HAL
open science

The second term for two-neighbour bootstrap percolation in two dimensions

Ivailo Hartarsky, Robert Morris

▶ To cite this version:

Ivailo Hartarsky, Robert Morris. The second term for two-neighbour bootstrap percolation in two dimensions. Transactions of the American Mathematical Society, 2019, 372 (9), pp.6465-6505. 10.1090/tran/7828 . hal-02275559

HAL Id: hal-02275559
https://hal.science/hal-02275559

Submitted on 28 Oct 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

arXiv:1806.08931v1 [math.PR] 23 Jun 2018

THE SECOND TERM FOR TWO-NEIGHBOUR BOOTSTRAP PERCOLATION IN TWO DIMENSIONS

IVAILO HARTARSKY AND ROBERT MORRIS

Abstract. In the r-neighbour bootstrap process on a graph G, vertices are infected (in each time step) if they have at least r already-infected neighbours. Motivated by its close connections to models from statistical physics, such as the Ising model of ferromagnetism, and kinetically constrained spin models of the liquid-glass transition, the most extensively-studied case is the two-neighbour bootstrap process on the two-dimensional grid $[n]^2$. Around 15 years ago, in a major breakthrough, Holroyd determined the sharp threshold for percolation in this model, and his bounds were subsequently sharpened further by Gravner and Holroyd, and by Gravner, Holroyd and Morris.

In this paper we strengthen the lower bound of Gravner, Holroyd and Morris by proving that the critical probability $p_c\big([n]^2, 2\big)$ for percolation in the two-neighbour model on $[n]^2$ satisfies

$$p_c\big([n]^2, 2\big) = \frac{\pi^2}{18 \log n} - \frac{\Theta(1)}{(\log n)^{3/2}} \,.$$

The proof of this result requires a very precise understanding of the typical growth of a critical droplet, and involves a number of technical innovations. We expect these to have other applications, for example, to the study of more general two-dimensional cellular automata, and to the r-neighbour process in higher dimensions.

1. Introduction

In this paper we study the phase transition of the two-neighbour bootstrap percolation process on the two-dimensional grid $[n]^2$. This model has been extensively studied over the past 30 years, most notably by Aizenman and Lebowitz [1] and by Holroyd [21], who determined the sharp threshold for percolation. In this paper we will improve the best known lower bound on the critical probability, proving a bound which matches the upper bound of Gravner and Holroyd [18] up to a constant factor in the second term.

The r-neighbour bootstrap percolation process, which was first introduced in 1979 by Chalupa, Leath and Reich [12], is a deterministic, monotone cellular automaton, defined on a graph G. Given a set $A \subset V(G)$ of initially 'infected' vertices, new vertices are infected at each time step if they have at least r infected neighbours, and infected vertices remain infected forever. Thus, writing A_t for the set of infected vertices at time t, we have $A_0 = A$, and

$$A_{t+1} = A_t \cup \Big\{ v \in V(G) : |N(v) \cap A_t| \geqslant r \Big\}$$

for each $t \geqslant 0$. We write $[A] := \bigcup_{t \geqslant 0} A_t$ for the *closure* of A under the bootstrap process (that is, set of eventually infected sites), and say that A *percolates* if $[A] = V(G)$.

2010 *Mathematics Subject Classification.* Primary 60C05; Secondary 60K35, 82B20.

Key words and phrases. bootstrap percolation, critical probability, finite-size scaling, sharp threshold.

Both authors were partially supported by ERC Starting Grant 680275 MALIG. The research of the second author was also partially supported by CNPq (Proc. 303275/2013-8), by FAPERJ (Proc. 201.598/2014), and by JSPS.

Motivated by applications to statistical physics, for example the Ising model [17, 23], kinetically constrained spin models of the liquid-glass transition [22, 24], and the abelian sandpile [16, 24], the bootstrap process has been most extensively studied on finite subsets of the lattice $\mathbb{Z}^d$, with the initial set A chosen randomly. More precisely, let G be the finite grid $[n]^d = \{1, \dots n\}^d$ with graph structure inherited from $\mathbb{Z}^d$, write $\mathbb{P}_p$ for the probability measure on subsets $A \subset [n]^d$ obtained by including each vertex in A independently at random with probability p (we call such a set *p-random*), and define the *critical probability* for percolation in the r-neighbour model on $[n]^d$ to be

$$p_c\Big([n]^d, r\Big) := \inf\Big\{p \in (0,1) \,:\, \mathbb{P}_p\Big([A] = [n]^d\Big) \geqslant 1/2\Big\}.$$

Foundational work on this problem was done by Aizenman and Lebowitz [1] in 1988, who determined $p_c([n]^d, 2)$ up to a constant factor for all fixed $d \geqslant 2$. However, the problem of determining a sharp threshold for $p_c([n]^2, 2)$ remained open until 2003, when Holroyd [21], in an important breakthrough, proved that

$$p_c\Big([n]^2, 2\Big) = \left(\frac{\pi^2}{18} + o(1)\right)\frac{1}{\log n}. \tag{1}$$

Holroyd's proof was particularly significant for its introduction of the so-called 'method of hierarchies' (see Section 5), which has played a crucial role in much of the subsequent progress in the area.

Building on Holroyd's proof, Gravner and Holroyd [18] and Gravner, Holroyd and Morris [19] improved (respectively) the upper and lower bounds obtained in [21], proving that

$$\frac{\pi^2}{18\log n} - \frac{C(\log\log n)^3}{(\log n)^{3/2}} \leqslant p_c\Big([n]^2, 2\Big) \leqslant \frac{\pi^2}{18\log n} - \frac{c}{(\log n)^{3/2}} \tag{2}$$

for some constants $C > c > 0$. The upper bound was obtained by considering a larger family of possible growth mechanisms than in [21], by allowing the shape of the so-called 'critical droplet' to vary a little (rather than being exactly square). The lower bound was obtained by repeating Holroyd's proof, but using a more refined notion of hierarchy, and counting these hierarchies more carefully (see Section 2 for a more detailed discussion).

In this paper we will prove a stronger lower bound on $p_c([n]^2, 2)$, which removes the $(\log\log n)^3$ term, and hence matches the upper bound up to a constant factor in the second term. To be precise, we will prove the following theorem.

Theorem 1.1. *There exist constants $C > c > 0$ such that*

$$\frac{\pi^2}{18\log n} - \frac{C}{(\log n)^{3/2}} \leqslant p_c\Big([n]^2, 2\Big) \leqslant \frac{\pi^2}{18\log n} - \frac{c}{(\log n)^{3/2}}.$$

Note that the upper bound follows from the theorem of Gravner and Holroyd [18]. We remark that, unlike the result of Gravner, Holroyd and Morris [19], our lower bound does not follow via a relatively minor modification of Holroyd's proof, and we will have to work much harder, and introduce a number of significant technical innovations. In particular, we will require a much finer understanding of the typical growth of a critical droplet, which we obtain using hierarchies that are chosen carefully in order to encode much more information about the growth of the droplet, and extremely precise bounds on the probability that each step of these hierarchies is 'satisfied' by the set A of initially infected sites. We remark that we will need to use non-monotone events in our hierarchies,

which creates additional technical difficulties in the proof. A more detailed sketch of the proof of Theorem 1.1 is given in Section 2.

Before embarking on the proof of Theorem 1.1, let us briefly discuss the state of knowledge in higher dimensions, and some exciting recent developments regarding more general models in two dimensions. As noted above, Aizenman and Lebowitz [1] determined $p_c\big([n]^d, 2\big)$ up to a constant factor for all $d \geqslant 2$. However, it took more than 10 years until a corresponding result was proved for all $d \geqslant r \geqslant 2$, by Cerf and Cirillo [10] (in the case $d = r = 3$) and Cerf and Manzo [11]. The sharp threshold for the r-neighbour process on $[n]^d$ was finally determined by Balogh, Bollobás, Duminil-Copin and Morris [4, 5], who determined, for each $d \geqslant r \geqslant 2$, an explicit constant $\lambda(d, r) > 0$ such that[1]

$$p_c\big([n]^d, r\big) = \left(\frac{\lambda(d,r) + o(1)}{\log_{(r-1)} n}\right)^{d-r+1}.$$

More recently, Uzzell [26] extended the upper bound of Gravner and Holroyd [18] to arbitrary $d \geqslant r \geqslant 2$, providing some hope that a result corresponding to Theorem 1.1 could eventually be proved in this more general setting.

In a different direction, an extremely general family of monotone cellular automata, called *$\mathcal{U}$-bootstrap percolation*, was recently introduced by Bollobás, Smith and Uzzell [9], and subsequently was also studied by Balister, Bollobás, Przykucki and Smith [2] and by Bollobás, Duminil-Copin, Morris and Smith [7]. The model is as follows: given an arbitrary finite collection $\mathcal{U} = \{X_1, \ldots, X_m\}$ of finite subsets of $\mathbb{Z}^2 \setminus \{\mathbf{0}\}$, and a set $A \subset \mathbb{Z}_n^2$ of initially infected sites, set $A_0 = A$ and define

$$A_{t+1} = A_t \cup \Big\{v \in \mathbb{Z}_n^2 \,:\, v + X \subset A_t \text{ for some } X \in \mathcal{U}\Big\}$$

for each $t \geqslant 0$. Confirming a conjecture of Bollobás, Smith and Uzzell, the main results of [9] and [2] together characterize for which of these families the critical probability is polynomial in n, for which it is polylogarithmic, and for which it is bounded away from zero (rather surprisingly, there are no other possibilities). In [7] the critical probability was determined up to a constant factor for families with polylogarithmic critical probability. The sharp threshold is only known for a certain class of centrally symmetric families [13], and for two specific non-symmetric models [8, 14]. For more details, and a discussion of the problem in higher dimensions, we refer the interested reader to [7] or [24]. We also remark that a sharper threshold, along the lines of Theorem 1.1, for a so-called 'unbalanced' model known as 'anisotropic bootstrap percolation' was recently proved by Duminil-Copin, van Enter and Hulshof [15].

The rest of the paper is organised as follows. In Section 2 we give a detailed outline of the proof of Theorem 1.1, and in Section 3 we recall some basic tools and facts that we will need later, and set up some useful notation and conventions used throughout the paper. In Section 4 we state (and give an extended sketch of the proof of) our key bounds on the probability that a rectangle is internally filled by A together with a sub-rectangle (the full details of the proof are postponed to the Appendix.) In Section 5 we introduce the hierarchies we will use in the proof, prove some standard facts about the family of hierarchies, and describe a partition of this family which plays an important role in the analysis. Finally, in Section 6, we prove Theorem 1.1. We finish the paper, in Section 7, by mentioning a couple of natural open problems.

[1] Here $\log_{(r)}$ is an r-times iterated natural logarithm.

2. An outline of the proof

The proof of Theorem 1.1 is very technical, so in this section we will attempt to give the reader an easily-digestible outline of the main ideas behind the proof. The main step will be to bound the probability that a 'critical droplet' R (a rectangle with sides of length between $1/p$ and $(1/p)\log(1/p)$) is 'internally filled' by the p-random set A. The claimed lower bound on $p_c([n]^2, 2)$ will follow easily from this bound via a standard argument (using a lemma due to Aizenman and Lebowitz [1] and the union bound). In order to state this theorem precisely, we will need to introduce a little notation.

A *rectangle* is a non-empty set $R \subset \mathbb{Z}^2$ of the form $[a,b] \times [c,d]$; we write $\dim(R) = (b-a+1, d-c+1)$ for the *dimensions* of R. We say that a rectangle R is *internally filled* by A if $[A \cap R] = R$. We also need the function

$$g(z) := -\log\Big(\beta\big(1-e^{-z}\big)\Big) \tag{3}$$

where $\beta(u) := \frac{1}{2}\Big(u + \sqrt{u(4-3u)}\Big)$, which was defined by Holroyd [21], who also proved that

$$\int_0^\infty g(z)\,\mathrm{d}z \;=\; \lambda := \frac{\pi^2}{18}\,. \tag{4}$$

Finally, set $q := -\log(1-p)$, and note that $q \geqslant p$, and that $q \sim p$ as $p \to 0$. (This notation is convenient, because the probability that a set of size a contains no element of the p-random set A is e^{-aq}. We will assume throughout that $p \to 0$.) We can now state our main bound on the probability that a critical droplet is internally filled.

Theorem 2.1. *There exists a constant $C > 0$ such that the following holds. Let R be a rectangle with dimensions $\dim(R) = (a,b)$, and suppose that $a \leqslant b$, and*

$$\frac{C}{q} \;\leqslant\; b \;\leqslant\; \frac{1}{2q}\log\frac{1}{q}\,. \tag{5}$$

Then

$$\mathbb{P}_p\Big([A \cap R] = R\Big) \leqslant \exp\left(-\min\left\{\frac{2\lambda}{q} + \frac{1}{q^{3/4}},\; (b-a)g(aq) + \frac{2}{q}\int_0^{aq} g(z)\,\mathrm{d}z - \frac{C}{\sqrt{q}}\right\}\right)\,.$$

We remark that the first term in the minimum is easily large enough for our purposes, and is only needed for technical reasons; the reader should therefore focus her attention on the second term. Let us write $\mathrm{long}(R)$ and $\mathrm{short}(R)$ for the maximum and minimum (respectively) of the dimensions of R. In order to deduce Theorem 1.1 from Theorem 2.1, we will need the following fundamental lemma of Aizenman and Lebowitz [1].

The Aizenman–Lebowitz lemma. *If $[A] = [n]^2$, then for each $1 \leqslant k \leqslant n$ there exists a rectangle R with*

$$k \leqslant \mathrm{long}(R) \leqslant 2k$$

that is internally filled by A.

To deduce a lower bound on $p_c([n]^2, 2)$, we simply apply the Aizenman–Lebowitz lemma with $k = (1/(4q))\log(1/q)$, and take a union bound over choices of R, using Theorem 2.1 to bound the probability that R is internally filled, and the (straightforward) fact that

$$(b-a)g(aq) + \frac{2}{q}\int_0^{aq} g(z)\,\mathrm{d}z \;\geqslant\; \frac{2\lambda}{q} - \frac{O(1)}{\sqrt{q}}$$

if $a \leqslant b$ and $b \geqslant (1/(4q)) \log(1/q)$, see Lemma 3.9.

Our main challenge will therefore be to prove Theorem 2.1. As has become standard in the area since their introduction by Holroyd [21], we will do so using hierarchies; however, our definition will differ in various important ways from that used in [21], and also from the various notions of hierarchy used in, for example, [7, 8, 14, 19].

In order to motivate the definition, let us begin by recalling the hierarchies used by Holroyd [21] to prove (1) and by Gravner, Holroyd and Morris [19] to prove (2). Roughly speaking, the basic idea of the proof in [21] is that, given an internally filled rectangle R, we would like to associate with R a constant-size rooted tree of sub-rectangles that 'encodes' the way in which the set $A \cap R$ grows to infect the rest of R. The leaves of this tree correspond to small internally filled rectangles ('seeds'), a vertex with two children corresponds to two (not too small) rectangles merging to form a larger rectangle, and a vertex with one child corresponds to a rectangle 'growing on its sides' to fill a slightly larger rectangle. Crucially, we would like all of these (increasing) events to occur disjointly, so that we can apply the van der Berg–Kesten inequality (see Section 3.3) to bound the probability of their intersection. Since there are few such hierarchies, and each is (roughly speaking) at least as as unlikely as a single 'seed' growing to fill R, one can deduce a sufficiently strong bound on the probability that R is internally filled.

Gravner, Holroyd and Morris [19] required two additional ideas in order to prove the lower bound in (2). First, they needed their seeds to be much smaller (of size $q^{-1/2}$, rather than $o(1/q)$), and to grow geometrically (rather than linearly) as a function of their height in the tree. As a result, the number of possible hierarchies became very large (too large to use a naive union bound), and to deal with this issue they partitioned the family of hierarchies according to the number of 'big' seeds. We will use refinements of both of these ideas in the proof of Theorem 2.1.

In order to prove Theorem 2.1 we can only afford to lose a factor of $\exp\big(O(q^{-1/2})\big)$ (in the expected number of 'satisfied' hierarchies), and since our hierarchies will typically have height $\Theta(q^{-1/2})$, this means that we can only allow ourselves a constant number of choices at each step, unless we 'pay' for extra choices via some unlikely event occurring. Fortunately, this is intuitively possible: the only things that could prevent us from choosing the next rectangle in an almost unique way are: (a) the existence of a 'double gap' of consecutive empty rows or columns blocking the growth of the critical droplet, or (b) the merging of two reasonably large internally filled rectangles. Our challenge will be to show that we gain enough from these events to compensate for the extra choices we are forced to make.

To do so, we will need to encode the existence of double gaps in our hierarchies, which causes two immediate problems: the events cease to be increasing, and cease to occur disjointly. To avoid these issues we only use the fact that the double gaps are empty in a single path through the hierarchy (which we call the 'trunk'); outside the trunk we use increasing events defined on the complement of the double gaps. In Section 4 we will state (and sketch the proofs of) a pair of technical 'crossing' lemmas which provide sufficiently strong bounds on the probabilities of these events. We remark that we gain from the existence of these double gaps in two distinct ways: they force us to find either two infected sites close together, or one infected site in a relatively small region, and when the rectangle is very large they are themselves unlikely to exist.

Bounding the expected number of 'satisfied' hierarchies with height $O(1/\sqrt{q})$ will then be relatively straightforward; unfortunately, however, this is not always the case. In Section 6 we will have to deal with various other types of hierarchy: those with too many vertices, with too many (or too large) seeds, and those whose growth deviates from the diagonal by a macroscopic amount (see Lemma 5.12). One additional innovation that we will need in order to deal with this last case is Lemma 5.16, which provides us with two 'pods', instead of the single pod required by Holroyd.

3. Basic facts and definitions

In this section we will recall a few basic facts about two-neighbour bootstrap percolation on $[n]^2$, state a few simple properties of the function $g(z)$, and introduce some further notation. For convenience, let us fix (for the rest of the paper) sufficiently large constants $B > 0$ and $C = C(B) > 0$, and a sufficiently small constant $\delta = \delta(B, C) > 0$.

3.1. **Preliminaries.** To begin, recall the following simple and well-known fact (see, e.g., [6, Problem 34]). We write $\phi(R)$ for the semi-perimeter of a rectangle R, so $\phi(R) = \mathrm{long}(R) + \mathrm{short}(R)$.

Lemma 3.1. *If $[A \cap R] = R$, then $|A \cap R| \geqslant \dfrac{\phi(R)}{2}$.*

Now, recall from (3) the definition of the function $g(z)$. The next lemma, which bounds the probability that a sufficiently small rectangle is internally filled, follows easily from Lemma 3.1 (see, e.g., [19, Lemma 2]).

Lemma 3.2. *There exists $\delta > 0$ such that for any $p > 0$ and any rectangle R with $\dim(R) = (a, b)$, where $a \leqslant b$ and $ap \leqslant \delta$,*

$$\mathbb{P}_p\Big([A \cap R] = R\Big) \leqslant 3^{\phi(R)} \exp\Big(- \phi(R) g(aq)\Big) .$$

In order to control the growth of a droplet, we will need to bound various probabilities relating to the existence of double gaps. To be precise, let us say that a rectangle $R = [a, b] \times [c, d]$ has a *vertical double gap* if there exists $j \in [a, b-1]$ such that

$$A \cap \Big([j, j+1] \times [c, d]\Big) = \emptyset ,$$

and similarly for a *horizontal double gap*. (We will say that R has a double gap if it has a horizontal or vertical double gap.) We will say that R is *crossed from left to right*[2] if it has no vertical double gap and the rightmost column $\{b\} \times [c, d]$ is *occupied*, that is, has non-empty intersection with A. Note that if the column to the left of R is already infected, and R is crossed from left to right, then R will also be infected by the process. The following simple estimates were proved in [21, Lemma 8].

Lemma 3.3. *If R is a rectangle with $\dim(R) = (a, b)$, then*

$$\mathbb{P}_p\Big(R \text{ has no vertical double gap}\Big) \leqslant e^{-(a-1)g(bq)}$$

and

$$\mathbb{P}_p\Big(R \text{ is crossed from left to right}\Big) \leqslant e^{-ag(bq)} .$$

[2]We define similarly the notions of being crossed from right to left, bottom to top, and top to bottom.

We remark that the function g is positive, decreasing, convex and differentiable on $(0,\infty)$, that $g(z) \sim e^{-2z}$ as $z \to \infty$, that

$$-\frac{1}{2}\log z - \sqrt{z} \leqslant g(z) \leqslant -\frac{1}{2}\log z + z \tag{6}$$

for all sufficiently small $z > 0$ (see [19, Observation 4]), that

$$e^{2g(z)} \leqslant \frac{C}{z} \tag{7}$$

for all $0 < z \leqslant 3e^{2B}$ (see [19, Observation 10]), and that

$$-g'(z) \leqslant \begin{cases} B/z & \text{if } z \leqslant B \\ 3e^{-2z} & \text{if } z \geqslant B/2 \end{cases} \tag{8}$$

since B and $C = C(B)$ were chosen sufficiently large.

3.2. **Analytic estimates.** We will use the following definition from [21] to control the growth of a droplet.

Definition 3.4. For each $\mathbf{a} \leqslant \mathbf{b} \in \mathbb{R}^2_+$, define

$$W(\mathbf{a},\mathbf{b}) = \inf_{\gamma\,:\,\mathbf{a}\to\mathbf{b}} \int_\gamma \Big(g(y)\,\mathrm{d}x + g(x)\,\mathrm{d}y\Big)\,, \tag{9}$$

where the infimum is taken over all piecewise linear increasing paths from $\mathbf{a}$ to $\mathbf{b}$ in $\mathbb{R}^2$.

Now, for any pair $S \subset R$ of rectangles, define

$$U(S,R) := W\Big(q\,\dim(S), q\,\dim(R)\Big)\,. \tag{10}$$

One of the key lemmas from [21] states that the integral in (9) is minimized when the path γ is chosen as close to the diagonal as possible. We will use the following immediate consequence of this fact.

Lemma 3.5 (Lemma 16 of [21])**.** *Let $S \subset R$ be rectangles with* $\mathrm{long}(S) \leqslant \mathrm{short}(R)$. *Then*

$$\frac{U(S,R)}{q} = (d-c)g(dq) + \frac{2}{q}\int_{dq}^{aq} g(z)\,\mathrm{d}z + (b-a)g(aq)\,,$$

where $a = \mathrm{short}(R)$, $b = \mathrm{long}(R)$, $c = \mathrm{short}(S)$ and $d = \mathrm{long}(S)$.

When $\mathrm{long}(S) > \mathrm{short}(R)$ we will use the following easy consequence of the fact that $g(z)$ is decreasing (it also follows immediately from [21, Lemma 16]).

Lemma 3.6. *Let $S \subset R$ be rectangles with* $\mathrm{long}(S) \geqslant \mathrm{short}(R)$. *Then*

$$\frac{U(S,R)}{q} \geqslant (b-d)g(aq)\,,$$

where $a = \mathrm{short}(R)$, $b = \mathrm{long}(R)$ and $d = \mathrm{long}(S)$.

We will also need the following straightforward bound from [19].

Lemma 3.7 (Lemma 14 of [19])**.** *Let $S \subset R$ be rectangles, with* $\mathrm{long}(S) \leqslant \mathrm{short}(R)$. *Then*

$$\frac{U(S,R)}{q} \geqslant \frac{2}{q}\int_0^{aq} g(z)\,\mathrm{d}z + (b-a)g(aq) - \frac{\phi(S)}{2}\log\left(1 + \frac{1}{\phi(S)q}\right) - O\Big(\phi(S)\Big)\,,$$

where $a = \mathrm{short}(R)$ and $b = \mathrm{long}(R)$.

In order to transition between $U(S,R)$ and the bounds proved in Section 4, below, we will also need the following simple upper bound. If S and R are rectangles with dimensions $\dim(R) = (a,b)$ and $\dim S = (a-s, b-t)$, then set

$$Q(S,R) := sg\Big((b-t)q\Big) + tg\Big((a-s)q\Big)\,, \tag{11}$$

The following lemma follows immediately from the fact that $g(z)$ is decreasing.

Lemma 3.8 (Proposition 13 of [21])**.** *Let $S \subset R$ be rectangles. Then*

$$\frac{U(S,R)}{q} \leqslant Q(S,R)\,.$$

We will also need a couple of additional technical lemmas, each of which follows easily from simple properties of the function g. The first is a variant of [19, Observation 19], with slightly weaker assumptions and conclusion.

Lemma 3.9. *If $a \leqslant b$ and $b \geqslant (1/(4q))\log(1/q)$, then*

$$\frac{2}{q}\int_0^{aq} g(z)\,\mathrm{d}z + (b-a)g(aq) \geqslant \frac{2\lambda}{q} - \frac{4e^4}{\sqrt{q}}\,.$$

Proof. Recall that $B > 0$ is a sufficiently large constant, and note that if $a \leqslant B/q$ then

$$(b-a)g(aq) \geqslant \frac{g(B)}{5q}\log\frac{1}{q} \geqslant \frac{2\lambda}{q}\,,$$

since g is decreasing and $q \to 0$. Let us therefore assume that $a \geqslant B/q$, and observe that therefore

$$\int_{aq}^{\infty} g(z)\,\mathrm{d}z \leqslant g(aq)\,, \tag{12}$$

since $g(z) \sim e^{-2z}$ as $z \to \infty$, and hence, recalling the definition (4) of λ,

$$\frac{2}{q}\int_0^{aq} g(z)\,\mathrm{d}z + (b-a)g(aq) \geqslant \frac{2}{q}\int_0^{\infty} g(z)\,\mathrm{d}z = \frac{2\lambda}{q}$$

if $b - a \geqslant 2/q$. Finally, if $b - a \leqslant 2/q$, then $a \geqslant (1/(4q))\log(1/q) - 2/q$, and so

$$\frac{2}{q}\int_0^{aq} g(z)\,\mathrm{d}z \geqslant \frac{2\lambda}{q} - \frac{2g(aq)}{q} \geqslant \frac{2\lambda}{q} - \frac{4e^4}{\sqrt{q}}\,,$$

by (12), and since $g(z) \leqslant 2e^{-2z}$ if $z \geqslant B$. □

The next lemma quantifies how much harder it is for a droplet to grow far from the diagonal. To state it, we need to introduce a further large constant $L_1 = L_1(B, C, \delta) > 0$.

Lemma 3.10. *If $L_1 a \leqslant b \leqslant B/q$, then*

$$\frac{2}{q}\int_{aq}^{bq} g(z)\,\mathrm{d}z \leqslant (b-a)\Big(g(aq) + g(bq)\Big) - 4Cb\,.$$

Proof. We claim first that $-B/z \leqslant g'(z) \leqslant -\delta/z$ for every $0 < z < B$. Indeed, this follows since $g'(z) \sim -1/(2z)$ as $z \to 0$ and $g'(z) \sim -2e^{-2z}$ as $z \to \infty$, and since B was chosen sufficiently large, and $\delta = \delta(B)$ sufficiently small. Now, integrating by parts, we obtain

$$\frac{2}{q}\int_{aq}^{bq} g(z)\,\mathrm{d}z \leqslant 2\Big(bg(bq) - ag(aq)\Big) + 2B(b-a)\,.$$

It follows that

$$\frac{2}{q}\int_{aq}^{bq} g(z)\,\mathrm{d}z - (b-a)\Big(g(aq)+g(bq)\Big) \leqslant (a+b)\Big(g(bq)-g(aq)+2B\Big)\,.$$

Now, since $g'(z) \leqslant -\delta/z$ for every $z < B$, and $b/a \geqslant L_1$, we have

$$g(aq) - g(bq) \;=\; -\int_{aq}^{bq} g'(z)\,\mathrm{d}z \;\geqslant\; \delta \log L_1 \;\geqslant\; 5C\,,$$

and so the claimed bound follows. □

We will also need some larger constants, which we will denote by $L_2, L_3, \ldots$, where each L_i is chosen to be sufficiently large depending on B, C, δ, and $L_1, \ldots, L_{i-1}$. We will use $O(\cdot)$ to denote the existence of an absolute constant, that is, a constant that does *not* depend on any of the aforementioned ones.

