x_\ell e^{O(\log^{-4/3} \frac{1}{p})} \\
 & \geq \left(\frac{8p}{3}\right)^{m-2r} \exp\left(3p \sum_{\ell=r}^{m-r} \ell\right) e^{(m-2r)O(\log^{-4/3} \frac{1}{p})} \\
 & = \left(\frac{8p}{3}\right)^{m-2r} \exp\left(\frac{3p}{2}((m-r)(m-r+1) - (r-1)r)\right) e^{o(m)} \\
 & = p^{-r} \left(\frac{8p}{3}\right)^{m-r} \exp\left(\frac{3p}{2}(m^2 - 2mr + m)\right) e^{o(m)} \\
 & = p^{-r} \left(\frac{8p}{3}\right)^{m-r} \exp\left(\frac{3p}{2}(m^2 - 2mr)\right) e^{o(m)}. \tag{2.7}
 \end{aligned}$$

Using Lemma 2.3(b) we can similarly bound the third product from below by

$$\begin{aligned}
 \prod_{\ell=m-r+1}^m \frac{8p^2 x_\ell}{5e} & = \left(\frac{8p}{15e}\right)^r \exp\left(\sum_{\ell=m-r+1}^m 3\ell p\right) \\
 & \geq \left(\frac{8p}{3}\right)^r \exp\left(\frac{3p}{2}(m(m+1) - (m-r)(m-r+1))\right) \left(\frac{1}{5e}\right)^r \\
 & = \left(\frac{8p}{3}\right)^r \exp\left(\frac{3p}{2}2mr\right) e^{o(m)}. \tag{2.8}
 \end{aligned}$$

Multiplying the bounds (2.6), (2.7), and (2.8), and using that $m = \frac{1}{3p} \log \frac{1}{p}$, we get

$$\begin{aligned}
 & \mathbb{P}_p \left(\left[\frac{e^{3rp}}{3p} \right] \times [r] \Rightarrow \left[\frac{e^{3mp}}{3p} \right] \times [m] \right) \\
 & \geq p^{-r} \left(\frac{8p}{3}\right)^m \exp\left(\frac{3p}{2}m^2 - m\right) e^{o(m)} \\
 & = p^{-r} \exp\left(\frac{3p}{2}m^2 - m \log \frac{1}{p} + m \log \frac{8}{3} - m\right) e^{o(m)} \\
 & = p^{-r} \exp\left(-\frac{1}{6p} \log^2 \frac{1}{p} + (1 - o(1)) \frac{1}{3p} \log \frac{8}{3e} \log \frac{1}{p}\right). \tag{2.9}
 \end{aligned}$$

Stage 3. The infected region can grow from $[\frac{1}{3p^2}] \times [m]$ to arbitrary size with good probability. Indeed, we claim that

$$\mathbb{P}_p \left(\left[\frac{1}{3p^2} \right] \times [m] \Rightarrow R \right) \geq e^{-O(1/p)}. \tag{2.10}$$

This bound is proved in [23, proof of Proposition 2.4]. We do not repeat the proof here, but let us indicate how this bound is established: Consider the case where the cluster first grows horizontally to width $1/p^2$. By Lemma 2.2(b) we have

$$\mathbb{P}_p \left(\left[\frac{1}{3p^2} \right] \times [m] \Rightarrow \left[\frac{1}{p^2} \right] \times [m] \right) \geq \exp \left(-\frac{2}{3p^2} \cdot p(1 + o(1)) \right) = e^{-O(1/p)}.$$

Now consider the case where it grows vertically, this time to height $3m$. This also occurs with probability at least $e^{-O(1/p)}$. As the infected region gets larger, the probability that it keeps growing converges to 1. The result is that (2.10) holds for any rectangle R that is large enough, as long as the dimensions of R are sufficiently balanced (which is guaranteed by the assumptions on x and y).

Now, by the FKG-inequality, we can multiply the bounds from the three stages (i.e., (2.2), (2.9), and (2.10)) to complete the proof of Proposition 2.1. □

3 An upper bound on the probability of up-traversability

The following bound is crucial for the proof of the upper bound of Theorem 1.3. Recall from (1.1) the definition of the bootstrap operator \mathcal{B} , and recall that $\mathcal{B}^{(t)}(\mathcal{S})$ is the t -th iterate of \mathcal{B} with initial set \mathcal{S} , and that $\langle \mathcal{S} \rangle = \lim_{t \rightarrow \infty} \mathcal{B}^{(t)}(\mathcal{S})$. Recall that a rectangle $R = [1, x] \times [1, y]$ is said to be up-traversable by a set \mathcal{S} if $R \subset \langle (\mathcal{S} \cap R) \cup ([1, x] \times \{0\}) \rangle$, and that we write \mathbb{P}_p to indicate that the elements of \mathcal{S} are chosen independently at random with probability p .

Lemma 3.1 *Let $1 \leq k \ll p^{-2/5}$ and let R be a rectangle with dimensions (x, y) such that $y < x$. Then, for p sufficiently small,*

$$\mathbb{P}_p(R \text{ is up-traversable}) \leq \begin{cases} p^{-k} e^{y/k} (24pk^2 + 8p)^y & \text{if } x < \frac{3k^2}{p}, \\ p^{-k} e^{y/k} (8p^2x + 8p)^y & \text{if } \frac{3k^2}{p} \leq x \leq \frac{1}{p^2}. \end{cases}$$

We will apply this lemma with $\frac{1}{p} \ll y \ll \frac{1}{p} \log^6 \frac{1}{p} \leq x$ and $k = \log^2 \frac{1}{p}$. Note that in this case the upper bound given by the lemma is not much larger than the lower bound given by Lemma 2.3. In particular, for these choices of x, y and k , the bound given by the lemma is of the form $((8 + o(1))p^2x)^y$.

We begin the proof of Lemma 3.1 with the following simple but important definition: let us say that a pair of sites \mathcal{P} is a *spanning pair* for the row $[a, b] \times \{c\}$ if

$$[a, b] \times \{c\} \subset \langle \mathcal{P} \cup [a, b] \times \{c - 1\} \rangle. \tag{3.1}$$

That is, \mathcal{P} is a spanning pair for $[a, b] \times \{c\}$ if the row becomes infected when \mathcal{P} and the row below it are infected. Note that for each spanning pair $\mathcal{P} = \{u, v\}$ there exists $z \in [a, b] \times \{c\}$ such that $u, v \in z + (\mathcal{N}_{(1,2)} \setminus \{(0, -1)\})$, and thus that any spanning pair is a translate of one of the eight pairs on the right-hand side of Fig. 2.

Lemma 3.2 *Let R be a rectangle such that R has $\mathbf{x}(R) \geq 2$ and $\mathbf{y}(R) \geq 1$, and let $S \subset R$. Then R is up-traversable by S if and only if $\langle S \rangle$ contains a spanning pair for every row of R .*

Proof Suppose that $R = [a, b] \times [c, d]$ with $b - a \geq 1$ and $d - c \geq 0$. It is easy to see that if $\langle S \rangle$ contains a spanning pair for every row of R , then R is up-traversable by S : if $\langle S \rangle$ contains a spanning pair for the bottom row of R , then the whole row becomes infected, i.e., $[a, b] \times \{c\} \subset \langle S \cup [a, b] \times \{c\} \rangle$. And given that the bottom row is infected, the row above the bottom row must also become infected, since $\langle S \rangle$ also contains a spanning pair for it, i.e., $[a, b] \times \{c + 1\} \subset \langle S \cup [a, b] \times \{c\} \rangle$. This argument can be repeated for all rows.

It will therefore suffice to prove that the converse holds. To do that, let $j \in [c, d]$ be the smallest j such that $\langle S \rangle$ does not contain a spanning pair for the row $[a, b] \times \{j\}$. We claim that the set

$$(\langle S \cup [a, b] \times \{j - 1\} \rangle \setminus \langle S \rangle) \cap ([a, b] \times \{j\})$$

is empty. Indeed, suppose that for some $t \geq 1$ there exists a site v such that

$$v \in \mathcal{B}^{(t)}(S \cup ([a, b] \times \{j - 1\})) \cap ([a, b] \times \{j\}),$$

Then there must be a pair of already-infected sites in $\mathcal{N}_{(1,2)}(v) \cap ([a, b] \times [j, j + 1])$ at time $t - 1$. But this pair lies in $\langle S \rangle$, and thus is a spanning pair for the row $[a, b] \times \{j\}$, a contradiction. Now, since $\langle S \rangle$ does not contain a spanning pair for $[a, b] \times \{j\}$, this implies that $R \not\subseteq \langle S \cup ([a, b] \times \{c - 1\}) \rangle$, as required. \square

We now make another important definition.

Definition 3.3 For each rectangle R such that $\mathbf{x}(R) \geq 2$ and $\mathbf{y}(R) \geq 1$, and each set $S \subset R$ such that R is up-traversable by S , let $\mathcal{A}(S) \subset S$ be a minimum-size subset such that R is up-traversable by $\mathcal{A}(S)$. (If more than one such subset exists, then choose one according to some arbitrary rule.) Define the *spanning time*

$$\begin{aligned} \tau &= \tau(R, S) \\ &:= \min \{t \geq 0 : \mathcal{B}^{(t)}(\mathcal{A}(S)) \text{ contains a spanning pair for each row of } R\}. \end{aligned}$$

In words, the spanning time τ is the first time t such that $\mathcal{B}^{(t)}(\mathcal{A}(S))$ spans all rows of R . Since R is up-traversable by $\mathcal{A}(S)$, it follows by Lemma 3.2 that τ must be finite. However, we emphasise that it is possible that $\tau > 0$, see Fig. 3 for some examples.

The central idea in the proof of Lemma 3.1 is to consider the cases $\tau = 0$ and $\tau > 0$ separately. When $\tau = 0$, the structure is significantly simpler than when $\tau > 0$, which allows for a very sharp estimate. When $\tau > 0$ more complex structures are possible, but more infected sites are required, and this allows us to use a less precise analysis.

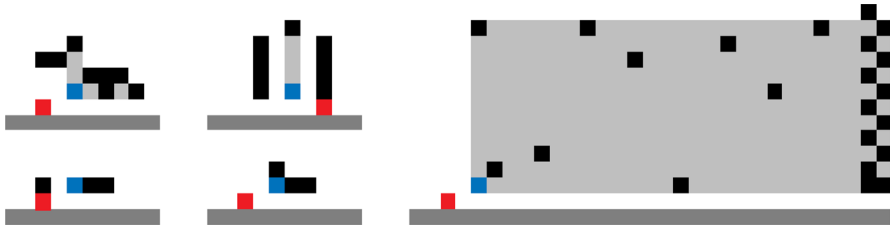


Fig. 3 Five configurations (the red and black sites) that do not have a spanning pair for the row above the dark grey row at time $t = 0$, but that create a spanning pair (the red and blue sites) at some time $t > 0$ by iteration with \mathcal{B} . The light grey sites indicate which sites must become infected to create the spanning pair. Note that in each case these sets have minimal cardinality (i.e., if we remove any black site, then iteration of \mathcal{B} will not create the spanning pair) (color figure online)

3.1 The case $\tau = 0$

Given a rectangle R , let $\mathcal{F}_0(R)$ and $\mathcal{F}_+(R)$ denote the families of all minimal sets $\mathcal{A} \subset R$ such that R is up-traversable by \mathcal{A} and $\tau(R, \mathcal{A}) = 0$ and $\tau(R, \mathcal{A}) > 0$, respectively. Let us write $\mathcal{U}_0(R)$ and $\mathcal{U}_+(R)$ for the upsets generated by $\mathcal{F}_0(R)$ and $\mathcal{F}_+(R)$, respectively, i.e., the collections of subsets of R that contain a set $\mathcal{A} \in \mathcal{F}_0(R)$ or $\mathcal{A} \in \mathcal{F}_+(R)$, respectively.

The following lemma gives a precise estimate of the probability that a rectangle is up-traversable and $\tau = 0$.

Lemma 3.4 *Let R be a rectangle with dimensions (x, y) , and let $p \in (0, 1)$. Then*

$$\mathbb{P}_p(\mathcal{S} \cap R \in \mathcal{U}_0(R)) \leq (8p^2x + 8p)^y. \tag{3.2}$$

We will prove Lemma 3.4 using the first moment method. To be precise, we will show that the expected number of members of $\mathcal{F}_0(R)$ that are contained in \mathcal{S} is at most the right-hand side of (3.2). This will follow easily from the following lemma.

Lemma 3.5 *Let R be a rectangle with dimensions (x, y) , and let $p \in (0, 1)$. Then*

$$|\mathcal{F}_0(R)| \leq \sum_{r=1}^y 8^y \binom{y-1}{r-1} x^r.$$

To count the sets in $\mathcal{F}_0(R)$, we will need to understand their structure. We will show that each set $\mathcal{A} \in \mathcal{F}_0(R)$ can be partitioned into “paths” as follows:

Lemma 3.6 *Let R be a rectangle with dimensions (x, y) , and let $\mathcal{A} \in \mathcal{F}_0(R)$. Then there exists a partition $\mathcal{A} = A_1 \cup \dots \cup A_r$, where $r = |\mathcal{A}| - y$, with the following property: For each $j \in [r]$, there exists an ordering $(u_1, \dots, u_{|A_j|})$ of the elements of A_j such that*

$$u_i - u_{i-1} \in \{(\pm 2, 1), (\pm 1, 1)\},$$

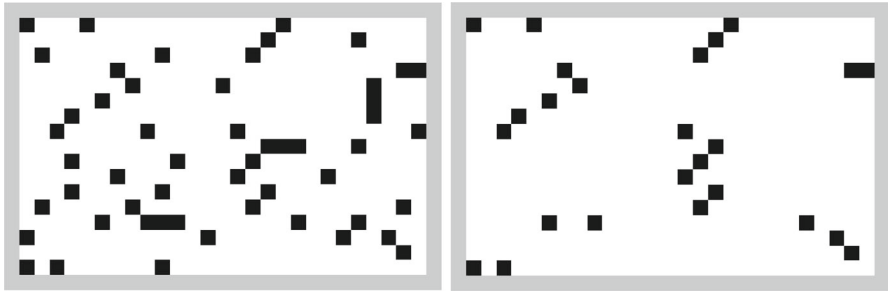


Fig. 4 On the left: an up-traversable rectangle. On the right: a minimal set \mathcal{A} . Note that \mathcal{A} is sufficient for up-traversability, that \mathcal{A} spans every row (so $\tau = 0$), and that \mathcal{A} consists of 8 paths (so $r = 8$)

for each $2 \leq i < |A_j|$, and

$$u_{|A_j|} - u_{|A_j|-1} \in \{(\pm 4, 0), (\pm 3, 0), (\pm 2, 0), (\pm 1, 0), (\pm 2, 1), (\pm 1, 1)\}.$$

See Fig. 4 for an illustration.

Proof Since \mathcal{A} is a minimal subset of R such that R is up-traversable by \mathcal{A} , and $\tau(R, \mathcal{A}) = 0$, it follows from Definition 3.3 that \mathcal{A} contains a spanning pair for each row of R , and hence (by minimality of \mathcal{A}) it follows that \mathcal{A} consists exactly of a union of spanning pairs (one pair for each row) and no other sites. Let these pairs be $\mathcal{P}_1, \dots, \mathcal{P}_y$, and define a graph on $[y]$ by placing an edge between i and j if $\mathcal{P}_i \cap \mathcal{P}_j$ is non-empty. The sets A_1, \dots, A_r are simply (the elements of \mathcal{A} corresponding to) the components of this graph.

Let the components of the graph be C_1, \dots, C_r , and note first that each component is a path, since a spanning pair for row $[a, b] \times \{c\}$ is contained in $[a, b] \times [c, c + 1]$. Moreover, it follows immediately from this simple fact that if $\mathcal{P}_i \cap \mathcal{P}_j$ is non-empty then \mathcal{P}_i and \mathcal{P}_j must be spanning pairs for adjacent rows (say, $[a, b] \times \{c\}$ and $[a, b] \times \{c + 1\}$), and that their common element must lie in $[a, b] \times \{c + 1\}$.

