Sharp asymptotics for finite point-to-plane connections in supercritical bond percolation in dimension at least three

Alexander Fribergh and Alan Hammond

August 29, 2024

Abstract

We consider supercritical bond percolation in $\mathbb{Z}^d$ for $d \geq 3$. The origin lies in a finite open cluster with positive probability, and, when it does, the diameter of this cluster has an exponentially decaying tail. For each unit vector $\boldsymbol{\ell}$, we prove sharp asymptotics for the probability that this cluster contains a vertex $x \in \mathbb{Z}^d$ that satisfies $x \cdot \boldsymbol{\ell} \geq u$. For an axially aligned $\boldsymbol{\ell}$, we find this probability to be of the form $\kappa \exp\{-\zeta u\}(1 + \mathrm{err})$ for $u \in \mathbb{N}$, where $|\mathrm{err}|$ is at most $C \exp\{-cu^{1/2}\}$; for general $\boldsymbol{\ell}$, the form of the asymptotic depends on whether $\boldsymbol{\ell}$ satisfies a natural lattice condition. To obtain these results, we prove that renewal points in long clusters are abundant, with a renewal block length whose tail is shown to decay as fast as $C \exp\{-cu^{1/2}\}$.

Contents

1   Introduction

We investigate the geometry and probability of finite connections in supercritical bond percolation on $\mathbb{Z}^d$ for dimension d at least three. We will present strongly quantified conclusions regarding the presence of renewal points in finite connections. These lead to precise estimates in Theorems 1.4 and 1.8 for the probability of long point-to-hyperplane connections running in a given direction $\boldsymbol{\ell}$.

1.1   The first main result, on the asymptotic probability of long finite clusters

Write $S^{d-1} = \left\{ u \in \mathbb{R}^d : ||u||_2 = 1 \right\}$ for the Euclidean unit ball.

Definition 1.1. Let $\boldsymbol{\ell} \in S^{d-1}$. This vector is called non-lattice if the $\mathbb{Z}$-linear span of its Cartesian coordinate projections $\boldsymbol{\ell} \cdot e_i$, $i \in [\![1,d]\!]$, is dense in $\mathbb{R}$. In the opposing case, where $\boldsymbol{\ell}$ is lattice, we define the span of ℓ to be the greatest positive s for which $\left| \boldsymbol{\ell} \cdot e_i \right| \in s \cdot \mathbb{N}$ for $i \in [\![1,d]\!]$. (For definiteness, $0 \in \mathbb{N}$.)

When $\boldsymbol{\ell}$ is lattice, the cover $\mathrm{Cover}(\boldsymbol{\ell})$ of $\boldsymbol{\ell}$ is $s \cdot \mathbb{N}$, with s the span of $\boldsymbol{\ell}$. When $\boldsymbol{\ell}$ is non-lattice, $\mathrm{Cover}(\boldsymbol{\ell})$ is simply $(0, \infty)$.

Definition 1.2. Let $f, g : \mathbb{N} \to (0, \infty)$. We will use the convention that $f \overset{\boldsymbol{\ell}}{\sim} g$ means that $f(u)/g(u) \to 1$ as $u \to \infty$ along the cover of $\boldsymbol{\ell}$. We will say that f and g are *weakly asymptotic* in this case. (When $\boldsymbol{\ell}$ is non-lattice, this notion is simply the usual asymptotic $\sim$; we will write $f \overset{\boldsymbol{\ell}}{\sim} g$ only in the lattice case.) Suppose that the pair of sequences (f, g) meets the stronger condition that there exist constants C and c in $(0, \infty)$ such that $\left| f(u)/g(u) - 1 \right| \leq C \exp\{-cu^{1/2}\}$ for u in the cover of $\boldsymbol{\ell}$. Then we say that f and g are *strongly* asymptotic, denoting this relation by $f \overset{\boldsymbol{\ell}}{\approx} g$.

Let $d \geq 2$. By $(\mathbb{Z}^d, \sim)$, we denote the lattice $\mathbb{Z}^d$ with the nearest-neighbour notion of adjacency. Writing $p_c = p_c(d)$ for the critical value for bond percolation on $\mathbb{Z}^d$, we let $p \in (p_c, 1)$ be a given supercritical value. Let $\mathbb{P}$ denote the law of bond percolation on $(\mathbb{Z}^d, \sim)$ with parameter p.

The height of $x \in \mathbb{Z}^d$ equals $x \cdot \boldsymbol{\ell}$.

Definition 1.3. A cluster is a finite open connected component under $\mathbb{P}$. The base of this cluster is a vertex therein attaining the minimum height. (If there is more than one minimizer, one is selected as the base in some definite way, as we will specify in Section 1.3.) When it exists, the cluster with base at $x \in \mathbb{Z}^d$ will be denoted $\mathcal{C}(x)$.

For $x \in \mathbb{Z}^d$ and $u \in (0, \infty)$, let $\mathsf{Cluster}_u(x)$ denote the event that x is the base of a cluster for which the maximum height attained among its vertices equals u.

The connection formed when $\mathsf{Cluster}_u(x)$ occurs is *finite*, because $|\mathcal{C}(x)| < \infty$, and of *point-to-plane* type, in the sense that an open path travels from x to the hyperplane of height u.

Here is one of the principal conclusions of this article.

Theorem 1.4. *Let $d \geq 3$, $p \in (p_c, 1)$ and $\boldsymbol{\ell} \in S^{d-1}$. Then there exist positive and finite constants $\kappa_{\mathrm{clust}} = \kappa_{\mathrm{clust}}(d, p, \boldsymbol{\ell})$ and $\zeta = \zeta(p, d, \boldsymbol{\ell})$ such that, if $\boldsymbol{\ell}$ is lattice, then*

$$\mathbb{P}\big(\mathsf{Cluster}_u\big) \overset{\boldsymbol{\ell}}{\approx} \kappa_{\mathrm{clust}} \exp\big\{ -\zeta u \big\};$$

while, if $\boldsymbol{\ell}$ is non-lattice, then

$$\mathbb{P}\big(\mathsf{Cluster}_{[u,\infty)}\big) \sim \kappa_{\mathrm{clust}} \exp\big\{-\zeta u\big\},$$

where $\mathsf{Cluster}_{[u,\infty)} = \cup_{u' \geq u}\mathsf{Cluster}_{u'}$.

The value of $\zeta(p, d, \boldsymbol{\ell})$ will be furnished by Proposition 1.7.

Fundamental to the derivation of Theorem 1.4 is the notation of a renewal point and a cousin of the cluster called a string. The introduction continues with suitable notation; Theorem 1.8, which is a string counterpart of Theorem 1.4; and Theorem 1.9, on the abundance of renewal points.

1.2 The basics

Let $x \cdot y$ denote the Cartesian scalar product of $x, y \in \mathbb{R}^d$.

We regard $\boldsymbol{\ell}$ as pointing vertically in the positive direction; so $x \in \mathbb{R}^d$ is upwards from $z \in \mathbb{R}^d$ if $x \cdot \boldsymbol{\ell} > z \cdot \boldsymbol{\ell}$.

For $h \in \mathbb{R}$, the *forward* and *backward* half-spaces of h are given by

$$F_h = \big\{u \in \mathbb{R}^d : u \cdot \boldsymbol{\ell} \geq h\big\} \text{ and } B_h = \big\{u \in \mathbb{R}^d : u \cdot \boldsymbol{\ell} < h\big\}.$$

For $e = (x_1, x_2) \in E(\mathbb{Z}^d)$, write $(e) = \big\{\lambda x_1 + (1 - \lambda)x_2 : \lambda \in (0, 1)\big\} \subset \mathbb{R}^d$ for the interior of the endpoint-interpolating interval. For $h \in \mathbb{R}$, we define the upper and lower half-spaces U_h and L_h of h to be subgraphs of $(\mathbb{Z}^d, \sim)$ whose edge sets are given by

$$E(U_h) = \big\{e \in E(\mathbb{Z}^d) : (e) \cap F_h \neq \emptyset\big\} \text{ and } E(L_h) = \big\{e \in E(\mathbb{Z}^d) : (e) \subset B_h\big\},$$

and whose vertex sets $V(U_h)$ and $V(L_h)$ are the sets of endpoints of elements of the respective edge sets.
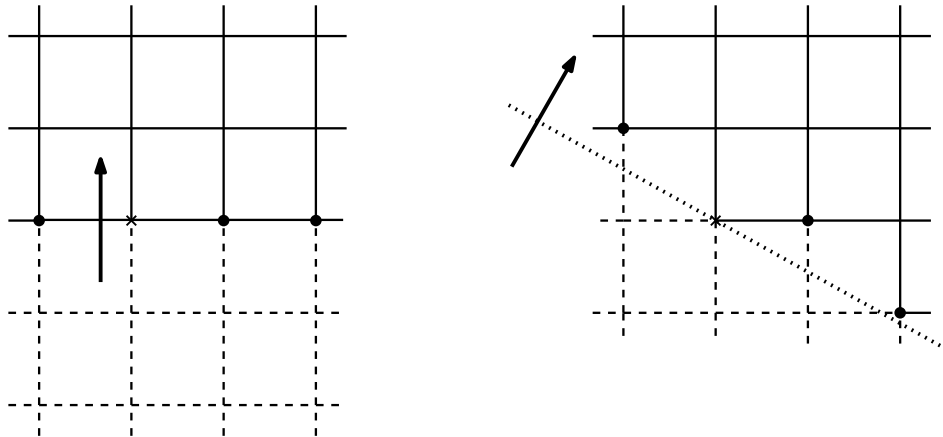
Figure 1: The partition $E(L_h) \cup E(L_h)$ of edges in $\mathbb{Z}^2$ for an axial and a more general choice of $\boldsymbol{\ell}$. The arrow represents $\boldsymbol{\ell}$, while the cross marks the origin, which is an element of $\mathrm{VertBdry}_h$ for $h = 0$. The discs mark the other elements of $\mathrm{VertBdry}_0$. Elements of $E(U_0)$ are drawn with solid lines; those of $E(L_0)$ are dashed.

The *vertex boundary* $\mathrm{VertBdry}_h$ of h equals the intersection of vertex sets $V(U_h) \cap V(L_h)$. The various characters appear in Figure 1.

For $m, n \in \mathbb{Z}$ with $m \leq n$, we denote by $[\![m, n]\!]$ the integer interval $\{k \in \mathbb{Z} : m \leq k \leq n\}$.

Let $G_i = \big(V(G_i), E(G_i)\big)$, $i \in [\![1, 2]\!]$, be two subgraphs of $(\mathbb{Z}^d, \sim)$. Their intersection $G_1 \cap G_2$ is the subgraph $\big(V(G_1) \cap V(G_2), E(G_1) \cap E(G_2)\big)$. For $y, z \in \mathbb{R}$ with $z \geq y$, we define the slab $\mathrm{Slab}_{y,z}$ from y to z to be the subgraph $U_y \cap L_z$.

An element $v \in V(\mathrm{Slab}_{y,z})$ is called lower if it belongs to $\mathrm{VertBdry}_y$, and upper if it belongs to $\mathrm{VertBdry}_z$.

1.3 Notation, strings, and renewal

1.3.1 The base and the head

For certain choices of $\boldsymbol{\ell}$, the hyperplane $H_s = \{z \in \mathbb{R}^d : z \cdot \boldsymbol{\ell} = s\}$, with $s \in \mathbb{R}$, may intersect a finite set $A \subset \mathbb{Z}^d$ at more than one point. To distinguish a point in the intersection, we consider the projection $H_s \to H_0$ orthogonal to $\boldsymbol{\ell}$. We fix an orthonormal basis for H_0. This basis induces a lexicographical ordering on H_0 and this ordering gives rise to one on H_s when we identify H_0 and H_s by translation by any given element in H_s. The latter ordering is a total order on H_s.

Let G be a finite subgraph of $(\mathbb{Z}^d, \sim)$. The base of G, denoted by $\mathrm{base}(G)$, is the lexicographically minimal element of $V(G)$ among those of minimum height. For the head $\mathrm{head}(G)$, we replace 'minimal' by 'maximal' and 'minimum' by 'maximum'. The usage of the term 'height' is extended so that the height of G is given by $\max\{(v - u) \cdot \boldsymbol{\ell} : u, v \in V(G)\}$.

1.3.2 Renewal

Next is the important definition of a renewal point.

Definition 1.5. Let U be a finite connected subgraph of $(\mathbb{Z}^d, \sim)$ and let $x \in V(U)$. The element x is called a *renewal point* of U if $V(U) \cap \mathrm{VertBdry}_{x \cdot \boldsymbol{\ell}}$ equals the singleton set $\{x\}$.

1.3.3 The string

Let $x, y \in \mathbb{Z}^d$, $x \cdot \boldsymbol{\ell} < y \cdot \boldsymbol{\ell}$. A string from x to y is a finite open connected component in $\mathrm{Slab}_{x \cdot \boldsymbol{\ell}, y \cdot \boldsymbol{\ell}}$ such that x, y are renewal points of S. A string crosses the slab $\mathrm{Slab}_{x \cdot \boldsymbol{\ell}, y \cdot \boldsymbol{\ell}}$ between locations x and y at which it has renewal points. The term 'string' is shorthand for 'string of sausages': as we will see, a string can be broken into sausages delimited by consecutive renewal points: see Figure 2.

For $u > 0$ and $x \in \mathbb{Z}^d$, let $\mathsf{Str}_u(x)$ denote the event there exists a string from x to z for some $z \in \mathbb{Z}^d$ with $(z - x) \cdot \boldsymbol{\ell} = u$. (Whatever the value of $\boldsymbol{\ell}$, this event is empty for some u.) When $\mathsf{Str}_u(x)$ occurs, it is straightforward, and implied by the upcoming Lemma 2.3, that there is a unique such z, and a unique string from x to z. We denote this string by $S_u(x)$. When $x = 0 \in \mathbb{Z}^d$, write $\mathsf{Str}_u = \mathsf{Str}_u(0)$ and $S_u = S_u(0)$.

1.3.4 Notation

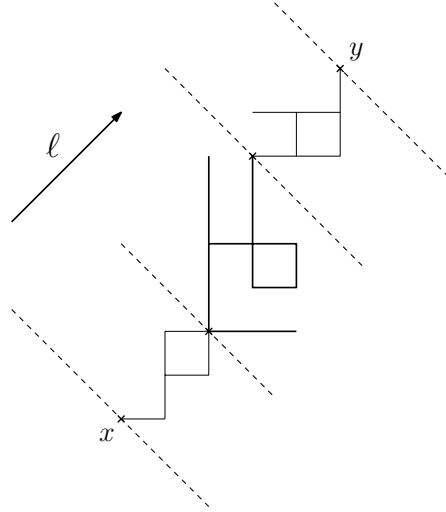The next definition develops the usage $\mathsf{Cluster}_{[u,\infty)}$ seen in Theorem 1.4.

4

Figure 2: Four crosses mark the renewal points in this string from x to y. The middle of its three sausages is drawn in bold.

Definition 1.6. Here the symbol E represents either Str or $\mathsf{Cluster}$. For $A \subseteq [0, \infty)$, we write $E_A = \cup_{u \in A} E_u$. This notation will be applied when A is a finite or semi-infinite interval.

Set $r = \max \{|e_i \cdot \boldsymbol{\ell}| : i \in [\![1, d]\!]\}$.

1.4   The asymptotic probability of long strings

The next result asserts that the string probability $\mathbb{P}\big(\mathsf{Str}_{[u,u+r)}\big)$ decays in the form $\exp\{-(\zeta+o(1))u\}$. In essence, the assertion is due to [FH14, Lemma 8.5], and we provide the straightforward derivation given this input at the end of Section 2.

Proposition 1.7. *There exists $\zeta = \zeta(p, d, \boldsymbol{\ell}) \in (0, \infty)$ such that*

$$\lim_u u^{-1} \log \mathbb{P}\big(\mathsf{Str}_{[u,u+r)}\big) = -\zeta \, .$$

Here is the string counterpart to Theorem 1.4.

Theorem 1.8. *Let $d \geq 3$, $p \in (p_c, 1)$ and $\boldsymbol{\ell} \in S^{d-1}$. Write ζ for the constant $\zeta(p, d, \boldsymbol{\ell}) \in (0, \infty)$ in Proposition 1.7. Then there exists a constant $\kappa_{\mathrm{string}} = \kappa_{\mathrm{string}}(d, p, \boldsymbol{\ell}) \in (0, \infty)$ such that, if $\boldsymbol{\ell}$ is lattice, then*

$$\mathbb{P}\big(\mathsf{Str}_u\big) \overset{\boldsymbol{\ell}}{\approx} \kappa_{\mathrm{string}} \exp\big\{ - \zeta u \big\} \, ;$$

while, if $\boldsymbol{\ell}$ is non-lattice, then

$$\mathbb{P}\big(\mathsf{Str}_{[u,\infty)}\big) \sim \kappa_{\mathrm{string}} \exp\big\{ - \zeta u \big\} \, .$$

1.5   Renewal results

Let U be a subgraph of $(\mathbb{Z}^d, \sim)$. A renewal level of U is the height of a renewal point of U. Let $R(U)$, and $\mathrm{RenLev}(U)$, denote the set of renewal points, and of renewal levels, of U.

5

We will write $\mathcal{C} = \mathcal{C}(0)$ denote the cluster with base at 0. Recall from Definition 1.6 that $\mathsf{Cluster}_{[u,u+3r]}$ denotes the event $\cup_{u \leq h \leq u+3r} \mathsf{Cluster}_h$, under which $\mathcal{C}$ has height in $[u, u+3r]$.

Theorem 1.9.

    1. There exist $c > 0$ and $u_0 \in \mathbb{N}$ such that, for $u \in [u_0, \infty)$,

$$\mathbb{P}\Big(\big|R(\mathcal{C})\big| \geq cu \,\Big|\, \mathsf{Cluster}_{[u,u+3r]} \Big) \geq c \,.$$

    2. For some $c > 0$, the condition $u^{1/2} \leq k \leq cu(\log u)^{-1}$ on $u \in (0, \infty)$ and $k \in \mathbb{N}$ implies that

$$\mathbb{P}\Big(\big|R(\mathcal{C})\big| \leq k \,\Big|\, \mathsf{Cluster}_{[u,u+3r]} \Big) \leq \exp\big\{ -cu/k \big\} \,.$$

In this theorem, and throughout, the positive constant $c > 0$ may depend on $d \geq 3$ and $p \in (p_c, 1)$.

Corollary 1.10. *Consider Theorem 1.9 with the event $\mathsf{Cluster}_{[u,u+3r]}$ in the conditioning replaced by $\mathsf{Cluster}_{[u,u+3r]} \cap \big\{ \{0, h\} \subseteq \mathrm{RenLev}(\mathcal{C}) \big\}$ where h denotes the height of $\mathrm{head}(\mathcal{C})$. In other words, condition not only on $\mathsf{Cluster}_{[u,u+3r]}$ but also on the resulting cluster having renewal points at the highest and lowest possible locations, namely $\mathrm{base}(\mathcal{C}) = 0$ and $\mathrm{head}(\mathcal{C})$. Alternatively, we may condition*

    1. Thus altered, Theorem 1.9(1) remains valid.

    2. And likewise Theorem 1.9(2).

    3. Theorem 1.9(1) remains valid if the conditioning is instead replaced by $\mathsf{Cluster}_{[u,u+3r)} \cap \big\{ \{0, h\} \subseteq \mathrm{RenLev}(\mathcal{C}) \big\}$, wherein the interval is open on the right.

1.6 Review of and comparison with prior work

Ornstein and Zernike [OZ15] presented a theory that offers a sharp asymptotic description of the density correlations in classical fluids by assuming the validity of a form of locality expressed via a renewal equation. Finding rigorous manifestations of their theory amounts to giving probabilistic descriptions of objects such as long clusters in terms of essentially one-dimensional chains of irreducible components. Early rigorous progress was made with perturbative arguments, such as [MZ96] for the Ising model at very high temperature. Self-avoiding walk with an exponential penalty for length was addressed for axial directions in [CC86] and it offered the setting for an early non-perturbative analysis in [Iof98]. Campanino and Ioffe [CI02] developed a non-perturbative approach for subcritical percolation on Euclidean lattices in work that provided sharp asymptotic estimates for the probability that distant pairs of points are connected. The analogous task for finite range Ising models above the critical temperature was undertaken by Campanino, Ioffe and Velenik in [CIV03] and for the subcritical random cluster model by these same authors in [CIV08]. The papers [CI02, CIV03, CIV08] employ a ball-and-point search algorithm that marks points on a lengthy connection and tracks where the connection leaves a ball of fixed radius (in a suitable norm) centred at the marked point. The algorithm's output offers a coarse-grained description of the long connection. In these analyses, the rare object, a long subcritical connection, is being detected and explored by a local procedure. Here a basic difference arises with our problem, where a long but finite open connection forms in supercritical percolation. For us, it is not the presence of a

long open path that is rare, but rather that this path is disjoint from the infinite open component. Indeed, the long finite path is enclosed by a surface formed of dual plaquettes that separates it from other open structure. This enclosing surface roughly plays the role analogous to the long subcritical connection in [CI02] and [CIV08]. But the surface in principle may lie far from a given point that it encloses, and there is no clear counterpart to locally operating ball-and-point algorithms that might roughly represent its form. This basic difference led us to consider very different techniques for proving asymptotics, and the resulting proofs that we present here offer new tools for deriving sharp Ornstein-Zernike asymptotics.

Campanino, Ioffe and Louidor [CIL10] find sharp asymptotics for finite point-to-point connections in supercritical bond percolation in dimension two. This counterpart to our work with $d = 2$ has a different flavour to ours, because the trapping surface whose interior contains two given distant points is a subcritical loop in the dual model, which is itself an instance of bond percolation in this dimension: thus [CIL10] may be viewed as a development of the Ornstein-Zernike theory of subcritical bond percolation.

The need for new tools is also illustrated by the way in which the coarse-grained algorithmic output description of a long connection is, in the percolation case, analysed in [CI02, (2.2)] by means of the BK inequality: disjoint connection probabilities are there bounded above by a product of marginal probabilities. In our context of finite connection, a comparable monotonicity is not available: if two points are finitely connected, then there is an open path between them which is surrounded by a surface of dual plaquettes, and these structures may or may not facilitate finite connection with a third point. In the random cluster case of [CIV08], an exponential decay of correlations is used to gain independence between modestly separated balls. A roughly counterpart assertion in our context might claim that, if a certain collection of points is finitely connected, then a dual plaquette surface will enclose all these points whose cardinality exceeds that of the minimum one necessary by a random amount whose law has a rapidly decaying tail. This however is a delicate assertion, for which the renewal abundance Theorem 1.9 may offer technical support.

Rigorous implementations of Ornstein-Zernike theory such as [CI02, CIV03, CIV08] reach conclusions including sharp asymptotics for the probability of point-to-point connections $0 \longleftrightarrow x$, in which a local central limit theorem style factor $||x||^{-(d-1)/2}$ multiplies an exponential decay rate, by establishing a *mass gap* condition. This condition asserts that a moment generating function associated to the connections decays at an exponential rate that is strictly lower than the counterpart function where connection occurs without intervening renewal points. The condition typically entails other favourable features such as the analyticity and strict convexity of the equidecay profile. For mass gap to hold, it is necessary that the conditional probability given a lengthy connection that renewal is absent decay exponentially in the distance travelled.

The principal conceptual progress made in the present article is the introduction and analysis of a slide resampling procedure to yield the renewal abundance Theorem 1.9, with the sharp asymptotic point-to-plane Theorems 1.4 and 1.8 resulting. We emphasise however that the just mentioned condition would correspond to a strengthening of Theorem 1.9(2) in which the decay rate $Ce^{-cu^{1/2}}$ is improved to Ce^{-cu}. We anticipate that the point-to-plane asymptotic will be sufficient for accompanying work, which we will mention next, concerning biased walk on the supercritical cluster. We present our results and the technical arguments that demonstrate them in the hope of attracting attention to a new technique in the Ornstein-Zernike vein as well as to offer a fundamental ingredient in a broader study of biased walk in disorder. The problem of proving mass gap is open, with

7

technical refinements of the slide resample offering one possible route to a solution.

1.6.1 Motivation from random walk in random environment

A random walk on the infinite open cluster of supercritical percolation on $(\mathbb{Z}^d, \sim)$ that has a bias towards a given preferred direction manifests a phase transition, as the strength of the bias rises, from a regime of linear progress in the long-run to one of sublinear progress in which the walk becomes waylaid for lengthy periods in traps: building on [BGP03, Szn03], [FH14] proved the existence of a critical point for this transition. In the sublinear regime, traps resemble lengthy finite percolation connections which have been opened up at one end, through which the walker may enter to become waylaid near the opposing end. Arguments in [FH14] were facilitated by percolation results of the like of Proposition 1.7, with asymptotic estimates on trap probabilities of the form $\epsilon^{-\zeta u(1+o(1))}$. The far stronger inferences made in the present article speak to the aim of analysing the sublinear regime at a finer level of detail, and we expect this article to facilitate an analysis of this regime and the scaling limit of the walk therein.

1.7 Structure of the paper

Strings are easier objects than clusters to work with in deriving asymptotic finite connection probabilities, so that the string Theorem 1.8 will be derived before the cluster Theorem 1.4, which is largely a consequence of the former result and its proof. In Section 2, we will decompose strings into irreducible building blocks called sausages, and use renewal theory to deliver Theorem 1.8 from the renewal abundance Theorem 1.9. The proof of the asymptotic cluster probability Theorem 1.4 finishes this section. Theorem 1.9 is conceptually central to this article, and it is proved by a novel *slide* resampling procedure, which is applied once to obtain Theorem 1.9(1) and in an iterative scheme to yield Theorem 1.9(2). A coarse-grained description of percolation involving a lattice of boxes of given width will facilitate these resampling arguments. In Section 3 we offer an overview of the slide resample and introduce basic aspects of the coarse-graining. In the more substantial Section 4, Theorem 1.9 is proved.

1.7.1 Acknowledgments

2 Asymptotic string and cluster probability via renewal plenitude

There are three subsections. In the first and principal Section 2.1, we use renewal theory to prove Theorem 1.8. This result has a weaker counterpart in which the strong asymptotic is replaced by the weak in the lattice case. We will first derive this weaker assertion, obtaining Proposition 2.13 (in the lattice case). The renewal abundance assertions Theorem 1.9(1) and Theorem 1.9(2) will be harnessed in our proofs: the former to obtain the weaker counterpart just mentioned; the latter, in deriving Theorem 1.8 itself. In fact, we will use the following consequence of Theorem 1.9(1), which is proved in Section 4.4.

Proposition 2.1. *There exists $c > 0$ such that, for u sufficiently high,*

$$\mathbb{P}\left(\left|R(S)\right| \geq \delta u \,\Big|\, \mathsf{Str}_{[u,u+r)}\right) \geq c\,.$$

This statement in fact raises a point about notation. We will shortly see, in Lemma 2.2(2), that the union event $\mathsf{Str}_{[u,u+r)}$ is characterized by the occurrence of $\mathsf{Str}_{u'}$ for a unique value of $u' \in [u, u+r)$. In the above, and later, we write S on $\mathsf{Str}_{[u,u+r)}$ for the string that realizes $\mathsf{Str}_{u'}$ for this value of u'.

Section 2.2 is devoted to deriving several consequences of our analysis of strings and their renewal that will be applied in the companion article. In Section 2.3, we derive the cluster probability asymptotic Theorem 1.4 by showing how a typical lengthy cluster splits into a long middle string that abuts on each side a shorter endpiece.

2.1 Deriving Theorem 1.8

Two basic lemmas concerning vertex boundary and renewal point sets will be useful. Here and later, it will be convenient to

> suppose that $e_i \cdot \boldsymbol{\ell} \geq 0$ for each $i \in [\![1,d]\!]$ and that $e_d \cdot \boldsymbol{\ell}$ attains $\max\left\{e_i \cdot \boldsymbol{\ell} : i \in [\![1,d]\!]\right\}$.   (1)

Reflecting in coordinate axes permits this assumption, which we will call the *axes ordering convention*. Note that, consequently, the quantity r specified in Subsection 1.3.4 equals $e_d \cdot \boldsymbol{\ell}$.

Lemma 2.2. *Let $h \in \mathbb{R}$ and $v \in \mathbb{Z}^d$.*

1. *The set* $\mathrm{VertBdry}_h$ *equals* $\left\{v \in \mathbb{Z}^d : v \cdot \boldsymbol{\ell} \in (h - r, h]\right\}$.

2. *Let S be a finite connected subgraph of $(\mathbb{Z}^d, \sim)$. The heights of distinct elements of $R(S)$ differ by at least r.*

Proof: (1). Let $v \in \mathrm{VertBdry}_h$. The interior of some edge abutting v lies in B_h, so that $v \cdot \boldsymbol{\ell} \leq h$; and the interior of another such edge intersects F_h, so that $(v + e_d) \cdot \boldsymbol{\ell} > h$ since $e_d \cdot \boldsymbol{\ell}$ attains $r = \max |e_i \cdot \boldsymbol{\ell}|$ by the axes ordering convention. Hence, $v \cdot \boldsymbol{\ell} > h - r$.

Now let $v \in \mathbb{Z}^d$ satisfy $v \cdot \boldsymbol{\ell} \in (h - r, h]$. The edge interior $(v, v + e_d)$ intersects F_h, so $v \in V(U_h)$. And $(v - e_d, v)$ lies in B_h, so $v \in V(L_h)$.

(2). If $y \in R(S)$, then $V(S) \cap \mathrm{VertBdry}_y = \{y\}$. Thus the preceding part implies that no element of $R(S)$ may adopt a height in $(y \cdot \boldsymbol{\ell} - r, y \cdot \boldsymbol{\ell})$. $\qquad\square$

Lemma 2.3. *A string S contains a unique element of lowest height and a unique element of greatest height. These respective vertices are $\mathrm{base}(S)$ and $\mathrm{head}(S)$. They are renewal points of S. The renewal point set $R(S)$ is strictly and totally ordered according to the elements' height.*

Proof. Recall that any given string S runs from y and to z, for certain $y, z \in \mathbb{Z}^d$, $y \cdot \boldsymbol{\ell} < z \cdot \boldsymbol{\ell}$, with $y, z \in R(S)$ and $S \subset \mathrm{Slab}_{x \cdot \boldsymbol{\ell}, y \cdot \boldsymbol{\ell}}$. Since $y \in R(S)$, $V(S) \cap \mathrm{VertBdry}_{y \cdot \boldsymbol{\ell}}$ equals $\{y\}$. Any element of $V(S)$ lies in $V(U_{y \cdot \boldsymbol{\ell}})$; and it is easy to check that any element of $V(U_{y \cdot \boldsymbol{\ell}}) \setminus \mathrm{VertBdry}_{y \cdot \boldsymbol{\ell}}$ has height strictly exceeding $y \cdot \boldsymbol{\ell}$. Thus, y is the unique element of $V(S)$ of lowest height, which forces $y = \mathrm{base}(S)$. Similar reasoning tells us that $z = \mathrm{head}(S)$ is the unique element of S of greatest height. Lemma 2.2(2) implies the strict ordering of renewal heights. $\qquad\square$

Lemma 2.3 prompts the next definition.

Definition 2.4. A sausage is a string whose vertex set contains no renewal point other than its head and base. Any string S is uniquely decomposed into sausages via its renewal point set $R(S)$. Indeed, any pair (u, v), $u \cdot \boldsymbol{\ell} < v \cdot \boldsymbol{\ell}$, of vertices of consecutive height in $R(S)$ specifies a slab $\mathrm{Slab}_{u \cdot \boldsymbol{\ell}, v \cdot \boldsymbol{\ell}}$, and the subgraph of S induced by $\mathrm{Slab}_{u \cdot \boldsymbol{\ell}, v \cdot \boldsymbol{\ell}}$ is a sausage in the decomposition of S.

Suppose that the list of sausages of a given string takes the form $\{S_i : i \in [\![1, \ell]\!]\}$ in increasing order of height. We will denote the decomposition $S = S_1 \circ S_2 \circ \cdots \circ S_\ell$.

Definition 2.5. The *slab* of a string S is $\mathrm{Slab}_{\mathrm{base}(S) \cdot \boldsymbol{\ell}, \mathrm{head}(S) \cdot \boldsymbol{\ell}}$. A string S is said to occur if, under $\mathbb{P}$, every edge in S is open, and every edge that is incident to S and lies in the slab of S is closed. We write $\mathsf{O}(S)$ for the event that S occurs.

Definition 2.6. Let Y denote the set of strings whose base equals $0 \in \mathbb{Z}^d$. Let $\mu : Y \to (0, \infty)$ be given by $\mu(S) = \exp\{\zeta \cdot \mathrm{hgt}(S)\} \mathbb{P}\big(\mathsf{O}(S)\big)$, where $\mathrm{hgt}(S)$ is the height of S as this notion is specified in Subsection 1.3.1, and where Proposition 1.7 provides the constant $\zeta > 0$.

Lemma 2.7. *Let $S = S_1 \circ S_2 \circ \cdots \circ S_\ell$ be a string. Then*

$$\mu(S) = \prod_{i=1}^{\ell} \mu(S_i) \quad and \quad \mathbb{P}\big(\mathsf{O}(S)\big) = \prod_{i=1}^{\ell} \mathbb{P}\big(\mathsf{O}(S_i)\big).$$

Proof. The event $\mathsf{O}(S)$ occurs precisely when all of the events $\mathsf{O}(S_i)$, $i \in [\![1, \ell]\!]$, do. The latter events are specified by the open-or-closed status of edges where a given edge is involved in determining the occurrence of at most one of the events $\mathsf{O}(S_i)$. Thus, these events are independent, and the latter formula in Lemma 2.7 is obtained. If we multiply this formula by

$$\exp\big\{\zeta \cdot \mathrm{hgt}(S)\big\} = \prod_{i=1}^{\ell} \exp\big\{\zeta \cdot \mathrm{hgt}(S_i)\big\},$$

then we obtain the former formula by distributing the terms in the product in the evident way. $\square$

Definition 2.8. Let $\mathscr{S}$ denote the set of sausages whose base equals $0 \in \mathbb{Z}^d$.

Proposition 2.9. *The sum $\sum_{S \in \mathscr{S}} \mu(S)$ equals one.*

A lemma will help to prove this result.

Let $u > 0$. Under $\mathbb{P}$, let $C(u)$ denote the open connected component of $0 \in \mathbb{Z}^d$ in $\mathrm{Slab}_{0,u}$. Let $\mathcal{S}_u$ denote the set of strings S with $\mathrm{base}(S) = 0$ and $\mathrm{head}(S) \cdot \boldsymbol{\ell} = u$.

