Maximal clusters in non-critical percolation and related models

Remco van der Hofstad* Frank Redig *

March 23, 2005

Abstract: We investigate the maximal non-critical cluster in a big box in various percolation-type models. We investigate its typical size, and the fluctuations around this typical size. The limit law of these fluctuations is related to maxima of *independent* random variables with law described by a single cluster.

Key-words: Maximal clusters, exponential law, Gumbel distribution, FKG inequality, second moment estimate.

1 Introduction and main results

In [5], Bazant studies the distribution of maximal subcritical clusters, both numerically and via a non-rigorous renormalization group argument. He finds that the cardinality of maximal clusters behaves like the maximum of independent geometrically distributed random variables, i.e., a "Gumbel-like" distribution. In [5], the role of the FKG inequality, which implies that clusters "repel each other", is emphasized in a subadditivity argument.

In this paper, we rigourously prove these claims for a broad class of non-critical percolation type models. In the FKG context, we obtain a discrete Gumbel law for maximal subcritical clusters. In a more general context, we can deal with dependent percolation models dominated by subcritical Bernoulli percolation. In the supercritical case we also obtain Gumbel laws under some extra assumption, which is satisfied e.g. for site percolation in $d = 2, 3$.

The key ingredient of the proof of the Gumbel law is the exponential law for the occurrence time of rare patterns. This idea is used by Wyner in [23] in the context of matching two random sequences. If a cluster bigger than u_n appears in a box $[-n, n]^d \cap \mathbb{Z}^d$ of volume $(2n + 1)^d$, then evidently the occurrence time $\mathbf{t}_{u_n}$ of such a cluster is less than $(2n + 1)^d$. Therefore, if $\mathbf{t}_{u_n}$ has approximately an exponential distribution, then the probability of having a cluster larger than u_n is approximately $1 - e^{-(2n+1)^d \mathbb{P}(\mathcal{C}_{u_n})}$, where $\mathcal{C}_{u_n}$ denotes the event that the cluster of the origin has cardinality at least u_n. If one can find a scale $u_n = u_n(x)$ such that $\mathbb{P}(\mathcal{C}_{u_n(x)}) \simeq e^{-x}/(2n + 1)^d$, then one obtains the Gumbel law. Assuming an exponential decay of the cluster cardinality, as expected for subcritical percolation, one obtains $u_n(x) = u_n + x$, where $u_n = c \log n(1 + o(1))$. For finite supercritical clusters, under the assumption that the cluster size has Weibull-tails, i.e., decays as a stretch exponential with exponent $\delta < 1$, we have $u_n(x) = (c \log n + c' \log \log n + x)^{1/\delta}$.

*Department of Mathematics and Computer Science, Eindhoven University of Technology and EURANDOM, P.O. Box 513, 5600 MB Eindhoven, The Netherlands. E-mail: rhofstad@win.tue.nl, f.h.j.redig@tue.nl

1.1 The model

We consider site percolation and related models on the lattice $\mathbb{Z}^d$. A configuration of occupied and vacant sites is an element $\omega \in \Omega = \{0,1\}^{\mathbb{Z}^d}$. A site x with $\omega(x) = 1$ is called occupied, and a site with $\omega(x) = 0$ is called vacant.

The configuration ω will be distributed according to a translation invariant probability measure $\mathbb{P}$ on the Borel-σ-field of Ω. Examples of $\mathbb{P}$ include the Bernoulli product measure $\mathbb{P}_p$ with $\mathbb{P}_p(\omega(x) = 1) = p$, but we also consider dependent random fields, such as the Ising model, below.

A set $A \subseteq \mathbb{Z}^d$ is *connected* if for any $x, y \in A$ there is a nearest-neighbor path γ joining x and y. The cluster $\mathcal{C}(x) = \mathcal{C}(x, \omega)$ of an occupied site x is the largest connected subset of occupied sites to which x belongs. By convention, $\mathcal{C}(x) = \varnothing$ if $\omega(x) = 0$. We also need the cluster $\mathcal{C}_{\mathrm{le}}(x)$ defined as follows

$$\mathcal{C}_{\mathrm{le}}(x) = \begin{cases} \mathcal{C}(x) & \text{if } x \text{ is the left endpoint of } \mathcal{C}(x), \\ \varnothing & \text{otherwise.} \end{cases} \tag{1.1}$$

Here, by the left-endpoint of a finite set $A \subseteq \mathbb{Z}^d$, we mean the minimum of A in the lexicographic order. By definition $\mathcal{C}_{\mathrm{le}}(x) \cap \mathcal{C}_{\mathrm{le}}(y) = \varnothing$ if $x \neq y$. In this paper, we work with site percolation. In the percolation community, it is more usual to consider bond percolation (see e.g. [19]). However, site percolation is more general than bond percolation, as shown e.g. in [19, Section 1.6]. We use results from [19] proved for bond percolation, but in general these results also hold for site percolation (as noted in [19, Section 12.1]).

Percolation has a *phase transition*, i.e., for $d \geq 2$, there exists a critical value $p_c \in (0,1)$ such that there exists an infinite cluster a.s. for $p > p_c$, while no such cluster exists a.s. for $p < p_c$. The goal of this paper is to investigate maximal clusters in a finite box for $p \neq p_c$.

1.2 Main results for site percolation

In this section, we describe our results in the simplest case, namely for site percolation, where all vertices are independently occupied with probability p and vacant with probability $1 - p$. In Section 3, we formulate our results in a more general context of possibly dependent percolation, and under weaker a priori conditions on the cluster tail behavior.

We study the maximal cluster inside a big box. To be able to state our result, we need some further notation. Let $B_n = [-n, n]^d \cap \mathbb{Z}^d$ be the cube of width $2n + 1$. We let

$$\omega_{B_n}(x) = \begin{cases} \omega(x) & \text{if } x \in B_n, \\ 0 & \text{otherwise,} \end{cases} \tag{1.2}$$

and

$$\mathcal{M}_n = \mathcal{M}_n(\omega) = \max_x |\mathcal{C}_{\mathrm{le}}(x, \omega_{B_n})|. \tag{1.3}$$

The random variable $\mathcal{M}_n$ is the maximal cluster inside B_n, with zero boundary conditions, i.e., where we do not consider connections outside B_n. The goal of this paper is to obtain an extreme value theorem such as

$$\mathbb{P}\left(\mathcal{M}_n \leq u_n + x\right) = e^{-a_n e^{-x}} + o(1) \tag{1.4}$$

for some $u_n \uparrow \infty$, and where a_n is a bounded sequence uniformly bounded from below. In words, this means that the distribution of the maximal cluster is "Gumbel-like", i.e., looks like the maximum of independent geometric random variables. We will prove (1.4) in the case of subcritical percolation and related models. For supercritical percolation, the results are different, in the sense that $u_n + x$ in (1.4) should be replaced with another sequence. The presence of the bounded sequence a_n in (1.4) is typical for the law of the maximum of independent geometric random variables, where we do not have an exact limiting extreme value distribution cannot (see e.g. [16, Corollary 2.4.1]). We now explain the basic idea in this paper, which applies both to sub- and supercritical percolation, in more detail.

The idea developed in this paper is that for any non-critical p, the law of $\mathcal{M}_n$ is asymptotically equal to the law of the maximum of $(2n+1)^d$ *independent* copies of a random variable X with law

$$\mathbb{P}(X = n) = \frac{1}{n}\mathbb{P}(|\mathcal{C}(0)| = n), \tag{1.5}$$

for $n \geq 1$, and

$$\mathbb{P}(X = 0) = 1 - \mathbb{E}(|\mathcal{C}(0)|^{-1}). \tag{1.6}$$

The law of X in (1.5–1.6) turns out to be equal to the law of the random variable $|\mathcal{C}_{\rm le}(0)|$ (see Lemma 4.1 below). Therefore, the law of $\mathcal{M}_n$ is equal to the law of $\max_x |\mathcal{C}(x, \omega_{B_n})|$, and thus the philosophy of the paper is to show that the clusters are only weakly dependent. We further use properties of the law of $\mathcal{C}(0)$ to derive the asymptotics of $\mathcal{M}_n$ in more detail.

We note that the cluster size distribution plays an essential part throughout the proof. We now state the results on this cluster size distribution which we need, in order to specialize the results. Since this law is crucially different for $p < p_c$ and $p > p_c$, we distinguish these two cases.

For $p < p_c$, it is shown in [19, Theorem (6.78)] that

$$\zeta(p, d) = \lim_{n\to\infty} -\frac{1}{n} \log \mathbb{P}_p(|\mathcal{C}(0)| \geq n) \tag{1.7}$$

exists, and that $\zeta = \zeta(p, d) > 0$ for all $p < p_c$. Moreover, there exists $C = C(p)$ such that

$$\mathbb{P}_p(|\mathcal{C}(0)| = n) \leq Cne^{-\zeta n}. \tag{1.8}$$

The above results imply that the cluster size distribution has *exponential tails*, similarly to geometric random variables. For i.i.d. geometric random variables, a law as in (1.4) holds, and we now state similar results for the maximal subcritical percolation cluster.

We sometimes work under the assumption that a somewhat stronger version of (1.7) holds, namely that

$$\lim_{n\to\infty} \frac{\mathbb{P}_p(|\mathcal{C}(0)| \geq n+1)}{\mathbb{P}_p(|\mathcal{C}(0)| \geq n)} = e^{-\zeta}. \tag{1.9}$$

Assumption (1.9) is stronger than (1.7), but weaker than the widely believed tail-behavior, namely that there exist $\theta = \theta(d) \in \mathbb{R}$ and $A = A(p, d)$ such that

$$\mathbb{P}_p(|\mathcal{C}(0)| \geq n) = An^\theta e^{-\zeta n}[1 + o(1)]. \tag{1.10}$$

Our main result for $p < p_c$ is the following theorem:

Theorem 1.1. *Fix $p < p_c$ and assume that (1.9) holds. Then there exists a sequence $u_n \in \mathbb{N}$, with $u_n \to \infty$, a real number $a > 0$ and a bounded sequence $a_n \in [a, 1]$, such that for all $x \in \mathbb{N}$*

$$\mathbb{P}(\mathcal{M}_n \leq u_n + x) = e^{-a_n e^{-x\zeta}} + o(1). \tag{1.11}$$

The more general version of this theorem is stated in Section 3, Theorem 3.6. In this theorem, upper and lower bounds are proved for $\mathbb{P}(\mathcal{M}_n \leq u_n + x)$ without assuming (1.9). Theorem 1.1, and also Theorem 3.6 below, shows that $\mathcal{M}_n$ is bounded above and below by Gumbel laws, and shows in particular that the sequence $\mathcal{M}_n - u_n$ is tight. Our proof will reveal that Theorem 1.1 can be extended to yield weak convergence along certain exponentially growing sequences.

The statement of our main result in terms of the sequence a_n is necessary, and, for instance, is also present when dealing with the maximum of n i.i.d. geometric random variables. The problem we are facing is that the fluctuations are uniformly bounded, while we also know that $\mathcal{M}_n$ is an integer, while we should think of u_n as being something like $u_n = \lfloor \frac{\log n}{\zeta} \rfloor$ (see e.g., Proposition 4.8 below). Thus, we can think of a_n as describing how far $\frac{\log n}{\zeta}$ is from an integer.

We now go to supercritical results, for which the cluster size distribution has quite different tails. Since $p_c(1) = 1$, we may assume that we are in dimension $d > 1$. When $p > p_c$, then it is known that the limit

$$\eta(p, d) = \lim_{n\to\infty} -\frac{1}{n^{\frac{d-1}{d}}} \log \mathbb{P}_p(n \leq |\mathcal{C}(0)| < \infty) \tag{1.12}$$

exists. The limit in (1.12) is related to the large deviations of large finite supercritical clusters, and can be written explicitly as a variational problem over possible cluster shapes. This variational problem involves the surface tension, and is maximized by the so-called Wulff shape. The result in $d = 2$ is in [3, 9], for $d = 3$, it is in [10], and for $d \geq 4$, it is in [11].

