Microsoft Technical Report
MSR-TR-98-26

UNIFORM BOUNDEDNESS OF CRITICAL CROSSING PROBABILITIES IMPLIES HYPERSCALING

C. Borgs [1,†], J. T. Chayes [1,‡], H. Kesten [2] and J. Spencer [3]

[1] *Microsoft Research*
One Microsoft Way, Redmond, WA 98052

[2] *Department of Mathematics, Cornell University*
Ithaca, NY 14853

[3] *Courant Institute of Mathematical Sciences*
New York University, 251 Mercer Street, New York, NY 10012

July 1998

Abstract. We consider bond percolation on the d-dimensional hypercubic lattice. Assuming the existence of a single critical exponent, the exponent ρ describing the decay rate of point-to-plane crossings at the critical point, we prove that hyperscaling holds whenever critical rectangle crossing probabilities are uniformly bounded away from 1.

1. Introduction

In this paper, we examine the relationship between boundedness of the critical rectangle crossing probabilities and hyperscaling in percolation. Rectangle crossing probabilites are fundamental quantities in percolation: differences in the scaling of these probabilites can be used to distinguish the subcritical, supercritical and critical regimes of the model. Moreover, as we will discuss, differences in the limiting behavior of rectangle crossing probabilities in the critical regime appear to distinguish systems which obey hyperscaling from those which do not.

Consider bond percolation on the hypercubic lattice $\mathbb{Z}^d$ with bond occupation density p. Let $R_{L,M}(p)$ denote the probability, at density p, that an $L \times M \times \cdots \times M$ rectangular box is crossed by a path of occupied bonds in the 1-direction. The box

[†] On leave from Institut für Theoretische Physik, Universität Leipzig, Germany
[‡] On leave from Department of Mathematics, University of California at Los Angles, USA

Typeset by $\mathcal{AMS}$-TeX

crossing probabilities we usually consider are of the form $R_{L,kL}(p)$, where k is some positive integer, independent of L (we often take $k = 3$).

It has been known for some time that if $k > 1$,

$$\begin{aligned} R_{L,kL}(p) &\le O(L^{2d-2}e^{-L/\xi}) \quad \text{if} \quad p < p_c, \\ 1 - R_{L,kL}(p) &\le e^{-c\sigma L^{d-1}} \qquad\qquad \text{if} \quad p > p_c \end{aligned} \tag{1.1}$$

where $\xi = \xi(p)$ is the correlation length of the percolation model, i.e., the rate of decay of the standard connectivity function, $\sigma = \sigma(p)$ is the the surface tension of the model, i.e., the rate of decay of a Wilson loop connectivity function [ACCFR83], and $c > 0$ is a constant that depends on k and the dimension d. While the first bound in (1.1) is an immediate consequence of the usual subadditivity argument for the connectivity function, the second is less trivial, see [CC87] for a proof (related results can be found in [Pis96]). For $p < p_c$, the methods of [CC84] can be used to complement the upper bound in (1.1) by a lower bound that allows to prove that for all $k > 0$ fixed, we have

$$R_{L,kL}(p) \sim e^{-L/\xi} \qquad \text{if} \quad p < p_c, \tag{1.2}$$

that is, $-(1/L)\log R_{L,kL}(p) \to 1/\xi$ as $L \to \infty$. In turn, the results of [CC87] give the lower bound

$$1 - R_{L,M}(p) \ge e^{-\sigma_M(p)L^{d-1}} \qquad \text{if} \quad p > p_c, \tag{1.3}$$

with a decay constant $\sigma_M(p)$ that converges to the surface tension $\sigma(p)$ as $M \to \infty$.

The fact that $R_{L,kL}(p) \sim e^{-L/\xi}$ if $p < p_c$ motivated the definition of a finite-size scaling correlation length in [CCF85], a concept which was further developed in [Kes87]. It is also known [CC86] that, when $p > p_c$, there are typically $O(L^{d-1})$ disjoint, but not necessarily disconnected crossings of the rectangle. The change in behavior (1.1) also motivated a numerical definition of a percolation threshold, $p_c(L)$, see e.g. [AS92], Section 4.1, or [BW95], [BW97] and references therein.