Lemma 2.10. *To each $A \subseteq \mathcal{S}_u$, we may associate the event $\mathsf{A} = \{C(u) \in A\}$. For $u > 0$ and any such A,*

$$\mathbb{P}\big(\mathsf{A}\big) = \exp\big\{-\zeta u\big\} \sum_{S \in A} \mu(S).$$

Proof. For $S \in \mathcal{S}_u$, $\mathsf{O}(S) = \{C(u) = S\}$. Thus,

$$\mathbb{P}\big(\mathsf{A}\big) = \sum_{S \in A} \mathbb{P}\big(\mathsf{O}(S)\big) = \sum_{S \in A} \prod_{i=1}^{\ell} \mathbb{P}\big(\mathsf{O}(S_i)\big),$$

where the generic $S \in A$ has been decomposed into sausages according to our convention. The quantity $\prod_{i=1}^{\ell} \exp \{\zeta \cdot \mathrm{hgt}(S_i)\}$ equals $\exp \{\zeta u\}$ for any string $S \in \mathcal{S}_u$. Thus, we may multiply the last display by this shared value and use the definition of μ to obtain

$$\exp \{\zeta u\} \mathbb{P}(\mathsf{A}) \;=\; \sum_{S \in A} \prod_{i=1}^{\ell} \mu(S_i) \,.$$

Since the right-hand summand equals $\mu(S)$ by Lemma 2.7, we obtain Lemma 2.10. $\qquad\square$

Proof of Proposition 2.9. For $k \geq 2$, let $\mathcal{S}_u(k)$ denote the set of $S \in \mathcal{S}_u$ for which $|R(S)| \geq k$. Recall Definition 1.5 and note that, by Lemma 2.10,

$$\mathbb{P}\Big(\mathsf{Str}_u, \big|R(S_u)\big| \geq cu\Big) \;=\; \exp \big\{ - \zeta u \big\} \sum_{S \in \mathcal{S}_u(cu)} \mu(S) \,.$$

Write $\mathcal{S}_{[u,u+r)}(k)$ for the union $\bigcup_{h \in [u,u+r)} \mathcal{S}_h(k)$. Summing the preceding equality over values of u in an interval of the form $[h, h+r)$, and relabelling h as u, we find that

$$\mathbb{P}\Big(\mathsf{Str}_{[u,u+r)}, \big|R(S)\big| \geq cu\Big) \;\leq\; \exp \big\{ - \zeta u \big\} \sum_{S \in \mathcal{S}_{[u,u+r)}(cu)} \mu(S) \,.$$

We have that

$$\sum_{S \in \mathcal{S}_{[u,u+r)}(k)} \mu(S) \leq \sum_{j=k-2}^{\infty} \mu(\mathscr{S})^j \,,$$

because Lemma 2.7 implies that the right-hand side equals the μ-value of the set of strings S with $\mathrm{base}(S) = 0$ and at least $k - 2$ non-trivial renewal points (or, equivalently, with a decomposition having at least $k - 1$ sausages). Suppose that $\mu(\mathscr{S}) < 1$. Choosing $k = cu$, we see from the last display that

$$\mathbb{P}\big(\mathsf{Str}_{[u,u+r)}, |R(S)| \geq cu\big) \leq C \exp \big\{ - (\zeta + \epsilon)u \big\} \,.$$

for some $\epsilon > 0$ and a positive constant C. By Proposition 2.1, this bound is equally valid for $\mathbb{P}\big(\mathsf{Str}_{[u,u+r)}\big)$. The bound for $\mathbb{P}\big(\mathsf{Str}_{[u,u+r)}\big)$ contradicts Proposition 1.7: thus, we learn that $\mu(\mathscr{S}) \geq 1$.

Suppose now that $\mu(\mathscr{S}) > 1$. For $\ell \in \mathbb{N}$, let $\mathscr{S}_\ell$ denote the set of sausages whose base is $0 \in \mathbb{Z}^d$ and whose head has Euclidean distance $|| \cdot ||$ at most ℓ from 0. When the condition on the head distance is suspended, the resulting set may naturally be denoted $\mathscr{S}_\infty$, and then we have $\mathscr{S} = \mathscr{S}_\infty$. Choose $\ell \in \mathbb{N}$ such that $\mu(\mathscr{S}_\ell) > 1$. For $n \in \mathbb{N}$, let $\mathscr{S}_\ell^n$ denote the set of strings whose base is 0 and whose decomposition into sausages has n elements, each of which lies in $\mathscr{S}_\ell$. Lemma 2.7 implies that

$$\sum_{S \in \mathscr{S}_\ell^n} \mu(S) = \mu(\mathscr{S}_\ell)^n \,.$$

For $S \in \mathscr{S}_\ell$, it follows readily from Lemma 2.2(2) that $||\mathrm{head}(S)|| \in [r, \ell]$, where $r = e_d \cdot \boldsymbol{\ell} > 0$. Thus $||\mathrm{head}(S)|| \in [nr, n\ell]$ whenever $S \in \mathscr{S}_\ell^n$. Choose $z \in \mathbb{Z}^d$ with $||z|| \leq n\ell$ such that

$$\sum_{\substack{S \in \mathscr{S}_\ell^n: \\ \mathrm{head}(S)=z}} \mu(S) \;\geq\; c(n\ell)^{-d} \mu(\mathscr{S}_\ell)^n \,,$$

where $c \in (0, \infty)$ satisfies $\limsup_{u \to \infty} u^{-d} \cdot \left| \{ z \in \mathbb{Z}^d : ||u|| \leq u \} \right| < c^{-1}$. Writing $m = ||z||$, this left-hand side is at most $\exp\{\zeta m\}\mathbb{P}(\mathsf{Str}_m)$ by Lemma 2.10. Since $rn \leq m \leq n\ell$ and $\mu(\mathscr{S}_\ell) > 1$,

$$\mathbb{P}(\mathsf{Str}_m) \geq cr^d(m\ell)^{-d}\Big(\exp\{-\zeta\}\mu(\mathscr{S}_\ell)^{1/\ell} \Big)^m.$$

From $\mu(\mathscr{S}_\ell) > 1$, we see that $\limsup_{m \to \infty} m^{-1} \log \mathbb{P}(\mathsf{Str}_m) > -\zeta$. But this bound contradicts Proposition 1.7. We learn then that $\mu(\mathscr{S}) \leq 1$ and thus complete the proof of Proposition 2.9. $\square$

Proposition 2.9 permits an abuse of notation whereby μ denotes a probability measure supported on $\mathcal{S}$. Let $\mu^{\otimes \mathbb{N}}$ denote the law on infinite strings S with base at $0 \in \mathbb{Z}^d$ given by the concatenation $S = S_1 \circ S_2 \circ \cdots$ of an independent sequence of sausages drawn from the law μ. Renewal notions naturally extend to infinite strings. Let RenLev denote the set of renewal levels in the random infinite string S specified under the law $\mu^{\otimes \mathbb{N}}$. In a notational abuse, we will permit $\mu^{\otimes \mathbb{N}}[a, b]$ to denote the mean value of $\mathrm{RenLev} \cap [a, b]$ under $\mu^{\otimes \mathbb{N}}$, for a real interval $[a, b]$.

Lemma 2.11. *Let $u \in (0, \infty)$.*

1. *We have that*
$$\exp\{\zeta u\}\mathbb{P}(\mathsf{Str}_u) = \mu^{\otimes \mathbb{N}}(u \in \mathrm{RenLev}).$$

2. *Let A denote a measurable collection of finite subsets of $(0, u)$. Then*
$$\exp\{\zeta u\}\mathbb{P}(\mathsf{Str}_u, \mathrm{RenLev}(S_u) \cap (0, u) \in A) = \mu^{\otimes \mathbb{N}}(u \in \mathrm{RenLev}, \mathrm{RenLev} \cap (0, u) \in A).$$

   (Elements of A lie in the union of the spaces $S_k = (0, u)^k$ for $k \in \mathbb{N}_+$. The union is a measurable space when it is embued with the σ-algebra associated to the Borel σ-algebras on each S_k.)

3. *For $k \in \mathbb{N}$ at least two,*
$$\exp\{\zeta u\}\mathbb{P}(\mathsf{Str}_u, |R(S_u)| = k) = \mu^{\otimes \mathbb{N}}(u \in \mathrm{RenLev}, |\mathrm{RenLev} \cap (0, u)| = k - 2).$$

Now suppose that $\boldsymbol{\ell}$ is non-lattice.

4. *For $\phi \in (0, \infty)$, $\lim_{a \to \infty} \mu^{\otimes \mathbb{N}}[a, a + \phi] = \phi r^{-1}\mu_\infty$, where $\mu_\infty = \lim_{u \to \infty} \mu^{\otimes \mathbb{N}}[u, u + r]$.*

5. *We have that*
$$e^{\zeta u}\mathbb{P}(\mathsf{Str}_{[u, u+r)}) \to \mu_\infty r \zeta^{-1}(1 - e^{-\zeta r}) \text{ and } e^{\zeta u}\mathbb{P}(\mathsf{Str}_{[u, \infty)}) \to \mu_\infty r \zeta^{-1}.$$

Proof: (1). For a string S, let $d(S) = |R(S)| - 1$ denote the number of sausages into which S decomposes. Note then that

$$
\begin{aligned}
\exp\{\zeta u\}\mathbb{P}(\mathsf{Str}_u) &= \sum_{S \in \mathcal{S}_u} \mu(S) = \sum_{\ell=1}^{\infty} \sum_{\substack{S \in \mathscr{S}_u: \\ d(S) = \ell}} \mu(S) \\
&= \mu^{\otimes \mathbb{N}}\Big(\exists \ell \in \mathbb{N} : \sum_{i=1}^{\ell} \mathrm{hgt}(S_i) = u \Big) = \mu^{\otimes \mathbb{N}}(u \in \mathrm{RenLev}),
\end{aligned}
$$

where the first equality is due to Lemma 2.10.

(2). Similarly, $\exp\{\zeta u\}\mathbb{P}(\mathsf{Str}_u, \mathrm{RenLev}(S_u) \in A)$ equals

$$\mu^{\otimes\mathbb{N}}\left(\exists\, \ell \in \mathbb{N}_+ : \sum_{i=1}^{\ell} h_i = u\,,\, \big(r_i : i \in [\![1, \ell - 1]\!]\big) \in A\right),$$

where here we write $h_i = \mathrm{hgt}(S_i)$ and $r_i = \sum_{j=1}^{i} h_j$ under $\mu^{\otimes\mathbb{N}}$; whence the result.

(3). This result arises from the preceding part by taking A to equal the collection of all subsets of $(0, u)$ of cardinality $k - 2$.

(4). This is due to the non-lattice renewal theorem [Fel71, Chapter XI].

(5). Recall from Definition 1.6 that $\mathsf{Str}_{[u,u+r)}$ equals $\cup_{h\in[u,u+r)}\mathsf{Str}_h$. Lemma 2.2(2) implies that at most one of the events Str_u in this union may occur; thus, $\mathbb{P}(\mathsf{Str}_{[u,u+r]})$ equals the sum of $\mathbb{P}(\mathsf{Str}_h)$ over those $h \in [u, u + r)$ for which this probability is non-zero. From Lemma 2.11(1), we find then that

$$e^{\zeta u}\mathbb{P}(\mathsf{Str}_{[u,u+r)}) = \mu^{\otimes\mathbb{N}}[u, u+r) \sum_{h} e^{-\zeta(h-u)}\frac{\mu^{\otimes\mathbb{N}}(h\in\mathrm{RenLev})}{\mu^{\otimes\mathbb{N}}[u,u+r)}\,,$$

where the sum is over those $h \in [u, u+r)$ for which the summand is non-zero. From Lemma 2.11(4) follows the claimed convergence for $e^{\zeta u}\mathbb{P}(\mathsf{Str}_{[u,u+r)})$. To finish the proof, note that

$$e^{\zeta u}\mathbb{P}(\mathsf{Str}_{[u,\infty)}) = \sum_{j=0}^{\infty} e^{-\zeta jr} \cdot e^{\zeta(u+jr)}\mathbb{P}(\mathsf{Str}_{[u+jr,u+(j+1)r)}) \to \mu_\infty r\zeta^{-1}\big(1 - e^{-\zeta r}\big)\sum_{j=0}^{\infty} e^{-\zeta jr} = \mu_\infty r\zeta^{-1}\,.$$

This completes the proof of Lemma 2.11(5). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Proposition 2.12. *The mean value of the head height* $\mathrm{head}(S)\cdot\boldsymbol{\ell}$ *under the sausage law* μ *is finite.*

Proof. Suppose instead that $\mathbb{E}_\mu\,\mathrm{head}(S)\cdot\boldsymbol{\ell} = \infty$. Let $\delta > 0$ be given. We may choose $k \in \mathbb{N}$ so that $\mathbb{E}_\mu\big[\min\{k, \mathrm{head}(S)\cdot\boldsymbol{\ell}\}\big] \geq 2/\delta$.

With S denoting a sample of $\mu^{\otimes\mathbb{N}}$, we write S_i for the ith sausage in the decomposition of S, translated in order that the base of S_i is the origin. Set

$$X_j = X_j(S) := \sum_{i=1}^{j} \min\{k, \mathrm{head}(S_i)\cdot\boldsymbol{\ell}\}\,.$$

Note then that, whatever the choice of $k \in \mathbb{N}$,

$$\mu^{\otimes\mathbb{N}}\Big(\big|\mathrm{RenLev}\cap[0,u]\big| \geq \delta u\Big) \leq \mu^{\otimes\mathbb{N}}\big(X_{\delta u} \leq u\big)\,. \tag{2}$$

Note however that, for $j \in \mathbb{N}$, X_j is a sum of j independent and identically distributed random variables whose shared mean is at least $2/\delta$. By the exponential Markov inequality, there exists $\epsilon > 0$ such that, for $u \in \mathbb{N}$,

$$\mu^{\otimes\mathbb{N}}\big(X_{\delta u} \leq u\big) \leq \exp\big\{-\epsilon u\big\}\,. \tag{3}$$

By Lemma 2.11(3),

$$\exp\{\zeta u\}\mathbb{P}\big(\mathsf{Str}_u, \big|R(S_u)\big| \geq \delta u + 2\big) = \mu^{\otimes\mathbb{N}}\Big(\big|\mathrm{RenLev}\cap(0,u)\big| \geq \delta u\,,\, u \in \mathrm{RenLev}\Big)\,.$$

13

Recalling that $\mathsf{Str}_{[u,u+r)} = \cup_{h \in [u,u+r)}\mathsf{Str}_h$, we now sum over those heights u lying in a given range $[h, h+r)$, and relabel h as u, to obtain

$$\exp\left\{\zeta u\right\}\mathbb{P}\left(\mathsf{Str}_{[u,u+r)}, \left|R(S)\right| \geq \delta u + 2\right) \leq \mu^{\otimes \mathbb{N}}\left(\left|\mathrm{RenLev} \cap (0,u)\right| \geq \delta u, \ [u, u+r) \cap \mathrm{RenLev} \neq \emptyset\right).$$

From this display alongside (2) and (3), we find that

$$\mathbb{P}\left(\mathsf{Str}_{[u,u+r)}, \left|R(S)\right| \geq \delta u + 2\right) \leq \exp\left\{-(\zeta + \epsilon)u\right\}$$

when u is high enough. Since $\delta u + 2 \leq 2\delta u$ for high u, we may select δ to be one-half of the value of this constant in Proposition 2.1 and learn from this result that

$$\mathbb{P}\left(\mathsf{Str}_{[u,u+r)}\right) \leq c^{-1}\exp\left\{-(\zeta + \epsilon)u\right\}.$$

This conclusion is in contradiction with Proposition 1.7. Thus, $\mathbb{E}_\mu\,\mathrm{head}(S)\cdot\boldsymbol{\ell} < \infty$, as we sought to show in demonstrating Proposition 2.12. □

We are ready to prove the weaker counterpart of Theorem 1.8 in the lattice case.

Proposition 2.13. *Suppose that $\boldsymbol{\ell} \in S^{d-1}$ is lattice. As $u \to \infty$ through the cover of $\boldsymbol{\ell}$, the limit $\mu^{\otimes \mathbb{N}}\left(u \in \mathrm{RenLev}\right)$ exists and lies in $(0,1)$. Denoting it by g, we have that*

$$\mathbb{P}\left(\mathsf{Str}_u\right) \overset{\boldsymbol{\ell}}{\sim} g\exp\left\{-\zeta u\right\}.$$

Proof. The height $\mathrm{head}(S)\cdot\boldsymbol{\ell}$ under μ has finite mean by Proposition 2.12, and the support of the distribution of $\mathrm{head}(S)\cdot\boldsymbol{\ell}$ under μ is readily seen in the lattice case to have greatest common denominator equal to the lattice span s. The renewal theorem [Fel71, Chapter XI] thus implies the existence of the concerned limit. The latter statement of Proposition 2.13 then follows from Lemma 2.11(1). □

Lemma 2.14. *Let $\boldsymbol{\ell} \in S^{d-1}$ be lattice with span s. There exist positive choices for C and c such that, for $u \in s \cdot \mathbb{N}$ (and g specified by Proposition 2.13), $\left|\mu^{\otimes \mathbb{N}}\left(u \in \mathrm{RenLev}\right) - g\right| \leq C\exp\left\{-cu^{1/2}\right\}$.*

The next definition and result will facilitate the proof of the one just stated.

Definition 2.15. Let G be a finite subgraph of $(\mathbb{Z}^d, \sim)$. Let $h_{\min}$ and $h_{\max}$ denote the minimum and maximum heights of elements of $V(G)$. Consider the sequence that begins with $h_{\min}$, ends with $h_{\max}$, and, in between, lists in increasing order the heights of renewal points of G that lie strictly between $h_{\min}$ and $h_{\max}$. Let $\mathrm{MaxRenGap}(G)$ denote the maximum difference of consecutive terms in this sequence.

For use a little later, take $k \in \mathbb{N}$ and let $\mathrm{MaxRenGap}(G, k)$ denote the maximum difference among the final k pairs of consecutive terms, where we take $\mathrm{MaxRenGap}(G, k) = \mathrm{MaxRenGap}(G)$ in the case that there are at most k terms.

Lemma 2.16. *Let $\boldsymbol{\ell} \in S^{d-1}$ be lattice with span s, and let $u \in s \cdot \mathbb{N}$.*

1. *We have that*

$$\mathbb{P}\left(\mathrm{RenLev}(S) \cap (0,u) = \emptyset, \mathsf{Str}_u\right) \leq (1-p)^{-d}\,\mathbb{P}\left(\mathrm{RenLev}(\mathcal{C}_u) \cap (0,u) = \emptyset, \mathsf{Cluster}_u\right).$$

2. *And that*

$$\mathbb{P}\big(\mathsf{Cluster}_{(u,u+r]}, \mathrm{MaxRenGap}(\mathcal{C}) \geq \ell\big) \leq p^{-2}(1-p)^{5-4d}\mathbb{P}\big(\mathsf{Str}_{(u+2r,u+3r]}, \mathrm{MaxRenGap}(S) \geq \ell\big),$$

where recall that the event $\mathsf{Cluster}_{(u,u+r]}$ *is specified by Definition 1.6 and* $\mathcal{C}$ *denotes the cluster containing* 0.

3. *And also that* $\mathbb{P}\big(\mathsf{Str}_{[u,u+r)}\big) \leq p^{-1}(1-p)^{1-2d}\mathbb{P}\big(\mathsf{Str}_{[u+r,u+2r)}\big).$

Proof: (1). Set $\mathbf{h} = \mathrm{head}(S_u)$ and $h = \mathbf{h} \cdot \boldsymbol{\ell}$. When Str_u occurs, and edges incident to $\mathbf{h}$ and to an element of $V(U_h) \setminus \mathrm{VertBdry}_h$ are closed, then $\mathsf{Cluster}_u$ occurs, with $\mathcal{C}_u$ equal to S_u. The event Str_u is independent of the status of the just mentioned edges, of which there are at most d, so that the desired bound results.

(2). The probability $\mathbb{P}\big(\mathsf{Cluster}_{(u,u+r]}, \mathrm{MaxRenGap}(\mathcal{C})$ may be expressed as a sum of terms p_D, the probability of the event E_D that the cluster of the origin equals D, as D ranges over all connected components whose base equals 0, whose head has height in $[u, u+r]$ and for which $\mathrm{MaxRenGap}(\mathcal{C}) \geq \ell$. The occurrence of E_D is characterized by the edges in D being open, and the neighbouring edges being closed. Consider instead the event F_D which is specified by the same set of open-and-closed conditions as is E_D, with the exception that we instead demand that the edges $e := (-e_d, 0)$ and $e' := \big(\mathrm{head}(D), \mathrm{head}(D)+e_d\big)$ be open, and edges that lie in $\mathrm{Slab}_{-e_d \cdot \boldsymbol{\ell}, (\mathrm{head}(D)+e_d) \cdot \boldsymbol{\ell}}$ but not in D and that border either of these two edges be closed. We have that $\mathbb{P}(F_D) = \alpha\beta\mathbb{P}(E_D)$ where the factor $\alpha = p^2(1-p)^{-2}$ addresses the switch in status for e and e' and $\beta = (1-p)^{\sigma}$, with $\sigma = \sigma(D) \in [\![0, 4d-3]\!]$, represents the closure of the further edges under F_D. When F_D occurs, the subgraph with edge-set $D \cup \{e\} \cup \{e'\}$ is a string with base $-e_d$ and head $\mathrm{head}(D) + e_d$. Indeed, $\mathrm{head}(D) + e_d$ has height at least r more than any other element of this subgraph, while $-e_d$ has height less than that of any such element by at least r; thus, Lemma 2.2(1) shows that these two vertices are renewal points in the subgraph and demonstrates that the subgraph is indeed a string.

Let $F_D^{\uparrow}$ denote the equiprobable event that F_D occurs after the configuration is shifted by $-e_d$. This collection of events, indexed by the above set of D, is disjoint, with union lying in $\mathsf{Str}_{(u+2r,u+3r]} \cap \big\{\mathrm{MaxRenGap}(S) \geq \ell\big\}$. In summary,

$$\begin{aligned}\mathbb{P}\big(\mathsf{Cluster}_{(u,u+r]}, \mathrm{MaxRenGap}(\mathcal{C}) \geq \ell\big) &= \sum_D \mathbb{P}(E_D) \leq p^{-2}(1-p)^{5-4d}\sum_D \mathbb{P}(F_D^{\uparrow}) \\ &\leq p^{-2}(1-p)^{5-4d}\mathbb{P}\big(\mathsf{Str}_{(u+2r,u+3r]}, \mathrm{MaxRenGap}(S) \geq \ell\big),\end{aligned}$$

so that Lemma 2.16(2) is obtained.

(3). This proof operates similarly to the preceding: open the e_d-directed edge bordering the string S that realizes $\mathsf{Str}_{(u,u+r]}$, and then close the neighbours of the edge's other endpoint that lie in $\mathrm{Slab}_{0,(\mathrm{head}(S)+e_d) \cdot \boldsymbol{\ell}}$. $\qquad\square$

Proof of Lemma 2.14. We will argue that, for some positive constants C and c,

$$\mu\big(\mathrm{hgt}(S) \geq u\big) \leq C\exp\big\{-cu^{1/2}\big\} \text{ for } n \in s\mathbb{N}. \tag{4}$$

Indeed, admitting this, we may invoke [SW67, Theorem 1] with the function $M(x)$ set equal to the expression $\exp\big\{-c/2 \cdot u^{1/2}\big\}$, and relabel $c > 0$, in order to obtain Lemma 2.14.

To derive (4), note that

$$\mu\big(\mathrm{hgt}(S) \geq u\big) \;=\; \sum_{v \geq u} \mu^{\otimes \mathbb{N}}\big(\mathrm{RenLev} \cap (0,v) = \emptyset\,,\, v \in \mathrm{RenLev}\big)$$

$$=\; \sum_{v \geq u} \mathbb{P}\Big(\mathrm{RenLev}(S) \cap (0,v) = \emptyset\,\Big|\,\mathsf{Str}_v\Big)\mathbb{P}(v \in \mathrm{RenLev})\,,$$

where v is understood to run values for which the summand at stake is non-zero. Here we used that $\mu^{\otimes \mathbb{N}}\big(\mathrm{RenLev} \cap (0,v) = \emptyset\,\big|\,v \in \mathrm{RenLev}\big) = \mathbb{P}\big(\mathrm{RenLev}(S) \cap (0,v) = \emptyset\big|\mathsf{Str}_v\big)$, a fact due to 0 being a renewal level in any of the concatenated sausage sequences sampled under $\mu^{\otimes \mathbb{N}}$. To obtain (4), it is thus enough to verify that

$$\mathbb{P}\Big(R(S) \cap (0,u) = \emptyset\,\Big|\,\mathsf{Str}_u\Big) \;\leq\; C\exp\big\{-cu^{1/2}\big\}\,. \tag{5}$$

To this end, note that, by considering the possibility that the edges not in $\mathrm{Slab}_{0,u}$ that are incident to either the base or the head of the string realizing Str_u are closed, we obtain the first bound as we write

$$\mathbb{P}\Big(R(S) \cap (0,u) = \emptyset\,,\,\mathsf{Str}_u\Big) \;\leq\; (1-p)^{-2d}\mathbb{P}\Big(R(\mathcal{C}) = \{0,u\}\,,\,\mathsf{Cluster}_u\Big)$$

$$\leq\; (1-p)^{-2d}\mathbb{P}\Big(R(\mathcal{C}) = \{0,h(\mathcal{C})\}\,,\,\mathsf{Cluster}_{[u-3r,u]}\Big)$$

$$\leq\; C\exp\big\{-cu^{1/2}\big\}\mathbb{P}\Big(\{0,h(\mathcal{C})\} \subseteq R(\mathcal{C})\,,\,\mathsf{Cluster}_{[u-3r,u]}\Big)\,,$$

where we write $h(\mathcal{C})$ for the height of $\mathrm{head}(\mathcal{C})$, with $\mathcal{C}$ the cluster of the origin specified when $\mathsf{Cluster}_{[u-3r,u]}$ occurs; and where the latter bound is due to Corollary 1.10(2).

By restriction of the configuration to a slab, we see that

$$\mathbb{P}\Big(\{0,h(\mathcal{C})\} \subseteq R(\mathcal{C})\,,\,\mathsf{Cluster}_{[u-3r,u]}\Big) \leq \mathbb{P}\Big(\bigcup_{v \in [u-3r,u]} \mathsf{Str}_v\Big)\,.$$

Now,

$$\mathbb{P}\Big(\bigcup_{v \in [u-3r,u]} \mathsf{Str}_v\Big) \;=\; \sum_{\substack{v \in s\mathbb{N} \\ v \in [u-3r,u]}} \mathbb{P}(\mathsf{Str}_v) \;\leq\; C\,\mathbb{P}(\mathsf{Str}_u)\,. \tag{6}$$

The equality invokes that $\boldsymbol{\ell}$ is lattice with span s. The bound, illustrated in Figure 3, arises by considering the configuration outside $\mathrm{Slab}_{0,v}$ and the event that an open path delimited by closed edges reaches to height u. Since this path has finite length, the constant C, while dependent on p, is finite.

Assembling these bounds, we learn that $\mathbb{P}\big(R(S) \cap (0,u) = \emptyset\,,\,\mathsf{Str}_u\big) \;\leq\; Ce^{-cu^{1/2}}\mathbb{P}(\mathsf{Str}_u)$ for a p-dependent constant C. From this conclusion, (5) naturally follows; in this way, we obtain Lemma 2.14. $\qquad\square$

Proof of Theorem 1.8. The non-lattice case has already been obtained, with

$$\kappa_{\mathrm{string}} = \zeta^{-1}r\lim_{u \to \infty} \mu^{\otimes \mathbb{N}}[u, u+r)\;:$$

it is Lemma 2.11(5). In the lattice case, the conclusion is due to Lemmas 2.11(1) and 2.14, with $\kappa_{\mathrm{string}} = \lim_{u \to \infty} \mu^{\otimes \mathbb{N}}(\{u\})$. $\qquad\square$
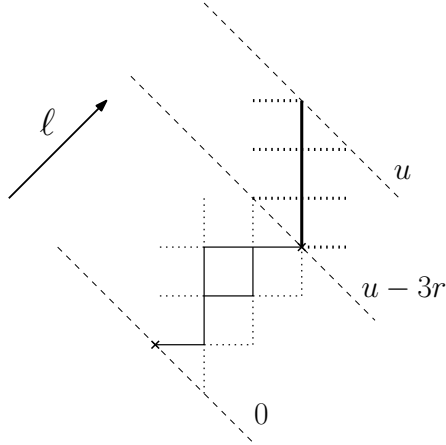
Figure 3: Illustrating the bound in (6). In this two dimensional example, $\boldsymbol{\ell}$ makes a forty-five degree angle with the axes, and the quantities s and r in (6) equal $2^{-1/.2}$. The two crosses mark the base and head of a string that realizes Str_{u-3r}, with closed edges bordering the string shown with dotted lines. This string may be extended to form a string realizing Str_u: three consecutive vertical edges emanating from the head are marked in bold, with bordering edges shown bold dotted; if the former are realized open, and the latter closed, then the desired extension occurs, and it does so with positive probability.

2.2 Further results on the renewal structure of strings

In separate subsections, we state and prove several results that address the renewal structure of strings and that will be applied in the companion paper. And then we prove Proposition 1.7.

2.2.1 Strings with a small constant fraction of renewal points are rare

Our first result is an assertion to this effect.

Proposition 2.17. *We have that*

$$\mathbb{P}\left(\bigcup_{w \in [u, u+r)} \mathsf{Str}_w \cap \left\{ |R(S_w)| \leq cw \right\} \right) \leq C \exp\left\{ -\zeta u - cu^{1/2} \right\}. \tag{7}$$

Just as after Theorem 1.9 we permitted generally $c > 0$ to depend on $d \geq 3$ and $p \in (p_c, 1)$, now we point out that in Proposition 2.17 and later, we also permit this dependence for $C > 0$.

For the proof of the proposition, let Z_i be specified under the law $\mu^{\otimes \mathbb{N}}$ to be the difference between the $(i+1)^{\text{st}}$ and the ith elements of RenLev for $i \in \mathbb{N}$.

Lemma 2.18. *The sequence $\left\{ Z_i : i \in \mathbb{N} \right\}$ is independent and identically distributed. For a p-dependent positive constant C, we have that $\mathbb{P}(Z_1 \geq k) \leq C \exp\left\{ -ck^{1/2} \right\}$.*

17

Proof. The first statement holds by the definition of RenLev. The second holds because

$$\mathbb{P}(Z_1 \geq u) = \mu^{\otimes \mathbb{N}}\Big(\exists v \geq u : v \in \mathrm{RenLev}\,,\ \mathrm{RenLev} \cap (0, u) = \emptyset\Big)$$

$$= \sum_{v \geq u} \mu^{\otimes \mathbb{N}}\Big(v \in \mathrm{RenLev}\,,\ \mathrm{RenLev} \cap (0, u) = \emptyset\Big)$$

$$= \sum_{v \geq u} e^{\zeta v} \mathbb{P}\Big(\mathsf{Str}_v\,,\ |R(S_v)| = 2\Big)$$

$$\leq C \sum_{v \geq u} e^{\zeta v} \mathbb{P}\Big(\mathsf{Cluster}_v\,,\ R(\mathcal{C}) = \{0, v\}\Big)$$

$$\leq C \sum_{v \geq u} e^{\zeta v - c v^{1/2}} \mathbb{P}\Big(\mathsf{Cluster}_v\,,\ \{0, v\} \subseteq R(\mathcal{C})\Big)$$

$$\leq C \sum_{v \geq u} e^{\zeta v - c v^{1/2}} \mathbb{P}\Big(\mathsf{Str}_v\Big)$$

$$\leq C \exp\big\{-c u^{1/2}\big\},$$

where each sum is taken over all $v \geq u$ for which the summand is non-zero. The first inequality is due to Lemma 2.16(1). To derive the second, we split the sum over $v \geq u$ by division of $[u, \infty)$ into intervals of length three, and apply Corollary 1.10(2) to each of the resulting probabilities. That the restriction of a configuration realizing $\mathsf{Cluster}_v$ to $\mathrm{Slab}_{0,v}$ realizes Str_v justifies the third bound. The fourth is due to Theorem 1.8. □

Proof of Proposition 2.17. By Lemma 2.11(1,3),

$$\mathbb{P}\Big(\mathsf{Str}_w, |R(S_w)| \leq cw\Big) \leq \exp\big\{-\zeta w\big\} \mu^{\otimes \mathbb{N}}\Big(w \in \mathrm{RenLev}\,,\ \big|\mathrm{RenLev} \cap (0, w)\big| \leq \lfloor cw \rfloor - 2\Big).$$

for $w > 0$ with $cw \geq 2$. The left-hand side in (7) is bounded above by the sum of the left-hand of the preceding display over all those $w \in [u, u+r)$ for which the event Str_w is non-empty. Thus, the concerned probability is also bounded above by the sum of the right-hand terms over such indices, and thus by

$$\exp\big\{-\zeta u\big\} \cdot \mu^{\otimes \mathbb{N}}\Big(\big|\mathrm{RenLev} \cap (0, u)\big| \leq \lfloor c(u+1) \rfloor - 2\Big) \cdot \kappa$$

where κ equals the mean value of $\big|\mathrm{RenLev}[u, u+r)\big|$ under $\mu^{\otimes \mathbb{N}}$ given $\big|\mathrm{RenLev} \cap (0, u)\big| \leq \lfloor c(u+1) \rfloor - 2$. The value of κ may be bounded above by instead conditioning on the maximum value of RenLev that is at most u; since the random variable Z_1 is almost surely positive, we see that κ is bounded above, uniformly in u. Further, we have that

$$\mu^{\otimes \mathbb{N}}\Big(\big|\mathrm{RenLev} \cap (0, u)\big| \leq \lfloor c(u+1) \rfloor - 2\Big) = \mathbb{P}\Big(\sum_{i=1}^{\lfloor c(u+1) \rfloor - 1} Z_i \geq u\Big).$$

We find from Lemma 2.18 and [Nag69] that $\mathbb{P}\big(\sum_{i=1}^{\lfloor cu \rfloor} Z_i > u\big) \leq C \exp\big\{-c u^{1/2}\big\}$ if $c > 0$ is chosen to be suitably small. In light of the bounds above, we obtain Proposition 2.17. □

2.2.2 The start and end of a long string

After presenting some notation, we state and prove Lemma 2.20, which addresses the joint form of the two ends of a long string.

Definition 2.19. For $k \in \mathbb{N}$, let $\mathcal{S}_k$ denote the set of strings that contain k sausages. Write $\mu^{\otimes k}(S) = \prod_{i=1}^{k} \mu(S_i)$ for $S \in \mathcal{S}_k$ of the form $S = S_1 \circ \cdots \circ S_k$.

For $k, \ell \in \mathbb{N}$, let $S_+ \in \mathcal{S}_k$ and $S_- \in \mathcal{S}_\ell$. Let $E(u, k, \ell, S_+, S_-)$ denote the event that Str_u occurs and that there exists a string S such that S_u has the form $S_+ \circ S \circ S_-$: in other words, that S_u contains at least $k + \ell$ sausages, with the initial string of length k being S_+ and the final string of length ℓ being S_-.