Now, consider a component $C_\ell = \{i_1, \dots, i_s\}$, set $A_\ell = \bigcup_{j=1}^s \mathcal{P}_{i_j}$, and note that $|A_\ell| = s + 1$. Let $A_\ell = \{u_1, \dots, u_{s+1}\}$, and assume (without loss of generality) that $\mathcal{P}_{i_j} = \{u_j, u_{j+1}\}$ for each $j \in \{1, \dots, s\}$. It now follows from the comments above, and the definition of a spanning pair in (3.1), that

$$u_i - u_{i-1} \in \{(\pm 2, 1), (\pm 1, 1)\},$$

for each $2 \leq i \leq s$, and that

$$u_{s+1} - u_s \in \{(\pm 4, 0), (\pm 3, 0), (\pm 2, 0), (\pm 1, 0), (\pm 2, 1), (\pm 1, 1)\},$$

as claimed. Finally, note that $|\mathcal{A}| = y + r$, since $|A_\ell| = |C_\ell| + 1$ for each $\ell \in \{1, \dots, r\}$. □

Proof of Lemma 3.5 To count the sets $\mathcal{A} \in \mathcal{F}_0(R)$, let us first fix $|\mathcal{A}|$, and the sizes of the sets A_1, \dots, A_r given by Lemma 3.6. Recall that $r = |\mathcal{A}| - y$ and that $\mathcal{A} =$

$A_1 \cup \dots \cup A_r$ is a partition, and note that $|A_j| \geq 2$ for each $j \in \{1, \dots, r\}$, since A_j is a union of spanning pairs. It follows that there are exactly

$$\binom{y-1}{r-1}$$

ways to choose the sequence $(|A_1|, \dots, |A_r|)$, where we order the sets A_j so that if $i < j$ then the top row of A_i is no higher than the bottom row of A_j . (Note that this is possible because each A_i is a union of spanning pairs for some set of consecutive rows of R .) Now, we claim that there are at most

$$x \cdot 8^{|A_j|-1}$$

ways of choosing the elements of $|A_j|$, given A_{j-1} and $|A_j|$. Indeed, given A_{j-1} we can deduce which is the bottom row of A_j , and we have at most x choices for the left-most element u_1 of A_j in that row. If $|A_j| = 2$ then (given u_1) there are then exactly 8 choices for the other element u_2 , since $u_2 - u_1 \in \{(4, 0), (3, 0), (2, 0), (1, 0), (\pm 2, 1), (\pm 1, 1)\}$. On the other hand, if $|A_j| \geq 3$, then there are at most $4^{|A_j|-2} \cdot 12 \leq 8^{|A_j|-1}$ choices for the remaining elements of A_j (given u_1), by Lemma 3.6, as required.

Now, multiplying together the (conditional) number of choices for each set A_j , it follows that

$$|\mathcal{F}_0(R)| \leq \sum_{r=1}^y \sum_{|A_1|, \dots, |A_r|} \prod_{j=1}^r (x \cdot 8^{|A_j|-1}) \leq \sum_{r=1}^y 8^y \binom{y-1}{r-1} x^r,$$

as claimed, since $\sum_{j=1}^r (|A_j| - 1) = y$. □

Lemma 3.4 now follows by Markov’s inequality:

Proof of Lemma 3.4 Define a random variable X to be the number of sets $\mathcal{A} \in \mathcal{F}_0(R)$ that are entirely infected at time zero, i.e., that are contained in our p -random set \mathcal{S} . By Markov’s inequality and Lemma 3.5, we have

$$\begin{aligned} \mathbb{P}_p(\mathcal{S} \cap R \in \mathcal{U}_0(R)) &\leq \mathbb{E}_p[X] \leq \sum_{r=1}^y 8^y \binom{y-1}{r-1} x^r p^{y+r} \\ &= \sum_{r=1}^y \binom{y-1}{r-1} (8p^2x)^r (8p)^{y-r} = \frac{px}{1+px} (8p^2x + 8p)^y \end{aligned}$$

as required. □

3.2 The case $\tau > 0$

In this section we analyse the event $\mathcal{S} \cap R \in \mathcal{U}_+(R)$. If R is up-traversable by \mathcal{S} , then let \mathcal{A} again denote a subset of \mathcal{S} of minimal cardinality such that R is up-traversable

by \mathcal{A} . By Lemma 3.2 above we know that if R is up-traversable by \mathcal{A} , then there must exist a time t at which there is spanning pair in $\mathcal{B}^{(t)}(\mathcal{A})$ for each row in R . The following definition isolates the sites that are responsible for the creation of such spanning pairs.

Definition 3.7 Given \mathcal{S} and a row ℓ , we say that $\mathcal{M} \subset \mathcal{S}$ is an *infector of the row ℓ* if

- there exists a $t \geq 0$ such that $\mathcal{B}^{(t)}(\mathcal{M})$ contains a spanning pair for the row ℓ , and
- there does not exist a subset $\mathcal{M}' \subset \mathcal{M}$ such that there exists a $t' \geq 0$ such that $\mathcal{B}^{(t')}(\mathcal{M}')$ contains a spanning pair for the row ℓ .

We call the bottom-most left-most site in \mathcal{M} the *root* of \mathcal{M} . Given \mathcal{S} we write $\mathbb{M}(\mathcal{S}, R)$ for the set of all infectors contained in \mathcal{S} for a row of R .

Note that spanning pairs are infectors, but that many other configurations are possible: see Fig. 3 for a few examples.

Lemma 3.8 (A property of the union of infectors) *Suppose $R = [1, x] \times [1, y]$ is up-traversable by \mathcal{S} and that \mathcal{A} is a subset of \mathcal{S} of minimal cardinality with the same property. For each $\ell \in \{1, \dots, y\}$ there exists an infector \mathcal{M}_ℓ of row ℓ in \mathcal{S} such that*

$$\bigcup_{\ell=1}^y \mathcal{M}_\ell = \mathcal{A}.$$

Proof Let \mathcal{A}' be a subset of \mathcal{S} such that R is up-traversable by \mathcal{A}' and such that \mathcal{A}' is a set with minimal cardinality for this property. By Lemma 3.2, the event that R is up-traversable by \mathcal{A}' is equivalent to the event that there exists a spanning pair for each row of R after some finite number of iterations of \mathcal{A}' by the bootstrap operator \mathcal{B} . This means that for each row \mathcal{A}' contains at least one infector. Note that it is a priori possible that the infectors in $\mathbb{M}(\mathcal{A}', R)$ overlap partially or that an infector for some row ℓ is contained in an infector for a row $\ell' \neq \ell$. Write $(\mathcal{M}^{(i)})_{i=1}^{|\mathbb{M}(\mathcal{A}', R)|}$ for some (arbitrary) ordered list of the infectors, and, for $1 \leq s \leq |\mathbb{M}(\mathcal{A}', R)|$ write

$$\mathbb{M}^b(s) := \bigcup_{i=1}^{s-1} \mathcal{M}^{(i)} \cup \bigcup_{i=s+1}^{|\mathbb{M}(\mathcal{A}', R)|} \mathcal{M}^{(i)}$$

for the union of the sites of all the infectors except those of $\mathcal{M}^{(s)}$. Now suppose that there exist $1 \leq s < t \leq |\mathbb{M}(\mathcal{A}', R)|$ such that $\mathcal{M}^{(s)}, \mathcal{M}^{(t)}$ are both infectors of the same row ℓ and suppose that $\mathcal{M}^{(s)} \setminus \mathbb{M}^b(s) \neq \emptyset$ and $\mathcal{M}^{(t)} \setminus \mathbb{M}^b(t) \neq \emptyset$. Then, since $\mathcal{M}^{(s)}$ is an infector for row ℓ and the sites in $\mathcal{M}^{(t)} \setminus \mathbb{M}^b(t)$ are not needed to create a spanning pair for any other row, R is also up-traversable by the set $\mathcal{A}' \setminus (\mathcal{M}^{(t)} \setminus \mathbb{M}^b(t))$, whose cardinality is strictly smaller than \mathcal{A}' . This gives a contradiction. Hence, for each row ℓ there must exist at most one infector $\mathcal{M}^{(s)}$ with the property that $\mathcal{M}^{(s)} \setminus \mathbb{M}^b(s) \neq \emptyset$. Taking their union we obtain \mathcal{A} (i.e., $\mathcal{A} = \mathcal{A}'$). □

Recall that for any set $\mathcal{Q} \subset \mathbb{Z}^2$ we write $\mathbf{x}(\mathcal{Q})$ and $\mathbf{y}(\mathcal{Q})$ for the horizontal and vertical dimensions of that set. We split the event $\{\mathcal{S} \cap R \in \mathcal{U}_+(R)\}$ according to whether there exists an infector \mathcal{M}_ℓ with $\mathbf{x}(\mathcal{M}_\ell) \geq 6k^2$ or not.

Lemma 3.9 (Wide infectors) *Let $R = [1, x] \times [1, k]$ with $k \geq 3$ such that $k \ll p^{-1}$, and x such that $k^5 \ll x \leq p^{-2}$, then*

$$\mathbb{P}_p\left(S \cap R \in \mathcal{U}_+(R), \max_{\ell=1}^k \mathbf{x}(\mathcal{M}_\ell) \geq 6k^2\right) = o((8p^2x + 8p)^k). \tag{3.3}$$

Proof Write \mathcal{M}_j for the first infector such that $\mathbf{x}(\mathcal{M}_j) \geq 6k^2$. Since $\mathcal{M}_j \subset [1, x] \times [1, k]$, $\mathbf{y}(\mathcal{M}_j) \leq k$. Moreover, \mathcal{M}_j is the minimal set responsible for the creation of the spanning pair in row j , so it must be the case that \mathcal{M}_j does not have a gap of more than three consecutive columns. There are at most xk possible positions for the root of the infector. We thus bound (3.3) for the range of x and our choice of k from above by

$$\begin{aligned} xk(1 - (1 - p)^{3k})^{2k^2} &\leq xk(3pk)^{2k^2} \ll (3pk)^{6k-3} \ll (800p^3k^3)^k \\ &\ll (8p^2x + 8p)^k. \end{aligned}$$

□

Lemma 3.10 (Small infectors) *There exist no infectors that are not a single spanning pair that intersect precisely one row, and there exist precisely two infectors that are not a single spanning pair that intersect precisely two rows, up to translations. The cardinality of these infectors is 4, and they span both rows they intersect.*

Proof Let \mathcal{M}_j be the infector for some row j . Write v for an element of the spanning pair for row j that becomes infected due to the bootstrap dynamics on \mathcal{M}_j . (It is easy to see that only one element of a spanning pair can arise after time $t = 0$, but we do not use this fact.) Suppose t is the first time such that $\mathcal{B}^{(t)}(\mathcal{M}_j)$ contains a spanning pair. Because \mathcal{M}_j is not a spanning pair, $t \geq 1$. Since v becomes infected at time t , it must be the case that $|\mathcal{N}_{(1,2)}(v) \cap \mathcal{B}^{(t-1)}(\mathcal{M}_j)| \geq 3$. Any configuration of three sites in $\mathcal{N}_{(1,2)}(v)$ contains a spanning pair for the row that v is in, so v cannot be in row j . By the definition of spanning pairs, (3.1), a site can either span the row that it is in, or the row below it, so v is in row $j + 1$. We conclude that there are no infectors that are not a spanning pair that intersect precisely one row.

By the same argument, if $t \geq 2$, then \mathcal{M}_j must contain a site in row $j + 2$, so only infectors that intersect two rows can have $t = 1$.

One can easily verify that the only infectors with $t = 1$ that intersect two rows are translations of the configurations $\{(0, 0), (0, 1), (3, 1), (4, 1)\}$ and $\{(0, 0), (0, 1), (-3, 1), (-4, 1)\}$ (see the configuration in the bottom-left corner of Fig. 3). These infectors both have cardinality 4, and span both rows they intersect. □

To analyse $\mathbb{P}_p(S \cap R \in \mathcal{U}_+(R))$ we again divide \mathcal{A} into the maximal number of disjoint, “causally independent” pieces, to which we may apply the BK-inequality. We have seen that when $\tau = 0$ these pieces can be described as paths. When $\tau > 0$ this is still the case, but now the path structure can be found at the level of the infectors. We partition \mathcal{A} as follows: let r be the largest integer such that there exist sets B_1, \dots, B_r that partition \mathcal{A} (i.e., $B_i \cap B_j = \emptyset$ for all $i \neq j$ and $\mathcal{A} = \bigcup_{i=1}^r B_i$) and such that there exist r pairs of integers $\{(a_i, b_i)\}_{i=1}^r$ such that

- $1 = a_1 \leq b_1 \leq a_2 \leq b_2 \leq \dots \leq a_r \leq b_r = k$, and
- the event

$$\{[1, x] \times [a_1, b_1] \text{ is up-trav by } B_1\} \circ \dots \circ \{[1, x] \times [a_r, b_r] \text{ is up-trav by } B_r\}$$

occurs.

Lemma 3.11 (Path structure of B_1, \dots, B_r) *Let $R = [1, x] \times [1, y]$ and suppose that R is up-traversable by \mathcal{S} . Let \mathcal{A} be the subset of \mathcal{S} with minimal cardinality such that R is up-traversable by \mathcal{A} . Let B_1, \dots, B_r be the division of \mathcal{A} into disjointly occurring pieces described above. Then the following hold:*

- (a) For any row $\ell \in \{1, \dots, y\}$ there exists a unique $i \in \{1, \dots, r\}$ such that $\mathcal{M}_\ell \subseteq B_i$.
- (b) If B_i spans rows $\ell, \dots, \ell + m$, then $B_i = \cup_{j=\ell}^{\ell+m} \mathcal{M}_j$.
- (c) If $\mathcal{M}_j \subseteq B_i$ and $j < b_i$, then at least one of the following holds: $\mathcal{M}_j = B_i$; or there exists a $j' < j$ such that $\mathcal{M}_j \subset \mathcal{M}_{j'} \subseteq B_i$; or $\mathcal{M}_j \cap \mathcal{M}_{j+1} \neq \emptyset$.
- (d) If $\mathcal{M}_j \subseteq B_i$ and $j = b_i$, then at least one of the following holds: $\mathcal{M}_j = B_i$; or there exists a $j' < j$ such that $\mathcal{M}_j \subset \mathcal{M}_{j'} \subseteq B_i$; or $\mathcal{M}_{j-1} \cap \mathcal{M}_j \neq \emptyset$.

Proof (a) By construction, $\mathcal{A} = \cup_{i=1}^r B_i$, and $B_i \circ B_j$ occurs if $i \neq j$. By Lemma 3.8, $\mathcal{A} = \cup_{\ell=1}^k \mathcal{M}_\ell$. Suppose that there exists an ℓ such that $\mathcal{M}_\ell \cap B_i \neq \emptyset$ and $\mathcal{M}_\ell \cap B_j \neq \emptyset$ for some $i \neq j$. Without loss of generality, we can further assume that $a_i \leq \ell \leq b_i$. Since \mathcal{M}_ℓ is the minimal set to create a spanning pair for row ℓ , and that $\mathcal{M}_\ell \cap B_i$ is a strict subset of \mathcal{M}_ℓ (since the latter intersects B_j , which is disjoint from B_i by assumption), we deduce that $\langle \mathcal{M}_\ell \cap B_i \rangle$ cannot contain a spanning pair for row ℓ . By Lemma 3.2, this means that $[1, x] \times [a_i, b_i]$ is not up-traversable by B_i , which is a contradiction.