Here we are interested in the *critical* behavior of the crossing probability, and a closely related quantity, the expected number of crossings. For simplicity of notation, let $R_L(p) = R_{L,3L}(p)$ be the probability of an easy-way crossing of an $L \times 3L \times \cdots \times 3L$ box in $\mathbb{Z}^d$ (that is, a crossing in the direction in which the box is shortest), and let $N_L(p)$ be the expectation of the maximal number of disjoint easy-way crossings in the box.

It is known (see Lemma 4.1 (v)) that

$$R_L(p_c) \ge c_1 > 0 \qquad \text{in all } d \ge 2 \tag{1.4}$$

uniformly in L, while

$$R_L(p_c) \le c_2 < 1 \qquad \text{in } d = 2 \tag{1.5}$$

uniformly in L. The behavior in (1.5) is expected to hold for all $d < d_c$, but not for $d \ge d_c$, where d_c is the so-called upper critical dimension, above which the critical exponents assume their mean-field values.

Using the relation (see Lemma 4.2)

$$R_L(p) \le N_L(p) \le \sum_{m=1}^{\infty} [R_L(p)]^m \,, \tag{1.6}$$

equations (1.4) and (1.5) give the rigorous results $N_L(p_c) \ge c_1 > 0$ in all $d \ge 2$, and $N_L(p_c) \le c_3 < \infty$ in $d = 2$, uniformly in L. It has been argued [Con85] that the latter behavior persists for all $d < d_c$, i.e.

$$N_L(p_c) \le c_3 < \infty \qquad \text{in all } d < d_c \tag{1.7}$$

uniformly in L, but that

$$N_L(p_c) \sim L^{d-d_c} \qquad \text{in all } d > d_c \,. \tag{1.8}$$

The behavior $N_L(p_c) \sim L^{d-d_c}$ has recently been rigorously established [Aiz97] under a natural, but as yet unproven, assumption on the decay of the two-point connectivity function at p_c, which is presumably true above d_c.

Returning to $R_L(p)$, as mentioned above, it is expected that

$$R_L(p_c) \le c_2 < 1 \qquad \text{in all } d < d_c \tag{1.9}$$

uniformly in L, while (1.6) and (1.8) would require that

$$R_L(p_c) \to 1 \quad \text{as} \quad L \to \infty \quad \text{in all } d > d_c \,. \tag{1.10}$$

Here we are interested in a different aspect of the critical behavior: namely, the relationship between the behavior of $R_L(p_c)$ and hyperscaling. A hyperscaling relation is a relationship among critical exponents that explicitly involves the spatial dimension d. Hyperscaling relations are expected to hold up to, but not above, the upper critical dimension. We will consider two explicit hyperscaling relations, which we will specify below.

We will prove that

$$R_L(p_c) < 1 \quad \text{uniformly in } L \qquad \Longrightarrow \qquad \text{hyperscaling} \tag{1.11}$$

(see Theorems 3.2 and 3.3 for a precise formulation). Actually, in the course of proving this result, we will prove the stronger statement

$$N_L(p_c) \le c_3 < \infty \quad \text{uniformly in } L \qquad \Longrightarrow \qquad \text{hyperscaling} \,. \tag{1.12}$$

This establishes half of the heuristic picture of Coniglio [Con85], who was the first to relate uniform boundedness of $N_L(p_c)$ to hyperscaling, and the behavior $N_L(p_c) \sim L^{d-d_c}$ to the breakdown of hyperscaling. In fact, our theorem requires even less than is indicated in (1.12): rather than uniform boundedness of $N_L(p_c)$, we need only that $N_L(p_c)$ grows more slowly than any power of L, a distinction which is presumably important in $d = d_c$.