We again formulate a different version of (1.12), namely that for every $x \geq 0$, we have

$$\lim_{n\to\infty} \mathbb{P}_p\Big(|\mathcal{C}(0)| \geq n + xn^{1/d} \Big| n \leq |\mathcal{C}(0)| < \infty\Big) = e^{-x\eta\frac{d-1}{d}}, \tag{1.13}$$

and change the definition of $\mathcal{M}_n$ slightly to

$$\mathcal{M}_n = \mathcal{M}_n(\omega) = \max_{x: |\mathcal{C}(x)| < \infty} |\mathcal{C}_{\text{le}}(x, \omega_{B_n})|, \tag{1.14}$$

i.e., we take the largest *finite* cluster. Of course, for $p < p_c$, (1.3) and (1.14) coincide.

Then we can prove the following scaling property:

Theorem 1.2. *Fix $p > p_c$ and assume that (1.12) hold. Then there exists a sequence $u_n \to \infty$ such that for all $x \in \mathbb{R}$*

$$\lim_{n\to\infty} \mathbb{P}(\mathcal{M}_n \leq u_n + xu_n^{1/d}) = e^{-e^{-x\eta\frac{d-1}{d}}}. \tag{1.15}$$

Theorem 1.2 is equivalent to the statement that $\frac{\mathcal{M}_n - u_n}{u_n^{1/d}}$ converges in distribution to a Gumbel random variable. The same statement is true when we consider the maximum of n i.i.d. random variables Y_i for which the tail is given by $\mathbb{P}(Y \geq n) = e^{-\eta n^{\frac{d-1}{d}}(1+o(1))}$.

Theorems 1.1 and 1.2 study fluctuations of $\mathcal{M}_n$ around their asymptotic mean under the Assumptions (1.9) and (1.13). The main difference between Theorems 1.1 and 1.2 is that Theorem 1.2 implies weak convergence of the rescaled $\mathcal{M}_n$ since the fluctuations grow with n, whereas in Theorem 1.1 this weak convergence does not hold due to the fact that the fluctuations are of order 1, so that the discrete nature of cluster sizes persists.

In Section 3 below, we will formulate a more general result that corresponds to Theorem 1.2 and holds *without* assumptions (1.12) and (1.13), but takes a form which is less elegant. It is not so hard to see that one can choose

$$u_n = O(\log n) \tag{1.16}$$

for $p < p_c$, while

$$u_n = O((\log n)^{\frac{d}{d-1}}) \tag{1.17}$$

when $p > p_c$. From Theorems 1.1 and 1.2 it immediately follows that $\mathcal{M}_n$ divided by $\log n$ for $p < p_c$, respectively, $(\log n)^{\frac{d}{d-1}}$ for $p > p_c$, converges in probability to a constant. In the next theorems, we will investigate the typical size of $\mathcal{M}_n$ in more detail and prove convergence almost surely without the assumptions of (1.9) and (1.12), respectively.

Theorem 1.3. *For $p < p_c$,*

$$\frac{\mathcal{M}_n}{\log n} \to d\zeta(p,d) \qquad a.s. \tag{1.18}$$

Theorem 1.4. *For $p > p_c$,*

$$\frac{\mathcal{M}_n}{(\log n)^{\frac{d}{d-1}}} \to d^{\frac{d-1}{d}}\eta(p,d) \qquad a.s. \tag{1.19}$$

These results appear to be "folklore", but we did not find an appropriate reference where they are stated and proved.

We close this section with a few observations concerning the role of the boundary conditions. In (1.3), we have taken the maximal cluster under the zero boundary condition, so that we can write $\mathcal{M}_n = \mathcal{M}_n^{(\text{zb})}$. Alternatively, we could defined $\mathcal{M}_n$ under free boundary conditions, i.e.,

$$\mathcal{M}_n^{(\text{fb})} = \max_{x\in B_n:|\mathcal{C}(x)|<\infty} |\mathcal{C}(x,\omega)|, \tag{1.20}$$

or under periodic boundary conditions, i.e.,

$$\mathcal{M}_n^{(\text{pb})} = \max_{x\in B_n} |\mathcal{C}(x,\omega'_{B_n})|, \tag{1.21}$$

where ω'_{B_n} is the site percolation configuration on the torus with vertex set B_n. We will finally show that this makes no difference whatsoever:

Theorem 1.5. *For $p < p_c$,*

$$\mathbb{P}_p(\mathcal{M}_n^{(\text{zb})} \neq \mathcal{M}_n^{(\text{fb})}) = o(1), \qquad \mathbb{P}_p(\mathcal{M}_n^{(\text{zb})} \neq \mathcal{M}_n^{(\text{pb})}) = o(1). \tag{1.22}$$

For $p > p_c$,

$$\mathbb{P}_p(\mathcal{M}_n^{(\text{zb})} \neq \mathcal{M}_n^{(\text{fb})}) = o(1). \tag{1.23}$$

Theorem 1.5 immediately shows that all results proved for $\mathcal{M}_n^{(\text{zb})}$ immediately also apply to $\mathcal{M}_n^{(\text{fb})}$ and $\mathcal{M}_n^{(\text{pb})}$ for $p < p_c$ and to $\mathcal{M}_n^{(\text{fb})}$ for $p > p_c$, i.e., that the boundary condition is irrelevant. For $p > p_c$, $\mathcal{M}_n^{(\text{pb})}$ is more difficult to work with since it is harder to 'exclude' the infinite cluster on the torus without looking outside the torus.

1.3 Discussion of the results

1.3.1 Runs and one-dimensional site percolation

In the case where $d = 1$, it easily follows that for any $p < p_c = 1$,

$$\mathbb{P}(|\mathcal{C}_{\text{le}}(0)| \geq n) = p^n. \tag{1.24}$$

In this simple case, the largest cluster is equal to the longest run of ones in n independent tosses. This is a classical problem, and the leading order asymptotics $\mathcal{M}_n = \log n / \log p(1 + o(1))$ is the celebrated Erdös-Rényi law [15]. Our results studies fluctuations around the Erdös-Rényi law. This problem has attracted considerable attention due to its relation to matching problems arising in sequence alignment (see e.g., [21] and the references therein).

1.3.2 Results for general subcritical FKG models and related Gumbel laws

Our results for subcritical clusters hold more generally than just for independent site percolation. The main technical ingredient in the proof are the FKG-inequality and bounds on the tails of the cluster size distribution. In Section 3 below, we will state a general result, that can be proved for site percolation and applied in the context of the following examples.

1. The two-dimensional Ising model at $\beta < \beta_c$.
2. The Ising model in general dimension, at high temperature and/or high enough magnetic field (see [18]).
3. Gibbs measures where the potential has a sufficiently small Dobrushin norm and a sufficiently high magnetic field.

See [18] for an introduction of the Ising model and Section 3 for more details.

We expect that related results hold for other maximal values of cluster characteristics. Examples are the maximal diameter of a supercritical finite cluster, or the maximal occupied line (i.e., a sequence of bonds) with any orientation for $p < 1$. We also expect that our results for maximal finite supercritical clusters continue to hold in the context of the Ising model in dimensions $d = 2, 3$ for $\beta > \beta_c$, where the Wulff crystal has been identified (see e.g. [6, 12]), and hence the exact behavior of the cluster tail is known.

1.3.3 Maximal clusters for critical percolation

Our results are only valid for non-critical percolation. In critical percolation, the behavior of the largest cluster in a box should be entirely different. Firstly, the scaling of the largest cluster in a box should be *polynomial* in the volume of the box, rather then *polylogarithmic* as in Corollaries 1.3 and 1.4. Secondly, when properly rescaled, the size of the largest cluster should converge to a proper random variable, rather than to a constant as in Corollaries 1.3 and 1.4. Thirdly, we expect that in some cases, the size of the largest cluster depends on the boundary conditions, which is not true off the critical point (see Theorem 1.5).

There have been results in the direction of the above claims. In [8], the largest critical cluster in a box was investigated and, under certain scaling assumptions, it was proved that the largest cluster with zero boundary conditions scales like $n^{\frac{d\delta}{\delta+1}}$, where δ is the critical exponent related to the critical cluster distribution

$$\frac{1}{\delta} = \lim_{n\to\infty} -\frac{\log \mathbb{P}_{p_c}(|\mathcal{C}(0)| \geq n)}{\log n}. \tag{1.25}$$

Of course, it is not obvious that this limit exists. The scaling assumptions are not expected to be true above the critical dimension $d_c = 6$.

1.3.4 Relation to random graphs

There is a wealth of related work for random graphs, which are finite graphs where edges are removed independently. This research topic was started by a seminal paper of Erdös and Rényi [14], which created the field of random graphs. Erdös and Rényi investigate what is called *the* random graph, i.e., the complete graph where edges are kept independently with fixed probability p and removed otherwise. See the books [4, 7, 20] and the references therein.

When dealing with random graphs, it is natural to investigate the largest connected component or cluster when the size of the graph tends to infinity. Therefore, results such as the ones presented in Section 1.2 have appeared in this field. In particular, detailed estimates of large subcritical and supercritical cluster have been obtained. Of course, for *finite* graphs, it is already non-trivial to define what a critical value is. Above the critical value, the largest cluster has a size of order of the size of the graph, while below the critical value, the largest cluster is logarithmic in the size of the graph.

In random graph theory, often there is a *discrete duality principle*, which means that when we remove the largest supercritical cluster, then the size and distribution of the remaining clusters is very much alike the size and distribution of subcritical clusters. See e.g. [4, Section 10.4] for an explanation of this principle for branching processes as well as for the random graph. We note that this principle is *false* for site percolation on $\mathbb{Z}^d$, as Theorems 1.3 and 1.4 show. This distinction arises from the fact that the classical random graph has no *geometry*, whereas the geometry is essential in the description of large finite supercritical clusters and appears prominently in the Wulff shape.

1.3.5 Organization

Our paper is organized as follows. In Section 2, we give heuristics for our results. In Section 3, we state our general results for FKG models under certain conditions. Section 4 is devoted to the proofs of the main results.

2 Extremes and rare events: heuristics

We are interested in the cardinality of maximal clusters inside a big box. Recall that $B_n = [-n, n]^d \cap \mathbb{Z}^d$. For $n \in \mathbb{N}$, define the σ-field $\mathcal{F}_n = \mathcal{F}_{B_n}$. A *pattern* A_n is a configuration with support on B_n, i.e., it is an element of $\{0,1\}^{B_n}$. We identify a pattern with its cylinder, i.e., we also denote A_n to be the set of those ω such that $\omega_{B_n} = A_n$. For a pattern A_n, we define its occurrence time to be

$$\mathbf{t}_{A_n}(\omega) = \min\Big\{|B_k| : \exists x \in \ B_k \text{ such that } \ B_n + x \subseteq B_k \text{ and } \theta_x \omega_{B_n} = A_n\Big\}, \tag{2.1}$$

where $\theta_x\omega$ denotes the configuration ω shifted over x, so that $(\theta_x\omega)(y) = \omega(x+y)$. In words, this is the volume of the minimal cube B_k which "contains" the pattern A_n. One expects that $\mathbf{t}_{A_n}$ is of the order $\mathbb{P}(A_n)^{-1}$. For a measurable event E_n, we write $E_n \in \mathcal{F}_n$ where, by convention, we always choose the minimal n such that this is the case. For $E_n \in \mathcal{F}_n$, there exists a unique set of patterns $\mathcal{A}(E_n)$ such that

$$E_n = \bigcup_{A_n \in \mathcal{A}(E_n)} A_n$$

The occurrence time of E_n is then defined as

$$\mathbf{t}_{E_n}(\omega) = \min_{A_n \in \mathcal{A}(E_n)} \mathbf{t}_{A_n} \tag{2.2}$$

In words, E_n is a set of patterns, and the occurrence time of E_n is the volume of the first cube B_k in which some pattern of E_n can be found. A sequence of $\mathcal{F}_n$-measurable events E_n is called a *sequence of rare events* if $\mathbb{P}(E_n) \to 0$ as $n \to \infty$. For sequences of rare events, one typically expects so-called *exponential laws*, i.e., limit theorems of the type

$$\mathbb{P}\left(\mathbf{t}_{E_n} \geq \frac{t}{\mathbb{P}(E_n)}\right) = e^{-\lambda_{E_n} t} + o(1). \tag{2.3}$$

Equation (2.3) has been proved for "high temperature Gibbsian random fields" and the parameter λ_{E_n} is bounded away from zero and infinity. In the case of patterns, the parameter depends on the self-repetitive structure of the pattern. For so-called good (meaning that there can be no fast returns) patterns, we even have that $\lambda_{E_n} = 1$. In [2] the exponential law for patterns is generalized to measurable events $E_n \in \mathcal{F}_n$, provided a second moment condition is satisfied. This second moment condition ensures that $1/\mathbb{P}(E_n)$ is the right time scale for the occurrence time, i.e., the parameter λ_{E_n} is bounded away from zero and infinity (see Theorem 3.5 below for the precise formulation).