3.3. **Correlation inequalities.** To finish this section, we will state the fundamental inequalities of van den Berg and Kesten [28] and Reimer [25], which we will use in Section 5 to bound the probability that a hierarchy is 'satisfied' by A, the p-random set of infected sites, see Definition 5.3 and Lemma 5.7.

In our setting, an *event* $\mathcal{E}$ is simply a family of subsets of $[n]^2$, and the event $\mathcal{E}$ is said to *occur* if $A \in \mathcal{E}$. Two events $\mathcal{E}$ and $\mathcal{F}$ are said to *occur disjointly* for A if there exist disjoint sets $X, Y \subset [n]^2$ depending on A such that $S \in \mathcal{E}$ for any S such that $S \cap X = A \cap X$, and $T \in \mathcal{F}$ for any T such that $T \cap Y = A \cap Y$. We write $\mathcal{E} \circ \mathcal{F}$ for the event that $\mathcal{E}$ and $\mathcal{F}$ occur disjointly.

Recall that we write $\mathbb{P}_p$ to indicate that A is a p-random subset of $[n]^2$. The following fundamental lemma was proved in 1985 by van den Berg and Kesten [28].

The van den Berg–Kesten Lemma. *Let $\mathcal{E}$ and $\mathcal{F}$ be any two increasing events and let $p \in (0,1)$. Then*

$$\mathbb{P}_p(\mathcal{E} \circ \mathcal{F}) \;\leqslant\; \mathbb{P}_p(\mathcal{E})\,\mathbb{P}_p(\mathcal{F})\,.$$

The authors of [28] also conjectured that their inequality holds in the following more general setting; this was proved 15 years later by Reimer [25].

Reimer's Theorem. *Let $\mathcal{E}$ and $\mathcal{F}$ be any two events and let $p \in (0,1)$. Then*

$$\mathbb{P}_p(\mathcal{E} \circ \mathcal{F}) \;\leqslant\; \mathbb{P}_p(\mathcal{E})\,\mathbb{P}_p(\mathcal{F})\,.$$

We remark that the events which we will need to consider will not all be increasing (or decreasing); however, they will all be obtained by intersecting an increasing event with a decreasing event. For such events the conclusion of Reimer's theorem was proved earlier, by van den Berg and Fiebig [27], and the proof is significantly simpler.

4. The key lemmas

In this section we will state our key bounds (Lemmas 4.3 and 4.4, below) on the probability that a rectangle R is internally filled by the union of A (chosen according to $\mathbb{P}_p$) and a rectangle $S \subset R$. In order to simplify the statement somewhat, we will begin by giving some rather technical definitions, which are illustrated in Figure 1.

Throughout this section, we will assume that $S \subset R$ are rectangles with $\mathrm{short}(S) \geqslant 2$.

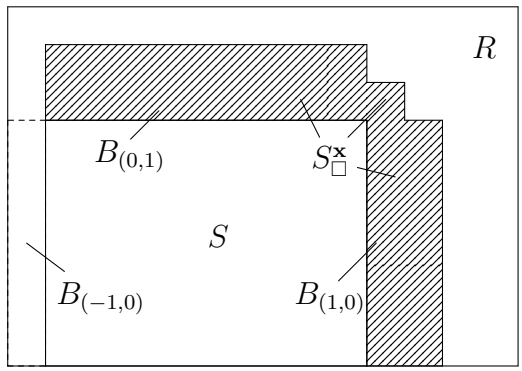


FIGURE 1. An example of a frame. The non-empty buffers and frame $S^{\mathbf{x}}_{\square}$ (hatched) of S in R with $x_{1,0} = x_{0,1} = 1$, $x_{-1,0} = x_{0,-1} = 0$. Note that the buffers may have width 1.

Definition 4.1. The *buffers* of S in R are the sets

$$B_{(i,j)}(S,R) := \Big\{ v \in R \setminus S \,:\, v - (2i, 2j) \in S \Big\},$$

where $(i,j) \in \mathcal{I} := \Big\{(1,0),(0,1),(-1,0),(0,-1)\Big\}$. We call the elements of $\mathcal{I}$ *directions*, define

$$Z(S,R) := \Big\{ \mathbf{d} \in \mathcal{I} \,:\, B_{\mathbf{d}}(S,R) \neq \emptyset \Big\}$$

to be the collection of non-empty buffers of S in R, and set $z(S,R) = |Z(S,R)|$.

Given $\mathbf{x} = (x_{\mathbf{d}})_{\mathbf{d}\in\mathcal{I}} \in \{0,1\}^{\mathcal{I}}$, define the $\mathbf{x}$-*buffer* of S in R to be

$$B^{\mathbf{x}}(S,R) := \bigcup_{\mathbf{d}\in\mathcal{I}\,:\,x_{\mathbf{d}}=1} B_{\mathbf{d}}(S,R),$$

and the $\mathbf{x}$-*frame* of S in R to be the set

$$S^{\mathbf{x}}_{\square} := B^{\mathbf{x}}(S,R) \cup \Big\{ v \in R \setminus S \,:\, |N(v) \cap B^{\mathbf{x}}(S,R)| \geqslant 2 \Big\}.$$

and set $S^{\mathbf{x}}_{\blacksquare} := S \cup S^{\mathbf{x}}_{\square}$. Thus $\mathbf{x}$ encodes the inclusion in $S^{\mathbf{x}}_{\square}$ of some of the (non-empty) buffers, and also the 'corner' site in between two selected buffers. We will write x and y for the number of non-empty horizontal and vertical buffers included in $B^{\mathbf{x}}(S,R)$, i.e.,

$$x := x'_{(1,0)} + x'_{(-1,0)} \qquad \text{and} \qquad y := x'_{(0,1)} + x'_{(0,-1)}\,, \tag{13}$$

where $\mathbf{x}' = \mathbf{x} \cdot \mathbb{1}_{Z(S,R)}$ (i.e., $x'_{\mathbf{d}} := x_{\mathbf{d}}$ if $\mathbf{d} \in Z(S,R)$, and $x'_{\mathbf{d}} := 0$ otherwise).

We are now ready to define our key technical events, which will appear in our hierarchies (see Section 5, below), and are designed to be sufficiently unlikely, and to occur disjointly.

Definition 4.2. Let the rectangles $S \subset R$, and $\mathbf{x} \in \{0,1\}^{\mathcal{I}}$, be as described above.

(**a**) $D^{\mathbf{x}}_1(S,R)$ denotes the event that

$$\Big[S \cup \Big(A \cap R \setminus S^{\mathbf{x}}_{\blacksquare} \Big) \Big] = R\,.$$

(**b**) $D^{\mathbf{x}}_2(S,R)$ denotes the event

$$D^{\mathbf{x}}_1(S,R) \cap \Big\{ A \cap S^{\mathbf{x}}_{\square} = \emptyset \Big\}.$$

The main results of this section are the following two lemmas, which provide us with close to best possible upper bounds on the probabilities of the events $D^{\mathbf{x}}_1(S,R)$ and $D^{\mathbf{x}}_2(S,R)$. The statements are designed to facilitate a proof by induction.

Lemma 4.3. *Let $S \subset R$ be rectangles with* $\dim(R) = (a, b)$ *and* $\dim(S) = (a - s, b - t)$, *let* $\mathbf{x} \in \{0,1\}^{\mathcal{I}}$ *and set* $z = z(S, R)$. *If*

$$L_1 \leqslant \mathrm{short}(R) \leqslant \frac{B}{q} \qquad \text{and} \qquad \mathrm{long}(R) \leqslant \frac{3e^{2B}}{q}\,, \tag{14}$$

and $s, t \leqslant 4\delta\sqrt{\mathrm{short}(R)}$, *then*

$$\mathbb{P}_p\Big(D_1^{\mathbf{x}}(S,R)\Big) \leqslant C^z \left(\frac{C}{\sqrt{a}}\right)^y \left(\frac{C}{\sqrt{b}}\right)^x \exp\Big(-sg(bq) - tg(aq)\Big)\,.$$

Lemma 4.4. *Let $S \subset R$ be rectangles with* $\dim(R) = (a, b)$ *and* $\dim(S) = (a - s, b - t)$, *let* $\mathbf{x} \in \{0,1\}^{\mathcal{I}}$, *and set* $z = z(S, R)$. *If*

$$\mathrm{short}(R) > \frac{B}{q} \qquad \text{and} \qquad \mathrm{long}(R) \leqslant \frac{1}{2q}\log\frac{1}{q} \tag{15}$$

and $s, t \leqslant \dfrac{4\delta}{\sqrt{q}} \cdot \exp\Big(\mathrm{short}(R)\cdot q\Big)$, *then*

$$\mathbb{P}_p\Big(D_2^{\mathbf{x}}(S,R)\Big) \leqslant \Big(Ce^{\mathrm{short}(R)q}\Big)^z \Big(C\sqrt{q}e^{-aq}\Big)^y \Big(C\sqrt{q}e^{-bq}\Big)^x \exp\Big(-sg(bq) - tg(aq)\Big)\,.$$

It will be convenient later when applying these lemmas to combine them as follows. First, the following function encodes the upper bounds on s and t:

$$f(R) := \begin{cases} \delta\sqrt{\mathrm{short}(R)} & \text{if } \mathrm{short}(R) \leqslant \dfrac{B}{q}, \\ \dfrac{\delta}{\sqrt{q}}\exp\Big(\mathrm{short}(R)\cdot q\Big) & \text{otherwise.} \end{cases} \tag{16}$$

Let us say that a rectangle R is 1-*critical* if it satisfies the bounds in (14), and 2-*critical* if it satisfies the bounds in (15). Recall that if S and R have dimensions $\dim(R) = (a, b)$ and $\dim S = (a - s, b - t)$, then

$$Q(S,R) = sg\Big((b-t)q\Big) + tg\Big((a-s)q\Big)\,,$$

and if $\mathbf{x} \in \{0,1\}^{\mathcal{I}}$ then write $\|\mathbf{x}\| := x + y = \sum_{\mathbf{d}\in Z(S,R)} x_{\mathbf{d}}$. The following corollary is an almost immediate consequence of Lemmas 4.3 and 4.4.

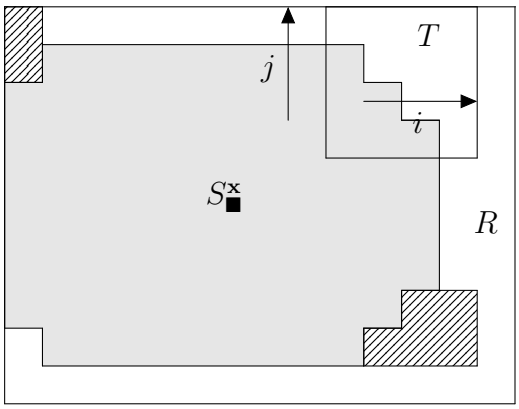
Corollary 4.5. *Let $S \subset R$ be rectangles with* $\dim(R) = (a, b)$ *and* $\dim(S) = (a-s, b-t)$, *let* $\mathbf{x} \in \{0,1\}^{\mathcal{I}}$. *Let* $j \in \{1, 2\}$, *and suppose that R is j-critical. If* $s, t \leqslant 4f(R)$, *then*

$$\mathbb{P}_p\Big(D_j^{\mathbf{x}}(S,R)\Big) \leqslant C^9 \left(\frac{\delta}{f(R)}\right)^{\|\mathbf{x}\|} \exp\Big(-Q(S,R) + 4\phi(R)q\Big)\,. \tag{17}$$

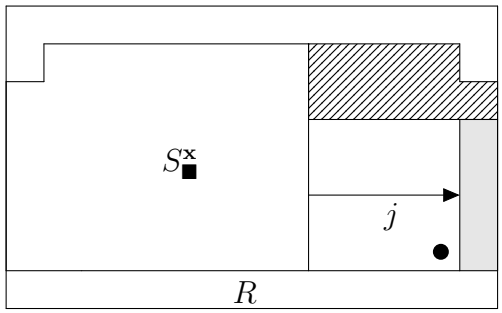
Proof. The claimed inequality follows from those given by Lemmas 4.3 and 4.4 using the bounds on $g'(z)$ given in (8), and noting that $x + y + z \leqslant 8$ and $z \leqslant 4$. To spell out the details, recall from (8) that $g'(z) \geqslant -B/z$ if $z \leqslant B$ and $g'(z) \geqslant -3e^{-2z}$ if $z \geqslant B/2$, and note that $g'(z)$ is increasing and that $f(R) \leqslant \delta/q$. It follows that

$$\begin{aligned} \exp\Big(-sg(bq) - tg(aq)\Big) &\leqslant \exp\Big(-Q(S,R) - 2stq\cdot g'\Big(\Big(\mathrm{short}(R) - 4f(R)\Big)\cdot q\Big)\Big) \\ &\leqslant \exp\Big(-Q(S,R) + \delta\Big) \end{aligned}$$

since $s, t \leqslant 4f(R)$ and $\delta = \delta(B)$ is sufficiently small. ∎



(A) Case 1: the rectangle T is internally filled outside the shaded $S^{\mathbf{x}}_{\blacksquare}$ and allows S to grow i to the right and j upwards.



(B) Case 2: S grows j to the right until it reaches a double gap (shaded). The last column before that is necessarily occupied. In both figures, the hatched region is assumed (in the sketch proof) to be unoccupied.

FIGURE 2. Two possible growth mechanisms.

Since the proofs of Lemmas 4.3 and 4.4 involve a significant amount of quite technical (and not especially illuminating) case analysis, we will give here only a sketch, and postpone the full details to the Appendix.

Sketch of the proof of Lemma 4.3. Let R be a 1-critical rectangle with dimensions $\dim(R) = (a, b)$, and for each x, y, z and each $s, t \leqslant 4f(R)$, set

$$F^{x,y,z}(s,t) := C^z \left(\frac{C}{\sqrt{a}}\right)^y \left(\frac{C}{\sqrt{b}}\right)^x \exp\Big(-sg(bq) - tg(aq)\Big).$$

We will prove, by induction on the pair $(s+t, -(x+y))$, that

$$\mathbb{P}_p\Big(D_1^{\mathbf{x}}(S,R)\Big) \leqslant F^{x,y,z}(s,t) \tag{18}$$

for every $0 \leqslant s, t \leqslant 4f(R)$ and $\mathbf{x} \in \{0,1\}^{\mathcal{I}}$, and every $S \subset R$ with $\dim(S) = (a-s, b-t)$, where x and y are as defined in (13), and $z = z(S,R)$.

The base of the induction is the case $\min\{s,t\} = 0$. Without loss of generality suppose that $t = 0$, and note that this implies that $x = y = 0$, since otherwise $\mathbb{P}_p\Big(D_1^{\mathbf{x}}(S,R)\Big) = 0$. It follows that $R \setminus S$ consists of two rectangles (one of which may be empty), one of which is crossed from left to right, and the other of which is crossed from right to left. By Lemma 3.3, it follows that

$$\mathbb{P}_p\Big(D_1^{\mathbf{x}}(S,R)\Big) \leqslant \exp\Big(-sg(bq)\Big) \leqslant F^{0,0,z}(s,0),$$

as required. We remark that, since $\mathrm{short}(R) \geqslant L_1$, the function $F^{x,y,z}(s,t)$ is increasing in z and decreasing in x, y, s and t.

For the induction step, fix $\mathbf{x} \in \{0,1\}^{\mathcal{I}}$ and $S \subset R$ with $\dim(S) = (a-s, b-t)$, and assume that (18) holds for all smaller values of the pair $(s+t, -(x+y))$ in lexicographical order. We partition into cases, depending on whether or not $z = x + y$.

Case 1: $z = x + y$, i.e., all of the non-empty buffers are included in $S^{\mathbf{x}}_{\square}$.

The key observation in this case is that if the event $D^{\mathbf{x}}_1(S, R)$ holds, then there exists a rectangle $T \subset R$ such that

$$\left[A \cap T \setminus S^{\mathbf{x}}_{\blacksquare}\right] = T \qquad \text{and} \qquad T \cap S^{\mathbf{x}}_{\blacksquare} \neq \emptyset$$

(see Figure 2a). For simplicity, we will assume that $\phi(T) \leqslant 36 f(R)$ (the other case is dealt with in the Appendix), which in particular implies that $\phi(T) \leqslant \delta \cdot \mathrm{short}(S)$.

We will sum over choices of T the probability that

$$\left[A \cap T \setminus S^{\mathbf{x}}_{\blacksquare}\right] = T \qquad \text{and} \qquad \left[S \cup T \cup \left(A \cap R \setminus S^{\mathbf{x}}_{\blacksquare}\right)\right] = R\,. \tag{19}$$

Note that these two events depend on disjoint sets of infected sites, and are therefore independent. To bound the probabilities of these events, we will partition according to $k := \phi(T)$, and the dimensions of $[S \cup T]$,

$$\dim\left([S \cup T]\right) = (a - s + i, b - t + j)\,.$$

Note that $4 \leqslant i + j \leqslant k$, and therefore, by Lemma 3.1, we have

$$|A \cap T \setminus S^{\mathbf{x}}_{\blacksquare}| \geqslant \frac{k}{2} \geqslant \frac{i+j}{2}\,.$$

Note also that, given i, j and k, we have at most $4k$ choices for the rectangle T (at most k per corner of S). Therefore, given i, j and k, the expected number of rectangles T satisfying the first event in (19) is at most

$$4k \binom{k^2}{\lceil k/2 \rceil} p^{\lceil k/2 \rceil} \leqslant (24kp)^{k/2}\,.$$

To bound the probability of the second event in (19), we use the induction hypothesis. To do so, however, we need to split into cases according to whether or not the buffers of $[S \cup T]$ that are not adjacent to T contain any elements of A. In this sketch we will assume that they do not; for the full details see the Appendix.

By the induction hypothesis (under the assumption that no additional infections are found in the buffers), it follows that[3]

$$\mathbb{P}_p\left(\left[S \cup T \cup \left(A \cap R \setminus S^{\mathbf{x}}_{\blacksquare}\right)\right] = R\right) \leqslant F^{x-1,y-1,z}(s - i, t - j)\,,$$

and hence, by the argument above, it will suffice (in this case) to show that

$$\sum_{i+j \geqslant 4} \sum_{k=i+j}^{36 f(R)} (24kp)^{k/2} \cdot F^{x-1,y-1,z}(s - i, t - j) \ll F^{x,y,z}(s, t)\,. \tag{20}$$

To see this, note first that $24kp \leqslant \delta$, since $k \leqslant 36 f(R)$, and that we may therefore assume that $k = i + j$. Now, observe that

$$\frac{F^{x-1,y-1,z}(s - i, t - j)}{F^{x,y,z}(s, t)} = \frac{\sqrt{ab}}{C^2} \exp\left(ig(bq) + jg(aq)\right) \leqslant \frac{\sqrt{ab}}{C^2} \left(\frac{C}{bq}\right)^{i/2} \left(\frac{C}{aq}\right)^{j/2} \tag{21}$$

[3]Note that we used here the bound $z([S \cup T], R) \leqslant z(S, R)$, and the fact that F is increasing in z.

by (7), since $\mathrm{long}(R) \leqslant 3e^{2B}/q$. Since $i+j=k$ and $i,j \geqslant 1$, and recalling that $p \leqslant q$, we have

$$\sum_{k=4}^{36f(R)} \sum_{i+j=k} (24kp)^{k/2} \cdot \frac{\sqrt{ab}}{C^2} \left(\frac{C}{bq}\right)^{i/2} \left(\frac{C}{aq}\right)^{j/2} \leqslant \sum_{k=4}^{36f(R)} \frac{k \cdot (C^2 k)^{k/2}}{\min\{a,b\}^{(k-2)/2}} \leqslant \frac{C^5}{\mathrm{short}(R)},$$

since $\mathrm{short}(R) \geqslant L_1$. Combining this with (21), we obtain (20), as claimed.

Case 2: $z > x+y$, i.e., some non-empty buffer is not included in $S_{\square}^{\mathbf{x}}$.

Without loss of generality, let $B_{(1,0)}(S,R)$ be a non-empty buffer that is not included in $S_{\square}^{\mathbf{x}}$, so $x_{(1,0)} = 0$. The idea is to 'grow' S to the right until we find a double gap, or reach the right-hand side of R, thus leading either to an increase in $x+y$, or a decrease in $s+t$. One significant complication is that before reaching a double gap we might find an infected site in one of the other buffers, which are growing with S (see Figure 2b). In this sketch we will assume that this does not occur, and also that we do not reach the right-hand side of R; the other cases are dealt with in the Appendix.

Let j be the distance to the first double gap to the right of S, that is

$$j := \min\Big\{ i \geqslant 0 \,:\, A \cap R \cap \big(S + (i+2,0)\big) \setminus \big(S + (i,0)\big) = \emptyset \Big\},$$

and denote by $\hat{S} := \bigcup_{i=0}^{j} \big(S + (i,0)\big)$ the rectangle formed by the growth of S to the right, until it reaches that double gap. As noted above, we will assume in this sketch that

$$B_{(1,0)}(\hat{S},R) \neq \emptyset \qquad \text{and} \qquad A \cap \hat{S}_{\square}^{\hat{\mathbf{x}}} \setminus S_{\square}^{\mathbf{x}} = \emptyset\,.$$

where $\hat{\mathbf{x}} := \mathbf{x} + \mathbb{1}_{(1,0)}$ (i.e., $\hat{x}_{(1,0)} = 1$ and $\hat{x}_{\mathbf{d}} = x_{\mathbf{d}}$ for each $(1,0) \neq \mathbf{d} \in \mathcal{I}$). In other words, we found a double gap before reaching the right-hand side of R, and no new infected site was found along the way in any of the buffers. We will sum over choices of j the probability that

$$\Big[S \cup \big(A \cap \hat{S}\big)\Big] = \hat{S} \qquad \text{and} \qquad \Big[\hat{S} \cup \big(A \cap R \setminus \hat{S}_{\blacksquare}^{\hat{\mathbf{x}}}\big)\Big] = R\,. \tag{22}$$

Note that these two events depend on disjoint sets of infected sites, and are therefore independent; we will bound the first using Lemma 3.3, and the second using the induction hypothesis. Indeed, by Lemma 3.3 (and since $g(z)$ is decreasing) we have

$$\mathbb{P}_p\Big(\Big[S \cup \big(A \cap \hat{S}\big)\Big] = \hat{S}\Big) \leqslant \exp\big(-jg(bq)\big),$$

and by the induction hypothesis (under the assumption that no additional infections are found in the buffers and that the right-hand side of R is not reached),

$$\mathbb{P}_p\Big(\Big[\hat{S} \cup \big(A \cap R \setminus \hat{S}_{\blacksquare}^{\hat{\mathbf{x}}}\big)\Big] = R\Big) \leqslant F^{x+1,y,z}(s-j,t)\,.$$

It follows that the probability that there exists $j \geqslant 0$ such that the events in (22) both hold is at most

$$\sum_{j=0}^{s-1} \exp\big(-jg(bq)\big) F^{x+1,y,z}(s-j,t) = \frac{Cs}{\sqrt{b}} \cdot F^{x,y,z}(s,t) \leqslant 4C\delta \cdot F^{x,y,z}(s,t)$$

since $s \leqslant 4\delta\sqrt{\mathrm{short}R}$. Since $\delta = \delta(C) > 0$ was chosen sufficiently small, this bound suffices in this case. For the full details of the proof, see the Appendix. $\square$

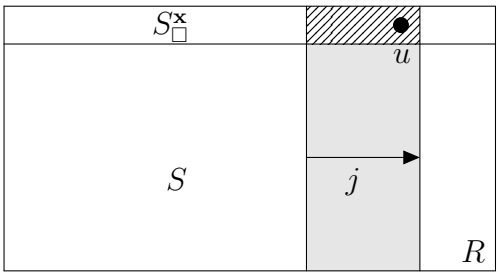


FIGURE 3. S grows j to the right and reaches the infected site u in the hatched region before a double gap to the right. Thus, the shaded region has no vertical double gap.

The proof of Lemma 4.4 is very similar to that of Lemma 4.3, and so we shall give here only a single calculation from the proof, which illustrates the main additional technicality that arises in this setting, and shows why the term $e^{\mathrm{short}(R)qz}$ is needed in the statement of the lemma. The full details can once again be found in the Appendix.

Sketch of the proof of Lemma 4.4. Recall that $D_2^{\mathbf{x}}(S,R)$ denotes the event that

$$\Big[S \cup \Big(A \cap R \setminus S^{\mathbf{x}}_{\blacksquare}\Big)\Big] = R \qquad \text{and} \qquad A \cap S^{\mathbf{x}}_{\square} = \emptyset\,.$$

Let R be a 2-critical rectangle with dimensions $\dim(R) = (a,b)$; as in the proof of Lemma 4.3, we use induction on the pair $(s+t, -(x+y))$, this time to prove that

$$\mathbb{P}_p\Big(D_2^{\mathbf{x}}(S,R)\Big) \leqslant \hat{F}^{x,y,z}(s,t)\,,$$

where

$$\hat{F}^{x,y,z}(s,t) := \Big(Ce^{\mathrm{short}(R)q}\Big)^z \Big(C\sqrt{q}e^{-aq}\Big)^y \Big(C\sqrt{q}e^{-bq}\Big)^x \exp\Big(-sg(bq) - tg(aq)\Big)\,,$$

for every $0 \leqslant s,t \leqslant 4f(R)$ and $\mathbf{x} \in \{0,1\}^{\mathcal{I}}$, and every $S \subset R$ with $\dim(S) = (a-s, b-t)$, where x and y are as defined in (13), and $z = z(S,R)$.

In this sketch we will only consider one very particular (but instructive) configuration, which is illustrated in Figure 3. In this example, the top buffer is of height 1, the left and bottom buffers are empty, and we attempt to grow S to the right in search of a double gap. However, before finding one, we pass an infected site $u \in A$ in the (new part of the) top buffer, which instead causes us to grow upwards by one step.