Lemma 2.20. *We have that*

$$\mathbb{P}\big(E(u, k, \ell, S_+, S_-)\big|\mathsf{Str}_u\big) = \mu^{\otimes k}(S_+)\mu^{\otimes \ell}(S_-)\frac{\mu^{\otimes \mathbb{N}}\big(u - \mathrm{height}(S_+) - \mathrm{height}(S_-) \in \mathrm{RenLev}\big)}{\mu^{\otimes \mathbb{N}}\big(u \in \mathrm{RenLev}\big)} .$$

Proof. Since $E(u, k, \ell, S_+, S_-) \subset \mathsf{Str}_u$ and

$$\mathbb{P}\big(E(u, k, \ell, S_+, S_-)\big) = \mu^{\otimes k}(S_+)\mathbb{P}\big(\mathsf{Str}_{u-\mathrm{height}(S_+)-\mathrm{height}(S_-)}\big)\mu^{\otimes \ell}(S_-) ,$$

the result follows from Lemma 2.11(1). $\square$

2.2.3 The maximum gap between renewals

Here we prove Proposition 2.21, which concerns the maximum displacement of consecutive renewal levels in a long string or cluster, and which is expressed in the terminology of Definition 2.15.

Proposition 2.21. *Let $\ell \in \mathbb{N}$ and $(\mathsf{E}, E) \in \big\{(\mathsf{Str}, S), (\mathsf{Cluster}, \mathcal{C})\big\}$.*

1. *Let $u \in (0, \infty)$ and suppose that $\ell \geq C(\log u)^2$. If $\boldsymbol{\ell}$ is non-lattice, then*

$$\mathbb{P}\big(\mathrm{MaxRenGap}(E_{(u,u+r]}) \geq \ell, \, \mathsf{E}_{(u,u+r]}\big) \leq C \exp\big\{-\zeta u - c\ell^{1/2}\big\} .$$

   If $\boldsymbol{\ell}$ is lattice, this statement holds when $E_{(u,u+r]}$ is replaced by E_u.

2. *Let $k \in \mathbb{N}$ and suppose that $\ell \geq C(\log k)^2$. If $\boldsymbol{\ell}$ is non-lattice, then*

$$\mathbb{P}\big(\mathrm{MaxRenGap}(E_{(u,u+r]}, k) \geq \ell, \, \mathsf{E}_{(u,u+r]}\big) \leq C \exp\big\{-\zeta u - ck^{1/2}\big\} .$$

   If $\boldsymbol{\ell}$ is lattice, this statement holds when $E_{(u,u+r]}$ is replaced by E_u.

Proof: (1). Call the claims the *string* and *cluster* bounds. The string bound is needed to prove the cluster bound and we derive it first. Suppose that $\boldsymbol{\ell}$ is non-lattice. When $\mathsf{Str}_{(u,u+r]}$ occurs, Lemma 2.2(2) implies that there is a unique string with base 0 whose height lies in $(u, u + r]$, which we will denote by S. Apply Lemma 2.11(2) with $u = h$ and A equal to the collection of finite subsets of $(0, h)$ for which some pair of consecutive terms in $\{0\} \cup A \cup \{h\}$ differ by at least ℓ. Sum the resulting conclusion over $h \in (u, u + r]$ to find that

$$\mathbb{P}\big(\mathrm{MaxRenGap}(S) \geq \ell, \mathsf{Str}_{(u,u+r]}\big) = \exp\{-\zeta u\}\mu^{\otimes \mathbb{N}}(u, u + r] \cdot \psi_{u,\ell} ,$$

where $\psi_{n,\ell}$ is the probability that there exist consecutive elements in $\mathrm{RenLev} \cap [0, u + r]$ of difference at least ℓ under the law $\mu^{\otimes \mathbb{N}}\big(\cdot \big|\mathrm{RenLev} \cap (u, u + r] \neq \emptyset\big)$. Since $\mu^{\otimes \mathbb{N}}(u, u + r]$ is uniformly positive by Lemma 2.11(3), we may apply Lemma 2.18 to find that $\psi_{\ell,u} \leq u \exp\big\{-c\ell^{1/2}\big\}$. The condition

$\ell \geq C(\log u)^2$ then implies the string bound in the non-lattice case. When $\boldsymbol{\ell}$ is lattice, we instead have

$$\mathbb{P}\big(\mathrm{MaxRenGap}(S) \geq \ell, \mathsf{Str}_u\big) = \exp\{-\zeta u\}\mu^{\otimes\mathbb{N}}\{u\} \cdot \psi_{u,\ell},$$

where we respecify $\psi_{n,\ell}$ by instead considering the conditional law $\mu^{\otimes\mathbb{N}}\big(\cdot \,\big|\, u \in \mathrm{RenLev}\big)$. In this lattice case, $\mu^{\otimes\mathbb{N}}(u \in \mathrm{RenLev})$ is uniformly positive by Proposition 2.13. The lattice string bound arises with these variations to the non-lattice case.

The cluster bound is due to Lemma 2.16(2).

(2). This proof is similar to the preceding one. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

2.2.4 Deriving Proposition 1.7

We conclude Section 2 by deriving Proposition 1.7 from [FH14, Lemma 8.5].

Proof of Proposition 1.7. In the notation of [FH14],

$$\mathcal{H}^+(k) = \big\{x \in \mathbb{Z}^d : x \cdot \boldsymbol{\ell} \geq k\big\} \text{ and } \mathcal{H}^-(k) = \big\{x \in \mathbb{Z}^d : x \cdot \boldsymbol{\ell} \leq k\big\}.$$

In [FH14], a notion of sausage connection is specified. In our terms, the event $\{0 \overset{\text{s.c.}}{\longleftrightarrow} \mathcal{H}^-(-u)\}$ is characterized by the existence of a point $b \in \mathbb{Z}^d$ with $-u - r \leq b \cdot \boldsymbol{\ell} \leq -u$ such that there exists a string S with base b and head 0 with the set of edges in S that are incident to either b or 0 equal to $\big\{(b, b+e_1), (-e_1, 0)\big\}$.

We *claim* that

$$\mathbb{P}\Big(0 \overset{\text{s.c.}}{\longleftrightarrow} \mathcal{H}^-(-u)\Big) \leq \mathbb{P}\big(\mathsf{Str}_{[u,u+r]}(0)\big). \tag{8}$$

Indeed, by translating the string that characterizes the occurrence of the sausage connection event by the vector $-b$, we see that the left-hand side of the above display is at most the probability that there exists a vertex $h \in \mathbb{Z}^d$ with $u \leq h \cdot \boldsymbol{\ell} \leq u + r$ such that there exists a string with base 0 and head h; the latter event is simply $\mathsf{Str}_{[u,u+r]}$, so the claim follows.

Consider the union U over $b \in \mathbb{Z}^d$ satisfying $-u \leq b \cdot \boldsymbol{\ell} < -u + r$ of the event $S(b, -e_1) \cap O^-(b) \cap O^+(-e_1)$, where $S(b, -e_1)$ is the event that there exists a string with base b and head $-e_1$; $O^-(b)$ is the event that the edge $(b - e_i, b)$ is open for $i \in [\![1, d]\!]$ if and only if $i = 1$; and $O^+(-e_1)$ is the event that the edge $(-e_1, -e_1 + e_i)$ is open for $i \in [\![1, d]\!]$ if and only if $i = 1$.

Note that $U \subseteq \{0 \overset{\text{s.c.}}{\longleftrightarrow} \mathcal{H}^-(-u)\}$. By translating configurations so that the point b is relocated to be the origin, we obtain the bound

$$\big(p(1-p)^{d-1}\big)^2 \mathbb{P}\Big(\mathsf{Str}_{[u-r,u)}(0)\Big) \leq \mathbb{P}\Big(0 \overset{\text{s.c.}}{\longleftrightarrow} \mathcal{H}^+(u)\Big).$$

Using the same translation on (8), we find that

$$\mathbb{P}\Big(0 \overset{\text{s.c.}}{\longleftrightarrow} \mathcal{H}^+(u)\Big) \leq \mathbb{P}\Big(\mathsf{Str}_{[u-r,u]}(0)\Big).$$

Since $\mathsf{Str}_{[u-r,u]}(0) \subset \mathsf{Str}_{[u-r,u)}(0) \cup \mathsf{Str}_{[u,u+r)}(0)$, the two preceding displays imply that

$$\mathbb{P}\Big(0 \overset{\text{s.c.}}{\longleftrightarrow} \mathcal{H}^+(u)\Big) \leq \mathbb{P}\Big(\mathsf{Str}_{[u-r,u]}(0)\Big) \leq \big(p(1-p)^{d-1}\big)^{-2} \left(\mathbb{P}\Big(0 \overset{\text{s.c.}}{\longleftrightarrow} \mathcal{H}^+(u)\Big) + \mathbb{P}\Big(0 \overset{\text{s.c.}}{\longleftrightarrow} \mathcal{H}^+(u+r)\Big)\right).$$

[FH14, Lemma 8.5] asserts that $\mathbb{P}\big(0 \overset{\text{s.c.}}{\longleftrightarrow} \mathcal{H}^+(u)\big) = \exp\big\{ - \zeta u(1 + o(1))\big\}$. The just derived bounds shows that this asymptotic passes to $\mathbb{P}\big(\mathsf{Str}_{[u-r,u]}(0)\big)$. We invoke Lemma 2.16(3) in the final inequality as we write

$$
\begin{aligned}
\mathbb{P}\Big(\mathsf{Str}_{[u,u+r)}(0)\Big) \le \mathbb{P}\Big(\mathsf{Str}_{[u,u+r]}(0)\Big) \;&\le\; \mathbb{P}\Big(\mathsf{Str}_{[u,u+r)}(0)\Big) + \mathbb{P}\Big(\mathsf{Str}_{[u+r,u+2r)}(0)\Big) \\
&\le\; (1 + C)\mathbb{P}\Big(\mathsf{Str}_{[u+r,u+2r)}(0)\Big).
\end{aligned}
$$

From $\mathbb{P}\big(\mathsf{Str}_{[u,u+r]}(0)\big) = e^{-\zeta u(1+o(1))}$, we learn that $\mathbb{P}\big(\mathsf{Str}_{[u,u+r)}(0)\big)$ also satisfies this bound. This completes the proof of Proposition 1.7. □

2.3 Deriving the asymptotic probability of clusters

Here we prove Theorem 1.4. A cluster typically has at least two renewal points. When it does so, it may be split into three pieces at its lowest and highest renewal points. The middle piece is a string, which is bookended by a cul-de-sac and an inverted cul-de-sac, each with unique renewal points. First we specify these *culs-de-sac*.

The cul-de-sac. Let $x \in U_0$. A cul-de-sac with entry x is a finite open connected component of the lower half-space $L_{x \cdot \boldsymbol{\ell}}$ of which 0 is the base and x is a renewal point. We denote this object by $\mathsf{CdS}[x]$ and write $\mathsf{CulDeSac}[x]$ for the event that it exists. For $u > 0$, we write $\mathsf{CulDeSac}_u = \cup\, \mathsf{CulDeSac}[x]$ where the union is over $x \in \mathbb{Z}^d$ with $x \cdot \boldsymbol{\ell} = u$. Note that, if u is such that $\mathsf{CulDeSac}_u$ is non-empty, then there is only one such value of x for which $\mathsf{CulDeSac}[x]$ occurs.

Now let $x \in L_0$. An inverted cul-de-sac with entry x is a finite open connected component of the upper half-space $U_{x \cdot \boldsymbol{\ell}}$ of which 0 is the head and x is a renewal point. We write $\mathsf{InvCulDeSac}$ and ICdS in place of $\mathsf{CulDeSac}$ and CdS.

The next definition and result speak to the decomposition of a cluster into three pieces.

Definition 2.22. On the event $\mathsf{Cluster}_u$, with $u \in (0,\infty)$ given, we write $r_{\max}$ and $r_{\min}$ for the maximum and minimum heights of renewal points of $\mathcal{C}_u$.

Lemma 2.23. *Let* $u \in (0,\infty)$ *and* $j, \ell \in [0, u]$ *with* $j + \ell \le u$. *We have that*

$$
\begin{aligned}
&\mathbb{P}\Big(\mathsf{Cluster}_u\,,\, r_{\min} = j\,,\, r_{\max} = u - \ell\Big) \\
=\;\; &\mathbb{P}\big(\mathsf{CulDeSac}_j, \mathrm{RenLev}(\mathsf{CdS}_j) = \{j\}\big)\mathbb{P}\big(\mathsf{Str}_{u-j-\ell}\big)\mathbb{P}\big(\mathsf{InvCulDeSac}_\ell, \mathrm{RenLev}(\mathsf{ICdS}_\ell) = \{0\}\big).
\end{aligned}
$$

Proof. Consider the percolation configurations restricted to the sets $L_{r_{\min}}$, $\mathrm{Slab}_{r_{\min}, r_{\max}}$ and $U_{r_{\max}}$ in the event $\mathsf{Cluster}_u \cap \{r_{\min} = j, r_{\max} = u - \ell\}$. As Figure 4 illustrates, the latter event is characterized by the occurrence of the three right-hand events in the above display, where in the second case, the configuration is translated so that $\mathrm{head}(\mathsf{CdS}_j)$ appears at the origin, and the third configuration is also shifted, so that the head of the string in the second event is also at the origin. These three events are respectively measurable with respect to percolation configurations restricted to $L_{r_{\min}}$, $\mathrm{Slab}_{r_{\min}, r_{\max}}$ and $U_{r_{\max}}$, so their probabilities multiply, as asserted in the lemma. □

Next we prove that unique-renewal culs-de-sac are rare when judged relative to the $e^{-\zeta u}$ decay in Theorems 1.4 and 1.8.
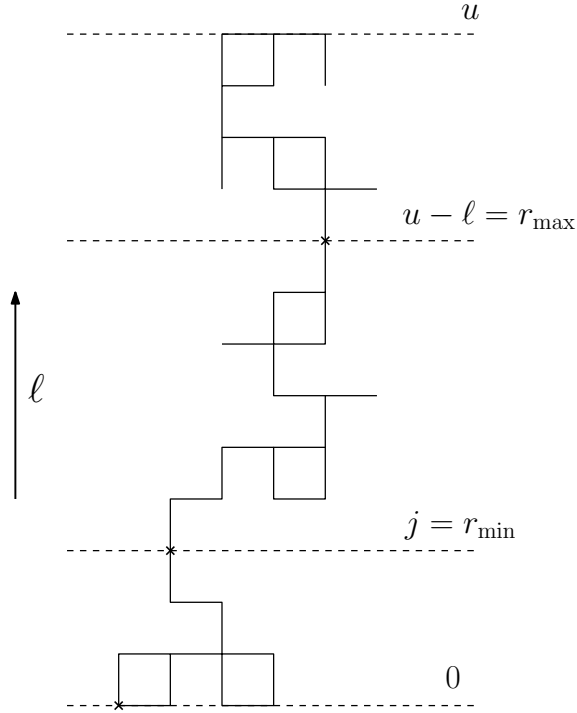
Figure 4: Illustrating Lemma 2.23. A cluster splits into an initial cul-de-sac, a middle string and final inverted cul-de-sac.

Theorem 2.24.

1. *When $\boldsymbol{\ell}$ is lattice,*

$$\mathbb{P}\Big(\mathsf{CulDeSac}_u\,,\,\mathrm{RenLev}(\mathrm{CdS}) = \{u\}\Big) \leq C \exp\big\{-\zeta u - cu^{1/2}\big\}$$

for $u \in \mathrm{Cover}(\boldsymbol{\ell})$. When $\boldsymbol{\ell}$ is non-lattice, this bound is valid for $u \in [0,\infty)$ when the left-hand expression is taken to be $\sum \mathbb{P}\Big(\mathsf{CulDeSac}_h\,,\,\mathrm{RenLev}(\mathrm{CdS}) = \{h\}\Big)$ where here the sum is performed over those $h \in [u, u+r)$ for which the summand is non-zero.

2. *A counterpart statement holds for inverted culs-de-sac. Indeed, the above assertion is equally valid after the changes $\mathsf{CulDeSac} \to \mathsf{InvCulDeSac}$ and $\mathrm{CdS} \to \mathrm{ICdS}$ are made.*

Proof: (1). In the case that $\boldsymbol{\ell}$ is lattice, we must show that there exist positive constants C and c such that, for $u \in (0,\infty)$,

$$\mathbb{P}\Big(\mathsf{CulDeSac}_u, R(\mathrm{CdS}_u) = \{\mathrm{head}(\mathrm{CdS}_u)\}\Big) \leq C \exp\big\{-\zeta u - cu^{1/2}\big\}.$$

To the end of doing so, we first claim that

$$\mathsf{CulDeSac}_u \cap \big\{R(\mathrm{CdS}_u) = \{\mathrm{head}(\mathrm{CdS}_u)\}\big\} \cap \mathsf{Q} \subseteq \mathsf{Cluster}_u \cap \{\mathcal{C} = \mathrm{CdS}_u\}, \tag{9}$$

where Q is the event that every element of U_u that is incident to $\mathrm{head}(\mathrm{CdS}_u)$ is closed under the percolation, and $\mathcal{C}$ is the open cluster containing 0. The inclusion is evident enough: an open path

22

from 0 in the cul-de-sac must leave L_u via the entry x, but it has no ongoing route from there if Q occurs.

If the edge running out of $\mathrm{head}(\mathrm{CdS}_u)$ in the direction e_i lies in $E(U_e)$, then the edge running out of this vertex in the direction $-e_i$ does not lie in this set. Thus, the number of elements of $E(U_e)$ that are incident to $\mathrm{head}(\mathrm{CdS}_u)$ is at most d. We see then that

$$\mathbb{P}\Big(\mathsf{Q}\,\Big|\,\mathsf{CulDeSac}_u \cap \big\{R(\mathrm{CdS}_u) = \{\mathrm{head}(\mathrm{CdS}_u)\}\big\}\Big) \geq (1-p)^d. \tag{10}$$

As a result of this and (9), we learn that

$$(1-p)^d\mathbb{P}\Big(\mathsf{CulDeSac}_u, R(\mathrm{CdS}_u) = \big\{\mathrm{head}(\mathrm{CdS}_u)\big\}\Big) \leq \mathbb{P}\Big(\mathsf{Cluster}_u(0), R(\mathcal{C}) = \big\{\mathrm{head}(\mathcal{C})\big\}\Big). \tag{11}$$

Suppose that $\mathsf{Cluster}_u(0)$ and $R(\mathcal{C}) = \{\mathrm{head}(\mathcal{C})\}$ occur. Consider a resampling of the edges incident to $-e_d$. Suppose this resampling results in $\{-e_d, 0\}$ being open, and the remaining edges being closed. And suppose further that we translate the resulting configuration by e_d. Then in the new configuration the event Str_{u+r} will occur with $R(S_{u+r}) = \{0, \mathrm{head}(S_{u+r})\}$. Thus,

$$p(1-p)^{2d-1}\mathbb{P}\Big(\mathsf{Cluster}_u(0), R(\mathcal{C}) = \big\{\mathrm{head}(\mathcal{C})\big\}\Big) \leq \mathbb{P}\Big(\mathsf{Str}_{u+r} \cap \big\{\mathrm{RenLev}(S_{u+r}) = \{0, u+r\}\big\}\Big). \tag{12}$$

The right-hand probability is at most $C\exp\big\{-\zeta u - cu^{1/2}\big\}$ by Proposition 2.17. What we have learnt is that

$$\mathbb{P}\Big(\mathsf{CulDeSac}_u, R(\mathrm{CdS}_u) = \big\{\mathrm{head}(\mathrm{CdS}_u)\big\}\Big) \leq p^{-1}(1-p)^{1-3d}C\exp\big\{-\zeta u - cu^{1/2}\big\}. \tag{13}$$

This completes the proof of Theorem 2.24(1) in the case that $\boldsymbol{\ell}$ is lattice.

Suppose now that $\boldsymbol{\ell}$ is non-lattice. Replace $u \to h$ in (11), and denote the right-hand side of the resulting bound summed over $h \in [u, u+r)$ by ρ. Analogously to (12),

$$p(1-p)^{2d-1}\rho \leq \mathbb{P}\Big(\mathsf{Str}_{[u,u+r)}\,, R(S_{[u,u+r)}) = \{0, \mathrm{head}(S_{[u,u+r)})\}\Big),$$

where $S_{[u,u+r)}$ denotes the string between 0 and some vertex of height in $[u, u+r)$ whose uniqueness is ensured by Lemma 2.2(1). Since the right-hand probability is at most $C\exp\big\{-\zeta u - cu^{1/2}\big\}$ by Proposition 2.17, assembling the estimates yields that the sum of

$$\mathbb{P}\big(\mathsf{CulDeSac}_h, R(\mathrm{CdS}_h) = \big\{\mathrm{head}(\mathrm{CdS}_h)\big\}\big)$$

over those $h \in [u, u+r)$ for which the summand is non-zero is bounded above by the right-hand expression in (13). Thus we obtain Theorem 2.24(1) in the non-lattice case.

(2). Write $-e = (-x, -y)$ and $e \oplus u = (x+u, y+u)$ for $e = (x, y) \in E(\mathbb{Z}^d)$ and $u \in \mathbb{Z}^d$. For $x \in \mathbb{Z}^d$, define a bijection $\omega \to \omega'$ on configurations by $\omega'(e) = \omega(-e \oplus x)$. Let $x \in U_0$. An inverted cul-de-sac with entry $-x$ occurs under ω' precisely when a cul-de-sac with entry x occurs under ω and any edge incident to x both of whose endpoints has height $x \cdot \boldsymbol{\ell}$ is ω-closed. (This ω-closed condition arises because the strict and non-strict inequalities in the definitions of forward and backward half-spaces in Section 1.2 change places under reflection through the origin.) The ratio of probabilities for an inverted and a standard cul-de-sac thus takes the form $(1-p)^\kappa$ for a $\boldsymbol{\ell}$-determined value of κ lying in $[\![0, 2(d-1)]\!]$. In this way, Theorem 2.24(2) is reduced to the preceding part. $\qquad\square$

Now we harness the just proved result to learn that the culs-de-sac that bookend a typical lengthy cluster are rather short.

Lemma 2.25.

1. *Suppose that $\boldsymbol{\ell}$ is lattice with span s. There exist positive C and c such that the quantities*

$$\alpha_u = \frac{\mathbb{P}\big(\mathsf{CulDeSac}_u, \mathrm{RenLev}(\mathrm{CdS}_u) = \{u\}\big)}{\mathbb{P}\big(\mathsf{Str}_u\big)}$$

   and

$$\beta_u = \frac{\mathbb{P}\big(\mathsf{InvCulDeSac}_u, \mathrm{RenLev}(\mathrm{ICdS}_u) = \{0\}\big)}{\mathbb{P}\big(\mathsf{Str}_u\big)}$$

   satisfy

$$\max\{\alpha_u, \beta_u\} \le C \exp\big\{ -cu^{1/2} \big\}$$

   whenever $u \in s \cdot \mathbb{N}$.

2. *Now suppose that $\boldsymbol{\ell} \in S^{d-1}$ is general, lattice or non-lattice. Write*

$$\alpha_{[u,u+r)} = \frac{\sum \mathbb{P}\big(\mathsf{CulDeSac}_h, \mathrm{RenLev}(\mathrm{CdS}_h) = \{h\}\big)}{\mathbb{P}\big(\mathsf{Str}_{[u,u+r)}\big)}$$

   and

$$\beta_{[u,u+r)} = \frac{\sum \mathbb{P}\big(\mathsf{InvCulDeSac}_h, \mathrm{RenLev}(\mathrm{ICdS}_h) = \{0\}\big)}{\mathbb{P}\big(\mathsf{Str}_{[u,u+r)}\big)}$$

   where the sums are taken over those $h \in [u, u+r)$ for which the summand is non-zero. Then $\max\{\alpha_{[u,u+r)}, \beta_{[u,u+r)}\} \le C \exp\big\{ -cu^{1/2} \big\}$ for $u \in (0, \infty)$.

(1). Theorem 1.8 and the lattice aspect of Theorem 2.24(1) imply the bound on α_u, while the bound on β_u arises when we instead invoke Theorem 2.24(2).

(2). Note that in the non-lattice case Theorem 1.8 implies that $\mathbb{P}\big(\mathsf{Str}_{[u,u+r)}\big) \sim \kappa_{\mathrm{string}}(1 - e^{-\zeta r})e^{-\zeta u}$. The sought result thus follows from the just given argument where we instead invoke the non-lattice assertions made in the two parts of Theorem 2.24. $\qquad\square$

We are ready to give the promised derivation.

Proof of Theorem 1.4. Suppose first that $\boldsymbol{\ell}$ is lattice, with span $s \in (0, \infty)$. Set $\gamma_i = \exp\{\zeta i\}\mathbb{P}\big(\mathsf{Str}_i\big)$ for $i \in s \cdot \mathbb{N}$. Theorem 1.8 implies the existence of $\gamma_\infty \in (0, 1)$ such that the strong asymptotic $\gamma_i \overset{\boldsymbol{\ell}}{\approx} \gamma_\infty$ holds (as $i \to \infty$) in the sense of Definition 1.2.

For $u \in s \cdot \mathbb{N}$, and in the notation of Lemma 2.25, set $\phi_u = \sum_{j,\ell \in [0,u]} \alpha_j \gamma_j \beta_\ell \gamma_\ell \gamma_{u-j-\ell} \mathbf{1}_{j+\ell \le u}$ for $u \in s \cdot \mathbb{N}$; here, and throughout the proof, summations over $s \cdot \mathbb{N}$ are understood. Also write $\phi_\infty = \gamma_\infty \sum_{j \ge 0} \alpha_j \gamma_j \sum_{\ell \ge 0} \beta_\ell \gamma_\ell$.

We claim that

$$\mathbb{P}\big(\mathsf{Cluster}_u, \mathrm{RenLev}(C_u) \ne \emptyset\big) = \phi_u. \tag{14}$$

Indeed, this follows from

$$\mathbb{P}\big(\mathsf{Cluster}_u, \mathrm{RenLev}(C_u) \neq \emptyset\big)$$

$$= \sum_{j,\ell \in [0,u]} \mathbb{P}\Big(\mathsf{Cluster}_u\,,\, r_{\min} = j\,,\, r_{\max} = u - \ell\Big)\mathbf{1}_{j+\ell \leq u}$$

$$= \sum_{j,\ell \in [0,u]} \mathbb{P}\big(\mathsf{CulDeSac}_j, \mathrm{RenLev}(\mathrm{CdS}_j) = \{j\}\big)\mathbb{P}\big(\mathsf{Str}_{u-j-\ell}\big)$$

$$\times\, \mathbb{P}\big(\mathsf{InvCulDeSac}_\ell, \mathrm{RenLev}(\mathrm{ICdS}_\ell) = \{0\}\big)\mathbf{1}_{j+\ell \leq u}$$

$$= \sum_{j,\ell \in [0,u]} \alpha_j \beta_\ell \mathbb{P}\big(\mathsf{Str}_j\big)\mathbb{P}\big(\mathsf{Str}_{u-j-\ell}\big)\mathbb{P}\big(\mathsf{Str}_\ell\big)\mathbf{1}_{j+\ell \leq u}\,,$$

where the second equality is due to Lemma 2.23.

We next argue in favour of the strong asymptotic relation $\phi_u \stackrel{\boldsymbol{\ell}}{\approx} \phi_\infty$. To derive this, note that $\phi_\infty - \phi_u$ equals $A_1 + A_2$, where

$$A_1 = \sum_{j,\ell \geq 0} \alpha_j \gamma_j \beta_\ell \gamma_\ell \big(\gamma_\infty - \gamma_{u-j-\ell}\big)\mathbf{1}_{j+\ell \leq u}$$

and

$$A_2 = \gamma_\infty \sum_{j,\ell \geq 0} \alpha_j \gamma_j \beta_\ell \gamma_\ell \mathbf{1}_{j+\ell > u}\,.$$

Note here that A_2, being at most $\gamma_\infty \sum_{j \geq u/2} \alpha_j \gamma_j \sum_{\ell \geq u/2} \beta_\ell \gamma_\ell$, satisfies $A_2 \stackrel{\boldsymbol{\ell}}{\approx} 0$ in view of Lemma 2.25(1). We also have that $|A_1| \leq A_{11} + A_{12} + A_{13}$, where the latter expressions are sums of

$$\alpha_j \gamma_j \beta_\ell \gamma_\ell \cdot \big|\gamma_\infty - \gamma_{u-j-\ell}\big| \cdot \mathbf{1}_{j+\ell \leq u}$$

over the respective index sets $0 \leq j, \ell \leq u/4$; $j \geq u/4$ with $\ell \geq 0$; and $\ell \geq u/4$ with $j \geq 0$. Since $\sup \gamma_i$ is bounded, the relations $\beta_\ell \gamma_\ell \stackrel{\boldsymbol{\ell}}{\approx} 0$ and $\alpha_j \gamma_j \stackrel{\boldsymbol{\ell}}{\approx} 0$ available from $\gamma_i \stackrel{\boldsymbol{\ell}}{\approx} \gamma_\infty$ and Lemma 2.25(1) serve to show that $A_{1i} \stackrel{\boldsymbol{\ell}}{\approx} 0$ for $i \in [\![1,3]\!]$. We may now return to $\phi_\infty - \phi_u = A_1 + A_2$ and find, as we sought to do, that $\phi_u \stackrel{\boldsymbol{\ell}}{\approx} \phi_\infty$. From (14), $\mathbb{P}\big(\mathsf{Cluster}_u, \mathrm{RenLev}(C_u) \neq \emptyset\big)$ is seen to be strongly asymptotic to $\kappa_{\mathrm{clust}} \exp\{-\zeta u\}$, where

$$\kappa_{\mathrm{clust}} = \sum_{j \in s \cdot \mathbb{N}} \alpha_j \gamma_j \sum_{\ell \in s \cdot \mathbb{N}} \beta_\ell \gamma_\ell\,.$$

But

$$\mathbb{P}\big(\mathsf{Cluster}_u\big) \stackrel{\boldsymbol{\ell}}{\approx} \mathbb{P}\big(\mathsf{Cluster}_u, \mathrm{RenLev}(C_u) \neq \emptyset\big) \tag{15}$$

by Theorem 1.9(2) and $\mathbb{P}\big(\mathsf{Cluster}_u\big) \geq c\mathbb{P}\big(\mathsf{Cluster}_{[u,u+3]}\big)$. In this way, we obtain Theorem 1.4 in the case that $\boldsymbol{\ell}$ is lattice.

When $\boldsymbol{\ell}$ is non-lattice, we seek to show that

$$\mathbb{P}\big(\mathsf{Cluster}_{[u,\infty)}\big) \sim \kappa \exp\big\{-\zeta u\big\}; \tag{16}$$

what has changed is that the event of height at least, rather than exactly, u is considered, and the claimed asymptotic relation is $\sim$ rather than the stronger $\approx$.

Let $\mathsf{CulDeSac}[x]$ denote the event that $\mathsf{CulDeSac}_{x \cdot \boldsymbol{\ell}}$ occurs, with $x \in V\big(\mathrm{CdS}_{x \cdot \boldsymbol{\ell}}\big)$; and similarly set $\mathsf{InvCulDeSac}[x] = \mathrm{ICdS}_{x \cdot \boldsymbol{\ell}} \cap \big\{ x \in V\big(\mathrm{ICdS}_{x \cdot \boldsymbol{\ell}}\big)\big\}$. On $\mathsf{CulDeSac}[x]$, set $\mathrm{CdS}[x] = \mathrm{CdS}_{x \cdot \boldsymbol{\ell}}$ and on $\mathsf{InvCulDeSac}[x]$, set $\mathrm{ICdS}[x] = \mathrm{ICdS}_{x \cdot \boldsymbol{\ell}}$. Write $\mathsf{Str}[x]$ for the event that $\mathsf{Str}_{x \cdot \boldsymbol{\ell}}$ occurs alongside $x \in R(S_x)$. We write

$$\alpha(x) = \frac{\mathbb{P}\big(\mathsf{CulDeSac}[x], R(\mathrm{CdS}[x]) = \{x\}\big)}{\mathbb{P}\big(\mathsf{Str}[x]\big)} \quad , \quad \beta(y) = \frac{\mathbb{P}\big(\mathsf{InvCulDeSac}[y], R(\mathrm{ICdS}[y]) = \{0\}\big)}{\mathbb{P}\big(\mathsf{Str}[y]\big)}$$

and $\gamma(x) = \mathbb{P}(\mathsf{Str}[x]) \exp\{\zeta x \cdot \boldsymbol{\ell}\}$. We now consider

$$\begin{aligned}
&\mathbb{P}\big(\mathsf{Cluster}_{[u,\infty)}, \mathrm{RenLev}(C_u) \neq \emptyset\big) \\
&= \sum_{x,y \in F_0} \mathbb{P}\Big(\mathsf{Cluster}_{[u,\infty)}, \, r_{\min} = x, \, r_{\max} = y\Big) \cdot \mathbf{1}_{(x+y) \cdot \boldsymbol{\ell} \leq u} \\
&= \sum_{x,y \in F_0} \mathbb{P}\big(\mathsf{CulDeSac}[x], R(\mathrm{ICdS}[x]) = \{x\}\big) \mathbb{P}\big(\mathsf{Str}_{[u-(x+y) \cdot \boldsymbol{\ell}, \infty)}\big) \\
&\qquad\qquad\qquad \times \mathbb{P}\big(\mathsf{InvCulDeSac}[y], \mathrm{RenLev}(\mathrm{ICdS}[y]) = \{0\}\big) \cdot \mathbf{1}_{(x+y) \cdot \boldsymbol{\ell} \leq u} \\
&= \sum_{x,y \in F_0} \alpha(x)\beta(y)\mathbb{P}\big(\mathsf{Str}[x]\big)\mathbb{P}\big(\mathsf{Str}_{[u-(x+y) \cdot \boldsymbol{\ell}, \infty)}\big)\mathbb{P}\big(\mathsf{Str}[y]\big) \cdot \mathbf{1}_{(x+y) \cdot \boldsymbol{\ell} \leq u} \, .
\end{aligned}$$

Writing $\psi : (0,\infty) \to (0,\infty)$, $\psi(u) = e^{\zeta u}\mathbb{P}(\mathsf{Str}_{[u,u+r]})$, we find that

$$\mathbb{P}\big(\mathsf{Cluster}_{[u,\infty)}, \mathrm{RenLev}(C_u) \neq \emptyset\big) \exp\{\zeta u\} = \sum_{x,y \in F_0} \alpha(x)\beta(y)\gamma(x)\gamma(y)\psi\big(u - (x+y) \cdot \boldsymbol{\ell}\big)\mathbf{1}_{(x+y) \cdot \boldsymbol{\ell} \leq u} \, .$$

Lemma 2.11(5) implies that $\psi(u) \sim \mu_\infty \zeta^{-1}(1 - e^{-\zeta})$ as $u \to \infty$, where $\mu_\infty = \lim_{u \to \infty} \mu^{\mathbb{N}}[u, u+r)$. We thus obtain the desired asymptotic relation (16) with $\kappa = \kappa_{\mathrm{clust}}$ given by

$$\kappa_{\mathrm{clust}} = \mu_\infty \zeta^{-1}(1 - e^{-\zeta}) \sum_{x \in F_0} \alpha(x)\gamma(x) \sum_{y \in F_0} \beta(y)\gamma(y) \, .$$

Thus we obtain Theorem 1.4 in the non-lattice case.

