Typical geometry of long subcritical clusters

— An introduction to the Ornstein–Zernike theory —

Yvan Velenik, Université de Genève

Abstract

In these notes, we review the description of typical long subcritical clusters in Bernoulli percolation on $\mathbb{Z}^d$. The analysis is based on the Ornstein–Zernike theory, as developed in [6, 8, 10]. Although the last two papers deal with more complicated models (the Ising model and the FK percolation, respectively), we only consider Bernoulli percolation for pedagogical reasons, but follow the arguments in these works.

1 Brief historical introduction

In this short section, we make a few brief historical comments, but emphasize the fact that this overview is very incomplete.

The Ornstein–Zernike (OZ) theory owes its name to a famous paper published in 1914 by the physicists Leonard Ornstein and Frits Zernike [29]. Their work introduced a new way to analyze (nonrigorously) the density-density correlations in fluids, and is still today an important part of the theory of fluids. Among the important applications of this theory, it was possible to determine for the first time the prefactor to the exponential decay of correlations away from the critical point [33].

A rigorous derivation of the OZ theory and its main consequences required more than 60 years, and was accomplished by Abraham and Kunz [1] and Paes-Leme [32] independently. Both works were however restricted to perturbative regimes (very low or very high temperatures, very low densities, etc.). Many further rigorous works, with the same limitations, were then obtained in the following decade, investigating in particular more general correlations.

In the 1980s, a first nonperturbative version of the OZ theory was developed. In [11], Chayes and Chayes derived sharp asymptotics for the 2-point function of the self-avoiding walk on $\mathbb{Z}^d$ valid in the whole subcritical regime, while Campanino, Chayes and Chayes did the same in [5] for subcritical Bernoulli percolation. The methods used, which relied on a detailed microscopic analysis, limited the applicability of this approach to very simple models and correlations along axes directions. An extension to general directions, in the case of the self-avoiding walk, was done by Ioffe in [20], using tilting techniques typical from large deviations theory.

The next breakthrough was made by Campanino and Ioffe in [6]. Using some ideas developed by Alexander [3], they were able to replace the delicate microscopic analysis used in the previous works by a much more robust procedure based on a coarse-graining of the microscopic object. In this way, they were able to obtain sharp asymptotics for connectivities in subcritical Bernoulli percolation, as well as additional information on equi-decay profiles, Wulff shape, etc.

The next stage required going beyond models with a built-in independence structure. This was first done by Campanino, Ioffe and Velenik for the Ising model above the critical temperature in [8], and then for FK percolation (thus including Potts models on particular) in [10]. A further improvement was made by Ott and Velenik in [31], by showing how the non-Markovian process studied in the previous two works could be replaced by a suitable random walk.

Since then, the (rigorous, nonperturbative) OZ theory has found many applications to a variety of problems, including: Brownian bridge asymptotics for the interface in two-dimensional Ising [15] and Potts [10] models; derivation of the prefactor to the exponential decay for supercritical Bernoulli percolation [7]; fluctuations of the boundary of the Wulff shape in two-dimensional Potts models [18, 19, 17]; analysis of the effect of a defect line on correlations in Bernoulli percolation [14] and in Potts models on $\mathbb{Z}^d$ above the critical temperature (as well as on the interface of the two-dimensional Potts model below the critical temperature) [31]; derivation of the prefactor for general odd-odd [9] and even-even [30] correlations in the Ising model; analysis of entropic repulsion for Ising-type polymers [22]; finite-volume versions of the Aizenman–Higuchi theorem for the two-dimensional Ising [13] and Potts [12] models; detailed analysis of stretched self-interacting polymers [23, 25], and of stretched or crossing polymers in random environment [24, 26, 27, 21].

2 Bernoulli percolation

Let us denote by E^d the set of all nearest-neighbor edges of the lattice $\mathbb{Z}^d$, $\mathsf{E}^d = \{\{i,j\} : \|j-i\|_1 = 1\}$. The Bernoulli (bond) percolation process on $\mathbb{Z}^d$ is a collection of i.i.d. random variables $(\omega_e)_{e\in\mathsf{E}^d}$, each variable following a Bernoulli distribution with parameter p. Their joint distribution will be denoted $\mathbb{P}_p$ and the corresponding expectation $\mathbb{E}_p$. An edge $e \in \mathsf{E}^d$ is said to be *open* in ω is $\omega_e = 1$; otherwise it is said to be *closed*.

We will be mainly interested in connectivity properties. It will thus be convenient to identify the set of all open edges of ω with the (random) graph G_ω with vertices $\mathbb{Z}^d$ and edges $\{e \in \mathsf{E}^d : \omega_e = 1\}$. We will then denote by C_x the *cluster* of $x \in \mathbb{Z}^d$, that is, the maximal connected component of G_ω containing x. Given $A_1, A_2 \subset \mathbb{Z}^d$ and $B \subset \mathsf{E}^d$, we will write $A_1 \overset{B}{\leftrightarrow} A_2$ for the event that there is a path of open edges starting in A_1, ending in A_2 and containing only edges of B. We will use the same notation with $A_1, A_2, B \subset \mathbb{Z}^d$ for the event that there is a path of open edges starting in A_1, ending in A_2 and containing only edges with endpoints in B. When $B = \mathbb{Z}^d$ or E^d, we simply write $A_1 \leftrightarrow A_2$. Moreover, when A_1 and/or A_2 are reduced to a single vertex x, we simply write x rather than $\{x\}$, as in $0 \leftrightarrow x$.

The measure $\mathbb{P}_p$ enjoys a number of useful properties; here, we will only need one: the *FKG inequality*. We consider the natural partial order on configurations ω: $\omega \le \omega'$ if and only if G_ω is a subgraph of $G_{\omega'}$. We then say that a function $f : \{0,1\}^{\mathsf{E}^d} \to \mathbb{R}$ is *increasing* if $\omega \le \omega' \implies f(\omega) \le f(\omega')$. f is said to be *decreasing* if $-f$ is increasing. An event $\mathcal{A}$ is *increasing (resp. decreasing)* if its indicator function is increasing (resp. decreasing). The FKG inequality then states that increasing events are positively correlated: for any increasing events $\mathcal{A}$ and $\mathcal{B}$, $\mathbb{P}_p(\mathcal{A} \cap \mathcal{B}) \ge \mathbb{P}_p(\mathcal{A})\mathbb{P}_p(\mathcal{B})$.

Conventions, notations: We will denote by $|A|$ the cardinality of the set A. We define the *external boundary* of a subset $A \subset \mathbb{Z}^d$ by $\partial^{\text{ext}} A = \{j \in \mathbb{Z}^d \setminus A : \exists i \in A, \{i,j\} \in \mathsf{E}^d\}$, and the *external boundary* of $A \subset \mathsf{E}^d$ by $\partial^{\text{ext}} A = \{\{i,j\} \in \mathsf{E}^d \setminus A : \exists k, \{i,k\} \in A\}$. In the sequel, c will denote a generic constant, whose value can change from place to place, including in a given equation. We will write $o_a(1)$ for a quantity that tends to 0 as a, $\|a\|$ or $|a|$ tends to infinity (depending on the context).

3 Inverse correlation length

Bernoulli percolation undergoes a phase transition at a specific value p_c of the parameter p; namely:

$$\mathbb{P}_p(|C_0| = \infty) \begin{cases} > 0 & \forall p > p_c, \\ = 0 & \forall p < p_c. \end{cases}$$

In particular, $\lim_{\|x\|\to\infty} \mathbb{P}_p(0 \leftrightarrow x) = 0$ when $p < p_c$ and a natural question is to determine how fast this decay actually is. This can be quantified using the inverse correlation length.

For any $x \in \mathbb{R}^d$, the *inverse correlation length* is defined by

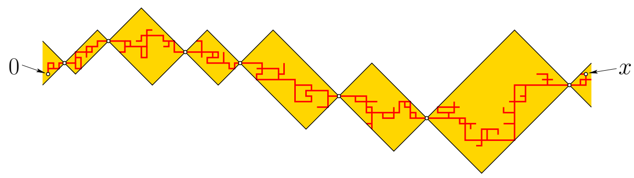
$$\xi_p(x) = -\lim_{k\to\infty} \frac{1}{k} \log \mathbb{P}_p(0 \leftrightarrow [kx]),$$

where, for any $y \in \mathbb{R}^d$, $[y] \in \mathbb{Z}^d$ is the component-wise integer part of y.

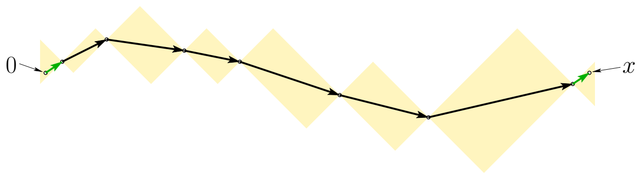
Exercise 1. *Show that the above limit exists and that $\mathbb{P}_p(0 \leftrightarrow x) \le e^{-\xi_p(x)}$ for all $x \in \mathbb{Z}^d$.*

It is a well known fact (first proved in [28, 2]) that $\xi_p(x) > 0$ for all $x \neq 0$ when $p < p_c$, while $\xi_p(x) = 0$ for all $x \in \mathbb{R}^d$ when $p > p_c$.

Exercise 2. *Show that ξ_p is a norm on $\mathbb{R}^d$ when $p < p_c$.*



(a) A typical realization of $C_{0,x}$ can be decomposed into a string of microscopic "irreducible" pieces. The size of these pieces has exponential tails.



(b) Large-scale statistical properties of $C_{0,x}$ can be derived from an associated effective directed random walk (the first and last increments have a different distribution, but they are small and thus do not affect large-scale properties).

Figure 1: Structure of a typical realization of the subcritical cluster $C_{0,x}$ and the associated effective random walk.

For $p < p_c$, we will denote by

$$\mathsf{U}_p = \left\{ x \in \mathbb{R}^d \,:\, \xi_p(x) \leq 1 \right\}$$

the unit ball in ξ_p-norm. Note that U_p is convex[1] and symmetric under lattice reflections (since ξ_p is).

4 Ornstein–Zernike analysis of typical long subcritical clusters

In this section, we provide a detailed description of the geometry of typical long subcritical clusters, showing that they take the form of a string of microscopic "irreducible" pieces as depicted in Fig. 1a. Moreover, we show that the large-scale statistical properties of this cluster can be obtained from those of an associated directed random walk on $\mathbb{Z}^d$ (whose steps are given by the increments of the irreducible pieces); see Fig. 1b. To do this, we follow the approach developed in [10] (based on the earlier works [6, 8]). We only discuss the case of Bernoulli percolation, as it is substantially simpler. Some remarks regarding the modifications required to cover the case of FK percolation will be given after the proof.

4.1 Coarse-graining

To slightly simplify the exposition and the notation, we will only consider the event $\{0 \leftrightarrow n\vec{e}_1\}$. It will then be convenient to write $y \in \mathbb{Z}^d$ as $y = (y^{\parallel}, y^{\perp}) \in \mathbb{Z} \times \mathbb{Z}^{d-1}$. We will also write $\hat{\xi}_p = \xi_p(\vec{e}_1)$.