- (b) By Lemma 3.8, $\mathcal{A} = \cup_{i=1}^k \mathcal{M}_i$. Combined with (a) this gives (b).
- (c) Suppose that B_i spans rows $\ell, \dots, \ell + m$ and suppose that there exists a $j < b_i$ such that neither $\mathcal{M}_j = B_i$ nor $\mathcal{M}_j \subset \mathcal{M}_{j'}$ for some $j' < j$, and such that $\mathcal{M}_j \cap \mathcal{M}_{j+1} = \emptyset$. Then we can partition

$$B_i = \left(\bigcup_{s=\ell}^j \mathcal{M}_s \right) \sqcup \left(\bigcup_{t=j+1}^{\ell+m} \mathcal{M}_t \right) =: B_{i,1} \sqcup B_{i,2}.$$

It then follows that

$$\{[1, x] \times [\ell, j] \text{ is up-trav by } B_{i,1}\} \circ \{[1, x] \times [j + 1, \ell + m] \text{ is up-trav by } B_{i,2}\}$$

occurs. This gives a contradiction, since by construction the sets B_1, \dots, B_r are the maximal partition of \mathcal{A} with this property, so such a j does not exist. So we conclude that if $\mathcal{M}_j \subset B_i$ but $\mathcal{M}_j \neq B_i$ and $\mathcal{M}_j \not\subseteq \mathcal{M}_{j'}$ for some $j' < j$, then $\mathcal{M}_j \cap \mathcal{M}_{j+1} \neq \emptyset$.

- (d) The proof is identical to that of (c), mutatis mutandis.

□

For all $k, \ell, m, x \in \mathbb{N}$, let $\mathcal{E}_{\ell+1, \ell+m}$ denote the event that a configuration of infected sites \mathcal{S} has the following properties:

- $\mathcal{S} \cap ([1, x] \times [\ell + 1, \ell + m]) \in \mathcal{U}_+([1, x] \times [\ell + 1, \ell + m])$,
- the minimal subset \mathcal{A} of \mathcal{S} such that $[1, x] \times [\ell + 1, \ell + m]$ is up-traversable by \mathcal{A} cannot be divided into two or more disjointly occurring pieces, i.e., $\mathcal{A} = B_1$ in the construction described above.
- $\max_{j=\ell+1}^m \mathbf{x}(\mathcal{M}_j) < 6k^2$.

Lemma 3.12 For $k \geq 3, \ell + m \leq k$ and all $p \in [0, 1]$,

$$\mathbb{P}_p(\mathcal{E}_{\ell+1, \ell+m}) \leq 4p^2x(12pk^2 + 7p)^{m-1}.$$

Proof There is at least one infected site in row $\ell + 1$, and it can be at x positions.

By Lemma 3.11, the event $\mathcal{E}_{\ell+1, \ell+m}$ implies that \mathcal{A} is the union of infectors that are not disjoint. Since, moreover, none of the infectors are wider than $6k^2 - 1$, for each of the rows $\ell + 2, \dots, \ell + m$ we then need to have at least 1 infected site in the line-segment $[-6k^2 - 3, 6k^2 + 3]$ directly above the infected site of the row below it. Finally, row $\ell + m$ must also be spanned, and by Lemma 3.10 its spanning pair must already be present at time $t = 0$, so there must be another infected site in that row, in one of the four positions that can create a spanning pair for line $\ell + m$. We thus bound

$$\mathbb{P}_p(\mathcal{E}_{\ell+1, \ell+m}) \leq px \cdot (p(12k^2 + 7))^{m-1}4p.$$

□

Write

$$\mathcal{V}_{a,b} := \{[1, x] \times [a, b] \text{ is up-traversable}\}$$

and

$$\mathcal{V}_{a,b}^+ := \{\mathcal{S} \cap ([1, x] \times [a, b]) \in \mathcal{U}_+([1, x] \times [a, b])\} \cap \{\max_{j=a}^b \mathbf{x}(\mathcal{M}_j) \leq 6k^2\}.$$

The following lemma states the key inequality for the induction:

Lemma 3.13 For $k \geq 2$,

$$\mathbb{P}_p(\mathcal{V}_{1,k}^+) \leq \sum_{m=2}^k \sum_{\ell=0}^{k-m} \mathbb{P}_p(\mathcal{V}_{1,\ell}) \mathbb{P}_p(\mathcal{E}_{\ell+1, \ell+m}) \mathbb{P}_p(\mathcal{V}_{\ell+m+1,k}).$$

Proof Since $\mathcal{V}_{1,k}^+$ occurs, $[1, x] \times [1, k]$ is up-traversable. Let \mathcal{A} be the minimal subset of \mathcal{S} such that $[1, x] \times [1, k]$ is up-traversable with respect to \mathcal{A} . Let B_1, \dots, B_r be the subdivision of \mathcal{A} described above. Let $u \in \mathcal{A}$ and $v \in \langle \mathcal{A} \rangle \setminus \mathcal{A}$ be such that $\{u, v\}$ form a spanning pair for the row i , while \mathcal{A} does not contain a spanning pair for row i . At least one such pair must exist since $\mathcal{V}_{1,k}^+$ occurs. Let j be such that $\mathcal{M}_i \subseteq B_j$

(we can find such a B_j by Lemma 3.11(a)). Suppose that B_j spans exactly the rows $\ell + 1, \dots, \ell + m$ (i.e., $a_j = \ell + 1$ and $b_j = \ell + m$). Then, by the construction of B_1, \dots, B_r and $\mathcal{E}_{\ell+1, \ell+m}$ we know that

$$\mathcal{V}_{1, \ell} \circ \mathcal{E}_{\ell+1, \ell+m} \circ \mathcal{V}_{\ell+m+1, k}$$

occurs for \mathcal{S} . Applying the BK-inequality and summing over ℓ and m gives the asserted inequality. The sum over m starts at 2 because by Lemma 3.10, B_j must span at least two rows. □

3.3 The proof of Lemma 3.1

To begin, assume that $\frac{3k^2}{p} \leq x \leq \frac{1}{p^2}$. We start by proving Lemma 3.1 for the cases where $y \leq k$. More precisely, we will prove that

$$\mathbb{P}_p(\mathcal{V}_{1, k}) \leq e(8p^2x + 8p)^k, \tag{3.4}$$

holds for $k \ll p^{-1}$. We use induction. The inductive hypothesis is that (3.4) holds for $k' \leq k - 1$ and $k^5 \ll x \leq p^{-2}$. To initialise the induction we observe that when $k = 1$ there exist four spanning pairs up to translations that intersect one row, so $\mathbb{P}_p(\mathcal{V}_{1, 1}) \leq 4p^2x < e(8p^2x + 8p)$. When $k = 2$ we use Lemma 3.10 to bound

$$\mathbb{P}_p(\mathcal{V}_{1, 2}^+) = 2xp^4, \tag{3.5}$$

which, combined with Lemma 3.4 yields that

$$\mathbb{P}_p(\mathcal{V}_{1, 2}) \leq (8p^2x + 8p)^2 + 4xp^4 \leq e(8p^2x + 8p)^2$$

when p is sufficiently small.

When $3 \leq k \ll p^{-1}$, by (3.5), Lemmas 3.4, 3.9, 3.12, and 3.13, and the induction hypothesis (3.5), when p is sufficiently small,

$$\begin{aligned} \mathbb{P}_p(\mathcal{V}_{1, k}) &\leq (8p^2x + 8p)^k + \frac{e-1}{3}(8p^2x + 8p)^k + 2xp^4(k-1)e^2(8p^2x + 8p)^{k-2} \\ &\quad + 4e^2xp^2k \sum_{m=3}^k (8p^2x + 8p)^{k-m+1} (12pk^2 + 7p)^{m-1}, \end{aligned} \tag{3.6}$$

where the second term on the right-hand side is due to Lemma 3.9, and the third and fourth correspond to the $m = 2$ and $m \geq 3$ terms in Lemma 3.13.

It is not difficult to show that

$$\sum_{m=3}^k a^{k-m+1} b^{m-1} = \frac{b^2 a^{k-1} - ab^k}{a - b}.$$

When $\frac{3k^2}{p} \leq x$ we have $12pk^2 + 7p \leq \frac{1}{2}(8p^2x + 8p)$, so this implies that

$$\begin{aligned} & 4e^2xp^2k \sum_{m=3}^k (8p^2x + 8p)^{k-m+1} (12pk^2 + 7p)^{m-1} \\ & \leq 8e^2xp^2(12pk^2 + 7p)^2(8p^2x + 8p)^{k-2}. \end{aligned}$$

Inspecting (3.6), it follows that the desired bound (3.4) holds if the following two inequalities hold for p sufficiently small:

$$\begin{aligned} 2e^2xp^4(k-1) &< \frac{e-1}{3}(8p^2x + 8p)^2, \\ 8e^2xp^2k(12pk^2 + 7p)^2 &< \frac{e-1}{3}(8p^2x + 8p)^2. \end{aligned}$$

The first inequality holds because $k \ll x$. It is easy to verify that the second inequality holds when $k^5 \ll x \leq p^{-2}$. Substituting the above inequalities into (3.6) proves the claim of Lemma 3.1 for $y \leq k$.

Now we consider $R = [1, x] \times [1, y]$ for y such that $k < y < x$ (still assuming that $\frac{3k^2}{p} \leq x \leq \frac{1}{p^2}$). We cover R with $\lceil y/k \rceil$ rectangles of height k . If y is not divisible by k the covering “overshoots”: it includes at most $k-1$ rows that are not in R . If R is up-traversable, and if the overshoot contains a connected upward path, then all these rectangles are also up-traversable. The probability that there is a connected path in the overshoot is at least p^k . It thus follows by the BK-inequality that

$$\begin{aligned} \mathbb{P}_p(R \text{ is up-trav}) &\leq p^{-k} \prod_{n=1}^{\lceil y/k \rceil} \mathbb{P}_p(\mathcal{V}_{(n-1)k+1, nk}) \\ &\leq p^{-k} e^{y/k} (8p^2x + 8p)^y, \end{aligned} \quad (3.7)$$

where these bounds again hold for p sufficiently small. This completes the proof of Lemma 3.1 for the case $\frac{3k^2}{p} \leq x \leq \frac{1}{p^2}$.

The case $x < \frac{3k^2}{p}$ is now easy. Note that if $[1, x] \times [1, y]$ is up-traversable by \mathcal{S} , then also $[1, x+a] \times [1, y]$ for any $a \geq 1$ is up-traversable by \mathcal{S} (i.e., up-traversability is a monotone increasing event in the width of the rectangle). Hence, $\mathbb{P}_p([x] \times [y] \text{ is up-trav})$ is a monotone increasing function in x . The bound thus follows by choosing $x = \frac{3k^2}{p}$ and applying the bound for the case $\frac{3k^2}{p} \leq x \leq \frac{1}{p^2}$. \square

4 The probability of simultaneous horizontal and vertical growth

The lemma below states an upper bound on the probability of an infected rectangle growing both vertically and horizontally, i.e., an upper bound on $\mathbb{P}_p(R \Rightarrow R')$ for certain $R \subset R'$.

Let

$$\xi := \left\lceil \log^2 \frac{1}{p} \right\rceil, \quad \text{and} \quad \delta_\xi := 1 - 2/\xi.$$

Recalling Lemma 3.1 and the bound on $f(p, y)$ in Lemma 2.2 above, let

$$\begin{aligned} \psi(x) := & - \left(\log(24p\xi^2 + 8p) + \xi^{-1} \right) \mathbb{1}_{\left\{x < \frac{3\xi^2}{p}\right\}} \\ & - \left(\log(8p^2x + 8p) + \xi^{-1} \right) \mathbb{1}_{\left\{\frac{3\xi^2}{p} \leq x \leq \frac{1}{p^2}\right\}}, \end{aligned} \tag{4.1}$$

and let

$$\phi(y) := e^{-3py} \mathbb{1}_{\left\{y > \frac{4}{p} \log \log \frac{1}{p}\right\}}. \tag{4.2}$$

For two rectangles $R \subset R'$ with dimensions (x, y) and $(x + s, y + t)$, let

$$U^P(R, R') := \delta_\xi (t\psi(x + s) + s\phi(y + t)). \tag{4.3}$$

Observe that ψ and ϕ are both positive, decreasing, and convex functions (where they are not zero).

Lemma 4.1 *Let $R \subset R'$, with dimensions (x, y) and $(x + s, y + t)$ respectively. Assume that $t \leq \frac{1}{p} \log^{-4} \frac{1}{p}$. Then, for p sufficiently small,*

$$\mathbb{P}_p(R \Rightarrow R') \leq 2p^{-\xi} \exp(-U^P(R, R')).$$

The proof uses a similar strategy as [23, Proof of Proposition 3.3]. Roughly speaking this strategy entails that we “decorrelate” the horizontal and vertical growth events needed for $\{R \Rightarrow R'\}$.

Proof If $y + t \leq \frac{4}{p} \log \log \frac{1}{p}$ and $x + s > 1/p^2$, then we use the trivial bound $\mathbb{P}_p(R \Rightarrow R') \leq 1$, corresponding to $U^P(R, R') = 0$, as required.

If $y + t \leq \frac{4}{p} \log \log \frac{1}{p}$ and $x + s \leq 1/p^2$, then we apply Lemma 3.1 (with $k = \xi$), again giving the required bound.

Therefore, we assume henceforth that $y + t > \frac{4}{p} \log \log \frac{1}{p}$ and $x + s \leq \frac{1}{p^2}$.

To start, suppose that $(1 - \delta_\xi)t\psi(x + s) > \delta_\xi s\phi(y + t)$, which corresponds to the vertical growth component t being disproportionately large compared to the horizontal growth component s . Then, we can simply ignore the horizontal growth and apply Lemma 3.1 to bound

$$\begin{aligned} \mathbb{P}_p(R \Rightarrow R') & \leq p^{-\xi} \exp(-t\psi(x + s)) \\ & \leq p^{-\xi} \exp(-\delta_\xi t\psi(x + s) - \delta_\xi s\phi(y + t)) \\ & = p^{-\xi} \exp(-U^P(R, R')), \end{aligned} \tag{4.4}$$

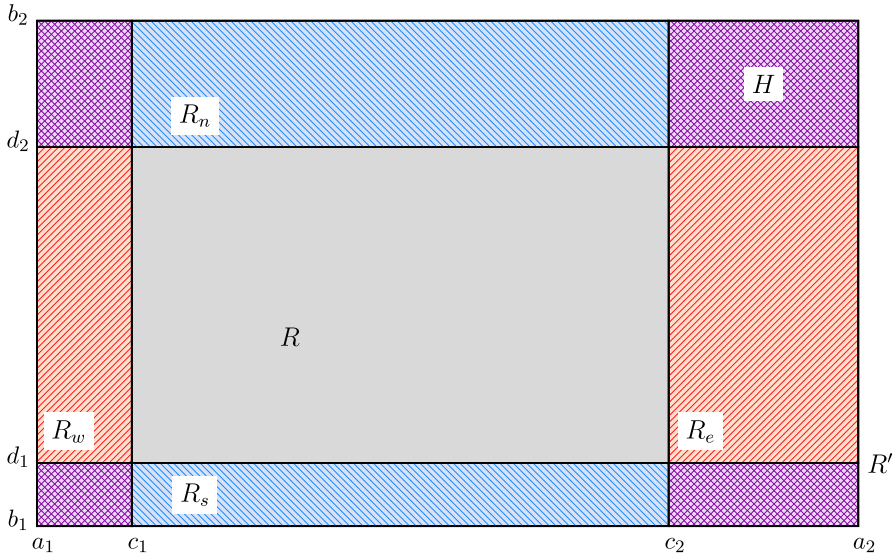


Fig. 5 The rectangles R_w and R_e are hatched red and the rectangles R_n and R_s are hatched blue. The regions where these four rectangles overlap, collectively called H , are cross hatched purple (color figure online)

and we are done. Therefore, let us henceforth also assume that

$$(1 - \delta_\xi)t\psi(x + s) \leq \delta_\xi s\phi(y + t). \tag{4.5}$$

We identify five (intersecting) regions within the area $R' \setminus R$: the North, South, West, and East regions R_n , R_s , R_w , and R_e , and the corner region H : for $R' = [a_1, a_2] \times [b_1, b_2]$ and $R = [c_1, c_2] \times [d_1, d_2]$, such that $a_1 \leq c_1 < c_2 \leq a_2$ and $b_1 \leq d_1 < d_2 \leq b_2$, we define the sets

$$\begin{aligned} R_w &:= [a_1, c_1 - 1] \times [b_1, b_2] & \text{and} & & R_e &:= [c_2 + 1, a_2] \times [b_1, b_2], \\ R_n &:= [a_1, a_2] \times [d_2 + 1, b_2] & \text{and} & & R_s &:= [a_1, a_2] \times [b_1, d_1 - 1], \\ H &:= R' \setminus \{(x, y) : x \in [c_1, c_2] \text{ or } y \in [d_1, d_2]\}, \end{aligned}$$

see Fig. 5. Observe that

$$\{R \Rightarrow R'\} \subset \{R_n \text{ is up-trav}\} \cap \{R_s \text{ is down-trav}\} \cap \{R_w \text{ and } R_e \text{ are hor-trav}\}.$$