(2). For $x \in \mathbb{Z}^d$, consider the operation that acts on a percolation configuration by reflecting it through the origin. The operation leads the percolation law invariant, and, neglecting a certain priviso, it maps a cul-de-sac with entry x to an inverted cul-de-sac with entry $-x$.

$\square$

3 Renewal abundance heuristics and basic tools

It remains to prove our key technical contribution on renewal abundance, Theorem 1.9. We will do so in the next section. Here, we explain in a first subsection some of the main ideas and in a second present some basic tools involving coarse-grained percolation needed to implement them.

3.1 Slide into renewal

How to prove that renewal points are abundant in a typical cluster of high diameter? Here we discuss an approach to this question; in so doing, we offer a heuristic overview of the proof of our result, Theorem 1.9(1), on the typical positive density of renewal in a lengthy finite cluster. In this presentation, we take $\boldsymbol{\ell} = e_d$ to be vertical, and consider the typical form of the cluster $\mathcal{C}$ under percolation $\mathbb{P}$ given the event $\mathsf{Cluster}_u$ that a finite cluster based at the origin runs to height equal to $u \in \mathbb{N}$. We will write $\mathcal{C}_u$ for $\mathcal{C}$ to emphasise this conditioning.

The cluster $\mathcal{C}_u$ is a finite open subgraph of vertical extension u that is encased away from the infinite open cluster $\mathcal{I}$ by dual plaquettes through whose midpoints run closed edges in the cluster boundary. In a crude but useful distortion, this dual plaquette surface may be viewed as a subcritical object, whose cardinality has an exponential tail under the unconditioned percolation law $\mathbb{P}$, because the opposite ends of the closed edge running through any constituent plaquette lie inside and outside of the cluster, so that the outside endpoint is at a significant (and possibly even infinite) remove from the inside endpoint in the chemical distance along open edges; and this event is a rarity that may be diagnosed locally.

In Proposition 3.2, we will develop this notion, by using a coarse-grained percolation and invoking results of Antal and Pisztora [AP96]. For now, we take the lesson of this result to be that, under $\mathbb{P}$ given $\mathsf{Cluster}_u$, $\mathcal{C}_u$ typically has volume of the order of the minimum possible, u. The cluster $\mathcal{C}_u$ runs between heights zero and u. Suppose that we slice through the cluster at a typical height h, around level $u/2$. The lesson indicates that the cross-section of $\mathcal{C}_u$ at height h is likely to have a cardinality that is bounded above independently of u. If this cross-section occurs at a unique vertex, that vertex would be a renewal point for the cluster; so in seeking to derive renewal abundance, we wish to argue that this circumstance is typical under the conditioned law. As it stands, though we know that the cross-section is typically bounded in cardinality, it may be high in diameter. The cluster could resemble a capital 'N', as depicted in Figure 5(left). Elements of $E(\mathcal{C}_u)$ are dashed, and have the rough shape of a path that runs upwards from the base, doubling back across level $h \approx u/2$, and then running upwards again to the head. Slicing through at height h decomposes the cluster into three pieces: the downcluster and upcluster, emboldened in the depiction, which are the respective open connected components of the cluster's base and head within the half-spaces delimited by the slice; and a remaining piece, the midcluster, which runs between the two vertices marked with crosses that lie one apiece in the downcluster and the upcluster.

The midcluster is a costly item in the example, because it is sheathed by a subcritical surface of dual plaquettes away from the two surfaces associated to the upcluster and the downcluster. Not only expensive but also, it would seem, needless: the downcluster covers a height of h; the upcluster, one of $u - h$; so together, all the distance needed to realize the event $\mathcal{C}_u$ is covered. How then to gauge the resulting rarity of this picture? It is tempting to shift the upcluster horizontally so that its base coincides with the head of the downcluster (which is to say, so that the two crosses coincide). Figure 5(right) shows the shifted picture. There is no expensive surface enclosing a midcluster and so the right picture seems of the two to be much more probable. A probability resampling experiment that retains enough information to quantity this intuition works as follows. Sample the law $\mathbb{P}$ given $\mathsf{Cluster}_u$. Record the downcluster; and also record the upcluster, but only up to a horizontal shift, so that only the equivalence class of the upcluster under identification by translation in $\mathbb{Z}^d$ is retained. Then work to reconstruct the law $\mathbb{P}$ given $\mathsf{Cluster}_u$ and also given the
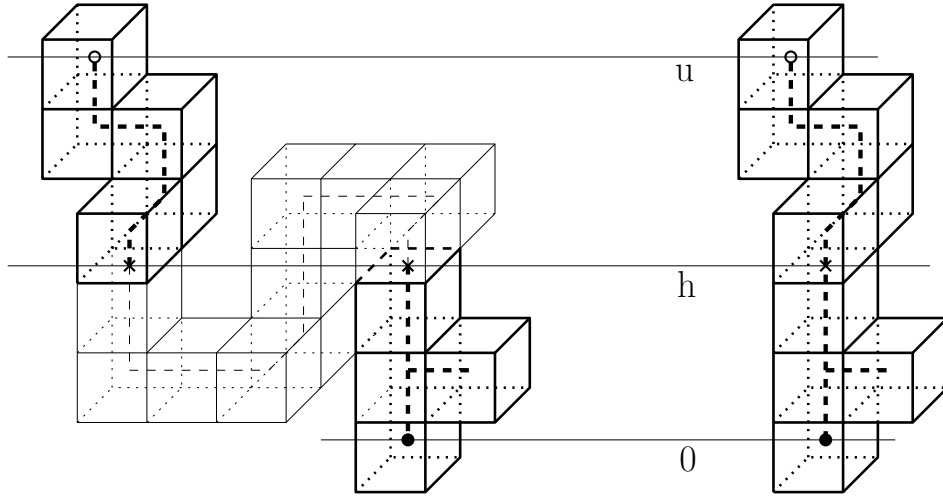
Figure 5: A slide eliminates from a capital N its midcluster.

two pieces of retained data. In the depicted case, this reconstruction job involves two tetris-like pieces, the lower one held in place, and the upper one resting on the level-h hyperplane, and capable of gliding over it, with any given spot being chosen with what we may call the *connection chance*, namely with a weight proportional to the probability that, were the rest of space to be completed according to the ordinary percolation law, a suitable (and perhaps trivial) finite midcluster would form to connect the two pieces and reconstitute a copy of the cluster $\mathcal{C}_u$.

When the upcluster is shifted so that the two crosses coincide, there is unit connection chance; whereas the connection chance for distant displacements may be expected to be of rapid decay.

These ideas may show some promise, but there are objections. One problem is that, if the slice height is such that the downcluster has more than one vertex at the height, no horizontal shift of the upcluster will lead to a renewal point at that height. To overcome this difficulty, we may simply cut the downcluster off from above at height h, while cutting the upcluster from below at height $h+2$; then, if we shift the upcluster so that its base hovers two units above the head of the downcluster, a resampling in which the two intervening vertical edges are declared open will, alongside a suitable set of other edges being declared closed, result in a resampled cluster $\mathcal{C}_u$ that has a renewal point at height at $h+1$; we say that a *simple join* occurs when the resampling acts in this way.

A second difficulty is represented by the 'collapsed capital N': the midcluster could be lengthy, but if its surface is shared with those that enclose the downcluster and the upcluster, it is hard to argue that the connection chance is small, since the formation of the midcluster surface is not an independent event.

Figure 6 shows a slide resample with a gap of two at the slide height. The illustration is schematic, depicting a two-dimensional example for simplicity. Simple join would occur were the upcluster on the left resampled by a shift to the left by one unit, with two open vertical midcluster edges running down from the resulting upcluster base, and suitable neighbouring edges being closed. The resampling on the right shows a dashed midcluster whose surface (which is not depicted) shares several facets with that of the downcluster. That the downcluster has several routes up to around level h broadens the ways in which the midcluster and downcluster surfaces may coincide
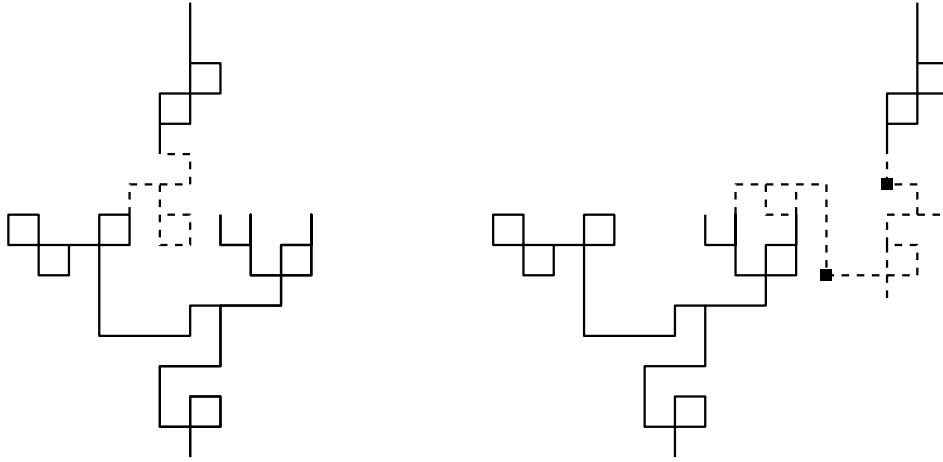
Figure 6: The slide resample pushes the upcluster some way to the right in this depiction. The downcluster has some presence near the slide height and this permits the midcluster in the right sketch to travel for a while around the downcluster, but the slide is big enough to force the midcluster to make a journey in virgin territory between the two square vertices. The disjoint surface that encloses this part of the midcluster reflects the low connection chance (and the high connection cost!) of the proposed slide location.

in part. Note however between the two square vertices runs a midcluster section whose surface is disjoint from those of the upcluster and the downcluster. In general, we may think of drawing a path in the midcluster that connects the downcluster head to the upcluster base and selecting the subpath—call it Connect!—from the last visit to a neighbour of the downcluster until the first visit a neighbour of the upcluster. Connect is plausibly enclosed by a disjoint surface and roughly speaking the connection chance for a given proposed upcluster shift is of the order of $e^{-c\kappa}$ where κ is the minimum length of Connect as we vary over viable choices of midcluster given the shift.

If the downcluster and upcluster cardinalities are bounded above suitably in a neighbourhood of the slice, then the number of shift locations for which the κ-value is bounded above is limited. The κ-value grows linearly in the distance away from such special sites. This means that a given shift that may lead to simple join and a renewal point (such as one unit to the left in Figure 6(left)) has a connection chance that is bounded below, while the sum over all shifts of the connection chance is bounded above, with a bounded population of sites making a unit-order contribution and others making one that is exponentially small as a function of distance between downcluster and shifted upcluster.

Our plan is to present the fairly weak (and in essence already known) estimate that we have mentioned, namely Proposition 3.2, to the effect that cluster volume under $\mathbb{P}$ given $\mathsf{Cluster}_u$ is typically linear in u, and to harness this to argue that downcluster and upcluster cardinalities at heights h' close to the slide height h typically grow at a suitably low rate as $|h' - h|$ rises. Then analysing a partition function given by the sum of connection chances will serve to show that the slide resample results with positive probability in renewal near the slide height. Since the slide resample leaves the conditioned law at equilibrium, we will thus prove Theorem 1.9(1), our assertion of positive density of renewal.

3.2 Coarse-grained percolation

This subsection is principally devoted to stating and proving Proposition 3.2 on the linear volume of the enclosing surface of a lengthy finite cluster. We use a coarse-graining of percolation to present the result, dividing $\mathbb{Z}^d$ into boxes.

Two distinct vertices $v, w \in \mathbb{Z}^d$ are *-neighbours* if $w_i - v_i \in \{-1, 0, 1\}$ for $i \in [\![1, d]\!]$. A *-path is a $\mathbb{Z}^d$-valued path whose pairs of consecutive elements are *-neighbours. The *-distance of $v, w \in \mathbb{Z}^d$ equals $\ell - 1$, where ℓ is the number of elements in the shortest *-path that connects v and w.

Let $K \in \mathbb{N}$ be a given *even* integer. For $v \in 2K \cdot \mathbb{Z}^d$, the K-box $B(v)$ with central vertex v is $v + [\![-K, K]\!]^d$. For A a subgraph of $(\mathbb{Z}^d, \sim)$, $\overline{A}$ will denote the set of K-boxes that contain an edge in A.

The set $\{B(v) : v \in \mathbb{Z}^d\}$ of K-boxes is naturally indexed by the central vertex $v \in \mathbb{Z}^d$, so that the notion of *-neighbour may be applied to such boxes just as it is to elements of $\mathbb{Z}^d$.

Recall from Section 1.2 that a path makes nearest neighbour steps.

Definition 3.1. Let $\partial_{\text{ext}}\overline{\mathcal{C}}$ denote the set of K boxes B such that B is a *-neighbour (in the sense of K-boxes) of an element of $\overline{\mathcal{C}}$ and there exists an infinite path Q in K boxes emanating from B that is disjoint from $\overline{\mathcal{C}}$.

Here is the promised result.

Proposition 3.2. *For $d \geq 2$, there exist positive constants C, c and D_0 such that, when $D \geq D_0$,*

$$\mathbb{P}\left(\left|\partial_{\text{ext}}\overline{\mathcal{C}}\right| \geq Du \,\middle|\, \mathsf{Cluster}_{[u,u+1]}\right) \leq C \exp\left\{-cDu\right\}.$$

3.2.1 Box percolation

As tools we now present will serve to show, the above proposition will be derived by arguing that the K-box elements of $\partial_{\text{ext}}\overline{\mathcal{C}}$ are unusual in a way that can be locally diagnosed.

The *boxlet* boxlet(v) with central vertex v is $v + [\![-K/2, K/2]\!]^d$ and the *big box* BigBox(v) with this central vertex is $v + [\![-5K, 5K]\!]^d$. We will refer to the K-box $B(v)$ as a function of the big box $D = \text{BigBox}(v)$ in the form Centre $= \text{Centre}(D)$.

For G a subgraph of $(\mathbb{Z}^d, \sim)$, we write $\partial G = \{v \in V(G) : \exists\, w \in \mathbb{Z}^d \setminus V(G),\, v \sim w\}$.

Let $D = 2Kv + [\![-5K, 5K]\!]^d$, $v \in \mathbb{Z}^d$, denote a big box. The K-box $B(v)$ will be called *bad* if at least one of two conditions is satisfied:

1. there exists no open path from boxlet(v) to $\partial B(v)$—we say that 'boxlet(v) is isolated'; or

2. there exists an open path P from boxlet(v) to $\partial B(v)$ (so that P lies in $B(v)$ and thus also in BigBox(v)), and an open path Q, also contained in BigBox(v), of Euclidean diameter at least $K/2$ that starts in a K-box that *-neighbours $B(v)$, such that P and Q lie in different open connected components of the percolation in BigBox(v).

When $\mathsf{Cluster}_u$ occurs, we write $\mathcal{C}_u$ for the cluster $\mathcal{C}$ containing 0, for emphasis.

Lemma 3.3. *Let $u \in (0, \infty)$ and $K \in \mathbb{N}$ satisfy $u \geq K > 0$. Suppose $\mathsf{Cluster}_u$ occurs. Let $B = B(v)$ be a K-box (with $v \in 2K \cdot \mathbb{Z}^d$) and let D be a K-box $*$-neighbour of B. If $B \cap \mathcal{C}_u = \emptyset$ and $D \cap \mathcal{C}_u \neq \emptyset$, then B is bad.*

Proof. We will show that if the first condition that would qualify B as bad is not met, then the second one is. Suppose then that boxlet(v) is not isolated, and let P denote an open path from this set to ∂B. Let x be a vertex in both $\mathcal{C}_u$ and D. The diameter of $\mathcal{C}_u$, being at least u, is at least K, so that $\max \{ d(x, v) : v \in V(\mathcal{C}_u) \} \geq K/2$ (where d is Euclidean distance). We may thus find an open path Q in $\mathcal{C}_u$ of diameter between $K/2$ and $K/2 + 1$ that emanates from x. We see that the path Q lies in BigBox(v), because x lies in this big box at distance at least $2K$ from its boundary. The path P lies in B, so that P is vertex disjoint from $\mathcal{C}_u$. Since $\mathcal{C}_u$ is an open cluster, there is no open path from $\mathcal{C}_u$ to P; and specifically, none in BigBox(v). We have shown what we sought to show: if boxlet(v) is not isolated, then B is bad via the second condition specifying this notion. This completes the proof of Lemma 3.3. $\qquad\square$

In the next result, we again rely on the indexing of K-boxes by $\mathbb{Z}^d$.

Lemma 3.4. *Define a random process on $\mathbb{Z}^d$ by declaring that $v \in \mathbb{Z}^d$ is bad if the K-box $B(v)$ is bad. For any $\epsilon > 0$, there exists $K \in \mathbb{N}$ such that the process of bad indices—a random subset of $\mathbb{Z}^d$—is stochastically dominated by site percolation of parameter ϵ.*

Proof. A given big box D can be contained in the set $B'_{\mathbf{i}}(N)$ specified in [AP96, Equation (2.7)], where the parameter N in this reference is set to be minimal to permit this containment. It is readily verified that, if Centre(D) is bad, then the complement of the event $R^{(N)}_{\mathbf{i}}$ in [AP96, Equation (2.9)] occurs. It is [AP96, Equation (2.24)] which establishes the rarity of this complementary event. Indeed, this bound shows that any given index in $\mathbb{Z}^d$ is bad with probability at most ϵ, where $\epsilon > 0$ tends to zero in the limit of high K. The process of bad indices is 5-dependent in the sense of [LSS97]. Thus Lemma 3.4 follows from [LSS97, Theorem 0.0(i)]. $\qquad\square$

We indicated in Section 3.1 how the slide resample will operate with a small gap, and next we present some related generalities of later use.

Definition 3.5. Let A denote a subgraph of $(\mathbb{Z}^d, \sim)$. We denote by $\overline{A}[2]$ the set of K-boxes whose $*$-distance from $\overline{A}$ is most two. And we write $\overline{A}[-2]$ for the set of K-boxes in $\overline{A}$ whose $*$-distance from the complement of $\overline{A}$ is at least three (so that at least *two* elements in any $*$-path in K-boxes between $\overline{A}[-2]$ and the complement of $\overline{A}$ lie outside of $\overline{A}[-2]$).

Lemma 3.6. *Let A be a subgraph of $(\mathbb{Z}^d, \sim)$. Let E denote a percolation event that is measurable with respect to the edges in A. Let Bad denote the random set of bad K-boxes that do not lie in $\overline{A}[2]$. Then the distribution of Bad is equal under the laws $\mathbb{P}$ and $\mathbb{P}(\cdot | \mathsf{E})$.*

Proof. The value of the random process Bad is determined by the open/closed status of the collection of edges lying in the union $\mathcal{U}$ of the big boxes whose Centre is a K-box that does not lie in $\overline{A}[2]$. Although $\mathcal{U}$ may contain edges that lie in the boundary of a K-box in $\overline{A}$, such edges do not belong to A, because, if they did, an adjacent K-box would also lie in $\overline{A}$ at $*$-distance two from Centre of a big box that lies in the union $\mathcal{U}$. Thus the process Bad is unperturbed by conditioning the law $\mathbb{P}$ on the status of edges in A. $\qquad\square$

3.2.2 Bounding the cardinality of the cluster boundary

Here we prove Proposition 3.2.

When $\mathsf{Cluster}_u$ occurs, we write $\overline{\mathcal{C}}_u$ in place of $\overline{\mathcal{C}_u}$.

Lemma 3.7. *When $\mathsf{Cluster}_u$ occurs, every element of $\partial_{\mathrm{ext}}\overline{\mathcal{C}}_u$ is a bad K-box.*

Proof. Let $B \in \partial_{\mathrm{ext}}\overline{\mathcal{C}}_u$. Let D be a K-box $*$-neighbour of B that lies in $\overline{\mathcal{C}}_u$. Lemma 3.3 then implies that B is bad. $\qquad\square$

Lemma 3.8. *For $d \geq 2$, there exist positive constants C and c such that, for $u \in (0,\infty)$ and $D \geq 1$,*

$$\mathbb{P}\Big(\mathsf{Cluster}_{[u,u+1]}, \big|\partial_{\mathrm{ext}}\overline{\mathcal{C}}\big| \geq Du\Big) \leq C \exp\big\{-cK^{-d}Du\big\}.$$

Proof. Suppose the occurrence of $\mathsf{Cluster}_{[u,u+1]}$ and that $\big|\partial_{\mathrm{ext}}\overline{\mathcal{C}}\big| = m$ for a given m whose value is at least Du. Since $0 \in V(\mathcal{C})$, we see that $\partial_{\mathrm{ext}}\overline{\mathcal{C}}$ contains a K-box whose $\mathbb{Z}^d$-index has Euclidean norm at most m/K.

By Lemmas 3.4 and 3.7, we see that $\mathsf{Cluster}_{[u,u+1]} \cap \big\{\big|\partial_{\mathrm{ext}}\overline{\mathcal{C}}\big| = m\big\}$ entails that a site percolation of parameter ϵ in $\mathbb{Z}^d$ contains a connected component of size at least $m(2K+1)^{-d}$ to which a vertex belongs whose distance from the origin is at most m. Here, the parameter $\epsilon > 0$ is a function of $K \in \mathbb{N}$ that may be chosen to be arbitrarily small by a high enough choice of K. The probability of the just mentioned eventuality is at most $Cm^d \exp\big\{-cmK^{-d}\big\}$ for suitable positive constants C and c in view of the exponential decay of the tail of the size of the connected component containing a given vertex in subcritical site percolation [AB87]. Summing this bound over $m \geq Du$, and using $D \geq 1$ to absorb the factor of m^d, we obtain Lemma 3.8. $\qquad\square$

Proof of Proposition 3.2. That there exists $c > 0$ such that $\mathbb{P}(\mathsf{Cluster}_{[u,u+1]}) > c^u$ is demonstrated by considering an explicit cluster, such as a concatenation of $\lceil u \cdot (e_d \cdot \boldsymbol{\ell})^{-1}\rceil$ e_d-directed edges rising from 0. Proposition 3.2 then follows from Lemma 3.8 by making a suitably high choice for the lower bound D_0 on D and by absorbing the term K^{-d}, which is constant for our given value of K, into the constant c in the argument of the exponential. Note that the value of $D_0 \in (0,\infty)$ may be chosen independently of $\boldsymbol{\ell} \in S^{d-1}$ since $e_d \cdot \boldsymbol{\ell} \geq d^{-1/2}$ in view of the axes ordering convention (1). $\qquad\square$

3.2.3 Latitude and isoperimetry for K boxes

It is useful to know that when an object such as the downcluster is large in a slice, the enclosing surface is also large there. Here we state and prove Proposition 3.10 to this effect, via Lemma 3.11.

Definition 3.9. Recall that a K-box takes the form $x + [\![-K,K]\!]^d$ with $x \in 2K \cdot \mathbb{Z}^d$. The latitude of B equals $\big\lfloor \frac{x \cdot \boldsymbol{\ell}}{2K e_d \cdot \boldsymbol{\ell}}\big\rfloor$. Latitude measures K-box location relative to the hyperplane $z \cdot \boldsymbol{\ell} = 0$ in integer units.

For $k \in \mathbb{Z}$ and $\ell \in \mathbb{N}_+$, let $\overline{N}_{k,k+\ell}$ denote the set of K boxes of latitude in $[\![k, k+\ell]\!]$.

Proposition 3.10. *For some $k \in \mathbb{Z}$, suppose that $\overline{A}$ and $\overline{B}$ are finite subsets of $\overline{N}_{k,k+\ell}$ such that $\overline{B}$ separates $\overline{A}$ from infinity in $\overline{N}_{k,k+\ell}$. Then*

$$|\overline{A}| \leq (\ell+1)|\overline{B}|^{\frac{d-1}{d-2}}.$$

Writing $\mathcal{B}$ for the set of K-boxes, we introduce the projection $P : \mathcal{B} \to \mathbb{Z}^{d-1}$ that sends $2Ku + [\![-K, K]\!]^d$ to $(u_1, \cdots, u_{d-1})$: namely, P sends any K-box to the projection on the first $d-1$ coordinates of its natural index in $\mathbb{Z}^d$.

Lemma 3.11. *Let $\overline{C} \subset \overline{N}_{k,k+\ell}$. Suppose that $x \in \mathbb{Z}^{d-1}$ is not separated from infinity by $P(C)$ in $\mathbb{Z}^{d-1}$. Then nor is any $B \in \overline{N}_{k,k+\ell}$ with $P(B) = x$ separated from infinity by $\overline{C}$ in $\overline{N}_{k,k+\ell}$.*

Proof. Let y and z be nearest neighbours in $\mathbb{Z}^{d-1}$, with $z = y + e_i$ for $i \in [\![1, d-1]\!]$. Let $\mathcal{B}[y]$ denote the set of K-boxes in $\overline{N}_{k,k+\ell}$ that project to y under P. Then $|\mathcal{B}[y]|$ equals $\ell + 1$; label the lowest and highest elements of $\mathcal{B}[y]$ in the form B_y and $B'_y = B_y + 2K\ell e_d$. Similarly for z. We *claim* that some element of $\mathcal{B}[z]$ is obtained from some element of $\mathcal{B}[y]$ by translation by the vector $2Ke_i$. Indeed, if this were not the case, then either B_z would be displaced from $B'_y + 2Ke_i$ by a positive multiple of $2Ke_d$; or B'_z would be displaced from $B_y + 2Ke_i$ by a negative multiple of $2Ke_d$. In the former case, $B'_y + 2Ke_i$ would have latitude less than k; in the latter, $B_y + 2Ke_i$ would have latitude more than $k + \ell$. However, $0 \leq e_i \cdot \boldsymbol{\ell} \leq e_d \cdot \boldsymbol{\ell}$ by the axes ordering convention; and thus the difference in latitude between boxes B and $\hat{B}$ satisfying $\hat{B} = B + 2Ke_i$ is seen to be at most one in absolute value. Since B'_y has latitude $k + \ell$ and B_y has latitude k, we have found a contradiction and proved the claim.

Whenever elements of $\overline{N}_{k,k+\ell}$ respectively project to nearest neighbours $y, z \in \mathbb{Z}^{d-1}$, we may invoke the claim to construct a path of at most $2(\ell + 1)$ K-boxes that connects these elements, with the path valued in $\overline{N}_{k,k+\ell}$ among K-boxes that themselves project to one or other of y and z. This local construction permits us to lift paths from $\mathbb{Z}^{d-1}$ to $\overline{N}_{k,k+\ell}$, and such a lifted path will prove Lemma 3.11. Indeed, since $x \in \mathbb{Z}^{d-1}$ is not separated from infinity by $P(\overline{C})$ in $\mathbb{Z}^{d-1}$, there exists a path Q in $\mathbb{Z}^{d-1}$ from x to infinity that does not visit $P(\overline{C})$. Arbitrarily select a sequence of elements of $\overline{N}_{k,k+\ell}$ that successively project to the consecutive elements of Q. Each consecutive pair of elements in this sequence may be interpolated by a path in K-boxes of length at most $2(\ell+1)$, by the local lift. When the resulting paths are concatenated, the path in K-boxes obtained reaches infinity without any visit to $\overline{C}$. This path validates the conclusion of Lemma 3.11. $\square$

Proof of Proposition 3.10. The first assertion of [Ham06, Lemma 2.3] is a simple consequence of the Loomis-Whitney inequality [LW49]: if $R \subseteq \mathbb{Z}^{d-1}$ denotes the collection of sites that a finite set $W \subseteq \mathbb{Z}^{d-1}$ separates from infinity, then $|R| \leq |W|^{(d-1)/(d-2)}$.

Since $\overline{A}$ is separated from infinity in $\overline{N}_{k,k+\ell}$ by $\overline{B}$, Lemma 3.11 implies that

$$P(\overline{A}) \text{ is separated from infinity in } \mathbb{Z}^{d-1} \text{ by } P(\overline{B}).$$

Taking $R = P(\overline{A})$ in the recalled bound, we find that $|P(\overline{A})| \leq |P(\overline{B})|^{\frac{d-1}{d-2}}$. Using $|\overline{A}| \leq (\ell+1)|P(\overline{A})|$ and $|P(\overline{B})| \leq |\overline{B}|$, we find that $|\overline{A}| \leq (\ell+1)|\overline{B}|^{\frac{d-1}{d-2}}$ and thus obtain Proposition 3.10. $\square$

4 Renewal levels via the slide resample

Here we prove both parts of Theorem 1.9 by slide resampling percolation conditional on the presence of the cluster. As we outlined in the preceding section, examining the slide resample when it operates at a single height is enough to prove Theorem 1.9(1). The resample will be applied iteratively at many heights in order to obtain Theorem 1.9(2).

There are four subsections. Critical to the utility of the slide for finding renewal levels is the *simple join* Proposition 4.11, which shows that, when the slide resample operates at a given height on a cluster with typical geometry, it often results in a resampled cluster with a clean join between the downcluster and the upcluster. In Section 4.1, we introduce the slide resample and demonstrate Theorem 1.9(1) given Proposition 4.11, whose proof is the subject of Section 4.2. As we turn to iterative use of the slide to reach the stronger Theorem 1.9(2), we face the problem that the slide at a given height may fail to produce a simple join. In principle, failure may undo renewal levels far from the present slide height that have been secured at earlier slide iterations. Section 4.3 is devoted to showing that, when the slide fails, damage to the cluster typically occurs only near the slide height. With this support, we introduce the iterated slide and prove Theorem 1.9(2) in Section 4.4.

The proof of Theorem 1.9 may be reprised to give Corollary 1.10, the counterpart result in which the conditioned cluster is supposed to have renewal points at both extreme heights. The proof of Corollary 1.10 is given at the end of Section 4.4, alongside the derivation of Proposition 2.1, a string counterpart to Corollary 1.10(1).

4.1 The single slide: positive density of renewal

As we begin to prove Theorem 1.9(1), we mention that we will develop the ideas of Section 3.1 closely enough, though since our context is now of a general rather than an axially aligned vector $\boldsymbol{\ell}$, certain adjustments are needed. In the axial case, we could speak of proposed slides as horizontal, and indexed clearly by a copy of $\mathbb{Z}^{d-1}$. The vector $\boldsymbol{\ell}$ may however be non-lattice, and we have to accept that there no natural substitute for the lattice $\mathbb{Z}^{d-1}$ as an index of possible shifts. It is inevitable that the height of the cluster head will change, albeit perhaps only by a bounded quantity, under the actions of reasonable shifts. This consideration even finds expression in the form of Theorem 1.9, where we condition on a cluster of height in a short interval $[u, u + 3r]$, rather than of the specific value u; the latter choice is a luxury that is unavailable in the non-lattice case. We will remark at suitable moments on how the proof of Theorem 1.9(1) varies from the template offered in Section 3.1, including in regard to the need to accomodate general values of $\boldsymbol{\ell}$.

Let $u, h \in (0, \infty)$ satisfy $u > h$. We introduce the upspace and the downspace, two half-spaces with boundaries near the slide height h. The *downspace* is simply the lower-half space L_h.

Recall from a specification made in Subsection 1.3.4 and from the axes ordering convention (1) that $r = e_d \cdot \boldsymbol{\ell}$.

The upspace is specified only under the event $\mathsf{Cluster}_{[u,u+3r]}$, and then by means of a little more notation. When this event occurs, there is a cluster $\mathcal{C}$, namely a finite open connected component with base 0, whose height lies in $[u, u + 3r]$; for emphasis, we denote it by $\mathcal{C}_{u,u+3r}$. On the same event, we write
$$h_{\max} = \max\{z \cdot \boldsymbol{\ell} : z \in V(\mathcal{C}_{u,u+3r})\}$$
for the maximum height among vertices in $\mathcal{C}_{u,u+3r}$. In this way, $\mathrm{head}(\mathcal{C}_{u,u+3r})$ is the lexicographically minimal vertex in $\mathcal{C}_{u,u+3r}$ of height $h_{\max}$. The upspace is then defined to be the upper half-space $U_{h_{\mathrm{adj}}}$, where the *adjusted height* h_{adj} denotes $h + (h_{\max} - u)$. When $\mathsf{Cluster}_{[u,u+3r]}$ occurs, $h_{\max} \in [u, u + 3r]$, so that the height $h_{\mathrm{adj}} \in [h, h + 3r]$ is adjusted from h by an addition of at most $3r$.

Also on $\mathsf{Cluster}_{[u,u+3r]}$, we define the upcluster and the downcluster to be two subgraphs of $\mathcal{C}_{u,u+3r}$.

First recall that, on this event, the cluster $\mathcal{C}_{u,u+3r}$ satisfies $\mathrm{base}(\mathcal{C}_{u,u+3r}) = 0$. The *downcluster*, denoted DownClust, is the connected component containing $0 \in \mathbb{Z}^d$ of the subgraph induced by $\mathcal{C}_{u,u+3r}$ in the downspace L_h. To wit, $v \in \mathbb{Z}^d$ lies in the vertex set of the downcluster if there exists an open path in L_h that starts at the origin and that ends at v. The edge set of the downcluster is the union of edges belonging to all such open paths. The top of the downcluster, Top(DownClust), is the set of vertices v in DownClust that lie in $\mathrm{VertBdry}_h$ (as this set is defined in Section 1.2).

The *upcluster*, denoted UpClust, is the open connected component containing $\mathrm{head}(\mathcal{C}_{u,u+3r})$ of the subgraph induced by $\mathcal{C}_{u,u+3r}$ in the upspace. Thus, an element $v \in \mathbb{Z}^d$ lies in the vertex set of the upcluster if and only if an open path in $U_{h_{\mathrm{adj}}}$ runs from $\mathrm{head}(\mathcal{C}_{u,u+3r})$ to v. The edge set of the upcluster is the union of the edges in such paths. Note that $\mathrm{head}(\mathrm{UpClust}) = \mathrm{head}(\mathcal{C}_{u,u+3r})$. The bed of the upcluster, Bed(UpClust), is the set of vertices v in UpClust that lie in $\mathrm{VertBdry}_{h_{\mathrm{adj}}}$.

The definition of the downcluster is largely consistent with the heuristics of Section 3.1, but the upcluster is specified a little differently: the hyperplane that delimits the upspace (and the upcluster) on the lower side is set at a height that is $u - h$ units below that of the upcluster head. This definition is made to face the challenge of setting up the slide resample when ℓ may be non-lattice and some counterpart must be found of the space $\mathbb{Z}^{d-1}$ of admissible shifts of the cluster (or the upcluster) head. By specifying the upcluster relative to its head, we obtain a definition of the upcluster that will be invariant under a suitable class of shifts.