Let j denote the $\|\cdot\|_\infty$-distance of u from S, and denote by

$$\hat{S}' := \bigcup_{i=0}^{j} \Big(S + (i,0)\Big) \qquad \text{and} \qquad \hat{S} := \hat{S}' \cup \Big(\hat{S}' + (0,1)\Big)$$

so $\hat{S}$ is the rectangle formed by the growth of S to the right, and one step upwards (using u). As noted above, we will assume in this sketch that all of the buffers of $\hat{S}$ are empty except $B_{(1,0)}(\hat{S},R)$. We will sum over choices of j the probability that $A \cap S^{\mathbf{x}}_{\square} = \emptyset$ (as in the definition of the event $D_2^{\mathbf{x}}(S,R)$), that $u \in A$, that there is no double gap to the right of S before it reaches u, and that

$$\Big[\hat{S} \cup \Big(A \cap R \setminus \hat{S}^{\hat{\mathbf{x}}}_{\square}\Big)\Big] = R\,,$$

where $\hat{\mathbf{x}} = \mathbf{x} - \mathbb{1}_{(0,1)} = \mathbf{0}$. Note that these four events are independent, and moreover the probability of each is easy to bound. Indeed, note that $\mathbb{P}_p(u \in A) = p$, that

$$\mathbb{P}_p\Big(A \cap S^{\mathbf{x}}_{\square} = \emptyset\Big) = (1-p)^{a-s} = \exp\Big(-(a-s)q\Big) \leqslant 2 \cdot e^{-aq}$$

since $\mathrm{long}(R) \leqslant (1/(2q))\log(1/q)$, and so $s \leqslant (4\delta/\sqrt{q}) \cdot e^{\mathrm{short}(R)q} \leqslant 4\delta/q$, that

$$\mathbb{P}_p\Big(\hat{S}' \setminus S \text{ has no vertical double gap}\Big) \leqslant \exp\Big(-(j-1)g(bq)\Big)$$

by Lemma 3.3 (and since $g(z)$ is decreasing), and that

$$\mathbb{P}_p\Big(\big[\hat{S} \cup \big(A \cap R \setminus \hat{S}^{\hat{\mathbf{x}}}_{\square}\big)\big] = R\Big) \leqslant \hat{F}^{0,0,1}(s-j,0) = \frac{e^{aq-\mathrm{short}(R)q+jg(bq)+g(aq)}}{C^2\sqrt{q}} \cdot \hat{F}^{0,1,2}(s,1)\,,$$

by the induction hypothesis. It follows that the probability that there exists $j \geqslant 0$ such that the four events above all hold is at most

$$\sum_{j=1}^{s-1} 2p \cdot \frac{e^{-\mathrm{short}(R)q+g(bq)+g(aq)}}{C^2\sqrt{q}} \cdot \hat{F}^{0,1,2}(s,1) \leqslant 4s\sqrt{q} \cdot \frac{e^{-\mathrm{short}(R)q}}{C^2} \cdot \hat{F}^{0,1,2}(s,1)\,, \tag{23}$$

since $p \leqslant q$, and since $a, b \geqslant B/q$ implies that $e^{g(aq)+g(bq)} \leqslant 2$. Finally, recall that $s\sqrt{q} \leqslant 4\delta \cdot e^{\mathrm{short}(R)q}$, so the right-hand side of (23) is at most $(16\delta/C^2) \cdot \hat{F}^{0,1,2}(s,1)$, as required. Once again, see the Appendix for the full details of the proof. □

5. Hierarchies

In this section we will define precisely the family of hierarchies that we will use in the proof of Theorem 2.1. Our definition is more complicated and restrictive than those used in [19, 21], and is designed to take advantage of the bounds proved in Section 4, and to allow us to prove a sufficiently strong upper bound on the number of hierarchies.

5.1. Good and satisfied hierarchies.

Definition 5.1. Let R be a rectangle. A *hierarchy* $\mathcal{H}$ for R is an oriented rooted tree $G_{\mathcal{H}}$ with edges pointing away from the root ("downwards"), with edges e labelled with vectors $\mathbf{x}(e) \in \{0,1\}^{\mathcal{I}}$ and vertices u labelled with rectangles $R_u \subset R$. Let $N_{G_{\mathcal{H}}}(u)$ denote the out-neighbourhood of u in $G_{\mathcal{H}}$. We require them to satisfy the following conditions.

(a) The label of the root is R.

(b) Each vertex has out-degree at most two.

(c) If $v \in N_{G_{\mathcal{H}}}(u)$, then $R_v \subset R_u$.

(d) If $N_{G_{\mathcal{H}}}(u) = \{v, w\}$, then $[R_v \cup R_w] = R_u$.

We will write $L(\mathcal{H})$ for the set of leaves of $G_{\mathcal{H}}$, and refer to the rectangles associated with leaves $u \in L(\mathcal{H})$ as *seeds* of the hierarchy. We will also refer to vertices with out-degree two as *split vertices*.

We next define a subclass of 'good' hierarchies that is sufficiently small to allow us to use the union bound (see Lemma 5.9), but sufficiently large so that every internally filled rectangle R can be associated with a good hierarchy that encodes the growth of the infected sites inside R (see Lemma 5.4). To do so, we will need one more piece of notation: if $S \subset R$ are rectangles with $R = [r_{(-1,0)}, r_{(1,0)}] \times [r_{(0,-1)}, r_{(0,1)}]$ and $S = [s_{(-1,0)}, s_{(1,0)}] \times [s_{(0,-1)}, s_{(0,1)}]$, then we define

$$d_{\mathbf{j}}(S,R) := |r_{\mathbf{j}} - s_{\mathbf{j}}|$$

for each $\mathbf{j} \in \mathcal{I}$, and $d(S,R) := \max\Big\{d_{\mathbf{j}}(S,R) : \mathbf{j} \in \mathcal{I}\Big\}$.

Definition 5.2. We say that a hierarchy $\mathcal{H}$ is *good* if the following conditions hold for every $u \in V(G_{\mathcal{H}})$:

(e) u is a leaf of $G_{\mathcal{H}}$ if and only if $\mathrm{short}(R_u) \leqslant q^{-1/2}$;

(f) if $N_{G_{\mathcal{H}}}(u) = \{v\}$, then
$$d(R_v, R_u) \leqslant 2f(R_u)\,;$$

(g) if $v \in N_{G_{\mathcal{H}}}(u)$ and either $|N_{G_{\mathcal{H}}}(u)| = 2$ or $|N_{G_{\mathcal{H}}}(v)| = 1$, then
$$d(R_v, R_u) \geqslant f(R_u)\,;$$

(h) if $N_{G_{\mathcal{H}}}(u) = \{v\}$ and $|N_{G_{\mathcal{H}}}(v)| = 1$, then $\mathbf{x}(uv) \leqslant \mathbb{1}_{Z(R_v,R_u)}$, and moreover either

(I) $\|\mathbf{x}(uv)\| = z(R_v, R_u)$, or

(II) $\|\mathbf{x}(uv)\| = z(R_v, R_u) - 1$ and $d(R_v, R_u) \in \{f(R_u), f(R_u) + 1\}$;

(i) if $v \in N_{G_{\mathcal{H}}}(u)$ and either $|N_{G_{\mathcal{H}}}(u)| \neq 1$ or $|N_{G_{\mathcal{H}}}(v)| \neq 1$, then $\mathbf{x}(uv) = \mathbf{0}$;

where the function $f(R)$ was defined in (16), and $\|\mathbf{x}(uv)\| = \sum_{\mathbf{d} \in Z(R_v,R_u)} x(uv)_{\mathbf{d}}$.

Finally, we need to define the family of events that we require to occur disjointly. In order to do so, let us first choose a path (the *trunk*) from the root to a leaf of a hierarchy $\mathcal{H}$ by choosing at each split vertex the out-neighbour whose associated rectangle has larger short side (if they are equal, choose arbitrarily). We will write $\mathrm{tr}(\mathcal{H})$ for the set of edges of the trunk.

Definition 5.3. A hierarchy $\mathcal{H}$ is *satisfied* by A if the following events all occur *disjointly*:

(j) If $u \in L(\mathcal{H})$, then the rectangle R_u is internally filled by A;

(k) If $N_{G_{\mathcal{H}}}(u) = \{v\}$ and $uv \notin \mathrm{tr}(\mathcal{H})$, then $D_1^{\mathbf{x}}(R_v, R_u)$ holds, where $\mathbf{x} = \mathbf{x}(uv)$;

(l) If $N_{G_{\mathcal{H}}}(u) = \{v\}$ and $uv \in \mathrm{tr}(\mathcal{H})$, then $D_2^{\mathbf{x}}(R_v, R_u)$ holds, where $\mathbf{x} = \mathbf{x}(uv)$.

We remark that the purpose of the trunk is to guarantee that the unoccupied frames in the events $D_2^{\mathbf{x}}(R_v, R_u)$ occur disjointly. For the sake of brevity, we will often say that a rectangle is in the trunk of a hierarchy $\mathcal{H}$, when we really mean that the associated vertex of $G_{\mathcal{H}}$ is in the trunk, and trust that this will cause no confusion.

5.2. **Fundamental properties.** The following deterministic lemma implies that every internally filled rectangle that satisfies the condition (5) of Theorem 2.1 has a good and satisfied hierarchy.

Lemma 5.4. *Let R be a rectangle that is internally filled by a set A, and suppose that*
$$\mathrm{long}(R) \leqslant \frac{1}{2q} \log \frac{1}{q}\,. \tag{24}$$
Then there exists a good hierarchy $\mathcal{H}$ for R that is satisfied by A.

A similar lemma was proved by Holroyd in [21] using the following lemma, which is a straightforward consequence of the 'rectangles process' of Aizenman and Lebowitz [1].

Lemma 5.5 (Proposition 30 of [21])**.** *Let R be a rectangle with* $\mathrm{long}(R) \geqslant 2$. *If R is internally filled by A, then there exist rectangles $S_1, S_2 \subsetneq R$, with $[S_1 \cup S_2] = R$, that are disjointly internally filled by A.*

We will need the following slight (and straightforward) strengthening of this lemma.

Lemma 5.6. *Let R be a rectangle with* $\mathrm{long}(R) \geqslant 2$. *If R is internally filled by A, then there exist rectangles $S_1, S_2 \subsetneq R$, with $[S_1 \cup S_2] = R$, such that*

$$\Big[A \cap (S_1 \setminus S_2)\Big] = S_1 \qquad \textit{and} \qquad \Big[A \cap (S_2 \setminus S_1)\Big] = S_2\,. \tag{25}$$

Proof. By taking a subset if necessary, we may assume that A is a minimal percolating set for R, i.e., that A is minimal such that $[A] = R$. We claim that for such a set A, the rectangles S_1 and S_2 given by Lemma 5.5 in fact satisfy (25). Indeed, suppose that A_1 and A_2 are disjoint subsets of A such that $S_1 = [A_1]$, $S_2 = [A_2]$ and $[S_1 \cup S_2] = R$, and observe that

$$\Big[A \setminus (A_1 \cap S_2)\Big] = R\,,$$

since $S_2 \subset [A \setminus A_1]$ and $[A] = R$. By the minimality of A, it follows that $A_1 \cap S_2 = \emptyset$, and similarly $A_2 \cap S_1 = \emptyset$ as required. □

Now we prove that any internally filled rectangle has a good and satisfied hierarchy.

Proof of Lemma 5.4. The proof is similar to that of [21, Proposition 32], but since there are several slightly subtle (and important) differences, we will give the details in full.

We prove the statement by induction on $\phi(R)$. If $\mathrm{short}(R) \leqslant q^{-1/2}$ then we can let $\mathcal{H}$ be the hierarchy with only one vertex, which is good by the bound on $\mathrm{short}(R)$, and satisfied since R is internally filled by A. So assume that $\mathrm{short}(R) > q^{-1/2}$, and that the lemma holds for all rectangles with semi-perimeter strictly smaller than $\phi(R)$.

We use Lemma 5.6 to construct a sequence of rectangles

$$R = T_0 \supsetneq T_1 \supsetneq \cdots \supsetneq T_m$$

for some $m \in \mathbb{N}$ as follows. For each $i \geqslant 0$, suppose that we have already constructed T_i, and let T_{i+1} and T_i' be the two rectangles given by Lemma 5.6 applied to T_i, where $d(T_{i+1}, R) \leqslant d(T_i', R)$. Now let m be minimal such that $d(T_m, R) \geqslant f(R)$, and note that m exists because (16) and (24) imply that $\mathrm{long}(R) > 2f(R)$. We consider three cases:

Case 1: $d(T_m, R) \leqslant 2f(R)$.

In this case, instead of applying the induction hypothesis to T_m (as in, e.g., [19, 21]), we let $T_m \subset S \subset R$ be a maximal internally filled rectangle with $d(S, R) \geqslant f(R)$, and apply the induction hypothesis to S. (We remark that this is a crucial step in our proof.) Observe that, by the maximality of S, one of the following two events holds:

(I) There is no element of A within distance two of S. In this case the event $D_2^{\mathbf{x}}(S, R)$ holds for $\mathbf{x} = \mathbb{1}_{Z(S,R)}$, since $A \cap S_\square^{\mathbf{x}} = \emptyset$, and R is internally filled by A.

(II) $d([S \cup \{u\}], R) < f(R)$ for each $u \in A$ within distance two of S, and therefore $d(S, R) \in \{f(R), f(R)+1\}$. Choose $\mathbf{d} \in \mathcal{I}$ such that $d_{\mathbf{d}}(S, R) \in \{f(R), f(R)+1\}$, and set $\mathbf{x} = \mathbb{1}_{Z(S,R)\setminus\{\mathbf{d}\}}$. We claim that the event $D_2^{\mathbf{x}}(S, R)$ holds. Indeed, R is internally filled by A, and if there exists an element $u \in A \cap S_\square^{\mathbf{x}}$, then we have $d([S \cup \{u\}], R) \geqslant d_{\mathbf{d}}(S, R) \geqslant f(R)$, contradicting the maximality of S.

Now, let $\mathcal{H}'$ be the good and satisfied hierarchy for S given by the induction hypothesis, and form a hierarchy $\mathcal{H}$ for R by adding an edge from a vertex u corresponding to R, to the (root) vertex v of $\mathcal{H}'$ corresponding to S. If $|N_{G_{\mathcal{H}}}(v)| \neq 1$, then set $\mathbf{x}(uv) = 0$, and otherwise define $\mathbf{x}(uv)$ as above, i.e., $\mathbf{x}(uv) = \mathbb{1}_{Z(S,R)}$ in (I) and $\mathbf{x}(uv) = \mathbb{1}_{Z(S,R)\setminus\{\mathbf{d}\}}$ in (II), where $d_{\mathbf{d}}(S, R) \in \{f(R), f(R)+1\}$.

We claim that $\mathcal{H}$ is good, and satisfied by A. To see that $\mathcal{H}$ is good, recall that $\mathcal{H}'$ is good, and note that $f(R) \leqslant d(S,R) \leqslant 2f(R)$, and that if $|N_{G_{\mathcal{H}}}(v)| = 1$ then either $\|\mathbf{x}(uv)\| = z(S,R)$ (if there is no element of A within distance two of S), or $\|\mathbf{x}(uv)\| = z(S,R) - 1$ and $d(S,R) \in \{f(R), f(R)+1\}$ (otherwise).

To see that $\mathcal{H}$ is satisfied by A, recall that $\mathcal{H}'$ is satisfied by A, and note that the event $D_2^{\mathbf{x}(uv)}(S,R)$ occurs (by the observations above). Moreover, the event $D_2^{\mathbf{x}(uv)}(S,R)$ depends only on sites in $R \setminus S$, whereas the events involved in $\mathcal{H}'$ depend only on sites inside S. The events involved in $\mathcal{H}$ therefore occur disjointly, as required.

Case 2: $d(T_1, R) > 2f(R)$.

Let $\{S_1, S_2\} = \{T_1, T_0'\}$, where the labelling is chosen so that $\mathrm{short}(S_1) \geqslant \mathrm{short}(S_2)$, and recall that $[S_1 \cup S_2] = R$, that (25) holds, i.e.,

$$\Big[A \cap (S_1 \setminus S_2)\Big] = S_1 \qquad \text{and} \qquad \Big[A \cap (S_2 \setminus S_1)\Big] = S_2\,,$$

and that $2f(R) < \min\{d(S_1,R), d(S_2,R)\}$. Set $A_1 := A \cap S_1$ and $A_2 := A \cap (S_2 \setminus S_1)$ and, applying the induction hypothesis, let $\mathcal{H}_1'$ and $\mathcal{H}_2'$ be good hierarchies for S_1 and S_2 that are satisfied by A_1 and A_2, respectively. Form a hierarchy $\mathcal{H}$ for R by adding edges from a vertex u corresponding to R, to the roots of $\mathcal{H}_1'$ and $\mathcal{H}_2'$, that is, the vertices v_1 and v_2 corresponding to S_1 and S_2 (respectively), and set $\mathbf{x}(uv_1) = \mathbf{x}(uv_2) = \mathbf{0}$.

We claim that $\mathcal{H}$ is good, and satisfied by A. To see that $\mathcal{H}$ is good, recall that $\mathcal{H}_1'$ and $\mathcal{H}_2'$ are good, and that $\min\{d(S_1,R), d(S_2,R)\} \geqslant 2f(R)$. To see that $\mathcal{H}$ is satisfied by A, recall that $\mathcal{H}_1'$ and $\mathcal{H}_2'$ are satisfied by A, and note that the trunk of $\mathcal{H}$ can be chosen to pass through S_1. Now, all of the increasing events involved in $\mathcal{H}_1'$ and $\mathcal{H}_2'$ are witnessed by disjoint subsets of A_1 and A_2, respectively, and A_1 and A_2 are disjoint sets (since $A_1 \subset S_1$ and $A_2 \cap S_1 = \emptyset$), so all of these events occur disjointly. Since S_2 is not in the trunk, the only remaining events are the decreasing events involved in $\mathcal{H}_1'$ (that the frames of rectangles in the trunk are empty), which all depend only on sites in $S_1 \setminus A_1$, and therefore occur disjointly from those that depend on A_1 and A_2. The events involved in $\mathcal{H}$ therefore occur disjointly, as required.

Case 3: $d(T_m, R) > 2f(R)$, and $m > 1$.

Set $S = T_{m-1}$, and let $\{S_1, S_2\} = \{T_m, T_{m-1}'\}$, where the labelling is chosen so that $\mathrm{short}(S_1) \geqslant \mathrm{short}(S_2)$, and recall that $[S_1 \cup S_2] = S$, that (25) holds, and that

$$d(S,R) < f(R) \qquad \text{and} \qquad d(S_i, S) \geqslant d(S_i, R) - d(S,R) > f(R) \tag{26}$$

for each $i \in \{1,2\}$. As in Case 2, set $A_1 := A \cap S_1$ and $A_2 := A \cap (S_2 \setminus S_1)$ and let $\mathcal{H}_1'$ and $\mathcal{H}_2'$ be good hierarchies for S_1 and S_2, satisfied by A_1 and A_2, respectively, given by the induction hypothesis. Form a hierarchy $\mathcal{H}$ for R by adding an edge from a vertex u corresponding to R, to a vertex v corresponding to S, and edges from v to the roots of $\mathcal{H}_1'$ and $\mathcal{H}_2'$, that is, the vertices w_1 and w_2 corresponding to S_1 and S_2 (respectively), and set $\mathbf{x}(uv) = \mathbf{x}(vw_1) = \mathbf{x}(vw_2) = \mathbf{0}$.

It is again easy to see that $\mathcal{H}$ is good, since $\mathcal{H}_1'$ and $\mathcal{H}_2'$ are good, and using the inequalities (26). We claim that moreover $\mathcal{H}$ is satisfied by A; this follows almost exactly as in Case 2, but for completeness we will spell out the details. Recall that $\mathcal{H}_1'$ and $\mathcal{H}_2'$ are satisfied by A, observe that the event $D_2^{\mathbf{x}}(S,R)$ holds (since R is internally filled by A, and $\mathbf{x}(uv) = \mathbf{0}$), and note that the trunk of $\mathcal{H}$ can be chosen to pass through S_1.

Now, all of the increasing events involved in $\mathcal{H}_1'$, $\mathcal{H}_2'$ and $D_2^{\mathbf{x}}(S,R)$ are witnessed by disjoint subsets of A_1, A_2 and $A \setminus S$, respectively, and A_1, A_2 and $A \setminus S$ are disjoint sets (since $A_1 \subset S_1$, $A_2 \cap S_1 = \emptyset$, and $A_1 \cup A_2 \subset S$), so all of these events occur disjointly. Now, S_2 is not in the trunk, and $\mathbf{x}(uv) = \mathbf{0}$, so the only remaining events are the decreasing events involved in $\mathcal{H}_1'$ (that the frames of rectangles in the trunk are empty), which all depend only on sites in $S_1 \setminus A_1$, and therefore occur disjointly from those that depend on A_1, A_2 and $A \setminus S$. The events involved in $\mathcal{H}$ therefore occur disjointly, as required. □

We are now ready to deduce our fundamental bound on the probability that a rectangle is internally filled, cf. [21, Section 10] or [19, Lemma 7]. Given a rectangle R, let us write $\mathcal{H}_R$ for the set of good hierarchies for R, and for each $\mathcal{H} \in \mathcal{H}_R$, set

$$G_{\mathcal{H}}^{(2)} = \mathrm{tr}(G_{\mathcal{H}}) \qquad \text{and} \qquad G_{\mathcal{H}}^{(1)} = E(G_{\mathcal{H}}) \setminus G_{\mathcal{H}}^{(2)}\,.$$

Recall also that $\mathbb{P}_p$ denotes the probability space obtained by choosing A to be a p-random subset of $[n]^2$, and let us write $I(R)$ for the event that R is internally filled by A.

Lemma 5.7. *If R is a rectangle with $\mathrm{long}(R) \leqslant (1/(2q))\log(1/q)$, then*

$$\mathbb{P}_p\big(I(R)\big) \leqslant \sum_{\mathcal{H}\in\mathcal{H}_R} \left(\prod_{j=1}^{2} \prod_{\substack{uv \in G_{\mathcal{H}}^{(j)} \\ N_{G_{\mathcal{H}}}(u)=\{v\}}} \mathbb{P}_p\big(D_j^{\mathbf{x}(uv)}(R_v,R_u)\big)\right)\left(\prod_{u\in L(\mathcal{H})} \mathbb{P}_p\big(I(R_u)\big)\right). \tag{27}$$

Proof. By Lemma 5.4, if R is internally filled by A then there exists a hierarchy $\mathcal{H} \in \mathcal{H}_R$ that is satisfied by A. By the union bound (over $\mathcal{H}_R$), it will therefore suffice to show that for each $\mathcal{H} \in \mathcal{H}_R$, the probability that $\mathcal{H}$ is satisfied by A is bounded above by the corresponding term of the right-hand side of (27). But this follows immediately from Definition 5.3 by Reimer's Theorem, and hence (27) holds, as claimed. □

We remark that we did not actually need the full power of Reimer's Theorem in the proof above, since our events are particularly simple: each is the intersection of an increasing and a decreasing event, and the decreasing events are moreover primitive (i.e., a fixed set must be empty). For events of this form, the conclusion of Reimer's theorem is actually a straightforward consequence of the van den Berg–Kesten lemma.

5.3. **Weighted counting.** Recall that we can bound the probabilities on the right-hand side of (27) using Lemma 3.2 and Corollary 4.5. It therefore remains to control the 'size' of the set $\mathcal{H}_R$; however, since hierarchies with many empty buffers are more numerous and less likely to be satisfied, we would like to give them lower 'weight' when measuring the size of $\mathcal{H}_R$. Due to the form of the right-hand side of (17) (in particular, its dependence on $\|\mathbf{x}\|$), we will find the following definition useful.

Definition 5.8. Given a rectangle R, the *weight* of a hierarchy $\mathcal{H} \in \mathcal{H}_R$ is defined to be

$$w(\mathcal{H}) := \prod_{N_{G_{\mathcal{H}}}(u)=\{v\}} \left(\frac{1}{f(R_u)}\right)^{\|\mathbf{x}(uv)\|}.$$

Given a hierarchy $\mathcal{H}$, we will write $v(\mathcal{H})$ for the number of vertices of $G_{\mathcal{H}}$, and $s(\mathcal{H}) = |L(\mathcal{H})|$ for the number of seeds of $\mathcal{H}$. (Note that $\mathcal{H}$ has exactly $s(\mathcal{H}) - 1$ split vertices.) Given a rectangle R, let us write

$$\mathcal{H}_R(N,M) := \Big\{\mathcal{H} \in \mathcal{H}_R \,:\, v(\mathcal{H}) = N,\, s(\mathcal{H}) = M\Big\}. \tag{28}$$

The following lemma bounds the total weight of $\mathcal{H}_R(N, M)$.

Lemma 5.9. *Let R be a rectangle, and let $N, M \in \mathbb{N}$. Then*

$$\sum_{\mathcal{H} \in \mathcal{H}_R(N,M)} w(\mathcal{H}) \leqslant \exp\Big(16\Big(N + M \log \phi(R)\Big)\Big) .$$

Proof. Let us first fix the tree $G_{\mathcal{H}}$ and the labels $\mathbf{x}(e)$ for each $e \in G_{\mathcal{H}}$. There are at most 3^N oriented rooted trees on N vertices with maximum out-degree at most two (and edges oriented away from the root), and at most 2^{4N} choices for the labels $\mathbf{x}(e) \in \{0,1\}^{\mathcal{I}}$. We will choose the rectangles one by one, starting at the root and working our way down the tree, counting the number of choices (given the earlier choices) at each step.

Let $u \in V(G_{\mathcal{H}})$, and suppose that we have already chosen the rectangle R_u. Suppose first that u is a split vertex, and let $N_{G_{\mathcal{H}}}(u) = \{v, w\}$. We clearly have at most $\phi(R)^4$ choices for each of R_v and R_w, and hence (recalling that there are $M - 1$ split vertices) the total number of choices for the rectangles associated with the out-neighbours of split vertices is at most $\phi(R)^{8M}$. Similarly, if $N_{G_{\mathcal{H}}}(u) = \{v\}$ and v is a split vertex or a seed, then we have at most $\phi(R)^4$ choices for R_v, so the total number of choices for the rectangles associated with such vertices is also at most $\phi(R)^{8M}$.

Suppose now that $N_{G_{\mathcal{H}}}(u) = \{v\}$ and $|N_{G_{\mathcal{H}}}(v)| = 1$, and recall from Definition 5.2 that $d(R_v, R_u) \leqslant 2f(R_u)$, and that either $\|\mathbf{x}(uv)\| = z(R_v, R_u)$, or

$$\|\mathbf{x}(uv)\| = z(R_v, R_u) - 1 \qquad \text{and} \qquad d(R_v, R_u) \in \{f(R_u), f(R_u) + 1\} .$$

In either case we have at most $2^{10} f(R_u)^{\|\mathbf{x}(uv)\|}$ choices for R_v, and it follows that

$$\sum_{\mathcal{H} \in \mathcal{H}_R(N,M)} w(\mathcal{H}) \leqslant (3 \cdot 2^{14})^N \cdot \phi(R)^{16M} \leqslant \exp\Big(16\Big(N + M \log \phi(R)\Big)\Big) ,$$

as claimed. □

5.4. **The height of a hierarchy.** Let us write $h(\mathcal{H})$ for the *height* of the hierarchy $\mathcal{H}$, that is, the number of vertices in the longest path from the root to a leaf of $G_{\mathcal{H}}$. In this subsection we will prove some straightforward (though sometimes slightly technical) properties of the height of a good hierarchy.

Let us begin with a simple lower bound on the size of a seed in a good hierarchy.