Let

$$E := \{R_n \text{ is up-trav}\} \cap \{R_s \text{ is down-trav}\}.$$

Recall from Definition 3.7 above that we write $\mathbb{M}(S, R_n)$ and $\mathbb{M}(S, R_s)$ for the sets of infectors of R_n and R_s (the latter being a set of infectors suitably defined for

down-traversability). By Lemma 3.8, we are able to determine whether E occurs by inspecting only sites in $\mathbb{M}(S, R_n)$ and $\mathbb{M}(S, R_s)$. So the event that R_w and R_e are horizontally traversable only depends on E through the information about the intersection of these sets with H , the region where the rectangles overlap. Define $\mathbb{M}_H^b(S)$ as the set of all sites in $S \cap H$ contained in either $\mathbb{M}(S, R_n)$ or $\mathbb{M}(S, R_s)$. Let Y denote the number of columns in H that contain at least one infected site in $\mathbb{M}_H^b(S)$. We split

$$\mathbb{P}_p(R \Rightarrow R') \leq \mathbb{P}_p(\{R \Rightarrow R'\} \cap \{Y \leq s/(2\xi)\}) + \mathbb{P}_p(Y > s/(2\xi)). \tag{4.6}$$

We start by bounding the first term in (4.6). Let $F := \{Y \leq s/(2\xi)\}$. We use Lemma 3.1 with $k = \xi = \lceil \log^2 \frac{1}{p} \rceil$ to bound

$$\begin{aligned} \mathbb{P}_p(\{R \Rightarrow R'\} \cap F) &\leq \mathbb{P}_p(R_e \text{ and } R_w \text{ are hor-trav} \mid E \cap F) \mathbb{P}_p(E) \\ &\leq \mathbb{P}_p(R_e \text{ and } R_w \text{ are hor-trav} \mid E \cap F) p^{-\xi} e^{-t\psi(x+s)}. \end{aligned} \tag{4.7}$$

Let \mathfrak{R}_n denote the set of all sets of n subrectangles of $R_e \cup R_w$ with heights $y + t$, total width n , and such that each pair of rectangles in a set $\tau \in \mathfrak{R}_n$ are separated by at least one column. I.e., for $\tau = \{\tau_i\}_{i=1}^{N(\tau)} \in \mathfrak{R}_n$ we have that τ is a collection of $N(\tau)$ strictly disjoint subrectangles with $\sum_{i=1}^{N(\tau)} x(\tau_i) = n$, and $y(\tau_i) = y + t$ for all $1 \leq i \leq N(\tau)$. For any $\tau = \{\tau_i\}_{i=1}^{N(\tau)} \in \mathfrak{R}_n$ define the following two events:

$$E_1(\tau) := \{\forall 1 \leq i \leq N(\tau) : \tau_i \text{ is horizontally traversable}\}$$

and

$$\begin{aligned} E_2(\tau) &:= \left\{ \mathbb{M}_H^b \cap \bigcup_{i=1}^{N(\tau)} \tau_i = \emptyset \right\} \\ &\cap \left\{ \nexists \tau' \in \mathfrak{R}_n : N(\tau') < N(\tau) \text{ and } \mathbb{M}_H^b \cap \bigcup_{i=1}^{N(\tau')} \tau'_i = \emptyset \right\} \\ &\cap \left\{ \nexists \tau'' \in \cup_{n'' > n} \mathfrak{R}_{n''} : \mathbb{M}_H^b \cap \bigcup_{i=1}^{N(\tau'')} \tau''_i = \emptyset \right\}, \end{aligned} \tag{4.8}$$

that is, $E_2(\tau)$ is the event that τ is the partition into the least number of rectangles of total width n that does not intersect \mathbb{M}_H^b , and that there is no partition of total width greater than n that also does not intersect \mathbb{M}_H^b . Observe that

$$\{R_e, R_w \text{ are hor-trav}\} \subseteq \bigsqcup_{m=0}^s \bigsqcup_{\tau \in \mathfrak{R}_{s-m}} (E_1(\tau) \cap E_2(\tau)).$$

Thus,

$$\begin{aligned} &\mathbb{P}_p(R_e, R_w \text{ are hor-trav} \mid E \cap F) \\ &\leq \sum_{m=0}^{s/(2\xi)} \sum_{\tau \in \mathfrak{R}_{s-m}} \mathbb{P}_p(E_1(\tau) \mid E_2(\tau) \cap E \cap F) \mathbb{P}_p(E_2(\tau) \mid E \cap F), \end{aligned} \tag{4.9}$$

where we used that the sum may be restricted to $m \leq s/(2\xi)$ by the conditioning on F . Now note that the events E and F can be verified by inspecting only \mathbb{M}_H^b , which, on the event $E_2(\tau)$ is contained in $H \setminus \tau$, while $E_1(\tau)$ by definition only depends on the sites in τ , so that conditionally on $E_2(\tau)$ the event $E_1(\tau)$ is independent of E and F . We may thus write

$$\begin{aligned} \mathbb{P}_p(E_1(\tau) \mid E_2(\tau) \cap E \cap F) &= \frac{\mathbb{P}_p(E_1(\tau) \cap E \cap F \mid E_2(\tau))}{\mathbb{P}_p(E \cap F \mid E_2(\tau))} \\ &= \frac{\mathbb{P}_p(E_1(\tau) \mid E_2(\tau)) \mathbb{P}_p(E \cap F \mid E_2(\tau))}{\mathbb{P}_p(E \cap F \mid E_2(\tau))} \\ &= \mathbb{P}_p(E_1(\tau) \mid E_2(\tau)). \end{aligned}$$

Observe that for any fixed τ the event $E_1(\tau)$ is increasing. Indeed, adding more sites to \mathcal{S} can either make horizontal traversal occur when it did not before, or else, have no effect. We claim that the event $E_2(\tau)$, on the other hand, is the intersection of three decreasing events, and hence itself a decreasing event. To see this, observe that the first event in (4.8) is decreasing because adding more sites to \mathcal{S} cannot decrease the total width of \mathbb{M}_H^b , since it is the union of *all* infectors intersecting H (not only those of minimal cardinality for a given row). The second event in (4.8) is likewise decreasing, because increasing \mathbb{M}_H^b cannot decrease the minimal number of rectangles of a partition that does not intersect \mathbb{M}_H^b , unless it also decreases the total width of that partition. The third event is decreasing because increasing \mathbb{M}_H^b cannot decrease its total width. Therefore, we may apply the FKG-inequality to obtain

$$\mathbb{P}_p(E_1(\tau) \mid E_2(\tau)) \leq \mathbb{P}_p(E_1(\tau)),$$

and we may thus further bound the right-hand side of (4.9) by

$$\sum_{m=1}^{s/(2\xi)} \sum_{\tau \in \mathfrak{R}_{s-m}} \mathbb{P}_p(E_1(\tau)) \mathbb{P}_p(E_2(\tau) \mid E \cap F).$$

Uniformly for any fixed $\tau \in \mathfrak{R}_{s-m}$ with $m \leq s/(2\xi)$, by Lemma 2.2,

$$\begin{aligned} \mathbb{P}_p(E_1(\tau)) &= \prod_{i=1}^{N(\tau)} \mathbb{P}_p(\tau_i \text{ is hor-trav}) \leq \exp\left(-\sum_{i=1}^{N(\tau)} (\mathbf{x}(\tau_i) - 2)f(p, y + t)\right) \\ &\leq \exp(-s(1 - 1/\xi)f(p, y + t)) \\ &\leq \exp(-\delta_\xi s\phi(y + t)), \end{aligned}$$

where the final inequality follows from Lemma 2.2(b) when p is sufficiently small. Inserting this bound in (4.9), we proceed by using that the events $E_2(\tau)$ are mutually disjoint for all τ to bound

$$\begin{aligned} &\mathbb{P}_p(R_e, R_w \text{ are hor-trav} \mid E \cap F) \\ &\leq \exp(-\delta_\xi s\phi(y + t)) \sum_{m=1}^{s/(2\xi)} \sum_{\tau \in \mathfrak{R}_{s-m}} \mathbb{P}_p(E_2(\tau) \mid E \cap F) \\ &\leq \exp(-\delta_\xi s\phi(y + t)). \end{aligned} \tag{4.10}$$

Combining (4.7) and (4.10) we bound the first term in (4.6) by $p^{-\xi} \exp(-U^p(R, R'))$.

Now we bound the second term in (4.6). If $Y > s/(2\xi)$ then at least $s/(2\xi)$ out of s columns are non-empty. The probability that a column is non-empty is $1 - (1 - p)^t \leq 2pt$ (when p is sufficiently small). Therefore, $\mathbb{P}(Y > s/(2\xi)) \leq \mathbb{P}(\text{Bin}(s, 2pt) > s/(2\xi))$. We use Chernoff’s bound that $\mathbb{P}(\text{Bin}(n, p) > q) \leq e^{-q}$ when $q > np$ to estimate

$$\mathbb{P}_p(Y > s/(2\xi)) \leq \exp(-s/(2\xi))$$

(here we used that $t \leq \frac{1}{p} \log^{-4} \frac{1}{p}$). Observe that since $\xi = \lceil \log^2 \frac{1}{p} \rceil$, $\delta_\xi = 1 - \xi^{-1}$, and $\phi(y + t) > \log^{-12} \frac{1}{p}$ by our assumption that $y + t > \frac{4}{p} \log \log \frac{1}{p}$, we have

$$\exp(-s/(2\xi)) \leq \exp\left(-\frac{\delta_\xi}{1 - \delta_\xi} s\phi(y + t)\right).$$

Now recall our assumption (4.5) that $(1 - \delta_\xi)t\psi(x + s) \leq \delta_\xi s\phi(y + t)$. Applying this inequality twice, it follows that

$$\frac{\delta_\xi}{1 - \delta_\xi} s\phi(y + t) \geq t\psi(x + s) \geq \delta_x t\psi(x + s) + \delta_\xi s\phi(y + t).$$

We thus have $\mathbb{P}_p(Y > s/(2\xi)) \leq \exp(-U^p(R, R'))$, as required.

Applying the bounds for the two cases to (4.6) completes the proof (using the crude upper bound $p^{-\xi} + 1 \leq 2p^{-\xi}$ for p sufficiently small). □

5 The upper bound of Theorem 1.3

Proposition 5.1 *Let $p > 0$ and $\frac{1}{3p} \log \frac{1}{p} \leq y \leq \frac{1}{p} \log \frac{1}{p}$ and $\frac{1}{p^2} \leq x \leq \frac{1}{p^5}$. Then*

$$\mathbb{P}_p([x] \times [y] \text{ is IF}) \leq \exp\left(-\frac{2C_1}{p} \log^2 \frac{1}{p} + (2C_2 + o(1))\frac{1}{p} \log \frac{1}{p}\right).$$

5.1 Notation and definitions

Before we proceed with the proof, we must introduce some more notation and a few definitions. Our proof uses *hierarchies*. The notion of hierarchies is due to Holroyd [33], and is common to much of the bootstrap percolation literature since. Here we use a definition of a hierarchy that is similar to the one in [23]:

Definition 5.2 (Hierarchies).

- (a) HIERARCHY, SEED, NORMAL VERTEX, AND SPLITTER: A *hierarchy* \mathcal{H} is a rooted tree with out-degrees at most three¹¹ and with each vertex v labeled by non-empty rectangle R_v such that R_v contains all the rectangles that label the descendants of v . If the number of descendants of a vertex is 0, we call the vertex a *seed*.¹² If the vertex has one descendant, we call it a *normal* vertex, and we write $u \mapsto v$ to indicate that u is a normal vertex with (unique) descendant v . If the vertex has two or more descendants, we call it a *splitter* vertex. We write $N(\mathcal{H})$ for the number of vertices in the tree \mathcal{H} .
- (b) PRECISION: A hierarchy of *precision* Z (with $Z \geq 1$) is a hierarchy that satisfies the following conditions:
 - (1) If w is a seed, then $\mathbf{x}(R_w) \geq 2$ and $\mathbf{y}(R_w) < 2Z$, while if u is a normal vertex or a splitter, then $\mathbf{y}(R_u) \geq 2Z$.
 - (2) If u is a normal vertex with descendant v , then $\mathbf{y}(R_u) - \mathbf{y}(R_v) \leq 2Z$.
 - (3) If u is a normal vertex with descendant v and v is either a seed or a normal vertex, then $\mathbf{y}(R_u) - \mathbf{y}(R_v) > Z$.
 - (4) If u is a splitter with descendants v_1, \dots, v_i and $i \in \{2, 3\}$, then there exists $j \in \{1, \dots, i\}$ such that $\mathbf{y}(R_u) - \mathbf{y}(R_{v_j}) > Z$.
- (c) PRESENCE: Given a set of infected sites \mathcal{S} we say that a hierarchy \mathcal{H} is *present in* \mathcal{S} if all of the following events occur *disjointly*:
 - (1) For each seed w , $R_w = \langle R_w \cap \mathcal{S} \rangle$ (i.e., R_w is internally filled by \mathcal{S}).
 - (2) For each normal u and every v such that $u \mapsto v$, $R_u = \langle (R_v \cup \mathcal{S}) \cap R_u \rangle$ (i.e., the event $\{R_v \Rightarrow R_u\}$ occurs on \mathcal{S}).

¹¹ In the original construction of a hierarchy by Holroyd [33] for the standard model, hierarchies have out-degree at most two. The fact that we need out-degree three corresponds to the fact that the anisotropic model requires three infected sites in a neighbourhood. As a result, it is possible that a rectangle is internally filled by a set of three but not two disjoint internally filled smaller rectangles. Our definition of hierarchies reflects this. See also [23].

¹² Note that although similar, this definition of a seed is different than the one used in the previous section.

- (d) **GOODNESS:** Similar to [30], we say that a seed w is *large* if $Z/3 \leq \mathbf{y}(R_w) \leq Z$. We call a hierarchy *good* if it has at most $\log^{11} \frac{1}{p}$ large seeds, and we call it *bad* otherwise.

5.2 Outline of the proof of Proposition 5.1

In this section we give the proof of Proposition 5.1 subject to Lemma 5.9 below. We prove Lemma 5.9 in Sect. 6.