The relation between maxima and rare events is intuitively obvious: if a cluster with cardinality bigger than m appears in a cube B_n, then the occurrence time for the appearance of a cluster bigger than m is not larger than $|B_n|$. More precisely, throughout the paper, we will work with the event

$$E_n = \{n \leq |\mathcal{C}_{\text{le}}(0)| < \infty\} \tag{2.4}$$

and define the random variable τ_{E_m} with values in $\{(2n+1)^d : n \in \mathbb{N}\}$ by

$$\{\tau_{E_m} \leq (2n+1)^d\} = \{\exists x \in B_n : \theta_x \omega_{B_n} \in E_m\} \tag{2.5}$$

The random variable τ_{E_m} is not exactly equal to the occurrence time $\mathbf{t}_{E_m}$, but we will see that asymptotically τ_{E_m} and $\mathbf{t}_{E_m}$ have the same distribution (see Lemma 4.5 below).

The advantage of working with τ_{E_m} lies in the equality

$$\{\mathcal{M}_n \geq m\} = \{\tau_{E_m} \leq (2n+1)^d\} \tag{2.6}$$

If we assume that the exponential law holds for the occurrence time, then, for all $u_n(x) \in \mathbb{N}$ to be determined later on,

$$\mathbb{P}\left(\mathcal{M}_n \geq u_n(x)\right) \approx 1 - e^{-\lambda_{E_n} \mathbb{P}(E_{u_n(x)})(2n+1)^d}. \tag{2.7}$$

Therefore, if we can choose $u_n(x)$ such that

$$\mathbb{P}(E_{u_n(x)}) \approx \frac{a_n e^{-x}}{(2n+1)^d}, \tag{2.8}$$

then we obtain (1.4). This is the guiding idea of this paper, and the proof of a result of the type (1.4) thus relies on the following three ingredients:

1. Verification of the validity of the exponential law for the events E_n. For this, we rely on the techniques developed in [2], which requires natural mixing conditions and a second moment estimate, see (cf. (3.9)).

2. Proof of the existence of the sequence $u_n(x)$ such that (2.8) holds.

3. Proof that $\lambda_{E_{u_n(x)}} = 1 + o(1)$.

3 General results

In this section, we introduce the conditions needed and state the precise form of (1.4). We start by defining the main conditions in Section 3.1, we state the exponential law proved in [2] in Section 3.2, and in Section 3.3, we state our main results valid under the formulated conditions.

3.1 The conditions

We need three main conditions, a non-uniformly exponentially φ-mixing condition, a finite energy condition, and a condition ensuring that clusters are subcritical or supercritical.

We first introduce the so-called "high mixing" condition which is adapted to the case of Gibbsian random fields. For $m > 0$ define

$$\varphi(m) = \sup \frac{1}{|A_1|} |\ \mathbb{P}\left(E_{A_1}|E_{A_2}\right) - \mathbb{P}\left(E_{A_1}\right)|, \tag{3.1}$$

where the supremum is taken over all finite subsets A_1, A_2 of $\mathbb{Z}^d$, with $d(A_1, A_2) \geq m$ and over all $E_{A_i} \in \mathcal{F}_{A_i}$ with $\mathbb{P}(E_{A_2}) > 0$.

Note that this $\varphi(m)$ differs from the usual φ-mixing function since we divide by the size of the dependence set of the event E_{A_1}. This is natural in the context of Gibbsian random fields, where the classical φ-mixing mostly fails (except for the simplest i.i.d. case and ad-hoc examples of independent copies of one-dimensional Gibbs measures).

We are now ready to formulate the non-uniformly exponentially φ-mixing (NUEM) condition:

Definition 3.1 (NUEM). *A random field is* **non-uniformly exponentially φ-mixing** *(NUEM) if there exist constants $C, c > 0$ such that*

$$\varphi(m) \leq C\exp(-cm) \quad \textit{for all} \quad m > 0. \tag{3.2}$$

Examples of random field satisfying the NUEM condition are Gibbs measures with exponentially decaying potential in the Dobrushin uniqueness regime, or local transformations of such measures. Of course, for site percolation, where we have independence, we have $\varphi = 0$.

We next define the finite energy property:

Definition 3.2 (Finite energy property). *A probability measure $\mathbb{P}$ has the* **finite energy property** *if there exists $\delta \in (0,1)$ such that*

$$\delta \leq \inf_{\omega\in\Omega} \mathbb{P}(\omega_x = 1|\omega_{\mathbb{Z}^d\setminus\{x\}}) \leq \sup_{\omega\in\Omega} \mathbb{P}(\omega_x = 1|\omega_{\mathbb{Z}^d\setminus\{x\}}) \leq 1-\delta. \tag{3.3}$$

Gibbs measures have the finite energy property (in particular, it holds of course for independent site percolation, for which (3.3) holds with $\delta = 1 - \delta = p$), but in general it suffices that there exists a bounded version of $\log \mathbb{P}(\sigma_0 = 1|\sigma_{\{0\}^c})$. A direct consequence of (3.3) is the existence of $C, C' > 0$ such that for any $\sigma \in \Omega$, $V \subseteq \mathbb{Z}^d$,

$$e^{-C|V|} \leq \mathbb{P}(\omega_V = \sigma_V) \leq e^{-C'|V|}. \tag{3.4}$$

Finally, we define what it means for a measure to have subcritical or supercritical clusters:

Definition 3.3 (Sub- and supercritical clusters). (i) *The probability measure $\mathbb{P}$ is said to have* **subcritical clusters** *if $\mathbb{P}(|\mathcal{C}_{\mathrm{le}}(0)| < \infty) = 1$ and if there exists $\zeta, \xi \in (0,\infty)$ such that*

$$e^{-\zeta} \leq \liminf_{n\to\infty} \frac{\mathbb{P}(|\mathcal{C}_{\mathrm{le}}(0)| \geq n+1)}{\mathbb{P}(|\mathcal{C}_{\mathrm{le}}(0)| \geq n)} \leq \limsup_{n\to\infty} \frac{\mathbb{P}(|\mathcal{C}_{\mathrm{le}}(0)| \geq n+1)}{\mathbb{P}(|\mathcal{C}_{\mathrm{le}}(0)| \geq n)} \leq e^{-\xi}. \tag{3.5}$$

(ii) *The probability measure* $\mathbb{P}$ *is said to have* **supercritical clusters** *if* $\mathbb{P}(|\mathcal{C}_{\rm le}(0)| = \infty) > 0$ *and if*

$$\lim_{n\to\infty} \frac{\mathbb{P}(n+1 \le |\mathcal{C}_{\rm le}(0)| < \infty)}{\mathbb{P}(n \le |\mathcal{C}_{\rm le}(0)| < \infty)} = 1. \tag{3.6}$$

Note that the condition in (3.5) implies that $\mathbb{P}(|\mathcal{C}_{\rm le}(0)| \ge n)$ is bounded above and below by sequences that converge to 0 exponentially. The main restriction is that $\xi < \infty$ and that $\zeta > 0$, so that we require something essentially stronger than the trivial bounds that $0 \le \frac{\mathbb{P}(|\mathcal{C}_{\rm le}(0)| \ge n+1)}{\mathbb{P}(|\mathcal{C}_{\rm le}(0)| \ge n)} \le 1$.

3.2 The exponential law

In order to have the exponential law, we need that the events E_n are somewhat localized. More precisely, the non-occurrence of the event in a big cube can be decomposed as an intersection of non-occurrence of the event in a union of small sub-cubes separated by corridors. Then mixing can be used to factorize the probabilities of non-occurrence in the sub-cubes, provided the corridors are sufficiently large. Optimization of this philosophy is the content of the Iteration Lemma in [2]. In our case, the events are not strictly localized but they can be replaced by local events, without affecting limit laws. This is made precise in the following definition:

Definition 3.4 (Localizability). (i) *Let* E_n *be a sequence of events such that* $\mathbb{P}(E_n) \to 0$. *The events* E_n *are called* **local** *with respect to* l_n *when* $E_n \in \mathcal{F}_{k_n}$, *where* k_n *is such that* $k_n(2n+1)^d l_n^{-\theta} \to 0$ *for all* $\theta > 0$.

(ii) *The events* E_n *satisfying* $\mathbb{P}(E_n) \to 0$ *are called* **localizable** *for* $m_n \uparrow \infty$ *if there exist events* E'_n *which are local with respect to* $l_n = 1/\mathbb{P}(E_n)$ *such that*

$$\lim_{n\to\infty} |\mathbb{P}(\mathbf{t}_{E'_n} \le m_n) - \mathbb{P}(\mathbf{t}_{E_n} \le m_n)| = 0$$

E'_n *is then called a local version of* E_n.

We use the following theorem which can be derived from [2], as we explain below.

Theorem 3.5 (Exponential law). *Suppose* $\mathbb{P}$ *has finite energy and satisfies the NUEM condition. Suppose further that* E_n *are localizable measurable events such that for some* $\delta, \gamma > 0$ *and all* $n \in \mathbb{N}$, $\mathbb{P}(E_n) \le e^{-\gamma n^\delta}$. *Assume furthermore that for all* $\alpha > 1$,

$$\limsup_{n\to\infty} \sum_{0<|x|\le n^\alpha} \frac{\mathbb{P}(E_n \cap \theta_x E_n)}{\mathbb{P}(E_n)} < \infty. \tag{3.7}$$

Then there exists $\Lambda_1, \Lambda_2, c, \rho \in (0,\infty)$, *such that for all* $n \in \mathbb{N}$ *there exists* $\lambda_{E_n} \in [\Lambda_1, \Lambda_2]$ *such that*

$$\left| \mathbb{P}\left(\mathbf{t}_{E_n} > \frac{t}{\lambda_{E_n}\mathbb{P}(E_n)} \right) - e^{-t} \right| \le \mathbb{P}(E_n)^\rho e^{-ct} \tag{3.8}$$

For the “local version” E'_n, the theorem follows from [2, Theorem 2.6 and Remark 2.8]. The extension to E_n is straightforward from Definition 3.4 and is formulated in detail in [2, Remark 4.13]. Note that there is some notational difference between the present paper and [2], since, in [2], the occurrence time $\mathbf{t}_{E_n}$ is the *width* of the first cube where E_n occurs, whereas in our setting, it is the *volume*.

Condition (3.7) is needed to apply Lemma 4.6 in [2], see also [1]. It ensures the existence of the lower bound Λ_1 on the parameter λ_{E_n} (which is obtained via a second moment estimate for the number of occurrences). It guarantees further that the parameter is bounded away from zero which means that in a cube of volume $\mathbb{P}(E_n)^{-1}$, the event E_n happens with a probability bounded away from zero (uniformly in n). This means that $\mathbb{P}(E_n)^{-1}$ is the right scale, i.e., a cube with this volume is such that the event E_n happens with probability bounded away from zero or one.