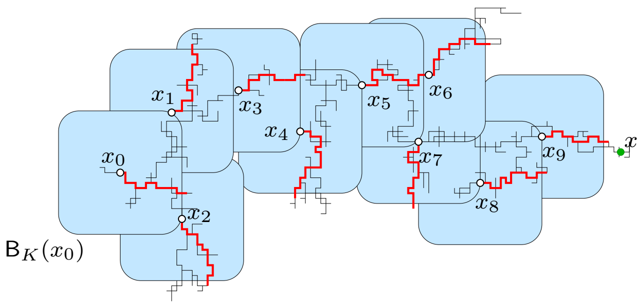
4.1.1 Skeleton

Given $0 \neq x \in \mathbb{Z}^d$, let us denote by $C_{0,x}$ the common cluster of 0 and x under $\mathbb{P}_p(\cdot \,|\, 0 \leftrightarrow x)$. Given $y \in \mathbb{Z}^d$ and $K > 0$, define

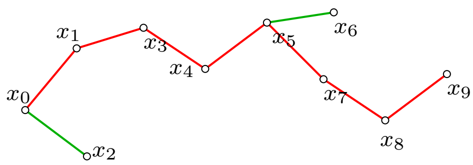
$$\mathsf{B}_K(y) = (y + K\mathsf{U}_p) \cap \mathbb{Z}^d.$$

The following algorithm is used to associate to any realization C of $C_{0,x}$ a tree that approximates C at scale K, which we will call its *skeleton tree* (see Fig. 2a).

[1]To see that, consider $x, y \in \mathsf{U}_p$ and $\lambda \in [0,1]$. Then, convexity of ξ_p (see Exercise 2) implies that $\xi_p(\lambda x + (1-\lambda)y) \leq \lambda\xi_p(x) + (1-\lambda)\xi_p(y) \leq 1$ and so $\lambda x + (1-\lambda)y \in \mathsf{U}_p$.



(a) Construction of the vertices $x_0 = 0, x_1, \ldots, x_9$. Observe that there are disjoint connexions from each of the vertices x_j to $\partial^{\text{ext}}\mathsf{B}_K(x_j)$ (highlighted in red).



(b) The resulting tree (the trunk is in red, the two non-empty branches in green).

Figure 2: Construction of the skeleton tree $\mathfrak{T}(C)$ associated to a realization C of the cluster $C_{0,x}$.

Algorithm 1: extraction of the vertices of the skeleton tree associated to C

---

let $x_0 = 0$, $\mathcal{C} = \mathsf{B}_K(x_0)$ and $i = 1$
while $\exists y \in \partial^{\text{ext}}\mathcal{C}$ *such that* $y \xleftrightarrow{C\setminus\mathcal{C}} \partial^{\text{ext}}\mathsf{B}_K(y)$ **do**
  let x_i be the minimal such vertex
  update $\mathcal{C} \leftarrow \mathcal{C} \cup \mathsf{B}_K(x_i), i \leftarrow i + 1$
end

The above algorithm yields a family of vertices $x_0, \ldots, x_N$. We consider these vertices as the vertices of a tree $\mathfrak{T} = \mathfrak{T}(C)$, which is obtained by adding edges according to the following rule (see Fig. 2b): for each $1 \le k \le N$, we add an edge between x_k and the minimal vertex x_j, $0 \le j \le k-1$, such that $x_k \in \partial^{\text{ext}}\mathsf{B}_K(x_i)$. We write $N(\mathfrak{T}) = N$ the number of vertices of the tree $\mathfrak{T}$ (minus 1).

Let us now make two important observations:

1. $\bigcup_{i=0}^{N(\mathfrak{T})} \mathsf{B}_K(x_i)$ is a connected set;

2. $C \subset \bigcup_{i=0}^{N(\mathfrak{T})} \mathsf{B}_{2K}(x_i)$.

The second observation explains in which sense the tree $\mathfrak{T}$ provides an approximation of the cluster C at scale K.

The main reason to construct such an approximation at scale K is that, once K is chosen large enough, one can implement a classical energy/entropy argument *for any* $p < p_c$.

The energy estimate. The probability of a particular tree $\mathfrak{T}$ is defined as

$$\mathbb{P}_p(\mathfrak{T}) = \sum_{C \sim \mathfrak{T}} \mathbb{P}_p(C_{0,x} = C),$$

where the sum is over all realizations C of $C_{0,x}$ to which Algorithm 1 associates the tree $\mathfrak{T}$. Set $A^0 = \mathsf{B}_K(x_0)$ and, for $1 \le j \le N(\mathfrak{T})$, $A^j = \mathsf{B}_K(x_j) \setminus \bigcup_{i=0}^{j-1} \mathsf{B}_K(x_i)$. We then have

$$\begin{aligned}\mathbb{P}_p(\mathfrak{T}) &\le \mathbb{P}_p(x_i \xleftrightarrow{A^i} \partial^{\mathrm{ext}}\mathsf{B}_K(x_i),\, 0 \le i \le N(\mathfrak{T})) \\ &= \prod_{i=0}^{N(\mathfrak{T})} \mathbb{P}_p(x_i \xleftrightarrow{A^i} \partial^{\mathrm{ext}}\mathsf{B}_K(x_i)) \\ &\le \exp\bigl(-K(1-o_K(1))(N(\mathfrak{T})+1)\bigr), \end{aligned} \tag{1}$$

where the equality in the second line follows from independence (since the sets A^i are disjoint) and the inequality in the third line from

$$\mathbb{P}_p(x_i \xleftrightarrow{A^i} \partial^{\mathrm{ext}}\mathsf{B}_K(x_i)) \le |\partial^{\mathrm{ext}}\mathsf{B}_K(x_i)| \max_{y\in\partial^{\mathrm{ext}}\mathsf{B}_K(x_i)} \mathbb{P}_p(x_i \leftrightarrow y) \le e^{-K(1-o_K(1))},$$

since $\mathbb{P}_p(x_i \leftrightarrow y) \le e^{\xi_p(y-x_i)} \le e^{-K}$ by Exercise 1 and the definition of the set $\mathsf{B}_K(x_i)$.

The entropy estimate. Let us now turn to the entropy (that is, combinatorial) estimate on the number of trees $\mathfrak{T}$ with prescribed number of vertices.

Lemma 4.1. *There exists $c = c(d)$ such that, for any $N \in \mathbb{Z}_{\ge 1}$,*

$$\#\{\mathfrak{T} : N(\mathfrak{T}) = N\} \le \exp(c\log(K)N).$$

Proof. Observe that the tree's branching number is bounded above by $|\partial^{\mathrm{ext}}\mathsf{B}_K(0)| \le C(d)K^{d-1}$. One way to upper bound the number of trees (rooted at 0) is to realize that each such tree $\mathfrak{T}$ can be seen as a subtree of the infinite $C(d)K^{d-1}$-regular tree. To count the number of such subtrees with N vertices, observe that one can associate to each such subtree a path starting at the root and crossing each edge of the subtree exactly twice (just perform a depth-first search through the subtree). The number of such subtrees is thus bounded above by the number of paths of length $2N$, which is itself bounded above by

$$\bigl(C(d)K^{d-1}\bigr)^{2N} = \exp\bigl(c(d)\log(K)N\bigr). \qquad \square$$

The important thing to notice is that our entropic estimate grows exponentially with $\log K$, while the energy estimate decreases exponentially with K; the latter will thus be the dominating factor when K is large. This observation will be used in Section 4.1.2.

Trunk and branches. It will be convenient to split the tree $\mathfrak{T}$ into a *trunk* and a set of *branches* (see Fig. 2b). The trunk $\mathfrak{t} = (\mathfrak{t}_0, \dots, \mathfrak{t}_{N(\mathfrak{t})})$ is the (unique) path in $\mathfrak{T}$ connecting 0 to x_F, where $x_F = \min\{x_i \in \mathfrak{T} : x \in \mathsf{B}_{2K}(x_i)\}$. The branches $\mathfrak{b}_0, \mathfrak{b}_1, \dots, \mathfrak{b}_{N(\mathfrak{t})}$ are the connected components of $\mathfrak{T}$ obtained after removing the edges of $\mathfrak{t}$; each $\mathfrak{b}_i$ is thus a tree rooted at a vertex $\mathfrak{t}_i$ of the trunk. The collection of all branches is denoted $\mathfrak{B} = (\mathfrak{b}_0, \dots, \mathfrak{b}_{N(\mathfrak{t})})$ and its size is written $N(\mathfrak{B}) = \sum_k N(\mathfrak{b}_k)$.

Similarly as before, we can define the probability of a trunk $\mathfrak{t}$ as

$$\mathbb{P}_p(\mathfrak{t}) = \sum_{C\sim\mathfrak{t}} \mathbb{P}_p(C_{0,x} = C),$$

where the sum is over all realizations C of $C_{0,x}$ whose tree has trunk $\mathfrak{t}$.

Our goal in the next two sections is to analyze the geometry of the trunk and branches associated to typical realizations of the cluster $C_{0,x}$.

4.1.2 Geometry of typical trees: bound on the size

Observe that, by construction, since the cluster C connects 0 and x, the associated trunk must contain at least $(1-o_K(1))\xi_p(x)/K$ vertices. The next result shows that it is very unlikely that the tree $\mathfrak{T}$ contains many more vertices than that (we will obtain much more precise information later on).

Lemma 4.2. *For any $\nu > 0$, there exist $c = c(\nu)$ and $K_0 = K_0(\nu)$ such that*

$$\mathbb{P}_p(N(\mathfrak{T}) \geq (1+\nu)\xi_p(x)/K \,|\, 0 \leftrightarrow x) \leq c\, e^{-\frac{1}{2}\nu\xi_p(x)},$$

uniformly in $x \in \mathbb{Z}^d$ and in scales $K \geq K_0$.

Proof. The proof is a simple consequence of the energy estimate (1) and the entropy estimate (**??**). Indeed, the latter imply that

$$\begin{aligned}\mathbb{P}_p(N(\mathfrak{T}) \geq (1+\nu)\xi_p(x)/K) &= \sum_{N \geq (1+\nu)\xi_p(x)/K} \mathbb{P}_p(N(\mathfrak{T}) = N) \\ &\leq \frac{1}{2} \sum_{N \geq (1+\nu)\xi_p(x)/K} e^{-K(1-o_K(1))N} \# \{\mathfrak{T} : N(\mathfrak{T}) = N\} \\ &\leq \frac{1}{2} \sum_{N \geq (1+\nu)\xi_p(x)/K} e^{-K(1-o_K(1))N} \\ &\leq e^{-(1+\nu)\xi_p(x)}.\end{aligned}$$

Therefore, since $\mathbb{P}_p(0 \leftrightarrow x) \geq e^{-\xi_p(x)(1+o_x(1))}$,

$$\mathbb{P}_p(N(\mathfrak{T}) \geq \nu\xi_p(x)/K \,|\, 0 \leftrightarrow x) \leq \frac{\mathbb{P}_p(N(\mathfrak{T}) \geq \nu\xi_p(x)/K)}{\mathbb{P}_p(0 \leftrightarrow x)} \leq e^{-\frac{1}{2}\nu\xi_p(x)},$$

for all $\|x\|$ large enough. □

We fix $\nu_1 > 1$ and call trunks such that $N(\mathfrak{t}) \leq \nu_1\xi_p(x)/K$ *admissible.* By Lemma 4.2, admissible trunks are typical:

$$\mathbb{P}_p\big(N(\mathfrak{t}) > \nu_1\xi_p(x)/K \;\big|\; 0 \leftrightarrow x\big) \leq e^{-\frac{1}{2}(\nu_1-1)\xi_p(x)}. \tag{2}$$

Moreover, for any $\epsilon > 0$,

$$\mathbb{P}_p\big(N(\mathfrak{B}) > \epsilon\xi_p(x)/K \;\big|\; 0 \leftrightarrow x\big) \leq \mathbb{P}_p\big(N(\mathfrak{t}) + N(\mathfrak{B}) > (1+\tfrac{1}{2}\epsilon)\xi_p(x)/K \;\big|\; 0 \leftrightarrow x\big) \leq e^{-\frac{1}{4}\epsilon\xi_p(x)}. \tag{3}$$

4.1.3 Geometry of typical trees: refined estimates

We are now going to obtain a much sharper description of typical trees $\mathfrak{T}$, showing in particular that the vast majority of the increments $\mathfrak{t}_{i+1} - \mathfrak{t}_i$ of the trunk $\mathfrak{t}$ are pointing roughly in the direction $\vec{e}_1$.