Observation 5.10. Let R be a rectangle, and suppose that $\mathrm{long}(R) \leqslant (1/(2q)) \log(1/q)$ and $\mathrm{short}(R) \geqslant q^{-1/2}$. If $\mathcal{H} \in \mathcal{H}_R$, and $v \in V(G_{\mathcal{H}})$, then

$$\phi(R_v) \geqslant \frac{\delta}{q^{1/4}} .$$

Proof. It suffices to prove the claimed bound for seeds of $\mathcal{H}$, so assume that v is a seed, and that $v \in N_{G_{\mathcal{H}}}(u)$ (if $\mathcal{H}$ has only one vertex then the result is trivial). Since u is not a seed and $\mathcal{H}$ is a good hierarchy (see Definition 5.2), we have $\mathrm{short}(R_u) > q^{-1/2}$. Thus

$$\phi(R_v) \geqslant \min\Big\{\phi(R_u) - 8f(R_u),\ f(R_u)\Big\} \geqslant \frac{\delta}{q^{1/4}} ,$$

as required, since if $|N_{G_{\mathcal{H}}}(u)| = 1$ then $d(R_v, R_u) \leqslant 2f(R_u)$, and if $|N_{G_{\mathcal{H}}}(u)| = 2$ then $d(R_v, R_u) \geqslant f(R_u)$. Note that in the first step we used the fact that if $[R_v \cup R_w] = R_u$, then $\phi(R_v) + \phi(R_w) \geqslant \phi(R_u)$, and in the second we used the definition (16) of $f(R)$. □

Next, let us recall a simple but key observation from [19]. Let us say that a seed S is *large* if $\mathrm{long}(S) \geqslant 1/(3\sqrt{q})$, and denote by $m(\mathcal{H})$ the number of large seeds of a hierarchy $\mathcal{H}$. Observe (or recall from [19, Observation 17]) that every non-leaf vertex of a good hierarchy lies above a large seed.

Observation 5.11. Let R be a rectangle with $\mathrm{short}(R) \geqslant q^{-1/2}$. If $\mathcal{H} \in \mathcal{H}_R$, then

$$v(\mathcal{H}) \leqslant 2 \cdot h(\mathcal{H}) \cdot m(\mathcal{H}) \,.$$

Proof. Since $\mathcal{H}$ is a good hierarchy for R, every non-leaf $u \in V(G_{\mathcal{H}})$ lies above a large seed. There are therefore are most $h(\mathcal{H}) \cdot m(\mathcal{H})$ vertices that are either large seeds or non-seeds. Since each small seed is adjacent to a non-seed, and each non-seed is adjacent to at most one small seed, the claimed bound follows. □

We will use Observation 5.11 together with the following lemma to bound the number of vertices in a 'typical' hierarchy $\mathcal{H} \in \mathcal{H}_R$.

Lemma 5.12. *Let R be a rectangle with $\mathrm{long}(R) \leqslant (1/(2q)) \log(1/q)$, and let $\mathcal{H} \in \mathcal{H}_R$. Then either*[4]

$$h(\mathcal{H}) \leqslant \frac{L_2}{\sqrt{q}} \,, \tag{29}$$

or there exists a vertex $u \in V(G_{\mathcal{H}})$ such that either

$$\mathrm{short}(R_u) \leqslant \frac{B}{q} \qquad \textit{and} \qquad \mathrm{long}(R_u) \geqslant 2L_1 \cdot \mathrm{short}(R_u) \,, \tag{30}$$

or

$$\mathrm{short}(R_u) \geqslant \frac{B}{q} \qquad \textit{and} \qquad \mathrm{long}(R_u) \geqslant 4 \cdot \mathrm{short}(R_u) \,. \tag{31}$$

Proof. Suppose that there exists no vertex $u \in V(G_{\mathcal{H}})$ satisfying either (30) or (31); we will show that $h(\mathcal{H}) \leqslant L_2/\sqrt{q}$. To do so, let v be the root of $\mathcal{H}$, let P be a longest path in $G_{\mathcal{H}}$ (from v to a seed w), and partition (the vertex set of) P into sets

$$P_1 := \{v\} \cup \Big\{u \in P : \mathrm{short}(R_u) > B/q\Big\} \qquad \text{and} \qquad P_2 := P \setminus P_1 \,.$$

Let u_1 be the lowest vertex of P_1, and let u_2 be the highest vertex of P_2.

We first claim that the distance (in $G_{\mathcal{H}}$) from u_2 to w is a most $L_1/(3\sqrt{q})$. To see this, note that any $w < y \leqslant u_2$ satisfies $\mathrm{long}(R_y) < 2L_1 \cdot \mathrm{short}(R_y)$, and so, by Definition 5.2, in the next two consecutive steps up P the semi-perimeter increases by at least

$$\frac{\delta}{\sqrt{2L_1}} \cdot \sqrt{\mathrm{long}(R_y)} \,.$$

Since L_1 is large, the claimed bound follows easily.

Similarly, we claim that the distance (in $G_{\mathcal{H}}$) from u_1 to v is a most $L_1/(3\sqrt{q})$. To see this, note that any $v > y \geqslant u_1$ satisfies $4 \cdot \mathrm{short}(R_y) > \mathrm{long}(R_y)$, so in the next two consecutive steps up P, either we reach v, or the semi-perimeter increases by at least

$$\frac{\delta}{\sqrt{q}} \exp\Big(q \cdot \phi(R_y)/5\Big) \,.$$

[4]Recall that $L_2 = L_2(B, C, \delta, L_1)$ is a sufficiently large constant. The lemma also holds with a smaller constant in (29), but this particular tripartition will be convenient in Section 6.

It is again not difficult to see that the claimed bound holds; indeed, the semi-perimeter takes at most $x = L_1/(6\sqrt{q})$ steps to increase by $5/q$, then at most $x/2$ steps to increase by $5/q$ again, and so on, until it has increased by $(5/q)\log_2 x$ in at most $2x$ steps. $\square$

The next lemma bounds the height of $\mathcal{H}$ in the case where only (30) is satisfied.

Lemma 5.13. *Let R be a rectangle with $\mathrm{long}(R) \leqslant (1/(2q))\log(1/q)$, and let $\mathcal{H} \in \mathcal{H}_R$. Suppose that neither (29) nor (31) holds for any vertex $u \in V(G_{\mathcal{H}})$. Then the vertex u satisfying (30) may be chosen so that*

$$h(\mathcal{H}) \leqslant L_1 q^{1/4} \cdot \mathrm{long}(R_u)\,. \tag{32}$$

Proof. Let $u \in V(G_{\mathcal{H}})$ be a vertex satisfying (30) with $\mathrm{long}(R_u)$ maximal, and set $c = \mathrm{short}(R_u)$ and $d = \mathrm{long}(R_u)$. Let P be the longest path in $\mathcal{H}$, and observe that P contains at most $L_1/(3\sqrt{q})$ vertices v with $\mathrm{short}(R_v) > B/q$, as in the proof of Lemma 5.12, since $\mathcal{H}$ contains no vertex such that (31) holds. Observe also that P contains at most $(L_1 q^{1/4}/2) \cdot d$ vertices v with $\mathrm{long}(R_v) \leqslant d$, since it follows from Definition 5.2 that in each two consecutive steps the semi-perimeter increases by at least $\delta q^{-1/4}$.

Finally, we claim that P contains at most $L_2/(3\sqrt{q})$ vertices v with $\mathrm{short}(R_v) \leqslant B/q$ and $\mathrm{long}(R_v) > d$. To see this, note that $2L_1 \cdot \mathrm{short}(R_v) > \mathrm{long}(R_v)$, by our choice of u, and therefore in each two consecutive steps up P, the semi-perimeter increases by at least

$$\frac{\delta}{\sqrt{2L_1}} \cdot \sqrt{\mathrm{long}(R_v)} \geqslant \frac{\delta}{2\sqrt{L_1}} \cdot \sqrt{\phi(R_v)}\,.$$

It now follows easily that after $L_2/(3\sqrt{q})$ steps we have $\phi(R_v) \geqslant 3L_1 B/q$, and hence $\mathrm{short}(R_v) \geqslant B/q$, as claimed.

Since (29) does not hold, it follows that

$$\frac{L_2}{\sqrt{q}} \leqslant h(\mathcal{H}) \leqslant \frac{L_1 q^{1/4}}{2} \cdot \mathrm{long}(R_u) + \frac{L_2}{2\sqrt{q}}\,,$$

and hence we obtain (32), as required. $\square$

Define the *upper trunk* of $\mathcal{H}$ to be the following set of vertices[5] of the trunk:

$$\mathrm{up}(\mathcal{H}) := \Big\{u \in V(G_{\mathcal{H}}) : u \text{ is in the trunk of } \mathcal{H} \text{ and } \mathrm{short}(R_u) \geqslant B/q\Big\}\,.$$

The final lemma of this subsection bounds the sum of the semi-perimeters of rectangles in the upper trunk when there does not exist a vertex $u \in V(G_{\mathcal{H}})$ satisfying (31).

Lemma 5.14. *Let R be a rectangle with $\mathrm{long}(R) \leqslant (1/(2q))\log(1/q)$, and let $\mathcal{H} \in \mathcal{H}_R$. Then either*

$$\sum_{u \in \mathrm{up}(\mathcal{H})} \phi(R_u) \leqslant \frac{L_2}{q^{3/2}},$$

or there exists a vertex $u \in V(G_{\mathcal{H}})$ such that

$$\mathrm{short}(R_u) \geqslant \frac{B}{q} \qquad \textit{and} \qquad \mathrm{long}(R_u) \geqslant 4 \cdot \mathrm{short}(R_u)\,. \tag{33}$$

[5]Recall that $\mathrm{tr}(\mathcal{H})$ denotes the set of *edges* of the trunk; we hope that this minor inconsistency in our notation (which will be quite convenient) will not confuse the reader.

Proof. If $u \in \mathrm{up}(\mathcal{H})$, and u does not satisfy (33), then by Definition 5.2 (as in the proof of Lemma 5.12), in the next two consecutive steps up the trunk either we reach the root v, or the semi-perimeter increases by at least

$$\frac{\delta}{\sqrt{q}} \exp\Big(q \cdot \phi(R_y)/5\Big) .$$

Set $x = L_1/\sqrt{q}$, and observe (cf. the proof of Lemma 5.12) that there are at most $2^{-k+1}x$ vertices $u \in \mathrm{up}(\mathcal{H})$ with $\phi(R_u) \geqslant (B+5k)/q$, for each $0 \leqslant k \leqslant \log_2 x$. It follows that

$$\sum_{u \in \mathrm{up}(\mathcal{H})} \phi(R_u) \leqslant \sum_{k=0}^{\infty} \frac{B+5k}{q} \cdot \frac{x}{2^{k-1}} \leqslant \frac{L_2}{q^{3/2}} ,$$

as required. □

5.5. **The pods of a hierarchy.** To finish this section, let us recall the following important lemma from [21], which is known as the 'pod lemma', and prove a generalization which we be useful in Section 6, below. Recall from (10) the definition of $U(S,R)$.

Lemma 5.15 (Lemma 38 of [21])**.** *Let $\mathcal{H} \in \mathcal{H}_R$. Then there exists a rectangle $S \subset R$ such that*

$$\dim(S) \leqslant \sum_{w \in L(\mathcal{H})} \dim(R_w)$$

and

$$\sum_{N_{G_{\mathcal{H}}}(u)=\{v\}} U(R_v, R_u) \geqslant U(S,R) - 2\Big(s(\mathcal{H})-1\Big) q g(\sqrt{q}) .$$

Holroyd called the rectangle S the *pod* of $\mathcal{H}$. Roughly speaking, Lemma 5.15 says that the 'cost' of the growth (given the size of the seeds) is minimized by placing all of the seeds near to one another, at the very bottom of the hierarchy. However, when we are in the case corresponding to (30) (or, more precisely, Lemma 5.13), we will need to make use of the special rectangle R_u, which is somewhere in the middle of $\mathcal{H}$. In order to use the fact that this rectangle appears in the hierarchy when minimizing the 'cost' of growth, we instead form two pods: one corresponding to the growth inside the special rectangle R_u, the other corresponding to the growth of this rectangle to fill R.

Lemma 5.16. *Let $\mathcal{H} \in \mathcal{H}_R$ and let $u \in V(G_{\mathcal{H}})$. Then there exist rectangles $S_1 \subset R_u$ and $R_u \subset S_2 \subset R$, such that*

$$\dim(S_1) + \dim(S_2) - \dim(R_u) \leqslant \sum_{w \in L(\mathcal{H})} \dim(R_w) \tag{34}$$

and

$$\sum_{N_{G_{\mathcal{H}}}(v)=\{w\}} U(R_w, R_v) \geqslant U(S_1, R_u) + U(S_2, R) - 2\Big(s(\mathcal{H})-1\Big) q g(\sqrt{q}) . \tag{35}$$

The proof of Lemma 5.16 is essentially the same as that of Lemma 5.15, and so we will give only a brief sketch here, and refer the reader to [21] for the details.

Sketch proof of Lemma 5.16. We will use induction on the distance from u to the root. Note first that when u is the root of $\mathcal{H}$, then the claimed conclusion follows from Lemma 5.15 by setting $S_1 = S$ and $S_2 = R$. For the induction step, we divide into cases according to whether the root has one or two neighbours.

Indeed, suppose first that the root has one neighbour, x, and apply the induction hypothesis to the sub-hierarchy of $\mathcal{H}$ rooted at x to obtain pods S_1 and S_2. Note that these pods satisfy (34), and also (35), since the inequality

$$U(R_x, R) \geqslant U(S_2, R) - U(S_2, R_x)\,,$$

follows immediately from the definition. On the other hand, if the root has two neighbours, x and y, and u is a descendant of x, then we apply the induction hypothesis to the sub-hierarchy of $\mathcal{H}$ rooted at x, giving pods S_1' and S_2', and Lemma 5.15 to the sub-hierarchy of $\mathcal{H}$ rooted at y, giving a pod T. Set $S_1 := S_1'$, and choose S_2, with

$$\dim(S_2') \leqslant \dim(S_2) \leqslant \dim(S_2') + \dim(T)$$

and

$$U(S_2', R_x) + U(T, R_y) \geqslant U(S_2, R) - 2qg(\sqrt{q})$$

by applying [21, Proposition 15], exactly as in the proof of [21, Lemma 38]. Noting that $s(\mathcal{H}) = s(\mathcal{H}_x) + s(\mathcal{H}_y)$, the inequalities (34) and (35) follow. □

6. Proof of Theorem 1.1

In this section we will put the pieces together and prove Theorem 1.1. The main step is the proof of Theorem 2.1, which we restate (this time with explicit constants) for convenience. Recall that $I(R)$ denotes the event that R is internally filled by A.

Theorem 6.1. *Let R be a rectangle with dimensions $\dim(R) = (a, b)$, and suppose that $a \leqslant b$, and*

$$\frac{3e^{2B}}{q} \leqslant b \leqslant \frac{1}{2q}\log\frac{1}{q}\,. \tag{36}$$

Then

$$\mathbb{P}_p\big(I(R)\big) \leqslant \exp\left(-\min\left\{\frac{2\lambda}{q} + \frac{1}{q^{3/4}},\ (b-a)g(aq) + \frac{2}{q}\int_0^{aq} g(z)\,\mathrm{d}z - \frac{L_6}{\sqrt{q}}\right\}\right)\,.$$

We will begin by giving an outline of the proof of Theorem 6.1, and proving a couple of straightforward technical lemmas. Let us fix a rectangle R as in the theorem until the end of its proof. The first step is to recall that

$$\mathbb{P}_p\big(I(R)\big) \leqslant \sum_{\mathcal{H}\in\mathcal{H}_R}\left(\prod_{j=1}^{2}\prod_{\substack{uv\in G_{\mathcal{H}}^{(j)}\\ N_{G_{\mathcal{H}}}(u)=\{v\}}}\mathbb{P}_p\big(D_j^{\mathbf{x}(uv)}(R_v, R_u)\big)\right)\left(\prod_{u\in L(\mathcal{H})}\mathbb{P}_p\big(I(R_u)\big)\right), \tag{37}$$

by Lemma 5.7, where we used the upper bound on $\mathrm{long}(R)$ given by (36). Recall that $g(z)$ is decreasing, and (from Definition 5.2) that $\mathrm{short}(R_u) \leqslant q^{-1/2}$ for every leaf $u \in L(\mathcal{H})$ of a good hierarchy $\mathcal{H}$. Therefore, if $\mathcal{H} \in \mathcal{H}_R$, then

$$\mathbb{P}_p\big(I(R_u)\big) \leqslant 3^{\phi(R_u)}\exp\Big(-\phi(R_u)g(\sqrt{q})\Big) \tag{38}$$

for each leaf $u \in L(\mathcal{H})$, by Lemma 3.2. Moreover, since $d(R_v, R_u) \leqslant 2f(R_u)$ whenever $N_{G_{\mathcal{H}}}(u) = \{v\}$, if R_u is j-critical for some $j \in \{1, 2\}$, then

$$\mathbb{P}_p\big(D_j^{\mathbf{x}(uv)}(R_v, R_u)\big) \leqslant C^9\left(\frac{\delta}{f(R_u)}\right)^{\|\mathbf{x}(uv)\|}\exp\Big(-Q(R_v, R_u) + 4\phi(R_u)q\Big) \tag{39}$$

by Corollary 4.5, where $Q(R_v, R_u)$ was defined in (11). Unfortunately, however, some rectangles are neither 1- nor 2-critical, and we must deal with these separately.

Lemma 6.2. *The probability that there exists an internally filled rectangle $S \subset R$ with*

$$\mathrm{short}(S) \leqslant \frac{B}{q} \qquad \textit{and} \qquad \mathrm{long}(S) \geqslant \frac{3e^{2B}}{q} \tag{40}$$

or an internally filled rectangle $S \subset R$ with

$$\mathrm{short}(S) \leqslant \frac{1}{q} \qquad \textit{and} \qquad \mathrm{long}(S) \geqslant \frac{B}{2q} \tag{41}$$

is at most $e^{-2/q}$.

Proof. Observe that if $S \subset R$ is internally filled, then it must be crossed from left to right, and from bottom to top. By Lemma 3.3, it follows that if S satisfies (40) then

$$\mathbb{P}_p\Big(I(S)\Big) \leqslant \exp\Big(-\mathrm{long}(S)\cdot g\Big(q\cdot \mathrm{short}(S)\Big)\Big) \leqslant \exp\left(-\frac{3e^{2B}g(B)}{q}\right).$$

Recalling that $g(z) \sim e^{-2z}$ as $z \to \infty$ (and that B is large), and applying the union bound, it follows that the probability that there exists such a rectangle S is at most

$$\Big(\mathrm{long}(R)\Big)^4 \cdot \exp\left(-\frac{3e^{2B}g(B)}{q}\right) \leqslant \frac{1}{2}\cdot e^{-2/q}.$$

The same bound (with the same proof, noting that $B > 4/g(1)$, since B is sufficiently large) holds if S satisfies (41). The result then follows by the union bound. □

Note that $\lambda = \pi^2/18 < 1$, so by Lemma 6.2 we may assume that R contains no internally filled rectangle S satisfying (40) or (41). It follows that each rectangle R_u (where $u \in V(G_{\mathcal{H}})$) either satisfies the condition (14) of Lemma 4.3 (and hence is 1-critical), or satisfies the condition (15) of Lemma 4.4 (and hence is 2-critical). Note also that, since $b \geqslant 3e^{2B}/q$, by (36), we may assume from now on that $a \geqslant B/q$.

The next problem is that we would like R_u to be j-critical when $uv \in G^{(j)}_{\mathcal{H}}$, and this is not necessarily the case. However, since $D_2^{\mathbf{x}(uv)}(R_v, R_u) \subset D_1^{\mathbf{x}(uv)}(R_v, R_u)$, it is not a problem if $uv \in G^{(2)}_{\mathcal{H}} = \mathrm{tr}(G_{\mathcal{H}})$ for some u with $\mathrm{short}(R_u) \leqslant B/q$. The next lemma bounds the probability that there exists $uv \in G^{(1)}_{\mathcal{H}}$ with $N_{G_{\mathcal{H}}}(u) = \{v\}$ and $\mathrm{short}(R_u) > B/q$.

Lemma 6.3. *The probability that there exist two disjointly internally filled rectangles $S_1, S_2 \subset R$ with*

$$\min\Big\{\mathrm{short}(S_1), \mathrm{short}(S_2)\Big\} \geqslant \frac{B}{q} \tag{42}$$

is at most $e^{-2/q+o(1/q)}$.

This lemma is an almost immediate consequence of Holroyd's theorem and the van den Berg–Kesten lemma. However, for convenience (since the version we need is not explicitly stated in [21]) we will deduce it from the following (very weak) consequence of [19, Proposition 15], which holds since $2\lambda = \pi^2/9 > 1$.

Proposition 6.4. *Let $S \subset R$ be a rectangle with* $\mathrm{short}(S) \geqslant B/q$. *Then*

$$\mathbb{P}_p\Big(I(S)\Big) \leqslant e^{-1/q}\,.$$

Proof of Lemma 6.3. By the van den Berg–Kesten inequality and Proposition 6.4, the probability that two given rectangles S_1 and S_2, each with short side at least B/q, are disjointly internally filled is at most $e^{-2/q}$. By the union bound, it follows that the probability that two such disjointly internally filled rectangles exist is at most

$$\Big(\mathrm{long}(R)\Big)^8 e^{-2/q} = e^{-2/q+o(1/q)}\,,$$

as claimed. □

Note that if there exists a vertex u with $\mathrm{short}(R_u) \geqslant B/q$ that is not in the trunk, then there must exist a split vertex above u whose neighbours are labelled with disjointly internally filled rectangles S_1 and S_2 satisfying (42). Hence, by Lemma 6.3, and recalling that $\lambda < 1$, we may assume that every vertex $u \in V(G_{\mathcal{H}})$ with $\mathrm{short}(R_u) \geqslant B/q$ is in the trunk, and hence $uv \in \mathrm{tr}(\mathcal{H})$ whenever $N_{G_{\mathcal{H}}}(u) = \{v\}$ and $\mathrm{short}(R_u) > B/q$. We may therefore apply the inequality (39) to bound the probability of the event $D_j^{\mathbf{x}(uv)}(R_v, R_u)$ for each $uv \in G_{\mathcal{H}}^{(j)}$ with $N_{G_{\mathcal{H}}}(u) = \{v\}$. Setting

$$X(\mathcal{H}) := \sum_{u \in L(\mathcal{H})} \phi(R_u)\,,$$

it follows from (37), (38) and (39), and Lemmas 6.2 and 6.3, that the probability that R is internally filled is bounded from above by $e^{-2/q+o(1/q)}$ plus

$$\sum_{\mathcal{H} \in \mathcal{H}_R^*} 3^{X(\mathcal{H})} e^{-X(\mathcal{H})g(\sqrt{q})} \prod_{N_{G_{\mathcal{H}}}(u)=\{v\}} C^9 \left(\frac{\delta}{f(R_u)}\right)^{\|\mathbf{x}(uv)\|} \exp\Big(-Q(R_v, R_u) + 4\phi(R_u)q\Big)\,,$$

where $\mathcal{H}_R^*$ denotes the set of hierarchies $\mathcal{H} \in \mathcal{H}_R$ that contain no rectangle satisfying either (40) or (41), and such that every vertex $u \in V(G_{\mathcal{H}})$ with $\mathrm{short}(R_u) \geqslant B/q$ is in the trunk. By Definition 5.8, this is at most

$$\sum_{\mathcal{H} \in \mathcal{H}_R^*} w(\mathcal{H}) \cdot C^{9v(\mathcal{H})} \cdot 3^{X(\mathcal{H})} e^{-X(\mathcal{H})g(\sqrt{q})} \prod_{N_{G_{\mathcal{H}}}(u)=\{v\}} \exp\Big(-Q(R_v, R_u) + 4\phi(R_u)q\Big)\,. \tag{43}$$

The rest of the proof of Theorem 6.1 is just a careful analysis of (43).

Proof of Theorem 6.1. As explained above, by Lemmas 3.2, 5.7, 6.2 and 6.3, and Corollary 4.5, to prove the theorem it will suffice to bound (43). Let us set

$$\Lambda(\mathcal{H}) := w(\mathcal{H}) \cdot C^{9v(\mathcal{H})} \cdot 3^{X(\mathcal{H})} e^{-X(\mathcal{H})g(\sqrt{q})} \prod_{N_{G_{\mathcal{H}}}(u)=\{v\}} \exp\Big(-Q(R_v, R_u) + 4\phi(R_u)q\Big)$$

for each $\mathcal{H} \in \mathcal{H}_R^*$, and write $\mathcal{H}_R^{(1)}$ for the set of $\mathcal{H} \in \mathcal{H}_R^*$ such that

$$h(\mathcal{H}) \leqslant \frac{L_2}{\sqrt{q}} \qquad \text{and} \qquad \sum_{u \in \mathrm{up}(\mathcal{H})} \phi(R_u) \leqslant \frac{L_2}{q^{3/2}}\,, \tag{44}$$

cf. Lemmas 5.12 and 5.14. Let us note that this is the most important class of hierarchies, since it will turn out that the remaining hierarchies $\mathcal{H}_R^* \setminus \mathcal{H}_R^{(1)}$ contribute only smaller order terms to (43). To slightly simplify the formulae below, let us write

$$\mathcal{J}(R) := \frac{2}{q}\int_0^{aq} g(z)\,\mathrm{d}z + (b-a)g(aq)\,,$$

where we recall that $a = \text{short}(R)$ and $b = \text{long}(R)$.

Claim 1: $\displaystyle\sum_{\mathcal{H}\in\mathcal{H}_R^{(1)}} \Lambda(\mathcal{H}) \leqslant \exp\left(-\mathcal{J}(R) + \frac{L_5}{\sqrt{q}}\right) + e^{-2/q}.$

Proof of Claim 1. The proof is a fairly standard (if somewhat complicated) calculation, similar to, e.g., [19], the main new ingredient being the weighted counting of Lemma 5.9. The first step is to deal with hierarchies with $X(\mathcal{H}) > 1/q$, and to do so we will first show that

$$\Lambda(\mathcal{H}) \leqslant w(\mathcal{H}) \cdot C^{10v(\mathcal{H})} \cdot \exp\left(-\frac{X(\mathcal{H})}{5}\log\frac{1}{q} + \frac{4L_2}{\sqrt{q}}\right) \tag{45}$$

for every $\mathcal{H} \in \mathcal{H}_R^{(1)}$. To see this, recall first that every vertex u with $\text{short}(R_u) \geqslant B/q$ is in the trunk, and no rectangle that appears in $\mathcal{H}$ satisfies (40). It follows that $\text{short}(R_u) \leqslant B/q$ and $\text{long}(R_u) \leqslant 3e^{2B}/q$ for every $u \notin \text{up}(\mathcal{H})$, and hence, by (44), we have

$$\prod_{N_{G_{\mathcal{H}}}(u)=\{v\}} \exp\left(4\phi(R_u)q\right) \leqslant C^{v(\mathcal{H})} \exp\left(4L_2/\sqrt{q}\right), \tag{46}$$

since $C = C(B) > 0$ was chosen sufficiently large. Next, observe that

$$3^{X(\mathcal{H})}e^{-X(\mathcal{H})g(\sqrt{q})} \leqslant \exp\left(-\frac{X(\mathcal{H})}{4}\left(\log\frac{1}{3^4 q} - 4q^{1/4}\right)\right) \leqslant \exp\left(-\frac{X(\mathcal{H})}{5}\log\frac{1}{q}\right), \tag{47}$$

since $g(\sqrt{q}) \geqslant \log(q^{-1/4}) - q^{1/4}$, by (6). Noting that $Q(R_v, R_u) \geqslant 0$ for every $R_v \subset R_u$, since $g(z)$ is positive, and using (46) and (47), we obtain (45), as claimed.