Let $\mathcal{H}_{Z,R}$ denote a hierarchy with root R and precision Z . Let $\mathbb{H}_{Z,R}$ denote the set of all $\mathcal{H}_{Z,R}$. Likewise, let $\mathbb{H}_{Z,R}^{\text{good}}$ and $\mathbb{H}_{Z,R}^{\text{bad}}$ denote the subsets of good and bad hierarchies in $\mathbb{H}_{Z,R}$. Lastly, given a set of hierarchies \mathbb{H} and a rectangle R , define the event

$$\mathcal{X}(R; \mathbb{H}) := \{\mathcal{S} \in \{0, 1\}^{\mathbb{Z}^2} : \exists \mathcal{H} \in \mathbb{H} \text{ such that } \mathcal{H} \text{ is present in } \mathcal{S} \cap R\}.$$

Lemma 5.3 *Let R be a rectangle with $\mathbf{x}(R) \geq 2$ and let $Z \geq 3$. If R is internally filled, then there exists a hierarchy $\mathcal{H}_{Z,R} \in \mathbb{H}_{Z,R}$ that is present, i.e., $\mathcal{X}(R; \mathbb{H}_{Z,R})$ occurs.*

The proof of this lemma is the same as the proof of [23, Proposition 3.8] so we do not repeat it here. (But note that it does not matter that our definition of hierarchies uses “internally filled” rather than “ k -occurs”.)

Throughout this paper, let

$$Z_p := \frac{1}{p} \log^{-8} \frac{1}{p}.$$

Conform the hypothesis of Proposition 5.1 we restrict ourselves to hierarchies with root label R_p of dimensions (x, y) such that

$$\frac{1}{p^2} \leq x \leq \frac{1}{p^5} \quad \text{and} \quad \frac{1}{3p} \log \frac{1}{p} \leq y \leq \frac{1}{p} \log \frac{1}{p}.$$

For the sake of simplicity we often suppress subscripts Z_p and R_p .

We bound the good and bad hierarchies separately:

$$\mathbb{P}_p(R_p \text{ is IF}) \leq \mathbb{P}_p(\mathcal{X}(R_p; \mathbb{H}^{\text{good}})) + \mathbb{P}_p(\mathcal{X}(R_p; \mathbb{H}^{\text{bad}})). \tag{5.1}$$

We bound the second term with the following lemma:

Lemma 5.4 *As p tends to 0 we have*

$$\mathbb{P}_p(\mathcal{X}(R_p; \mathbb{H}^{\text{bad}})) \leq \exp\left(-\Omega\left(\frac{1}{p} \log^3 \frac{1}{p}\right)\right).$$

Proof We claim that if R is a large seed, i.e., $Z_p/3 \leq \mathbf{y}(R) \leq Z_p$, then

$$\mathbb{P}_p(R \text{ is IF}) \leq \exp\left(-\Omega\left(\frac{1}{p} \log^{-8} \frac{1}{p}\right)\right).$$

To see that this is indeed the case we consider the cases $x \geq 1/p$ and $x < 1/p$ separately. For the case $x \geq 1/p$, the bound follows from Lemma 2.2(c):

$$\begin{aligned} \mathbb{P}_p(R \text{ is IF}) &\leq \mathbb{P}_p(R \text{ is hor-trav}) \leq \exp\left(- (x - 2) \left(\frac{1}{2}py - 3p^2y^2\right)\right) \\ &\leq \exp\left(- \left(\frac{1}{p} - 2\right) \left(\frac{1}{2}p \cdot \frac{1}{3p} \log^{-8} \frac{1}{p} - 3p^2 \cdot \frac{1}{p^2} \log^{-16} \frac{1}{p}\right)\right) \\ &\leq \exp\left(-\Omega\left(\frac{1}{p} \log^{-8} \frac{1}{p}\right)\right). \end{aligned}$$

For the case $x < 1/p$, the bound follows from Lemma 3.1 with $k = 2$ and p sufficiently small:

$$\begin{aligned} \mathbb{P}_p(R \text{ is IF}) &\leq \mathbb{P}_p(R \text{ is up-trav}) \leq p^{-2}e^{y/2}(104p)^y \\ &= \exp\left(2 \log \frac{1}{p} + \frac{y}{2} + y \log(108) - y \log \frac{1}{p}\right) \\ &= \exp\left(- (1 + o(1))y \log \frac{1}{p}\right) = \exp\left(-\Omega\left(\frac{1}{p} \log^{-8} \frac{1}{p}\right)\right). \end{aligned}$$

Now consider the event $\mathcal{X}(R_p; \mathbb{H}^{\text{bad}})$. This event implies that there exists a hierarchy \mathcal{H} that is present and bad, which by definition means that more than $\log^{11} \frac{1}{p}$ rectangles of size between $Z_p/3$ and Z_p are internally filled disjointly. Since R_p contains at most $\frac{1}{p^6} \log \frac{1}{p}$ sites, the probability of this event is smaller than

$$\left(\frac{1}{p^6} \log \frac{1}{p} \cdot e^{-\frac{c}{p} \log^{-8} \frac{1}{p}}\right)^{\log^{11} \frac{1}{p}} \leq \exp\left(-\Omega\left(\frac{1}{p} \log^3 \frac{1}{p}\right)\right).$$

□

We bound the first term of (5.1) as follows:

$$\mathbb{P}_p(\mathcal{X}(R_p; \mathbb{H}^{\text{good}})) \leq |\mathbb{H}^{\text{good}}| \max_{\mathcal{H} \in \mathbb{H}^{\text{good}}} \mathbb{P}_p(\mathcal{H} \text{ is present}). \tag{5.2}$$

Now we apply the following lemma.

Lemma 5.5 *The number of good hierarchies satisfies*

$$|\mathbb{H}^{\text{good}}| \leq e^{O(1/p)}.$$

Proof We start by observing that any good hierarchy $\mathcal{H} \in \mathbb{H}^{\text{good}}$ has root R_p such that $x \leq \frac{1}{p^5}$ and $\mathbf{y}(R_p) \leq \frac{1}{p} \log \frac{1}{p}$ and precision $Z_p = \frac{1}{p} \log^{-8} \frac{1}{p}$, so its height $h(\mathcal{H})$ is bounded from above by

$$h(\mathcal{H}) \leq \frac{\mathbf{y}(R_p)}{Z_p} \leq \log^9 \frac{1}{p}.$$

Moreover, since there are at most $\log^{11} \frac{1}{p}$ large seeds in a good hierarchy, the number of vertices $N_{\mathcal{H}}$ in the hierarchy \mathcal{H} obeys

$$N_{\mathcal{H}} \leq \log^{11} \frac{1}{p} \cdot \log^9 \frac{1}{p} = \log^{20} \frac{1}{p}. \tag{5.3}$$

Each vertex of a hierarchy has 0, 1, 2 or 3 descendants, so there are at most $4^{\log^{20} \frac{1}{p}}$ unlabelled trees corresponding to the good hierarchies. Finally, since each vertex of a hierarchy is labelled by a sub-rectangle of R_p with $\mathbf{x}(R_p) \leq p^{-5}$, the number of choices for each label is bounded from above by

$$\mathbf{x}(R_p)^2 \cdot \mathbf{y}(R_p)^2 \leq p^{-10} \cdot \frac{1}{p^2} \log^2 \frac{1}{p} \leq p^{-13},$$

so

$$|\mathbb{H}^{\text{good}}| \leq \left(4p^{-13}\right)^{\log^{20} \frac{1}{p}} = e^{13 \log^{21} \frac{1}{p} + \log 4 \log^{20} \frac{1}{p}} \leq e^{O(1/p)}.$$

□

By Lemma 5.5 it suffices to give a uniform bound on the probability that a given hierarchy is present, if the hierarchy is good. Indeed, it remains to show that

$$\max_{\mathcal{H} \in \mathbb{H}^{\text{good}}} \mathbb{P}_p(\mathcal{H} \text{ is present}) \leq \exp\left(-\frac{2C_1}{p} \log^2 \frac{1}{p} + (2C_2 + o(1)) \frac{1}{p} \log \frac{1}{p}\right).$$

Before we proceed, let us deal with a small technical issue: the possibility of “wide” seeds (Lemma 5.9 below does not work in their presence). Observe that if the hierarchy \mathcal{H} that maximises the probability of being present contains a seed with label R_s such that $\mathbf{x}(R_s) > \frac{1}{3p} \log^{12} \frac{1}{p}$ (i.e., the seed is extremely wide), then the probability that \mathcal{H} is present is bounded by the probability that R_s is horizontally traversable, which, by Lemma 2.2(c) can be bounded as follows:

$$\begin{aligned} \mathbb{P}_p(R_s \text{ is IF}) &\leq \mathbb{P}_p(R_s \text{ is hor-trav}) \\ &\leq \exp\left(-(\mathbf{x}(R_s) - 2) \left(\frac{1}{2} p \mathbf{y}(R_s) - 3p^2 \mathbf{y}(R_s)^2\right)\right) \end{aligned}$$

$$\begin{aligned} &\leq \exp\left(-\left(\frac{1}{p} \log^{12} \frac{1}{p} - 2\right)\left(\frac{1}{2} \log^{-8} \frac{1}{p} - 3 \log^{-16} \frac{1}{p}\right)\right) \\ &= \exp\left(-\Omega\left(\frac{1}{p} \log^3 \frac{1}{p}\right)\right), \end{aligned}$$

where for the third inequality we used the assumption on $\mathbf{x}(R_s)$ and that $\mathbf{y}(R_s) < 2Z = \frac{2}{p} \log^{-8} \frac{1}{p}$. Proposition 5.1 thus holds for hierarchies with wide seeds. Let us therefore assume from here on that $\mathbf{x}(R_s) \leq \frac{1}{3p} \log^{12} \frac{1}{p}$ for all seeds.

By the BK-inequality we have

$$\mathbb{P}_p(\mathcal{H} \text{ is present}) \leq \prod_{u \text{ seed}} \mathbb{P}_p(R_u \text{ is IF}) \prod_{v \mapsto w} \mathbb{P}_p(R_v \Rightarrow R_w) \tag{5.4}$$

(we ignore here the contributions from splitter vertices).

The following lemma is used to determine a bound for the product of the seeds:

Lemma 5.6 *Given a hierarchy $\mathcal{H}_{Z,R}$, let N_{seed} denote the number of seeds of the hierarchy $\mathcal{H}_{Z,R}$, and let $u_1, \dots, u_{N_{\text{seed}}}$ be an arbitrary ordering of the seeds of $\mathcal{H}_{Z,R}$. Then*

$$\prod_{u \text{ seed}} \mathbb{P}_p(R_u \text{ is IF}) \leq \prod_{n=1}^{N_{\text{seed}}} \mathbb{P}_p(\tilde{R}_n \Rightarrow \tilde{R}_{n+1})$$

where

$$\tilde{R}_n = [1, \mathbf{x}(R_{u_1}) + \dots + \mathbf{x}(R_{u_n})] \times [1, \mathbf{y}(R_{u_1}) + \dots + \mathbf{y}(R_{u_n})].$$

Proof For any $p > 0$ and any rectangle R' with dimensions (x, y) with $\min\{x, y\} \geq 2$ and any $a \geq 2, b \geq 1$, we have

$$\mathbb{P}_p(R' \text{ is IF}) \leq \mathbb{P}_p([1, a] \times [1, b] \Rightarrow [1, a + x] \times [1, b + y]). \tag{5.5}$$

Indeed, if the rectangle $[a + 1, a + x] \times [b + 1, b + y]$ is internally filled, then the event $\{[1, a] \times [1, b] \Rightarrow [1, a + x] \times [1, b + y]\}$ occurs. An application of the FKG-inequality thus gives (5.5).

Any seed of a hierarchy must have dimensions at least $(2, 1)$ by definition, so an iterated application of (5.5) completes the proof. \square

Recall the definition of $U^p(R, R')$ in (4.3) above. We use Lemmas 4.1 and 5.6 to bound the first product on the right-hand side of (5.4):

$$\begin{aligned} \prod_{u \text{ seed}} \mathbb{P}_p(R_u \text{ is IF}) &\leq \prod_{n=1}^{N_{\text{seed}}} \mathbb{P}_p(\tilde{R}_n \Rightarrow \tilde{R}_{n+1}) \\ &\leq 2p^{-\xi N_{\text{seed}}} \exp\left(-\sum_{n=0}^{N_{\text{seed}}} U^p(\tilde{R}_n, \tilde{R}_{n+1})\right), \end{aligned}$$

where $\tilde{R}_0 = [1] \times [1]$.

To bound the second product of (5.4), we use the following lemma:

Lemma 5.7 *Let $p > 0$. Let N_{splitter} denote the number of splitter vertices of the hierarchy \mathcal{H} . Then there exists an integer $\hat{N} = \hat{N}(\mathcal{H}) \geq 1$ and a sequence of nested rectangles $\hat{R}_0 \subset \dots \subset \hat{R}_{\hat{N}}$ with the following properties:*

- $\hat{R}_0 = \tilde{R}_{N_{\text{seed}}}$ (with $\tilde{R}_{N_{\text{seed}}}$ as defined in Lemma 5.6 above),
- $\hat{R}_{\hat{N}}$ has dimensions larger than R ,
- $\mathbf{y}(\hat{R}_{n+1}) - \mathbf{y}(\hat{R}_n) \leq \frac{1}{p} \log^{-8} \frac{1}{p}$ for every $0 \leq n \leq \hat{N} - 1$,
- for p sufficiently small,

$$\prod_{v \mapsto w} \mathbb{P}_p(R_w \Rightarrow R_v) \leq 2p^{-\xi N_{\text{splitter}}} \prod_{n=0}^{\hat{N}-1} \exp\left(-U^p(\hat{R}_n, \hat{R}_{n+1})\right).$$

The proof of this lemma goes by induction, using Lemma 4.1, and it is essentially the same as the proof of [23, Lemma 3.11], so we omit it here.

We use Lemma 5.7 to determine that there exist rectangles $\hat{R}_1 \subset \dots \subset \hat{R}_{\hat{N}}$ satisfying the conditions of the lemma such that

$$\prod_{v \mapsto w} \mathbb{P}_p(R_v \Rightarrow R_w) \leq 2p^{-\xi N_{\text{splitter}}} \prod_{n=0}^{\hat{N}-1} \exp\left(-U^p(\hat{R}_n, \hat{R}_{n+1})\right).$$

Using Lemmas 4.1 and 5.7 and writing $(R_n)_{n=0}^N$ for the concatenation of the sequences $(\tilde{R}_n)_{n=0}^{N_{\text{seed}}}$ and $(\hat{R}_n)_{n=0}^{\hat{N}}$, i.e.,

$$(R_n)_{n=0}^N := (\tilde{R}_0, \tilde{R}_1, \dots, \tilde{R}_{N_{\text{seed}}}, \hat{R}_1, \dots, \hat{R}_{\hat{N}})$$

with $N := N_{\text{seed}} + \hat{N}$, we bound

$$\mathbb{P}_p(\mathcal{H} \text{ is present}) \leq 4p^{-\xi(N_{\text{seed}} + N_{\text{splitter}})} \exp\left(-\sum_{n=0}^{N-1} U^p(R_n, R_{n+1})\right). \tag{5.6}$$

To bound the first factor in (5.6) we use the following lemma:

Lemma 5.8 *Any good hierarchy satisfies*

$$4p^{-\xi(N_{\text{seed}} + N_{\text{splitter}})} \leq e^{O(1/p)}.$$

Proof By (5.3) there are at most $\log^{20} \frac{1}{p}$ vertices in a good hierarchy, and $\xi = \lceil \log^2 \frac{1}{p} \rceil$, so for any $\mathcal{H} \in \mathbb{H}^{\text{good}}$,

$$4p^{-\xi(N_{\text{seed}} + N_{\text{splitter}})} \leq 4p^{-\lceil \log \frac{1}{p} \rceil^{22}} = 4\left(e^{\log \frac{1}{p}}\right)^{\lceil \log \frac{1}{p} \rceil^{22}} \leq e^{O(1/p)}.$$

□

The final ingredient of the proof is the following lemma:

Lemma 5.9 *Let $\mathcal{R}_N = \{R_n\}_{n=0}^N$ be a sequence of increasing, nested rectangles such that*

- $R_0 = [1] \times [1]$,
- $y_1 \leq \frac{2}{p} \log^{-8} \frac{1}{p}$ and $\frac{1}{3p} \log \frac{1}{p} \leq y_N \leq \frac{1}{p} \log \frac{1}{p}$,
- $x_1 \leq \frac{1}{3p} \log^{12} \frac{1}{p}$ and $x_N \geq \frac{1}{p^2}$.