The surcharge function. An important tool in the following derivation is the surcharge function. The latter was introduced in [6], building on ideas from [3].

The *surcharge functional* associated to a trunk $\mathfrak{t} = (\mathfrak{t}_0, \dots, \mathfrak{t}_N)$ is given by

$$\mathfrak{s}(\mathfrak{t}) = \sum_{k=1}^{N} \mathfrak{s}(\mathfrak{t}_k - \mathfrak{t}_{k-1}),$$

where the *surcharge function* $\mathfrak{s} : \mathbb{R}^d \to \mathbb{R}$ is defined by

$$\mathfrak{s}(y) = \xi_p(y) - \hat{\xi}_p y^{\|}.$$

Since[2] ξ_p is convex and satisfies $\xi_p(y^{\|}, y^{\perp}) = \xi_p(y^{\|}, -y^{\perp})$,

$$\hat{\xi}_p y^{\|} \leq \xi_p(y), \tag{4}$$

for all $y \in \mathbb{R}^d$. In particular, $\mathfrak{s}$ is nonnegative. Moreover, $\mathfrak{s}(y) = 0$ whenever $y^{\|} \geq 0$ and $y^{\perp} = 0$. Observe also that $\mathfrak{s}$ is a convex function and that $\mathfrak{s}(\lambda y) = \lambda\mathfrak{s}(y)$ for all $\lambda \in \mathbb{R}_{\geq 0}$.

Given $\delta \in (0,1)$, we define the *forward cone* $\mathcal{Y}_\delta^{\blacktriangleleft}$ by

$$\mathcal{Y}_\delta^{\blacktriangleleft} = \big\{y \in \mathbb{Z}^d \,:\, \mathfrak{s}(y) \leq \delta\xi_p(y)\big\}$$

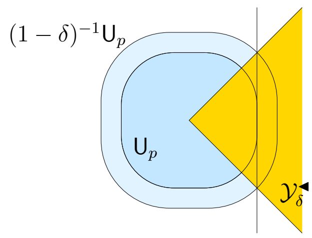


Figure 3: Construction of the forward cone $\mathcal{Y}^{\blacktriangleleft}_\delta$. (To see why this is true, observe that, for all points y such that $y^\parallel = \hat{\xi}_p^{-1}$, one has $y \in \mathcal{Y}^{\blacktriangleleft}_\delta \Leftrightarrow \xi_p(y) \le (1-\delta)^{-1} \Leftrightarrow y \in (1-\delta)^{-1}\mathsf{U}_p$.)

(see Fig. 3) and the *backward cone* by $\mathcal{Y}^{\blacktriangleright}_\delta = -\mathcal{Y}^{\blacktriangleleft}_\delta$. We will use the notation $\mathcal{Y}^{\blacktriangleleft}_\delta(z) = z + \mathcal{Y}^{\blacktriangleleft}_\delta$ and $\mathcal{Y}^{\blacktriangleright}_\delta(z) = z + \mathcal{Y}^{\blacktriangleright}_\delta$.

Note that

$$y = (y^\parallel, y^\perp) \in \mathcal{Y}^{\blacktriangleleft}_\delta \quad \text{if and only if} \quad \hat{\xi}_p y^\parallel \ge (1-\delta)\xi_p(y). \tag{5}$$

Now, observe that there exists a constant c such that, for any K and any $y \in \partial^{\text{ext}}\mathsf{B}_K(0)$, $K \le \xi_p(y) \le K + c$. Since, for all $0 \le i \le N(\mathfrak{t}) - 1$, $\mathfrak{t}_{i+1} - \mathfrak{t}_i \in \partial^{\text{ext}}\mathsf{B}_K(0)$, we can bound

$$\prod_{i=0}^{N(\mathfrak{t})-1} \max_{y \in \partial^{\text{ext}}\mathsf{B}_K(\mathfrak{t}_i)} e^{-\xi_p(y-\mathfrak{t}_i)} \le \exp\Big(cN(\mathfrak{t}) - \sum_{i=1}^{N(\mathfrak{t})} \xi_p(\mathfrak{t}_i - \mathfrak{t}_{i-1})\Big) \le \exp\big(cN(\mathfrak{t}) + 2K - \hat{\xi}_p x^\parallel - \mathfrak{s}(\mathfrak{t})\big).$$

Thus, proceeding similarly as we did in (1), we obtain, for any admissible trunk $\mathfrak{t}$,

$$\mathbb{P}_p(\mathfrak{t}) \le \prod_{i=0}^{N(\mathfrak{t})} \mathbb{P}_p(x_i \overset{A^i}{\longleftrightarrow} \partial^{\text{ext}}\mathsf{B}_K(x_i)) \le \exp\big(-\hat{\xi}_p x^\parallel - \mathfrak{s}(\mathfrak{t}) + c(\log K/K)\xi_p(x)\big), \tag{6}$$

for some constant c, whenever $\xi_p(x) \gg K \gg 1$. This inequality will allow us to reduce the probabilistic problem of eliminating classes of atypical trunks to a purely geometric issue, thanks to the following lemma.

Proposition 4.3. *Let $\epsilon \in (0,1)$. There exist $c > 0$ and a scale $K_0(\epsilon)$ such that*

$$\mathbb{P}_p\big(\mathfrak{s}(\mathfrak{t}) > \epsilon\xi_p(x) \bigm| 0 \leftrightarrow x\big) \le e^{-c\epsilon\xi_p(x)},$$

uniformly in $x \in \mathcal{Y}^{\blacktriangleleft}_{\epsilon/2}$ and scales $K > K_0$.

Proof. Thanks to (2), we can reduce our attention to admissible trunks. The number of such trunks is bounded above by

$$(C(d)K^{d-1})^{\nu_1\xi_p(x)/K} \le \exp\big(c(\log K/K)\xi_p(x)\big)$$

for some $c = c(d)$. It thus follows from (6) that

$$\begin{aligned}\mathbb{P}_p\big(\mathfrak{s}(\mathfrak{t}) > \epsilon\|x\|, N(\mathfrak{t}) \le \nu_1\xi_p(x)/K, 0 \leftrightarrow x\big) &\le \exp\big\{-\hat{\xi}_p x^\parallel - \big(\epsilon - c'(\log K/K)\big)\xi_p(x)\big\} \\ &\le \exp\big\{-(1 + \tfrac{1}{4}\epsilon)\xi_p(x)\big\},\end{aligned}$$

when K is large enough and $\xi_p(x) \gg K$. For the last inequality, we used (5) and the assumption that $x \in \mathcal{Y}^{\blacktriangleleft}_{\epsilon/2}$. The conclusion follows. □

Typical trunks have many cone-points. Given $\delta \in (0, \frac{1}{3})$, we say that the vertex $\mathfrak{t}_k$ is a *forward cone-point of the trunk* $\mathfrak{t} = (\mathfrak{t}_0, \ldots, \mathfrak{t}_{N(\mathfrak{t})})$ if

$$\{\mathfrak{t}_{k+1}, \ldots, \mathfrak{t}_{N(\mathfrak{t})}\} \subset \mathcal{Y}^{\blacktriangleleft}_\delta(\mathfrak{t}_k).$$

Similarly, $\mathfrak{t}_k$ is a *backward cone-point of* $\mathfrak{t}$ if

$$\{\mathfrak{t}_0, \ldots, \mathfrak{t}_{k-1}\} \subset \mathcal{Y}^{\blacktriangleright}_\delta(\mathfrak{t}_k).$$

[2]Just note that $\hat{\xi}_p y^\parallel = \xi_p(y^\parallel, 0) = \xi_p\big(\frac{1}{2}(y^\parallel, y^\perp) + \frac{1}{2}(y^\parallel, -y^\perp)\big) \le \frac{1}{2}\xi_p(y^\parallel, y^\perp) + \frac{1}{2}\xi_p(y^\parallel, -y^\perp) = \xi_p(y^\parallel, y^\perp)$.

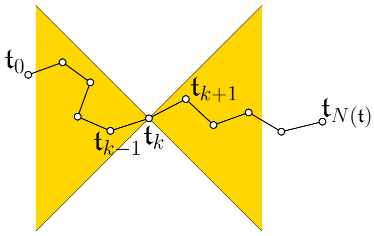


Figure 4: $\mathfrak{t}_k$ is a cone-point of the trunk $\mathfrak{t} = (\mathfrak{t}_0, \ldots, \mathfrak{t}_{N(\mathfrak{t})})$.

Finally, $\mathfrak{t}_k$ is a *cone-point of the trunk* $\mathfrak{t}$ if it is both a forward and a backward cone-point of $\mathfrak{t}$ (see Fig. 4).

We are going to show that most vertices of a typical trunk $\mathfrak{t}$ are cone-points of $\mathfrak{t}$ (for large enough scales K). In order to do this, we are going to mark bad vertices, in such a way that all unmarked vertices will necessarily be cone-points. Namely, if $\mathfrak{t}$ contains vertices that are not forward cone-points, we define

$$\begin{aligned}
\ell_1^{\blacktriangleleft} &= \min\{j : \mathfrak{t}_j \text{ is not a forward cone-point of } \mathfrak{t}\},\\
r_1^{\blacktriangleleft} &= \min\left\{j > \ell_1^{\blacktriangleleft} : \mathfrak{t}_j \notin \mathcal{Y}_\delta^{\blacktriangleleft}(\mathfrak{t}_{\ell_1}^{\blacktriangleleft})\right\},\\
\ell_2^{\blacktriangleleft} &= \min\left\{j \geq r_1^{\blacktriangleleft} : \mathfrak{t}_j \text{ is not a forward cone-point of } \mathfrak{t}\right\},\\
r_2^{\blacktriangleleft} &= \min\left\{j > \ell_2^{\blacktriangleleft} : \mathfrak{t}_j \notin \mathcal{Y}_\delta^{\blacktriangleleft}(\mathfrak{t}_{\ell_2}^{\blacktriangleleft})\right\},\\
&\ \ \vdots
\end{aligned}$$

and, similarly, if $\mathfrak{t}$ contains vertices that are not backward cone-points, we set

$$\begin{aligned}
\ell_1^{\blacktriangleright} &= \max\{j : \mathfrak{t}_j \text{ is not a backward cone-point of } \mathfrak{t}\},\\
r_1^{\blacktriangleright} &= \max\left\{j < \ell_1^{\blacktriangleright} : \mathfrak{t}_j \notin \mathcal{Y}_\delta^{\blacktriangleright}(\mathfrak{t}_{\ell_1}^{\blacktriangleright})\right\},\\
\ell_2^{\blacktriangleright} &= \max\left\{j \leq r_1^{\blacktriangleright} : \mathfrak{t}_j \text{ is not a backward cone-point of } \mathfrak{t}\right\},\\
r_2^{\blacktriangleright} &= \max\left\{j < \ell_2^{\blacktriangleright} : \mathfrak{t}_j \notin \mathcal{Y}_\delta^{\blacktriangleright}(\mathfrak{t}_{\ell_2}^{\blacktriangleright})\right\},\\
&\ \ \vdots
\end{aligned}$$

We then say that $\mathfrak{t}_j$ is a *marked point* of $\mathfrak{t}$ if

$$j \in \bigcup_k \{\ell_k^{\blacktriangleleft}, \ldots, r_k^{\blacktriangleleft} - 1\} \cup \bigcup_k \{r_k^{\blacktriangleright} + 1, \ldots, \ell_k^{\blacktriangleright}\}.$$

Let us denote by $\#_\delta^{\text{marked}}(\mathfrak{t})$ the number of marked points of $\mathfrak{t}$. The next lemma shows that, typically, most vertices of $\mathfrak{t}$ are cone-points.