Now, recall that $m(\mathcal{H})$ denotes the number of large seeds in a hierarchy $\mathcal{H}$, and that

$$v(\mathcal{H}) \leqslant 2 \cdot h(\mathcal{H}) \cdot m(\mathcal{H})$$

by Observation 5.11, and observe that therefore

$$X(\mathcal{H}) \geqslant \frac{m(\mathcal{H})}{3\sqrt{q}} + \frac{\delta(s(\mathcal{H}) - m(\mathcal{H}))}{q^{1/4}} \geqslant \frac{v(\mathcal{H})}{L_3} + \frac{s(\mathcal{H})}{q^{1/5}}, \tag{48}$$

by Observation 5.10 and (44). We claim next that

$$\sum_{\mathcal{H}\in\mathcal{H}_R^{(1)}:\, X(\mathcal{H})>1/q} \Lambda(\mathcal{H}) \leqslant e^{-2/q}. \tag{49}$$

To prove (49), let us write

$$\mathcal{H}_R^{(1)}(N, M) := \left\{\mathcal{H} \in \mathcal{H}_R^{(1)} : v(\mathcal{H}) = N,\ s(\mathcal{H}) = M\right\},$$

as in (28), and recall that, by Lemma 5.9,

$$\sum_{\mathcal{H}\in\mathcal{H}_R^{(1)}(N,M)} w(\mathcal{H}) \leqslant \exp\left(16\left(N + M\log\phi(R)\right)\right). \tag{50}$$

Combining (45) with (48) and (50), it follows that

$$\begin{aligned}\sum_{\substack{\mathcal{H}\in\mathcal{H}_R^{(1)}\\ X(\mathcal{H})>1/q}}\Lambda(\mathcal{H}) &\leqslant \sum_{N,M}\sum_{\substack{\mathcal{H}\in\mathcal{H}_R^{(1)}(N,M)\\ X(\mathcal{H})>1/q}} w(\mathcal{H})\cdot C^{10v(\mathcal{H})}\cdot\exp\Big(-\frac{X(\mathcal{H})}{5}\log\frac{1}{q}+\frac{4L_2}{\sqrt{q}}\Big)\\ &\leqslant e^{-2/q}\sum_{N,M}\exp\big(-CN-Mq^{-1/5}\big)\sum_{\mathcal{H}\in\mathcal{H}_R^{(1)}(N,M)} w(\mathcal{H})\\ &\leqslant e^{-2/q}\sum_{N,M}\exp\big(-N-M\big)\leqslant e^{-2/q},\end{aligned}$$

as claimed, where in the second step we used the bound $X(\mathcal{H})>1/q\gg 1$.

We will therefore assume from now on that

$$X(\mathcal{H})\leqslant\frac{1}{q}\,. \tag{51}$$

We next claim that

$$\Lambda(\mathcal{H})\leqslant w(\mathcal{H})\cdot C^{10v(\mathcal{H})}\cdot\exp\Big(-\frac{U(S,R)}{q}-\frac{X(\mathcal{H})}{4}\log\frac{1}{q}+13L_2\cdot X(\mathcal{H})\Big)\,. \tag{52}$$

To prove this we repeat the proof of (45), being slightly less wasteful in (47), and replacing the trivial bound $Q(S,R)\geqslant 0$ by a more complicated argument. To be more precise, recall that $U(R_v,R_u)\leqslant q\cdot Q(R_v,R_u)$ for every $R_v\subset R_u$, by Lemma 3.8, and that therefore, by Lemma 5.15, there exists a pod S, with $\phi(S)\leqslant X(\mathcal{H})$, such that

$$\sum_{N_{G_{\mathcal{H}}}(u)=\{v\}} Q(R_v,R_u)\geqslant\sum_{N_{G_{\mathcal{H}}}(u)=\{v\}}\frac{U(R_v,R_u)}{q}\geqslant\frac{U(S,R)}{q}-2s(\mathcal{H})g(\sqrt{q})\,. \tag{53}$$

Now, recall from (48) that we have $X(\mathcal{H})\geqslant s(\mathcal{H})\cdot q^{-1/5}\gg s(\mathcal{H})g(\sqrt{q})$, and note that $X(\mathcal{H})\geqslant 1/(3\sqrt{q})$, since every hierarchy in $\mathcal{H}_R$ has at least one large seed. Hence, by (46), (47) and (53), we obtain (52), as claimed.

Now, by Lemma 3.7, we have

$$\frac{U(S,R)}{q}\geqslant\frac{2}{q}\int_0^{aq}g(z)\,\mathrm{d}z+(b-a)g(aq)-\frac{\phi(S)}{2}\log\Big(1+\frac{1}{\phi(S)q}\Big)-O\big(\phi(S)\big)\,,$$

since (51) implies[6] that $\phi(S)\leqslant X(\mathcal{H})\leqslant 1/q\leqslant a$, and hence

$$\frac{U(S,R)}{q}\geqslant\mathcal{J}(R)-\frac{X(\mathcal{H})}{2}\log\Big(1+\frac{1}{X(\mathcal{H})q}\Big)-O\big(X(\mathcal{H})\big)\,,$$

since the function $x\mapsto x\log\big(1+\frac{1}{x}\big)$ is increasing. Combining this with (52), recalling that $v(\mathcal{H})\leqslant L_3\cdot X(\mathcal{H})$, by (48), and noting that $1+\frac{1}{X(\mathcal{H})q}\leqslant\frac{2}{X(\mathcal{H})q}$, by (51), we obtain

$$\Lambda(\mathcal{H})\leqslant w(\mathcal{H})\cdot\exp\Big(-\mathcal{J}(R)-\frac{X(\mathcal{H})}{2}\log\frac{X(\mathcal{H})\sqrt{q}}{L_4}\Big) \tag{54}$$

where $L_4=C^{O(L_3)}$. Finally, observe that, by (48) and (50), we have

$$\sum_{\mathcal{H}\in\mathcal{H}_R^{(1)}:\,X(\mathcal{H})=x} w(\mathcal{H})\leqslant\exp\big(O\big(L_3\cdot x\big)\big)$$

[6]Recall that, by Lemma 6.2, we may assume that $a\geqslant B/q$.

for any $x \in \mathbb{N}$. It follows that

$$\sum_{\substack{\mathcal{H}\in\mathcal{H}_R^{(1)}\\ X(\mathcal{H})\leqslant 1/q}} \Lambda(\mathcal{H}) \leqslant e^{-\mathcal{J}(R)} \sum_{x=1/(3\sqrt{q})}^{1/q} \exp\left(-\frac{x}{2}\log\frac{x\sqrt{q}}{L_4}\right) \sum_{\substack{\mathcal{H}\in\mathcal{H}_R^{(1)}\\ X(\mathcal{H})=x}} w(\mathcal{H})$$

$$\leqslant e^{-\mathcal{J}(R)} \sum_{x=1/(3\sqrt{q})}^{1/q} \exp\left(-\frac{x}{2}\log\frac{x\sqrt{q}}{L_4} + O\big(L_3\cdot x\big)\right) \leqslant \exp\left(-\mathcal{J}(R) + \frac{L_5}{\sqrt{q}}\right),$$

where $L_5 = L_4 \cdot e^{O(L_3)}$, since the summand decreases super-exponentially quickly once $x\sqrt{q}$ is larger than this. This completes the proof of Claim 1. □

If $\mathcal{H} \in \mathcal{H}_R^* \setminus \mathcal{H}_R^{(1)}$ then, by Lemmas 5.12 and 5.14, there exists a vertex $u \in V(G_{\mathcal{H}})$ satisfying either (30) or (31). The rest of the proof consists of bounding the contribution to (43) of hierarchies containing such a vertex. In order to simplify the argument, it will be convenient to first (in Claims 2 and 3) deal with those hierarchies in which either $v(\mathcal{H})$ or $X(\mathcal{H})$ is unusually large. We then (in Claims 4 and 5) consider the remaining hierarchies with an 'abnormal' vertex, i.e., one satisfying either (30) or (31).

We begin by considering hierarchies with unusually many vertices. Let us write $\mathcal{H}_R^{(2)}$ for the set of $\mathcal{H} \in \mathcal{H}_R^* \setminus \mathcal{H}_R^{(1)}$ such that

$$v(\mathcal{H}) \geqslant 8\cdot s(\mathcal{H}) + \frac{4L_1}{q^{3/4}}\,. \tag{55}$$

For such hierarchies we will prove the following stronger bound.

Claim 2: $\displaystyle\sum_{\mathcal{H}\in\mathcal{H}_R^{(2)}} \Lambda(\mathcal{H}) \leqslant e^{-2/q}$.

Proof of Claim 2. The first step is to prove the following bounds,

$$|\mathrm{up}(\mathcal{H})| \leqslant \frac{L_1}{\sqrt{q}}\log\frac{1}{q} \qquad \text{and} \qquad \sum_{u\in\mathrm{up}(\mathcal{H})} \phi(R_u) \leqslant \frac{L_1}{q^{3/2}}\left(\log\frac{1}{q}\right)^2 \tag{56}$$

which replace those in (44), and hold for any $\mathcal{H} \in \mathcal{H}_R$. Both follow immediately from the upper bound on $b = \mathrm{long}(R)$ in (36), and the observation that in two consecutive steps of $\mathrm{up}(\mathcal{H})$, the semi-perimeter of the corresponding rectangles grows by at least $\delta/\sqrt{q}$.

Now, for each $\mathcal{H} \in \mathcal{H}_R^{(2)}$, consider the set $\mathcal{Y}(\mathcal{H})$ of edges $uv \in G_{\mathcal{H}}$ such that

$$\mathrm{short}(R_u) \leqslant \frac{B}{q}, \qquad N_{G_{\mathcal{H}}}(u) = \{v\} \qquad \text{and} \qquad |N_{G_{\mathcal{H}}}(v)| = 1\,.$$

We claim that

$$|\mathcal{Y}(\mathcal{H})| \geqslant v(\mathcal{H}) - 4s(\mathcal{H}) - |\mathrm{up}(\mathcal{H})| \geqslant \frac{v(\mathcal{H})}{2} + \frac{L_1}{q^{3/4}}\,. \tag{57}$$

To see this, recall that $\mathcal{H}$ has $s(\mathcal{H})$ seeds and $s(\mathcal{H})-1$ split vertices, and so there are at most $4s(\mathcal{H})-2$ vertices $u \in V(G_{\mathcal{H}})$ that are either seeds, or split-vertices, or have a single out-neighbour that is a seed or a split vertex. Moreover, $\mathcal{H} \in \mathcal{H}_R^*$ implies that every vertex $u \in V(G_{\mathcal{H}})$ with $\mathrm{short}(R_u) \geqslant B/q$ is in the trunk. The second inequality follows from (55) and (56).

We next claim that

$$\Lambda(\mathcal{H}) \leqslant w(\mathcal{H}) \cdot C^{10v(\mathcal{H})} \cdot \exp\left(\frac{4L_1}{\sqrt{q}}\left(\log\frac{1}{q}\right)^2 - \frac{\delta^2|\mathcal{Y}(\mathcal{H})|}{q^{1/4}}\right). \tag{58}$$

The proof of this is similar to that of (45). Indeed, we obtain a slightly weaker bound in place of (46) by using (56) instead of (44), and (47) still holds, and the right-hand side is at most 1. Moreover, $\mathcal{H} \in \mathcal{H}_R^*$ implies that $\mathrm{long}(R_u) \leqslant 3e^{2B}/q$ for each edge $uv \in \mathcal{Y}(\mathcal{H})$, and therefore

$$Q(R_u, R_v) \geqslant \frac{\delta g(3e^{2B})}{q^{1/4}} \geqslant \frac{\delta^2}{q^{1/4}},$$

for each such edge, since $d(R_u, R_v) \geqslant f(R_u) \geqslant \delta\sqrt{\mathrm{short}(R_u)} \geqslant \delta q^{-1/4}$, by Definition 5.2 and (16), and since $\delta = \delta(B)$ was chosen sufficiently small. Plugging these bounds into the definition of $\Lambda(\mathcal{H})$, we obtain (58).

By (57), it follows that

$$\Lambda(\mathcal{H}) \leqslant w(\mathcal{H}) \cdot \exp\left(-\frac{\delta^3|\mathcal{Y}(\mathcal{H})|}{q^{1/4}}\right),$$

and hence, by (57) and Lemma 5.9, and since $\phi(R) \leqslant 1/q^2$, we obtain

$$\begin{aligned}
\sum_{\mathcal{H}\in\mathcal{H}_R^{(2)}} \Lambda(\mathcal{H}) &\leqslant \sum_{y \geqslant L_1 q^{-3/4}} \exp\left(-\frac{\delta^3 y}{q^{1/4}}\right) \sum_{M=1}^{2y} \sum_{N=M}^{2y} \sum_{\substack{\mathcal{H}\in\mathcal{H}_R^{(2)}(N,M)\\ |\mathcal{Y}(\mathcal{H})|=y}} w(\mathcal{H}) \\
&\leqslant \sum_{y \geqslant L_1 q^{-3/4}} \exp\left(-\frac{\delta^3 y}{q^{1/4}}\right) \sum_{M=1}^{2y} \sum_{N=M}^{2y} \exp\left(16\left(N + 2M\log\frac{1}{q}\right)\right) \\
&\leqslant \sum_{y \geqslant L_1 q^{-3/4}} \exp\left(-\frac{\delta^4 y}{q^{1/4}}\right) \leqslant e^{-2/q},
\end{aligned}$$

as required. □

We will next deal with those hierarchies for which $X(\mathcal{H})$ is unusually large. To be precise, let us define $\mathcal{H}_R^{(3)}$ to be the set of $\mathcal{H} \in \mathcal{H}_R^* \setminus \left(\mathcal{H}_R^{(1)} \cup \mathcal{H}_R^{(2)}\right)$ such that

$$X(\mathcal{H}) \geqslant \frac{1}{L_2 q^{3/4}}. \tag{59}$$

For this class of hierarchies we will prove the following bound.

Claim 3: $\displaystyle\sum_{\mathcal{H}\in\mathcal{H}_R^{(3)}} \Lambda(\mathcal{H}) \leqslant \exp\left(-\mathcal{J}(R) - \frac{1}{q^{3/4}}\right) + e^{-2/q}.$

Proof of Claim 3. Let $\mathcal{H} \in \mathcal{H}_R^{(3)}$, and observe that

$$v(\mathcal{H}) \leqslant 8 \cdot s(\mathcal{H}) + \frac{4L_1}{q^{3/4}} \qquad \text{and} \qquad \sum_{u\in\mathrm{up}(\mathcal{H})} \phi(R_u) \leqslant \frac{L_1}{q^{3/2}}\left(\log\frac{1}{q}\right)^2, \tag{60}$$

where the first inequality holds since $\mathcal{H} \notin \mathcal{H}_R^{(2)}$, and the second holds for any $\mathcal{H} \in \mathcal{H}_R$, by (56). We will repeat the proof of Claim 1, using the bounds (60) instead of (44).

Indeed, note (cf. (46)) that

$$\prod_{N_{G_{\mathcal{H}}}(u)=\{v\}} \exp\Big(4\phi(R_u)q\Big) \leqslant C^{v(\mathcal{H})} \exp\left(\frac{4L_1}{\sqrt{q}}\left(\log\frac{1}{q}\right)^2\right), \tag{61}$$

and hence, using (47), we obtain

$$\Lambda(\mathcal{H}) \leqslant w(\mathcal{H}) \cdot C^{10v(\mathcal{H})} \cdot \exp\left(-\frac{X(\mathcal{H})}{5}\log\frac{1}{q} + \frac{4L_1}{\sqrt{q}}\left(\log\frac{1}{q}\right)^2\right).$$

Now, note that, by Observation 5.10 and the bounds (59) and (60), we have

$$X(\mathcal{H}) \geqslant \frac{1}{L_2} \cdot \max\left\{\frac{s(\mathcal{H})}{q^{1/4}}, \frac{1}{q^{3/4}}\right\} \geqslant \frac{v(\mathcal{H})}{L_3},$$

It follows, exactly as in the proof of Claim 1 (cf. the proof of (49)), that

$$\sum_{\mathcal{H}\in\mathcal{H}_R^{(3)}:\, X(\mathcal{H})>1/q} \Lambda(\mathcal{H}) \leqslant e^{-2/q}.$$

We will therefore assume from now on that $X(\mathcal{H}) \leqslant 1/q$.

We now simply repeat the remainder of the proof of Claim 1, using the bounds $1/(L_2q^{3/4}) \leqslant X(\mathcal{H}) \leqslant 1/q$, to obtain

$$\Lambda(\mathcal{H}) \leqslant w(\mathcal{H}) \cdot \exp\left(-\mathcal{J}(R) - \frac{X(\mathcal{H})}{2}\log\frac{X(\mathcal{H})\sqrt{q}}{L_4}\right)$$

for each $\mathcal{H} \in \mathcal{H}_R^{(3)}$, and hence

$$\begin{aligned}\sum_{\substack{\mathcal{H}\in\mathcal{H}_R^{(3)}\\ X(\mathcal{H})\leqslant 1/q}} \Lambda(\mathcal{H}) &\leqslant e^{-\mathcal{J}(R)} \sum_{x=1/(L_2q^{3/4})}^{1/q} \exp\left(-\frac{x}{2}\log\frac{x\sqrt{q}}{L_4}\right) \sum_{\substack{\mathcal{H}\in\mathcal{H}_R^{(3)}\\ X(\mathcal{H})=x}} w(\mathcal{H}) \\ &\leqslant e^{-\mathcal{J}(R)} \sum_{x=1/(L_2q^{3/4})}^{1/q} \exp\left(-\frac{x}{2}\log\frac{x\sqrt{q}}{L_4} + O\Big(L_3\cdot x\Big)\right) \leqslant \exp\left(-\mathcal{J}(R) - \frac{1}{q^{3/4}}\right),\end{aligned}$$

as claimed. □

We are now ready to deal with those hierarchies that travel 'far from the diagonal', i.e., that contain a vertex u satisfying either (30) or (31). We will first consider the (easier) case in which u is in the upper trunk, i.e.,

$$\mathrm{short}(R_u) \geqslant \frac{B}{q} \qquad \text{and} \qquad \mathrm{long}(R_u) \geqslant 4\cdot\mathrm{short}(R_u). \tag{62}$$

Let $\mathcal{H}_R^{(4)}$ be the set of hierarchies $\mathcal{H} \in \mathcal{H}_R^* \setminus \bigcup_{i=1}^3 \mathcal{H}_R^{(i)}$ containing a vertex u such that (62) holds. For these hierarchies we will prove the following bound.

Claim 4: $\displaystyle\sum_{\mathcal{H}\in\mathcal{H}_R^{(4)}} \Lambda(\mathcal{H}) \leqslant \exp\left(-\frac{2\lambda}{q} - \frac{2}{q^{3/4}}\right).$

Proof of Claim 4. Given $\mathcal{H} \in \mathcal{H}_R^{(4)}$ and $u \in V(\mathcal{H})$ satisfying (62), we claim that

$$\Lambda(\mathcal{H}) \leqslant w(\mathcal{H}) \cdot \exp\left(-\mathcal{J}(R_u) + \frac{L_3}{q^{3/4}}\right). \tag{63}$$

To prove this, we will repeat the proof of Claim 1, with some minor changes. First, note that (51) holds, and moreover

$$\frac{s(\mathcal{H})}{L_1 q^{1/4}} \leqslant X(\mathcal{H}) \leqslant \frac{1}{L_2 q^{3/4}}, \tag{64}$$

the first holding by Observation 5.10, and the second since $\mathcal{H} \notin \mathcal{H}_R^{(3)}$. Now, applying Lemma 5.15 to $\mathcal{H}_u$, the sub-hierarchy of $\mathcal{H}$ rooted at u, and using (61) instead of (46), we obtain (cf. the proof of (52))

$$\Lambda(\mathcal{H}) \leqslant w(\mathcal{H}) \cdot C^{10v(\mathcal{H})} \cdot \exp\left(-\frac{U(S,R_u)}{q} - \frac{X(\mathcal{H})}{5}\log\frac{1}{q} + \frac{4L_1}{\sqrt{q}}\left(\log\frac{1}{q}\right)^2\right)$$

for some pod S with $\phi(S) \leqslant X(\mathcal{H}_u) \leqslant X(\mathcal{H})$. Using (60) again (this time to bound $v(\mathcal{H})$ from above), and continuing to follow the proof of Claim 1, we obtain

$$\Lambda(\mathcal{H}) \leqslant w(\mathcal{H}) \cdot \exp\left(-\mathcal{J}(R_u) - \frac{X(\mathcal{H})}{2}\log\frac{X(\mathcal{H})q^{3/5}}{L_4} + \frac{L_2}{q^{3/4}}\right)$$

instead of (54), which implies (63).

Now, let $c = \mathrm{short}(R_u)$, and observe that

$$\mathcal{J}(R_u) \geqslant \frac{2}{q}\int_0^{cq} g(z)\,\mathrm{d}z + 3cg(cq) \geqslant \frac{2}{q}\int_0^{\infty} g(z)\,\mathrm{d}z + 2cg(cq),$$

by (62) and (12), and since $c \geqslant B/q$, and that

$$cg(cq) \geqslant \frac{c}{2}\cdot e^{-2cq} \geqslant \frac{1}{4q^{3/4}}\log\frac{1}{q},$$

where the first inequality holds since $c \geqslant B/q$, and the second since $4c \leqslant \mathrm{long}(R) \leqslant (1/(2q))\log(1/q)$. Combining this with (63), it follows that

$$\Lambda(\mathcal{H}) \leqslant w(\mathcal{H}) \cdot \exp\left(-\frac{2\lambda}{q} - \frac{L_3}{q^{3/4}}\right).$$

Hence, recalling that $v(\mathcal{H}) \leqslant 5L_1/q^{3/4}$ and $s(\mathcal{H}) \leqslant q^{-1/2}$ for every $\mathcal{H} \in \mathcal{H}_R^{(4)}$, by (60) and (64), and applying Lemma 5.9, we obtain

$$\begin{aligned}\sum_{\mathcal{H}\in\mathcal{H}_R^{(4)}} \Lambda(\mathcal{H}) &\leqslant \exp\left(-\frac{2\lambda}{q} - \frac{L_3}{q^{3/4}}\right)\sum_{M=1}^{1/q^{1/2}}\sum_{N=M}^{5L_1/q^{3/4}}\sum_{\mathcal{H}\in\mathcal{H}_R^{(4)}(N,M)} w(\mathcal{H}) \\ &\leqslant \exp\left(-\frac{2\lambda}{q} - \frac{L_3}{q^{3/4}}\right)\sum_{M=1}^{1/q^{1/2}}\sum_{N=M}^{5L_1/q^{3/4}} \exp\left(\frac{L_2}{q^{3/4}}\right) \leqslant \exp\left(-\frac{2\lambda}{q} - \frac{2}{q^{3/4}}\right),\end{aligned}$$

as claimed. □

Finally, we come to most technically challenging family of hierarchies: those which contain a vertex $u \in V(G_{\mathcal{H}})$ such that

$$\text{short}(R_u) \leqslant \frac{B}{q} \qquad \text{and} \qquad \text{long}(R_u) \geqslant 2L_1 \cdot \text{short}(R_u)\,. \tag{65}$$

Let $\mathcal{H}_R^{(5)}$ denote the set of hierarchies $\mathcal{H} \in \mathcal{H}_R^* \setminus \bigcup_{i=1}^4 \mathcal{H}_R^{(i)}$ containing a vertex u such that (65) holds. For this final class of hierarchies we will prove the following bound.

Claim 5: $\displaystyle\sum_{\mathcal{H}\in\mathcal{H}_R^{(5)}} \Lambda(\mathcal{H}) \leqslant \exp\left(-\mathcal{J}(R) - \frac{1}{q^{3/4}}\right).$

Proof of Claim 5. Given a hierarchy $\mathcal{H} \in \mathcal{H}_R^{(5)}$, let $u \in V(\mathcal{H})$ be a vertex satisfying (65) with $\text{long}(R_u)$ maximal, and set $c = \text{short}(R_u)$ and $d = \text{long}(R_u)$. We will prove that

$$\Lambda(\mathcal{H}) \leqslant w(\mathcal{H}) \cdot \exp\left(-\mathcal{J}(R) - 3Cd + X(\mathcal{H}) \log \frac{1}{X(\mathcal{H})q^{3/4}}\right)\,, \tag{66}$$

from which the claim will follow easily, using Lemma 5.9.

In order to prove (66), we will need various bounds on c, d, $h(\mathcal{H})$, $v(\mathcal{H})$ and $X(\mathcal{H})$. Note first that $\mathcal{H}$ does not contain a vertex satisfying (62) since $\mathcal{H} \notin \mathcal{H}_R^{(4)}$. It follows, by Lemmas 5.13 and 5.14, and since $\mathcal{H} \notin \mathcal{H}_R^{(1)}$, that

$$\frac{L_2}{\sqrt{q}} \leqslant h(\mathcal{H}) \leqslant L_1 q^{1/4} d\,. \tag{67}$$

Indeed, the lower bound holds since $\mathcal{H} \notin \mathcal{H}_R^{(1)}$ implies that one of the inequalities in (44) must fail to hold, and by Lemma 5.14, it must be the bound on $h(\mathcal{H})$. The upper bound then follows by Lemma 5.13, and by our choice of u (i.e., with $\text{long}(R_u)$ maximal).

We next claim that

$$v(\mathcal{H}) \leqslant d, \qquad c \leqslant \frac{1}{q} \qquad \text{and} \qquad \frac{1}{q^{3/4}} \leqslant d \leqslant \frac{B}{2q}\,. \tag{68}$$

Indeed, the lower bound on d follows immediately from (67), since $L_1 \leqslant L_2$. To prove the other bounds, recall first that (since $\mathcal{H} \in \mathcal{H}_R^*$) the rectangle R_u does not satisfy (40) or (41). Since $c \leqslant B/q$, by (65), it follows that $d \leqslant 3e^{2B}/q$, and hence

$$c \leqslant \frac{d}{L_1} \leqslant \frac{3e^{2B}}{L_1 q} \leqslant \frac{1}{q}\,.$$

Now, since R_u does not satisfy (41), it follows that $d \leqslant B/(2q)$, as claimed. Finally, to prove the bound on $v(\mathcal{H})$, recall that

$$\frac{m(\mathcal{H})}{3\sqrt{q}} \leqslant X(\mathcal{H}) \leqslant \frac{1}{L_2 q^{3/4}}\,, \tag{69}$$

by the definition of large seeds, and since $\mathcal{H} \notin \mathcal{H}_R^{(3)}$. Hence, by (67) and Observation 5.11, we obtain

$$v(\mathcal{H}) \leqslant 2 \cdot h(\mathcal{H}) \cdot m(\mathcal{H}) \leqslant 6L_1 q^{3/4} d \cdot X(\mathcal{H}) \leqslant d\,,$$

as claimed.

We now apply Lemma 5.16 to obtain two pods $S_1 \subset R_u$ and $R_u \subset S_2 \subset R$, such that

$$\phi(S_1) + \phi(S_2) - \phi(R_u) \leqslant X(\mathcal{H}) \tag{70}$$

and

$$\sum_{N_{G_{\mathcal{H}}}(v)=\{w\}} U(R_w, R_v) \geqslant U(S_1, R_u) + U(S_2, R) - 2s(\mathcal{H})qg(\sqrt{q})\,.$$

Let $S_1 \subset S \subset S_2$ be a rectangle with

$$\dim(S) = \dim(S_1) + \dim(S_2) - \dim(R_u)\,, \tag{71}$$

so $\phi(S) \leqslant X(\mathcal{H})$, by (70), and moreover $U(S_1, R_u) \geqslant U(S, S_2)$, since $g(z)$ is decreasing.