The parameter λ_{E_n} measures the "self-repetitive" nature of the event E_n, i.e., whether the event appears typically isolated or in clusters. See also [1] for one-dimensional examples of $\lambda_{E_n} \neq 1+o(1)$ and conditions ensuring $\lambda_{E_n} = 1+o(1)$. For the events E_n of our paper, we show that $\lambda_{E_n} = 1+o(1)$ in Section 4 below.

3.3 Main results

In our context, Condition (3.7) is satisfied as soon as for all $\alpha > 1$, we have

$$\limsup_{n\to\infty} \sum_{0<|x|<n^\alpha} \frac{\mathbb{P}\left(\{n \le |\mathcal{C}_{\rm le}(x)| < \infty\} \cap \{n \le |\mathcal{C}_{\rm le}(0)| < \infty\}\right)}{\mathbb{P}(n \le |\mathcal{C}_{\rm le}(0)| < \infty)} < \infty. \tag{3.9}$$

The value of α which we will need later is related to the localization of the event $|\mathcal{C}_{\rm le}(0)| > n$ to the event $n < |\mathcal{C}_{\rm le}(0)| < n^\alpha$ (see the proof in Section 4 for more details).

Now we can state our main result for the subcritical case:

Theorem 3.6 (Subcritical Gumbel law). *Suppose $\mathbb{P}$ has finite energy, is NUEM, has subcritical clusters and satisfies (3.9). Then there exists a sequence u_n with $u_n \in \mathbb{N}$ and $u_n \uparrow \infty$, and a bounded sequence $a_n \in [e^{-\zeta}, 1]$, such that for $x \ge 0$*

$$e^{-a_n e^{-x\zeta}} \le \mathbb{P}(\mathcal{M}_n \le u_n + x) \le e^{-a_n e^{-x\xi}}. \tag{3.10}$$

When $x < 0$, the upper and lower bound are reversed.

Moreover, if $\xi = \zeta$, then there exists a constant $\rho > 0$ such that

$$|\mathbb{P}\left(\mathcal{M}_n \le u_n + x\right) - e^{-a_n e^{-\zeta x}}| \le \frac{1}{n^\rho}. \tag{3.11}$$

We now turn to examples where we can apply Theorem 3.6. The following proposition yields a class of non-trivial examples:

Proposition 3.7. *If $\mathbb{P}$ is a subcritical Markov measure satisfying the FKG inequality, then (3.9) is satisfied.*

This gives the following applications:

1. Subcritical site percolation $\mathbb{P} = \mathbb{P}_p$ where $\mathbb{P}_p$ is the Bernoulli measure with $\mathbb{P}_p(\omega_0) = p$ and $p < p_c$.

2. In $d = 2$: Ising model at $\beta < \beta_c$. In general dimension, Ising model at high temperature and/or high enough magnetic field (see [18]).

In particular, Theorem 3.6 implies Theorem 1.1, because if (1.9) holds, then $\xi = \zeta$.

In very general context we have (3.9) in high enough magnetic field. The idea is that as soon as for any V, and any $\omega \in \Omega$, the conditional probabilities $\mathbb{P}_V(\cdot|\omega_{V^c})$ can be dominated by a Bernoulli measure with subcritical clusters, then of course, for any $x \neq 0$,

$$\mathbb{P}\big(|\mathcal{C}_{\text{le}}(x)| \geq n \big| |\mathcal{C}_{\text{le}}(0)| \geq n\big) \leq \mathbb{P}_p(|\mathcal{C}_{\text{le}}(0)| \geq n), \tag{3.12}$$

and hence (3.9) is satisfied.

We now formulate another class of examples. We say that $\mathbb{P}$ is *dominated by a Bernoulli measure in the sense of Holley,* when for all $\omega \in \Omega$

$$\mathbb{P}(\omega_0 = 1|\omega_{\{0\}^c}) < p. \tag{3.13}$$

This condition implies that $\mathbb{P}$ is stochastically dominated by the Bernoulli measure $\mathbb{P}_p$. For measures that are dominated by a subcritical Bernoulli measure, our results also apply:

Proposition 3.8. *Let p_c denote the critical value for Bernoulli site percolation. If (3.13) is satisfied for some $p < p_c$, then (3.9) holds true.*

This proposition can be applied to Gibbs measures such that the potential has a Dobrushin norm which is small enough (to guarantee mixing condition), with magnetic field high enough such that (3.13) holds, see [18] for more details.

Our last theorem applies for independent supercritical site percolation. Recall (1.14). Then we have the following result for supercritical site percolation:

Theorem 3.9 (Supercritical Gumbel law for independent site percolation). *Let $p > p_c$ and let $\mathbb{P}$ denote the measure of supercritical site percolation with percolation probability p. Then there exists a constant $a > 0$, a sequence $a_n \in (a, 1]$ and a sequence $u_n(x)$ with $u_n(x) \in \mathbb{N}$ and $u_n(x) \uparrow \infty$ for all $x \in \mathbb{R}$ as $n \uparrow \infty$, such that for all $x \in \mathbb{R}$*

$$\mathbb{P}_p(\mathcal{M}_n \leq u_n(x)) = e^{-a_n e^{-x}} + o(1). \tag{3.14}$$

If $\mathbb{P}$ has supercritical clusters, then $a_n = 1 + o(1)$.

Theorem 3.9 implies Theorem 1.2. Indeed, (3.6) holds if (1.13) is satisfied, and in that case, it is easy to verify that $u_n(x) = u_n + x\eta^{(1-d)/d} u_n^{1/d}(1 + o(1))$, where $u_n = u_n(0) \uparrow \infty$.

4 Proofs

In this section, we prove the main results stated in Sections 1 and 3.

4.1 Preparations

In this section, we state some general results for non-critical clusters. In Proposition 4.2 and Lemma 4.3, we investigate the cluster size distribution in more detail.

The following Lemma 4.1, identifies the law of $|\mathcal{C}_{\text{le}}(0)|$ for non-critical clusters in terms of the law of $|\mathcal{C}(0)|$. It follows immediately from (6.104) in [19] in the context of percolation, but the proof applies to the more general setting as well.

Lemma 4.1 (The law of $|\mathcal{C}_{\rm le}(0)|$). *For all $n \geq 1$,*

$$\mathbb{P}(|\mathcal{C}(0)| = n) = n\mathbb{P}(|\mathcal{C}_{\rm le}(0)| = n). \tag{4.1}$$

Before we formulate our next proposition, we remark that the cluster $\mathcal{C}_{\rm le}(0)$ is finite with probability one, since it has 0 as its left endpoint. Therefore, we have $\mathbb{P}(n \leq |\mathcal{C}_{\rm le}(0)| < \infty) = \mathbb{P}(|\mathcal{C}_{\rm le}(0)| \geq n)$ in the supercritical case, and we can drop the restriction that the cluster is finite in the notation. Naturally, we also drop this restriction in the subcritical case.

Proposition 4.2 (Lower bound on the cluster tail). *If $\mathbb{P}$ has finite energy, then there exists $\zeta > 0$ such that*

$$\liminf_{n\to\infty} \frac{\mathbb{P}(|\mathcal{C}_{\rm le}(0)| \geq n+1)}{\mathbb{P}(|\mathcal{C}_{\rm le}(0)| \geq n)} \geq e^{-\zeta}. \tag{4.2}$$

Proof. We start with the subcritical case. We abbreviate $X = |\mathcal{C}_{\rm le}(0)|$. Define $A_n = \{X \geq n\}$. We have to estimate the ratio

$$\frac{\mathbb{P}(A_{n+1})}{\mathbb{P}(A_{n+1}) + \mathbb{P}(A_n \setminus A_{n+1})}$$

from below, i.e., the ratio

$$\frac{\mathbb{P}(A_n \setminus A_{n+1})}{\mathbb{P}(A_{n+1})}$$

from above. To do so, we will modify a configuration $\omega \in A_n \setminus A_{n+1}$ into a configuration $T(\omega) \in A_{n+1}$ such that only a finite number of occupancies are flipped and such that T is at most K to one where $K = 2^{2d-1}$. The map T is described as follows. Look for the "right endpoint" y_ω of $\mathcal{C}_{\rm le}(0,\omega)$, i.e., the largest point of $\mathcal{C}_{\rm le}(0,\omega)$ in the lexicographic order, and define $x_\omega = y_\omega + (1, 0, \ldots, 0)$. We then define $T(\omega)$ by putting $T(\omega)(x_\omega) = 1$, $T(\omega)(z) = 0$ for all neighbors of y_ω which are not in $\mathcal{C}_{\rm le}(0,\omega)$, and $T(\omega)(x) = \omega(x)$ for all other sites x. Clearly, the left endpoint remains unchanged in this construction, and there correspond at most 2^{2d-1} configurations $\omega \in A_n \setminus A_{n+1}$ to each configuration in the image $T(A_n \setminus A_{n+1})$. Since T only modifies ω in a finite number of sites, by the finite energy property of $\mathbb{P}$, the ratio

$$0 < C_2 < \frac{\mathbb{P}(\omega|\omega^{A^c})}{\mathbb{P}(T\omega|\omega^{A^c})} < C_1,$$

where A denotes y_ω and all its neighbours, is uniformly bounded away from zero and infinity. Therefore,

$$\frac{\mathbb{P}(A_n \setminus A_{n+1})}{\mathbb{P}(A_{n+1})} \leq C_1 2^{2d-1}$$

The proof of the supercritical case is the same, because the transformation T can, from a configuration $\omega \in A_n \setminus A_{n+1}$, never create a configuration where the origin is in an infinite cluster. $\square$

Lemma 4.3 (Existence of v_n). *Suppose that $\mathbb{P}$ has finite energy. There exists a sequence $v_n \in \mathbb{N}$ with $v_n \uparrow \infty$ and a sequence b_n satisfying $e^{-\zeta} < b_n \leq 1$ such that*

$$\mathbb{P}(|\mathcal{C}_{\rm le}(0)| > v_n) = \frac{b_n}{n}. \tag{4.3}$$

Where ζ is defined in (4.2). If $\mathbb{P}$ has supercritical clusters, then $b_n = 1 + o(1)$.

Proof. We again abbreviate $X = |\mathcal{C}_{\rm le}(0)|$, and recall that $X < \infty$ both above and below criticality. We define

$$\begin{aligned} v_n^+ &= \inf\{x \in \mathbb{N} : \mathbb{P}(X \geq x) \leq \frac{1}{n}\}, \\ v_n^- &= \sup\{x \in \mathbb{N} : \mathbb{P}(X \geq x) \geq \frac{1}{n}\}. \end{aligned} \tag{4.4}$$

Then $v_n^+ = v_n^- + 1$ or $v_n^+ = v_n^-$. Put $v_n = v_n^+$. By definition

$$\mathbb{P}(X \geq v_n) \leq \frac{1}{n},$$

so that $b_n \leq 1$.

Moreover,

$$\begin{aligned} n\mathbb{P}(X \geq v_n) &= n\mathbb{P}(X \geq v_n^-)\frac{\mathbb{P}(X \geq v_n^+)}{\mathbb{P}(X \geq v_n^-)} \\ &\geq \frac{\mathbb{P}(X \geq v_n^- + 1)}{\mathbb{P}(X \geq v_n^-)}. \end{aligned} \tag{4.5}$$

We note that $v_n \uparrow \infty$ when $n \to \infty$. Therefore, by Proposition 4.2,

$$\liminf_{n\to\infty} n\mathbb{P}(X \geq v_n) \geq \liminf_{n\to\infty} \frac{\mathbb{P}(X \geq n+1)}{\mathbb{P}(X \geq n)} = e^{-\zeta}. \tag{4.6}$$

Thus, we obtain $\liminf_{n\to\infty} b_n \geq e^{-\zeta}$. On the other hand, by definition, if $\mathbb{P}$ has supercritical clusters, then

$$1 \geq \lim_{n\to\infty} n\mathbb{P}(X \geq v_n) \geq \lim_{n\to\infty} \frac{\mathbb{P}(X \geq n+1)}{\mathbb{P}(X \geq n)} = 1. \tag{4.7}$$

Thus, we obtain $\lim_{n\to\infty} b_n = 1$. □

We next verify that the events

$$E_n = \{|\mathcal{C}_{\rm le}(0)| \geq n\} \tag{4.8}$$

are localizable. This is the content of the next lemma.