Lemma 4.4. *Fix $\delta \in (0, \frac{1}{3})$. Then, for any $\epsilon > 0$, there exist $c = c(\epsilon, \delta) > 0$ and $K_0 = K_0(\epsilon, \delta)$ such that*

$$\mathbb{P}_p\big(\#_\delta^{\text{marked}}(\mathfrak{t}) \geq \epsilon N(\mathfrak{t}) \,\big|\, 0 \leftrightarrow x\big) \leq e^{-c\xi_p(x)}, \tag{7}$$

uniformly in $x \in \mathcal{Y}_{\epsilon\delta/14}^{\blacktriangleleft}$ and scales $K \geq K_0$.

Proof. We are going to prove that

$$\mathfrak{s}(\mathfrak{t}) \geq \frac{1}{7}\delta K \#_\delta^{\text{marked}}(\mathfrak{t}). \tag{8}$$

Since $N(\mathfrak{t}) \geq \xi_p(x)/K$ for any trunk $\mathfrak{t}$, the desired conclusion will follow from Proposition 4.3 and (8).

Let us now prove (8). Observe that it is sufficient to show that, for every marked interval $\{\ell_k^{\blacktriangleleft}, \ldots, r_k^{\blacktriangleleft} - 1\}$,

$$\sum_{j=\ell_k^{\blacktriangleleft}+1}^{r_k^{\blacktriangleleft}} \mathfrak{s}(\mathfrak{t}_j - \mathfrak{t}_{j-1}) \geq \frac{1}{7}\delta K (r_k^{\blacktriangleleft} - \ell_k^{\blacktriangleleft}),$$

(a) Large displacement along the first coordinate.

(b) Small displacement along the first coordinate (and hence many increments pointing "in the wrong direction").

Figure 5: The two cases in the proof of Lemma 4.4.

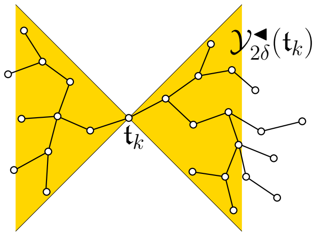


Figure 6: $\mathfrak{t}_k$ is a cone-point of the tree.

the intervals $\bigcup_k\{r^{\blacktriangleright}_k+1,\dots,\ell^{\blacktriangleright}_k\}$ being treated in exactly the same way. To show this, we consider two cases (see Fig. 5):

Case 1: $(\mathfrak{t}_{r^{\blacktriangleleft}_k}-\mathfrak{t}_{\ell^{\blacktriangleleft}_k})^{\parallel}\geq\frac{2}{7}\hat{\xi}_p^{-1}K(1-\delta)(r^{\blacktriangleleft}_k-\ell^{\blacktriangleleft}_k)$. By construction, $\mathfrak{t}_{r^{\blacktriangleleft}_k}-\mathfrak{t}_{\ell^{\blacktriangleleft}_k}\notin\mathcal{Y}^{\blacktriangleleft}_\delta$ and therefore, by convexity and positive homogeneity of $\mathfrak{s}$, (4) and the fact that $0<\delta<\frac{1}{3}$,

$$\sum_{j=\ell^{\blacktriangleleft}_k+1}^{r^{\blacktriangleleft}_k}\mathfrak{s}(\mathfrak{t}_j-\mathfrak{t}_{j-1})\;\geq\;\mathfrak{s}(\mathfrak{t}_{r^{\blacktriangleleft}_k}-\mathfrak{t}_{\ell^{\blacktriangleleft}_k})\;>\;\delta\xi_p(\mathfrak{t}_{r^{\blacktriangleleft}_k}-\mathfrak{t}_{\ell^{\blacktriangleleft}_k})\;\geq\;\delta\hat{\xi}_p(\mathfrak{t}_{r^{\blacktriangleleft}_k}-\mathfrak{t}_{\ell^{\blacktriangleleft}_k})^{\parallel}\;\geq\;\frac{1}{7}\delta K(r^{\blacktriangleleft}_k-\ell^{\blacktriangleleft}_k).$$

Case 2: $(\mathfrak{t}_{r^{\blacktriangleleft}_k}-\mathfrak{t}_{\ell^{\blacktriangleleft}_k})^{\parallel}<\frac{2}{7}\hat{\xi}_p^{-1}K(1-\delta)(r^{\blacktriangleleft}_k-\ell^{\blacktriangleleft}_k)$. Observe that, on the one hand, all increments satisfy

$$(\mathfrak{t}_j-\mathfrak{t}_{j-1})^{\parallel}\geq-2K/\hat{\xi}_p,$$

provided that K be large enough. On the other hand, for any increment satisfying $\mathfrak{t}_j-\mathfrak{t}_{j-1}\in\mathcal{Y}^{\blacktriangleleft}_\delta$, one has

$$(\mathfrak{t}_j-\mathfrak{t}_{j-1})^{\parallel}\geq(1-\delta)\xi_p(\mathfrak{t}_j-\mathfrak{t}_{j-1})/\hat{\xi}_p\geq(1-\delta)K/\hat{\xi}_p.$$

Using this, we can bound from below the number M_k of increments $\mathfrak{t}_j-\mathfrak{t}_{j-1}$, $\ell^{\blacktriangleleft}_k<j\leq r^{\blacktriangleleft}_k$, such that $\mathfrak{t}_j-\mathfrak{t}_{j-1}\notin\mathcal{Y}^{\blacktriangleleft}_\delta$. Indeed,

$$\frac{2}{7}\hat{\xi}_p^{-1}K(1-\delta)(r^{\blacktriangleleft}_k-\ell^{\blacktriangleleft}_k)>(\mathfrak{t}_{r^{\blacktriangleleft}_k}-\mathfrak{t}_{\ell^{\blacktriangleleft}_k})^{\parallel}\geq-2KM_k/\hat{\xi}_p+(r^{\blacktriangleleft}_k-\ell^{\blacktriangleleft}_k-M_k)(1-\delta)K/\hat{\xi}_p,$$

and therefore

$$M_k\geq\frac{1}{7}(r^{\blacktriangleleft}_k-\ell^{\blacktriangleleft}_k).$$

Since each of these increments contributes at least δK to the surcharge, the conclusion follows. □

Typical trees have many cone-points. Now that we know that most vertices of a typical trunk are cone-points, we want to extend this to the full tree, taking branches into account. Let us thus consider a tree $\mathfrak{T}=(\mathfrak{t},\mathfrak{B})$. We say that a vertex $\mathfrak{t}_k$ of the trunk $\mathfrak{t}=(\mathfrak{t}_0,\dots,\mathfrak{t}_{N(\mathfrak{t})})$ is a *cone-point of the tree* $\mathfrak{T}$ if

$$\mathfrak{T}\subset\mathcal{Y}^{\blacktriangleright}_{2\delta}(\mathfrak{t}_k)\cup\mathcal{Y}^{\blacktriangleleft}_{2\delta}(\mathfrak{t}_k).$$

(Note the use of 2δ in this definition and see Fig. 6.) Let us say that a cone-point $\mathfrak{t}_i$ of the trunk $\mathfrak{t}$ is *blocked* if it is not a cone-point of the tree $\mathfrak{T}$. We denote by $\#_\delta^{\text{blocked}}(\mathfrak{T})$ the number of such blocked vertices.

Lemma 4.5. *Fix* $\delta \in (0, \frac{1}{3})$. *Then, there exists* $c = c(\delta) > 0$ *such that, for any* $\epsilon > 0$ *and any* $K \geq K_2(\epsilon, \delta)$,

$$\mathbb{P}_p(\#_\delta^{\text{blocked}}(\mathfrak{T}) \geq \epsilon N(\mathfrak{t}) \,|\, 0 \leftrightarrow x) \leq e^{-c\epsilon\xi_p(x)}, \tag{9}$$

uniformly in $x \in \mathbb{Z}^d$.

Proof. Notice first that, for any blocked cone-point $\mathfrak{t}_i$ of $\mathfrak{t}$, there must exist at least one index $0 \leq j \leq N(\mathfrak{t})$ such that

$$\mathfrak{b}_j \not\subset \mathcal{Y}_{2\delta}^{\blacktriangleright}(\mathfrak{t}_i) \cup \mathcal{Y}_{2\delta}^{\blacktriangleleft}(\mathfrak{t}_i).$$

Moreover, since $\mathfrak{t}_i$ is a cone-point of the trunk, we have in that case that $\mathfrak{t}_j \in \mathcal{Y}_\delta^{\blacktriangleright}(\mathfrak{t}_i) \cup \mathcal{Y}_\delta^{\blacktriangleleft}(\mathfrak{t}_i)$. These observations imply that there must exist a constant $c = c(\delta) > 0$ such that $N(\mathfrak{b}_j) \geq c|(\mathfrak{t}_j - \mathfrak{t}_i)^{\parallel}|/K$. In particular, since two cone-points $\mathfrak{t}_k$ and $\mathfrak{t}_{k'}$, $k > k'$, of the trunk always satisfy

$$(\mathfrak{t}_k - \mathfrak{t}_{k'})^{\parallel} \geq (1-\delta)\xi_p(\mathfrak{t}_k - \mathfrak{t}_{k'})/\hat{\xi}_p \geq (k-k')(1-\delta)K/\hat{\xi}_p,$$

the branch $\mathfrak{b}_j$ can block at most $2N(\mathfrak{b}_j)\hat{\xi}_p/c(1-\delta)$ vertices. This implies that

$$\#_\delta^{\text{blocked}}(\mathfrak{T}) \leq \frac{2\hat{\xi}_p}{c(1-\delta)} \sum_{j=0}^{N(\mathfrak{t})} N(\mathfrak{b}_j) = \frac{2\hat{\xi}_p}{c(1-\delta)} N(\mathfrak{B}).$$

Since $N(\mathfrak{t}) \geq \xi_p(x)/K$, we obtain

$$\mathbb{P}_p\big(\#_\delta^{\text{blocked}}(\mathfrak{T}) \geq \epsilon N(\mathfrak{t}) \,\big|\, 0 \leftrightarrow x\big) \leq \mathbb{P}_p\big(N(\mathfrak{B}) \geq \big(c(1-\delta)\epsilon/2K\hat{\xi}_p\big)\xi_p(x) \,\big|\, 0 \leftrightarrow x\big)$$

and the conclusion thus follows from (3). □

We can now combine the last two lemmas in order to show that most vertices of the trunk of a typical tree $\mathfrak{T}$ are cone-points of $\mathfrak{T}$. Let us write $\#_\delta^{\text{cone}}(\mathfrak{T})$ the number of cone-points of $\mathfrak{T}$.