Next, some simple height facts.

Lemma 4.1. *Let H_d and H_u denote the heights of* DownClust *and* UpClust. *Almost surely under $\mathbb{P}$ given* $\mathsf{Cluster}_{[u,u+3r]}$, *$h - r \leq H_d \leq h$ and $u - h \leq H_u \leq u - h + r$.*

Proof. In view of Lemma 2.2(1), we find that $h - r \leq H_d \leq h$ since $V(\mathrm{DownClust})$ consists of vertices in L_h of non-negative height, one of which lies in $\mathrm{VertBdry}_h$. In regard to H_u, note that elements in $V(\mathrm{UpClust})$ have height at most $h_{\max}$ and lie in the upspace $U_{h_{\mathrm{adj}}}$, with at least one lying in $\mathrm{VertBdry}_h$. Any element of the upspace has height at least $h_{\mathrm{adj}} - r$, because, in the opposing case and by the axes ordering convention, the interior of each incident edge would lie in $B_{h_{\mathrm{adj}}}$. And an element that lies in $\mathrm{VertBdry}_h$ has height at most h_{adj}, because, were it to be higher, the interior of each incident edge would intersect $F_{h_{\mathrm{adj}}}$. Thus, we confirm that $u - h \leq H_u \leq u - h + r$. $\qquad \square$

We now present the decomposition of the cluster into three pieces that will allow us shortly to specify the midcluster MidClust.

Definition 4.2. Under $\mathsf{Cluster}_{[u,u+3r]}$, the *residual cluster*, denoted ResidualClust, is in essence the residue of the cluster $\mathcal{C}_{u,u+3r}$ after the removal of the upcluster and the downcluster. Formally, it is the graph whose vertices are endpoints of edges in $\mathcal{C}_{u,u+3r}$ that are edges in neither the upcluster nor the downcluster; the edge-set of ResidualClust is the collection of such edges.

The residual cluster is comprised of possibly several connected components, as we now describe.

Lemma 4.3.

1. *The graph* ResidualClust *contains at least one connected component G such that $V(G)$ intersects* Bed(UpClust) *and* Top(DownClust).

2. *The vertex set of any connected component of* ResidualClust *intersects the union of the vertex sets of* DownClust *and* UpClust. *Any vertex in this intersection lies in the union of*

Top(DownClust) *and* Bed(UpClust).

Proof: (1). Since base$(\mathcal{C}_{u,u+3r})$ and head$(\mathcal{C}_{u,u+3r})$ lie in the connected subgraph $\mathcal{C}_{u,u+3r}$, we may find a path in $\mathcal{C}_{u,u+3r}$ that runs from base$(\mathcal{C}_{u,u+3r})$ to head$(\mathcal{C}_{u,u+3r})$. This path begins in DownClust and ends in UpClust. Consider the subpath between the final vertex in DownClust and the first in UpClust. Since these two vertices lie in VertBdry$_h$, we may set G to be the connected component of ResidualClust that contains this subpath.

(2). Let H be a connected component of ResidualClust. Since the cluster $\mathcal{C}_{u,u+3r}$ is connected, a path in $\mathcal{C}_{u,u+3r}$ runs from any given element of H to base$(\mathcal{C}_{u,u+3r})$. When this path visits H for the final time, it does so at a vertex which also lies in the vertex set of either UpClust or DownClust—in the former case, at an element of Bed(UpClust); in the latter, at an element of Top(DownClust). $\square$

We are ready to select and label as MidClust a connected component of ResidualClust that connects the upcluster and the downcluster. The residual cluster may also contain certain further fragments, which touch either Bed(UpClust) or Top(DownClust) and are not central players in our analysis, that we label MidFragment.

Definition 4.4.

1. Among the components G specified in Lemma 4.3(1), we select one and call it the *midcluster* MidClust. For definiteness, $V(\text{MidClust})$ contains the lexicographically minimal element among vertices in any of the concerned components.

2. The union of connected components of ResidualClust excluding MidClust will be denoted by MidFragment.

For $A \in \mathbb{N}$, we identify three slabs of width $2A$ in the neighbourhood of height h: a north zone, a south zone, and a tropical zone sandwiched between them. (Here, latitude may be identified with height: it acts as an alternative verbal marker for the value of the coordinate in the direction $\boldsymbol{\ell}$.)

Definition 4.5. The subgraphs NorthZone, Tropics and SouthZone are each slabs, respectively set equal to Slab$_{h+A,h+3A}$, Slab$_{h-A,h+A}$ and Slab$_{h-3A,h-A}$.

Definition 4.6. The *upwell configuration space* UCS consists of all values that can be adopted by the subgraph UpClust in a percolation configuration that realizes the event $\mathsf{Cluster}_{[u,u+3r]}$.

We now introduce an equivalence relation on this configuration space by identifying pairs of elements if one element of the pair may be obtained from the other by translation.

Definition 4.7. Two elements U and U' of UCS are *shift equivalent* if there exists $v \in \mathbb{Z}^d$ such that U' is the graph obtained from U by shifting by v (so that the vertex and edge sets of U' are obtained from the counterparts for U by shifting by this vector).

We will write [UpClust] for the shift equivalence class of UpClust. Thus, [UpClust] is a random variable that is defined whenever $\mathsf{Cluster}_{[u,u+3r]}$ occurs.

In Section 3.1, we indicated how the slide resample would be liable to act in a desired fashion if the upcluster and downcluster cardinalities are suitably limited at heights close to the resample height h. We now set down conditions that express these favourable features.

Definition 4.8. Let R, C and L be three positive parameters. Here we introduce several conditions of smallness for the cardinality of $\mathcal{C}_{u,u+3r}$. The conditions receive labels **T**, **S** and **G** according to

whether they concern the size of $\mathcal{C}_{u,u+3r}$ in the *tropics* (with heights in $[h-A, h+A]$); the *subtropics* (the southzone or the northzone); or *globally* (when the total size of $\mathcal{C}_{u,u+3r}$ is considered). Here are the conditions.

$\mathbf{T}_1(R)$: for $i \in [\![1, A]\!]$, the number of vertices in $\mathcal{C}_{u,u+3r}$ with height in $[h+i-1, h+i]$ is at most $i^3 R$.

$\mathbf{T}_2(R)$: for $i \in [\![1, A]\!]$, the number of vertices in $\mathcal{C}_{u,u+3r}$ with height in $[h-i, h-i+1]$ is at most $i^3 R$.

$\mathbf{S}_1(C)$: the number of vertices in $\mathcal{C}_{u,u+3r}$ with height in $(i-1, i]$ is at most C for at least one-third of indices i in $[\![h+1+A, h+3A]\!]$.

$\mathbf{S}_2(C)$: the number of vertices in $\mathcal{C}_{u,u+3r}$ with height in $(i-1, i]$ is at most C for at least one-third of indices i in $[\![h+1-3A, h-A]\!]$.

$\mathbf{G}(L)$: $|V(\mathcal{C}_{u,u+3r})| \le Lu^d$.

We further consider three adjusted conditions, marked with a prime.

$\mathbf{T}'_1(R)$: for $i \in [\![1, A]\!]$, the number of vertices in UpClust whose height lies in $[h_{\mathrm{adj}}+i-1, h_{\mathrm{adj}}+i]$ is at most $i^3 R$.

$\mathbf{T}'_2(R)$: for $i \in [\![1, A]\!]$, the number of vertices in DownClust with height in $[h-i, h-i+1]$ is at most $i^3 R$.

$\mathbf{S}'_1(C)$: the number of vertices in $\mathcal{C}_{u,u+3r}$ with height in $(i-1, i]$ is at most C for at least one-quarter of indices i in $\mathbb{Z} \cap [h_{\mathrm{adj}}+1+A, h_{\mathrm{adj}}+3A]$.

Notably, the north tropical and subtropical conditions $\mathbf{T}'_1(R)$ and $\mathbf{S}'_1(C)$ are expressed in terms of the adjusted equatorial height h_{adj} rather than the ordinary equatorial height h. Equivalent elements of UCS are equal when viewed in these coordinates: this makes the coordinates useful, because it means that the primed conditions will be automatically maintained in the slide resample. In other regards, the changes of the primed events relative to the unprimed counterparts are weakenings of conditions: the replacement of $\mathcal{C}_{u,u+3r}$ by its subgraphs UpClust and DownClust; and of one-third by one-quarter.

(The global condition $|V(\mathcal{C}_{u,u+3r})| \le Lu^d$ may seem absurdly weak, with $|V(\mathcal{C}_{u,u+3r})| \le Lu$ being more natural. The weak form is adequate for our purpose and avoids some technical encumbrances.)

In terms of positive parameters K, R and L, we further define

two subevents $\mathsf{VeryFine}_h = \mathsf{VeryFine}_h(K, R, L)$ and $\mathsf{Fine}_h = \mathsf{Fine}_h(K, R, L)$ of $\mathsf{Cluster}_{[u,u+3r]}$,

using primed (as well as some unprimed) conditions in the latter case.

Indeed (as Figure 7 illustrates), $\mathsf{VeryFine}_h$ occurs when the following conditions are met:

$$\mathbf{T}_1(2^{-1}3^{-3}R), \ \mathbf{T}_2(2^{-1}3^{-3}R), \ \mathbf{S}_1(R), \ \mathbf{S}_2(R) \text{ and } \mathbf{G}(L).$$

And Fine_h occurs when it is these conditions that are met:

$$\mathbf{T}'_1(R), \ \mathbf{T}'_2(R), \ \mathbf{S}'_1(R), \ \mathbf{S}_2(R) \text{ and } \mathbf{G}(L).$$

Lemma 4.9. *When $A \ge 36$, we have that $\mathsf{VeryFine}_h \subset \mathsf{Fine}_h$.*
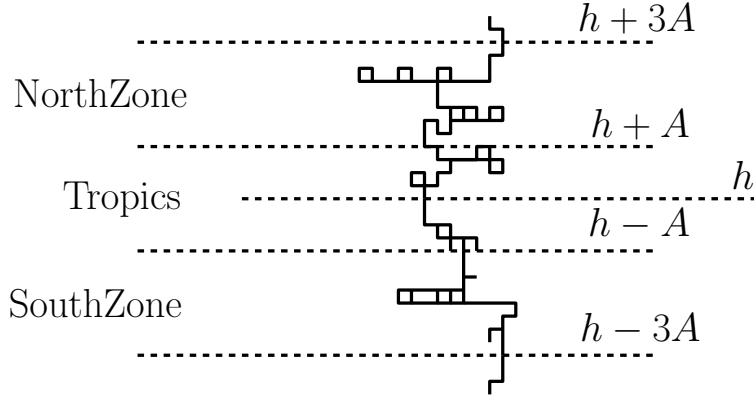
Figure 7: A schematic instance of the event $\mathsf{VeryFine}_h$ viewed in the region around the tropics. The cardinality of the intersection of the depicted cluster $\mathcal{C}_{u,u+3r}$ with horizontal lines in Tropics is bounded above quadratically in the distance from the equatorial line at height h, while a positive density of such cardinalities must be bounded above in the neighbouring subtropical regions.

Proof. In regard to the tropical conditions, note that the number of vertices in UpClust with height in $[h_{\mathrm{adj}}+i-1, h_{\mathrm{adj}}+i]$ is at most the number of vertices in $\mathcal{C}_{u,u+3r}$ with height in $[h+\kappa-2, h+\kappa]$, where $\kappa = \lfloor h_{\mathrm{adj}} - h \rfloor \in [\![0,2]\!]$. If $\mathsf{VeryFine}_h$ occurs, then this upper bound is at most $(2^3+3^3)2^{-1}3^{-3}R \leq R$. So conditions $\mathbf{T}'_1(R)$ and $\mathbf{T}'_2(R)$ of Fine_h have been verified. In regard to the northzone condition, note that the symmetric difference between the index sets $[\![h+A, h+3A]\!]$ and $\mathbb{Z} \cap [h_{\mathrm{adj}}+A, h_{\mathrm{adj}}+3A]$ is at most six. So if a set of integers occupies a proportion of at least one-third of the first of these intervals, it will occupy at least one-quarter of the second, provided that $A \geq 36$. Thus $\mathbf{S}_1(R)$ implies $\mathbf{S}'_1(R)$. Since other conditions are either shared between the events Fine_h and $\mathsf{VeryFine}_h$, or are weaker in the former case, we obtain Lemma 4.9. $\qquad\square$

We will obtain Theorem 1.9(1) by showing that a typical cluster has $\Theta(u)$ levels h for which $\mathsf{VeryFine}_h$ occurs; and that, no matter the form of $([\mathrm{UpClust}], \mathrm{DownClust})$, the thus typical Fine_h often entails a simple join event, in which UpClust sits directly on top of DownClust, ensuring renewal at height h. The next two propositions, whose proofs we defer, show $\Theta(u)$ cardinality of very fine levels and that fine levels often present simple joins.

Proposition 4.10. *Let N denote the number of $h \in \mathbb{N} \cap [0,u]$ for which $\mathsf{VeryFine}_h$ occurs. With the relation $L = 2^{\frac{d-2}{d-1}} R K^{-\frac{2d-3}{d-1}}$ set between the parameters in the definition of the event $\mathsf{VeryFine}_h$, there exist positive constants C and c such that*

$$\mathbb{P}\Big(N < (1-a)u \,\Big|\, \mathsf{Cluster}_{[u,u+3r]}\Big) \leq C \exp\Big\{-cR^{\frac{d-2}{d-1}}K^{-1}a\Big\}.$$

Recall that base(UpClust) and head(DownClust) are certain elements in the subgraphs Bed(UpClust) and Top(DownClust). For the purpose of the ensuing definition, take $v = \mathrm{head}(\mathrm{DownClust}) + e_d$.

The event $\mathsf{SimpleJoin}$ occurs when $\mathsf{Cluster}_{[u,u+3r]}$ does alongside the conditions that

1. MidClust has merely three vertices and two edges, the pair of edges having the form $[v, v+e_d]$ and $[v-e_d, v]$; and

2. the graph MidFragment is empty.

On SimpleJoin, the vertex $v = \text{head}(\text{DownClust}) + e_d$ will be called the *midvertex*.

We write $\mathbb{P}^{\mathsf{Fine}_h} = \mathbb{P}\big(\cdot \,\big|\, \mathsf{Fine}_h \big)$. And we further

denote by $\mathbb{P}^{\mathsf{Fine}_h}_{[\text{UpClust}],\text{DownClust}}$ the law $\mathbb{P}^{\mathsf{Fine}_h}$ conditionally on $([\text{UpClust}], \text{DownClust})$. (17)

We have yet to define the slide resample, but it is perhaps helpful to use it to informally locate the above notation. Indeed, the slide resample at height h concerns the just denoted law, in which conditioning occurs not only on the event Fine_h but also on the random variable $([\text{UpClust}], \text{DownClust})$. Note a variation here on the theme of Section 3.1: the slide is holding invariant the law of $\mathbb{P}$ given $\mathsf{Fine}_h \subset \mathsf{Cluster}_{[u,u+3r]}$, alongside the data $([\text{UpClust}], \text{DownClust})$, so that in addition to the occurrence of $\mathsf{Cluster}_{[u,u+3r]}$, the properties of Fine_h (which are conditionally fairly typical) must be respected when the slide resample operates.

The next result, which Figure 8 illustrates, gives a lower bound on the probability that the slide resample realizes a simple join at height h.

Proposition 4.11. *For any positive R and K, there exist $c = c(R,K) > 0$ and $C = C(K,d)$ such that the condition that h and $u - h$ exceed $C \log u$ implies that*

$$\mathbb{P}^{\mathsf{Fine}_h}\Big(\mathsf{SimpleJoin} \,\Big|\, [\text{UpClust}], \text{DownClust} \Big) \geq c \,.$$
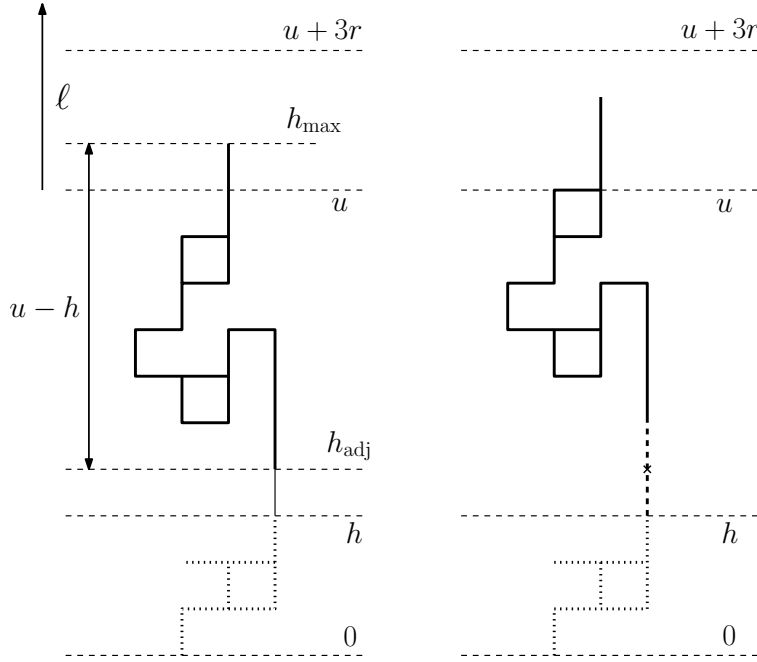
Figure 8: The slide resample illustrated, with $\mathsf{SimpleJoin}$ occurring as a result. *Left*: UpClust is bold, with height $u - h = h_{\max} - h_{\mathrm{adj}}$; DownClust is bold dotted. *Right*: UpClust moves up by one unit, to accomodate the two bold dashed edges that meet at the midvertex, marked with a cross.

And now an assertion, that a simple join is a renewal point, which will permit us to obtain Theorem 1.9(1) from the two preceding propositions.

Lemma 4.12. *When* SimpleJoin *occurs, the midvertex is a renewal vertex of the cluster* $\mathcal{C}_{u,u+3r}$.

The next claim will be used in proving this result. It is expressed using notation that develops the notion of vertex boundary $\mathrm{VertBdry}_h$ for $h \in \mathbb{R}$ that is specified in Section 1.2. Indeed, we set the *edge boundary* $\mathrm{EdgeBdry}_h$ equal to the set of elements of $E(\mathbb{Z}^d)$ that are incident to at least one element of $\mathrm{VertBdry}_h$.

Lemma 4.13. *Let* $v \in \mathbb{Z}^d$ *and* $e \in E(\mathbb{Z}^d)$.

1. *If* $e \in E(U_{v+e_d})$ *then* $e \notin \mathrm{EdgeBdry}_{v \cdot \boldsymbol{\ell}}$.

2. *If* $e \in E(L_{v-e_d})$ *then* $e \notin \mathrm{EdgeBdry}_{v \cdot \boldsymbol{\ell}}$.

Proof. Let w and x denote the endpoints of e. By the axes ordering convention, $e_d \cdot \boldsymbol{\ell} \geq e_i \cdot \boldsymbol{\ell} \geq 0$ for $i \in [\![1, d]\!]$. We also suppose that $w \cdot \boldsymbol{\ell} \geq x \cdot \boldsymbol{\ell}$. By these assumptions, x equals $w - e_i$ for some $i \in [\![1, d]\!]$.

(1). We need to show that neither w nor x are elements of $V(L_v)$. Note that, since $e \in E(U_{v+e_d})$, and $w \cdot \boldsymbol{\ell} \geq x \cdot \boldsymbol{\ell}$, we have that $w \cdot \boldsymbol{\ell} \geq (v + e_d) \cdot \boldsymbol{\ell}$. Note then that $x \cdot \boldsymbol{\ell} \geq w \cdot \boldsymbol{\ell} - e_d \cdot \boldsymbol{\ell} \geq (v + e_d) \cdot \boldsymbol{\ell} - e_d \cdot \boldsymbol{\ell} \geq v \cdot \boldsymbol{\ell}$.

We know that $w \cdot \boldsymbol{\ell}$ either equals, or exceeds, $(v + e_d) \cdot \boldsymbol{\ell}$. Suppose that equality holds. Then $x \cdot \boldsymbol{\ell}$ also equals $(v + e_d) \cdot \boldsymbol{\ell}$. Indeed, in the opposite case, we would have $x \cdot \boldsymbol{\ell} < (v + e_d) \cdot \boldsymbol{\ell}$. But then the edge interior (x, w) would lie in the backward space B_{v+e_d} (as specified in Section 1.2), which would imply, contrary to our assumption, that $e \in E(L_{v+e_d})$. We see then that, in the case that $w \cdot \boldsymbol{\ell} = (v + e_d) \cdot \boldsymbol{\ell}$, the heights of both w and x exceed $v \cdot \boldsymbol{\ell}$. This implies that neither w nor x lies in $V(L_v)$, so that the conclusion of Lemma 4.13(1) holds in this case.

Suppose instead that $w \cdot \boldsymbol{\ell} > (v + e_d) \cdot \boldsymbol{\ell}$. From $x = w - e_i$ and $e_i \cdot \boldsymbol{\ell} \leq e_d \cdot \boldsymbol{\ell}$, we find that $x \cdot \boldsymbol{\ell} > v \cdot \boldsymbol{\ell}$. In this case, then, w and x both have height exceeding v's. Thus, $w, x \notin V(L_v)$, so that Lemma 4.13(1) holds in this case also.

(2). Here, we need to show that $w, x \notin V(U_v)$.

Since $e \in E(L_{v-e_d})$ and the height of w is at least that of x, we have that $w \cdot \boldsymbol{\ell} \leq (v - e_d) \cdot \boldsymbol{\ell}$.

Note that, if $z \in V(U_v)$, then $(z, z + e_j) \cap F_v \neq \emptyset$ for some $j \in [\![1, d]\!]$. Thus, $z \cdot \boldsymbol{\ell} > v \cdot \boldsymbol{\ell} - e_d \cdot \boldsymbol{\ell}$.

For our purpose, it suffices then to show that $w \cdot \boldsymbol{\ell}$ and $x \cdot \boldsymbol{\ell}$ are at most $(v - e_d) \cdot \boldsymbol{\ell}$. Above we showed this for w, and by assumption $x \cdot \boldsymbol{\ell} \leq w \cdot \boldsymbol{\ell}$. Thus the proof of Lemma 4.13(2) is complete. □

Proof of Lemma 4.12. We see that $\mathcal{C}_{u,u+3r} = \mathrm{DownClust} \circ \mathrm{MidClust} \circ \mathrm{UpClust}$ fragments at v into two subgraphs whose edge sets are the union of $[v, v + e_d]$ and $E(\mathrm{UpClust})$; and the union of $[v - e_d, v]$ and $E(\mathrm{DownClust})$. To obtain Lemma 4.12, it is enough to argue that $e \in E(\mathrm{UpClust}) \cup E(\mathrm{DownClust})$ implies that $e \notin \mathrm{EdgeBdry}_{v \cdot \boldsymbol{\ell}}$. Given that $E(\mathrm{UpClust}) \subset E(U_{v+e_d})$ and $E(\mathrm{DownClust}) \subset E(L_{v-e_d})$, this statement follows from Lemma 4.13. □

Proof of Theorem 1.9(1). For $h \in [0, u]$, let $v = \mathrm{head}(\mathrm{DownClust}) + e_d$. Denote by $\mathscr{R}_u(h) \subseteq \mathsf{Cluster}_{[u,u+3r]}$ the event that v is a renewal point in $\mathcal{C}_{u,u+3r}$. We claim that there exist positive C and c such that, when h and $u - h$ exceed $C \log u$,

$$\mathbb{P}\big(\mathscr{R}_u(h)\big) \geq \mathbb{P}\big(\mathsf{Cluster}_{[u,u+3r]}, \mathsf{SimpleJoin}_h\big) \geq \mathbb{P}\big(\mathsf{Fine}_h, \mathsf{SimpleJoin}_h\big)$$

$$= \mathbb{E}\left[\mathbb{P}^{\mathsf{Fine}_h}_{[\mathrm{UpClust}_h], \mathrm{DownClust}_h}\big(\mathsf{SimpleJoin}\big) \cdot \mathbf{1}_{\mathsf{Fine}_h}\right] \geq c\,\mathbb{P}\big(\mathsf{Fine}_h\big),$$

where we have used the notation (17). Indeed, Lemma 4.12, $\mathsf{Fine}_h \subseteq \mathsf{Cluster}_{[u,u+3r]}$ and Proposition 4.11 deliver the respective inequalities.

For each $h \in [0, u]$, the associated value of head(DownClust) is an element of $V(L_h)$ that lies in $\mathsf{VertBdry}_h$. By Lemma 2.2(1), we see that, if two such values of h differ by at least $r = e_d \cdot \boldsymbol{\ell}$, the associated vertices v cannot be equal. Thus, if we ascertain the occurrence of a certain number of events $\mathscr{R}_u(h)$ for $h \in r\mathbb{N} \cap [0, u]$, then we can be sure that there are at least this many elements of $R(\mathcal{C}_{u,u+3r})$.

We may thus divide the bound derived above by $\mathbb{P}(\mathsf{Cluster}_{[u,u+3r]})$ and sum the result over values $h \in r\mathbb{N} \cap [0, u]$ to find that $\mathbb{E}\big[|R(\mathcal{C}_{u,u+3r})|\big|\mathsf{Cluster}_{[u,u+3r]}\big]$ is at least $c\,\mathbb{E}[N_0|\mathsf{Cluster}_{[u,u+3r]}]$, where N_0 denotes the number of values of $h \in r\mathbb{N} \cap [0, u]$ for which Fine_h occurs. Lemma 4.9 implies that N_0 is at least the quantity N specified in Proposition 4.10; thus, the latter result implies that $\mathbb{E}[N_0|\mathsf{Cluster}_{[u,u+3r]}] \geq c_1 u/r$ for some $c_1 > 0$. Whence,

$$ur^{-1}\,\mathbb{P}\big(|R(\mathcal{C}_{u,u+3r})| \geq c_2 ur^{-1}\big|\mathsf{Cluster}_{[u,u+3r]}\big) + c_2 ur^{-1}\,\mathbb{P}\big(|R(\mathcal{C}_{u,u+3r})| < c_2 ur^{-1}\big|\mathsf{Cluster}_{[u,u+3r]}\big)$$
$$\geq\ \ \mathbb{E}\big[|R(\mathcal{C}_{u,u+3r})|\big|\mathsf{Cluster}_{[u,u+3r]}\big]\ \geq\ cc_1 ur^{-1}$$

and $\mathbb{P}\big(|R(\mathcal{C}_{u,u+3r})| \geq c_2 ur^{-1}\big|\mathsf{Cluster}_{[u,u+3r]}\big) \geq cc_1 - c_2$, for any $c_2 > 0$. Taking $c_2 = cc_1/2$ and relabelling $c > 0$, we obtain Theorem 1.9(1). $\qquad\square$

We close this section by deriving Proposition 4.10. The more substantial Proposition 4.11 will be proved in the next.

Proposition 4.10 is reduced to Lemma 4.14 and the latter is then proved by means of Proposition 3.10.

Lemma 4.14. *There exists a positive choice of the constant c determined solely by $d \geq 3$ such that, when $\mathsf{Cluster}_{[u,u+3r]}$ occurs, the condition*

$$\big|\partial_{\mathrm{ext}}\overline{\mathcal{C}}_u\big| \ \leq\ \min\left\{cR^{\frac{d-2}{d-1}}K^{-1}au,\, 4^{-\frac{d-2}{d-1}}(LK)^{\frac{d-2}{d-1}}u^{d-2}\right\} \tag{18}$$

ensures that $N \geq (1 - a)u$.

Proof of Proposition 4.10. Set $D = cR^{\frac{d-2}{d-1}}K^{-1}a$. The hypothesised condition $L = 2^{\frac{d-2}{d-1}}RK^{-\frac{2d-3}{d-1}}$ assures alongside $u \geq 1$ that the minimum in (18) is attained by the first term, whatever the value of $u \in [1, \infty)$. We then obtain the sought result from Lemma 4.14 and Proposition 3.2 by suitably taking $C \geq D_0$ high and $c > 0$ low. $\qquad\square$

Proof of Lemma 4.14. First we argue that the global condition $\mathbf{G}(L)$ in Fine_h, namely that $|V(\mathcal{C}_{u,u+3r})| \leq Lu^d$, is implied by $\big|\partial_{\mathrm{ext}}\overline{\mathcal{C}}_u\big| \leq 4^{-\frac{d-2}{d-1}}(LK)^{\frac{d-2}{d-1}}u^{d-2}$. Suppose that $\mathbf{G}(L)$ fails. Since $\mathcal{C}_{u,u+3r}$ occupies heights in $[0, u]$, latitudes of elements of $\overline{\mathcal{C}}_u$ are integers in $[0, u/(2Kr)]$, so they number at most $u/(2Kr) + 1$, which is at most $u/(Kr)$ since $u \geq 2Kr$. There are thus more than $Lu^{d-1}Kr$ elements of $\overline{\mathcal{C}}_u$ living at the same latitude. Calling this latitude $k \in \mathbb{Z}$, note that $\partial_{\mathrm{ext}}\overline{C}_u \cap \overline{N}_{k,k+1}$ separates $\overline{\mathcal{C}}_u \cap \overline{N}_{k,k+1}$ from infinity in $\overline{N}_{k,k+1}$, so that Proposition 3.10 implies that

$$\big|\overline{\mathcal{C}}_u \cap \overline{N}_{k,k+1}\big| \ \leq\ 2\big|\overline{N}_{k,k+1} \cap \partial_{\mathrm{ext}}\overline{\mathcal{C}}_u\big|^{\frac{d-1}{d-2}},$$

where the notation $\overline{N}_{k,k+1}$ is introduced in Subsection 3.2.3. Since the left-hand side is at least $Lu^{d-1}Kr$, $\partial_{\mathrm{ext}}\overline{\mathcal{C}}_u$ contains more than $4^{-\frac{d-2}{d-1}}(LK)^{\frac{d-2}{d-1}}u^{d-2}$ elements. This is as we sought to argue. Hence, condition $\mathbf{G}(L)$ is satisfied, whatever the value of h.

Lemma 4.15. *For $q \in \mathbb{R}$, let Γ_q denote the set of latitudes of K boxes that contain a vertex with $v \cdot \boldsymbol{\ell} \in [q, q+1]$. Then Γ_q is an integer interval $[\![a,b]\!]$ with $a = \lfloor (q - Kd^{1/2})(2Kr)^{-1} \rfloor$ and $a + d^{1/2}r^{-1} \leq b \leq a + d^{1/2}r^{-1} + 1$.*

Proof. Suppose that the K-box $B = 2Ku + [\![-K, K]\!]^d$ contains v. Then $|v \cdot \boldsymbol{\ell} - 2Ku \cdot \boldsymbol{\ell}| \leq Kd^{1/2}$, so that $v \cdot \boldsymbol{\ell} \in [q, q+1]$ implies that $q - Kd^{1/2} \leq 2Ku \cdot \boldsymbol{\ell} \leq q + Kd^{1/2} + 1$. The quantities $\lfloor (q - Kd^{1/2})(2Kr)^{-1} \rfloor$ and $\lfloor (q + Kd^{1/2} + 1)(2Kr)^{-1} \rfloor$ are thus lower and upper bounds on the latitude of B. Lemma 4.15 will follow once we conform that Γ_q is an integer interval. If Γ_q^+ and Γ_q^- denote the counterparts to Γ_q in whose definition $[q, q+1]$ is respectively replaced by $[q, \infty)$ and $(-\infty, q+1]$ then Γ_q may be written $\Gamma_q^+ \cap \Gamma_q^-$ as an intersection of two semi-infinite integer intervals (that extend to infinity in opposing directions); whence we learn that Γ_q is an integer interval indeed. $\square$

Let F_i denote the set of $h \in \mathbb{N} \cap [0, u]$ such that there are at least i^3R elements $x \in V(\mathcal{C}_{u,u+3r})$ such that $x \cdot \boldsymbol{\ell} \in [h+i-1, h+i]$. Let $G_i \subseteq F_i$ be chosen so that the distance between any pair of elements in G_i exceeds $Kd^{1/2} + 2$, with $|G_i| \geq |F_i|(Kd^{1/2} + 3)^{-1}$.

Lemma 4.15 permits us to associate to each $j \in G_i$ a value $\ell_j \in \Gamma_{j+i-1}$ such that at least $i^3R(d^{1/2}r^{-1} + 2)^{-1}$ elements of $\overline{\mathcal{C}}_{u,u+3r}$ have latitude ℓ_j. Elements of $\{\Gamma_{j+i-1} : j \in G_i\}$ are at pairwise distance at least two, so that the differences between consecutive ℓ_j as indexed by $j \in G_i$ are also at least two. Thus, the sets $\overline{\mathcal{C}}_{u,u+3r} \cap \overline{N}_{\ell_j,\ell_j+1}$ indexed by $j \in G_i$ are pairwise disjoint. Since $\overline{\mathcal{C}}_{u,u+3r} \cap \overline{N}_{\ell_j,\ell_j+1}$ is separated from infinity in $\overline{N}_{\ell_j,\ell_j+1}$ by $\partial_{\mathrm{ext}}\overline{\mathcal{C}}_{u,u+3r}$, Proposition 3.10 implies that

$$\left|\overline{\mathcal{C}}_{u,u+3r} \cap \overline{N}_{\ell_j,\ell_j+1}\right| \leq 2\left|\partial_{\mathrm{ext}}(\overline{\mathcal{C}}_{u,u+3r}) \cap \overline{N}_{\ell_j,\ell_j+1}\right|^{\frac{d-1}{d-2}}.$$

But

$$\left|\overline{\mathcal{C}}_{u,u+3r} \cap \overline{N}_{\ell_j,\ell_j+1}\right| \geq \frac{i^3R}{d^{1/2}r^{-1} + 2}$$

so that

$$\left|\partial_{\mathrm{ext}}(\overline{\mathcal{C}}_{u,u+3r}) \cap \overline{N}_{\ell_j,\ell_j+1}\right| \geq 2^{\frac{2-d}{d-1}}\left(\frac{i^3R}{d^{1/2}r^{-1} + 2}\right)^{\frac{d-2}{d-1}}$$

for $j \in G_i$. The disjointness of $\overline{N}_{\ell_j,\ell_j+1}$ for $j \in G_i$ implies that

$$\left|\partial_{\mathrm{ext}}(\overline{\mathcal{C}}_{u,u+3r})\right| \geq 2^{\frac{2-d}{d-1}}\left(\frac{i^3R}{d^{1/2}r^{-1} + 2}\right)^{\frac{d-2}{d-1}}|G_i|.$$

By assumption $\left|\partial_{\mathrm{ext}}(\overline{\mathcal{C}}_{u,u+3r})\right| \leq cR^{\frac{d-2}{d-1}}K^{-1}au$. Recall that $|G_i| \geq |F_i|(Kd^{1/2} + 3)^{-1}$. Thus,

$$|F_i| \leq cR^{\frac{d-2}{d-1}}K^{-1}2^{\frac{d-2}{d-1}}\left(\frac{i^3R}{d^{1/2}r^{-1} + 2}\right)^{-\frac{d-2}{d-1}}(Kd^{1/2} + 3)u,$$

where the right-hand term here is bounded above by $c\alpha i^{-3(d-2)/(d-1)}au$, with α a constant depending on d but not on R or K. Thus $\sum_{i=1}^{A}|F_i| \leq c\alpha\zeta(3/2)au$ for any $d \geq 3$. For a suitably small

d-dependent choice of the positive constant c, we see that $\sum_{i=1}^{A}|F_i| \leq au/4$. A quantity F_i' may be defined as F_i was, with the height interval $[h+i-1, h+i]$ replaced by $[h-i, h-i+1]$. An essentially identical argument to that just given yields the bound $\sum_{i=1}^{A}|F_i'| \leq au/4$. The expression $\sum_{i=1}^{A}\left(|F_i| + |F_i'|\right)$ is an upper bound on the cardinality of the set of $h \in \mathbb{N} \cap [0,u]$ that violate at least one of the tropical conditions $\mathbf{T}_j(R)$, $(j,i) \in [\![1,2]\!] \times [\![1,A]\!]$, that specify $\mathsf{VeryFine}_h$. Thus we see that the number of tropical condition violating values of h is at most $au/2$.