Recalling that $U(R_w, R_v) \leqslant q \cdot Q(R_w, R_v)$, by Lemma 3.8, and that $g(\sqrt{q}) \leqslant \log(1/q)$, by (6), it follows that

$$\sum_{N_{G_{\mathcal{H}}}(v)=\{w\}} Q(R_w, R_v) \geqslant \frac{1}{q}\Big(U(S, S_2) + U(S_2, R)\Big) - 2s(\mathcal{H}) \log \frac{1}{q}\,.$$

Hence, using (47), (61) and (64), we obtain

$$\Lambda(\mathcal{H}) \leqslant w(\mathcal{H}) \cdot C^{10v(\mathcal{H})} \exp\left(-\frac{1}{q}\Big(U(S, S_2) + U(S_2, R)\Big) - \frac{X(\mathcal{H})}{4} \log \frac{1}{q} + \frac{1}{q^{3/4}}\right)\,. \tag{72}$$

It only remains to bound $U(S, S_2)$ and $U(S_2, R)$; controlling $U(S, S_2)$ will take some work, but we obtain a suitable bound on $U(S_2, R)$ simply by applying Lemma 3.5. Indeed, by (68), (69) and (71), we have

$$\mathrm{long}(S_2) \;\leqslant\; \phi(S) + \mathrm{long}(R_u) \;\leqslant\; X(\mathcal{H}) + d \;\leqslant\; \frac{1}{q^{3/4}} + \frac{B}{2q} \;\leqslant\; a \;=\; \mathrm{short}(R)\,, \tag{73}$$

and therefore, setting $s_2 = \mathrm{short}(S_2)$ and $t_2 = \mathrm{long}(S_2)$, we may apply Lemma 3.5, which gives

$$\frac{U(S_2, R)}{q} \;\geqslant\; (t_2 - s_2)g(t_2 q) + \frac{2}{q}\int_{t_2 q}^{aq} g(z)\,\mathrm{d}z + (b-a)g(aq)\,. \tag{74}$$

As noted above, we will have to work harder to obtain a suitable bound on $U(S, S_2)$; in particular, the bound we obtain will depend on whether or not $\mathrm{long}(S) \leqslant \mathrm{short}(S_2)$.

Case 1: $\mathrm{long}(S) \leqslant \mathrm{short}(S_2)$.

This case is also straightforward, since we may apply Lemma 3.7, which gives

$$\frac{U(S, S_2)}{q} \;\geqslant\; \frac{2}{q}\int_0^{s_2 q} g(z)\,\mathrm{d}z + (t_2 - s_2)g(s_2 q) - \frac{X(\mathcal{H})}{2} \log \frac{2}{X(\mathcal{H})q} - O\Big(X(\mathcal{H})\Big)\,, \tag{75}$$

where we used the inequalities $\phi(S) \leqslant X(\mathcal{H}) \leqslant 1/q$. To show that this is sufficient to deduce (66), we will use Lemma 3.10. Indeed, observe that

$$L_1 \cdot \mathrm{short}(S_2) \leqslant L_1\Big(c + X(\mathcal{H})\Big) \leqslant d \leqslant \mathrm{long}(S_2) \leqslant \frac{B}{q}\,,$$

where the first inequality follows from (70) (since $R_u \subset S_2$), the second follows since $2L_1 \cdot c \leqslant d$, by (65), and $L_2 \cdot X(\mathcal{H}) \leqslant q^{-3/4} \leqslant d$, by (68) and (69), the third since $R_u \subset S_2$, and the last by (73). Since $d \leqslant t_2$ it follows, by Lemma 3.10, that

$$\frac{2}{q}\int_{s_2 q}^{t_2 q} g(z)\,\mathrm{d}z \;\leqslant\; (t_2 - s_2)\Big(g(s_2 q) + g(t_2 q)\Big) - 4Cd\,.$$

Combining this with (74) and (75), it follows that

$$\frac{1}{q}\Big(U(S, S_2) + U(S_2, R)\Big) \geqslant \mathcal{J}(R) + 4Cd - \frac{X(\mathcal{H})}{2} \log \frac{2}{X(\mathcal{H})q} - O\Big(X(\mathcal{H})\Big)\,.$$

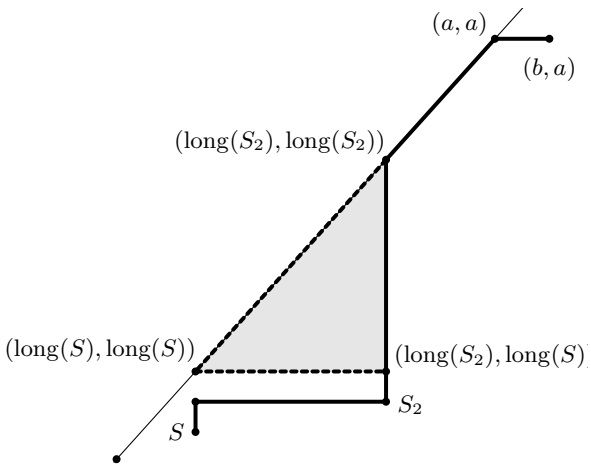


FIGURE 4. The easiest path via which the rectangle S can grow first to S_2, and then to R, in Case 2 of Claim 5. In the proof, the lower two thick segments are replaced with the lower dashed one, and Lemma 3.10 is applied to the shaded triangle.

Hence, by (72), and recalling from (68) that $d \geqslant \max\{v(\mathcal{H}), q^{-3/4}\} \geqslant X(\mathcal{H})$, we obtain

$$\Lambda(\mathcal{H}) \leqslant w(\mathcal{H}) \cdot \exp\left(-\mathcal{J}(R) - 3Cd + \frac{X(\mathcal{H})}{2} \log \frac{1}{X(\mathcal{H})\sqrt{q}}\right),$$

which is slightly stronger than (66).

Case 2: $\mathrm{long}(S) > \mathrm{short}(S_2)$.

This case is a bit more tricky, as the easiest path from S to S_2 does not reach the diagonal, see Figure 4. As a consequence, we cannot apply Lemma 3.7 directly to bound $U(S, S_2)$, nor can we apply Lemma 3.10 directly to the dimensions of S_2. Instead, we observe that, setting $t := \mathrm{long}(S)$, we have

$$\frac{U(S, S_2)}{q} \geqslant (t_2 - t)g(s_2 q) \geqslant (t_2 - t)g(tq),$$

by Lemma 3.6, and since $g(z)$ is decreasing. Combining this with (74), we obtain

$$\frac{1}{q}\Big(U(S, S_2) + U(S_2, R)\Big) \geqslant \mathcal{J}(R) - \frac{2}{q}\int_0^{t_2 q} g(z)\,\mathrm{d}z + (t_2 - s_2)g(t_2 q) + (t_2 - t)g(tq).$$

We will now apply Lemma 3.10 to the pair (t, t_2). Indeed, we have

$$L_1 \cdot \mathrm{long}(S) \leqslant L_1 \cdot X(\mathcal{H}) \leqslant d \leqslant \mathrm{long}(S_2) \leqslant \frac{B}{q},$$

where the first inequality follows since $\phi(S) \leqslant X(\mathcal{H})$, and the others follow as in Case 1. By Lemma 3.10, and since $d \leqslant t_2$, it follows that

$$\frac{2}{q}\int_{tq}^{t_2 q} g(z)\,\mathrm{d}z \leqslant (t_2 - t)\Big(g(tq) + g(t_2 q)\Big) - 4Cd.$$

Note also that, by integrating (6), we have

$$\frac{2}{q}\int_0^{tq} g(z)\,\mathrm{d}z \leqslant t \log \frac{1}{tq} + O(t).$$

Hence, recalling that $s_2 < t \leqslant \phi(S) \leqslant X(\mathcal{H})$, it follows that

$$\frac{1}{q}\Big(U(S, S_2) + U(S_2, R)\Big) \geqslant \mathcal{J}(R) + 4Cd - X(\mathcal{H}) \log \frac{1}{X(\mathcal{H})q} - O\Big(X(\mathcal{H})\Big).$$

Hence, by (72), and since $d \geqslant \max\{v(\mathcal{H}), q^{-3/4}\} \geqslant X(\mathcal{H})$, by (68), we obtain (66).

It now only remains to deduce the claim from (66); we do so using Lemma 5.9. Indeed, recalling that $q^{-1/5}s(\mathcal{H}) \leqslant X(\mathcal{H}) \leqslant 1/(L_2 q^{3/4}) \leqslant \max\{v(\mathcal{H}), q^{-3/4}\} \leqslant d \leqslant B/q$, by (64) and (68), we obtain

$$\begin{aligned}\sum_{\mathcal{H}\in\mathcal{H}_R^{(5)}} \Lambda(\mathcal{H}) &\leqslant e^{-\mathcal{J}(R)} \sum_{x=1}^{1/(L_2 q^{3/4})} \left(\frac{1}{q^{3/4}x}\right)^x \sum_{d=q^{-3/4}}^{B/q} e^{-3Cd} \sum_{M=1}^{q^{1/5}d} \sum_{N=M}^{d} \sum_{\mathcal{H}\in\mathcal{H}_R^{(5)}(N,M)} w(\mathcal{H}) \\ &\leqslant e^{-\mathcal{J}(R)+q^{-3/4}} \sum_{d=q^{-3/4}}^{B/q} e^{-3Cd} \sum_{M=1}^{q^{1/5}d} \sum_{N=M}^{d} \exp\Big(O\Big(N + M\log(1/q)\Big)\Big) \\ &\leqslant e^{-\mathcal{J}(R)+q^{-3/4}} \sum_{d=q^{-3/4}}^{B/q} e^{-2Cd} \leqslant \exp\left(-\mathcal{J}(R) - \frac{1}{q^{3/4}}\right),\end{aligned}$$

as required. This concludes the proof of Claim 5. □

Now, combining Claims 1–5, it follows that

$$\sum_{\mathcal{H}\in\mathcal{H}_R^*} \Lambda(\mathcal{H}) \leqslant 3\cdot\exp\left(-\min\left\{\frac{2\lambda}{q} + \frac{2}{q^{3/4}}, \mathcal{J}(R) - \frac{L_5}{\sqrt{q}}\right\}\right).$$

As was observed before the proof (see the discussion leading up to (43)), this completes the proof of Theorem 6.1. □

We are finally ready to deduce Theorem 1.1 from Theorem 6.1.

Proof of Theorem 1.1. Recall that the upper bound was proved in [18]; we will prove the lower bound. Let $n \in \mathbb{N}$ be sufficiently large and set

$$q := \frac{\lambda}{\log n} - \frac{4e^4 + L_6}{(\log n)^{3/2}},$$

and note that the same is satisfied by $p = q + \Theta(q^2)$ with a slightly smaller constant. Suppose that $[n]^2$ is (internally) filled. Then by the Aizenman–Lebowitz lemma there exists an internally filled rectangle $R \subset [n]^2$ with

$$\frac{1}{4q}\log\frac{1}{q} \leqslant \mathrm{long}(R) \leqslant \frac{1}{2q}\log\frac{1}{q}.$$

There are at most $n^2(\log n)^3$ rectangles satisfying those conditions, and each one satisfies the conditions of Theorem 6.1. Hence, by the union bound and Lemma 3.9 we have

$$\mathbb{P}_p\Big([A] = [n]^2\Big) \leqslant n^2(\log n)^3 \exp\left(-\frac{2\lambda}{q} + \frac{4e^4 + L_6}{\sqrt{q}}\right) \to 0$$

as $n \to \infty$, as required. □

7. Open problems

The most obvious problem suggested by Theorem 1.1 is to determine even more precise bounds on $p_c\Big([n]^2, 2\Big)$. By a theorem of Balogh and Bollobás [3], it is known that the 'critical window' in which the probability of percolation increases from $o(1)$ to $1-o(1)$ has size at most $(\log n)^{-2+o(1)}$, and it is therefore natural to make the following conjecture.

Conjecture 7.1. *There exists a constant $\mu > 0$ such that*

$$p_c\big([n]^2, 2\big) = \frac{\pi^2}{18 \log n} - \frac{\mu + o(1)}{(\log n)^{3/2}}$$

as $n \to \infty$.

Another natural direction for future research would be to extend the results of this paper to higher dimensions. The following conjecture was made by by Uzzell [26], who also established the upper bound.

Conjecture 7.2 (Conjecture 7.1 of [26])**.**

$$p_c\left([n]^d, r\right) = \left(\frac{\lambda(d,r)}{\log_{(r-1)} n} - \frac{\Theta(1)}{\left(\log_{(r-1)} n\right)^{3/2}}\right)^{d-r+1}.$$

As a first step, it would be interesting to determine whether or not this conjecture holds in the case $r = 2$. In particular, one might hope that the conjecture in this case would follow from a suitable generalization of the technique used in this paper.

However, perhaps the most interesting avenue for further research would be to prove corresponding 'sharp' and 'sharper' thresholds for other two-dimensional models, cf. the discussion of $\mathcal{U}$-bootstrap percolation in the Introduction. It would be very interesting (and, most likely, very challenging) to determine a sharp threshold for all families with polylogarithmic critical probability, or a 'sharper' threshold for either some large class of models (e.g., that studied in [13], or a corresponding class of 'unbalanced' models), or for other specific interesting examples, such as the Duarte model (see [8]). The problem in higher dimensions is also extremely interesting, but much more difficult, and proving even much weaker bounds on the critical probability for general $\mathcal{U}$-bootstrap models (see, for example, [24, Conjecture 1.6]) is an important open problem.

Appendix A. Proof of Lemmas 4.3 and 4.4

In this appendix we complete the proofs of Lemmas 4.3 and 4.4, by dealing with the cases that were omitted from the sketch proofs given in Section 4. The details are somewhat tedious but, for the convenience of the diligent (or skeptical) reader, we will attempt to spell everything out slowly and carefully.

For the entire appendix we fix the rectangle $R = [(0,0),(a-1,b-1)]$. Recall from Definition 4.1 that if $S \subset R$ is a rectangle, then we write

$$z(S,R) = |Z(S,R)| = \left|\left\{\mathbf{d} \in \mathcal{I} \,:\, B_{\mathbf{d}}(S,R) \neq \emptyset\right\}\right|,$$

where $\mathcal{I} = \{(-1,0),(1,0),(0,-1),(0,1)\}$ is the set of directions, and $B_{\mathbf{d}}(S,R)$ is the buffer in direction $\mathbf{d}$. Recall also that if $\mathbf{x} \in \{0,1\}^{\mathcal{I}}$ and $S \subset R$ is a rectangle, then

$$x = x'_{(1,0)} + x'_{(-1,0)} \qquad \text{and} \qquad y = x'_{(0,1)} + x'_{(0,-1)}\,, \tag{76}$$

where $\mathbf{x}' := \mathbf{x} \cdot \mathbb{1}_{Z(S,R)}$. For convenience, let us begin by restating the two key lemmas.

Lemma A.1 (Lemma 4.3 restated)**.** *Let $\mathbf{x} \in \{0,1\}^{\mathcal{I}}$ and $S \subset R$ be a rectangle with $\dim S = (a-s, b-t)$ and set $z = z(S,R)$. If*

$$L_1 \leqslant \operatorname{short}(R) \leqslant \frac{B}{q} \qquad \textit{and} \qquad \operatorname{long}(R) \leqslant \frac{3e^{2B}}{q}\,, \tag{77}$$

and $s,t \leqslant 4\delta\sqrt{\mathrm{short}(R)}$*, then*

$$\mathbb{P}_p\Big(D_1^{\mathbf{x}}(S,R)\Big) \leqslant C^z \left(\frac{C}{\sqrt{a}}\right)^y \left(\frac{C}{\sqrt{b}}\right)^x \exp\Big(-sg(bq)-tg(aq)\Big).$$

Lemma A.2 (Lemma 4.4 restated)**.** *Let* $\mathbf{x} \in \{0,1\}^{\mathcal{I}}$ *and* $S \subset R$ *be a rectangle with* $\dim S = (a-s, b-t)$ *and set* $z = z(S,R)$*. If*

$$\mathrm{short}(R) > \frac{B}{q} \qquad \textit{and} \qquad \mathrm{long}(R) \leqslant \frac{1}{2q}\log\frac{1}{q} \tag{78}$$

and $s,t \leqslant \dfrac{4\delta}{\sqrt{q}} \cdot \exp\Big(\mathrm{short}(R)\cdot q\Big)$*, then*

$$\mathbb{P}_p\Big(D_2^{\mathbf{x}}(S,R)\Big) \leqslant \Big(Ce^{\mathrm{short}(R)q}\Big)^z \Big(C\sqrt{q}e^{-aq}\Big)^y \Big(C\sqrt{q}e^{-bq}\Big)^x \exp\Big(-sg(bq)-tg(aq)\Big).$$

We begin with a straightforward technical lemma, which will be required in both proofs.

Lemma A.3. *Let* $15 \leqslant D \leqslant 4\delta/q$*, and suppose that* $a,b \leqslant (1/(2q))\log(1/q)$*, and that* $s,t \leqslant D$ *and* $t \leqslant \min\{a,b\}$*. Then, for any* $x,y \in \{0,1,2\}$*, we have*

$$\exp\Big(-3Dg(tq)\Big) \leqslant \left(\frac{1}{\sqrt{a}}\right)^y \left(\frac{1}{\sqrt{b}}\right)^x \exp\Big(-sg(bq)-tg(aq)\Big)$$

and

$$\exp\Big(-3Dg(tq)\Big) \leqslant \Big(\sqrt{q}e^{-aq}\Big)^y \Big(\sqrt{q}e^{-bq}\Big)^x \exp\Big(-sg(bq)-tg(aq)\Big).$$

Proof. Observe first that $-3g(tq) \leqslant \log(tq) \leqslant \log(Dq)$, by (6) and since $t \leqslant D \leqslant 4\delta/q$. Noting that $D\log(Dq)$ is decreasing in D, it follows that

$$\exp\Big(-Dg(tq)\Big) \leqslant \exp\Big(5\log(15q)\Big) \leqslant q^4 \leqslant \min\left\{\frac{1}{ab},\, q^2e^{-2q(a+b)}\right\},$$

where in the last step we used the bound $a,b \leqslant (1/(2q))\log(1/q)$. Moreover, we have

$$\exp\Big(-2Dg(tq)\Big) \leqslant \exp\Big(-sg(bq)-tg(aq)\Big),$$

since $s,t \leqslant D$ and $t \leqslant \min\{a,b\}$, and recalling that g is decreasing. Combining these two inequalities we obtain the two claimed bounds. □

Proof of Lemma A.1. Recall that $D_1^{\mathbf{x}}(S,R)$ denotes the event that

$$\Big[S \cup \Big(A \cap R \setminus S_{\blacksquare}^{\mathbf{x}}\Big)\Big] = R.$$

Let a,b satisfy (77), and for each x,y,z and each $s,t \leqslant 4\delta\sqrt{\min\{a,b\}}$, set

$$F^{x,y,z}(s,t) := C^z \left(\frac{C}{\sqrt{a}}\right)^y \left(\frac{C}{\sqrt{b}}\right)^x \exp\Big(-sg(bq)-tg(aq)\Big).$$

We will prove, by induction on the pair $(s+t, -(x+y))$, that

$$\mathbb{P}_p\Big(D_1^{\mathbf{x}}(S,R)\Big) \leqslant F^{x,y,z}(s,t) \tag{79}$$

for every $0 \leqslant s,t \leqslant 4\delta\sqrt{\min\{a,b\}}$ and $\mathbf{x} \in \{0,1\}^{\mathcal{I}}$, and every $S \subset R$ with $\dim(S) = (a-s, b-t)$, where x and y are as defined in (76), and $z = z(S,R)$.

The base of the induction, the case $\min\{s,t\} = 0$, was dealt with in Section 4, so let us fix $\mathbf{x} \in \{0,1\}^{\mathcal{I}}$ and $S \subset R$ with $\dim(S) = (a-s, b-t)$, and assume that (79) holds for all smaller values of the pair $(s+t, -(x+y))$ in lexicographical order.

Note first that, since $\mathrm{short}(R) \geqslant L_1$, the function $F^{x,y,z}(s,t)$ is increasing in z and decreasing in x, y, s and t. Note also that we may assume, without loss of generality, that $\mathbf{x} = \mathbf{x} \cdot \mathbb{1}_{Z(S,R)}$ (i.e., that $\mathbf{x_d} = 0$ whenever the buffer $B_{\mathbf{d}}(S,R)$ is empty), since neither side of the inequality (79) depends on the value of $\mathbf{x_d}$ if $\mathbf{d} \notin Z(S,R)$.

We partition into cases, depending on whether or not $z = x + y$.

Case 1: $z = x + y$, i.e., all of the non-empty buffers are included in $S^{\mathbf{x}}_{\square}$.

The key observation in this case is that if the event $D^{\mathbf{x}}_1(S,R)$ holds, then there exists a rectangle $T \subset R$ such that

$$\left[A \cap T \setminus S^{\mathbf{x}}_{\blacksquare}\right] = T \qquad \text{and} \qquad T \cap S^{\mathbf{x}}_{\blacksquare} \neq \emptyset \tag{80}$$

(see Figure 6a). In Section 4 we assumed that $\phi(T) \leqslant 36 f(R) = 36\delta\sqrt{\min\{a,b\}}$, so let us begin by dealing with the other case. To do so, the key observation is that, since T is internally filled by the infected sites in $T \setminus S^{\mathbf{x}}_{\blacksquare}$, one of the eight rectangles in Figure 5 has no double gap crossing it in the 'short' direction. To be more precise, set $D := 4\delta\sqrt{\min\{a,b\}}$ (so, in particular, $s,t \leqslant D$ and $9D \leqslant \min\{a,b\}$) and consider the $(3D+1) \times (\leqslant t)$ rectangle

$$\Big([D, 4D] \times [0, t-1]\Big) \cap \Big(R \setminus S\Big),$$

which is located at the bottom and to the left of Figure 5.

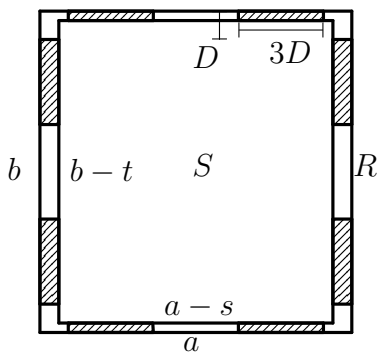


FIGURE 5. The 8 rectangles, one of which must have a 'short' double gap if $\phi(T)$ is large. Each has width at most D and length $3D+1$. Note that $9D \leqslant \mathrm{short}(R)$, so the rectangles do not overlap.

By Lemmas 3.3 and A.3, the probability that this rectangle contains no vertical double gap is at most

$$\exp\Big(-3Dg(tq)\Big) \leqslant \left(\frac{1}{\sqrt{a}}\right)^y \left(\frac{1}{\sqrt{b}}\right)^x \exp\Big(-sg(bq) - tg(aq)\Big) \leqslant \frac{1}{C} \cdot F^{x,y,z}(s,t)\,.$$

Applying the same argument to the other seven rectangles in Figure 5, we may assume that each is either empty, or contains a double gap crossing it in the short direction. But in this case any rectangle satisfying (80) must be contained in a square of size $(4D+1) \times (4D+1)$ in one of the four corners of R, and it is therefore not possible that it has semi-perimeter greater than $9D$, as required.

We will now sum over choices of T with $\phi(T) \leqslant 9D$ the probability that

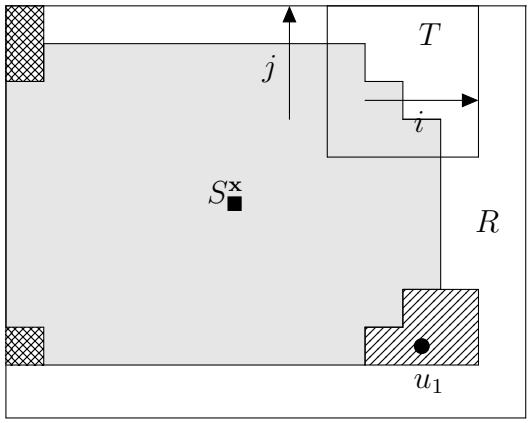
$$\left[A \cap T \setminus S^{\mathbf{x}}_{\blacksquare}\right] = T \qquad \text{and} \qquad \left[S \cup T \cup \Big(A \cap R \setminus S^{\mathbf{x}}_{\blacksquare}\Big)\right] = R\,. \tag{81}$$

Note that these two events depend on disjoint sets of infected sites, and are therefore independent. To bound the probabilities of these events, we will partition according to

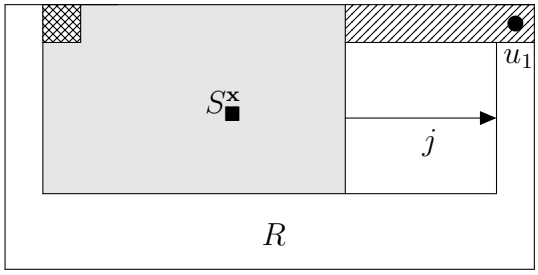
$k := \phi(T)$, and the dimensions of $[S \cup T]$,

$$\dim\Big([S \cup T]\Big) = (a - s + i, b - t + j)\,.$$

It was proved in Section 4 that, given i, j and k, the expected number of rectangles T satisfying the first event in (81) is at most $(24kp)^{k/2}$. Bounding the probability of the second event in (81) is unfortunately rather more complicated, and will require the induction hypothesis, and a careful analysis of the possible positions of elements of A in the buffers of $[S \cup T]$. Recall that the case in which there are no infected sites in the buffers was analyzed in Section 4.



(A) The rectangle T is internally filled outside the shaded $S^{\mathbf{x}}_{\blacksquare}$ and allows S to grow i to the right and j upwards. Since there is an infected site $u_1 \in J$ (in the hatched region), but not in the double hatched set (so $J(u_1)$ is empty), in this case we have $E = \{u_1\}$.



(B) In Algorithm A.5, the rectangle S grows j steps to the right until it reaches either (a) the right-hand side of R, (b) a double gap, or (c) an infected site u_1 in the hatched region (here $x_{(0,1)} = x_{(-1,0)} = 1$ and $x_{(1,0)} = x_{(0,-1)} = 0$). In the figure case (c) occurs, and hence $\ell_{(0,1)} = 1$. The double hatched site in the top-left corner is not occupied, so the set $J_{(-1,0)}$ is empty and hence $\ell_{(-1,0)} = 0$. It follows that $E = \{u_1\}$ in this case.

FIGURE 6. The two possible growth mechanisms.

In order to deal systematically with all of the possible cases, we run the following algorithm. For simplicity we assume that T contains the top right corner of $S^{\mathbf{x}}_{\square}$ (as in Figure 6a); the same bound follows in the other three cases by symmetry.

Algorithm A.4. We define a set $E \subset R \setminus [S \cup T]$ of size 0, 1 or 2, as follows:

1. If the set
$$J := A \cap \Big(B_{(-1,0)}([S \cup T], R) \cup B_{(0,-1)}([S \cup T], R)\Big) \setminus S^{\mathbf{x}}_{\square}$$
is empty, then set $E := \emptyset$.
2. Otherwise, choose $\mathbf{d} \in \{(-1,0),(0,-1)\}$ and $u_1 \in J \cap B_{\mathbf{d}}([S \cup T], R)$.[7] Now, if
$$J(u_1) := A \cap B_{\mathbf{d}'}\Big([S \cup T \cup \{u_1\}], R\Big) \setminus S^{\mathbf{x}}_{\square}$$
is empty, where $\{\mathbf{d}, \mathbf{d}'\} = \{(0,-1),(-1,0)\}$, then set $E := \{u_1\}$.
3. Otherwise, choose $u_2 \in J(u_1)$, and set $E := \{u_1, u_2\}$.