Theorem 4.6. *Fix* $\delta \in (0, \frac{1}{3})$. *Then, for any* $\epsilon > 0$, *there exist* $c = c(\epsilon, \delta) > 0$ *and* $K_0 = K_0(\epsilon, \delta)$ *such that*

$$\mathbb{P}_p(\#_\delta^{\text{cone}}(\mathfrak{T}) \leq (1-\epsilon)N(\mathfrak{t}) \,|\, 0 \leftrightarrow x) \leq e^{-c\xi_p(x)},$$

uniformly in $x \in \mathcal{Y}_{\epsilon\delta/28}^{\blacktriangleleft}$ *and scales* $K \geq K_0$.

Proof. Since $\#_\delta^{\text{cone}}(\mathfrak{T}) \leq (1-\epsilon)N(\mathfrak{t})$ implies that either $\#_\delta^{\text{marked}}(\mathfrak{t}) \geq \frac{1}{2}\epsilon N(\mathfrak{t})$ or $\#_\delta^{\text{blocked}}(\mathfrak{t}) \geq \frac{1}{2}\epsilon N(\mathfrak{t})$, the claim immediately follows from Lemmas 4.4 and 4.5. □

4.2 Geometry of typical clusters

We now know that the tree associated to typical realizations of long subcritical clusters has a very regular geometry (essentially ballistic trunk, few small branches), at least at large enough scale K. By construction, the cluster remains close to the associated tree, so one should be able to derive similar results for the cluster itself. This is what we do in the present section.

4.2.1 Cone-points of C

Given $\delta \in (0, \frac{1}{3})$, let us say that a vertex $y \in C$ is a *cone-point of the cluster* C if

$$C \subset \mathcal{Y}_{3\delta}^{\blacktriangleright}(y) \cup \mathcal{Y}_{3\delta}^{\blacktriangleleft}(y).$$

(Note the use of 3δ in this definition and see Fig. 7.) Let $\#_\delta^{\text{cone}}(C)$ be the number of cone-points of C.

Theorem 4.7. *Fix* $\delta \in (0, \frac{1}{3})$. *There exist* $K_0 = K_0(\delta)$, $\nu_7 = \nu_7(\delta) > 0$, $\nu_8 > 0$ *and* $c = c(\delta) > 0$ *such that*

$$\mathbb{P}_p\big(\#_\delta^{\text{cone}}(C_{0,x}) \leq \nu_7\|x\| \,\big|\, 0 \leftrightarrow x\big) \leq e^{-c\|x\|},$$

uniformly in $x \in \mathcal{Y}_{\nu_8\delta}^{\blacktriangleleft}$ *and scales* $K \geq K_0$.

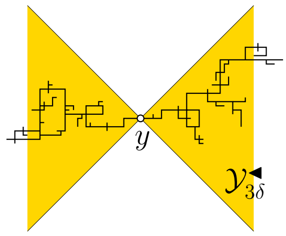


Figure 7: y is a cone-point of the cluster.

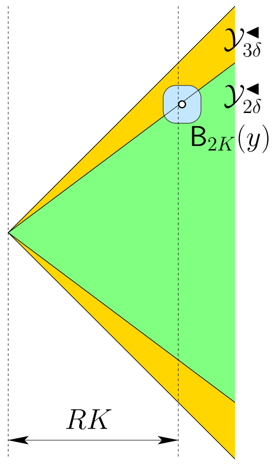


Figure 8: For any point $y \in \mathcal{Y}_{2\delta}^{\blacktriangleleft}$ such that $y^{\parallel} > RK$, $\mathsf{B}_{2K}(y)$ is included inside $\mathcal{Y}_{3\delta}^{\blacktriangleleft}$.

Proof. Let C be a realization of $C_{0,x}$. Of course, cone-points of the tree $\mathfrak{T}$ associated to C are prime candidates to be cone-points of C. We are going to show that, typically, a positive density of the latter are.

Let $R = R(\delta)$ be the smallest integer such that (see Fig. 8)

$$y \in \mathcal{Y}_{2\delta}^{\blacktriangleleft} \text{ and } y^{\parallel} \geq RK \implies \mathsf{B}_{2K}(y) \subset \mathcal{Y}_{3\delta}^{\blacktriangleleft}.$$

Let $L = 2\lceil (R\hat{\xi}_p + 2)/(1-\delta)\rceil) + 1$ and split the set of all vertices of the trunk into packets of size L:

$$\mathcal{S}_k = \{\mathsf{t}_{(k-1)L}, \ldots, \mathsf{t}_{kL-1}\} \qquad (1 \leq k \leq \lfloor N(\mathsf{t})/L \rfloor).$$

Let us say that $\mathcal{S}_k$ is *clean* if it is entirely composed of cone-points of $\mathfrak{T}$ from which no branch originates. Choosing K_0 large enough (depending on δ), we can assume that at least half of the sets $\mathcal{S}_k$ are clean. Indeed, when this is not the case, either there are at most $(1-1/4L)N(\mathsf{t})$ cone-points of $\mathfrak{T}$, which has exponentially small probability by Theorem 4.6, or $N(\mathfrak{B}) \geq (1/4L)N(\mathsf{t}) \geq (1/4L)\xi_p(x)/K$, which has exponentially small probability by (3).

Consider now one clean set $\mathcal{S}_k$ and let us denote by $\mathsf{t}_k^* = \mathsf{t}_{(k-1)L+(L-1)/2}$ the vertex in the middle of this set. Let us also introduce the slab

$$\Delta_k = \bigl\{x \in \mathbb{Z}^d \,:\, |(x - \mathsf{t}_k^*)^{\parallel}| \leq RK\bigr\}.$$

Observe that, by definition of R,

$$C \setminus \Delta_k \subset \mathcal{Y}_{3\delta}^{\blacktriangleright}(\mathsf{t}_k^*) \cup \mathcal{Y}_{3\delta}^{\blacktriangleleft}(\mathsf{t}_k^*),$$

so that, for t_k^* to be a cone-point of C, only the behavior of C inside Δ_k matters. To control the latter it will be convenient to consider the slightly enlarged slab

$$\bar{\Delta}_k = \bigl\{x \in \mathbb{Z}^d \,:\, |(x - \mathsf{t}_k^*)^{\parallel}| \leq RK + 2K/\hat{\xi}_p\bigr\}.$$

We are now going to construct a new cluster C' by modifying C inside $\bar{\Delta}_k$ in such a way that t^* becomes a cone-point of C', while ensuring that $C' \sim \mathfrak{T}$.

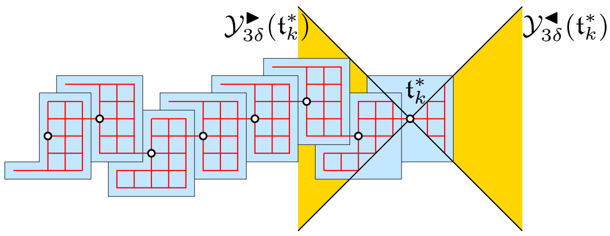


Figure 9: Construction of the cluster C'. Only some of the balls $\mathsf{B}_K(\mathsf{t}_\ell)$ are shown (balls have been represented as squares to simplify drawing). Note that the open edges (in red in the picture) are the only ones kept inside Δ_k. In $\bar{\Delta}_k \setminus \Delta$, these edges are added when absent, but no edge is removed (this is to ensure that C' is connected).

Observe now that, by construction, $\mathsf{t}_{(k-1)L}$ and t_{kL-1} lie outside $\bar{\Delta}_k$. Write

$$A^\ell = \big(\mathsf{B}_K(t_\ell) \setminus \mathsf{B}_K(t_{\ell-1})\big) \cap \big(\mathcal{Y}^{\blacktriangleleft}_{3\delta}(t^*_k) \cup \mathcal{Y}^{\blacktriangleright}_{3\delta}(t^*_k)\big).$$

To define C', we proceed as follows (see Fig. 9):

- For each $(k-1)L < \ell < kL-1$: add to C all edges inside A^ℓ.
- For each $(k-1)L < \ell < kL-1$ such that $\mathsf{t}_\ell \in \Delta_k$: remove from C all edges $\partial^{\mathrm{ext}} A^\ell$, except those incoming on $\mathsf{t}_{\ell-1}$ or $\mathsf{t}_{\ell+1}$.
- the cluster C' is the resulting connected component containing 0.

Observe now that the previous construction ensures that

- The vertices t_ℓ, $(k-1)L < \ell < kL-1$, are connected to each other.
- C' coincides with C outside $\bar{\Delta}_k$.
- Each connected components of $C' \setminus \bar{\Delta}_k$ is necessarily connected outside Δ_k to some vertex $t_j \in \bar{\Delta}_k$ (this follows from the fact that $t_{(k-1)L}$ and t_{kL-1} are cone-points of $\mathfrak{T}$ and from the fact that the cluster C remains within a ξ_p-distance at most $2K$ from $\mathfrak{T}$).
- $C' \sim \mathfrak{T}$.
- t^*_k is a cone-point of C'.

We will write $C \rightsquigarrow C'$ to mean that applying the above construction to the cluster C yields the cluster C'. Since at most cK^d edges have been modified, $\mathbb{P}_p(C_{0,x} = C') \geq e^{-cK^d}\mathbb{P}_p(C_{0,x} = C)$ for any $C \rightsquigarrow C'$. This implies that

$$\begin{aligned}\mathbb{P}_p(\mathfrak{T}) = \sum_{C\sim\mathfrak{T}} \mathbb{P}_p(C_{0,x} = C) &\leq e^{cK^d} \sum_{C\sim\mathfrak{T}} \mathbb{P}_p(C_{0,x} = C') \\ &\leq e^{cK^d} \sum_{\substack{C'\sim\mathfrak{T} \\ t^*_k \text{ cone-point of } C'}} \mathbb{P}_p(C_{0,x} = C')\,\#\{C \sim \mathfrak{T} : C \rightsquigarrow C'\} \\ &\leq e^{c'K^d} \sum_{\substack{C'\sim\mathfrak{T} \\ t^*_k \text{ cone-point of } C'}} \mathbb{P}_p(C_{0,x} = C') \\ &= e^{c'K^d}\mathbb{P}_p(\mathfrak{T}, t^*_k \text{ cone-point of } C_{0,x}),\end{aligned}$$

for some $c' > 0$. Thus, $\mathbb{P}_p(t^*_k \text{ cone-point of } C_{0,x} \,|\, \mathfrak{T}) \geq e^{-c'K^d} \equiv \alpha(K)$. Of course, we can apply the same argument for an arbitrary collection of clean sets $\mathcal{S}_{k_1}, \ldots, \mathcal{S}_{k_m}$; we thus conclude that

$$\mathbb{P}_p(t^*_{k_1}, \ldots, t^*_{k_m} \text{ are cone-points of } C_{0,x} \,|\, \mathfrak{T}) \geq \alpha(K)^m.$$

It thus follows that the collection of random variables $\mathbf{1}_{\{t^*_k \text{ is a cone-point of } C_{0,x}\}}$, where k runs overs all values of k for which $\mathcal{S}_k$ is clean, stochastically dominates a collection of i.i.d. Bernoulli random variables of parameter $\alpha(K)$. The number of clean sets being at least $\lfloor \frac{1}{2}N(\mathsf{t})/L \rfloor$, the conclusion follows. □

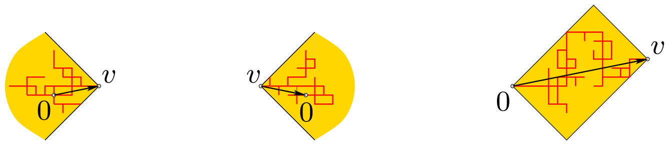


Figure 10: Backward-irreducible, forward-irreducible and irreducible clusters and their displacement.