Lemma 4.4 (Localizability of E_n). *Suppose that $\mathbb{P}$ has subcritical clusters, or that $\mathbb{P}$ is the measure of supercritical site percolation. Then the events $E_n = \{|\mathcal{C}_{\rm le}(0)| \geq n\}$ are localizable for $m_n = t/\mathbb{P}(E_n)$ and any $t \geq 0$, and their local versions can be chosen as*

$$E_n' = \{n \leq |\mathcal{C}_{\rm le}(0)| < n^\theta\} \tag{4.9}$$

for some $\theta \in (1, \infty)$ with $k_n = n^\theta$ in Definition 3.4.

Proof. Clearly, $E_n' \subseteq E_n$ and $E_n \backslash E_n' \subseteq E_{n^\theta}$. Consequently, $\mathbf{t}_{E_n'} \geq \mathbf{t}_{E_n}$ and

$$\mathbb{P}(\mathbf{t}_{E_n} \leq m_n) - \mathbb{P}(\mathbf{t}_{E_n'} \leq m_n) = \mathbb{P}(\mathbf{t}_{E_n} \leq m_n, \mathbf{t}_{E_n'} > m_n) \leq \mathbb{P}(\mathbf{t}_{E_{n^\theta}} \leq m_n).$$

Clearly, we have that, with $m_n = t/\mathbb{P}(E_n)$,

$$\mathbb{P}(\mathbf{t}_{E_{n^\theta}} \leq m_n) \leq (2m_n + 1)^d \mathbb{P}(E_{n^\theta}) \leq (3t)^d \frac{\mathbb{P}(E_{n^\theta})}{\mathbb{P}(E_n)^d}. \tag{4.10}$$

The right-hand side is $o(1)$ for any $t \geq 0$ fixed by (3.5) in the subcritical case and (1.13) in the supercritical case. The fact that E_n' is local w.r.t. $1/\mathbb{P}(E_n)$ with $k_n = n^\theta$ is obvious. □

We finish this section with a lemma showing the asymptotic equivalence of τ_{E_n} introduced in (2.5) and the occurrence time $\mathbf{t}_{E_n}$.

More precisely, we have the following lemma:

Lemma 4.5 (Occurence times). *Let $m = m_n \uparrow \infty$ be such that $m_n n^{\epsilon-1}$ converges to zero as $n \to \infty$ for some $\epsilon \in (0,1)$, and such that $\mathbb{P}(E_{m_n}) \leq n^{-d+\epsilon}$, where, as before,*

$$E_{m_n} = \{|\mathcal{C}_{\rm le}(0)| \geq m_n\}.$$

Then

$$\mathbb{P}\Big(\tau_{E_{m_n}} \leq (2n+1)^d\Big) = \mathbb{P}\Big(\mathbf{t}_{E_{m_n}} \leq (2n+1)^d\Big) + o(1). \tag{4.11}$$

where $\tau_{E_{m_n}}$ $\mathbf{t}_{E_{m_n}}$, respectively, are defined in (2.5) and (2.2), respectively.

Proof. First we remark that

$$\{\mathbf{t}_{E_m} \leq (2n+1)^d\} \subseteq \{\tau_{E_m} \leq (2n+1)^d\}, \tag{4.12}$$

and

$$\{\tau_{E_m} \leq (2n+1)^d\} \setminus \{\mathbf{t}_{E_m} \leq (2n+1)^d\} \subseteq \{\exists x \in B_n : x + B_m \not\subseteq B_n : |\mathcal{C}_{\rm le}(x)| > m\}. \tag{4.13}$$

We estimate

$$\begin{aligned} \mathbb{P}\left(\{\exists x \in B_n : x + B_{m_n} \not\subseteq B_n, |\mathcal{C}_{\rm le}(x)| > m_n\}\right) &\leq \mathbb{P}(|\mathcal{C}_{\rm le}(0)| \geq m_n)|\{x \in B_n : x + B_{m_n} \not\subseteq B_n\}| \\ &\leq \frac{m_n n^{d-1}}{n^{d-\epsilon}} = m_n n^{\epsilon-1}. \end{aligned} \tag{4.14}$$

This converges to zero as $n \to \infty$ by the assumption on m_n. □

4.2 Maximal subcritical clusters

In this section, we prove Theorems 1.1 and 3.6. We study the tails of the cluster size distribution, subject to (3.5). The main result is the following lemma:

Lemma 4.6 (Identification of $u_n(x)$). *Suppose $\mathbb{P}$ has finite energy and has subcritical clusters, then there exists a sequence a_n with $0 < a_n \leq 1$ such that $\liminf_{n\to\infty} a_n \geq e^{-\zeta}$, such that for all $x > 0$ and for all $n \geq 1$,*

$$\frac{a_n}{(2n+1)^d} e^{-\zeta x}(1+o(1)) \leq \mathbb{P}(|\mathcal{C}_{\rm le}(0)| \geq u_n + x) \leq \frac{a_n}{(2n+1)^d} e^{-\xi x}(1+o(1)), \tag{4.15}$$

where $u_n = v_{(2n+1)^d}$ with v_n the sequence appearing in Lemma 4.3. For $x < 0$, the same inequality holds with ζ and ξ interchanged.

Proof. Let $x > 0$. We again abbreviate $X = |\mathcal{C}_{\rm le}(0)|$. We define, using $u_n = v_{(2n+1)^d}$,

$$a_n = b_{(2n+1)^d} = (2n+1)^d \mathbb{P}(|\mathcal{C}_{\rm le}(0)| \geq u_n). \tag{4.16}$$

Then, the bounds on a_n follow from the bounds on b_n in Lemma 4.3. Furthermore,

$$\frac{\mathbb{P}(X \geq u_n + x)}{\mathbb{P}(X \geq u_n)} = \prod_{i=1}^{x} \frac{\mathbb{P}(X \geq u_n + i)}{\mathbb{P}(X \geq u_n + i - 1)}. \tag{4.17}$$

Hence, with the choice of $u_n = v_{(2n+1)^d}$ where v_n is as in Lemma 4.3, and all $x \in \mathbb{N}$ fixed,

$$\liminf_{n\to\infty} \frac{\mathbb{P}(X \ge u_n + x)}{a_n(2n+1)^d} \ge \liminf_{n\to\infty} \prod_{i=1}^{x} \frac{\mathbb{P}(X \ge u_n + i)}{\mathbb{P}(X \ge u_n + i - 1)} = e^{-\zeta x}, \tag{4.18}$$

so that the lower bound in (4.15) follows. For the upper bound, a similar argument gives

$$\limsup_{n\to\infty} \frac{\mathbb{P}(X \ge u_n + x)}{a_n(2n+1)^d} \ge \limsup_{n\to\infty} \prod_{i=1}^{x} \frac{\mathbb{P}(X \ge u_n + i)}{\mathbb{P}(X \ge u_n + i - 1)} = e^{-\xi x}. \tag{4.19}$$

This proves the claim for $x > 0$. The proof for $x < 0$ is similar. □

We now verify Condition (3.7) for subcritical FKG measures:

Proof of Proposition 3.7.
We have to prove that for any $\alpha > 0$,

$$\limsup_{n\to\infty} \sum_{0<|x|\le n^\alpha} \frac{\mathbb{P}(|\mathcal{C}_{\mathrm{le}}(0)| \ge n, |\mathcal{C}_{\mathrm{le}}(x)| \ge n)}{\mathbb{P}(|\mathcal{C}_{\mathrm{le}}(0)| \ge n)} < \infty. \tag{4.20}$$

In fact, we will show that the left-hand side of (4.20) equals 0 when $n \to \infty$.

We denote by $\mathbb{P}^\eta_\Lambda(\omega_\Lambda)$ the conditional probability to find ω inside Λ, given η outside Λ. For a Markov random field, the dependence on η is only through the boundary of Λ, i.e.,

$$\mathbb{P}^\eta_\Lambda(\omega_\Lambda) = \mathbb{P}^{\eta_{\partial\Lambda}}_\Lambda(\omega_\Lambda), \tag{4.21}$$

where $\partial\Lambda$ denotes the exterior boundary of Λ, i.e., the set of those sites not belonging to Λ which have at least one neighbor inside Λ. Thus, we can think of η as describing the boundary condition. By the FKG-property, we have that if $\eta \le \zeta$ for $\eta, \zeta \in \{0,1\}^{\mathbb{Z}^d}$, then

$$\mathbb{P}^\eta_\Lambda \le \mathbb{P}^\zeta_\Lambda \tag{4.22}$$

Moreover, by definition of the clusters $\mathcal{C}_{\mathrm{le}}(x)$, we have that $\mathcal{C}_{\mathrm{le}}(0) \cap \mathcal{C}_{\mathrm{le}}(x) = \varnothing$ for $x \ne 0$. Therefore, we can write, for $x \ne 0$,

$$\begin{aligned}
&\mathbb{P}(|\mathcal{C}_{\mathrm{le}}(0)| \ge n, |\mathcal{C}_{\mathrm{le}}(x)| \ge n) \\
&\quad= \sum_{A:|A|\ge n, le(A)=0} \mathbb{P}(\mathcal{C}_{\mathrm{le}}(0) = A, |\mathcal{C}_{\mathrm{le}}(x)| \ge n, \mathcal{C}_{\mathrm{le}}(x) \cap A = \varnothing) \\
&\quad= \sum_{A:|A|\ge n, le(A)=0} \mathbb{P}\big(|\mathcal{C}_{\mathrm{le}}(x)| \ge n, \mathcal{C}_{\mathrm{le}}(x) \cap A = \varnothing \big| \omega_A = 1, \omega_{\partial A} = 0\big)\mathbb{P}(\mathcal{C}_{\mathrm{le}}(0) = A) \\
&\quad= \sum_{A:|A|\ge n, le(A)=0} \mathbb{P}(\mathcal{C}_{\mathrm{le}}(0) = A)\mathbb{P}^{0_{\partial A}}_{\mathbb{Z}^d\setminus\bar{A}}(|\mathcal{C}_{\mathrm{le}}(x)| \ge n),
\end{aligned} \tag{4.23}$$

where in the last step we have used the Markov property, with the notation $\bar{A} = A \cup \partial A$ and ∂A is the exterior boundary of A, i.e., the neighbours of elements in A that are not in A. Using (4.22), we thus arrive at

$$\mathbb{P}^{0_{\partial A}}_{\mathbb{Z}^d\setminus\bar{A}}(n \le |\mathcal{C}_{\mathrm{le}}(x)| \le n^\theta) \le \mathbb{P}_{\mathbb{Z}^d\setminus\bar{A}}(|\mathcal{C}(x)| \ge n) \le \mathbb{P}(|\mathcal{C}(x)| \ge n). \tag{4.24}$$

Equation (4.24) combined with (4.23) leads to the correlation inequality

$$\mathbb{P}(|\mathcal{C}_{\mathrm{le}}(0)| \geq n, |\mathcal{C}_{\mathrm{le}}(x)| \geq n) \leq \mathbb{P}(|\mathcal{C}_{\mathrm{le}}(0)| \geq n)\mathbb{P}(|\mathcal{C}(0)| \geq n). \tag{4.25}$$

Therefore,

$$\sum_{0<|x|\leq n^{\alpha}} \frac{\mathbb{P}(|\mathcal{C}_{\mathrm{le}}(0)| \geq n, |\mathcal{C}_{\mathrm{le}}(x)| \geq n)}{\mathbb{P}(|\mathcal{C}_{\mathrm{le}}(0)| \geq n)} \leq (2n^{\alpha}+1)^d \mathbb{P}(|\mathcal{C}(0)| \geq n) \to 0, \tag{4.26}$$

because the decay of the probability $\mathbb{P}(|\mathcal{C}(0)| \geq n)$ is faster than $\frac{1}{n^{\beta}}$ for any $\beta > 0$, since $\mathbb{P}$ has subcritical clusters. □