In order to specify the hyperscaling relations in (1.11) and (1.12), we define several fundamental quantities in percolation. Let

$$\tau(\mathbf{0}, v; p) = Pr_p(\mathbf{0} \leftrightarrow v) \tag{1.13}$$

be the probability that the origin is connected to v (the point-to-point connectivity function), let

$$\widetilde{\pi}_n(p) = Pr_p\{\exists\ v = (n, \cdot)\ \text{such that}\ \mathbf{0} \leftrightarrow v\} \tag{1.14}$$

be the probability that the origin is connected to a plane a distance n away (the point-to-plane connectivity function), and let

$$P_s = P_s(p) = Pr_p(|\mathcal{C}(\mathbf{0})| = s) \tag{1.15}$$

be the probability that the cluster of the origin is of size s (the cluster size distribution). At p_c, it is believed that these quantities decay with the power laws:

$$\sup_{v: |v|_\infty \geq n} \tau(\mathbf{0}, v; p_c) \approx \frac{1}{n^{d-2+\eta}}\,, \tag{1.16}$$

$$\widetilde{\pi}_n(p_c) \approx n^{-1/\rho}\,, \tag{1.17}$$

and

$$P_s(p_c) \approx s^{-(1+1/\delta)}\,, \tag{1.18}$$

where $\approx$ means equality up to a slowly varying function (e.g. a logarithm), see Section 2 for the precise definition. The hyperscaling relations referred to in (1.11) and (1.12) are

$$d\rho = \delta + 1 \tag{1.19}$$

and

$$2 - \eta = d\,\frac{\delta - 1}{\delta + 1}\,. \tag{1.20}$$

The relation (1.19) is standard. Assuming usual scaling relations $\gamma + 2\beta = \beta(\delta + 1)$ and $\gamma = \nu(2 - \eta)$ (see Section 2 for definitions of β, γ and ν), the relation (1.20) is equivalent to the standard hyperscaling relation $d\nu = \gamma + 2\beta$.

The hyperscaling relations (1.19) and (1.20) will be proved in two steps. First, assuming only the existence of the exponents, we will prove that

$$d\rho \geq \delta + 1 \qquad \text{and} \qquad d - 2 + \eta \geq 2/\rho. \tag{1.21}$$

Then, using the uniform boundedness assumptions in (1.11) and (1.12), respectively, we will prove equality in (1.21), and hence the hyperscaling relations (1.19) and (1.20).

The results of this paper came out of our attempt to understand the finite-size scaling of the largest cluster in a finite box [BCKS97], see [CPS96] for an announcement of our results. In fact, the main technical step in the proof of the upper bound on $d\rho$ is Lemma 5.1 below, which bounds the expected size of the largest cluster in a box of side length $6n$ in terms of $n^d \pi_{3n}(p_c)$ and $N_n(p_c)$. Assuming uniform boundedness of $N_n(p_c)$, this bound is then converted into a bound on $\pi_n(p_c)$ that implies the bound $d\rho \leq \delta + 1$, see Proposition 5.2 and its proof for details.

The organization of the paper is as follows: In Section 2, we give definitions and notation. In Section 3, we give a precise statement of the results reviewed here, as well as some additional results. The proofs are in Sections 4 – 7; see Section 3.6 for a detailed directory of the proofs. Sections 5 and 6 contain results which may be of interest in their own right. Section 5 establishes scaling properties of some of the fundamental quantities in percolation, and Section 6 has a general moment estimate. Section 7 uses the results of Section 6 to prove an exponential tail estimate for the subcritical cluster size distribution.

Acknowledgment: The authors wish to thank the the Forschungsinstitut of the ETH in Zürich and the Institute for Advanced Study in Princeton for their hospitality and partial support of the research in this paper. The authors are also grateful for partial support from other sources: C.B. was supported by the Commission of the European Union under the grant CHRX-CT93-0411, J.T.C. by NSF grant DMS-9403842, and H.K. by an NSF grant to Cornell University.