4.2.2 Decomposition of C into irreducible pieces

Theorem 4.7 leads naturally to a decomposition of the cluster $C_{0,x}$ into irreducible pieces. We define three notions of irreducibility and a *displacement* application X associating to an irreducible piece a vector in $\mathbb{Z}^d$ (see Fig. 10):

- A cluster $C \ni 0$ is *backward-irreducible* if it possesses a cone-point v such that $C \subset \mathcal{Y}^{\blacktriangleright}_{3\delta}(v)$ and no other cone-points y with $y^{\parallel} \geq 0$. In this case, $X(C) = v$.
- A cluster $C \ni 0$ is *forward-irreducible* if it possesses a cone-point v such that $C \subset \mathcal{Y}^{\blacktriangleleft}_{3\delta}(v)$ and no other cone-points y with $y^{\parallel} \leq 0$. In this case, $X(C) = -v$.
- A cluster C is *irreducible* if it possesses exactly two cone-points 0 and v, and $C \subset D_{3\delta}(0, v)$, where, as before, $D_{3\delta}(u, u') = \mathcal{Y}^{\blacktriangleleft}_{3\delta}(u) \cap \mathcal{Y}^{\blacktriangleright}_{3\delta}(u')$ is the *diamond* with endpoints u, u'. In this case, $X(C) = v$.

(For convenience, we allow backward-irreducible and forward-irreducible clusters to consist of the single vertex 0.) We denote the sets of all backward-irreducible, forward-irreducible and irreducible pieces $\mathcal{I}_{\rm b}$, $\mathcal{I}_{\rm f}$ and $\mathcal{I}$ respectively.

A cluster $C_{\rm b} \in \mathcal{I}_{\rm b}$ can be naturally concatenated with $m \geq 0$ clusters $C_1, \ldots, C_m \in \mathcal{I}$ by first translating each C_k by $X(C_{\rm b}) + X(C_1) + \cdots + X(C_{k-1})$. We can then also concatenate a cluster $C_{\rm f} \in \mathcal{I}_{\rm f}$ by first translating it by $X(C_{\rm b}) + \sum_{k=1}^{m} X(C_k) + X(C_{\rm f})$. The resulting cluster is denoted $C_{\rm b} \sqcup C_1 \sqcup \cdots \sqcup C_m \sqcup C_{\rm f}$.

Let now $x \in \mathbb{Z}^d$ be such that $x^{\perp} = 0$ and $x^{\parallel} > 0$ large. In view of Theorem 4.7, we can restrict our attention to realizations C of $C_{0,x}$, with at least two cone-points with first component in $[0, x^{\parallel}]$. Such realizations can then be decomposed in a unique way as

$$C = C_{\rm b} \sqcup C_1 \sqcup \cdots \sqcup C_N \sqcup C_{\rm f},$$

where $C_{\rm b} \in \mathcal{I}_{\rm b}$, $C_{\rm f} \in \mathcal{I}_{\rm f}$, $C_1, \ldots, C_N \in \mathcal{I}$ and $N \geq 1$ depends on C. Observe that, in this case,

$$\begin{aligned} \mathbb{P}_p(C_{0,x} = C) &= \mathbb{P}_p(\omega_e = 1 \text{ for all } e \in C, \omega_e = 0 \text{ for all } e \in \partial^{\rm ext} C) \\ &= \tilde{\mathsf{w}}(C_{\rm b})\tilde{\mathsf{w}}(C_{\rm f}) \prod_{k=1}^{N} \tilde{\mathsf{w}}(C_k), \end{aligned}$$

where we have introduced the weights

$$\begin{aligned} \tilde{\mathsf{w}}(C_k) &= (1-p)^{-2d}\mathbb{P}_p(\omega_e = 1 \text{ for all } e \in C_k, \omega_e = 0 \text{ for all } e \in \partial^{\rm ext} C_k), \\ \tilde{\mathsf{w}}(C_{\rm b}) &= (1-p)^{-d}\mathbb{P}_p(\omega_e = 1 \text{ for all } e \in C_{\rm b}, \omega_e = 0 \text{ for all } e \in \partial^{\rm ext} C_{\rm b}), \\ \tilde{\mathsf{w}}(C_{\rm f}) &= (1-p)^{-d}\mathbb{P}_p(\omega_e = 1 \text{ for all } e \in C_{\rm f}, \omega_e = 0 \text{ for all } e \in \partial^{\rm ext} C_{\rm f}). \end{aligned}$$

It turns out that these are not the most natural weights to work with. Namely, it is much more convenient to consider the weights

$$\mathsf{w}(C_k) = e^{\hat{\xi}_p X(C_k)^{\parallel}} \tilde{\mathsf{w}}(C_k), \quad \mathsf{w}(C_{\rm b}) = e^{\hat{\xi}_p X(C_{\rm b})^{\parallel}} \tilde{\mathsf{w}}(C_{\rm b}) \quad \text{and} \quad \mathsf{w}(C_{\rm f}) = e^{\hat{\xi}_p X(C_{\rm f})^{\parallel}} \tilde{\mathsf{w}}(C_{\rm f}),$$

for which we have the identity

$$e^{\xi_p(x)}\, \mathbb{P}_p(C_{0,x} = C) = \mathsf{w}(C_{\rm b})\mathsf{w}(C_{\rm f}) \prod_{k=1}^{N} \mathsf{w}(C_k).$$

The reasons these weights are better are summarized in the following

Proposition 4.8. *1. The weights* w *have exponential tails: there exist* $c = c(\delta) > 0$ *and* $c' > 0$ *such that, for any* $y \in \mathbb{Z}^d$,

$$\sum_{\substack{C\in\mathcal{I}\cup\mathcal{I}_{\mathrm{b}}\cup\mathcal{I}_{\mathrm{f}}:\\ X(C)=y}} \mathsf{w}(C) \leq c' e^{-c\|y\|}. \tag{10}$$

2. w *is a probability mass function on* $\mathcal{I}$: $\sum_{C\in\mathcal{I}} \mathsf{w}(C) = 1$.

Proof. 1. Assume first that $y \notin \mathcal{Y}^{\blacktriangleleft}_{\nu_8\delta}$, then

$$\sum_{\substack{C\in\mathcal{I}\cup\mathcal{I}_{\mathrm{b}}\cup\mathcal{I}_{\mathrm{f}}:\\ X(C)=y}} \mathsf{w}(C) \leq (1-p)^{-2d} e^{\hat{\xi}_p y^{\|}} \mathbb{P}_p(0 \leftrightarrow y) \leq (1-p)^{-2d} e^{\hat{\xi}_p y^{\|}} e^{-\xi_p(y)} \leq (1-p)^{-2d} e^{-\nu_8\delta\xi_p(y)},$$

where the last inequality follows from (5).

Suppose now that $y \in \mathcal{Y}^{\blacktriangleleft}_{\nu_8\delta}$. In this case, it follows from Theorem 4.7 that

$$\sum_{\substack{C\in\mathcal{I}:\\ X(C)=y}} \mathsf{w}(C) \leq (1-p)^{-2d} e^{\hat{\xi}_p y^{\|} - \xi_p(y)} \mathbb{P}_p\big(\#^{\mathrm{cone}}_{\delta} \leq \nu_7\|y\| \bigm| 0 \leftrightarrow y\big) \leq (1-p)^{-2d} e^{-c\|y\|},$$

since $\mathbb{P}_p(0 \leftrightarrow y) \leq e^{-\xi_p(y)} \leq e^{-\hat{\xi}_p y^{\|}}$ (remember (4)). The corresponding claims for $C \in \mathcal{I}_{\mathrm{b}} \cup \mathcal{I}_{\mathrm{f}}$ are proved in a similar manner. The only additional issue is that the cluster C might extend "outward"; namely, when $C \in \mathcal{I}_{\mathrm{b}}$, for instance, C may have many cone-points on the left of 0, so that the above argument must be amended. To solve this minor problem, first observe that the above argument does apply if there are no such cone-point. Therefore, suppose that there is at least one such cone-point, and denote by $z \in \mathcal{Y}^{\blacktriangleright}_{3\delta}$ the one with $z^{\|}$ maximal. Since z is a cone-point of C, one can disconnect the part of C at the left of z at a cost at most $p/(1-p)$ (by closing at most one edge). One is then reduced to a sum over a cluster $C \in \mathcal{I}$ with $X(C) = y + z$, for which the above analysis applies. Summing over z yields the desired conclusion.

2. Let us consider the generating functions

$$\mathsf{G}(s) = \sum_{n\in\mathbb{Z}_{\geq 0}} s^n e^{\hat{\xi}_p n} \mathbb{P}_p\big(0 \leftrightarrow x_n\big) \quad\text{and}\quad \mathsf{H}(s) = \sum_{C\in\mathcal{I}} s^{X(C)^{\|}} \mathsf{w}(C),$$

where we wrote x_n for the point in $\mathbb{Z}^d$ with $x_n^{\|} = n$ and $x_n^{\perp} = 0$. Let us write ρ_{G} and ρ_{H} for the radii of convergence of G and H. Observe that $\rho_{\mathsf{G}} = 1$, since $\mathbb{P}_p\big(0 \leftrightarrow x_n\big) = e^{-\hat{\xi}_p n(1+o_n(1))}$, while (10) implies that $\rho_{\mathsf{H}} > 1$.