[7]Whenever we have a choice to make in either algorithm, we choose the first direction / site of A in some (arbitrary) pre-defined order on $\mathcal{I}$ / the sites in R.

We now partition the second event in (81) according to the set E, and apply the induction hypothesis to the rectangle

$$\hat{S}(E) := [S \cup T \cup E]\,.$$

Recall that the case $E = \emptyset$ was dealt with in Section 4, so we may assume that $|E| \in \{1, 2\}$. It is possible to deal with these two cases at the same time, but to simplify the notation we will take them one at a time.

Indeed, suppose first that $|E| = 2$. Then, by the induction hypothesis, it follows that[8]

$$\mathbb{P}_p\Big(\big[\hat{S}(E) \cup \big(A \cap R \setminus S^{\mathbf{x}}_{\blacksquare}\big)\big] = R\Big) \leqslant F^{x-2,y-2,z}(s-i-2, t-j-2)\,.$$

Now, recalling that $k = \phi(T)$, note that there are at most $(2k)^2$ choices for the set E. Hence, recalling that the expected number of rectangles T satisfying the first event in (81) is at most $(24kp)^{k/2}$, it will suffice (in this case) to show (cf. (20)) that

$$\sum_{i+j \geqslant 4} \sum_{k=i+j}^{36f(R)} 4k^2p^2 \cdot (24kp)^{k/2} \cdot F^{x-2,y-2,z}(s-i-2, t-j-2) \ll F^{x,y,z}(s,t)\,. \tag{82}$$

To see this, note first that $24kp \leqslant \delta$, since $k \leqslant 36\delta\sqrt{\min\{a,b\}} \leqslant q^{-1/2}$, and that we may therefore assume that $k = i + j$. Now, observe that

$$\frac{F^{x-2,y-2,z}(s-i-2, t-j-2)}{F^{x,y,z}(s,t)} = \frac{ab}{C^4} \cdot e^{(i+2)g(bq)+(j+2)g(aq)} \leqslant \frac{ab}{C^4}\left(\frac{C}{bq}\right)^{(i+2)/2}\left(\frac{C}{aq}\right)^{(j+2)/2}$$

by (7), since $\max\{a,b\} \leqslant 3e^{2B}/q$. Since $i + j = k$, and recalling that $p \leqslant q$ and that $\min\{a,b\} \geqslant \max\{L_1, C^3k\}$, we obtain

$$\sum_{k=4}^{36f(R)} \sum_{i+j=k} 4k^2 \cdot (24kp)^{k/2} \cdot \frac{1}{C^2}\left(\frac{C}{bq}\right)^{i/2}\left(\frac{C}{aq}\right)^{j/2} \leqslant \sum_{k=4}^{36f(R)} \frac{4k^3 \cdot (C^2k)^{k/2}}{\min\{a,b\}^{k/2}} \leqslant \frac{C^5}{\min\{a,b\}^2}\,,$$

which implies (82).

The argument for the case $|E| = 1$ is almost the same. Observe first that, by symmetry, we may assume that $u_1 \in B_{(0,-1)}([S \cup T], R)$, as in Figure 6a. Noting that the set $J(u_1)$ (in Algorithm A.4) is empty, it follows from the induction hypothesis that

$$\mathbb{P}_p\Big(\big[\hat{S}(E) \cup \big(A \cap R \setminus S^{\mathbf{x}}_{\blacksquare}\big)\big] = R\Big) \leqslant F^{x-1,y-2,z}(s-i, t-j-2)\,.$$

There are at most $2k$ choices for the vertex u_1, so it will suffice (in this case) to show that

$$\sum_{i+j \geqslant 4} \sum_{k=i+j}^{36f(R)} 2kp \cdot (24kp)^{k/2} \cdot F^{x-1,y-2,z}(s-i, t-j-2) \ll F^{x,y,z}(s,t)\,. \tag{83}$$

Since $24kp \leqslant \delta$, we may again assume that $k = i + j$. Now, observe that

$$\frac{F^{x-1,y-2,z}(s-i, t-j-2)}{F^{x,y,z}(s,t)} = \frac{a\sqrt{b}}{C^3} \cdot e^{ig(bq)+(j+2)g(aq)} \leqslant \frac{a\sqrt{b}}{C^3}\left(\frac{C}{bq}\right)^{i/2}\left(\frac{C}{aq}\right)^{(j+2)/2}$$

by (7), since $\max\{a,b\} \leqslant 3e^{2B}/q$. Since $i \geqslant 1$, we obtain

$$\sum_{k=4}^{36f(R)} \sum_{i+j=k} 2k \cdot (24kp)^{k/2} \cdot \frac{\sqrt{b}}{C^2}\left(\frac{C}{bq}\right)^{i/2}\left(\frac{C}{aq}\right)^{j/2} \leqslant \sum_{k=4}^{36f(R)} \frac{2k^2 \cdot (C^2k)^{k/2}}{\min\{a,b\}^{(k-1)/2}} \leqslant \frac{C^5}{\min\{a,b\}^{3/2}}\,,$$

[8]Note that we used here the bound $z(\hat{S}, R) \leqslant z(S, R)$, and the monotonicity of F in z, s, t.

which implies (83), since $\min\{a,b\} \geqslant \max\{L_1, C^3k\}$. This completes the proof in Case 1.

Case 2: $z > x + y$, i.e., some non-empty buffer is not included in $S^{\mathbf{x}}_{\square}$.

Without loss of generality, let $B_{(1,0)}(S,R)$ be a non-empty buffer that is not included in $S^{\mathbf{x}}_{\square}$, so $x_{(1,0)} = 0$. As explained in the sketch proof, the idea is to 'grow' S to the right until we find a double gap, an infected site in one of the buffers, or reach the right-hand side of R, thus leading either to an increase in $x + y$, or a decrease in $s + t$. In Section 4 we dealt with the cases in which we find a double gap before reaching the right-hand side, and that before doing so we do not find any infected sites in the buffers above or below S. Here we will deal with the other cases.

Algorithm A.5. We define a set $E \subset R \setminus S^{\mathbf{x}}_{\square}$ of size 0, 1, 2 or 3, and also an integer $j \in \{0, \dots, s\}$ and a variables $\ell_{\mathbf{d}} \in \{0,1\}$ for each direction $\mathbf{d} \in \mathcal{I}$, as follows:[9]

1. Set $\tilde{S} := \bigcup_{i=0}^{j} \big(S + (i,0)\big)$ and $\tilde{\mathbf{x}} := \mathbf{x} + \mathbb{1}_{(1,0)}$, where
$$j := \min\Big\{ i \geqslant 0 \,:\, A \cap R \cap \big(S + (i+2,0)\big) \setminus \big(S + (i,0)\big) = \emptyset \Big\}.$$
If the set $A \cap \tilde{S}^{\tilde{\mathbf{x}}}_{\square} \setminus S^{\mathbf{x}}_{\square}$ is empty, then go to Step 8.
2. Set $\tilde{S} := \bigcup_{i=0}^{j} \big(S + (i,0)\big)$, where j is minimal such that $\tilde{S} \setminus S$ is crossed from left to right, and
$$\big(A \cap \tilde{S}^{\tilde{\mathbf{x}}}_{\square}\big) \setminus \big(S^{\mathbf{x}}_{\square} \cup B_{(1,0)}(\tilde{S},R)\big) \neq \emptyset\,.$$
3. Now, if $x_{(0,1)} = 1$ and the set[10]
$$J_{(0,1)} := A \cap \Big(\big(B_{(0,1)}(\tilde{S},R) \setminus B_{(0,1)}(S,R)\big) \cup \{v_{(1,1)}\} \Big)$$
is non-empty, then add a site $u_{(0,1)} \in J_{(0,1)}$ to E, and set $\ell_{(0,1)} := 1$.
4. Similarly, if $x_{(0,-1)} = 1$ and the set
$$J_{(0,-1)} := A \cap \Big(\big(B_{(0,-1)}([\tilde{S} \cup E],R) \setminus B_{(0,-1)}(S,R)\big) \cup \{v_{(1,-1)}\} \Big)$$
is non-empty, then add a site $u_{(0,-1)} \in J_{(0,-1)}$ to E, and set $\ell_{(0,-1)} := 1$.
5. If $x_{(0,1)} = 1$ and $E = \{v_{(1,-1)}\}$, and the set[11]
$$J'_{(0,1)} := A \cap \Big(\big(B_{(0,1)}([\tilde{S} \cup E],R) \setminus B_{(0,1)}(S,R)\big) \Big)$$
is non-empty, then add a site $u_{(0,1)} \in J'_{(0,1)}$ to E, and set $\ell_{(0,1)} := 1$.
6. If $x_{(-1,0)} = 0$ then go to Step 8. Otherwise, if the set
$$J_{(-1,0)} := A \cap B_{(-1,0)}([\tilde{S} \cup E],R) \setminus S^{\mathbf{x}}_{\square}$$
is non-empty, then add a site $u_{(-1,0)} \in J_{(-1,0)}$ to E, and set $\ell_{(-1,0)} := 1$.

[9] Initially $E = \emptyset$ and $\ell_{\mathbf{d}} = 0$ for every direction $\mathbf{d} \in \mathcal{I}$.

[10] For each $i \in \{1,-1\}$, let $v_{(1,i)}$ denote the unique 'corner' site in $R \setminus \tilde{S}$ with a neighbour in each of the buffers $B_{(1,0)}(\tilde{S},R)$ and $B_{(0,i)}(\tilde{S},R)$.

[11] Note that this set has at most one element.

7. If either $\ell_{(0,1)}+\ell_{(0,-1)} = x_{(0,1)}+x_{(0,-1)}$ or $\ell_{(-1,0)} = 0$ then go to Step 8. Otherwise:
   (a) If $\ell_{(0,1)} = 1$ and $\ell_{(0,-1)} = 0$, and the set
   $$J_{(-1,-1)} := A \cap B_{(0,-1)}([\tilde{S} \cup E], R) \setminus S^{\mathbf{x}}_{\square}$$
   is non-empty, then add a site $u_{(-1,-1)} \in J_{(-1,-1)}$ to E and set $\ell_{(0,-1)} := 1$.
   (b) If $\ell_{(0,1)} = 0$ and $\ell_{(0,-1)} = 1$, and the set
   $$J_{(-1,1)} := A \cap B_{(0,1)}([\tilde{S} \cup E], R) \setminus S^{\mathbf{x}}_{\square}$$
   is non-empty, then add a site $u_{(-1,1)} \in J_{(-1,1)}$ to E and set $\ell_{(0,1)} := 1$.
8. Set $\hat{S} := [\tilde{S} \cup E]$ and STOP.

Observe that Algorithm A.5 outputs a set E and an integer j (which together determine the sets $\tilde{S}$ and $\hat{S}$, and the variable $\ell_{\mathbf{d}}$ for each $\mathbf{d} \in \mathcal{I}$) with the following properties:

(a) $E \subset A$;

(b) $[S \cup (A \cap \tilde{S})] = \tilde{S}$;

(c) On the event $D_1^{\mathbf{x}}(S, R)$ the event $D_1^{\hat{\mathbf{x}}}(\hat{S}, R)$ occurs, i.e.,
$$\Big[\hat{S} \cup \Big(A \cap R \setminus \hat{S}^{\hat{\mathbf{x}}}_{\blacksquare}\Big)\Big] = R\,,$$
where $\hat{\mathbf{x}} = \mathbf{x} - \boldsymbol{\ell}$, except in the case treated in Section 4, in which $\hat{\mathbf{x}} = \mathbf{x} + \mathbb{1}_{(1,0)}$.

We will analyse each case individually, and sum over all possible sets E and $j \in \{0, \dots, s\}$.

Suppose first that $E = \emptyset$. In Section 4 we dealt with this case, under the additional assumption that $B_{(1,0)}(\hat{S}, R) \neq \emptyset$ (i.e., we found a double gap before reaching the right-hand side of R); here we will deal with the other case (i.e., that $\hat{S}$ reaches the right-hand side of R). To do so, we need to bound the probability that
$$[S \cup (A \cap \tilde{S})] = \tilde{S} \qquad \text{and} \qquad D_1^{\hat{\mathbf{x}}}(\hat{S}, R)\,,$$
where in this case $\hat{S} = \tilde{S}$ and $\hat{\mathbf{x}} = \mathbf{x}$. Note that these two events depend on disjoint sets of sites, and are therefore independent; we will bound the first using Lemma 3.3, and the second using the induction hypothesis. Indeed, by Lemma 3.3 (and since $g(z)$ is decreasing) we have
$$\mathbb{P}_p\Big([S \cup (A \cap \tilde{S})] = \tilde{S}\Big) \leqslant \exp\Big(-jg(bq)\Big)\,,$$
and by the induction hypothesis (since $B_{(1,0)}(\hat{S}, R) = \emptyset$ and $E = \emptyset$),
$$\mathbb{P}_p\Big(D_1^{\mathbf{x}}(\hat{S}, R)\Big) \leqslant F^{x,y,z-1}(s-j, t)\,.$$
Thus, the probability that $E = \emptyset$, $B_{(1,0)}(\hat{S}, R) = \emptyset$ and $D_1^{\mathbf{x}}(S, R)$ is at most[12]
$$\exp\Big(-jg(bq)\Big)F^{x,y,z-1}(s-j, t) = \frac{1}{C} \cdot F^{x,y,z}(s, t)$$
which suffices since C is sufficiently large.

We will therefore assume from now on that $E \neq \emptyset$, which means that before reaching a double gap (when trying to grow S rightwards, forming $\tilde{S}$), we find an infected site in

[12]Note that there is a unique possible value of j for which $B_{(1,0)}(\hat{S}, R) = \emptyset$.

either the buffer above or below $\tilde{S}$. More precisely, recall that $\tilde{S} := \bigcup_{i=0}^{j} \big(S + (i,0)\big)$, where j is minimal such that $\tilde{S} \setminus S$ is crossed from left to right, and

$$\big(A \cap \tilde{S}_{\square}^{\tilde{\mathbf{x}}}\big) \setminus \big(S_{\square}^{\mathbf{x}} \cup B_{(1,0)}(\tilde{S}, R)\big) \neq \emptyset\,,$$

where $\tilde{\mathbf{x}} = \mathbf{x} + \mathbb{1}_{(1,0)}$. Note that $x_{(0,1)} + x_{(0,-1)} \geqslant 1$, and that moreover $\ell_{(0,1)} + \ell_{(0,-1)} \geqslant 1$ (since the set above is non-empty).

Suppose first that $|E| = 1$, and therefore (without loss of generality) we have $\ell_{(0,1)} = 1$, $E = \{u_{(0,1)}\}$ and $\ell_{(0,-1)} = \ell_{(-1,0)} = 0$. This implies that the (independent) events

$$u_{(0,1)} \in A, \qquad [S \cup (A \cap \tilde{S})] = \tilde{S} \qquad \text{and} \qquad D_1^{\hat{\mathbf{x}}}(\hat{S}, R)$$

occur, where $\hat{\mathbf{x}} = \mathbf{x} - \mathbb{1}_{(0,1)}$. Given j, it follows from the minimality of j that there are at most four choices for $u_{(0,1)}$, so

$$\mathbb{P}_p\Big(\Big\{u_{(0,1)} \in A\Big\} \cap \Big\{[S \cup (A \cap \tilde{S})] = \tilde{S}\Big\}\Big) \leqslant 4p \exp\big(-jg(bq)\big)\,.$$

Suppose first that $u_{(0,1)} \neq v_{(1,1)}$. In this case, by the induction hypothesis, we have

$$\mathbb{P}_p\Big(D_1^{\hat{\mathbf{x}}}(\hat{S}, R)\Big) \leqslant F^{x,y-1,z}(s-j, t-2)\,,$$

and so the probability in this case can be bounded by

$$\sum_{j=0}^{s} 4p \exp\big(-jg(bq)\big) F^{x,y-1,z}(s-j, t-2) \leqslant \frac{4(s+1)p\sqrt{a}}{C} e^{2g(aq)} \cdot F^{x,y,z}(s,t)\,.$$

Recalling that $s \leqslant 4\delta\sqrt{a}$ and that $e^{2g(aq)} \leqslant C/aq$, by (7), and recalling that δ is sufficiently small, we obtain a suitably strong bound in this case. Similarly, if $u_{(0,1)} = v_{(1,1)}$, then the induction hypothesis gives

$$\mathbb{P}_p\Big(D_1^{\hat{\mathbf{x}}}(\hat{S}, R)\Big) \leqslant F^{x,y-1,z}(s-j-1, t-1)\,,$$

and so the probability in this case can be bounded by

$$\sum_{j=0}^{s} 4p \exp\big(-jg(bq)\big) F^{x,y-1,z}(s-j-1, t-1) \leqslant \frac{4(s+1)p\sqrt{a}}{C} e^{g(aq)+g(bq)} \cdot F^{x,y,z}(s,t)\,.$$

Since $s \leqslant 4\delta\sqrt{b}$ and $e^{g(aq)+g(bq)} \leqslant C/(q\sqrt{ab})$, we again obtain a suitable bound.

In all remaining cases we win easily, since each extra infected site is extremely expensive. Nevertheless, we will go carefully through each case. Indeed, suppose next that $|E| = 2$, so (without loss of generality) either

(*a*) $\ell_{(0,1)} = \ell_{(0,-1)} = 1$, $\ell_{(-1,0)} = 0$, and $E = \{u_{(0,1)}, u_{(0,-1)}\}$, or

(*b*) $\ell_{(0,1)} = \ell_{(-1,0)} = 1$, $\ell_{(0,-1)} = 0$, and $E = \{u_{(0,1)}, u_{(-1,0)}\}$.

In case (*a*), the (independent) events

$$E \subset A, \qquad [S \cup (A \cap \tilde{S})] = \tilde{S} \qquad \text{and} \qquad D_1^{\hat{\mathbf{x}}}(\hat{S}, R)$$

occur, where $\hat{\mathbf{x}} = \mathbf{x} - \mathbb{1}_{(0,1)} - \mathbb{1}_{(0,-1)}$. Given j, there are at most five choices for each element of E, so

$$\mathbb{P}_p\Big(E \subset A \text{ and } [S \cup (A \cap \tilde{S})] = \tilde{S}\Big) \leqslant (5p)^2 \exp\big(-jg(bq)\big)\,.$$

Suppose first that the set $E \cap \{v_{(1,1)}, v_{(1,-1)}\}$ is empty. In this case, by the induction hypothesis, we have

$$\mathbb{P}_p\Big(D_1^{\hat{\mathbf{x}}}(\hat{S}, R)\Big) \leqslant F^{x,y-2,z}(s-j, t-4)\,,$$

and so the probability in this case can be bounded by

$$\sum_{j=0}^{s}(5p)^2 e^{-jg(bq)} \cdot F^{x,y-2,z}(s-j, t-4) \leqslant \frac{25(s+1)p^2 a}{C^2} e^{4g(aq)} \cdot F^{x,y,z}(s,t)\,.$$

Since $s \leqslant 4\delta\sqrt{a}$ and $e^{2g(aq)} \leqslant C/aq$, we win (easily) in this case. Similarly, if the set $E \cap \{v_{(1,1)}, v_{(1,-1)}\}$ is non-empty, then the induction hypothesis gives

$$\mathbb{P}_p\Big(D_1^{\hat{\mathbf{x}}}(\hat{S}, R)\Big) \leqslant F^{x,y-2,z}(s-j-1, t-3)\,,$$

and so the probability in this case can be bounded by

$$\sum_{j=0}^{s}(5p)^2 e^{-jg(bq)} \cdot F^{x,y-2,z}(s-j-1, t-3) \leqslant \frac{25(s+1)p^2 a}{C^2} e^{3g(aq)+g(bq)} \cdot F^{x,y,z}(s,t)\,.$$

Since $s \leqslant 4\delta\sqrt{b}$ and $e^{3g(aq)+g(bq)} \leqslant C^2/(q^2\sqrt{a^3 b})$, we again obtain a suitable bound.

In case (b), the (independent) events

$$E \subset A, \qquad [S \cup (A \cap \tilde{S})] = \tilde{S} \qquad \text{and} \qquad D_1^{\hat{\mathbf{x}}}(\hat{S}, R)$$

occur, where $\hat{\mathbf{x}} = \mathbf{x} - \mathbb{1}_{(0,1)} - \mathbb{1}_{(-1,0)}$. Given j, there are at most four choices for each element of E, so

$$\mathbb{P}_p\Big(E \subset A \text{ and } [S \cup (A \cap \tilde{S})] = \tilde{S}\Big) \leqslant (4p)^2 \exp\Big(-jg(bq)\Big)\,.$$

Suppose first that $u_{(0,1)} \neq v_{(1,1)}$. In this case, by the induction hypothesis, we have

$$\mathbb{P}_p\Big(D_1^{\hat{\mathbf{x}}}(\hat{S}, R)\Big) \leqslant F^{x-1,y-1,z}(s-j-2, t-2)\,,$$

and so the probability in this case can be bounded by

$$\sum_{j=0}^{s}(4p)^2 e^{-jg(bq)} \cdot F^{x-1,y-1,z}(s-j-2, t-2) \leqslant \frac{16(s+1)p^2\sqrt{ab}}{C^2} e^{2g(aq)+2g(bq)} \cdot F^{x,y,z}(s,t)\,.$$

Since $s \leqslant 4\delta\sqrt{a}$ and $e^{2g(aq)+2g(bq)} \leqslant C^2/(abq^2)$, we are done in this case, as before. Similarly, if $u_{(0,1)} = v_{(1,1)}$, then the induction hypothesis gives

$$\mathbb{P}_p\Big(D_1^{\hat{\mathbf{x}}}(\hat{S}, R)\Big) \leqslant F^{x-1,y-1,z}(s-j-3, t-1)\,,$$

and so the probability in this case can be bounded by

$$\sum_{j=0}^{s}(4p)^2 e^{-jg(bq)} \cdot F^{x-1,y-1,z}(s-j-3, t-1) \leqslant \frac{16(s+1)p^2\sqrt{ab}}{C^2} e^{g(aq)+3g(bq)} \cdot F^{x,y,z}(s,t)\,.$$

Since $s \leqslant 4\delta\sqrt{b}$ and $e^{g(aq)+3g(bq)} \leqslant C^2/(q^2\sqrt{ab^3})$, we again obtain a suitable bound.

Finally, suppose that $|E| = 3$, and observe that $\ell_{(0,1)} = \ell_{(0,-1)} = \ell_{(-1,0)} = 1$ and, without loss of generality, either

(a) $E = \{u_{(0,1)}, u_{(0,-1)}, u_{(-1,0)}\}$, or

(b) $E = \{u_{(0,1)}, u_{(-1,0)}, u_{(-1,-1)}\}$.

In either case, the (independent) events

$$E \subset A, \qquad [S \cup (A \cap \tilde{S})] = \tilde{S} \qquad \text{and} \qquad D_1^{\hat{\mathbf{x}}}(\hat{S}, R)$$

occur, where $\hat{\mathbf{x}} = \mathbf{x} - \mathbb{1}_{(0,1)} - \mathbb{1}_{(0,-1)} - \mathbb{1}_{(-1,0)} = \mathbf{0}$. Given j, there are at most six choices for each element of E, so

$$\mathbb{P}_p\Big(E \subset A \text{ and } [S \cup (A \cap \tilde{S})] = \tilde{S}\Big) \leqslant (6p)^3 \exp\Big(-jg(bq)\Big).$$

Suppose first that the set $E \cap \{v_{(1,1)}, v_{(1,-1)}\}$ is empty. In this case, by the induction hypothesis, we have

$$\mathbb{P}_p\Big(D_1^{\hat{\mathbf{x}}}(\hat{S}, R)\Big) \leqslant F^{x-1,y-2,z}(s-j-2, t-4),$$

and so the probability in this case can be bounded by

$$\sum_{j=0}^{s-1} (6p)^3 e^{-jg(bq)} \cdot F^{x-1,y-2,z}(s-j-2, t-4) \leqslant \frac{6^3 s p^3 a\sqrt{b}}{C^3} e^{4g(aq)+2g(bq)} \cdot F^{x,y,z}(s,t).$$

Since $s \leqslant 4\delta\sqrt{a}$ and $e^{4g(aq)+2g(bq)} \leqslant C^3/(q^3a^2b)$, we again win easily in this case. Similarly, if $E \cap \{v_{(1,1)}, v_{(1,-1)}\}$ is non-empty, then the induction hypothesis gives

$$\mathbb{P}_p\Big(D_1^{\hat{\mathbf{x}}}(\hat{S}, R)\Big) \leqslant F^{x-1,y-2,z}(s-j-3, t-3),$$

and so the probability in this case can be bounded by

$$\sum_{j=0}^{s-1} (6p)^3 e^{-jg(bq)} \cdot F^{x-1,y-2,z}(s-j-3, t-3) \leqslant \frac{6^3 s p^3 a\sqrt{b}}{C^3} e^{3g(aq)+3g(bq)} \cdot F^{x,y,z}(s,t).$$

Since $s \leqslant 4\delta\sqrt{b}$ and $e^{3g(aq)+3g(bq)} \leqslant C^3/(q^3\sqrt{a^3b^3})$, we again obtain a suitable bound.

Summing over the various cases completes the proof of Lemma A.1. □

The proof of Lemma A.2 is very similar to that of Lemma A.1, and we will be able to reuse large parts of the proof (in particular Algorithms A.4 and A.5).

Proof of Lemma A.2. Recall that $D_2^{\mathbf{x}}(S, R)$ denotes the event that

$$\Big[S \cup \Big(A \cap R \setminus S_{\blacksquare}^{\mathbf{x}}\Big)\Big] = R \qquad \text{and} \qquad A \cap S_{\square}^{\mathbf{x}} = \emptyset.$$

As in the proof of Lemma A.1, we use induction on the pair $(s+t, -(x+y))$, this time to prove that

$$\mathbb{P}_p\Big(D_2^{\mathbf{x}}(S, R)\Big) \leqslant \hat{F}^{x,y,z}(s,t),$$

where

$$\hat{F}^{x,y,z}(s,t) := \Big(Ce^{\mathrm{short}(R)q}\Big)^z \Big(C\sqrt{q}e^{-aq}\Big)^y \Big(C\sqrt{q}e^{-bq}\Big)^x \exp\Big(-sg(bq) - tg(aq)\Big),$$

for every $0 \leqslant s, t \leqslant 4\delta \cdot q^{-1/2} \cdot \exp\Big(\min\{a,b\} \cdot q\Big)$ and $\mathbf{x} \in \{0,1\}^{\mathcal{I}}$, and every $S \subset R$ with $\dim(S) = (a-s, b-t)$, where x and y are as defined in (76), and $z = z(S,R)$. The base of the induction remains unchanged from Lemma A.1.

As in the proof of Lemma A.1, we partition into cases depending on whether or not $z = x+y$, the function $\hat{F}^{x,y,z}(s,t)$ is increasing in z and decreasing in x, y, s and t, and we may assume without loss of generality that $\mathbf{x} = \mathbf{x} \cdot \mathbb{1}_{Z(S,R)}$.