Let χ_i equal the cardinality of the set of $x \in V(\mathcal{C}_{u,u+3r})$ for which $x \cdot \boldsymbol{\ell} \in (i-1, i]$ and let

$$M = \left\{ h \in \mathbb{N} \cap [0,u] : \chi_i > R \text{ for at least } 4A/3 \text{ values } i \in [\![h+1+A, h+3A]\!] \right\}.$$

Note that $4A/3$ equals two-thirds of the cardinality of $[\![h+1+A, h+3A]\!]$.

Let M denote the set of $h \in \mathbb{N} \cap [0,u]$ such that the number of $x \in \mathcal{C}_{u,u+3r}$ with $x \cdot \boldsymbol{\ell} \in (i-1, 1]$ exceeds R for at least two-thirds of the indices i in $[\![h+1+A, h+3A]\!]$. For $h \in M$, write

$$O_h = \left[h + A - 3Kd^{1/2}r^{-1}, h + 3A + 3Kd^{1/2}r^{-1} \right].$$

Given a finite collection $\mathcal{C}$ of real intervals of equal length, a pairwise disjoint subset $\mathcal{I}$ may be selected such that $\sum_{I \in \mathcal{I}} |I| \geq \frac{1}{2} \left| \bigcup_{I \in \mathcal{C}} I \right|$. We may thus let $U \subseteq M$ be chosen so that $\{ O_u : u \in U \}$ is a pairwise disjoint collection that satisfies

$$\sum_{u \in U} |O_u| \geq \frac{1}{2} \cdot \left| \bigcup_{m \in M} O_m \right|.$$

Let $h \in U$ and write $K_h = \left\{ i \in [\![h+1+A, h+3A]\!] : \chi_i \geq R \right\}$. We form a subset J_h of K_h by first including in it the minimum value in K_h, and then iteratively including every $\lceil 5Kd^{1/2} \rceil^{\text{th}}$ value of K_h as these are encountered in increasing order. The constructed set J_h has the properties that $j \in J_h$ implies that $\chi_j > R$; that the distance between consecutive elements of J_h is at least $5Kd^{1/2}$; and that $|J_h| \geq \frac{4A}{3(5Kd^{1/2}+1)}$.

Let $u \in U$ and $j \in J_u$. Since $\chi_j \geq R$, Lemma 4.15 permits us to find $\ell_j \in \Gamma_{j-1}$ such that $\overline{\chi}_j \geq R(d^{1/2}r^{-1} + 2)^{-1}$, where $\overline{\chi}_j$ equals the number of K-boxes in $\overline{\mathcal{C}}_{u,u+3r}$ of latitude ℓ_j.

We *claim* that there is no coincidence among ℓ_j as u and j run over U and J_u, and indeed that the pairwise distance between these quantities is at least two. To confirm this, we begin by arguing that when $j \in J_u$ and $j' \in J_{u'}$ for $u, u' \in U$, equality $j = j'$ is possible only if $u = u'$, and, if $j \neq j'$, then $|j - j'| \geq 5Kd^{1/2}$. Indeed, if it happens that $j, j' \in J_u$ for a common value of u, then, if $j \neq j'$, $|j - j'|$ is indeed at least $5Kd^{1/2}$. If $j \in J_u$ and $j \in J_{u'}$ for distinct u and u', then the distances $d(j, O_u^c)$ and $d(j', O_{j'}^c)$ are both at least $3Kd^{1/2}r^{-1}$, so that the disjointness of O_u and $O_{u'}$ implies that $|j - j'| \geq 6Kd^{1/2}r^{-1}$.

To arrive at the claim, we next argue that, if $k, k' \in \mathbb{Z}$ differ by at least $5Kd^{1/2}$, then Γ_k and $\Gamma_{k'}$ lie at distance at least two. To see this, note that, since $d \geq 3$ and $r \leq 1$, the difference between k and k' is at least $2Kr(d^{1/2}r^{-1} + 2)$; from here, Lemma 4.15 shows the assertion just made.

We know that $\ell_j \in \Gamma_{j-1}$ and $\ell_{j'} \in \Gamma_{j'-1}$. Taking $k = j - 1$ and $k' = j' - 1$, note that we have confirmed the hypothesis that $d(k, k') \geq 5Kd^{1/2}$, so that the assertion in the preceding paragraph yields the sought claim.

Next we will show that

$$\Big| \bigcup_{h \in M} O_h \Big| \geq 2M/5 \,. \tag{19}$$

To see this, set $\kappa = |\mathbb{Z} \cap O_h|$ for any $h \in M$. For $x \in \mathbb{Z}$, let $\phi_x = \big|\{h \in M : x \in O_h\}\big|$. Since any $x \in \mathbb{Z}$ belongs to at most $2\kappa + 1$ intervals O_h with $h \in M$, we see that $\max\{\phi_x : x \in \mathbb{Z}\} \leq 2\kappa + 1$. Note then that

$$\kappa|M| = \sum_{h \in M} |O_h| = \sum_{x \in \mathbb{Z}} \phi_x \leq (2\kappa + 1)|\mathrm{supp}(\phi)| \leq (2\kappa + 1)\Big| \bigcup_{h \in M} O_h \Big| \,,$$

since $\mathrm{supp}(\phi) = \cup_{h \in M} O_h$. Thus, $\sum_{h \in M} |O_h| \geq \frac{\kappa}{2\kappa+1}|M|$. Since

$$\kappa = \big|\mathbb{Z} \cap \big[h + A - 3Kd^{1/2}r^{-1}, h + 3A + 3Kd^{1/2}r^{-1}\big]\big|$$

$\kappa \geq 2A - 1 + 6Kd^{1/2}r^{-1} \geq 2A$, where the latter bound depends on $d \geq 3$, $K \geq 1$ and $r \leq 1$. We confirm that (19) holds when $A \geq 1$.

Note that

$$(2A + 6Kd^{1/2}r^{-1})|U| = \sum_{u \in U} |O_u| \geq \tfrac{1}{2} \cdot \Big| \bigcup_{m \in M} O_m \Big| \geq M/5 \,,$$

where (19) provides the latter bound. Proposition 3.10 implies that, for each $u \in U$ and $j \in J_u$,

$$\Big| \partial_{\mathrm{ext}} \overline{\mathcal{C}}_{u,u+3r} \cap \overline{N}_{\ell_j, \ell_j+1} \Big| \geq 2^{\frac{2-d}{d-1}} \Big(\frac{R}{d^{1/2}r^{-1}+2} \Big)^{\frac{d-2}{d-1}} \,.$$

Since each pair of these latitudes differ by at least two, we find that the the number of K-boxes in $\partial_{\mathrm{ext}} \overline{\mathcal{C}}_{u,u+3r}$ at latitudes $\{\ell_{u,j}, \ell_{u,j} + 1 : u \in U, j \in J_u\}$ is at least $2^{\frac{2-d}{d-1}} \Big(\frac{R}{d^{1/2}r^{-1}+2} \Big)^{\frac{d-2}{d-1}} |U| \frac{4A}{3(5Kd^{1/2}+1)}$. But $\big| \partial_{\mathrm{ext}} \overline{\mathcal{C}}_{u,u+3r} \big| \leq cR^{\frac{d-2}{d-1}} K^{-1} au$, which implies that

$$|U| \leq cR^{\frac{d-2}{d-1}} K^{-1} 2^{\frac{d-2}{d-1}} \Big(\frac{d^{1/2}r^{-1}+2}{R} \Big)^{\frac{d-2}{d-1}} \frac{3(5Kd^{1/2}+1)}{4A} au \,.$$

We find that $|M| \leq 10(A + 3Kd^{1/2}r^{-1})|U|$. Using $A \geq 3Kd^{1/2}r^{-1}$, we obtain

$$|M| \leq 15r^{-1} c(2R)^{\frac{d-2}{d-1}} K^{-1} \big((d^{1/2}r^{-1} + 2)R^{-1} \big)^{\frac{d-2}{d-1}} (5Kd^{1/2} + 1)au \,;$$

whence $|M| \leq c\beta u$, where β is a positive constant determined solely by d. We may then choose $c = \beta/4$ to ensure that $|M| \leq au/4$. Thus, the set of h for which the northzone condition $\mathbf{S}_1(R)$ specifying $\mathsf{VeryFine}_h$ is violated is at most $au/4$. Similarly, the same upper bound holds in regard to the southzone condition $\mathbf{S}_2(R)$.

In summary, the cardinality of the set of $h \in \mathbb{N} \cap [0, u]$ for which $\mathsf{VeryFine}_h$ does not occur is at most au. We have proved Lemma 4.14. $\qquad\square$

4.2 Simple joins are often found at fine levels: deriving Proposition 4.11

To prove Proposition 4.11, we need to gain a more explicit understanding of how the slide resample operates. A first subsection presents Lemma 4.18, to this effect. In this lemma appears a denominator or partition function. This is a sum over head locations $v \in \mathbb{Z}^d \cap F_u \cap B_{u+3r}$ of what we have called connection chances: conditional probabilities for cluster formation given, roughly speaking, the construction of the downcluster, and of the shift of the upcluster so that its head lies at v. In Subsection 4.2.2, we reduce Proposition 4.11 to Proposition 4.20, which offers an upper bound on the partition function. The proof of the latter result, which involves coarse-grained percolation ideas seen in Section 3.2, is the subject of Subsection 4.2.3.

4.2.1 An explicit description of the conditional law in the slide resample

Two definitions concerning the resample permit us to present Lemma 4.18, which rigorously expresses the notion advanced in Section 3.1 that any given shift value is selected under resampling with probability proportional to its connection chance.

Recall the random shift-equivalence class [UpClust] specified under $\mathbb{P}$ given $\mathsf{Cluster}_{[u,u+3r]}$ by Definition 4.7.

Definition 4.16. Let $v \in \mathbb{Z}^d \cap F_u \cap B_{u+3r}$ (so that $v \cdot \boldsymbol{\ell} \in [u, u+3r]$). On the event $\mathsf{Cluster}_{[u,u+3r]}$, we denote by $\mathrm{UpClust}^v$ the unique element of [UpClust] whose head equals v. Note that the equivalence class [UpClust] is equal to $\{\mathrm{UpClust}^v : v \in \mathbb{Z}^d \cap F_u \cap B_{u+3r}\}$. Under the percolation law $\mathbb{P}$ conditionally on the occurrence of $\mathsf{Cluster}_{[u,u+3r]}$, the cluster $\mathcal{C}_{u,u+3r}$ is well defined but random; thus, so is UpClust. We may thus define, under the event $\mathsf{Cluster}_{[u,u+3r]}$, the random vector $V \in \mathbb{Z}^d \cap F_u \cap B_{u+3r}$ such that UpClust is equal to $\mathrm{UpClust}^V$.

Definition 4.17. Let Ω denote the power set of the collection of nearest-neighbour edges in $\mathbb{Z}^d$. The set Ω inherits a topology by identifying this set with a unit interval in $\mathbb{R}$, and we set $\mathcal{B}$ equal to the associated Lebesgue σ-algebra on Ω. The law $\mathbb{P}$ of bond percolation on the graph $(\mathbb{Z}^d, \sim)$, in which each edge is independently declared *open* with probability p, and *closed* in the opposing event, will be treated as being defined on the measurable space $(\Omega, \mathcal{B})$.

We select a second copy $(\Omega', \mathcal{B}', \mathbb{P}')$ of this probability space $(\Omega, \mathcal{B}, \mathbb{P})$. The product space $(\Omega \times \Omega', \mathcal{B} \times \mathcal{B}', \mathbb{P} \times \mathbb{P}')$ describes two independent bond percolations on $\mathbb{Z}^d$ of parameter p. The first may be called the *original* percolation and the second, associated to the prime variable, the *auxiliary* percolation.

Let Ω'' denote a further copy of the power set of nearest-neighbour edges in $\mathbb{Z}^d$. For each $v \in \mathbb{Z}^d \cap F_u \cap B_{u+3r}$, we specify a further configuration $\omega''_v \in \Omega''$ as a function of $(\omega, \omega') \in \Omega \times \Omega'$.

An edge is called ω-open or ω-closed if it is, or is not, an element of ω; naturally the same usage may be made for ω' and ω''_v. Set ω'' to be a definite choice—the identically closed configuration, say—if $\omega \notin \mathsf{Cluster}_{[u,u+3r]}$. Suppose now that $\omega \in \mathsf{Cluster}_{[u,u+3r]}$. In what follows, every random variable, such as $\mathrm{UpClust}^v$ or DownClust, is understood to be specified by the configuration ω. To specify the configuration ω''_v, we classify edges in $\mathbb{Z}^d$.

1. Elements of $E(\mathrm{UpClust}^v) \cup E(\mathrm{DownClust})$ are called *discovered open*.

2. Elements of $E(L_h) \cap \partial_E(\mathrm{DownClust})$ and $E(U_{h_{\mathrm{adj}}}) \cap \partial_E(\mathrm{UpClust}^v)$ are *discovered closed*.

45

3. Every further nearest-neighbour edge in $\mathbb{Z}^d$ is called *undiscovered*.

Here, $\partial_E(G)$, for $G \subseteq E(\mathbb{Z}^d)$, denotes the set of edges in $\mathbb{Z}^d$ that do not belong to G but each of which is incident to some element of G.

Each discovered edge is ω_v''-open precisely when it is ω-open. Each undiscovered edge is ω_v''-open precisely when it is ω'-open. Note that, in so specifying ω_v'', discovered open edges are ω_v''-open, and discovered closed edges are ω_v''-closed.

Let $[\mathfrak{u}]$ and $\mathfrak{d}$ denote possible values for [UpClust] and DownClust as ω ranges over $\mathsf{Cluster}_{[u,u+3r]}$. Set $R^v_{[\mathfrak{u}],\mathfrak{d}}(\omega')$ equal to the value of ω_v'' given $\mathsf{Cluster}_{[u,u+3r]}$, [UpClust] $= [\mathfrak{u}]$ and DownClust $= \mathfrak{d}$. The dependence is on ω' because it is the values of ω_v'' on undiscovered edges which dictate the value of ω_v'' in this circumstance.

Set $\mathbb{R}^v_{[\mathfrak{u}],\mathfrak{d}} : \mathcal{B} \to [0,1]$, $\mathbb{R}^v_{[\mathfrak{u}],\mathfrak{d}}(A) = \mathbb{P}'\big(\omega' \in \Omega' : R^v_{[\mathfrak{u}],\mathfrak{d}}(\omega') \in A\big)$.

In the next result, we use this framework to express conditional probability given a subevent of $\mathsf{Cluster}_{[u,u+3r]}$. Just as random variables under $\mathbb{P}$ are understood to be given by the ω-configuration, so are events: for example, in the left-hand side of Lemma 4.18.

Lemma 4.18. *Let $B, H \subseteq \Omega$ denote two $\mathcal{B}$-measurable events, with $\mathbb{P}(B|\mathsf{Cluster}_{[u,u+3r]}) > 0$. Let $([\mathfrak{u}], \mathfrak{d})$ be a value that $([\mathrm{UpClust}], \mathrm{DownClust})$ assumes with positive probability under $\mathbb{P}$ given $B \cap \mathsf{Cluster}_{[u,u+3r]}$. Then*

$$
\mathbb{P}\left(H \,\middle|\, [\mathrm{UpClust}] = [\mathfrak{u}], \mathrm{DownClust} = \mathfrak{d}, B, \mathsf{Cluster}_{[u,u+3r]} \right)
$$
$$
= \frac{\sum_{v \in \mathbb{Z}^d \cap F_u \cap B_{u+3r}} \mathbb{R}^v_{[\mathfrak{u}],\mathfrak{d}}\big(H \cap B \cap \mathsf{Cluster}_{[u,u+3r]} \cap \{V = v\}\big)}{\sum_{v \in \mathbb{Z}^d \cap F_u \cap B_{u+3r}} \mathbb{R}^v_{[\mathfrak{u}],\mathfrak{d}}\big(B \cap \mathsf{Cluster}_{[u,u+3r]} \cap \{V = v\}\big)} \,,
$$

where V is the random vector specified under $\mathbb{P}(\cdot \,|\, \mathsf{Cluster}_{[u,u+3r]})$ in Definition 4.16.

Proof. Let $G \subseteq \Omega$ be an arbitrary $\mathcal{B}$-measurable event. Take $v \in \mathbb{Z}^d \cap F_u \cap B_{u+3r}$, and consider the event $\mathsf{Cluster}_{[u,u+3r]} \cap G \cap \{\mathrm{UpClust} = \mathfrak{u}^v\} \cap \{\mathrm{DownClust} = \mathfrak{d}\}$. Does this event occur? We may find the answer in two stages. First, we check a necessary condition, an event that we label N: namely, that elements of $E(\mathfrak{u}^v) \cup E(\mathfrak{d})$ are ω-open and elements of $E(L_u) \cap \partial_E(\mathfrak{d})$ and $E\big(U_{h_{\max}-(u-h)}\big) \cap \partial_E(\mathfrak{u}^v)$ are ω-closed. If this test is passed, we move to the second stage, and realize all further edges in $E(\mathbb{Z}^d)$. Check that the event $\mathsf{Cluster}_{[u,u+3r]} \cap G$ occurs. If this happens, then DownClust must equal $\mathfrak{d}$, but UpClust is not yet assuredly $\mathfrak{u}^v$. We need to confirm the latter condition, and this is done by checking (at the end of the second stage) that the random vector V equals v.

Writing $\alpha_v = \mathbb{P}(N) > 0$, note that α_v and $\mathbb{R}^v_{[\mathfrak{u}],\mathfrak{d}}\big(G \cap \mathsf{Cluster}_{[u,u+3r]} \cap \{V = v\}\big)$ equal the respective probabilities that the checks at the first and second stages are passed. In fact, α_v may be written α because translation invariance of percolation in the upspace implies that this quantity has no dependence on v. What we have learnt is that, for $v \in \mathbb{Z}^d \cap F_u \cap B_{u+3r}$,

$$
\mathbb{P}\left(G, \mathrm{UpClust} = \mathfrak{u}^v, \mathrm{DownClust} = \mathfrak{d}, \mathsf{Cluster}_{[u,u+3r]} \right) = \alpha \cdot \mathbb{R}^v_{[\mathfrak{u}],\mathfrak{d}}\big(G \cap \mathsf{Cluster}_{[u,u+3r]} \cap \{V = v\}\big).
$$
(20)

Now set $G = B \cap \mathsf{Cluster}_{[u,u+3r]}$ and sum (20) out over $v \in \mathbb{Z}^d \cap F_u \cap B_{u+3r}$. The resulting value is positive by our assumptions. Then take G equal to $H \cap B \cap \mathsf{Cluster}_{[u,u+3r]}$ and sum out similarly. Dividing the latter of the resulting equalities by the former, we obtain Lemma 4.18. $\qquad\square$

4.2.2 Reducing Proposition 4.11 to a partition function upper bound

We reduce Proposition 4.11 to two results the first of which we then prove.

Lemma 4.19. *There exists $c > 0$ such that $\sum_{v \in \mathbb{Z}^d \cap F_u \cap B_{u+3r}} \mathbb{R}^v_{[\mathrm{UpClust}],\mathrm{DownClust}}(\mathsf{Fine}, V = v) \geq c$ for $\omega \in \mathsf{VeryFine}$. Indeed, for such ω there exists $x \in \mathbb{Z}^d \cap F_u \cap B_{u+3r}$ for which*

$$\mathbb{R}^x_{[\mathrm{UpClust}],\mathrm{DownClust}}\big(\mathsf{Fine} \cap \mathsf{SimpleJoin} \cap \{V = x\}\big) \geq c \,.$$

(The value of h that specifies $\mathsf{Fine} = \mathsf{Fine}_h$ is given in Proposition 4.11 and we will for now omit it from this notation.)

The second result is our partition function upper bound, which we prove in the next subsection.

Proposition 4.20. *Suppose A/K is bounded below by a suitably high constant multiple of $\log u$. Then there exists $C > 0$ such that $\sum_{v \in \mathbb{Z}^d \cap F_u \cap B_{u+3r}} \mathbb{R}^v_{[\mathrm{UpClust}],\mathrm{DownClust}}(\mathsf{Fine}) \leq C$ whenever $\omega \in \mathsf{Fine}$.*

Proof of Proposition 4.11. Take $H = \mathsf{SimpleJoin}$ and $B = \mathsf{Fine}_u$ in Lemma 4.18. We seek to bound the right-hand numerator below, and its denominator above. The assumption that h and $u - h$ exceed a large constant multiple of $\log u$ permits a choice of A of the same order, with $\min\{h, u - h\} > 3A$. This choice of A allows us to apply Proposition 4.20, to bound above the denominator. Bounding the numerator below via Lemma 4.19, we obtain the sought assertion. $\square$

Lemma 4.21. *The event Fine entails that the vertex sets $\mathrm{Bed}(\mathrm{UpClust})$ and $\mathrm{Top}(\mathrm{DownClust})$ have cardinality at most R.*

Proof. By Lemma 2.2(1), every element of $\mathrm{Top}(\mathrm{DownClust}) \subset \mathrm{VertBdry}_h$ is a vertex in $\mathcal{C}_{u,u+3r}$ whose height lies in $[h - r, h]$ for $r = e_d \cdot \boldsymbol{\ell}$. Recalling that $h_{\mathrm{adj}} = h_{\max} - (u - h)$, note that every element of $\mathrm{Bed}(\mathrm{UpClust})$ lies in $\mathrm{VertBdry}_{h_{\mathrm{adj}}}$, so that the just quoted result implies that each such element has height in $[h_{\mathrm{adj}} - r, h_{\mathrm{adj}}]$. Since $r \leq 1$, we use properties $\mathbf{T}'_1(R)$ and $\mathbf{T}'_2(R)$ (with $i = 1$) of the event Fine to obtain Lemma 4.21. $\square$

Proof of Lemma 4.19. Select a vector $x \in \mathbb{Z}^d$ such that some vertex in $\mathrm{Bed}(\mathrm{UpClust}^x)$ is displaced from some vertex $v \in \mathrm{Top}(\mathrm{DownClust})$ by $2e_d$. (If for given x there is a choice of admissible v, we pick v from the admissible set to be the lexicographically minimal vertex among those of minimum height.) The height $x \cdot \boldsymbol{\ell}$ of x equals the sum of the heights of $\mathrm{DownClust}$ and $\mathrm{UpClust}$ and the value $2e_d \cdot \boldsymbol{\ell}$. By Lemma 4.1, $x \cdot \boldsymbol{\ell} \in [u + e_d \cdot \boldsymbol{\ell}, u + 3e_d \cdot \boldsymbol{\ell}] \subset [u, u + 3r]$. We have confirmed that $x \in \mathbb{Z}^d \cap F_u \cap B_{u+3r}$.

Consider the event E that

- the edges $[v, v + e_d]$ and $[v + e_d, v + 2e_d]$ are ω'-open, while all other edges incident to $v + e_d$ are ω'-closed, as are all other undiscovered edges that are incident to $\mathrm{Bed}(\mathrm{UpClust}^u) \cup \mathrm{Top}(\mathrm{DownClust})$.

For each $\omega \in \mathsf{Fine}$,

$$\mathbb{R}^x_{[\mathrm{UpClust}],\mathrm{DownClust}}(E) \geq p^2 (1-p)^{2(d-1)} (1-p)^{(2d-1) \cdot 2(R-1)} \,.$$

Indeed, the right-hand term bounds the probability of E: p^2 opens the two e_d-oriented edges, $(1-p)^{2(d-1)}$ closes the other edges incident to the shared endpoint of this pair, and $(1-p)^{(2d-1) \cdot 2(R-1)}$ makes the other needed closures in view of Lemma 4.21.

When the configuration $R^x_{[\text{UpClust}],\text{DownClust}}(\omega')$ realizes the event E, MidClust contains precisely two edges, these of the form $[v, v + e_d]$ and $[v - e_d, v]$ for some given $v \in \mathbb{Z}^d$ with $v \cdot \boldsymbol{\ell} = h$; and the graph MidFragment is empty. Thus, $R^x_{[\text{UpClust}],\text{DownClust}}(\omega')$ realizes the event SimpleJoin in this circumstance. Moreover, we *claim* that, when $\omega \in$ Fine and $R^x_{[\text{UpClust}],\text{DownClust}}(\omega') \in E$, we have that $R^x_{[\text{UpClust}],\text{DownClust}}(\omega') \in$ Fine $\cap \{V = u\}$. To confirm that $R^x_{[\text{UpClust}],\text{DownClust}}(\omega') \in$ Fine, we need to check the tropical, subtropical and global conditions for Fine. To do so, note that any vertex in $\mathcal{C}_{u,u+3r}$ is an element of precisely one of the upcluster, the midcluster and the downcluster; in the transition from ω to $R^x_{[\text{UpClust}],\text{DownClust}}(\omega')$, the upcluster may have shifted; the midcluster has collapsed, to be confined to its minimum vertex cardinality of three, with vertices $v - e_d$, v and $v + e_d$; and DownClust has remained unaltered. The various upper bounds on vertex cardinalities in the three conditions are thus seen to pass from ω to $R^x_{[\text{UpClust}],\text{DownClust}}(\omega')$. Thus, the latter configuration indeed verifies Fine. To confirm that $V = x$, it is enough to check that UpClust under $R^x_{[\text{UpClust}],\text{DownClust}}(\omega')$ is equal to $\text{UpClust}^x(\omega)$. There is a unique open connected component of $\text{Cluster}_{[u,u+3r]}$ in $U_{h_{\max}-(u-h)}$ under this configuration, because the only element of $E(\text{Cluster}_{[u,u+3r]})$ that our construction permits that has one endpoint in VertBdry_h and the other in $V(U_h)\backslash\text{VertBdry}_h$ is the edge $[v, v + e_d]$. Since this unique component equals $\text{UpClust}^x(\omega)$, we indeed find that $V = x$. This completes the proof of the claim. We learn that, for $\omega \in$ Fine,

$$\mathbb{R}^x_{[\text{UpClust}],\text{DownClust}}\big(\text{Fine} \cap \text{SimpleJoin}\big) \geq p^2(1-p)^{2(2dR-d-R)}.$$

The latter statement of Lemma 4.19 has been derived; the former is a trivial consequence. $\qquad\square$

4.2.3 Proving the partition function upper bound

Here we prove Proposition 4.20. To argue that $\mathbb{R}^v_{[\text{UpClust}],\text{DownClust}}(\text{Fine})$ is small when $v \in \mathbb{Z}^d \cap F_u \cap B_{u+3r}$ is large, we will note that the tropical conditions in Fine entail that UpClust^v and DownClust have a sparse presence near the equator. Thus, if v is large, these graphs visit the tropical region at significant horizontal separation. The occurrence of Fine $\subset \text{Cluster}_{[u,u+3r]}$ under $\mathbb{R}^v_{[\text{UpClust}],\text{DownClust}}$ thus requires the formation of new, independent, cluster boundaries to bridge the gap, either horizontal or vertical, between UpClust^v and DownClust. The cost of these cluster boundaries is high, so that $\mathbb{R}^v_{[\text{UpClust}],\text{DownClust}}(\text{Fine})$ decays rapidly as the norm of v rises.

To show that such costs are indeed high, we will use the coarse-graining of percolation seen in Section 3.2. Indeed, we let $K \in \mathbb{N}$ be a positive parameter, and recall from that section that the collection of K-boxes tesselates $\mathbb{Z}^d$, and has an index set which may be naturally identified with $\mathbb{Z}^d$. Recall that a path in K-boxes makes nearest-neighbour steps given this identification; and also that the notion of $*$-neighbour specified at the start of Section 3.2 passes naturally from $\mathbb{Z}^d$ to the set of K-boxes.

The notation $\overline{A}$ comes from Definition 3.5. Since MidClust is nearest-neighbour connected, so is $\overline{\text{MidClust}}$; the latter set thus lies in a single $*$-connected component of $\overline{\text{ResidualClust}}$. We call this component $\mathcal{H}$. Using Definition 3.1, we set $\text{Surface}_K = \partial_{\text{ext}}\mathcal{H}$.

Lemma 4.22.

1. *The collection* Surface_K *is a nearest-neighbour connected set of K-boxes.*

2. *Every element of* Surface_K *is disjoint from* ResidualClust.

3. *Every infinite path in K-boxes lying in $\overline{S}_{0,n}$ that begins at a box that intersects* MidClust *visits* Surface$_K$.

Proof: (1). This is due to [Tim13, Theorem 4].

(2,3). These are direct from definitions. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Lemma 4.23. *Suppose that* Cluster$_{[u,u+3r]}$ *occurs. Let B be a K-box that lies in* Surface$_K$ *and whose vertex set is disjoint from the union of the vertex sets of* UpClust *and* DownClust. *Then B is a bad K-box.*

Proof. Note that B is disjoint from $\mathcal{C}_{u,u+3r}$. Indeed, any vertex in $\mathcal{C}_{u,u+3r}$ lies in at least one of DownClust, UpClust and ResidualClust. A vertex in B may not lie in any of these graphs: the first and second are excluded by hypothesis; the third, by Lemma 4.22(2).

Let D denote a K-box $*$-neighbour of B that intersects ResidualClust $\subseteq \mathcal{C}_{u,u+3r}$. Lemma 3.3 now implies that B is a bad K-box. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

It is useful to consider slight fattenings and shrinkings of collections of K-boxes. The next result and its proof utilize the notions of K-box latitude from Definition 3.9, and of the K-box fattening $\overline{A}[2]$ and shrinking $\overline{A}[-2]$ from Definition 3.5.

Lemma 4.24. *Suppose that* Fine $\subset$ Cluster$_{[u,u+3r]}$ *occurs.*

1. *A K-box lying in $\overline{\text{DownClust}}[2]$ has latitude at most $5d/2$.*

2. *Such a box lying in $\overline{\text{Tropics}}[-2]$ has latitude at least $-(2K)^{-1}(A + Kd^{1/2})$.*

3. *For $(A - 5Kd^{1/2})(2Kr)^{-1} \geq i \geq -5d/2$, the number of K-boxes of latitude $-i$ lying in $\overline{\text{DownClust}}[2]$ is at most $5^d\left(6Kd^{1/2} + 2Kr + |i|\right)^4 R$.*

4. *For $(A - 5Kd^{1/2})(2Kr)^{-1} \geq j \geq -5d/2$, the number of K-boxes of latitude j lying in $\overline{\text{UpClust}}[2]$ is at most $5^d\left(6Kd^{1/2} + 2Kr + |j|\right)^4 R$.*

Proof: (1). Since DownClust is a subgraph of L_h, $v \in V(\text{DownClust})$ satisfies $v \cdot \boldsymbol{\ell} \leq h$. If x is the central index of any element $\overline{\text{DownClust}}$, we thus have $x \cdot \boldsymbol{\ell} \leq h + Kd^{1/2}$; and thus, if w is the central index of any element $B \in \overline{\text{DownClust}}[2]$ (so that $\max_{i \in [\![1,d]\!]} |x_i - w_i| \leq 4K$), $w \cdot \boldsymbol{\ell} \leq h + K(1+2^2)d^{1/2}$. We find then that

$$\text{Lat}(B) = \left\lfloor \tfrac{w \cdot \boldsymbol{\ell} - h}{2Kr} \right\rfloor \leq \frac{5Kd^{1/2}}{2Kr} \leq 5d/2$$

where $r = e_d \cdot \boldsymbol{\ell} \geq d^{-1/2}$ by axes ordering (1).

(2). Let $2Kw + [\![-K,K]\!]^d \in \overline{\text{Tropics}}[-2]$. Let B be the K-box with central vertex $2Kw - 4Ke_d$. Then $B \in \overline{\text{Tropics}}$, so B intersects U_{h-A}, with the argument for Lemma 2.2(1) yielding that the height of every vertex in U_{h-A} being at least $h - A - r$. Thus, $(2Kw - 4Ke_d) \cdot \boldsymbol{\ell} \geq h - A - r - Kd^{1/2}$. Writing $B' = 2Kw + [\![-K,K]\!]^d$,

$$\text{Lat}(B') = \left\lfloor \tfrac{w \cdot \boldsymbol{\ell} - h}{2Kr} \right\rfloor \geq \frac{(2K-1)r - A - Kd^{1/2}}{2Kr} \geq -\frac{A + d^{1/2}K}{2K}$$

where we used $K \geq 1/2$ and $r \leq 1$ in the latter bound.