Now, we can write

$$e^{\hat{\xi}_p n} \mathbb{P}_p\big(0 \leftrightarrow x_n\big) = \sum_{N\geq 1} \sum_{C_{\mathrm{b}},C_{\mathrm{f}},C_1,\dots,C_N} \mathbf{1}_{\{X(C_{\mathrm{b}})+X(C_{\mathrm{f}})+\sum_{k=1}^N X(C_k)=x_n\}} \mathsf{w}(C_{\mathrm{b}})\mathsf{w}(C_{\mathrm{f}}) \prod_{k=1}^N \mathsf{w}(C_k) + \mathsf{r}_n,$$

where the sum is over possible decompositions into irreducible pieces and $|\mathsf{r}_n| \leq e^{-cn}$ for some $c > 0$. Writing $\mathsf{R}(s) = \sum_{n\in\mathbb{Z}_{\geq 0}} s^n \mathsf{r}_n$ and substituting the above into the expression for G yields, for $s \geq 0$,

$$\begin{aligned}
\mathsf{G}(s) &\leq \sum_{C_{\mathrm{b}},C_{\mathrm{f}}} s^{X(C_{\mathrm{b}})^{\|}+X(C_{\mathrm{f}})^{\|}} \mathsf{w}(C_{\mathrm{b}})\mathsf{w}(C_{\mathrm{f}}) \sum_{N\geq 1} \Big(\sum_{C\in\mathcal{I}} s^{X(C)^{\|}} \mathsf{w}(C)\Big)^N + \mathsf{R}(s) \\
&= \sum_{C_{\mathrm{b}},C_{\mathrm{f}}} s^{X(C_{\mathrm{b}})^{\|}+X(C_{\mathrm{f}})^{\|}} \mathsf{w}(C_{\mathrm{b}})\mathsf{w}(C_{\mathrm{f}}) \sum_{N\geq 1} \mathsf{H}(s)^N + \mathsf{R}(s) \\
&= \frac{\mathsf{H}(s)}{1-\mathsf{H}(s)} \sum_{C_{\mathrm{b}},C_{\mathrm{f}}} s^{X(C_{\mathrm{b}})^{\|}+X(C_{\mathrm{f}})^{\|}} \mathsf{w}(C_{\mathrm{b}})\mathsf{w}(C_{\mathrm{f}}) + \mathsf{R}(s).
\end{aligned}$$

Note that, by (10), the sums over C_{b} and C_{f} are convergent when s is slightly larger than 1. The radius of convergence of R is at least $e^c > 1$, and we already saw that $\rho_{\mathsf{H}} > 1$. Since $\rho_{\mathsf{G}} = 1$, we conclude that $\mathsf{H}(1) = 1$, which proves the claim. □

To summarize, there exists a constant $c>0$ such that, for any function f of the cluster C_{0,x_n},

$$\Big|e^{\hat\xi_p n}\,\mathbb{E}_p(f(C_{0,x_n})\mathbf{1}_{\{0\leftrightarrow x_n\}}) - \sum_{\substack{N\geq 1, C_\mathrm{b}, C_\mathrm{f}, C_1,\dots,C_N:\\ X(C_\mathrm{b})+X(C_\mathrm{f})+\sum_{k=1}^N X(C_k)=n}} f(C_\mathrm{b}\sqcup C_1\sqcup\dots\sqcup C_N\sqcup C_\mathrm{f})\,\mathsf{w}(C_\mathrm{b})\mathsf{w}(C_\mathrm{f})\prod_{k=1}^N \mathsf{w}(C_k)\Big| \leq \|f\|_\infty\, e^{-cn}. \quad (11)$$

4.2.3 The effective directed random walk

Let us now define three positive measures on $\mathbb{Z}^d$: for any $y\in\mathbb{Z}^d$,

$$\rho_\mathrm{b}(y) = \sum_{\substack{C_\mathrm{b}\in\mathcal{I}_\mathrm{b}\\ X(C_\mathrm{b})=y}} \mathsf{w}(C_\mathrm{b}), \qquad \rho_\mathrm{f}(y) = \sum_{\substack{C_\mathrm{f}\in\mathcal{I}_\mathrm{f}\\ X(C_\mathrm{f})=y}} \mathsf{w}(C_\mathrm{f}), \qquad \mathsf{p}(y) = \sum_{\substack{C\in\mathcal{I}\\ X(C)=y}} \mathsf{w}(C).$$

By Proposition 4.8 and symmetries of the model, we know that

- $\rho_\mathrm{b}, \rho_\mathrm{f}$ and p have exponential tails;
- ρ_b, ρ_f are supported on $\mathcal{Y}^{\blacktriangleleft}_{3\delta}$ and $\rho_\mathrm{b}(0), \rho_\mathrm{f}(0)>0$;
- p is a probability measure supported on $\mathcal{Y}^{\blacktriangleleft}_{3\delta}\setminus\{0\}$;
- $\mathsf{p}(y^\parallel, y^\perp) = \mathsf{p}(y^\parallel, -y^\perp)$.

We denote by P_u the law of the directed random walk $(S_n)_{n\geq 0}$ on $\mathbb{Z}^d$, starting at $u\in\mathbb{Z}^d$ and with increments of law p, and by E_u the corresponding expectation. Given $v\in\mathbb{Z}^d$, let us consider the event

$$\mathcal{R}_v = \{\exists N\geq 1 \,:\, S_N = v\} \quad (12)$$

and let us write $\mathsf{P}_{u,v}(\cdot) = \mathsf{P}_u(\cdot\,|\,\mathcal{R}_v)$.

Note that, by the local limit theorem, for any fixed $\alpha>0$,

$$\mathsf{P}_0(S_k = y) = \frac{1+o_k(1)}{\sqrt{(2\pi k)^d \det D}} \exp\bigl\{-(y-k\vec\mu)\cdot D^{-1}(y-k\vec\mu)/2k\bigr\},$$

uniformly in y such that $\|y-k\vec\mu\|_2 \leq k^{2/3-\alpha}$. In the above formula,

- $\vec\mu = \mathsf{E}_0(S_1) = \mu\vec e_1$,
- $D = (D_{ij})_{1\leq i,j\leq d}$ is the covariance matrix: $D_{ij} = \mathrm{Cov}_{\mathsf{P}_0}(S_1(i), S_1(j))$, where $S_1(\ell) = S_1\cdot\vec e_\ell$.

It is now straightforward to derive asymptotic estimates for $\mathsf{P}_0(\mathcal{R}_y)$ from (12). Let $n\in\mathbb{Z}^d_0$ be large. Assume that $\|y-n\vec\mu\|_2\leq n^{1/2-\alpha}$. Fix $\epsilon>0$ small. Since the increments of the random walk have exponential tails, a standard large deviation upper bound shows that

$$\mathsf{P}_0(\exists N \,:\, |N-n| > n^{1/2+\epsilon}, S_N = y) \leq e^{-cn^{2\epsilon}},$$

for some $c>0$. Let us now consider the remaining values of N: by (12) and our assumption on y,

$$\begin{aligned}
\sum_{N:\,|N-n|\leq n^{1/2+\epsilon}} \mathsf{P}_0(S_N=y) &= \frac{1+o_n(1)}{\sqrt{(2\pi n)^d\det D}} \sum_{N:\,|N-n|\leq n^{1/2+\epsilon}} \exp\bigl\{-(y-N\vec\mu)\cdot D^{-1}(y-N\vec\mu)/2N\bigr\}\\
&= \frac{1+o_n(1)}{\sqrt{(2\pi n)^d\det D}} \sum_{N:\,|N-n|\leq n^{1/2+\epsilon}} \exp\bigl\{-\tfrac12\bigl(\tfrac{N-n}{\sqrt n}\bigr)^2\vec\mu\cdot D^{-1}\vec\mu\bigr\}\\
&= \frac{(1+o_n(1))}{\sqrt{(2\pi)^d n^{d-1}\det D}} \sum_{N:\,|N-n|\leq n^{1/2+\epsilon}} n^{-1/2}\exp\bigl\{-\tfrac12\bigl(\mu\tfrac{N-n}{\sqrt n}\bigr)^2 (D^{-1})_{1,1}\bigr\}\\
&= \frac{(1+o_n(1))}{\sqrt{(2\pi)^d n^{d-1}\det D}} \int_{-\infty}^{\infty} \exp\bigl\{-\tfrac12 s^2\mu^2(D^{-1})_{1,1}\bigr\}\mathrm{d}s\\
&= \frac{(1+o_n(1))}{\sqrt{(2\pi n)^{d-1}\mu^2(D^{-1})_{1,1}\det D}}.
\end{aligned}$$

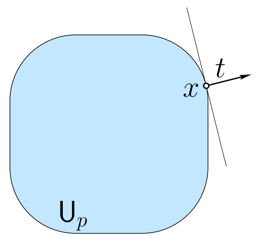


Figure 11: The polars to a point $x \in \partial\mathsf{U}_p$ are given by all vectors normal to a supporting hyperplane to U_p at x.

Therefore,

$$\mathsf{P}_0(\mathcal{R}_y) = \check{C} n^{-(d-1)/2}\,(1+o_n(1)), \tag{13}$$

with $\check{C} = \big((2\pi)^{d-1}\mu^2(D^{-1})_{1,1}\det D\big)^{-1/2}$, uniformly in y such that $\|y - n\vec{\mu}\|_2 \leq n^{1/2-\alpha}$.

4.2.4 One application: Ornstein–Zernike asymptotics for connectivities

Using the notations from Section 4.2.3, it follows from (11) that

$$\Big| e^{\hat{\xi}_p n}\,\mathbb{P}_p(0 \leftrightarrow x_n) - \sum_{\substack{u \in \mathcal{Y}^{\blacktriangleleft}_{3\delta} \\ v \in x_n + \mathcal{Y}^{\blacktriangleright}_{3\delta}}} \rho_{\mathrm{b}}(u)\rho_{\mathrm{f}}(x_n - v)\,\mathsf{P}_u(\mathcal{R}_v) \Big| \leq e^{-cn},$$

for some $c > 0$. Fix $\alpha > 0$. On the one hand, since ρ_{b} and ρ_{f} have exponential tails, the contributions of u, v such that $\|u\|_2, \|x_n - v\|_2 > n^{1/2-\alpha}$ is of order $e^{-O(n^{1/2-\alpha})}$. On the other hand, it follows from (13) that, uniformly in u, v such that $\|u\|_2, \|x_n - v\|_2 \leq n^{1/2-\alpha}$,

$$\frac{\mathsf{P}_u(\mathcal{R}_v)}{\mathsf{P}_0(\mathcal{R}_{x_n})} = 1 + o_n(1) \qquad \text{and} \qquad \mathsf{P}_0(\mathcal{R}_{x_n}) = (1+o_n(1))\,\check{C}\,\mu^{(d-1)/2}\,\|x_n\|_2^{-(d-1)/2}.$$

In particular, we obtain the celebrated *Ornstein–Zernike asymptotics*:

$$\mathbb{P}_p(0 \leftrightarrow x_n) = \frac{\Psi_p}{\|x_n\|_2^{(d-1)/2}}\, e^{-\xi_p(x_n)}\,(1+o_n(1)),$$

with $\Psi_p = \check{C}\mu^{(d-1)/2} \sum_{\substack{u \in \mathcal{Y}^{\blacktriangleleft}_{3\delta} \\ v \in x_n + \mathcal{Y}^{\blacktriangleright}_{3\delta}}} \rho_{\mathrm{b}}(u)\,\rho_{\mathrm{f}}(x_n - v)$.

4.3 Generalizations: some remarks

In this short subsection, we make some remarks concerning extensions of the results proved in this section.