Case 1: $z = x + y$, i.e., all of the non-empty buffers are included in $S^{\mathbf{x}}_{\square}$.

As in Lemma A.1, the event $D^{\mathbf{x}}_2(S, R)$ requires the existence of a rectangle T such that

$$\Big[A \cap T \setminus S^{\mathbf{x}}_{\blacksquare}\Big] = T \qquad \text{and} \qquad T \cap S^{\mathbf{x}}_{\blacksquare} \neq \emptyset\,.$$

The first step is to apply Lemma A.3, as in the proof of Lemma A.1, to exclude rectangles T with $\phi(T) > 9D$, where this time we set $D := 4\delta \cdot q^{-1/2} \cdot \exp\Big(\min\{a, b\} \cdot q\Big)$. It follows from (78) that $s, t \leqslant D \leqslant 4\delta/q$ and $9D \leqslant \min\{a, b\}$, and we may therefore argue exactly as before, except using the second inequality in Lemma A.3, which gives

$$\exp\Big(-3Dg(tq)\Big) \leqslant \Big(\sqrt{q}e^{-aq}\Big)^y \Big(\sqrt{q}e^{-bq}\Big)^x \exp\Big(-sg(bq) - tg(aq)\Big) \leqslant \frac{1}{C} \cdot \hat{F}^{x,y,z}(s, t)\,.$$

We will therefore assume from now on that $\phi(T) \leqslant 9D$, and sum over choices of T with $\phi(T) \leqslant 9D$ the probability that

$$\Big[A \cap T \setminus S^{\mathbf{x}}_{\blacksquare}\Big] = T, \quad \Big[S \cup T \cup \Big(A \cap R \setminus S^{\mathbf{x}}_{\blacksquare}\Big)\Big] = R \quad \text{and} \quad A \cap S^{\mathbf{x}}_{\square} = \emptyset\,. \tag{84}$$

Note that these events depend on disjoint sets of sites and are therefore independent. It was proved in Section 4 that, given $k := \phi(T)$ and the dimensions of $[S \cup T]$, the expected number of rectangles T satisfying the first event in (84) is at most $(24kp)^{k/2}$. For the intersection of the second and third events, we will partition the space according to the set E given by Algorithm A.4, and apply the induction hypothesis to the set

$$\hat{S}(E) := [S \cup T \cup E]\,.$$

Suppose first that $E = \emptyset$, and recall that this means that the set

$$\Big(B_{(-1,0)}([S \cup T], R) \cup B_{(0,-1)}([S \cup T], R)\Big) \setminus S^{\mathbf{x}}_{\square}$$

contains no element of A. Together with the second and third events in (84), this implies that the event $D^{\hat{\mathbf{x}}}_2(\hat{S}(E), R)$ occurs, where $\hat{\mathbf{x}} = \mathbf{x} - \mathbb{1}_{(1,0)} - \mathbb{1}_{(0,1)}$. By the induction hypothesis, we have

$$\mathbb{P}_p\Big(D^{\hat{\mathbf{x}}}_2(\hat{S}(E), R)\Big) \leqslant \hat{F}^{x-1,y-1,z}(s - i, t - j)\,,$$

where $\dim([S \cup T]) = (a - s + i, b - t + j)$. Now, since $B/q \leqslant a, b \leqslant (1/(2q)) \log(1/q)$, we have

$$\frac{\hat{F}^{x-1,y-1,z}(s - i, t - j)}{\hat{F}^{x,y,z}(s, t)} = \frac{e^{(a+b)q}}{C^2 q} \exp\Big(ig(bq) + jg(aq)\Big) \leqslant \frac{2^{i+j}}{C^2 q^2}$$

since $e^{g(aq)+g(bq)} \leqslant 2$. Recalling that $k \leqslant 9D \leqslant 36\delta/q$, it follows that the probability of this case is at most

$$\sum_{i+j \geqslant 4} \sum_{k=i+j}^{9D} (24kp)^{k/2} \cdot \frac{2^{i+j}}{C^2 q^2} \cdot \hat{F}^{x,y,z}(s, t) \leqslant \frac{1}{C} \cdot \hat{F}^{x,y,z}(s, t)\,,$$

as required.

Suppose next that $|E| = 1$, and observe that, by symmetry, we may assume that $u_1 \in B_{(0,-1)}([S \cup T], R)$, as in Figure 6a. Recalling that the set $J(u_1)$ (in Algorithm A.4) is empty, it follows that the event $D^{\hat{\mathbf{x}}}_2(\hat{S}(E), R)$ occurs, where $\hat{\mathbf{x}} = \mathbf{x} - \mathbb{1}_{(1,0)} - \mathbb{1}_{(0,1)} - \mathbb{1}_{(0,-1)}$. By the induction hypothesis, we have

$$\mathbb{P}_p\Big(D^{\hat{\mathbf{x}}}_2(\hat{S}(E), R)\Big) \leqslant \hat{F}^{x-1,y-2,z}(s - i, t - j - 2)$$

and, since $B/q \leqslant a, b \leqslant (1/(2q))\log(1/q)$, we have (as before)

$$\frac{\hat{F}^{x-1,y-2,z}(s-i,t-j-2)}{\hat{F}^{x,y,z}(s,t)} = \frac{e^{(2a+b)q}}{C^3 q^{3/2}} \exp\Big(ig(bq) + (j+2)g(aq)\Big) \leqslant \frac{2^{i+j+2}}{C^3 q^3}\,.$$

Noting that there are at most $2k$ choices for the vertex u_1, and recalling that $k \leqslant 9D \leqslant 36\delta/q$, it follows that the probability of this case is at most

$$\sum_{i+j\geqslant 4} \sum_{k=i+j}^{9D} 2kp \cdot (24kp)^{k/2} \cdot \frac{2^{i+j+2}}{C^3 q^3} \cdot \hat{F}^{x,y,z}(s,t) \leqslant \frac{1}{C^2} \cdot \hat{F}^{x,y,z}(s,t)\,,$$

as required.

Finally, suppose that $|E| = 2$, and observe that in this case the event $D_2^{\hat{\mathbf{x}}}(\hat{S}(E), R)$ occurs, where $\hat{\mathbf{x}} = \mathbf{0}$, and that $\mathbf{x} = \mathbf{1}$. By the induction hypothesis, we have

$$\mathbb{P}_p\Big(D_2^{\hat{\mathbf{x}}}(\hat{S}(E), R)\Big) \leqslant \hat{F}^{x-2,y-2,z}(s-i-2, t-j-2)$$

and, since $B/q \leqslant a, b \leqslant (1/(2q))\log(1/q)$, we have (as before)

$$\frac{\hat{F}^{x-2,y-2,z}(s-i-2,t-j-2)}{\hat{F}^{x,y,z}(s,t)} = \frac{e^{(2a+2b)q}}{C^4 q^2} \exp\Big((i+2)g(bq) + (j+2)g(aq)\Big) \leqslant \frac{2^{i+j+4}}{C^4 q^4}\,.$$

Noting that there are at most $4k^2$ choices for E, and since $k \leqslant 9D \leqslant 36\delta/q$, it follows that the probability of this case is at most

$$\sum_{i+j\geqslant 4} \sum_{k=i+j}^{9D} (2kp)^2 \cdot (24kp)^{k/2} \cdot \frac{2^{i+j+4}}{C^4 q^4} \cdot \hat{F}^{x,y,z}(s,t) \leqslant \frac{1}{C^3} \cdot \hat{F}^{x,y,z}(s,t)\,,$$

as required. This completes the proof in Case 1.

Case 2: $z > x + y$.

As in the proof of Lemma A.1, let $B_{(1,0)}(S,R)$ be a non-empty buffer that is not included in $S_\square^{\mathbf{x}}$, so $x_{(1,0)} = 0$, and define a set E using Algorithm A.5.

Suppose first that $E = \emptyset$, and recall that $\tilde{S} = \bigcup_{i=0}^{j} \Big(S + (i,0)\Big)$, where

$$j = \min\Big\{i \geqslant 0 \,:\, A \cap R \cap \Big(S + (i+2,0)\Big) \setminus \Big(S + (i,0)\Big) = \emptyset\Big\},$$

and that $\tilde{S}_\square^{\tilde{\mathbf{x}}} \setminus S_\square^{\mathbf{x}}$ contains no elements of A, where $\tilde{\mathbf{x}} = \mathbf{x} + \mathbb{1}_{(1,0)}$. There are two sub-cases, depending on whether or not $B_{(1,0)}(\tilde{S}, R) = \emptyset$, that is, whether or not we reached the right-hand side without finding a double gap. Suppose first that we did find a double gap (i.e., $B_{(1,0)}(\tilde{S}, R) \neq \emptyset$). We will sum over choices of j the probability that

$$[S \cup (A \cap \tilde{S})] = \tilde{S}, \qquad \Big[\tilde{S} \cup \Big(A \cap R \setminus \tilde{S}_\blacksquare^{\tilde{\mathbf{x}}}\Big)\Big] = R \qquad \text{and} \qquad A \cap \tilde{S}_\square^{\tilde{\mathbf{x}}} = \emptyset\,. \tag{85}$$

Note that these three events depend on disjoint sets of sites, and are therefore independent; we will bound the first using Lemma 3.3, and the intersection of the second and third using the induction hypothesis. Indeed, by Lemma 3.3 (and since $g(z)$ is decreasing) we have

$$\mathbb{P}_p\Big([S \cup (A \cap \tilde{S})] = \tilde{S}\Big) \leqslant \exp\Big(-jg(bq)\Big)\,.$$

Moreover, the second and third events imply that the event $D_2^{\tilde{\mathbf{x}}}(\tilde{S}, R)$ occurs, and by the induction hypothesis we have

$$\mathbb{P}_p\Big(D_2^{\tilde{\mathbf{x}}}(\tilde{S}, R)\Big) \leqslant \hat{F}^{x+1,y,z}(s-j,t)\,.$$

It follows that the probability that there exists $j \geqslant 0$ such that the events in (85) all hold is at most

$$\sum_{j=0}^{s-1} e^{-jg(bq)} \cdot \hat{F}^{x+1,y,z}(s-j,t) = Cs\sqrt{q}e^{-bq} \cdot \hat{F}^{x,y,z}(s,t) \leqslant 4C\delta \cdot \hat{F}^{x,y,z}(s,t)$$

as required, since $s \leqslant 4\delta \cdot q^{-1/2} \cdot \exp\Big(\min\{a,b\} \cdot q\Big)$ and $\delta = \delta(C) > 0$ is sufficiently small.

We next deal with the case $E = \emptyset$ and $B_{(1,0)}(\tilde{S}, R) = \emptyset$ (i.e., we reached the right-hand side without finding a double gap). In this case, the event $D_2^{\mathbf{x}}(\tilde{S}, R)$ occurs, and by the induction hypothesis we have

$$\mathbb{P}_p\Big(D_2^{\mathbf{x}}(\tilde{S}, R)\Big) \leqslant \hat{F}^{x,y,z-1}(s-j,t)\,,$$

since $z(\tilde{S}, R) = z(S, R) - 1$. It follows that the probability that the events in (85) all hold, and $B_{(1,0)}(\tilde{S}, R) = \emptyset$, is at most

$$\exp\Big(-jg(bq)\Big)\hat{F}^{x,y,z-1}(s-j,t) \leqslant \frac{1}{C} \cdot \hat{F}^{x,y,z}(s,t)$$

which suffices since C is sufficiently large.

We will therefore assume from now on that $E \neq \emptyset$, so $\tilde{S} = \bigcup_{i=0}^{j}\Big(S + (i,0)\Big)$, where j is minimal such that $\tilde{S} \setminus S$ is crossed from left to right, and

$$\Big(A \cap \tilde{S}_{\square}^{\tilde{\mathbf{x}}}\Big) \setminus \Big(S_{\square}^{\mathbf{x}} \cup B_{(1,0)}(\tilde{S}, R)\Big) \neq \emptyset\,.$$

In other words, before reaching a double gap we found an infected site in either the buffer above or below $\tilde{S}$. In particular, note that $\ell_{(0,1)} + \ell_{(0,-1)} \geqslant 1$.

Suppose first that $|E| = 1$, and therefore that (without loss of generality) we have $\ell_{(0,1)} = 1$, $E = \{u_{(0,1)}\}$ and $\ell_{(0,-1)} = \ell_{(-1,0)} = 0$. Then the events

$$u_{(0,1)} \in A, \quad [S \cup (A \cap \tilde{S})] = \tilde{S}, \quad A \cap B_{(0,1)}(S, R) = \emptyset \quad \text{and} \quad D_2^{\hat{\mathbf{x}}}(\hat{S}, R)$$

occur, where $\hat{S} = [\tilde{S} \cup E]$ and $\hat{\mathbf{x}} = \mathbf{x} - \mathbb{1}_{(0,1)}$. There is an important subtlety in this case, since these events might not be independent: the buffer $B_{(0,1)}(S, R)$ might 'stick out' of the top of $\hat{S}$, and therefore intersect the set of sites that the event $D_2^{\hat{\mathbf{x}}}(\hat{S}, R)$ depends on. However, the only dependence is between the decreasing event $\{A \cap B_{(0,1)}(S, R) = \emptyset\}$ and the increasing part of the event $D_2^{\hat{\mathbf{x}}}(\hat{S}, R)$ (since $\hat{x}_{(0,1)} = 0$), so by Harris' inequality [20][13] the probability that all four events occur is at most the product of their probabilities.

Given j, there are at most four choices for $u_{(0,1)}$, so

$$\mathbb{P}_p\Big(\Big\{u_{(0,1)} \in A\Big\} \cap \Big\{[S \cup (A \cap \tilde{S})] = \tilde{S}\Big\}\Big) \leqslant 4p\exp\Big(-jg(bq)\Big)\,.$$

Suppose first that $z(\hat{S}, R) < z(S, R)$. In this case, by the induction hypothesis, we have

$$\mathbb{P}_p\Big(\Big\{A \cap B_{(0,1)}(S, R) = \emptyset\Big\} \cap D_2^{\hat{\mathbf{x}}}(\hat{S}, R)\Big) \leqslant (1-p)^{a-s} \cdot \hat{F}^{x,y-1,z-1}(s-j-1, t-2)\,,$$

[13] Harris' inequality states that increasing events in a product space are positively correlated. It is often referred to as the FKG inequality, which is a generalization that was proved somewhat later.

and so the probability in this case can be bounded by[14]

$$\sum_{j=0}^{s} 4p(1-p)^{a-s}e^{-jg(bq)} \cdot \hat{F}^{x,y-1,z-1}(s-j-1,t-2)$$
$$\leqslant 4(s+1)p \cdot e^{-(a-s)q} \cdot \frac{e^{aq-\min\{a,b\}q}}{C^2\sqrt{q}} \cdot e^{2g(aq)+g(bq)} \cdot \hat{F}^{x,y,z}(s,t) \leqslant \frac{\delta}{C} \cdot \hat{F}^{x,y,z}(s,t)\,,$$

since $s \leqslant 4\delta \cdot q^{-1/2} \cdot \exp\Big(\min\{a,b\} \cdot q\Big) \leqslant 4\delta/q$, and hence $e^{sq+2g(aq)+g(bq)} \leqslant 2$.

On the other hand, if $z(\hat{S},R) = z(S,R)$ then the buffer $B_{(0,1)}(S,R)$ must have height at least two, and hence

$$\mathbb{P}_p\Big(\Big\{A \cap B_{(0,1)}(S,R) = \emptyset\Big\} \cap D_2^{\hat{\mathbf{x}}}(\hat{S},R)\Big) \leqslant (1-p)^{2(a-s)} \cdot \hat{F}^{x,y-1,z}(s-j-1,t-2)\,.$$

The probability in this case can therefore be bounded, as above, by

$$\sum_{j=0}^{s} 4p(1-p)^{2(a-s)}e^{-jg(bq)} \cdot \hat{F}^{x,y-1,z}(s-j-1,t-2)$$
$$\leqslant 4(s+1)p \cdot e^{-2(a-s)q} \cdot \frac{e^{aq}}{C\sqrt{q}} \cdot e^{2g(aq)+g(bq)} \cdot \hat{F}^{x,y,z}(s,t) \leqslant \delta \cdot \hat{F}^{x,y,z}(s,t)\,,$$

as required.

The remaining cases are similar but easier, since each extra infected site is extremely expensive. We will therefore be able to be use slightly weaker bounds, which simplifies the analysis somewhat. Suppose next that $|E| = 2$, so either

(*a*) $\ell_{(0,1)} = \ell_{(0,-1)} = 1$, $\ell_{(-1,0)} = 0$, and $E = \{u_{(0,1)}, u_{(0,-1)}\}$, or

(*b*) $\ell_{(0,1)} = \ell_{(-1,0)} = 1$, $\ell_{(0,-1)} = 0$, and $E = \{u_{(0,1)}, u_{(-1,0)}\}$.

In either case, given j there are at most five choices for each element of E, so

$$\mathbb{P}_p\Big(E \subset A \text{ and } [S \cup (A \cap \tilde{S})] = \tilde{S}\Big) \leqslant (5p)^2 \exp\Big(-jg(bq)\Big)\,.$$

Moreover, in case (*a*), by the induction hypothesis and Harris' inequality, we have

$$\mathbb{P}_p\Big(\Big\{A \cap B = \emptyset\Big\} \cap D_2^{\hat{\mathbf{x}}}(\hat{S},R)\Big) \leqslant (1-p)^{2(a-s)} \cdot \hat{F}^{x,y-2,z}(s-j-1,t-4)\,,$$

where $B = B_{(0,1)}(S,R) \cup B_{(0,-1)}(S,R)$ and $\hat{\mathbf{x}} = \mathbf{x} - \mathbb{1}_{(0,1)} - \mathbb{1}_{(0,-1)}$.[15] The probability in this case can therefore be bounded by

$$\sum_{j=0}^{s} (5p)^2(1-p)^{2(a-s)}e^{-jg(bq)} \cdot \hat{F}^{x,y-2,z}(s-j-1,t-4) \leqslant \frac{\delta}{C} \cdot \hat{F}^{x,y,z}(s,t)\,,$$

since $s \leqslant 4\delta/q$ and $e^{2sq+4g(aq)+g(bq)} \leqslant 2$. Similarly, in case (*b*) we have

$$\mathbb{P}_p\Big(\Big\{A \cap B = \emptyset\Big\} \cap D_2^{\hat{\mathbf{x}}}(\hat{S},R)\Big) \leqslant (1-p)^{a-s+b-t} \cdot \hat{F}^{x-1,y-1,z}(s-j-3,t-2)\,,$$

[14] We remark that this is the only point in the proof where we will need the term $e^{\mathrm{short}(R)qz}$ in the bound in Lemma 4.4. This term gives rise to the term $\mathrm{up}(\mathcal{H})$ in the proof of Theorem 2.1 and the corresponding precision needed in Lemma 5.14.

[15] Note that here (and also below) we could have gained substantially by dividing into cases, as above, depending on whether or not $z(\hat{S},R) < z(S,R)$. In this case, however, this weaker bound will suffice.

where $B = B_{(0,1)}(S,R) \cup B_{(-1,0)}(S,R)$ and $\hat{\mathbf{x}} = \mathbf{x} - \mathbb{1}_{(0,1)} - \mathbb{1}_{(-1,0)}$, which allows us to bound the probability in this case by

$$\sum_{j=0}^{s} (5p)^2 (1-p)^{a-s+b-t} e^{-jg(bq)} \cdot \hat{F}^{x-1,y-1,z}(s-j-3, t-2) \leqslant \frac{\delta}{C} \cdot \hat{F}^{x,y,z}(s,t)\,,$$

exactly as before. The calculation when $|E| = 3$ is almost the same. Recall that $\ell_{(0,1)} = \ell_{(0,-1)} = \ell_{(-1,0)} = 1$ and, without loss of generality, either

(a) $E = \{u_{(0,1)}, u_{(0,-1)}, u_{(-1,0)}\}$, or

(b) $E = \{u_{(0,1)}, u_{(-1,0)}, u_{(-1,-1)}\}$.

In either case, given j there are at most six choices for each element of E, so

$$\mathbb{P}_p\Big(E \subset A \text{ and } [S \cup (A \cap \tilde{S})] = \tilde{S}\Big) \leqslant (6p)^3 \exp\Big(-jg(bq)\Big)\,.$$

Moreover, in either case, by the induction hypothesis and Harris' inequality, we have

$$\mathbb{P}_p\Big(\Big\{A \cap B = \emptyset\Big\} \cap D_2^{\hat{\mathbf{x}}}(\hat{S}, R)\Big) \leqslant (1-p)^{2(a-s)+b-t} \cdot \hat{F}^{x-1,y-2,z}(s-j-3, t-4)\,,$$

where $B = B_{(0,1)}(S,R) \cup B_{(0,-1)}(S,R) \cup B_{(-1,0)}(S,R)$ and $\hat{\mathbf{x}} = \mathbf{0}$. The probability in this case can therefore be bounded by

$$\sum_{j=1}^{s} (6p)^3 (1-p)^{2(a-s)+b-t} e^{-jg(bq)} \cdot \hat{F}^{x-1,y-2,z}(s-j-3, t-4) \leqslant \frac{\delta}{C^2} \cdot \hat{F}^{x,y,z}(s,t)\,,$$

since $s \leqslant 4\delta/q$ and $e^{2sq+tq+4g(aq)+3g(bq)} \leqslant 2$. This completes the proof of Lemma A.2. □

Acknowledgements

The authors would like to thank Béla Bollobás for introducing them to the problem (several years apart), and for many interesting and inspiring conversations on the topic of bootstrap percolation. The second author would also like to thank Ander Holroyd for a useful conversation about the ideas behind the proof of our main theorem.

This research was partly carried out while the authors were visiting IMT Lucca, and partly while the first author was visiting the University of Cambridge, and we are grateful to both institutions for their hospitality, and for providing a wonderful working environment. Our main theorem was first announced by the second author in 2009, and he would like to thank Murray Edwards College, Cambridge, the Japan Society for the Promotion of Science, and Keio University, which supported his research and hosted him during the early stages of this project.

References

1. M. Aizenman and J. L. Lebowitz, *Metastability effects in bootstrap percolation*, J. Phys. A **21** (1988), no. 19, 3801–3813. MR 968311
2. P. Balister, B. Bollobás, M. Przykucki, and P. Smith, *Subcritical $\mathcal{U}$-bootstrap percolation models have non-trivial phase transitions*, Trans. Amer. Math. Soc. **368** (2016), no. 10, 7385–7411. MR 3471095
3. J. Balogh and B. Bollobás, *Sharp thresholds in bootstrap percolation*, Physica A: Statistical Mechanics and its Applications **326** (2003), no. 3-4, 305–312.
4. J. Balogh, B. Bollobás, H. Duminil-Copin, and R. Morris, *The sharp threshold for bootstrap percolation in all dimensions*, Trans. Amer. Math. Soc. **364** (2012), no. 5, 2667–2701. MR 2888224
5. J. Balogh, B. Bollobás, and R. Morris, *Bootstrap percolation in three dimensions*, Ann. Probab. **37** (2009), no. 4, 1329–1380. MR 2546747
6. B. Bollobás, *The art of Mathematics: Coffee time in Memphis*, Cambridge University Press, 2006.

7. B. Bollobás, H. Duminil-Copin, R. Morris, and P. Smith, *Universality of two-dimensional critical cellular automata*, Proc. Lond. Math. Soc., To appear.
8. ———, *The sharp threshold for the Duarte model*, Ann. Probab. **45** (2017), no. 6B, 4222–4272. MR 3737910
9. B. Bollobás, P. Smith, and A. Uzzell, *Monotone cellular automata in a random environment*, Combin. Probab. Comput. **24** (2015), no. 4, 687–722. MR 3350030
10. R. Cerf and E. N. M. Cirillo, *Finite size scaling in three-dimensional bootstrap percolation*, Ann. Probab. **27** (1999), no. 4, 1837–1850. MR 1742890
11. R. Cerf and F. Manzo, *The threshold regime of finite volume bootstrap percolation*, Stochastic Process. Appl. **101** (2002), no. 1, 69–82. MR 1921442
12. J. Chalupa, P. L. Leath, and G. R. Reich, *Bootstrap percolation on a Bethe lattice*, J. Phys. C **12** (1979), no. 1, L31–L35.
13. H. Duminil-Copin and A. Holroyd, *Finite volume bootstrap percolation with balanced threshold rules on* $\mathbb{Z}^2$, (2012), Preprint available at http://www.ihes.fr/~duminil/.
14. H. Duminil-Copin and A. C. D. Van Enter, *Sharp metastability threshold for an anisotropic bootstrap percolation model*, Ann. Probab. **41** (2013), no. 3A, 1218–1242. MR 3098677
15. H. Duminil-Copin, A. C. D. van Enter, and T. Hulshof, *Higher order corrections for anisotropic bootstrap percolation*, Probability Theory and Related Fields (2017).
16. A. Fey, L. Levine, and Y. Peres, *Growth rates and explosions in sandpiles*, J. Stat. Phys. **138** (2010), no. 1-3, 143–159. MR 2594895
17. L. R. Fontes, R. H. Schonmann, and V. Sidoravicius, *Stretched exponential fixation in stochastic Ising models at zero temperature*, Comm. Math. Phys. **228** (2002), no. 3, 495–518. MR 1918786
18. J. Gravner and A. E. Holroyd, *Slow convergence in bootstrap percolation*, Ann. Appl. Probab. **18** (2008), no. 3, 909–928. MR 2418233
19. J. Gravner, A. E. Holroyd, and R. Morris, *A sharper threshold for bootstrap percolation in two dimensions*, Probab. Theory Related Fields **153** (2012), no. 1-2, 1–23. MR 2925568
20. T. E. Harris, *A lower bound for the critical probability in a certain percolation process*, Proc. Cambridge Philos. Soc. **56** (1960), 13–20. MR 0115221
21. A. E. Holroyd, *Sharp metastability threshold for two-dimensional bootstrap percolation*, Probab. Theory Related Fields **125** (2003), no. 2, 195–224. MR 1961342
22. F. Martinelli, R. Morris, and C. Toninelli, *Universality results for kinetically constrained spin models in two dimensions*, ArXiv e-prints (2018).
23. R. Morris, *Zero-temperature Glauber dynamics on* $\mathbb{Z}^d$, Probab. Theory Related Fields **149** (2011), no. 3-4, 417–434. MR 2776621
24. ———, *Bootstrap percolation, and other automata*, European J. Combin. **66** (2017), 250–263. MR 3692148
25. D. Reimer, *Proof of the van den Berg-Kesten conjecture*, Combin. Probab. Comput. **9** (2000), no. 1, 27–32. MR 1751301
26. A. Uzzell, *An improved upper bound for bootstrap percolation in all dimensions*, ArXiv e-prints (2012).
27. J. van den Berg and U. Fiebig, *On a combinatorial conjecture concerning disjoint occurrences of events*, Ann. Probab. **15** (1987), no. 1, 354–374. MR 877608
28. J. van den Berg and H. Kesten, *Inequalities with applications to percolation and reliability*, J. Appl. Probab. **22** (1985), no. 3, 556–569. MR 799280

Département de Mathématiques et Applications, École Normale Supérieure, CNRS, PSL Research University, Sorbonne Universités, 45 rue d'Ulm, Paris, France
E-mail address: ivailo.hartarsky@ens.fr

IMPA, Estrada Dona Castorina 110, Jardim Botânico, Rio de Janeiro, 22460-320, Brazil
E-mail address: rob@impa.br