Then

$$\exp\left(-\sum_{n=0}^{N-1} U^p(R_n, R_{n+1})\right) \leq \exp\left(-\frac{2C_1}{p} \log^2 \frac{1}{p} + (2C_2 + o(1)) \frac{1}{p} \log \frac{1}{p}\right).$$

The proof involves a longer computation, so we defer it to Sect. 6.

Proof of Proposition 5.1 subject to Lemma 5.9 We combine the above lemmas and the bounds derived in the discussion to conclude that

$$\begin{aligned} \mathbb{P}_p(R \text{ is IF}) &\leq \mathbb{P}_p(\mathcal{X}(R_p; \mathbb{H}^{\text{good}})) + \mathbb{P}_p(\mathcal{X}(R_p; \mathbb{H}^{\text{bad}})) \quad \text{[Lemma 5.3 and (5.1)]} \\ &\stackrel{\text{[(5.2) \& Lemma 5.4]}}{\leq} |\mathbb{H}^{\text{good}}| \max_{\mathcal{H} \in \mathbb{H}^{\text{good}}} \mathbb{P}_p(\mathcal{H} \text{ is present}) + e^{-\Omega\left(\frac{1}{p} \log^3 \frac{1}{p}\right)} \\ &\stackrel{\text{[(5.4) and Lemma 5.5]}}{\leq} e^{O(1/p)} \prod_{u \text{ seed}} \mathbb{P}_p(R_u \text{ is IF}) \prod_{v \mapsto w} \mathbb{P}_p(R_v \Rightarrow R_w) + e^{-\Omega\left(\frac{1}{p} \log^3 \frac{1}{p}\right)} \\ &\stackrel{\text{[Lemma 5.6]}}{\leq} e^{O(1/p)} \prod_{n=1}^{N_{\text{seed}}} \mathbb{P}_p(\tilde{R}_n \Rightarrow \tilde{R}_{n+1}) \prod_{v \mapsto w} \mathbb{P}_p(R_v \Rightarrow R_w) + e^{-\Omega\left(\frac{1}{p} \log^3 \frac{1}{p}\right)} \\ &\stackrel{\text{[Lemmas 4.1 and 5.7]}}{\leq} e^{O(1/p)} 2p^{-\xi N_{\text{seed}}} \exp\left(-\sum_{n=0}^{N_{\text{seed}}} U^p(\tilde{R}_n, \tilde{R}_{n+1})\right) \\ &\quad \times 2p^{-\xi N_{\text{splitter}}} \exp\left(-\sum_{n=0}^{\hat{N}-1} U^p(\hat{R}_n, \hat{R}_{n+1})\right) + e^{-\Omega\left(\frac{1}{p} \log^3 \frac{1}{p}\right)} \\ &\stackrel{\text{[Lemma 5.8]}}{\leq} e^{O(1/p)} \exp\left(-\sum_{n=0}^{N-1} U^p(R_n, R_{n+1})\right) + e^{-\Omega\left(\frac{1}{p} \log^3 \frac{1}{p}\right)} \\ &\stackrel{\text{[Lemma 5.9]}}{\leq} \exp\left(-\frac{2C_1}{p} \log^2 \frac{1}{p} + (2C_2 + o(1)) \frac{1}{p} \log \frac{1}{p}\right), \end{aligned}$$

as claimed. □

It remains to prove Lemma 5.9. We do this in the upcoming section.

6 Variational principles: proof of Lemma 5.9

To prove Lemma 5.9 we will start by setting up some variational principles, similar to [33, Sect. 6]. We start with a few general lemmas.

Assume throughout this section that $f(x)$ and $g(y)$ are positive, non-increasing, convex, Riemann-integrable functions. Let $\mathbb{R}_+ = (0, \infty)$ and for $\underline{a} = (a_1, a_2) \in \mathbb{R}_+^2$ and $\underline{b} = (b_1, b_2) \in \mathbb{R}_+^2$, write $\underline{a} \leq \underline{b}$ if $a_1 \leq b_1$ and $a_2 \leq b_2$. For $\underline{a}, \underline{b} \in \mathbb{R}_+^2$ with $\underline{a} \leq \underline{b}$ and any path γ from \underline{a} to \underline{b} , define

$$w_{f,g}(\gamma) := \int_{\gamma} f(x)dy + g(y)dx, \tag{6.1}$$

and

$$W_{f,g}(\underline{a}, \underline{b}) := \inf_{\gamma: \underline{a} \rightarrow \underline{b}} \int_{\gamma} f(x)dy + g(y)dx. \tag{6.2}$$

To start, an elementary lemma:

Lemma 6.1 *If $\underline{a} \leq \underline{b} \leq \underline{c}$, then $W_{f,g}(\underline{a}, \underline{b}) + W_{f,g}(\underline{b}, \underline{c}) \geq W_{f,g}(\underline{a}, \underline{c})$.*

The proof is easy (see [33, Sect. 6]).

Let

$$\Delta_{f,g} := \left\{ (x, y) \in \mathbb{R}_+^2 : f'(x) \neq 0, g'(y) \neq 0, f'(x) = g'(y) \right\}. \tag{6.3}$$

Note that since $f(x)$ and $g(y)$ are assumed to be convex decreasing functions, $\Delta_{f,g}$ describes a simple curve in $[a, b] \times [c, d] \subset \mathbb{R}_+^2$ if $f'(x) \neq 0$ for all $x \in [a, b]$ and $g'(y) \neq 0$ for all $y \in [c, d]$.

For sets $A, B \subseteq \mathbb{R}_+^2$ we say that A lies Northwest of B and we write $A \succcurlyeq B$ if for any $\underline{a} \in A$ and any $\underline{b} \in B$ that satisfy $a_1 + a_2 = b_1 + b_2$ we have $a_2 \geq b_2$.

Lemma 6.2 *If γ_1 and γ_2 are paths from \underline{a} to \underline{b} , and we have either $\gamma_1 \succcurlyeq \gamma_2 \succcurlyeq \Delta_{f,g}$ or $\Delta_{f,g} \succcurlyeq \gamma_2 \succcurlyeq \gamma_1$, then $w_{f,g}(\gamma_1) \geq w_{f,g}(\gamma_2)$.*

Proof To start, assume that $\gamma_1 \succcurlyeq \gamma_2 \succcurlyeq \Delta_{f,g}$. Let H be the region between γ_1 and γ_2 :

$$H := \{ \underline{u} : \underline{a} \leq \underline{u} \leq \underline{b} \text{ and } \gamma_1 \succcurlyeq \{ \underline{u} \} \succcurlyeq \gamma_2 \}.$$

By Green’s Theorem in the plane we have

$$w_{f,g}(\gamma_1) - w_{f,g}(\gamma_2) = \iint_H (g'(y) - f'(x)) \, dx dy.$$

Now, since $\gamma_1 \succcurlyeq \gamma_2 \succcurlyeq \Delta_{f,g}$ we have $H \succcurlyeq \Delta_{f,g}$, and since moreover f and g are convex decreasing functions, we have $g'(y) - f'(x) \geq 0$ for all $(x, y) \in H$. It follows that $w_{f,g}(\gamma_1) - w_{f,g}(\gamma_2) \geq 0$.

By the same reasoning we have $w_{f,g}(\gamma_1) - w_{f,g}(\gamma_2) \geq 0$ when $\Delta_{f,g} \succcurlyeq \gamma_2 \succcurlyeq \gamma_1$. □

Lemma 6.3 For $\underline{a}, \underline{b} \in \Delta_{f,g}$ with $\underline{a} \leq \underline{b}$, let $\gamma_0 := \Delta_{f,g} \cap ([a_1, b_1] \times [a_2, b_2])$, then $W_{f,g}(\underline{a}, \underline{b}) = w_{f,g}(\gamma_0)$.

Proof Suppose by contradiction that $\gamma_1 \neq \gamma_0$ is a minimiser of $W_{f,g}(\underline{a}, \underline{b})$, and γ_0 is not. Then γ_1 must intersect γ_0 in at least two points (counting \underline{a} and \underline{b} as intersection points as well). So we can find a set of disjoint curves $\{\eta_i^1\}$ with $\eta_i^1 \subset \gamma_1$ and a set of disjoint curves $\{\eta_i^0\}$ with $\eta_i^0 \subset \gamma_0$ so that $\gamma_1 \setminus \cup_i \eta_i^1 = \gamma_0 \setminus \cup_i \eta_i^0$ and $\eta_i^1 \succ \eta_i^0 \succ \Delta_{f,g}$ or $\Delta_{f,g} \succ \eta_i^0 \succ \eta_i^1$ for each i , and so that η_i^0 and η_i^1 have the same end-points. By Lemma 6.2, replacing the curve η_i^1 by η_i^0 in γ_1 does not increase the value of the line integral. Repeating this procedure for each such interval, we end up replacing the minimiser γ_1 by γ_0 without increasing the value of the integral, contradicting the assumption that γ_0 was not a minimiser. \square

Given a set of points $\{\underline{a}_{(i)}\}$, with $\underline{a}_{(i)} \in \mathbb{R}_+^2$, we write $\underline{a}_{(1)} \rightarrow \underline{a}_{(2)} \rightarrow \dots \rightarrow \underline{a}_{(n)}$ for the path that linearly interpolates between successive points $\underline{a}_{(i)}$ and $\underline{a}_{(i+1)}$. Given a path γ and two points $\underline{a}, \underline{b} \in \gamma$, we write $\underline{a} \xrightarrow{\gamma} \underline{b}$ for the part of γ between \underline{a} and \underline{b} .

Lemma 6.4 If $g(y) = c$ for some constant $c \geq 0$ and $f(x)$ is a positive, monotone decreasing function, then for $\underline{a} \leq \underline{b}$,

$$W_{f,g}(\underline{a}, \underline{b}) = c(b_1 - a_1) + f(b_1)(b_2 - a_2),$$

and the path from \underline{a} to \underline{b} that minimises $W_{f,g}(\underline{a}, \underline{b})$ is $\underline{a} \rightarrow (b_1, a_2) \rightarrow \underline{b}$.

Proof This follows directly from the definition of $W_{f,g}$ and the assumptions on f and g . \square

Recall the definitions of $\psi(y)$ and $\phi(x)$ from (4.1) to (4.2), and the definition of $\Delta_{f,g}$ in (6.3). Observe that

$$\psi'(x) = \frac{1}{x} \mathbb{1}_{\left\{\frac{3\xi^2}{p} \leq x \leq \frac{1}{p^2}\right\}} \quad \text{and} \quad \phi'(y) = -3pe^{-3py} \mathbb{1}_{\left\{y > \frac{4}{p} \log \log \frac{1}{p}\right\}},$$

so $\psi'(x) = \phi'(y)$ is solved by

$$x(y) = \frac{e^{3py}}{3p} \quad \text{and} \quad y(x) = \frac{1}{3p} \log(3px)$$

when both $\psi(x) \neq 0$ and $\phi(y) \neq 0$. Observe that

$$\begin{aligned} y\left(\frac{3\xi^2}{p}\right) &= \frac{1}{3p} \log(9\xi^2) < \frac{4}{p} \log \log \frac{1}{p}, \\ y\left(\frac{1}{p^2}\right) &= \frac{1}{3p} \log \frac{3}{p}, \\ x\left(\frac{4}{p} \log \log \frac{1}{p}\right) &= \frac{1}{3p} \log^{12} \frac{1}{p}. \end{aligned}$$

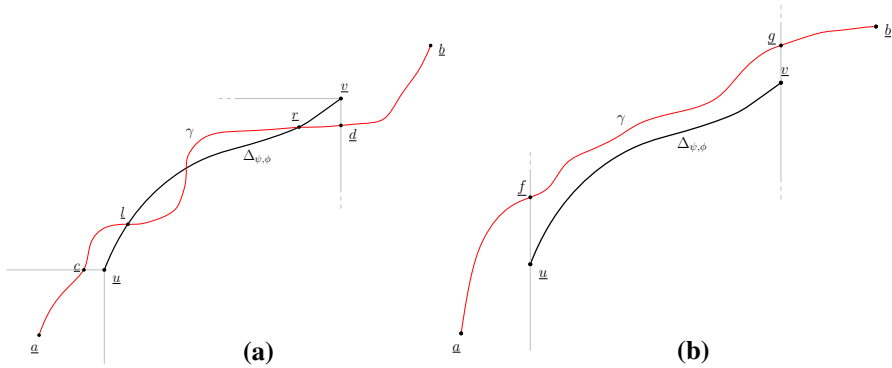


Fig. 6 The partitioning of γ used in Lemma 6.5 for the case $\gamma \cap \Delta_{\psi,\phi} \neq \emptyset$ in (a), and for the case $\gamma \cap \Delta_{\psi,\phi} = \emptyset$ in (b)

We can thus write

$$\begin{aligned} \Delta_{\psi,\phi} &= \left\{ (x, y) \in \mathbb{R}_+^2 : \psi'(x) \neq 0, \phi'(y) \neq 0, \psi'(x) = \phi'(y) \right\} \\ &= \left\{ \left(\frac{e^{3py}}{3p}, y \right) : y \in \left(\frac{4}{p} \log \log \frac{1}{p}, \frac{1}{3p} \log \frac{3}{p} \right] \right\}. \end{aligned}$$

The leftmost and rightmost points of $\Delta_{\psi,\phi}$ are given by

$$\underline{u} = \left(\frac{1}{3p} \log^{12} \frac{1}{p}, \frac{4}{p} \log \log \frac{1}{p} \right) \quad \text{and} \quad \underline{v} = \left(\frac{1}{p^2}, \frac{1}{3p} \log \frac{3}{p} \right). \tag{6.4}$$

Lemma 6.5 *Let \underline{u} and \underline{v} be as in (6.4), and let \underline{a} be such that $\underline{a} \leq \underline{u}$, and let \underline{b} be such that $\underline{v} \leq \underline{b}$. Then*

$$W_{\psi,\phi}(\underline{a}, \underline{b}) = W_{\psi,\phi}(\underline{a}, \underline{u}) + W_{\psi,\phi}(\underline{u}, \underline{v}) + W_{\psi,\phi}(\underline{v}, \underline{b}).$$

Proof By Lemma 6.1, the right-hand side is an upper bound on $W_{\psi,\phi}(\underline{a}, \underline{b})$. It remains to prove that it is also a lower bound.

Since ψ and ϕ are decreasing, positive, continuous functions, any path that minimises $W_{\psi,\phi}(\underline{a}, \underline{b})$ must be a coordinate-wise increasing path. Fix a coordinate-wise increasing path $\gamma \subset \mathbb{R}_+^2$ from \underline{a} to \underline{b} . Then either

- (a) $\gamma \cap \Delta_{\psi,\phi} \neq \emptyset$, or
- (b) $\gamma \cap \Delta_{\psi,\phi} = \emptyset$.