(3). We take $i \geq 0$; the case where $i \in [\![-\lfloor 5d/2 \rfloor, -1]\!]$ is similar. Let $B = 2Kw + [\![-K, K]\!]^d$ be a K-box of latitude $-i$ lying in $\overline{\text{DownClust}}[2]$. Let $x \in V(\text{DownClust}) \cap V(\text{BigBox}(B))$. Note that $x \cdot \boldsymbol{\ell} - h = 2Kw \cdot \boldsymbol{\ell} - h + \phi$ with $|\phi| \leq 5Kd^{1/2}$. Since $2Kw \cdot \boldsymbol{\ell} \in h + 2Kr \cdot [i, i+1)$, we find that

$$x \cdot \boldsymbol{\ell} - h \in \left[2Kri - 5Kd^{1/2}, 2Kr(i+1) + 5Kd^{1/2} \right).$$

We claim that $h \geq x \cdot \boldsymbol{\ell} \geq h - A$. The first bound is due to $x \in V(\text{DownClust})$. To confirm the second, note that $\left\lfloor \frac{2Kw \cdot \boldsymbol{\ell} - h}{2Kr} \right\rfloor = -i$ implies that $2Kw \cdot \boldsymbol{\ell} - h \geq h - 2Kri$. But then $x \cdot \boldsymbol{\ell} \geq 2Kw \cdot \boldsymbol{\ell} - 5Kd^{1/2}$ and

$$x \cdot \boldsymbol{\ell} \geq h - 2Kri - 5Kd^{1/2}, \tag{21}$$

so that the sought bound $x \cdot \boldsymbol{\ell} \geq h - A$ follows from the hypothesised condition $i \leq \frac{A - 5Kd^{1/2}}{2Kr}$. For $j \in [\![1, A]\!]$, set $\Phi_j = \left\{ x \in V(\text{DownClust}) : x \cdot \boldsymbol{\ell} \in (h - j, h - j + 1] \right\}$. In this way, condition $\mathbf{T}_2'(R)$ in Fine_h asserts that $|\Phi_j| \leq j^3 R$ for $j \in [\![1, A]\!]$. Since $h \geq x \cdot \boldsymbol{\ell} \geq h - A$, there exists $j \in [\![1, A]\!]$ such that $x \in \Phi_j$. Moreover, if $x \in \Phi_j$, then $x \cdot \boldsymbol{\ell} \leq h - j + 1$; alongside (21), we find that $j \leq 5Kd^{1/2} + 2Kri + 1$. Which is to say: x is necessarily a member of a set Φ_j of cardinality at most $j^3 R$ for an index j which is at least one and at most $5Kd^{1/2} + 2Kr + i$. Moreover, given x, there are at most 5^d compatible choices of the K-box B. We thus see that the number of choices of B is at most

$$5^d \sum_{j=1}^{6kd^{1/2} + 2Kr + i} j^3 R,$$

whose right-hand side is bounded above by $5^d \left(6Kd^{1/2} + 2Kr + i \right)^4 R$. That is, we have found that

$$\left| \overline{N}_{-i} \cap \overline{\text{DownClust}}[2] \right| \leq 5^d \left(6Kd^{1/2} + 2Kr + i \right)^4 R.$$

(4). Evident changes of notation in the preceding proof yield the argument. $\qquad\square$

Definition 4.25. Suppose given a finite path $P = \left(P_i : i \in [\![0, k]\!] \right)$ whose ending point P_k coincides with the starting point Q_0 of a path $Q = \left(Q_i : i \in [\![0, m]\!] \right)$. Here, $k \in \mathbb{N}$ and $m \in \mathbb{N} \cup \{\infty\}$, with $[\![0, \infty]\!] = \mathbb{N}$ understood. The concatenation of P and Q is the sequence $P \circ Q : [\![0, k + m]\!] \to \mathbb{Z}^d$ with $(P \circ Q)_i = P_i$ for $i \in [\![0, k]\!]$ and $(P \circ Q)_i = Q_{i-k}$ for $i \in [\![k, k + m]\!]$.

Proof of Proposition 4.20. Recall from Definition 4.8 that $\mathsf{Fine} \subset \mathsf{Cluster}_{[u, u+3r]}$ and that $V \in \mathbb{Z}^d \cap F_u \cap B_{u+3r}$ is a random head location specified by Definition 4.16 when $\mathsf{Cluster}_{[u, u+3r]}$ occurs. Let $v \in \mathbb{Z}^d \cap F_u \cap B_{u+3r}$. In the ensuing paragraphs, we will specify some aspects of the geometry of the cluster under the event $\mathsf{Fine} \cap \{V = v\}$; we thus suppose that this event occurs. Recall that MidClust contains an open path that connects a vertex in Top(DownClust) to a vertex in Bed(UpClust). Let Q denote the lexicographically minimal among the set of such open paths. Suppose that Q runs from $w \in \text{Top}(\text{DownClust})$ to $w' \in \text{Bed}(\text{UpClust})$. Let Q^- denote the lexicographically minimal path in DownClust from $\text{base}(\mathcal{C}_{u,u+3r}) = 0 \in \mathbb{Z}^d$ to w; and let Q^+ denote the similarly minimal path in UpClust from w' to $\text{head}(\mathcal{C}_{u,u+3r}) = \text{head}(\text{UpClust})$. Denote by P_1 the concatenation $Q^- \circ Q \circ Q^+$. Note that P_1 is an open path in the cluster $\mathcal{C}_{u,u+3r}$ that begins at $\text{base}(\mathcal{C}_{u,u+3r})$ and ends at $\text{head}(\mathcal{C}_{u,u+3r})$ and that runs successively through the down-, mid- and up-clusters.

Prefix to P_1 an infinite path P_0 of consecutive positively oriented e_1-parallel edges that runs directly into $\text{base}(\mathcal{C}_{u,u+3r})$, in a direction that we may call 'rightwards' (and which is increasing for height

because $e_1 \cdot \boldsymbol{\ell} \geq 0$ by axes ordering). Postfix such a path P_2 that starts at head($\mathcal{C}_{u,u+3r}$) and that also runs directly rightwards. The resulting path $P = P_0 \circ P_1 \circ P_2$ is a bi-infinite path in $\mathbb{Z}^d$ that runs directly rightwards through heights at most zero into an open subpath that runs from base($\mathcal{C}_{u,u+3r}$) and takes a passage through MidClust to end at head($\mathcal{C}_{u,u+3r}$), and then pursues a directly rightward trajectory through heights at least u. We now wish to specify $P[K]$ to be the K-box valued bi-infinite path that visits the K-boxes consecutively visited by P. Since K-boxes overlap (albeit only on their boundaries), this specification needs to be clarified. Recall that elements $x = 2Ku$, $u \in \mathbb{Z}^d$, of the K-box index space index K-boxes $x + [\![-K, K]\!]^d$.

Definition 4.26. Let ind $: \mathbb{Z}^d \to \mathbb{Z}^d$ be the function that sends any element of $2Ku + [\![-K, K-1]\!]^d$ to the index $2Ku$. To the bi-infinite path $P : \mathbb{Z} \to \mathbb{Z}^d$, we may associate the mapping $\hat{P}$ from $\mathbb{Z}$ to the K-box index space via $\hat{P}(i) = \text{ind}\big(P(i)\big)$ for $i \in \mathbb{Z}$. Consider the bi-infinite path S in K-boxes $i \to \hat{P}(i) + [\![-K, K]\!]^d$, $i \in \mathbb{Z}$. Let $R \subset \mathbb{Z}$ denote the set of indices i at which $S(i) \neq S(i-1)$ (and note that R has infinite intersection with both $\mathbb{N}$ and $\mathbb{Z} \setminus \mathbb{N}$). Let $\rho : \mathbb{Z} \to R$ be an arbitrary increasing bijection. We then set $P[K]$ to be the bi-infinite path in K-boxes $i \to S(\rho(i))$, $i \in \mathbb{Z}$.

Note that indeed $P[K]$ is a nearest-neighbour path in K-boxes. This path visits MidClust in correspondence with the subpath Q of P. It is Lemma 4.22(3) that permits the next definition.

Definition 4.27. Let B_1 denote the last element of Surface$_K$ along $P[K]$ before the first visit of $P[K]$ to MidClust; and let B_2 denote the first element of Surface$_K$ encountered along $P[K]$ after the final visit of this path to MidClust.

Let $\mathsf{Long}^+ \subseteq \mathsf{Fine} \cap \{V = v\}$ denote the event that Surface$_K$ contains a K-box path that starts at a box in $\overline{\text{Equator}}[2]$ and that ends at a box in $\overline{\text{NorthZone}}[2]$ with no element in the path belonging to $\overline{\text{UpClust}}[2]$. Let $\mathsf{Long}^- \subseteq \mathsf{Fine} \cap \{V = v\}$ denote the event that Surface$_K$ contains a K-box path that also starts at a box in $\overline{\text{Equator}}[2]$ and that ends at a box that intersects $\overline{\text{SouthZone}}[2]$, with no element in the path belonging to $\overline{\text{DownClust}}[2]$.

Set $\mathsf{Long} = \mathsf{Long}^+ \cup \mathsf{Long}^-$ and $\mathsf{Short} = \mathsf{Fine} \cap \{V = v\} \cap \mathsf{Long}^c$.

Lemma 4.28. *Suppose that* Short *occurs. At least one of two alternatives holds:*

1. *there exists a K-box in $\overline{\text{Tropics}}[-2]$ that lies in both $\overline{\text{DownClust}}[2]$ and $\overline{\text{UpClust}}[2]$; or*

2. *there exists a K-box path Ψ in Surface$_K \cap \overline{\text{Tropics}}[-2] \cap \overline{\text{DownClust}}[2]^c \cap \overline{\text{UpClust}}[2]^c$ whose first element is a nearest neighbour of an element of $\overline{\text{DownClust}}[2]$ and whose final element is a nearest neighbour of an element of $\overline{\text{UpClust}}[2]$.*

Lemma 4.28 will permit us to specify a K-box path Ψ that plays the counterpart role in rigorous analysis that the path Connect through virgin territory did in heuristic discussion in Section 3.1. Indeed, when the latter alternative in the lemma occurs, we will set Ψ equal to some thus named path of minimum length, chosen in some definite manner, and $D[\Psi]$ and $U[\Psi]$ for the elements of $\overline{\text{DownClust}}[2]$ and $\overline{\text{UpClust}}[2]$ that neighbour the start and end of Ψ. Note also that, for the event in question, we admit the possibility that Ψ is empty. In this case, the condition amounts to the existence of nearest neighbour K-boxes $D[\Psi]$ and $U[\Psi]$ in $\overline{\text{Tropics}}[-2]$ such that $D[\Psi] \in \overline{\text{DownClust}}[2]$ and $U[\Psi] \in \overline{\text{UpClust}}[2]$. In fact, even the event in the first alternative in Lemma 4.28 may be subsumed in this framework, because we may take $D[\Psi]$ and $U[\Psi]$ to be equal to the K-box posited in this case.

Proof of Lemma 4.28. Recall the K-boxes B_1 and B_2 in Definition 4.27. Note that B_1 and B_2 are elements of Surface_K. Suppose that these two boxes are equal: $B_1 = B_2 = B$. We will argue that the first alternative presented in the lemma holds. Since B_1 intersects either P_0 or DownClust, it also intersects L_h. Likewise, since B_2 intersects either P_2 or UpClust, it intersects U_h. Thus B intersects Equator. Since $u - h$ and h exceed K, B can intersect neither P_2 (a path at height at least u) nor P_0 (a path at height at most zero). Thus $B = B_1 \in \overline{\text{DownClust}} \subset \overline{\text{DownClust}}[2]$ and $B = B_2 \in \overline{\text{UpClust}} \subset \overline{\text{UpClust}}[2]$. And indeed B must lie in $\overline{\text{Tropics}}[-2]$, because $A - 3K > K$. Thus we confirm the first alternative in the lemma.

Suppose now that B_1 and B_2 are not equal. Since Surface_K is a connected set, we may find a K-box path valued in Surface_K that starts at B_1 and ends at B_2. Select such a path Γ of minimal length, chosen in some definite manner. We will argue that there exists a box B_3 on Γ that lies in $\overline{\text{DownClust}}[2] \cap \overline{\text{Tropics}}[-2]$ such that the subpath of Γ strictly beyond B_3 makes no visit to $\overline{\text{DownClust}}[2]$ and does not leave $\overline{\text{Tropics}}[-2]$ via its south side. To establish this, note that if Γ begins (at B_1) to the south of $\overline{\text{Tropics}}[-2]$, or if it leaves $\overline{\text{Tropics}}[-2]$ via the south side, then it will, after reentering $\overline{\text{Tropics}}[-2]$, make a visit to $\overline{\text{DownClust}}[2]$ before ending at B_2: this is because we are supposing that Short occurs and B_2 lies in, or to the north of, $\overline{\text{Equator}}[2]$. If we take B_3 to be the last visit of Γ to $\overline{\text{DownClust}}[2]$, it thus has the sought property.

Now consider the reversal $\Gamma^\downarrow$ of Γ from B_2, stopped at B_3. We assert that there is a box B_4 in $\Gamma^\downarrow$ that lies in $\overline{\text{UpClust}}[2] \cap \overline{\text{Tropics}}[-2]$ such that the subpath of $\Gamma^\downarrow$ strictly beyond B_4 makes no visit to $\overline{\text{UpClust}}[2]$ and does not leave $\overline{\text{Tropics}}[-2]$ via its north side. Indeed, and similarly to above, if $\Gamma^\downarrow$ begins to the north of $\overline{\text{Tropics}}[-2]$, or if it leaves $\overline{\text{Tropics}}[-2]$ via the north side, then it will, after reentering $\overline{\text{Tropics}}[-2]$, make a visit to $\overline{\text{UpClust}}[2]$ before ending at B_3. Naturally we then take B_4 to be the last visit of $\Gamma^\downarrow$ to $\overline{\text{UpClust}}[2]$.

It may be that $B_3 = B_4$: the first case in the lemma then holds. If not, we take Ψ to be the subpath of Γ strictly between B_3 and B_4. This path remains in $\overline{\text{Tropics}}[-2]$ and thus verifies the second case in the lemma. This completes the proof of Lemma 4.28. $\qquad\square$

The projected distance d_{proj} between two K-boxes B_1 and B_2 is given by translating the two boxes in the e_d-direction so that they have latitude zero, and setting $d_{\text{proj}}(B_1, B_2)$ to be the ℓ_1-distance between these translates.

Let $i, j \in \mathbb{Z}$ be at least $-(d/2 + 2)$ and let $k \in \mathbb{N}$. We set $J_v[i, j, k]$ to be the event that Short occurs alongside the latter alternative in Lemma 4.28, with the K-box path Ψ in that alternative being such that $D[\Psi]$ has latitude $-i$; $U[\Psi]$ has latitude j; and $d_{\text{proj}}\big(D[\Psi], U[\Psi]\big) = k$. The event that the first alternative in Lemma 4.28 obtains, in which $D[\Psi]$ and $U[\Psi]$ may be viewed as being equal, can in fact be included in this classification, by taking $k = 0$ and $i = -j$.

Lemma 4.29. *With $\epsilon \in \big(0, 1/(2d-1)\big)$ denoting a parameter that may be chosen arbitrarily small by making a sufficiently high choice of $K \in \mathbb{N}$, we have that, for $\omega \in \mathsf{Fine}$,*

1. and for $v \in \mathbb{Z}^d \cap F_u \cap B_{u+3r}$, $i, j \in [\![-\lfloor d/2 + 2\rfloor, \lfloor A/K \rfloor]\!]$ and $k \in \mathbb{N}$,

$$\mathbb{R}^v_{[\text{UpClust}], \text{DownClust}}\big(J_v[i, j, k]\big) \leq 5^d \big(6Kd^{1/2} + 2Kr + |i|\big)^4 R(2d-1)^{|i+j|+k} \epsilon^{|i+j|+k} \, ;$$

2. for such (i, j, k), the set of $v \in \mathbb{Z}^d \cap F_u \cap B_{u+3r}$ for which $\mathbb{R}^v_{[\text{UpClust}], \text{DownClust}}\big(J_v[i, j, k]\big)$ is

non-zero has cardinality at most

$$50^d \big(6Kd^{1/2} + 2Kr + |i|\big)^4 \big(6Kd^{1/2} + 2Kr + |j|\big)^4 \big(2(|i+j|+k)+1\big)^d R^2 K^d;$$

3. and, for a universal constant $C > 0$,

$$\mathbb{R}^v_{[\mathrm{UpClust}],\mathrm{DownClust}}(\mathsf{Long}) \le C(Lu^d K^{-1} + 2)^{d-1}\big((2d-1)\epsilon\big)^{A/K-6}.$$

Proof: (1). For given $\omega \in \mathsf{Fine}$ and $v \in \mathbb{Z}^d \cap F_u \cap B_{u+3r}$, the set of nearest-neighbour edges in $\mathbb{Z}^d$ is partitioned in Definition 4.17 into three disjoint subsets: discovered open edges; discovered closed edges; and undiscovered edges. We have $R^v_{[\mathrm{UpClust}],\mathrm{DownClust}}(e) = \omega'(e)$ for any undiscovered edge e. Recall also, from the final sentence of Definition 4.17, that

$$\mathbb{R}^v_{[\mathrm{UpClust}],\mathrm{DownClust}}\big(J_v[i,j,k]\big) = \mathbb{P}'\Big(\omega' \in \Omega' : R^v_{[\mathrm{UpClust}],\mathrm{DownClust}}(\omega') \in J_v[i,j,k]\Big).$$

It is perhaps useful to recall that, under the probability measure $\mathbb{R}^v_{[\mathrm{UpClust}],\mathrm{DownClust}}(\cdot)$, $\mathrm{UpClust}^v$, when it is viewed, as it will be, as a function of ω, is a well-defined and non-random subset of edges, specified by Definition 4.16. Moreover, when a sample of this probability measure realizes $V = v$, the random variable UpClust defined by the sample equals $\mathrm{UpClust}^v = \mathrm{UpClust}^v(\omega)$. In what follows, we will consider this sample when various subsets of the event $\{V = v\}$ occur, so that UpClust is well-defined and takes the just stated value.

Call a K-box B *undiscovered* if every nearest-neighbour edge in the big box with Centre B is undiscovered for the given $\omega \in \mathsf{Fine}$ and $v \in \mathbb{Z}^d \cap F_u \cap B_{u+3r}$. If the configuration $R^v_{[\mathrm{UpClust}],\mathrm{DownClust}}$ realizes $\mathsf{Cluster}_{[u,u+3r]} \cap \{V = v\}$, then any box in $\mathrm{Surface}_K$ that lies in $\overline{\mathrm{Tropics}}[-2]$—recall Definition 4.5—but in neither $\overline{\mathrm{UpClust}}[2]$ nor $\overline{\mathrm{DownClust}}[2]$ is undiscovered: indeed, in the case that $R^v_{[\mathrm{UpClust}],\mathrm{DownClust}}$ realizes $\mathsf{Cluster}_{[u,u+3r]}$, any discovered edge in Tropics is incident to a vertex in either UpClust or DownClust; since no edge in a big box with Centre equal to such a K-box is thus incident, each such edge is undiscovered.

When $R^v_{[\mathrm{UpClust}],\mathrm{DownClust}}(\omega') \in J_v[i,j,k]$, the K-box path Ψ offered by Lemma 4.28 lies in $\mathrm{Surface}_K \cap \overline{\mathrm{Tropics}}[-2]$ but in neither $\overline{\mathrm{UpClust}}[2]$ nor $\overline{\mathrm{DownClust}}[2]$. Thus, the K-boxes along this path are undiscovered. Lemma 4.23 implies that the K-boxes along Ψ are bad. The configuration $R^v_{[\mathrm{UpClust}],\mathrm{DownClust}}(\omega')$ belongs to $J_v[i,j,k]$ and thus realizes Fine. The K-box $D[\Psi]$ lies in $\overline{\mathrm{Tropics}}[-2]$ at latitude $-i$ and intersects $\overline{\mathrm{DownClust}}[2]$. Lemma 4.24(1,3) thus implies that $D[\Psi]$ may adopt one of at most $5^d\big(6Kd^{1/2} + 2Kr + |i|\big)^4 R$ locations. By Lemma 3.4, we see that the path Ψ, which begins at a neighbour of $D[\Psi]$, and has length $|i+j|+k$ (the absolute value is operative when one or other of i and j lies in $[\![-\lfloor 3d/2 \rfloor, -1]\!]$), lies in the set of open vertices of a site percolation on $\mathbb{Z}^d$ of parameter $\epsilon > 0$, where this parameter may be chosen to be arbitrarily small provided that a sufficiently high choice of $K \in \mathbb{N}$ is made. And indeed this percolation is independent of the conditioning that specifies the law $\mathbb{R}^v_{[\mathrm{UpClust}],\mathrm{DownClust}}(\cdot)$: this conditioning is measurable with respect to the configuration in DownClust, UpClust, NorthZone and SouthZone, a region that no big box with Centre in Ψ intersects, so that Lemma 3.6 delivers the claimed independence. The probability $p(\ell, \epsilon)$ that such a site percolation contains an open path of length $\ell \in \mathbb{N}$ emanating from the origin is readily seen to satisfy

$$p(\ell, \epsilon) \le (2d-1)^{\ell-1}\epsilon^\ell. \tag{22}$$

Thus, we see that, with $\ell = |i + j| + k$,

$$\mathbb{R}^v_{[\text{UpClust}],\text{DownClust}}\big(J_v[i,j,k]\big) \leq 5^d\big(6Kd^{1/2} + 2Kr + |i|\big)^4 R(2d-1)(2d-1)^{\ell-1}\epsilon^\ell.$$

Using $K \geq 1$, we obtain Lemma 4.29(1).

(2). We seek an upper bound on the cardinality of the set of $v \in \mathbb{Z}^d \cap F_u \cap B_{u+3r}$ for which $\omega' \in \Omega'$ exists such that $R^v_{[\text{UpClust}],\text{DownClust}}(\omega') \in J_v[i,j,k]$. It is useful to partition the set of $v \in \mathbb{Z}^d \cap F_u \cap B_{u+3r}$ into $(2K)^d$ classes, with members of a given class differing by an element of $(2K)^d \cdot \mathbb{Z}^d$. We do so because, as v varies over the elements in a given class, the collection $\overline{\text{UpClust}}[2]$ shifts in a straightforward fashion, by vector displacement.

In any case, if such ω' is to exist, then there must exist a pair $\big(D(\Psi), U(\Psi)\big)$ of K-boxes, the former in $\overline{\text{DownClust}}[2]$ at latitude $-i$, and the latter in $\overline{\text{UpClust}}[2]$ at latitude j, at K-box nearest-neighbour distance at most $|i+j| + k + 1$. The number of admissible pairs of K-boxes is at most

$$\alpha_1 = 5^d\big(6Kd^{1/2} + 2Kr + |i|\big)^4 R \cdot 5^d\big(6Kd^{1/2} + 2Kr + |j|\big)^4 R$$

by Lemma 4.24(1,2,3,4). When such a pair is given, the variation of v within a given class yields various outcomes at which the elements of the pair lie at K-box distance at most $|i+j| + k + 1$. An upper bound on the number of such v is $\alpha_2 = \big(2(|i+j| + k) + 1\big)^d$, because this is the number of K-boxes at such a distance from a given K-box. The number of classes into which the set of v is divided is at most $\alpha_3 = (2K)^d$. The product $\alpha_1 \alpha_2 \alpha_3$ is the upper bound in Lemma 4.29(2).

(3). Under $\text{Fine} \cap \text{Long}$, Ψ is a path in bad boxes of length at least $A/K - 6$ that begins at a box whose distance (in the lattice $\mathbb{Z}^d$) from the vertical line through $0 \in \mathbb{Z}^d$ is at most $Lu^d + 2K$, since this box $*$-neighbours an element of $\overline{\mathcal{C}}_u$, and $|V(\mathcal{C}_{u,u+3r})| \leq Lu^d$ by condition $\mathbf{G}(L)$. Elements in the path Ψ lie in $\overline{\text{Tropics}}[-2]$ and in neither $\overline{\text{UpClust}}[2]$ nor $\overline{\text{DownClust}}[2]$, so the constituent boxes are undiscovered for the given $\omega \in \text{Fine}$. The probability of such a path emanating from a given box is at most $\big((2d-1)\epsilon\big)^{A/K-6}$ by Lemma 3.4 and the bound (22). Lemma 4.29(3) follows from $\epsilon < (2d-1)^{-1}$. $\qquad\square$

Note that $\mathbb{R}^v_{[\text{UpClust}],\text{DownClust}}(\text{Fine})$ equals $\mathbb{R}^v_{[\text{UpClust}],\text{DownClust}}(\text{Fine}, V = v)$ when v belongs to the set $[\![-Lu^d, Lu^d]\!]^{d-1} \times \mathbb{Z}$; and equals zero when v is any other element of $\mathbb{Z}^d$. This justifies the first inequality as we write

$$\sum_{v \in \mathbb{Z}^d \cap F_u \cap B_{u+3r}} \mathbb{R}^v_{[\text{UpClust}],\text{DownClust}}(\text{Fine})$$

$$\leq \sum_{v \in \mathbb{Z}^d \cap F_u \cap B_{u+3r}} \sum_{i,j,k} \mathbb{R}^v_{[\text{UpClust}],\text{DownClust}}\big(J_v[i,j,k]\big) + \sum_{v \in [\![-Ln^d, Ln^d]\!]^{d-1} \times \mathbb{Z}} \mathbb{R}^v_{[\text{UpClust}],\text{DownClust}}(\text{Long})$$

$$\leq 250^d R^3 K^d \sum_{i,j,k} \big(6Kd^{1/2} + 2Kr + |i|\big)^8 \big(6Kd^{1/2} + 2Kr + |j|\big)^4$$

$$\times \big(2(|i+j| + k) + 1\big)^d (2d-1)^{|i+j|+k} \epsilon^{|i+j|+k}$$

$$+ (2Lu^d + 1)^{d-1} C(Lu^d K^{-1} + 2)^{d-1}\big((2d-1)\epsilon\big)^{A/K-6},$$

where Lemma 4.29(1,2) is invoked to bound above the first term in the latter inequality, and Lemma 4.29(3) offers control on the second such term. The final sum is over $i, j \geq -3$ and $k \geq 0$; it

converges due to $(2d-1)\epsilon < 1$. Recall that, in Proposition 4.20, we suppose that A/K is bounded below by a suitably high multiple of $\log u$. We find then that $\sum_{v \in \mathbb{Z}^d \cap F_u \cap B_{u+3r}} \mathbb{R}^v_{[\mathrm{UpClust}], \mathrm{DownClust}}(\mathsf{Fine})$ is bounded above, uniformly in $\omega \in \mathsf{Fine}$. This completes the proof of Proposition 4.20. $\square$

4.3 Controlling the damage when the single slide fails

The event Catastrophe occurs when ResidualClust contains a vertex whose height differs from h by more than $3A$; which is to say, there exists a vertex in ResidualClust that lies in none of NorthZone, Tropics and SouthZone. This tall structure prevents renewal levels at all the heights that it occupies. This section is devoted to proving the next result, which shows the rarity of this untoward outcome of slide resampling.

Proposition 4.30. *For any positive R and K, there exist $c = c(R, K) > 0$ and $C = C(K, d)$ such that the condition that $A \geq C \log u$ implies that*

$$\mathbb{P}^{\mathsf{Fine}}\Big(\mathsf{Catastrophe} \,\Big|\, [\mathrm{UpClust}], \mathrm{DownClust} \Big) \leq \exp\big\{ -cA \big\}.$$

The slide resample has been the centre of attention, but we have never explicitly defined it. In order to prove Proposition 4.30, it is useful to do so. The slide operation Slide at height h associates to each configuration $\omega \in \Omega$ a random configuration Slide(ω). Let the positive parameters R and K that specify the event Fine be given.

Definition 4.31. If $\omega \notin \mathsf{Fine}$, set Slide$(\omega) = \omega$. If $\omega \in \mathsf{Fine}$, we set Slide(ω) to be a random element of Ω whose law equals $\mathbb{P}^{\mathsf{Fine}}\big(\cdot \,\big|[\mathrm{UpClust}], \mathrm{DownClust}\big)$.

Lemma 4.32. *The probability space $(\Omega, \mathcal{B}, \mathbb{P})$ may be augmented in order to support the system of random variables $\big\{ \mathrm{Slide}(\omega) : \omega \in \Omega \big\}$ specified in Definition 4.31.*

Proof. The value of Slide(ω) may be chosen so that it is determined by the evaluation at ω of $([\mathrm{UpClust}], \mathrm{DownClust})$. There are only finitely many possible values for this pair. For each, the value of the concerned Slide(ω) may be set by sampling the probability measure in Lemma 4.18 with $B = \mathsf{Fine}$ and $([\mathfrak{u}], \mathfrak{d}) = \big([\mathrm{UpClust}](\omega), \mathrm{DownClust}(\omega)\big)$. $\square$

Definition 4.33. Let $\ell, \ell' \in \mathbb{Z}$ with $\ell \leq \ell'$. A connected subgraph S of the slab $\mathrm{Slab}_{\ell,\ell'}$ is called *crossing* for this slab if $V(S)$ contains an element in $\mathrm{VertBdry}_\ell$ and another in $\mathrm{VertBdry}_{\ell'}$.

A *north island* is a finite open connected component in NorthZone that is disjoint from UpClust. A north island is called *crossing* if it is crossing for NorthZone. A north island is called *meagre* if for at least one-quarter of north-zone integral height indices $i \in [\![h + A, h + 3A - 1]\!]$ there at most KR vertices in the north island with height in $[i, i+1)$. A north island is called *near* if the distance from the vertical line running through $0 \in \mathbb{Z}^d$ of all of its vertices is at most Lu^d.

A *south island* is a finite open connected component in SouthZone that is disjoint from DownClust. The definitions above may equally be made for south islands, with the replacements NorthZone $\to$ SouthZone, $h + A \to h - 3A$ and $h + 3A \to h - A$ made.

Let CrossMeagNearIsland denote the event that there exists a north or south island that is crossing, meagre and near.

Lemma 4.34. *There exist positive C and c such that, whenever the parameter $A \in \mathbb{N}$ appearing in Definition 4.5 satisfies $A \geq C \log u$,*

$$\mathbb{R}^v_{[\mathrm{UpClust}],\mathrm{DownClust}}\big(\mathsf{CrossMeagNearIsland}\big) \leq \exp\big\{ - cA\big\}$$

for any $v \in \mathbb{Z}^d \cap F_u \cap B_{u+3r}$ and $\omega \in \mathsf{Cluster}_{[u,u+3r]}$.

Let $([\mathfrak{u}], \mathfrak{d})$ denote a possible value of $([\mathrm{UpClust}], \mathrm{DownClust})$ as ω ranges over $\mathsf{Cluster}_{[u,u+3r]}$, and let $v \in \mathbb{Z}^d \cap F_u \cap B_{u+3r}$. Let Seed denote the set of $x \in \mathrm{VertBdry}_{h+A} \cap [\![-Lu^d, Lu^d]\!]^d$ such that x is an endpoint of an edge in NorthZone that is undiscovered given the data $([\mathfrak{u}], \mathfrak{d}, v)$ according to Definition 4.17. The vertex set of any near crossing island in NorthZone intersects Seed. (The set Seed is so named because an iterative procedure will grow a cluster from a starting point therein in the proof of the next result.)

Lemma 4.35. *There exists $c > 0$ such that, for $\omega \in \mathsf{Cluster}_{[u,u+3r]}$, $x \in D$ and $v \in \mathbb{Z}^d \cap F_u \cap B_{u+3r}$, the conditional probability under $\mathbb{R}^v_{[\mathrm{UpClust}],\mathrm{DownClust}}(\cdot)$ given that this value is adopted that x is an element in the vertex set of a meagre crossing island is at most e^{-cA}.*

Proof of Lemma 4.34. We sum the estimate in Lemma 4.35 over $x \in$ Seed to obtain the lemma with a relabelling of $c > 0$. $\qquad\square$

Proof of Lemma 4.35. It is perhaps helpful to recall the meaning of the sought assertion. We take $\omega \in \mathsf{Cluster}_{[u,u+3r]}$ given and consider the conditional probability

$$\mathbb{R}^v_{[\mathrm{UpClust}],\mathrm{DownClust}}\big(\mathsf{CrossMeagNearIsland}\big)$$
$$= \ \mathbb{P}'\Big(\omega' \in \Omega' : R^v_{[\mathrm{UpClust}],\mathrm{DownClust}}(\omega') \text{ realizes the event } \mathsf{CrossMeagNearIsland}\Big),$$

where ω determines the value of $([\mathrm{UpClust}], \mathrm{DownClust})$. The sought claim is that this probability decays exponentially in the north- and south-zone width A, independently of ω. To prove it, take given $x \in$ Seed. First an heuristic overview. Our task is to examine the north island that contains x—which is to say, the open connected component in NorthZone containing the vertex x under the auxiliary percolation ω'. We will iteratively form this component by locally realizing its edges in a process that begins with edges neighbouring x. If the resulting component is to cross NorthZone, the process must run for at least $2A$ steps. At each step, certain edges at the boundary of the thus realized cluster remain to be examined. If, at a given step, there are at most a given bounded number of such edges—call this the 'bounded case'—there is a uniformly positive probability that the process will end at this very step, because it may be that each such edge will turn out to be closed when it is realized. On the other hand, if there are a greater number of unexamined edges at the given step—'the unbounded case'—then it is likely that a large number of open edges will enter the forming cluster when they are examined. Thus, if the bounded case is more typical, the forming north island is likely to have died out before the $(2A)^{\mathrm{th}}$ step is reached, so that the island cannot be crossing. And if the unbounded case is more prevalent, then typically the growing island will often welcome in many new edges and its resulting form will not be meagre. In summary, an island can be crossing and meagre only by performing the rare feat of surviving for long without thriving.

To make this scheme precise, we specify the searching procedure in terms of some notation. Let $i \in [\![0, 2A-1]\!]$, and set S_i equal to the subgraph $\mathrm{Slab}_{h+A+i,h+A+i+1}$. Thus, the edge set of NorthZone $=$ $\mathrm{Slab}_{h+A,h+3A}$ is partitioned into the sets $\big\{E(S_i) : i \in [\![0, 2A-1]\!]\big\}$.

We will iteratively specify edge sets U_j and O_j for $j \in [\![0, T]\!]$. The value of the index j will rise until its terminal value T, which is at most $2A$. Write $\mathrm{UnDisc} \subset E(\mathbb{Z}^d)$ for the set of undiscovered edges in the sense of Definition 4.17, and $\mathcal{O} \subset E(\mathbb{Z}^d)$ for the set of open edges. Set

$$C_0 = \{x\} \ , \ U_0 = \{e \in \mathrm{UnDisc} \cap E(S_0) : e \text{ is incident to } x\} \text{ and } O_0 = U_0 \cap \mathcal{O} \,.$$

For $j \in [\![1, 2A]\!]$, set $T = j$ and terminate the procedure if $O_{j-1} = \emptyset$. In the opposing case, set C_{j-1} equal to the open connected component of x in $\mathrm{Slab}_{h+A,h+A+j} = \cup_{i=0}^{j-1} S_i$; then further set

$$U_j = \{e \in \mathrm{UnDisc} \cap E(S_j) : e \text{ is incident to } V(C_{j-1})\} \text{ and } O_j = U_j \cap \mathcal{O} \,.$$

Noting that $e \in E(\mathbb{Z}^d)$ belongs to S_i precisely when $(e) \cap F_{h+A+i} \neq \emptyset$ and $(e) \subset B_{h+A+i+1}$, we see that $e \in S_i$ implies that the minimum height of the endpoints of e lies in $[h + A + i, h + A + i + 1)$. We thus find that, whenever $O_j \neq \emptyset$,

$$\mathcal{C}_{u,u+3r} \text{ contains at least } |O_j|/(2d-1) \text{ vertices whose height lies in } [h + A + i, h + A + i + 1) \,. \tag{23}$$

Write $O = \cup_{i=1}^{T} O_i$. Note that $O \subseteq E(N_x)$, where N_x is the north island that contains the vertex x under $R^v_{[\mathrm{UpClust}],\mathrm{DownClust}}$. We analyse how O forms. Use the shorthand

$$D = 2(2d - 1)p^{-1}KR \,.$$

We say that step j is *fecund* if $|U_j| \geq D$ and *unproductive* if $|O_j| \leq D\frac{p}{2(2d-1)}$. Let T denote the number of steps for which the procedure runs. Let M denote the number of steps $j \in [\![0, T]\!]$ that are not fecund. To express the theme of the proof sketch in these terms, a meagre crossing island emanating from x must arise from a procedure with many unproductive steps. For such a procedure, there are either many steps that fail to be fecund, or many steps that are fecund and unproductive. Next is a result showing that each of these eventualities is a rarity.