Proposition 4.7 (The subcritical intensity is one). *For u_n as in Lemma 4.6 and for every x finite, there exists a $\beta > 0$ such that*

$$1 - \mathbb{P}(E_{u_n+x})^{\beta} \leq \lambda_{E_{u_n+x}} \leq 1. \tag{4.27}$$

Proof. We will first identify $\lambda_{E_{u_n+x}}$. We use [2, (2.6)], which states that

$$\lambda_E = -\frac{\log \mathbb{P}\big(\mathbf{t}_E > f_E\big)}{f_E \mathbb{P}(E)}, \tag{4.28}$$

where, for some $\gamma \in (0,1)$,

$$f_E = \lceil \mathbb{P}(E)^{-\gamma} \rceil. \tag{4.29}$$

We will show that $\mathbb{P}\big(\mathbf{t}_E \leq f_E\big)$ is quite small (as proved in the sequel), so that we can approximate

$$-\log \mathbb{P}\big(\mathbf{t}_E > f_E\big) = \mathbb{P}\big(\mathbf{t}_E \leq f_E\big) + O\Big(\mathbb{P}\big(\mathbf{t}_E \leq f_E\big)^2\Big). \tag{4.30}$$

Therefore,

$$\lambda_E = \frac{\mathbb{P}\big(\mathbf{t}_E \leq f_E\big)}{f_E \mathbb{P}(E)} + o(1). \tag{4.31}$$

We proceed by computing $\mathbb{P}\big(\mathbf{t}_E \leq f_E\big)$. To do so, we write

$$\mathbb{P}\big(\mathbf{t}_{E_{u_n+x}} \leq f_{E_{u_n+x}}\big) = \mathbb{P}\Big(\bigcup_{y \in B_{m_{n,x}}} \{|\mathcal{C}_{\mathrm{le}}(y)| \geq u_n + x\}\Big), \tag{4.32}$$

where we abbreviate $m_{n,x} = f_{E_{u_n+x}}^{1/d}$. By Boole's inequality,

$$\mathbb{P}\big(\mathbf{t}_{E_{u_n+x}} \leq f_{E_{u_n+x}}\big) \leq \sum_{y \in B_{m_{n,x}}} \mathbb{P}(|\mathcal{C}_{\mathrm{le}}(y)| \geq u_n + x) = f_{E_{u_n+x}} \mathbb{P}(E_{u_n+x}). \tag{4.33}$$

Thus,

$$\lambda_{E_{u_n+x}} \leq 1. \tag{4.34}$$

For the lower bound, use

$$\begin{aligned} \mathbb{P}\big(\mathbf{t}_{E_{u_n+x}} \leq f_{E_{u_n+x}}\big) \quad \geq \quad & \sum_{y \in B_{m_{n,x}}} \mathbb{P}(|\mathcal{C}_{\mathrm{le}}(y)| \geq u_n + x) \\ & - \sum_{y,z \in B_{m_{n,x}}: y \neq z} \mathbb{P}(|\mathcal{C}_{\mathrm{le}}(y)| \geq u_n + x, |\mathcal{C}_{\mathrm{le}}(z)| \geq u_n + x). \end{aligned} \tag{4.35}$$

The first term is identical to the first term in the upper bound, and we need to bound the second term only. For this, we use (4.25), and thus obtain

$$\mathbb{P}\big(\mathbf{t}_{E_{u_n+x}} \le f_{E_{u_n+x}}\big) \ge f_{E_{u_n+x}}\mathbb{P}(E_{u_n+x}) - f^2_{E_{u_n+x}}\mathbb{P}(E_{u_n+x})\mathbb{P}(|\mathcal{C}(0)| \ge u_n + x). \tag{4.36}$$

Thus,

$$\lambda_{E_{u_n+x}} \ge 1 - f_{E_{u_n+x}}\mathbb{P}(|\mathcal{C}(0)| \ge u_n + x) \ge 1 - \mathbb{P}(E_{u_n+x})^\beta \tag{4.37}$$

for some $\beta > 0$. □

We finally identify the sequence u_n under the hypothesis of a "classical" subcritical cluster tail behavior in Proposition 4.8, and under the hypothesis of a "classical" supercritical cluster tail behavior in Proposition 4.12.

Proposition 4.8 (Identification $u_n(x)$ for classical subcritical tails).
Suppose that there exists $\alpha \in \mathbb{R}, \zeta > 0$ and $0 < C < \infty$, such that

$$\mathbb{P}(|\mathcal{C}_{\mathrm{le}}(0)| \ge n) = Cn^\alpha e^{-\zeta n}[1 + o(1)]. \tag{4.38}$$

Then u_n of Lemma 4.6 can be chosen as:

$$u_n = \Big\lfloor \frac{\log n}{\zeta} + \frac{\alpha \log\log n}{\zeta} \Big\rfloor. \tag{4.39}$$

Proof. This is a simple computation, using $u_n = v_{(2n+1)^d}$ where v_n is introduced in Lemma 4.3. □

Proof of Theorem 3.6 and Theorem 1.1. We first finish the proof of Theorem 3.6. We use the equality

$$\{\mathcal{M}_n \ge m\} = \{\tau_{E_m} \le (2n+1)^d\}. \tag{4.40}$$

Then we use Lemma 4.5 to obtain that as long as $\mathbb{P}(E_{m_n}) \le n^{-d+\epsilon}$, we have

$$\mathbb{P}(\mathcal{M}_n \ge m_n) = \mathbb{P}(\mathbf{t}_{E_{m_n}} \le (2n+1)^d) + o(1). \tag{4.41}$$

We wish to apply Theorem 3.5, and will first check that the conditions are fulfilled. We note from Lemma 4.4 that the events E_n are localizable for $t/\mathbb{P}(E_n)$ with local versions E_n'. Furthermore, from Proposition 3.7, it follows that Condition (3.7) is fulfilled for E_n. Therefore, we may apply Theorem 3.5.

We choose $u_n(x) = u_n + x$ as in Lemma 4.6, and the event E_{u_n+x} as before in (4.8). Note that for this $u_n(x)$, we indeed have that for every x fixed, and using the bound $a_n \le 1$ in Lemma 4.6, to obtain

$$\mathbb{P}(E_{u_n+x}) = \frac{e^{-x}}{(2n+1)^d}a_n \le n^{-d+\epsilon}, \tag{4.42}$$

so that we can use (4.41).

Assume that $x \ge 0$. For $x < 0$ some inequalities reverse sign. Then we apply Theorem 3.5 to obtain:

$$\mathbb{P}(\mathcal{M}_n \ge u_n + x) = \mathbb{P}(\mathbf{t}_{E_{u_n+x}} \le (2n+1)^d) + o(1) = 1 - \exp\big(-\lambda_{E_{u_n+x}}(2n+1)^d\mathbb{P}(E_{u_n+x})\big) + o(1). \tag{4.43}$$

We need to investigate the exponent. By Lemma 4.6, we have that

$$\frac{a_n}{(2n+1)^d}e^{-\zeta x} \leq \mathbb{P}(E_{u_n+x}) \leq \frac{a_n}{(2n+1)^d}e^{-\zeta x}, \tag{4.44}$$

and this inequality is reversed for $x < 0$. By Proposition 4.7, we have that

$$\lambda_{E_{u_n+x}} = 1 + o(1). \tag{4.45}$$

Therefore, for any $x \in \mathbb{N}$,

$$1 - \exp(-a_n e^{-\xi x}) + o(1) \leq \mathbb{P}(\mathcal{M}_n \geq u_n + x) \leq 1 - \exp(-a_n e^{-\xi x}) + o(1). \tag{4.46}$$

This completes the proof of Theorem 3.6. When $\zeta = \xi$, the statement in Theorem 1.1 is a direct consequence of Theorem 3.6, combined with Lemma 4.5.

□

Remark. The examples mentioned in Section 1.3.2 fit into the context of Theorem 3.6. Indeed, for the Ising model, the inequality (3.5) is verified above the critical temperature in $d = 2$ and at high enough temperature in any dimension. The mixing condition (3.2) is verified at high temperature in the Dobrushin uniqueness regime, and in $d = 2$ above the critical temperature, by complete analyticity. For general Gibbs measures with a potential with a finite Dobrushin norm, one can choose the magnetic field high enough such that the Dobrushin uniqueness condition and hence condition (3.2) is satisfied (see e.g. [17]), and such that (3.5) follows from a domination with Bernoulli measures (see [18]).

4.3 Maximal supercritical clusters

In this section we prove Theorems 3.9 and 1.2.

In the following proposition we show that we can still find a sequence $u_n(x)$, but not necessarily of the form $u_n + x$, if we omit the subcriticality condition. This will be useful when we study the supercritical percolation clusters.

Lemma 4.9 (Existence of $u_n(x)$). *Suppose $\mathbb{P}$ has finite energy, is NUEM and furthermore*

$$\mathbb{E}(|\mathcal{C}(0)|I[|\mathcal{C}(0)| < \infty]) < \infty.$$

Then there exists a function $u_n(x)$ with $u_n(x) \in \mathbb{N}$, $u_n(x) \uparrow \infty$ as $n \to \infty$, for all $x \in \mathbb{R}$, such that

$$\mathbb{P}(|\mathcal{C}_{\text{le}}(0)| \geq u_n(x)) = \frac{e^{-x}}{(2n+1)^d}a_n, \tag{4.47}$$

where $a_n \leq 1$ is a sequence for which $\liminf_{n\to\infty} a_n > 0$ independent of x. Furthermore, if $\mathbb{P}$ has supercritical clusters, then $a_n = 1 + o(1)$.

Proof. Since $\mathbb{E}(|\mathcal{C}(0)|I[|\mathcal{C}(0)| < \infty]) < \infty$, we can use Lemma 4.1. As in the proof of Lemma 4.3, we define

$$\begin{aligned} v_n^+(x) &= \inf\{k : \mathbb{P}(X \geq k) \geq \frac{e^{-x}}{n}\}, \\ v_n^-(x) &= \sup\{k : \mathbb{P}(X \geq k) \leq \frac{e^{-x}}{n}\}. \end{aligned} \tag{4.48}$$

We can then choose $u_n(x) = v^+_{(2n+1)^d}(x)$ and

$$a_n = (2n+1)^d \mathbb{P}(|\mathcal{C}(0)| \geq u_n(x)) \leq 1. \tag{4.49}$$

Then we can repeat the proof of Lemma 4.3 to see that $\liminf_{n\to\infty} a_n \geq e^{-\zeta}$. Finally, we can use (3.6) to conclude that $a_n = 1 + o(1)$. □

We continue with the following proposition which will guarantee Condition (3.7) for finite supercritical clusters.