Consider first case (a). Write $\underline{c} \in \gamma$ for the first point in γ such that either $c_1 = u_1$ or $c_2 = u_2$ and $\underline{d} \in \gamma$ for the first point in γ such that either $d_1 = v_1$ or $d_2 = v_2$. Write \underline{l} and \underline{r} for the first and last point along γ where γ and $\Delta_{\psi,\phi}$ intersect. Since γ is coordinate-wise increasing we have $\underline{a} \leq \underline{c} \leq \underline{l} \leq \underline{r} \leq \underline{d} \leq \underline{b}$. See Fig. 6a. We split the integral along γ into five parts:

$$w_{\psi,\phi}(\gamma) = w_{\psi,\phi}(\underline{a} \xrightarrow{\gamma} \underline{c}) + w_{\psi,\phi}(\underline{c} \xrightarrow{\gamma} \underline{l}) + w_{\psi,\phi}(\underline{l} \xrightarrow{\gamma} \underline{r}) + w_{\psi,\phi}(\underline{r} \xrightarrow{\gamma} \underline{d}) + w_{\psi,\phi}(\underline{d} \xrightarrow{\gamma} \underline{b}).$$

Using Lemma 6.3 we split the minimising integral from \underline{u} to \underline{v} into three parts:

$$W_{\psi,\phi}(\underline{u}, \underline{v}) = w_{\psi,\phi}(\underline{u} \xrightarrow{\Delta_{\psi,\phi}} \underline{l}) + w_{\psi,\phi}(\underline{l} \xrightarrow{\Delta_{\psi,\phi}} \underline{r}) + w_{\psi,\phi}(\underline{r} \xrightarrow{\Delta_{\psi,\phi}} \underline{v}).$$

By Lemma 6.4,

$$w_{\psi,\phi}(\underline{a} \xrightarrow{\gamma} \underline{c}) \geq w_{\psi,\phi}(\underline{a} \rightarrow (c_1, a_2) \rightarrow \underline{c}).$$

By Lemma 6.2 and the fact that either

$$\underline{c} \xrightarrow{\gamma} \underline{l} \succ \underline{c} \rightarrow \underline{u} \xrightarrow{\Delta_{\psi,\phi}} \underline{l} \succ \Delta_{\psi,\phi} \quad \text{or} \quad \Delta_{\psi,\phi} \succ \underline{c} \rightarrow \underline{u} \xrightarrow{\Delta_{\psi,\phi}} \underline{l} \succ \underline{c} \xrightarrow{\gamma} \underline{l},$$

we have

$$w_{\psi,\phi}(\underline{c} \xrightarrow{\gamma} \underline{l}) \geq w_{\psi,\phi}(\underline{c} \rightarrow \underline{u}) + w_{\psi,\phi}(\underline{u} \xrightarrow{\Delta_{\psi,\phi}} \underline{l}).$$

By Lemma 6.1,

$$w_{\psi,\phi}(\underline{a} \rightarrow (c_1, a_2) \rightarrow \underline{c}) + w_{\psi,\phi}(\underline{c} \rightarrow \underline{u}) = w_{\psi,\phi}(\underline{a} \rightarrow (c_1, a_2) \rightarrow \underline{c} \rightarrow \underline{u}) \geq W_{\psi,\phi}(\underline{a}, \underline{u}).$$

Moreover, by Lemma 6.3,

$$w_{\psi,\phi}(\underline{l} \xrightarrow{\gamma} \underline{r}) \geq w_{\psi,\phi}(\underline{l} \xrightarrow{\Delta_{\psi,\phi}} \underline{r}).$$

By Lemma 6.2 and the fact that either

$$\underline{r} \xrightarrow{\gamma} \underline{d} \succ \underline{r} \xrightarrow{\Delta_{\psi,\phi}} \underline{v} \rightarrow \underline{d} \succ \Delta_{\psi,\phi} \quad \text{or} \quad \Delta_{\psi,\phi} \succ \underline{r} \xrightarrow{\Delta_{\psi,\phi}} \underline{v} \rightarrow \underline{d} \succ \underline{r} \xrightarrow{\gamma} \underline{d},$$

we have

$$w_{\psi,\phi}(\underline{r} \xrightarrow{\gamma} \underline{d}) \geq w_{\psi,\phi}(\underline{r} \xrightarrow{\Delta_{\psi,\phi}} \underline{v}) + w_{\psi,\phi}(\underline{v} \rightarrow \underline{d}).$$

And finally, since $\underline{v} \rightarrow \underline{d} \xrightarrow{\gamma} \underline{b}$ is a path from \underline{v} to \underline{b} ,

$$W_{\psi,\phi}(\underline{v}, \underline{b}) \leq w_{\psi,\phi}(\underline{v} \rightarrow \underline{d}) + w_{\psi,\phi}(\underline{d} \xrightarrow{\gamma} \underline{b}).$$

Combining the above inequalities we obtain

$$\begin{aligned}
 w_{\psi,\phi}(\gamma) &\geq W_{\psi,\phi}(\underline{a}, \underline{u}) + w_{\psi,\phi}(\underline{u} \xrightarrow{\Delta_{\psi,\phi}} \underline{l}) + w_{\psi,\phi}(\underline{l} \xrightarrow{\Delta_{\psi,\phi}} \underline{r}) \\
 &\quad + w_{\psi,\phi}(\underline{r} \xrightarrow{\Delta_{\psi,\phi}} \underline{v}) + w_{\psi,\phi}(\underline{v} \rightarrow \underline{d}) + W_{\psi,\phi}(\underline{v}, \underline{b}) - w_{\psi,\phi}(\underline{v} \rightarrow \underline{d}) \\
 &\geq W_{\psi,\phi}(\underline{a}, \underline{u}) + W_{\psi,\phi}(\underline{u}, \underline{v}) + W_{\psi,\phi}(\underline{v}, \underline{b}). \tag{6.5}
 \end{aligned}$$

Now we consider case (b), that $\gamma \cap \Delta_{\psi,\phi} = \emptyset$. Let \underline{f} be the first point on γ such that $f_1 = u_1$, and let \underline{g} be the first point on γ such that $g_1 = v_1$. See Fig. 6b. We divide the integral along γ into three parts:

$$w_{\psi,\phi}(\gamma) = w_{\psi,\phi}(\underline{a} \xrightarrow{\gamma} \underline{f}) + w_{\psi,\phi}(\underline{f} \xrightarrow{\gamma} \underline{g}) + w_{\psi,\phi}(\underline{g} \xrightarrow{\gamma} \underline{b}).$$

By Lemma 6.4,

$$\begin{aligned}
 w_{\psi,\phi}(\underline{a} \xrightarrow{\gamma} \underline{f}) &\geq w_{\psi,\phi}(\underline{a} \rightarrow (f_1, a_2) \rightarrow \underline{f}) \\
 &= w_{\psi,\phi}(\underline{a} \rightarrow (u_1, a_2) \rightarrow \underline{u}) + w_{\psi,\phi}(\underline{u} \rightarrow \underline{f}) \\
 &= W_{\psi,\phi}(\underline{a}, \underline{u}) + w_{\psi,\phi}(\underline{u} \rightarrow \underline{f}).
 \end{aligned}$$

Since $\gamma \cap \Delta_{\psi,\phi} = \emptyset$, either $f_2 \geq u_2$ and $g_2 \geq v_2$ or $f_2 \leq u_2$ and $g_2 \leq v_2$, so that either

$$\underline{u} \rightarrow \underline{f} \xrightarrow{\gamma} \underline{g} \rightarrow \underline{v} \succcurlyeq \Delta_{\psi,\phi} \quad \text{or} \quad \Delta_{\psi,\phi} \succcurlyeq \underline{u} \rightarrow \underline{f} \xrightarrow{\gamma} \underline{g} \rightarrow \underline{v}.$$

It thus follows by Lemmas 6.2 and 6.3 that

$$w_{\psi,\phi}(\underline{f} \xrightarrow{\gamma} \underline{g}) \geq W_{\psi,\phi}(\underline{u}, \underline{v}) - w_{\psi,\phi}(\underline{u} \rightarrow \underline{f}) - w_{\psi,\phi}(\underline{g} \rightarrow \underline{v}).$$

Finally, since $\underline{v} \rightarrow \underline{g} \xrightarrow{\gamma} \underline{b}$ is a path from \underline{v} to \underline{b} ,

$$W_{\psi,\phi}(\underline{v}, \underline{b}) \leq w_{\psi,\phi}(\underline{v} \rightarrow \underline{g}) + w_{\psi,\phi}(\underline{g} \xrightarrow{\gamma} \underline{b}).$$

Combining the above inequalities we obtain

$$\begin{aligned}
 w_{\psi,\phi}(\gamma) &\geq W_{\psi,\phi}(\underline{a}, \underline{u}) + w_{\psi,\phi}(\underline{u} \rightarrow \underline{f}) + W_{\psi,\phi}(\underline{u}, \underline{v}) - w_{\psi,\phi}(\underline{u} \rightarrow \underline{f}) \\
 &\quad - w_{\psi,\phi}(\underline{g} \rightarrow \underline{v}) + W_{\psi,\phi}(\underline{v}, \underline{b}) - w_{\psi,\phi}(\underline{v} \rightarrow \underline{g}) \\
 &= W_{\psi,\phi}(\underline{a}, \underline{u}) + W_{\psi,\phi}(\underline{u}, \underline{v}) + W_{\psi,\phi}(\underline{v}, \underline{b}). \tag{6.6}
 \end{aligned}$$

From (6.5) and (6.6) we conclude that the lower bound $w_{\psi,\phi}(\gamma) \geq W_{\psi,\phi}(\underline{a}, \underline{u}) + W_{\psi,\phi}(\underline{u}, \underline{v}) + W_{\psi,\phi}(\underline{v}, \underline{b})$ holds uniformly for any path γ , and therefore, it also holds for the infimum, completing the proof. □

The following lemma now states the crucial bound:

Lemma 6.6 *Let \underline{u} and \underline{v} be as in (6.4), and let $\underline{a} \leq \underline{u}$, $a_2 = o(1/p)$, and $\underline{b} \geq \underline{v}$. Then*

$$W_{\psi,\phi}(\underline{a}, \underline{b}) \geq \frac{1}{6p} \log^2 \frac{1}{p} - (1 + o(1)) \frac{1}{3p} \log \left(\frac{8}{3e} \right) \log \frac{1}{p}.$$

Proof Since we assumed $\underline{a} \leq \underline{u}$ and $\underline{b} \geq \underline{v}$, Lemma 6.5 gives

$$W_{\psi,\phi}(\underline{a}, \underline{b}) \geq W_{\psi,\phi}(\underline{a}, \underline{u}) + W_{\psi,\phi}(\underline{u}, \underline{v}). \tag{6.7}$$

Recall that we have set $\xi = \lceil \log^2 \frac{1}{p} \rceil$ and $\delta_\xi = 1 - 2/\xi$. We use Lemma 6.4, that $a_1 \leq u_1$, and that $a_2 = o(1/p)$ to bound

$$\begin{aligned} W_{\psi,\phi}(\underline{a}, \underline{u}) &\geq - \left(\log(8p^2 u_1 + 8p) + 1/\xi \right) (u_2 - a_2) \\ &= \frac{4}{p} \log \frac{1}{p} \log \log \frac{1}{p} - o \left(\frac{1}{p} \log \frac{1}{p} \right). \end{aligned} \tag{6.8}$$

Now we bound $W_{\psi,\phi}(\underline{u}, \underline{v})$. It follows by Lemma 6.3 that

$$\begin{aligned} W_{\psi,\phi}(\underline{u}, \underline{v}) &= \int_{\frac{4}{p} \log \log \frac{1}{p}}^{\frac{1}{3p} \log \frac{3}{p}} \psi(x(y)) dy + \phi(y) x'(y) dy \\ &\geq \int_{\frac{4}{p} \log \log \frac{1}{p}}^{\frac{1}{3p} \log \frac{1}{p}} \left(-\log \left(\frac{8p}{3} e^{3py} \right) - 1/\xi + e^{-3py} \cdot e^{3py} \right) dy \\ &\geq \int_{\frac{4}{p} \log \log \frac{1}{p}}^{\frac{1}{3p} \log \frac{1}{p}} \left(1 - 3/\xi - \log \left(\frac{8p}{3} \right) - 3py \right) dy. \end{aligned} \tag{6.9}$$

The integral on the right-hand side evaluates to

$$\begin{aligned} &\left(1 - 3/\xi - \log \frac{8p}{3} \right) \left[y \right]_{\frac{4}{p} \log \log \frac{1}{p}}^{\frac{1}{3p} \log \frac{1}{p}} - \left[\frac{3py^2}{2} \right]_{\frac{4}{p} \log \log \frac{1}{p}}^{\frac{1}{3p} \log \frac{1}{p}} \\ &= \frac{1}{6p} \log^2 \frac{1}{p} - \frac{4}{p} \log \frac{1}{p} \log \log \frac{1}{p} - (1 + o(1)) \frac{1}{3p} \log \frac{8}{3e} \log \frac{1}{p}. \end{aligned}$$

(Observe that the first term in the first integral in (6.9) thus gives a complementary bound to (2.7), while the second term is complementary to (2.6).) It follows that

$$\begin{aligned}
 W_{\psi,\phi}(u, v) &\geq \frac{1}{6p} \log^2 \frac{1}{p} - \frac{4}{p} \log \frac{1}{p} \log \log \frac{1}{p} \\
 &\quad - (1 + o(1)) \frac{1}{3p} \log \frac{8}{3e} \log \frac{1}{p}.
 \end{aligned}
 \tag{6.10}$$

Substituting (6.8) and (6.10) into (6.7) completes the proof. □

Proof of Lemma 5.9 Given a sequence of increasing rectangles $(R_n)_{n=0}^N$, let $(x_n, y_n) \in \mathbb{R}_+^2$ denote the dimensions of R_n . Construct the path $\gamma \subset \mathbb{R}_+^2$ by linearly interpolating between successive points (x_n, y_n) , i.e.,

$$\gamma = (x_0, y_0) \rightarrow (x_1, y_1) \rightarrow \dots \rightarrow (x_N, y_N).$$

Recall the definition of $U^p(R, R')$ from (4.3) and recall that $\delta_\xi = 1 - \log^{-2} \frac{1}{p}$. It follows from Lemma 6.1 that

$$\sum_{n=0}^{N-1} U^p(R_n, R_{n+1}) \geq \delta_\xi W_{\psi,\phi}((x_0, y_0), (x_N, y_N)),$$

and it follows from Lemma 6.6 that the right-hand side is bounded from below by

$$\frac{1}{6p} \log^2 \frac{1}{p} - (1 + o(1)) \frac{1}{3p} \log \left(\frac{8}{3e} \right) \log \frac{1}{p},$$

completing the proof. □

7 The critical probability: proof of Theorem 1.1

We start with the upper bound. From [23] we know that if $L \ll e^{p^{-1+\varepsilon}}$ for any $\varepsilon > 0$, then $\mathbb{P}_p([L]^2 \text{ is IF}) = o(1)$, so we assume that $L \geq e^{p^{-1+\varepsilon}}$. Let $m = p^{-5}$. The probability that $[L]^2$ is internally filled is bounded from below by the probability that $[L]^2$ contains exactly one internally filled translate of $[m]^2$, and that $\{[m]^2 \Rightarrow [L]^2\}$ occurs. Indeed, let $R_{(i,j)} = [(i - 1)m + 1, im] \times [(j - 1)m + 1, jm]$, and let

$$A_{(i,j)} := \{R_{(i,j)} \text{ is IF}\} \cap \{\forall (i', j') \neq (i, j) : R_{(i',j')} \text{ is not IF}\} \cap \{R_{(i,j)} \Rightarrow [L]^2\}.$$

Then

$$\{[L]^2 \text{ is IF}\} \supseteq \bigsqcup_{i,j=1}^{\lfloor L/m \rfloor} A_{(i,j)},$$

so that

$$\begin{aligned} \mathbb{P}_p([L]^2 \text{ is IF}) \geq & \sum_{i,j=1}^{\lfloor L/m \rfloor} \left(\mathbb{P}_p(R_{(i,j)} \text{ is IF}) - (1 - \mathbb{P}_p(R_{(i,j)} \Rightarrow [L]^2)) \right. \\ & \left. - \sum_{\substack{i',j'=1: \\ i' \neq i, j' \neq j}}^{\lfloor L/m \rfloor} \mathbb{P}_p(R_{(i,j)} \text{ and } R_{(i',j')} \text{ are IF}) \right). \end{aligned} \tag{7.1}$$