Lemma 4.36.

1. *The random variable M is stochastically dominated by a geometric random variable X with law $\mathbb{P}(X = k) = \rho^{k-1}(1 - \rho)$ for $k \geq 1$, where $\rho = 1 - (1 - p)^{K-1}$.*

2. *There exists $q = q(p) \in (0, 1)$ such that, for each $k \in \mathbb{N}_+$, the probability that there are at least k fecund steps, and that, among the first k such, there are at least $k/2$ unproductive steps, is at most q^k.*

Proof: (1). At any given step that is not fecund, the size of the unexamined set is at most $D - 1$. With probability at least $(1 - p)^{D-1}$, each edge will be found to be closed when it is examined. The procedure will then terminate; otherwise, it will continue, with further independent examinations of unexamined edges. Thus M, the number of steps that are not fecund, is stochastically dominated by a geometric random variable that is at least one and whose success probability equals $(1 - p)^{D-1}$.

(2). The probability that a given fecund step is unproductive is a constant c that is strictly less than one-half. These outcomes are independent, so that, if the list of fecund steps is recorded, and the unproductive elements are marked, the collection of marked indices is stochastically dominated by a Bernoulli-c sequence truncated at a random time. From this, Lemma 4.36(2) follows. $\qquad \square$

We *claim* that, if x is a vertex in a crossing island, then the procedure duration T is at least $2A-1$. To see this, let m_i denote the maximum height of a vertex in C_i. Note that $x \cdot \boldsymbol{\ell} \leq h + A$ by $x \in \mathrm{VertBdry}_{h+A}$ and Lemma 2.2(1); thus, $m_0 \leq h + A$. Note that $m_{i+1} - m_i \leq r$, with $r = \max_{1 \leq k \leq d} |e_k \cdot \boldsymbol{\ell}|$ so that $m_{i+1} - m_i \leq 1$. If x is a vertex in a crossing island, then this island intersects $\mathrm{VertBdry}_{h+3A}$, so that $m_T \geq h + 3A - r$ by Lemma 2.2(1). We find then that $T \geq 2A/r - 1$, so that the claim follows from $r \leq 1$.

The probability that there are among the procedure's steps at least A that are not fecund is at most $\mathbb{P}(M \geq A) \leq \rho^{A-1}$ by Lemma 4.36(1). If x is a vertex in a crossing island and there are fewer than A steps that are not fecund, it must be, in view of $T \geq 2A - 1$, that there are at least A fecund steps. The probability that there are among these fecund steps at least $A/2$ that are unproductive is at most q^A by Lemma 4.36(2). Moreover, (23) shows that, when there are at least $A/2$ steps that are not unproductive, there are at least $A/2$ values $i \in [\![h + A, h + 3A - 1]\!]$ such that the crossing island of which x is an element has more than $D \frac{p}{2(2d-1)} = KR$ vertices whose height lies in $[i, i+1)$. In this event, this crossing island is thus not meagre. Making a suitably small choice of $c > 0$ completes the proof of Lemma 4.35. $\qquad\square$

Proof of Proposition 4.30. We will first argue that

$$\mathsf{Fine} \cap \mathsf{Catastrophe} \subseteq \mathsf{CrossMeagNearIsland}\,. \tag{24}$$

Indeed, suppose that the left-hand event occurs. Since $\mathsf{ResidualClust}$ contains a vertex v with $|v \cdot \boldsymbol{\ell} - h| > 3A$, Lemma 4.3(2) implies that there exists an open path in $\mathsf{ResidualClust}$ that lies in one or other of $\mathsf{NorthZone}$ or $\mathsf{SouthZone}$ and that crosses the slab in question between its lower and upper sides. The open path thus lies in a crossing island. This island is a subset of $\mathsf{ResidualClust}$ and thus of the well $\mathcal{C}_{u,u+3r}$. Every vertex in this island lies at distance at most Lu^d from the vertical line through $0 \in \mathbb{Z}^d$, because $|V(\mathcal{C}_{u,u+3r})| \leq Lu^d$ (by condition $\mathbf{G}(L)$ of Fine) and contains $0 \in \mathbb{Z}^d$. Thus, the island is near. Since Fine occurs, this event's subtropical conditions $\mathbf{S}_1'(R)$ and $\mathbf{S}_2(R)$—recall Definition 4.8—ensure that the island must also be meagre. We thus see that $\mathsf{CrossMeagNearIsland}$ occurs, so that (24) has been verified.

From (24) and Lemma 4.18, and using the notation (17), we note that, for $\omega \in \mathsf{Fine}$,

$$\mathbb{P}^{\mathsf{Fine}}_{[\mathsf{UpClust}],\mathrm{DownClust}}\big(\mathsf{Catastrophe}\big)$$
$$\leq \frac{\sum_{v \in [\![-Lu^d, Lu^d]\!]^{d-1}} \mathbb{R}^v_{[\mathsf{UpClust}],\mathrm{DownClust}}\big(\mathsf{CrossMeagNearIsland} \cap \mathsf{Fine}\big)}{\sum_{v \in \mathbb{Z}^d \cap F_u \cap B_{u+3r}} \mathbb{R}^v_{[\mathsf{UpClust}],\mathrm{DownClust}}\big(\mathsf{Fine}, V = v\big)}\,,$$

where the summand is thus restricted in the right-hand numerator because summands indexed by other values of $v \in \mathbb{Z}^d \cap F_u \cap B_{u+3r}$ report the probability of an event that entails that $\mathsf{UpClust}$ contains a vertex whose distance from the vertical line containing the origin exceeds Lu^d, which event is inconsistent with the global condition $\mathbf{G}(L)$ in Fine. Proposition 4.30 then follows from Lemmas 4.19 and 4.34. $\qquad\square$

4.4 The iterated slide: fast decay for inter-renewal distance

Here we prove Theorem 1.9(2). To do so, we will iterate the slide resample that led to Theorem 1.9(1). In analysing the steps of the iteration, we will not invoke Theorem 1.9(1) but rather

the underlying Propositions 4.11 and 4.30, which show that each step often acts favourably with a negligible probability of catastrophic failure.

The level $h \in [0, u]$ has been given, and we have often omitted it in describing the slide resample. Now, however, we will treat h as variable, and indicate its role by placing it in a subscript of relevant notation, in usages such as Fine_h, $\mathsf{UpClust}_h$, Slide_h and $\mathsf{Catastrophe}_h$. In the cases where the subscript role is already reserved, namely $\mathsf{Cluster}_{[u,u+3r]}$ and $\mathcal{C}_{u,u+3r}$, this event and random variable have no h-dependence, and it should be understood that these notions remain unaltered.

For convenience, recall that Fine_h from Definition 4.8 occurs when $\mathsf{Cluster}_{[u,u+3r]}$ does and

1. for $i \in [\![-A, -1]\!] \cup [\![1, A]\!]$, $\left| \left\{ v \in V(\mathcal{C}_{u,u+3r}) : v \cdot \boldsymbol{\ell} \in [h+i, h+i+1) \right\} \right| \leq i^3 R$;

2. for at least one-quarter of $i \in I$, $\left| \left\{ v \in V(\mathcal{C}_{u,u+3r}) : v \cdot \boldsymbol{\ell} \in [h+i, h+i+1) \right\} \right| \leq KR$, when $I = [\![h+A, h+3A-1]\!]$; and likewise when $I = [\![h-3A, h-A-1]\!]$; and

3. $|V(\mathcal{C}_{u,u+3r})| \leq Lu^d$.

Let $h \in [0, u]$. The random operator Slide_h has been specified in Definition 4.31. This operator maps the event $\mathsf{Cluster}_{[u,u+3r]}$ to itself and preserves the data ([UpClust], DownClust). If ω has the law $\mathbb{P}\left(\cdot \,|\, \mathcal{C}_{u,u+3r} \cap \mathsf{Fine}_h \right)$, then $\mathsf{Slide}_h(\omega)$ is conditionally independently of ω given the value of ([UpClust], DownClust) at ω. We now offer a definition of an iterated slide resample, in which this resampling is undertaken at several heights.

Definition 4.37. Let $\overline{h} = \left\{ h_i : i \in [\![1, L]\!] \right\}$ be a decreasing vector of elements of $[0, u]$. We now specify an iterative resampling procedure $\mathrm{IterateSlide}_{\overline{h}} : \mathsf{Cluster}_{[u,u+3r]} \to \mathsf{Cluster}_{[u,u+3r]}$,

$$\mathrm{IterateSlide}_{\overline{h}} = \mathrm{Slide}_{h_L} \circ \cdots \circ \mathrm{Slide}_{h_2} \circ \mathrm{Slide}_{h_1}.$$

In this definition, the randomness used at each stage is independent of all earlier or later used randomness.

It is perhaps helpful to review this definition, and we set down a little notation by doing so; the reader may consult Figure 9.

The operator $\mathrm{IterateSlide}_{\overline{h}}$ acts on an input configuration ω that realizes $\mathsf{Cluster}_{[u,u+3r]}$. The form of its action is a composition of L steps. Set $S_0 = \omega$ and $S_i = \mathrm{Slide}_h(S_{i-1})$ for $i \in [\![1, L]\!]$, evoking randomness in each step that is independent of that used at other steps. Then we take $\mathrm{IterateSlide}_{\overline{h}}(\omega) = S_L$. The first step is illustrative of the general one. The form of this step depends on whether ω realizes the event Fine_{h_1}. If it does not, the first step is dormant, so that ω remains the state of the system at the end of the first step: which is to say, $S_1 = \omega$. If $\omega \in \mathsf{Fine}_{h_1}$, then the first-step ending-state S_1 is set equal to a realization of the law of the percolation measure $\mathbb{P}$ subject to the event $\mathsf{Cluster}_{[u,u+3r]} \cap \mathsf{Fine}_{h_1}$ occurring and to the data $\mathrm{DownClust}_{h_1}$ and $\left[\mathrm{UpClust}_{h_1} \right]$ coinciding with its specification under $S_0 = \omega$. The height-h hyperplane—the equator—moves downwards at each step, visiting the successive levels in the vector $\overline{h}$, with a dormant or active step taking place at each level. The cluster $\mathcal{C}_{u,u+3r}$ is a function of S_i that is being held at equilibrium while being modified at each step. It is our aim to show that each modification introduces a renewal point at the concerned height with a uniformly positive probability, and further that earlier successes, in which renewal points have been secured higher up in the well, are unlikely to be undone by the ongoing resampling dynamic.
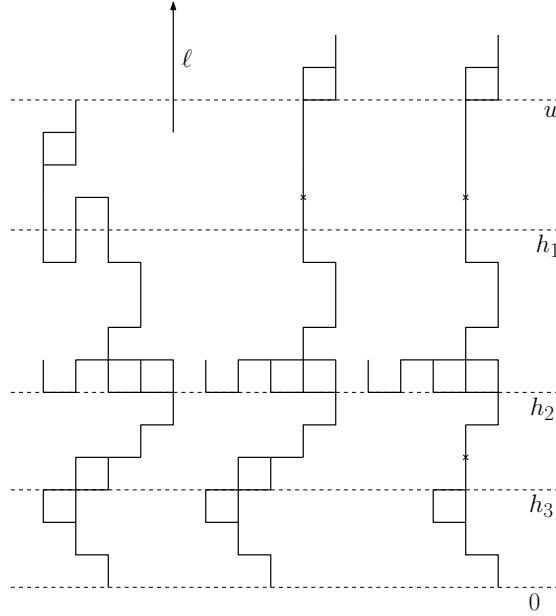
Figure 9: The slide resample iterated, here with $L = 3$ steps. The first and third steps are active; while the second is dormant, due to the high cardinality of the cluster around height h_2. On the left, the cluster under S_0 is shown; in the middle, this cluster under $S_1 = S_2$; on the right, the output cluster, under S_3. The active steps generate simple joins, resulting in two renewal points, marked with crosses, in the output.

Let $i \in [\![1, L]\!]$. The ith step of the procedure is *regular* if either $S_{i-1} \notin \mathsf{Fine}_{h_i}$ or $S_{i-1} \in \mathsf{Fine}_{h_i}$ and $S_i \notin \mathsf{Catastrophe}_{h_i}$. The operator $\text{IterateSlide}_{\overline{h}}$ is said to *act normally* if every step, indexed by $i \in [\![1, L]\!]$, is regular.

Let $j \in [\![0, L]\!]$. An element h_i, with $i \in [\![1, L]\!]$, in the vector $\overline{h}$ is called *fine in* S_j if the configuration S_j realizes the event Fine_{h_i}.

For $L \in \mathbb{N}$, let $\{h_i : i \in [\![1, L]\!]\}$ be a decreasing sequence of elements of $(0, u)$. We further set $h_0 = u$ and $h_{L+1} = 0$. Such a sequence is *well-spaced* if $h_{i+1} - h_i < -6A$ for $i \in [\![0, L]\!]$.

Proposition 4.38. *Let* $L \in \mathbb{N}$ *and let* $\overline{h} = \{h_i : i \in [\![1, L]\!]\}$ *be well-spaced. Suppose that* $\text{IterateSlide}_{\overline{h}}$ *acts normally. Let* $j \in [\![0, L]\!]$. *If* h_j *is very fine in* S_0, *then* h_j *is very fine in* S_j.

For $j \in [\![0, L]\!]$ and $h \in \mathbb{R}$, let $C_{j,h}$ denote the cardinality of the intersection of VertBdry_h and the set $V(\mathcal{C}_{u,u+3r})$ as specified by the configuration S_j.

Lemma 4.39. *Suppose that* $\text{IterateSlide}_{\overline{h}}$ *acts normally. Let* $i \in [\![1, L]\!]$ *and* $\ell \in \mathbb{N}$, *with* $|\ell| \leq 3A$. *Then the sequence* $[\![0, i-1]\!] : j \to C_{j,h_i+\ell}$ *is non-increasing.*

We have seen that the value of the well $\mathcal{C}_{u,u+3r}$ is variable during the L-step procedure that specifies $\text{IterateSlide}_{\overline{h}}$. We write $\mathcal{C}_{u,u+3r}(S_i)$ to indicate the form at the end of the ith step when the algorithm reports state S_i. Naturally we make analogous uses such as $\text{UpClust}_{h_i}(S_j)$: here, the denoted object is the upcluster specified by equatorial height h_i for the well $\mathcal{C}_{u,u+3r}(S_j)$.

Proof of Lemma 4.39. Let $j \in [\![0, L]\!]$. At the jth step of the procedure, an input configura-

tion S_{j-1} is transformed into an output S_j. In this change, the upcluster $\mathrm{UpClust}_{h_j}(S_{j-1})$ shifts to become $\mathrm{UpClust}_{h_j}(S_j)$. Indeed, we may locate a vector $V_j \in \mathbb{Z}^d$ such that $\mathrm{UpClust}_{h_j}(S_j) = \mathrm{UpClust}_{h_j}(S_{j-1}) + V_j$. Let $i \in [\![1, L]\!]$ and $j \in [\![1, i]\!]$. We *claim* that

- when $S_j \notin \mathsf{Catastrophe}_{h_j}$, every $v \in V\big(\mathcal{C}_{u,u+3r}(S_j)\big)$ with $v \cdot \boldsymbol{\ell} < h_j - 3A$ is a vertex in $\mathrm{DownClust}_{h_j}(S_j)$.

To see this, note that a vertex v in $\mathcal{C}_{u,u+3r}(S_j)$ is a vertex of at least one of $\mathrm{UpClust}_{h_j}(S_j)$, $\mathrm{DownClust}_{h_j}(S_j)$ and $\mathrm{ResidualClust}_{h_j}(S_j)$. Any vertex w in $\mathrm{UpClust}_{h_j}(S_j)$ satisfies $w \cdot \boldsymbol{\ell} \geq h_j$. Because $S_j \notin \mathsf{Catastrophe}_{h_j}$, any vertex w in $\mathrm{ResidualClust}_{h_j}(S_j)$ satisfies $w \cdot \boldsymbol{\ell} \geq h_j - 3A$. Thus we see that a vertex v in $\mathcal{C}_{u,u+3r}(S_j)$ with $v \cdot \boldsymbol{\ell} < h_j - 3A$ is indeed a vertex in $\mathrm{DownClust}_{h_j}(S_j)$, as we claimed.

The case $i = 1$ in Lemma 4.39 being trivial, we let $i \in [\![2, L]\!]$ and $j \in [\![1, i-1]\!]$. Because the vector $\overline{h}$ is well-spaced, $h_i + \ell < h_j - 3A$ whenever $|\ell| \leq 3A$. Since $\mathrm{IterateSlide}_{\overline{h}}$ acts normally, either $S_{j-1} \notin \mathsf{Fine}_{h_j}$ or $S_j \notin \mathsf{Catastrophe}_{h_j}$. In the former case, $S_j = S_{j-1}$ and thus $C_{j,h_i+\ell} = C_{j-1,h_i+\ell}$. In the latter case, we invoke the claim to find that any vertex v in $\mathcal{C}_{u,u+3r}(S_j)$ that lies in $\mathrm{VertBdry}_{h_i+\ell}$ is a vertex in $\mathrm{DownClust}_{h_j}(S_j)$. Since $\mathrm{DownClust}_{h_j}(S_j) = \mathrm{DownClust}_{h_j}(S_{j-1})$ is a subgraph of $\mathcal{C}_{u,u+3r}(S_{j-1})$, we see that the number of vertices in $\mathcal{C}_{u,u+3r}(S_j)$ that lie in $\mathrm{VertBdry}_{h_i+\ell}$ is at most the number of vertices in $\mathcal{C}_{u,u+3r}(S_{j-1})$ that lie in $\mathrm{VertBdry}_{h_i+\ell}$: which is to say, $C_{j,h_i+\ell} \leq C_{j-1,h_i+\ell}$. Lemma 4.39 has been proved. $\qquad\square$

Proof of Proposition 4.38. Let $j \in [\![0, i-1]\!]$. It is enough to argue that, if h_i is very fine in S_j, then it is very fine in S_{j+1}. If $j = i-1$, note that $S_i = \mathrm{Slide}_{h_i}(S_{i-1})$ and that the definition of the random map Slide_{h_i} entails the desired assertion for this value of j. Suppose then that $j \leq i-2$. The event $\mathsf{VeryFine}_{h_i}$ from Definition 4.8 entails tropical, subtropical and global conditions: that the first or second hold under S_{j+1} when the respective assertion does under S_j follows from Lemma 4.39. Note that either $S_{j+1} = S_j$, or S_{j+1} realizes the event $\mathsf{VeryFine}_{h_j}$: either way, the global condition is verified for S_{j+1} when it is for S_j. $\qquad\square$

Let $i \in \mathbb{N}$ and set $h_i = i \cdot 12A$. Let Slab_i denote $\big\{ z \in r\mathbb{Z} : |z - h_i| < 3A \big\}$. Set $L = \lfloor u(12A)^{-1} \rfloor - 1$. The set $\big\{ \mathrm{Slab}_i : 1 \leq i \leq L \big\}$ is a collection of consecutive elements of $r \cdot \mathbb{N}$ that is contained in $[0, u]$. Each element has cardinality $6Ar^{-1}$ (up to a discrepancy of two), and the gap between consecutive members, in the sense of the cardinality of the interval in $r\mathbb{Z}$ lying to the right of the lower and to the left of the higher, is also $6Ar^{-1}$ (up to the same discrepancy).

Let $\overline{H} = \big(H_i : i \in [\![1, L]\!] \big)$ denote a random vector whose components are selected independently. The law of H_i is uniform on Slab_i.

We will consider the iterative random procedure $\mathrm{IterateSlide}_{\overline{H}}$, where the initial condition S_0 has the law $\mathbb{P}(\cdot | \mathsf{Cluster}_{[u,u+3r]})$, the random vector $\overline{H}$ is selected independently of S_0, and the randomness that is invoked during the running of the procedure is independent of this other randomness. Let $\mathbb{Q}$ denote the law that governs the procedure in this form.

Let $\alpha \in (0, 1)$. We say that the procedure's *input is α-satisfactory* if, in the configuration S_0, there are at least αL indices $i \in [\![1, L]\!]$ for which $\mathsf{VeryFine}_{H_i}$ occurs.

Proposition 4.40. *There exist $\alpha \in (0, 1)$ and $c > 0$ such that, under $\mathbb{Q}$, input is α-satisfactory with probability at least $1 - \exp\big\{ -cL \big\}$.*

Proof. Let $a \in (0, 1/5)$. Choose $K \in \mathbb{N}$ to be at least Ca^{-1} for the constant C in Proposition 4.10. Under $\mathbb{Q}$, the input configuration S_0 has the law $\mathbb{P}(\cdot | \mathsf{Cluster}_{[u,u+3r]})$. Recall from the same proposition that N denotes the number of $h \in \mathbb{N} \cap [0, u]$ for which $\mathsf{VeryFine}_h$ occurs. By this result, $N \geq (1-a)u$ except with probability at most $\exp\{-cKau\}$. In the event $N \geq (1-a)u$, at most au/r levels $i \in r\mathbb{N} \cap (0, u]$ fail to realize the event $\mathsf{VeryFine}_i$. A slab Slab_i may be called *grim* if at least one-half of its members fail in this sense. Since the cardinality of a slab and the gap between one slab and the next both equal $6Ar^{-1}$ up to an error of two, the union $\cup_{i=1}^L \mathrm{Slab}_i$, which is a subset of $r\mathbb{Z} \cap [0, u]$, comprises a proportion of at least $1/2 - o(1)$ of elements of $r\mathbb{N} \cap [0, u]$, where $o(1) \to 0$ in high A. By choosing A to be at least a suitably high constant A_0, we thus ensure that, when $N \geq (1-a)u$, the number of grim slabs is at most $5aL$, and the number of slabs that are not grim is at least $(1-5a)L$. If i indexes a not grim slab, $\mathsf{VeryFine}_{H_i}$ occurs with probability at least one-half, independently of other such outcomes. In the event $N \geq (1-a)u$, then, the number of indices $i \in [\![1, L]\!]$ for which $\mathsf{VeryFine}_{H_i}$ occurs stochastically dominates a binomial random variable with parameters $(1-5a)L$ and $1/2$. Let α be a given positive value less than $(1-5a)/2$. Proposition 4.40 follows from binomial tail bounds. $\qquad\square$

We fix a value for $\alpha \in (0, 1)$ in accordance with Proposition 4.40.

Proposition 4.41. *Consider the law $\mathbb{Q}$ conditionally on input being α-satisfactory. There exist positive constants c_0, c_1 and c_2 such that, when $L \leq c_0 u(\log u)^{-1}$, the probability that the output S_L contains at least $c_1 L$ renewal levels is at least $1 - \exp\{-c_2 L\} - L \exp\{-c_2 u/L\}$.*

Proof. Note first that a small enough choice of $c_0 > 0$ yields the hypothesis $\min\{h, u-h, A\} \geq C \log u$ that permits the application of Propositions 4.11 and 4.30 in the upcoming argument.

Let $S \subset [\![1, L]\!]$ denote the set of indices i for which $\mathsf{VeryFine}_{H_i}$ occurs in the input configuration S_0. Under the conditional law that we consider, $|S|$ is at least αL. Let $i \in S$. Suppose that, in the first $i-1$ steps of the procedure that specifies $\mathrm{IterateSlide}_{\overline{H}}$, catastrophe has not occurred: that is, for $j \in [\![1, i-1]\!]$, $S_{j-1} \notin \mathsf{Catastrophe}_{H_j}$. From the proof of Proposition 4.38, we learn that $\mathsf{VeryFine}_{H_i}$ occurs in S_{i-1}. By Lemma 4.9 and Proposition 4.11, there is probability at least some positive quantity c_0 that H_i is a renewal level in S_i.

If $S_k \notin \mathsf{Catastrophe}_{H_k}$ for all $k \in [\![1, L]\!]$, then, for any $i \in [\![1, L]\!]$ for which H_i is a renewal level in S_i, the value H_i is also a renewal level in the output configuration S_L. Indeed, in any update during the running of the procedure after the ith step, the downcluster and the residual cluster will lie below the level H_i. Since the upcluster is preserved at each of these later steps, the renewal status of level H_i measured relative to the top cannot be jeopardised during this later part of the procedure.

We see then that the number of renewal levels in the output of the procedure is stochastically dominated by $B\mathbf{1}_{\mathsf{N}}$, where B is a binomial random variable with parameters αL and c_0, and N is the event that $S_i \notin \mathsf{Catastrophe}_{H_i}$ for $i \in [\![1, L]\!]$. By Proposition 4.30, the quantity $\mathbb{Q}(\mathsf{N})$ is at least $1 - L \exp\{-c_2 u/L\}$ for some positive value of c_2. Set $c_1 = \alpha c_0/2$. We have that $B \geq c_1 L$ except with probability at most $\exp\{-c\alpha L\}$ for some positive c (that does not depend on α). Adjusting $c_2 > 0$ to be at most $c\alpha$, we obtain Proposition 4.41. $\qquad\square$

Proposition 4.42. *There exists $c > 0$ such that, for $k \leq cu(\log u)^{-1}$, the probability under $\mathbb{P}$ given $\mathsf{Cluster}_{[u,u+3r]}$ that $\mathcal{C}_{u,u+3r}$ has fewer than k renewal levels is at most $\exp\{-ck\} + k\exp\{-cu/k\}$.*

Proof. We consider the copy of the law $\mathbb{P}(\cdot | \mathsf{Cluster}_{[u,u+3r]})$ offered by the distribution of the output

of the iterative procedure IterateSlide$_{\overline{H}}$. We aim to set the parameter L close to the value k. By choosing $A = \lfloor u/(12(k+1)) \rfloor$, we see that

$$L = \lfloor u(12A)^{-1} \rfloor - 1 \text{ lies between } k \text{ and } k + 1 + 12(k+1)^2/u + k^3 u^{-2} O(1)\,.$$

Since our assumption entails that $k \leq cu$ for a suitably small $c > 0$, we have that $L \in [\![k, 2k]\!]$. Thus, input is α-satisfactory with probability at least $1 - \exp\{-ck\}$ by Proposition 4.40. Provided that input is α-satisfactory, the output configuration will contain at least $c_1 k$ renewal levels with probability at least $1 - \exp\{-c_2 k\} - 2k \exp\{-2^{-1}c_2 u/k\}$ by Proposition 4.41. By relabelling k, we obtain Proposition 4.42 for a suitably small choice of $c > 0$. $\qquad\square$

Proof of Theorem 1.9(2:Cluster). The statement is obtained from Proposition 4.42 by lowering the value of $c > 0$ if need be. $\qquad\square$

We end the section by deriving Proposition 2.1, which is a string counterpart to Theorem 1.9(1).

Let $\mathsf{Cluster}^*_{[u,u+3r]}$ denote the event $\mathsf{Cluster}_{[u,u+3r]}$ given that $\mathrm{head}(\mathcal{C})$ and $\mathrm{base}(\mathcal{C})$ are renewal points in $R(\mathcal{C})$ (and we also permit semi-open intervals $[u, u+3r)$ in using this notation). Clusters with such renewal points are close cousins of strings; Corollary 1.10 is the counterpart of Theorem 1.9 for such clusters, and Corollary 1.10(1) is the principal tool for proving Proposition 2.1.

Proof of Corollary 1.10. The derivation coincides with that of Theorem 1.9 by considering the event $\mathsf{Cluster}^*_{[u,u+3r]}$ (or $\mathsf{Cluster}^*_{[u,u+3r)}$, for the third part) in place of $\mathsf{Cluster}_{[u,u+3r]}$. $\qquad\square$

A further straightforward claim is needed.

Lemma 4.43. *We have that*

$$\mathbb{P}\big(\mathsf{Cluster}^*_{[u-2r,u+r)}\big) \geq c\,\mathbb{P}\big(\mathsf{Str}_{[u,u+r)}\big)$$

Proof. This follows from $\mathbb{P}\big(\mathsf{Cluster}^*_{[u-2r,u+r)}\big) \geq \mathbb{P}\big(\mathsf{Cluster}^*_{[u,u+r]}\big) \geq c\mathbb{P}\big(\mathsf{Str}_{[u,u+r]}\big)$, the latter bound due to a bounded number of independent open/closed choices rendering a realization of $\mathsf{Str}_{[u,u+r]}$ into an instance of $\mathsf{Cluster}^*_{[u,u+r]}$. $\qquad\square$

Proof of Proposition 2.1. Lemma 4.43 implies that

$$\mathbb{P}\Big(|R(S)| \geq \delta u \,\Big|\, \mathsf{Str}_{[u,u+r)}\Big) \geq c \cdot \frac{\mathbb{P}\Big(|R(S)| \geq \delta u\,,\, \mathsf{Str}_{[u,u+r)}\Big)}{\mathbb{P}\big(\mathsf{Cluster}^*_{[u-2r,u+r)}\big)}\,. \tag{25}$$

It is straightforward to see that

$$\mathbb{P}\Big(|R(S)| \geq \delta u\,,\, \mathsf{Str}_{[u,u+r)}\Big) \geq c \cdot \mathbb{P}\Big(|R(S)| \geq \delta u - 3\,,\, \mathsf{Str}_{[u-2r,u+r)}\Big)$$

where on the right-hand side, S denotes the string of minimal length that manifests the event $\mathsf{Str}_{[u+r-3,u+r)}$. Indeed, the occurrence of the above right-hand event entails that on the left provided that a bounded set of open/closed decisions outside the slab implicated in the right-hand event fall the right way.

By restricting to a slab, we see that the last displayed right-hand probability is at least

$$\mathbb{P}\Big(|R(\mathcal{C})| \geq \delta u - 3\,,\, \mathsf{Cluster}^*_{[u-2r,u+r)}\Big)\,.$$

Returning to (25), we find that

$$\mathbb{P}\Big(|R(S)| \geq \delta u \,\Big|\, \mathsf{Str}_{[u,u+r)}\Big) \;\geq\; c \cdot \mathbb{P}\Big(|R(\mathcal{C}) \geq \delta u - 3 \,\Big|\, \mathsf{Cluster}^*_{[u-2r,u+r)}\Big).$$

Corollary 1.10(3) shows that the right-hand side is at least c^2. This completes the proof of Proposition 2.1. $\qquad\square$

References

[AB87]   Michael Aizenman and David J. Barsky. Sharpness of the phase transition in percolation models. *Communications in Mathematical Physics*, 108(3):489–526, 1987.

[AP96]   Peter Antal and Agoston Pisztora. On the chemical distance for supercritical Bernoulli percolation. *The Annals of Probability*, 24(2):1036–1048, 1996.

[BGP03] Noam Berger, Nina Gantert, and Yuval Peres. The speed of biased random walk on percolation clusters. *Probab. Theory Relat. Fields*, 126:221–242, 2003.

[CC86]   J. T. Chayes and L. Chayes. Ornstein-Zernike behavior for self-avoiding walks at all noncritical temperatures. *Comm. Math. Phys.*, 105(2):221–238, 1986.

[CI02]   Massimo Campanino and Dmitry Ioffe. Ornstein-Zernike theory for the Bernoulli bond percolation on $\mathbb{Z}^d$. *The Annals of Probability*, 30(2):652 – 682, 2002.

[CIL10]   M. Campanino, D. Ioffe, and O. Louidor. Finite connections for supercritical Bernoulli bond percolation in $2D$. *Markov Proc. Rel. Fields*, 16(2):225–266, 2010.

[CIV03] Massimo Campanino, Dmitry Ioffe, and Yvan Velenik. Ornstein-Zernike theory for finite range Ising models above t_c. *Probab. Theory Relat. Fields*, 125:305–349, 2003.

[CIV08] Massimo Campanino, Dmitry Ioffe, and Yvan Velenik. Fluctuation theory of connectivities for subcritical random cluster models. *Ann. Probab.*, 36(4):1287–1321, 2008.

[Fel71]   William Feller. *An introduction to probability theory and its applications. Vol. II.* Second edition. John Wiley & Sons Inc., New York, 1971.

[FH14]   Alexander Fribergh and Alan Hammond. Phase transition for the speed of the biased random walk on the supercritical percolation cluster. *Communications on pure and applied mathematics*, 67(2):173–245, 2014.

[Ham06] Alan Hammond. Greedy lattice animals: Geometry and criticality. *The Annals of Probability*, 34(2):593–637, 2006.

[Iof98]   Dmitry Ioffe. Ornstein-Zernike behaviour and analyticity of shapes for self-avoiding walks on $\mathbb{Z}^d$. *Markov Proc. Rel. Fields*, 4(3):323–350, 1998.

[LSS97] T. M. Liggett, R. H. Schonmann, and A. M. Stacey. Domination by product measures. *The Annals of Probability*, 25(1):71–95, 1997.

[LW49]   L. H. Loomis and H. Whitney. An inequality related to the isoperimetric inequality. *Bulletin of the American Mathematical Society*, 55(10):961–962, 1949.

[MZ96]   R. A. Minlos and E. A. Zhizhina. Asymptotics of decay of correlations for lattice spin fields at high temperatures. I. the Ising model. *J. Statist. Phys.*, 84(1/2):85–118, 1996.

[Nag69]   A. V. Nagaev. Integral limit theorems taking large deviations into account when Cramér's condition does not hold. ii. *Theory of Probability & Its Applications*, 14(2):193–208, 1969.

[OZ15]   L. S. Ornstein and F. Zernike. Accidental deviations of density and opalescence at the critical point of a single substance. *Proc. Acad. Sci. (Amst.)*, 17:793–806, 1915.

[SW67]   Charles Stone and Stephen Wainger. One-sided error estimates in renewal theory. *Journal d'analyse mathématique (Jerusalem)*, 20(1):325–352, 1967.

[Szn03]   Alain-Sol Sznitman. On the anisotropic walk on the supercritical percolation cluster. *Commun. Math. Phys.*, 240:123–148, 2003.

[Tim13]   Ádám Timár. Boundary-connectivity via graph theory. *Proceedings of the American Mathematical Society*, 141(2):475–480, 2013.

ALEXANDER FRIBERGH
MONTREAL
CANADA
Email: `alex@montreal.canada`

ALAN HAMMOND
DEPARTMENTS OF MATHEMATICS AND STATISTICS, U.C. BERKELEY
899 EVANS HALL, BERKELEY, CA, 94720-3840, U.S.A.
Email: `alanmh@berkeley.edu`