Proposition 4.10 (Supercritical second moment condition). *Suppose that $\mathbb{P}$ is the measure of supercritical site percolation. Then for every $\alpha > 1$,*

$$\limsup_{n\to\infty} \sum_{0<|x|<n^\alpha} \frac{\mathbb{P}(E_n \cap \theta_x E_n)}{\mathbb{P}(E_n)} = 0. \tag{4.50}$$

Proof. Denote $\tilde{E}_n = E_n \cap E^c_{n^\alpha}$. We rewrite, using that E_n is the disjoint union of $\tilde{E}_n$ and E_{n^α},

$$\mathbb{P}(E_n \cap \theta_x E_n) = \mathbb{P}(\tilde{E}_n \cap \theta_x \tilde{E}_n) + \mathbb{P}(E_{n^\alpha} \cap \theta_x E_n) + \mathbb{P}(\tilde{E}_n \cap \theta_x E_{n^\alpha}). \tag{4.51}$$

The last two terms are simple. We bound their contribution to the left-hand side of (4.50) by

$$2(2n^\alpha + 1)^d \frac{\mathbb{P}(E_{n^\alpha})}{\mathbb{P}(E_n)}. \tag{4.52}$$

We further compute

$$\mathbb{P}(\tilde{E}_n \cap \theta_x \tilde{E}_n) \quad = \sum_{\Gamma_1 \in \mathcal{G}_n(0)} \sum_{\Gamma_2 \in \mathcal{G}_n(x)} \mathbb{P}(\mathcal{C}_{\mathrm{le}}(0) = \Gamma_1, \mathcal{C}_{\mathrm{le}}(x) = \Gamma_2), \tag{4.53}$$

where

$$\mathcal{G}_n(x) = \{\Gamma : x = le(\Gamma), n \leq |\Gamma| < n^\alpha\}. \tag{4.54}$$

For $\Gamma \subseteq \mathbb{Z}^d$, we denote $\overline{\Gamma} = \Gamma \cup \partial\Gamma$. When $\overline{\Gamma}_1 \cap \overline{\Gamma}_2 = \varnothing$, then the events $\{\mathcal{C}_{\mathrm{le}}(0) = \Gamma_1\}$ and $\{\mathcal{C}_{\mathrm{le}}(x) = \Gamma_2\}$ are independent. We can bound this contribution to the left-hand side of (4.53) by

$$\sum_{\Gamma_1 \in \mathcal{G}_n(0), \Gamma_2 \in \mathcal{G}_n(x) : \overline{\Gamma}_1 \cap \overline{\Gamma}_2 = \varnothing} \mathbb{P}(\mathcal{C}_{\mathrm{le}}(0) = \Gamma_1)\mathbb{P}(\mathcal{C}_{\mathrm{le}}(x) = \Gamma_2) \leq \mathbb{P}(\tilde{E}_n)^2 \leq \mathbb{P}(E_n)^2, \tag{4.55}$$

so that the contribution to the left-hand side of (4.50) is bounded by

$$(2n^\alpha + 1)^d \mathbb{P}(E_n). \tag{4.56}$$

We are left to deal with Γ_1, Γ_2 for which $\overline{\Gamma}_1 \cap \overline{\Gamma}_2 \neq \varnothing$. By construction, we have that $\Gamma_1 \cap \Gamma_2 = \varnothing$, so that there must be at least one $y \in \overline{\Gamma}_1 \cap \overline{\Gamma}_2$ such that $y \notin \Gamma_1 \cup \Gamma_2$. Clearly, we have that y is vacant. We follow the construction in the proof of Proposition 4.2. When we make y occupied, then we create a cluster of size at least $2n$ with left endpoint $0 \wedge x$, the minimum of 0 and x in the lexicographic order. Moreover, when we also flip the occupied neighbors of y that are not in $\Gamma_1 \cup \Gamma_2$ to be vacant, then we guarantee that $|\mathcal{C}_{\mathrm{le}}(0 \wedge x)| < 2n^\alpha \leq (2n)^\alpha$, and $|\mathcal{C}_{\mathrm{le}}(0 \wedge x)| = 2n$. This map is at most 2^{2d-2} to one. Therefore, similarly to the Proof of Proposition 4.2, the change in probability is at most $C_{p,d} = 2^{2d-2}(\max\{\frac{1-p}{p}, \frac{p}{1-p}\})^{2d+1}$, so that this contribution to the left-hand side of (4.53) is bounded by

$$C_{p,d} \sum_y \sum_{(\Gamma_1,\Gamma_2)\in\mathcal{G}_n(0,x,y)} \mathbb{P}\big(\mathcal{C}_{\rm le}(0\wedge x) = \Gamma_1 \cup \Gamma_2 \cup \{y\}\big), \tag{4.57}$$

where

$$\mathcal{G}_n(0,x,y) = \{(\Gamma_1,\Gamma_2) : \Gamma_1 \in \mathcal{G}_n(0), \Gamma_2 \in \mathcal{G}_n(x), y \in \overline{\Gamma}_1 \cap \overline{\Gamma}_2\}. \tag{4.58}$$

Therefore, this contribution to the left-hand side of (4.53) is bounded above by

$$\begin{aligned} C_{p,d} \sum_y \mathbb{P}\big(y \in \mathcal{C}_{\rm le}(0\wedge x), 2n \le |\mathcal{C}_{\rm le}(0\wedge x)| < (2n)^\alpha\big) \\ \le (2n)^\alpha C_{p,d} \mathbb{P}(\theta_{0\wedge x}\tilde{E}_{2n}) \le (2n^\alpha+1) C_{p,d}\mathbb{P}(E_{2n}), \end{aligned} \tag{4.59}$$

so that the contribution to the left-hand side of (4.50) is bounded by

$$C_{p,d}(2n^\alpha+1)^{d+1}\frac{\mathbb{P}(E_{2n})}{\mathbb{P}(E_n)}. \tag{4.60}$$

In particular, for every $x \neq 0$,

$$\mathbb{P}\left(E_n \cap \theta_x E_n\right) \le C(2n^\alpha+1)^{d+1}[\mathbb{P}(E_{2n}) + \mathbb{P}(E_n)^2]. \tag{4.61}$$

We conclude that

$$\sum_{0<|x|<n^\alpha} \frac{\mathbb{P}\left(E_n \cap \theta_x E_n\right)}{\mathbb{P}\left(E_n\right)} \le 2(2n^\alpha+1)^d \frac{\mathbb{P}(E_{n^\alpha})}{\mathbb{P}(E_n)} + (2n^\alpha+1)^d \mathbb{P}(E_n) + C_{p,d}(2n^\alpha+1)^{d+1}\frac{\mathbb{P}(E_{2n})}{\mathbb{P}(E_n)}. \tag{4.62}$$

By [19, Theorem (8.65)], there exists $\eta = \eta(p,d) > 0$ such that

$$e^{-\gamma n^{\frac{d-1}{d}}} \le \mathbb{P}\left(E_n\right) \le \mathbb{P}\left(n \le |\mathcal{C}(0)| < \infty\right) \le e^{-\eta n^{\frac{d-1}{d}}}. \tag{4.63}$$

From (4.62) and (4.63), we conclude that the first two terms converge to zero, while by (1.13), the second term also converges to zero, for every $\alpha > 1$. Thus, (4.50) follows. □

Proposition 4.11 (The supercritical intensity is one). *Suppose $\mathbb{P}$ is the measure of supercritical site percolation. For $u_n(x)$ as in Lemma 4.9 and for every x bounded, there exists a $\beta > 0$ such that*

$$1 - \mathbb{P}(E_{u_n(x)})^\beta \le \lambda_{E_{u_n(x)}} \le 1. \tag{4.64}$$

Proof. We follow the proof of Proposition 4.7. We will first identify $\lambda_{E_{u_n(x)}}$. Recall (4.28) and (4.31). The upper bound in (4.33) applies verbatim.

For the lower bound, use

$$\begin{aligned} \mathbb{P}\big(\mathbf{t}_{E_{u_n(x)}} \le f_{E_{u_n(x)}}\big) \quad \ge \quad & \sum_{y\in B_{m_{n,x}}} \mathbb{P}(u_n(x) \le |\mathcal{C}_{\rm le}(y)| < \infty) \\ & - \sum_{y,z\in B_{m_{n,x}}: y\neq z} \mathbb{P}(u_n(x) \le |\mathcal{C}_{\rm le}(y)| < \infty, u_n(x) \le |\mathcal{C}_{\rm le}(z)| < \infty), \end{aligned} \tag{4.65}$$

where now $m_{n,x} = f_{E_{u_n(x)}}^{1/d}$. The first term is identical to the first term in the upper bound, and we need to bound the second term only.

Using (4.61) in conjunction with (1.13), we arrive at

$$\sum_{y,z\in B_{m_{n,x}}:y\neq z} \mathbb{P}(u_n(x) \leq |\mathcal{C}_{\rm le}(y)| < \infty, u_n(x) \leq |\mathcal{C}_{\rm le}(z)| < \infty) \tag{4.66}$$

$$\leq f^2_{E_{u_n(x)}} \mathbb{P}(E_{u_n(x)})^\kappa \leq [f_{E_{u_n(x)}} \mathbb{P}(E_{u_n(x)})] \mathbb{P}(E_{u_n(x)})^{\kappa-\gamma-1},$$

so that we arrive at the claim when $\kappa > \gamma + 1$ with $\beta = \kappa - \gamma - 1$.

□

Finally, for supercritical clusters we expect that

$$\mathbb{P}(n \leq |\mathcal{C}_{\rm le}(0)| < \infty) = Cn^\alpha e^{-\eta n^\delta}[1 + o(1)], \tag{4.67}$$

i.e., when the cluster size distribution has Weibull tails (possibly with polynomial corrections), and with $\delta = \frac{d-1}{d}$.

So far, (4.67) has not been proved rigorously, but if we assume such a tail behavior, then we can infer the precise form of the sequence $u_n(x)$ in Lemma 4.9.

Proposition 4.12 (Identification $u_n(x)$ for classical supercritical tails). *If (4.67) is satisfied, then the sequence $u_n(x)$ of Lemma 4.9 can be chosen of the form*

$$u_n(x) = \Big\lfloor \big(\frac{\log n}{\eta} + \frac{\alpha \log\log n}{\eta\delta} + x\big)^{1/\delta} \Big\rfloor. \tag{4.68}$$

Proof. Under the condition (4.67), it is a simple computation to verify that

$$\mathbb{P}(u_n(x) \leq |\mathcal{C}_{\rm le}(0)| < \infty) = \frac{e^{-x}}{(2n+1)^d}(1 + o(1)). \tag{4.69}$$

□

Proof of Theorem 3.9 and Theorem 1.2. We first finish the proof of Theorem 3.9. We follow the line of argument in the proof of Theorem 3.6. We first use (4.40). Then we use Lemma 4.5 to obtain that as long as $\mathbb{P}(E_{m_n}) \leq n^{-d+\epsilon}$, we have (4.41).

We again apply Theorem 3.5, and now check the conditions. We note from Lemma 4.4 that the events E_n are localizable for $m_n = t/\mathbb{P}(E_n)$ with local versions E_n'. Furthermore, from Proposition 4.10, it follows that Condition (3.7) is fulfilled for E_n. Therefore, we may apply Theorem 3.5.

We choose $u_n(x)$ as in Lemma 4.9, and the event $E_{u_n(x)}$ as before. Note that for this $u_n(x)$, we indeed have that

$$\mathbb{P}(E_{u_n(x)}) = \frac{e^{-x}}{(2n+1)^d} a_n \leq n^{-d+\epsilon}, \tag{4.70}$$

so that we can use (4.41).

Assume that $x \geq 0$. For $x < 0$ some inequalities reverse sign. Then we apply Theorem 3.8 to obtain:

$$\mathbb{P}(\mathcal{M}_n \geq u_n(x)) = \mathbb{P}(\mathbf{t}_{E_{u_n(x)}} \leq (2n+1)^d) + o(1) = 1 - \exp\Big(-\lambda_{E_{u_n(x)}}(2n+1)^d \mathbb{P}(E_{u_n(x)})\Big) + o(1). \tag{4.71}$$

We need to investigate the exponent. By Lemma 4.6, we have that

$$\mathbb{P}(E_{u_n(x)}) = \frac{a_n}{(2n+1)^d} e^{-x}. \tag{4.72}$$

By Proposition 4.11, we have that

$$\lambda_{E_{u_n(x)}} = 1 + o(1). \tag{4.73}$$

Therefore, for any x,

$$\mathbb{P}(\mathcal{M}_n \geq u_n(x)) = 1 - \exp(-a_n e^{-x}) + o(1). \tag{4.74}$$

This completes the proof of Theorem 3.9.