To bound $\mathbb{P}_p(R_{(i,j)} \Rightarrow [L]^2)$, observe that if every horizontal and vertical line segment of length p^{-5} intersecting $[L]^2$ contains a pair of adjacent infected sites, then it must be the case that $\{R_{(i,j)} \Rightarrow [L]^2\}$ occurs. We can bound the probability of this event from below by

$$(1 - (1 - p^2)^{\frac{m}{2}})^{4L^2} = \exp(4L^2 e^{-\frac{1}{2p^3}(1+o(1))}) \geq \exp(ce^{-c'p^{-3}}),$$

for some $c, c' > 0$ and p sufficiently small. Note that by Proposition 2.1, the right-hand side is $1 - o(\mathbb{P}_p([m]^2 \text{ is IF}))$. Inserting this bound into (7.1) and summing over the indices, we obtain

$$\begin{aligned} \mathbb{P}_p([L]^2 \text{ is IF}) \geq & \frac{L^2}{m^2} \left(\mathbb{P}_p([m]^2 \text{ is IF}) - (1 - \exp(ce^{-c'p^{-3}})) - \frac{L^2}{m^2} \mathbb{P}_p([m]^2 \text{ is IF})^2 \right) \\ \geq & \frac{L^2}{2m^2} \mathbb{P}_p([m]^2 \text{ is IF}) \left(1 - \frac{2L^2}{m^2} \mathbb{P}_p([m]^2 \text{ is IF}) \right), \end{aligned}$$

where the second inequality holds for p sufficiently small. Taking minus the logarithm on both sides and applying the above bound and Proposition 2.1 again we obtain the inequality

$$\begin{aligned} & -\log(\mathbb{P}_p([L]^2 \text{ is IF})) \\ & \leq 2C_1 \frac{1}{p} \log^2 \frac{1}{p} - 2C_2 \frac{1}{p} \log \frac{1}{p} - 2 \log L + 10 \log \frac{1}{p} + \log 2 \\ & \quad - \log \left(1 - \frac{2L^2}{p^{10}} \exp \left(-2C_1 \frac{1}{p} \log^2 \frac{1}{p} + (2C_2 - o(1)) \frac{1}{p} \log \frac{1}{p} \right) \right). \end{aligned}$$

Observe that if the right-hand side tends to 0 from above when we let $p \rightarrow 0$, then $\mathbb{P}_p([L]^2 \text{ is IF}) \rightarrow 1$. To make the right-hand side vanish we will fix L to be equal to Λ , where

$$\Lambda = \Lambda(p) := \exp \left(\frac{C_1}{p} \log^2 \frac{1}{p} - \frac{C_2 - \eta_u}{p} \log \frac{1}{p} \right),$$

where $\eta_u := p(10 + \log 2 \log^{-1} \frac{1}{p}) = o(1)$. Indeed, with this choice all the leading order terms cancel, and we obtain

$$-\log(\mathbb{P}_p([L]^2 \text{ is IF})) \leq -\log\left(1 - 4 \exp\left(-o\left(\frac{1}{p} \log \frac{1}{p}\right)\right)\right) \rightarrow 0 \text{ as } p \rightarrow 0,$$

as desired.

Now we invert $\Lambda(p)$ to find $p_u = p_u(L)$, an asymptotically minimal sequence in p such that

$$\mathbb{P}_{p_u(L)}([L]^2 \text{ is IF}) \rightarrow 1 \quad \text{as } L \rightarrow \infty.$$

We can express p_u and $1/p_u$ in terms of Λ :

$$p_u = \frac{C_1 \log^2 \frac{1}{p_u} - (C_2 - \eta_u) \log \frac{1}{p_u}}{\log \Lambda}, \tag{7.2}$$

and

$$\frac{1}{p_u} = \frac{\log \Lambda}{C_1 \log^2 \frac{1}{p_u} - (C_2 - \eta_u) \log \frac{1}{p_u}}. \tag{7.3}$$

Substituting (7.3) into (7.2), we get

$$p_u = \frac{1}{\log \Lambda} \left(C_1 \left((\log \log \Lambda) - \log \left(C_1 \log^2 \frac{1}{p_u} - (C_2 - \eta_u) \log \frac{1}{p_u} \right) \right)^2 - (C_2 - \eta_u) \left((\log \log \Lambda) - \log \left(C_1 \log^2 \frac{1}{p_u} - (C_2 - \eta_u) \log \frac{1}{p_u} \right) \right) \right).$$

For sufficiently large Λ (and hence for small p_u) and $\delta > 0$, the following inequalities hold:

$$\frac{1}{p_u} \leq \log \Lambda \leq \frac{1}{p_u^{1+\delta}}, \quad \text{and} \quad \log \frac{1}{p_u} \leq \log \log \Lambda \leq (1 + \delta) \log \frac{1}{p_u}.$$

Whence we obtain the asymptotic formula

$$p_u = \frac{C_1 (\log \log \Lambda)^2}{\log \Lambda} - \frac{4C_1 \log \log \Lambda \log \log \log \Lambda}{\log \Lambda} + \frac{(C_2 + 2C_1 \log C_1 + \delta) \log \log \Lambda}{\log \Lambda},$$

for $\delta > 0$, giving the upper bound in Theorem 1.1.

Now we prove the lower bound. Again let $m = p^{-5}$ and let $n = \frac{1}{p} \log \frac{1}{p}$ and let $L \gg m, n$. Let \mathcal{R} denote the set of all rectangles $R \subset [L]^2$ with dimensions (x, y)

such that $m/3 \leq x \leq m$ and $n/3 \leq y \leq n$. It is a straightforward consequence of the proof of [23, Lemma 3.7] that if $[L]^2$ is internally filled, then there must exist a rectangle $R \in \mathcal{R}$ such that $\{R \Rightarrow [L]^2\}$ occurs.¹³ The number of rectangles in \mathcal{R} is bounded by mnL^2 , so by Proposition 5.1,

$$\begin{aligned} \mathbb{P}_p([L]^2 \text{ is IF}) &= \mathbb{P}_p\left(\bigcup_{R \in \mathcal{R}} \{R \text{ is IF}\} \cap \{R \Rightarrow [L]^2\}\right) \\ &\leq mnL^2 \exp\left(-2C_1 \frac{1}{p} \log^2 \frac{1}{p} + (2C_2 + \zeta) \frac{1}{p} \log \frac{1}{p}\right) \end{aligned}$$

for any $\zeta > 0$. Taking the logarithm of both sides gives

$$\log \mathbb{P}_p([L]^2 \text{ is IF}) \leq 2 \log L - 2C_1 \frac{1}{p} \log^2 \frac{1}{p} + 2(C_2 + \eta_\ell) \frac{1}{p} \log \frac{1}{p},$$

where $\eta_\ell := p \log(mn) + \zeta$. Observe that if the upper bound tends to $-\infty$ as $p \rightarrow 0$, then $\mathbb{P}_p([L]^2 \text{ is IF}) \rightarrow 0$. To minimise the right-hand side, we will fix L to be equal to λ , where

$$\lambda = \lambda(p) := \exp\left(\frac{C_1}{p} \log^2 \frac{1}{p} - \frac{C_2 + \eta_\ell}{p} \log \frac{1}{p}\right).$$

Again, we invert $\lambda(p)$, now to find $p_\ell = p_\ell(L)$, an asymptotically maximal sequence in p such that

$$\mathbb{P}_{p_\ell(L)}([L]^2 \text{ is IF}) \rightarrow 0 \quad \text{as} \quad L \rightarrow \infty.$$

Using the same steps as we used in the proof of the upper bound, we can now determine that p_ℓ satisfies

$$\begin{aligned} p_\ell &= \frac{C_1 (\log \log \lambda)^2}{\log \lambda} - \frac{4C_1 \log \log \lambda \log \log \log \lambda}{\log \lambda} \\ &\quad + \frac{(C_2 + 2C_1 \log C_1 - \delta) \log \log \lambda}{\log \lambda} \end{aligned}$$

for some $\delta > 0$ that can be chosen arbitrarily small but depends on ζ . This concludes the proof of Theorem 1.1. □

Acknowledgements The authors would like to thank Robert Morris for his involvement in the earlier stages of the project, and for the many crucial insights he provided. We thank the anonymous referee for their careful reading and comments. The third author would like to thank Robert Fitzner for useful discussions about computer simulations. The first author was supported by the IDEX Chair funded by Paris-Saclay and the NCCR SwissMap funded by the Swiss NSF. The third author was supported by the Netherlands Organisation for Scientific Research (NWO) through Gravitation—Grant networks-024.002.003.

¹³ This follows if we stop the algorithm in the proof of [23, Lemma 3.7] when a set \mathcal{S} such that $\langle \mathcal{S} \rangle \subset \mathcal{R}$ is first constructed, and by observing that if $[L]^2$ is internally filled, then we must construct such a set \mathcal{S} .

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Adler, J.: Bootstrap percolation. *Physica A* **171**(3), 453–470 (1991)
- Adler, J., Lev, U.: Bootstrap percolation: visualizations and applications. *Braz. J. Phys.* **33**(3), 641–644 (2003)
- Adler, J., Stauffer, D., Aharony, A.: Comparison of bootstrap percolation models. *J. Phys. A Math. Gen.* **22**(7), L297 (1989)
- Aizenman, M., Lebowitz, J.L.: Metastability effects in bootstrap percolation. *J. Phys. A* **21**(19), 3801–3813 (1988)
- Amini, H.: Bootstrap percolation in living neural networks. *J. Stat. Phys.* **141**(3), 459–475 (2010)
- Balister, P., Bollobás, B., Przykucki, M., Smith, P.: Subcritical \mathcal{U} -bootstrap percolation models have non-trivial phase transitions. *Trans. A. Math. Soc.* **368**(10), 7385–7411 (2016)
- Balogh, J., Bollobás, B.: Sharp thresholds in bootstrap percolation. *Physica A* **326**(3), 305–312 (2003)
- Balogh, J., Bollobás, B., Duminil-Copin, H., Morris, R.: The sharp threshold for bootstrap percolation in all dimensions. *Trans. Am. Math. Soc.* **364**(5), 2667–2701 (2012)
- Balogh, J., Bollobás, B., Morris, R.: Bootstrap percolation in three dimensions. *Ann. Probab.* **37**(4), 1329–1380 (2009)
- Balogh, J., Bollobás, B., Morris, R.: Bootstrap percolation in high dimensions. *Comb. Probab. Comput.* **19**(5–6), 643–692 (2010)
- Boerma-Klooster, S.: A sharp threshold for an anisotropic bootstrap percolation model. University of Groningen Bachelor Thesis (2011)
- Bollobás, B., Duminil-Copin, H., Morris, R., Smith, P.: Universality of two-dimensional critical cellular automata. To appear in *Proc. Lond. Math. Soc.*, Preprint. [arXiv:1406.6680](https://arxiv.org/abs/1406.6680) (2014)
- Bollobás, B., Duminil-Copin, H., Morris, R., Smith, P.: The sharp threshold for the Duarte model. To appear in *Ann. Probab.*, Preprint. [arXiv:1603.05237](https://arxiv.org/abs/1603.05237) (2016)
- Bollobás, B., Smith, P., Uzzell, A.J.: Monotone cellular automata in a random environment. *Comb. Prob. Comput.* **24**, 687–722 (2015)
- Bringmann, K., Mahlborg, K.: Improved bounds on metastability thresholds and probabilities for generalized bootstrap percolation. *Trans. Am. Math. Soc.* **364**(7), 3829–3859 (2012)
- Bringmann, K., Mahlborg, K., Mellit, A.: Convolution bootstrap percolation models, Markov-type stochastic processes, and mock theta functions. *Int. Math. Res. Not. IMRN* **5**, 971–1013 (2013)
- Cancrini, N., Martinelli, F., Roberto, C., Toninelli, C.: Kinetically constrained spin models. *Probab. Theory Relat. Fields* **140**(3–4), 459–504 (2008)
- Cerf, R., Cirillo, E.N.M.: Finite size scaling in three-dimensional bootstrap percolation. *Ann. Probab.* **27**(4), 1837–1850 (1999)
- Cerf, R., Manzo, F.: The threshold regime of finite volume bootstrap percolation. *Stoch. Process. Appl.* **101**(1), 69–82 (2002)
- Chalupa, J., Leath, P.L., Reich, G.R.: Bootstrap percolation on a Bethe lattice. *J. Phys. C* **12**(1), L31 (1979)
- De Gregorio, P., Lawlor, A., Bradley, P., Dawson, K.A.: Clarification of the bootstrap percolation paradox. *Phys. Rev. Lett.* **93**(2), 025501 (2004)
- Duarte, J.: Simulation of a cellular automaton with an oriented bootstrap rule. *Physica A* **157**(3), 1075–1079 (1989)
- Duminil-Copin, H., van Enter, A.C.D.: Sharp metastability threshold for an anisotropic bootstrap percolation model. *Ann. Probab.* **41**(3A), 1218–1242 (2013)
- Duminil-Copin, H., van Enter, A.C.D.: Erratum to “Sharp metastability threshold for an anisotropic bootstrap percolation model”. *Ann. Probab.* **44**(2), 1599 (2016)
- Duminil-Copin, H., Holroyd, A.E.: Finite volume bootstrap percolation with threshold rules on \mathbb{Z}^2 . Preprint at <http://www.ihes.fr/~duminil/publi.html> (2012)
- Fontes, L.R., Schonmann, R.H., Sidoravicius, V.: Stretched exponential fixation in stochastic Ising models at zero temperature. *Commun. Math. Phys.* **228**(3), 495–518 (2002)

27. Gravner, J., Griffeath, D.: First passage times for threshold growth dynamics on \mathbb{Z}^2 . *Ann. Probab.* **24**(4), 1752–1778 (1996)
28. Gravner, J., Holroyd, A.E.: Slow convergence in bootstrap percolation. *Ann. Appl. Probab.* **18**(3), 909–928 (2008)
29. Gravner, J., Holroyd, A.E.: Local bootstrap percolation. *Electron. J. Probab.* **14**(14), 385–399 (2009)
30. Gravner, J., Holroyd, A.E., Morris, R.: A sharper threshold for bootstrap percolation in two dimensions. *Probab. Theory Relat. Fields* **153**(1–2), 1–23 (2012)
31. Gray, L.: A mathematician looks at Wolfram’s new kind of science. *Not. AMS* **50**, 200–211 (2003)
32. Grimmett, G.: *Percolation*, 2nd edn. Springer, Berlin (1999)
33. Holroyd, A.E.: Sharp metastability threshold for two-dimensional bootstrap percolation. *Probab. Theory Relat. Fields* **125**(2), 195–224 (2003)
34. Holroyd, A.E., Liggett, T.M., Romik, D.: Integrals, partitions, and cellular automata. *Trans. Am. Math. Soc.* **356**(8), 3349–3368 (2004). (electronic)
35. Janson, S.: Poisson approximation for large deviations. *Random Struct. Algorithms* **1**(2), 221–229 (1990)
36. Morris, R.: Zero-temperature Glauber dynamics on \mathbb{Z}^d . *Probab. Theory Relat. Fields* **149**(3–4), 417–434 (2011)
37. Morris, R.: The second term for bootstrap percolation in two dimensions. Manuscript in preparation. Available at <http://w3.impa.br/~rob/> (2014)
38. Mountford, T.S.: Critical length for semi-oriented bootstrap percolation. *Stoch. Process. Appl.* **56**(2), 185–205 (1995)
39. Rudin, W.: *Real and Complex Analysis*, 3rd edn. McGraw-Hill Book Co., New York (1987)
40. Schonmann, R.H.: Critical points of two-dimensional bootstrap percolation-like cellular automata. *J. Stat. Phys.* **58**(5–6), 1239–1244 (1990)
41. van Enter, A.C.D., Hulshof, T.: Finite-size effects for anisotropic bootstrap percolation: logarithmic corrections. *J. Stat. Phys.* **128**(6), 1383–1389 (2007)