4.3.1 General directions

The analysis for a general cluster $C_{0,x}$, with $x \in \mathbb{Z}^d$ not necessarily on the first coordinate axis, follows the same pattern. In fact, apart from minor notational adjustments, the only real change occurs when defining the surcharge function $\mathfrak{s}$. In order to state the proper generalization of the surcharge function, it is convenient to introduce the *polar set*[3] K_p of the unit ball U_p, defined by

$$\mathsf{K}_p = \big\{ t \in \mathbb{R}^d \,:\, \max_{y \in \mathsf{U}_p} \langle t, y \rangle \leq 1 \big\}.$$

The set K_p is a compact convex set with nonempty interior ($0 \in \mathring{\mathsf{K}}_p$). As is easily checked, for any $x \in \partial\mathsf{U}_p$ and $t \in \partial\mathsf{K}_p$,

$$\max_{y \in \mathsf{U}_p} \langle t, y \rangle = \max_{s \in \mathsf{K}_p} \langle s, x \rangle = \xi_p(x) = 1.$$

[3] Often, K_p is also called the *Wulff shape*. This is due to the fact that, in dimension 2, the relation between K_p and ξ_p precisely mirrors that of the Wulff shape and the surface tension. See [4] for additional information on this topic.

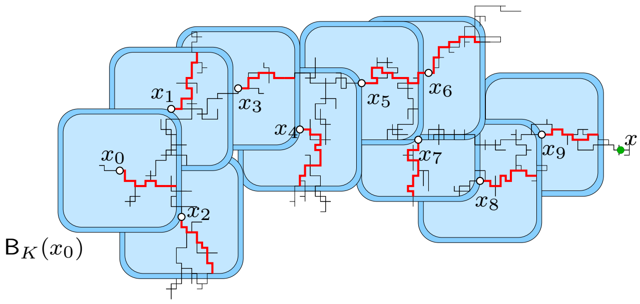


Figure 12: Construction of the skeleton tree $\mathfrak{T}(C)$ associated to a realization C of the cluster $C_{0,x}$ in FK percolation. The introduction of an additional layer at the boundary of the balls ensures good decoupling properties of the connection events.

Finally, given $x \in \mathbb{R}^d \setminus \{0\}$, a point $t \in \partial \mathsf{K}_p$ is said to be *polar* to x if $\langle t, x\rangle = \max_{s\in\mathsf{K}_p} \langle s, x\rangle = \xi_p(x)$; see Fig. 11 for an illustration[4].

Given $x \in \mathbb{Z}^d$, let $t \in \partial\mathsf{K}_p$ be polar to x. The surcharge function is then defined by

$$\mathfrak{s}(y) = \xi_p(y) - \langle t, y\rangle\,.$$

Observe that, when $x^\perp = 0$, we recover the definition used in Section 4.1.3, since $\hat{\xi}_p\vec{e}_1 \in \mathsf{K}_p$ is polar to x by symmetry of U_p. Moreover, the second term remains linear in y and satisfies $\langle t, y\rangle \le \xi_p(y)$ for all $y \in \mathbb{R}^d$ (by definition of K_p), which is the analogue of (4) here. All our analysis thus holds in the same way as before. We refer the reader to [10] for the details.

4.3.2 FK percolation

Ising and Potts models are intimately related to a generalization of Bernoulli percolation known as FK percolation (or random-cluster model), and the latter plays an essential role in numerous analyses of these models. We will not discuss it here and refer the interested reader to [16] for more information. FK percolation shares many properties with Bernoulli percolation (including the availability of the FKG inequality, at least for the versions relevant to the Ising and Potts models). One major difference, however, is the lack of independence. This makes it necessary to amend some of the above arguments when extending the Ornstein–Zernike analysis to this setting. We will not discuss the details of the construction here (see [10] for that), but will point out the two main points at which independence played a crucial role: (i) in the derivation of the energy estimate (1), (ii) in order to have independent increments in the resulting effective random walk. Both of these issues are solved by exploiting suitable exponential mixing estimates.

First, the coarse-graining procedure must be slightly modified, by adding layers of width $r \log K$ (for some fixed, large enough r) around the balls $\mathsf{B}_K(x_i)$ in Algorithm 1; see Fig. 12. This ensures that the connection events occur far enough from each other to recover a factor $e^{-K(1-o_K(1))}$ for each vertex of the skeleton tree, while not modifying substantially the ξ_p-distance between successive vertices.

The analysis then goes through essentially verbatim up until Section 4.2.2. One still obtains a string of irreducible pieces, but it is no longer true that the probability nicely factorizes over the irreducible pieces (the property that led ultimately to the appearance of the effective random walk). Nevertheless, one can then exploit exponential mixing in at least two ways in order to extract the relevant information. The first way, used in [8, 10], is to observe that the resulting process, while not Markovian, is amenable to an analysis in terms of a suitable Ruelle–Perron–Frobenius operator. One can then derive local limit results for the process that have the same qualitative features as those obtained for processes with independent increments. The second approach, developed in the Appendix of [31], relies on ideas from perfect simulation and shows that it is possible to recover a genuine random walk (with independent increments), at the cost of randomly aggregating consecutive

[4]Note that, a priori, there may be a continuum of polars to a given point x. A consequence of the Ornstein–Zernike analysis, which we will not discuss here, is that the sets U_p and K_p have locally analytic boundaries with Gaussian curvature uniformly bounded away from 0. This implies, a posteriori, the existence of a one-to-one mapping between points $x \in \partial\mathsf{U}_p$ and their polar. See [6] for more information on this topic.

irreducible pieces in a suitable manner. The resulting random walk still enjoys all the properties that we derived for Bernoulli percolation, including exponential tails for the increments.

Acknowledgments. This set of informal notes was written for the EMS-IAMP Summer School in Mathematical Physics "Universality in Probability Theory and Statistical Mechanics", which took place in Ischia, 11-15 June 2018. The author would like to express his gratitude to the organizers, Alessandro Giuliani, Serena Cenatiempo, Rafael Greenblatt and Domenico Monaco, for this very enjoyable meeting, as well as to Sébastien Ott for corrections and remarks.

References

[1] D. B. Abraham and H. Kunz. Ornstein-Zernike theory of classical fluids at low density. *Phys. Rev. Lett.*, 39(16):1011–1014, 1977.

[2] M. Aizenman and D. J. Barsky. Sharpness of the phase transition in percolation models. *Comm. Math. Phys.*, 108(3):489–526, 1987.

[3] K. S. Alexander. Mixing properties and exponential decay for lattice systems in finite volumes. *Ann. Probab.*, 32(1A):441–487, 2004.

[4] T. Bodineau, D. Ioffe, and Y. Velenik. Rigorous probabilistic analysis of equilibrium crystal shapes. *J. Math. Phys.*, 41(3):1033–1098, 2000.

[5] M. Campanino, J. T. Chayes, and L. Chayes. Gaussian fluctuations of connectivities in the subcritical regime of percolation. *Probab. Theory Related Fields*, 88(3):269–341, 1991.

[6] M. Campanino and D. Ioffe. Ornstein-Zernike theory for the Bernoulli bond percolation on $\mathbb{Z}^d$. *Ann. Probab.*, 30(2):652–682, 2002.

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

[8] M. Campanino, D. Ioffe, and Y. Velenik. Ornstein-Zernike theory for finite range Ising models above T_c. *Probab. Theory Related Fields*, 125(3):305–349, 2003.

[9] M. Campanino, D. Ioffe, and Y. Velenik. Random path representation and sharp correlations asymptotics at high-temperatures. In *Stochastic analysis on large scale interacting systems*, volume 39 of *Adv. Stud. Pure Math.*, pages 29–52. Math. Soc. Japan, Tokyo, 2004.

[10] M. Campanino, D. Ioffe, and Y. Velenik. Fluctuation theory of connectivities for subcritical random cluster models. *Ann. Probab.*, 36(4):1287–1321, 2008.

[11] 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.

[12] L. Coquille, H. Duminil-Copin, D. Ioffe, and Y. Velenik. On the Gibbs states of the noncritical Potts model on $\mathbb{Z}^2$. *Probab. Theory Related Fields*, 158(1-2):477–512, 2014.

[13] L. Coquille and Y. Velenik. A finite-volume version of Aizenman-Higuchi theorem for the 2d Ising model. *Probab. Theory Related Fields*, 153(1-2):25–44, 2012.

[14] S. Friedli, D. Ioffe, and Y. Velenik. Subcritical percolation with a line of defects. *Ann. Probab.*, 41(3B):2013–2046, 2013.

[15] L. Greenberg and D. Ioffe. On an invariance principle for phase separation lines. *Ann. Inst. H. Poincaré Probab. Statist.*, 41(5):871–885, 2005.

[16] G. Grimmett. *The random-cluster model*, volume 333 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2006.

[17] A. Hammond. Phase separation in random cluster models III: circuit regularity. *J. Stat. Phys.*, 142(2):229–276, 2011.

[18] A. Hammond. Phase separation in random cluster models I: Uniform upper bounds on local deviation. *Comm. Math. Phys.*, 310(2):455–509, 2012.

[19] A. Hammond. Phase separation in random cluster models II: The droplet at equilibrium, and local deviation lower bounds. *Ann. Probab.*, 40(3):921–978, 2012.

[20] D. Ioffe. Ornstein-Zernike behaviour and analyticity of shapes for self-avoiding walks on $\mathbf{Z}^d$. *Markov Process. Related Fields*, 4(3):323–350, 1998.

[21] D. Ioffe. Multidimensional random polymers: a renewal approach. In *Random walks, random fields, and disordered systems*, volume 2144 of *Lecture Notes in Math.*, pages 147–210. Springer, Cham, 2015.

[22] D. Ioffe, S. Shlosman, and F. L. Toninelli. Interaction versus entropic repulsion for low temperature Ising polymers. *J. Stat. Phys.*, 158(5):1007–1050, 2015.

[23] D. Ioffe and Y. Velenik. The statistical mechanics of stretched polymers. *Braz. J. Probab. Stat.*, 24(2):279–299, 2010.

[24] D. Ioffe and Y. Velenik. Crossing random walks and stretched polymers at weak disorder. *Ann. Probab.*, 40(2):714–742, 2012.

[25] D. Ioffe and Y. Velenik. Self-attractive random walks: the case of critical drifts. *Comm. Math. Phys.*, 313(1):209–235, 2012.

[26] D. Ioffe and Y. Velenik. Stretched polymers in random environment. In *Probability in complex physical systems*, volume 11 of *Springer Proc. Math.*, pages 339–369. Springer, Heidelberg, 2012.

[27] D. Ioffe and Y. Velenik. An almost sure CLT for stretched polymers. *Electron. J. Probab.*, 18:No. 97, 20, 2013.

[28] M. V. Menshikov. Coincidence of critical points in percolation problems. *Dokl. Akad. Nauk SSSR*, 288(6):1308–1311, 1986.

[29] L. S. Ornstein and F. Zernike. Accidental deviations of density and opalescence at the critical point of a single substance. *Proc. Akad. Sci.*, 17:793–806, 1914.

[30] S. Ott and Y. Velenik. Asymptotics of even-even correlations in the Ising model. *Preprint, arXiv:1804.08323.*

[31] S. Ott and Y. Velenik. Potts models with a defect line. *Commun. Math. Phys., to appear; arXiv:1706.09130.*

[32] P. J. Paes-Leme. Ornstein-Zernike and analyticity properties for classical lattice spin systems. *Ann. Physics*, 115(2):367–387, 1978.

[33] F. Zernike. The clustering-tendency of the molecules in the critical state and the extinction of light caused thereby. *Koninklijke Nederlandse Akademie van Wetenschappen Proceedings Series B Physical Sciences*, 18:1520–1527, 1916.