If we further assume that $\mathbb{P}$ has supercritical clusters, then by Lemma 4.9 we can take $a_n = 1 + o(1)$. For Theorem 1.2, we note that the further assumption (1.13) implies that $\mathbb{P}$ has supercritical clusters, and that we can choose $u_n(x) = \lceil u_n + x u_n^{1/d} \rceil$, where $u_n = u_n(0)$. Hence, we obtain that

$$\mathbb{P}(E_{u_n(x)}) = \mathbb{P}(E_{u_n}) \frac{\mathbb{P}(E_{u_n + x u_n^{1/d}})}{\mathbb{P}(E_{u_n})} = \mathbb{P}(E_{u_n}) e^{-x\eta \frac{d-1}{d}} [1 + o(1)] = n^{-d} e^{-x\eta \frac{d-1}{d}} [1 + o(1)]. \tag{4.75}$$

The conclusion then follows from (4.74). □

4.4 Proof of Theorems 1.3, 1.4 and 1.5

Proof of Theorems 1.3 and 1.4. We will prove Theorems 1.3 and 1.4 simultaneously. In order to do so, we let $\delta = 1$ for $p < p_c$ and $\delta = \frac{d-1}{d}$ for $p > p_c$. We then assume that

$$-\lim_{n\to\infty} \frac{1}{n^\delta} \log \mathbb{P}(|\mathcal{C}(0)| \geq n) = \xi \tag{4.76}$$

exists. The main ingredient is the following lemma:

Lemma 4.13 (Convergence in probability). *For any $\varepsilon > 0$, there exists $\kappa > 0$ such that as $n \to \infty$,*

$$\mathbb{P}\big(\big|\frac{\mathcal{M}_n}{(\log n)^{1/\delta}} - C\big| > \varepsilon\big) \leq n^{-\kappa}, \tag{4.77}$$

where $C = d\zeta$ for $p < p_c$ and $C = d^{\frac{d-1}{d}} \eta$ for $p > p_c$.

Before proving Lemma 4.13, we will complete the proofs of Theorems 1.3 and 1.4 subject to Lemma 4.13.

Take $n_k = 2^k$. As a consequence of Lemma 4.13, and the fact that for every $\kappa > 0$,

$$n_k^{-\kappa} = 2^{-\kappa k}$$

is summable in k, we obtain that $\frac{\mathcal{M}_{n_k}}{(\log(n_k))^{1/\delta}}$ converges to C a.s. Thus, we have a.s. convergence along the subsequence $(n_k)_{k\geq 0}$. Moreover, we have that a.s. $n \mapsto \mathcal{M}_n$ is non-decreasing. Therefore, for any $n_k < n \leq n_{k+1}$ we can bound

$$\frac{\mathcal{M}_{n_k}}{(\log(n_k))^{1/\delta}} \big(\frac{\log(n_k)}{\log(n_{k+1})}\big)^{1/\delta} \leq \frac{\mathcal{M}_n}{(\log n)^{1/\delta}} \leq \frac{\mathcal{M}_{n_{k+1}}}{(\log(n_{k+1}))^{1/\delta}} \big(\frac{\log(n_{k+1})}{\log(n_k)}\big)^{1/\delta}. \tag{4.78}$$

As $n \to \infty$, also $n_k, n_{k+1} \to \infty$. Thus, $\frac{\mathcal{M}_{n_k}}{(\log(n_k))^{1/\delta}}$ and $\frac{\mathcal{M}_{n_{k+1}}}{(\log(n_{k+1}))^{1/\delta}}$ converge a.s. to C. Furthermore,

$$\lim_{k\to\infty} \frac{\log(n_{k+1})}{\log(n_k)} = \lim_{k\to\infty} \frac{k+1}{k} = 1, \tag{4.79}$$

so that both upper and lower bound in (4.78) converge to C almost surely. This completes the proofs of Theorems 1.3 and 1.4. □

Proof of Lemma 4.13. Let C be the constant such that

$$\mathbb{P}\Big(\frac{|\mathcal{C}(0)|}{(\log n)^{1/\delta}} > C\Big) = n^{-d(1+o(1))}, \tag{4.80}$$

This constant exists by (1.7) in the case $p < p_c$, and by assumption (1.12) (proved in $d = 2, 3$) for $p > p_c$. With this choice of C, for $\varepsilon > 0$, there exists a $\kappa' \in (0, d)$ such that

$$\mathbb{P}\Big(\frac{|\mathcal{C}(0)|}{(\log n)^{1/\delta}} > C + \varepsilon\Big) \leq n^{-d-\kappa'}, \tag{4.81}$$

while

$$\mathbb{P}\Big(\frac{|\mathcal{C}(0)|}{(\log n)^{1/\delta}} < C - \varepsilon\Big) \leq 1 - n^{-d+\kappa'}. \tag{4.82}$$

Fix $\varepsilon > 0$. We will prove

$$\mathbb{P}\Big(\frac{\mathcal{M}_n}{(\log n)^{1/\delta}} > C + \varepsilon\Big) \leq n^{-\kappa}, \tag{4.83}$$

and

$$\mathbb{P}\Big(\frac{\mathcal{M}_n}{(\log n)^{1/\delta}} < C - \varepsilon\Big) \leq n^{-\kappa}. \tag{4.84}$$

To prove (4.83), we use that

$$\begin{aligned}
\mathbb{P}\Big(\frac{\mathcal{M}_n}{(\log n)^{1/\delta}} > C + \varepsilon\Big) &= \mathbb{P}\Big(\bigcup_{x\in B_n} \{|\mathcal{C}(x)| > (C+\varepsilon)(\log n)^{1/\delta}\}\Big) \\
&\leq \sum_{x\in B_n} \mathbb{P}\Big(|\mathcal{C}(0)| > (C+\varepsilon)(\log n)^{1/\delta}\Big) \\
&\leq |B_n| n^{-d-\kappa'} \leq n^{-\kappa},
\end{aligned} \tag{4.85}$$

where we use (4.81).

To prove (4.84), we use that the events $\{|\mathcal{C}(x)| \leq (C+\varepsilon)(\log n)^{1/\delta}\}_{x\in A_n}$ are independent when

$$A_n = (K_n\mathbb{Z})^d \cap B_n. \tag{4.86}$$

and

$$K_n = \lceil (C+\varepsilon)(\log n)^{1/\delta} \rceil \tag{4.87}$$

Therefore,

$$\begin{aligned}
\mathbb{P}\Big(\frac{\mathcal{M}_n}{(\log n)^{1/\delta}} < C - \varepsilon\Big) &= \mathbb{P}\Big(\bigcap_{x\in A_n} \{|\mathcal{C}(x)| \leq (C-\varepsilon)(\log n)^{1/\delta}\}\Big) \\
&= \prod_{x\in A_n} \mathbb{P}\Big(|\mathcal{C}(x)| < (C-\varepsilon)(\log n)^{1/\delta}\Big) \\
&\leq \mathbb{P}\Big(|\mathcal{C}(0)| < (C-\varepsilon)(\log n)^{1/\delta}\Big)^{|A_n|}.
\end{aligned} \tag{4.88}$$

We next use (4.82) and the fact that

$$|A_n| \geq \Big(\frac{n}{\lceil (C+\varepsilon)(\log n)^{1/\delta}\rceil}\Big)^d, \tag{4.89}$$

so arrive at a bound, for every $\kappa \in (0, \kappa')$,

$$\mathbb{P}\Big(\frac{\mathcal{M}_n}{(\log n)^{1/\delta}} < C - \varepsilon\Big) \;\leq\; \big(1 - n^{-d+\kappa'}\big)^{|A_n|} \leq n^{-\kappa}, \tag{4.90}$$

which completes the proof. □

Proof of Theorem 1.5. We again use (4.80) together with the observation that the events $\{\mathcal{M}_n^{(\mathrm{zb})} \neq \mathcal{M}_n^{(\mathrm{fb})}\}$ and $\{\mathcal{M}_n^{(\mathrm{zb})} \neq \mathcal{M}_n^{(\mathrm{pb})}\}$ are contained in the event that there exists a cluster on the boundary (either with free or periodic boundary conditions) such that there exists an $x \in \partial B_n$ such that $|\mathcal{C}(x)| \geq \mathcal{M}_n^{(\mathrm{zb})}$. By Theorems 1.3 and 1.4, we have that $\mathcal{M}_n^{(\mathrm{zb})} \geq (C-\varepsilon)(\log n)^{1/\delta}$ a.s. By (4.80) and when $\varepsilon > 0$ is sufficiently small, this probability is thus bounded above by

$$n^{d-1}\mathbb{P}\Big(|\mathcal{C}(x)| \geq (C-\varepsilon)(\log n)^{1/\delta}\Big) \leq n^{-\kappa}$$

for some $\kappa > 0$. □

Acknowledgements

The work of RvdH was supported in part by Netherlands Organisation for Scientific Research (NWO). The authors thank Harry Kesten for stimulating discussions during the early stages of this work, Olivier Couronné for useful remarks, and Redu-Jürg Messikh for a discussion of the Wulff shape results.

References

[1] M. Abadi. Sharp error terms and necessary conditions for exponential hitting times in mixing processes. Preprint. To appear in *Ann. Probab.*, (2003)

[2] M. Abadi, J.R. Chazottes, F. Redig and E. Verbitskyi. Exponential distribution for the occurrence of rare patterns in Gibbsian random fields. Preprint available at xxx.lanl.gov, to appear in Comm. Math. Phys., (2003).

[3] K.S. Alexander, J.T. Chayes and L. Chayes. The Wulff construction and asymptotics of the finite cluster distribution for two-dimensional Bernoulli percolation. *Commun. Math. Phys.* **131** no.1: 1–10, (1990).

[4] N. Alon and J.H. Spencer. *The Probabilistic Method.* Wiley, New York, 2nd edition, (2000).

[5] Bazant. The largest cluster in subcritical percolation. Phys. Rev. E **62**, 1660– (2000).

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

[7] B. Bollobás. *Random Graphs.* Cambridge University Press, Cambridge, 2nd edition, (2001).

[8] C. Borgs, J.T. Chayes, H. Kesten, and J. Spencer. The birth of the infinite cluster: finite-size scaling in percolation. *Commun. Math. Phys.*, **224**:153–204, (2001).

[9] R. Cerf. Large deviations of the finite cluster shape for two-dimensional percolation in the Hausdorff and L^1 topology. *J. Theor. Probab.* **12** no.4: 1137–1163, (1999).

[10] R. Cerf. *Large deviations for three dimensional supercritical percolation. Astérisque No. 267* (2000).

[11] R. Cerf. *The Wulff crystal in Ising and Percolation models.* Saint Flour lecture notes, Preprint (2004).

[12] R. Cerf and A. Pisztora. Phase coexistence in Ising, Potts and percolation models. *Ann. Inst. H. Poincare Probab. Statist.*, **37**(6):643–724, (2001).

[13] J.R. Chazottes and F. Redig. Occurrence, repetition and matching of patterns in the low temperature Ising model. Preprint available at xxx.lanl.gov (2003).

[14] P. Erdös and A. Rényi. On the evolution of random graphs. *Bull. Inst. Internat. Statist.* **38**, 343–347 (1961).

[15] P. Erdös and A. Rényi. On a new law of large numbers. *J. Analyse Math.* **23** ,103–111 (1970).

[16] J. Galambos. *The Asymptotic Theory of Extreme Order Statistics.* Wiley, New-York (1978).

[17] H.O. Georgii. *Gibbs Measures and Phase Transitions.* Walter de Gruyter, Berlin, (1988).

[18] H.O. Georgii, O. Haggstrom and C. Maes, *The random geometry of equilibrium phases*, In *Phase Transitions and Critical Phenomena*, Vol **18**, Eds. C. Domb and J.L. Lebowitz (Academic Press, London), 1-142 (2001).

[19] G. Grimmett. *Percolation*, Second edition. Springer-Verlag, 1999.

[20] S. Janson, T. Łuczak, and A. Ruciński. *Random Graphs.* John Wiley and Sons, New York, (2000).

[21] S. Karlin and A. Dembo. Limit distributions of maximal segmental score emong Markov-dependent partial sums. *Adv. Appl. Prob.*, **24**: 113-140, (1992).

[22] M.R. Leadbetter, G. Lindgren and H. Rootzén. *Extremes and Related Properties of Random Sequences and Processes.* Springer-Verlag (1982).

[23] A. Wyner, More on recurrence and waiting times. Ann. Appl. Probab. **9** (1999), 780–796.