2. Definitions and Background

Consider the hypercubic lattice $\mathbb{Z}^d$. Bond percolation on $\mathbb{Z}^d$ is defined by choosing each bond between adjacent sites of $\mathbb{Z}^d$ to be *occupied* with probability p and *vacant* with probability $1-p$, independently of all other bonds. The corresponding product measure on configurations of occupied and vacant bonds is denoted by Pr_p, and expectation with respect to the measure Pr_p is denoted by E_p. A generic configuration is denoted by ω. If $S_1, S_2, S_3 \subset \mathbb{Z}^d$, we say that S_1 is connected to S_2 in S_3, denoted by $\{S_1 \leftrightarrow S_2 \text{ in } S_3\}$, if there exists an occupied path with vertices in S_3 from some site of S_1 to some site of S_2. Maximal connected subsets of sites are called *(occupied) clusters*. The occupied cluster (in the configuration ω) containing the site x is denoted by $\mathcal{C}(x) = \mathcal{C}(x;\omega)$. The size of the cluster $\mathcal{C}$, denoted by $|\mathcal{C}|$, is the number of sites in $\mathcal{C}$. $\mathcal{C}_\infty$ denotes the (unique) infinite cluster, i.e., the occupied cluster with $|\mathcal{C}| = \infty$. We also consider connected clusters $\mathcal{C}_\Lambda(x) = \mathcal{C}_\Lambda(x;\omega)$ in a finite box $\Lambda \subset \mathbb{Z}^d$, defined as the set of all points y in Λ which are connected to x by an occupied path with vertices in Λ. We say that $\mathcal{C}$ is a cluster in Λ, if $\mathcal{C} = \mathcal{C}_\Lambda(x)$ for some $x \in \Lambda$. The origin in $\mathbb{Z}^d$ is denoted by $\mathbf{0}$.

The cluster size distribution is characterized by

$$P_s = P_s(p) = Pr_p(|\mathcal{C}(\mathbf{0})| = s) \tag{2.1}$$

or alternatively by

$$P_{\geq s} = P_{\geq s}(p) = Pr_p(|\mathcal{C}(\mathbf{0})| \geq s)\,, \tag{2.2}$$

and the order parameter of the model is the *percolation probability* or infinite-cluster density

$$P_\infty(p) = Pr_p(|\mathcal{C}(\mathbf{0})| = \infty)\,. \tag{2.3}$$

The *critical probability* is

$$p_c = \inf\{p : P_\infty(p) > 0\}\,. \tag{2.4}$$

We consider several connectivity functions: the *(point-to-point) connectivity function*

$$\tau(v, w; p) = Pr_p(v \leftrightarrow w)\,, \tag{2.5}$$

the *finite-cluster (point-to-point) connectivity function*

$$\tau^{\mathrm{fin}}(v, w; p) = Pr_p(v \leftrightarrow w, |\mathcal{C}(v)| < \infty)\,, \tag{2.6}$$

the *point-to-hyperplane connectivity function*

$$\widetilde{\pi}_n(p) = Pr_p\{\exists\ v = (n, \cdot)\,\text{such that}\,\mathbf{0} \leftrightarrow v\}\,, \tag{2.7}$$

and the *point-to-box connectivity function*

$$\pi_n(p) = Pr_p\{\mathbf{0} \leftrightarrow \partial B_n(\mathbf{0})\}\,, \tag{2.8}$$

where

$$B_n(v) = \{w \in \mathbb{Z}^d : |v - w|_\infty \le n\} = [-n, n]^d \cap \mathbb{Z}^d \tag{2.9}$$

with $|\cdot|_\infty$ denoting the ℓ_∞-norm. B_n wil be short for $B_n(\mathbf{0})$. Notice that $\pi_n(p)$ and $\widetilde{\pi}_n(p)$ are equivalent, in the sense that

$$\widetilde{\pi}_n(p) \le \pi_n(p) \le 2d\widetilde{\pi}_n(p)\,. \tag{2.10}$$

We also consider the susceptibilities

$$\chi(p) = E_p(|\mathcal{C}(\mathbf{0})|) = \sum_v \tau(\mathbf{0}, v; p) \tag{2.11}$$

and

$$\chi^{\mathrm{fin}}(p) = E_p(|\mathcal{C}(\mathbf{0})|, |\mathcal{C}(\mathbf{0})| < \infty) = \sum_v \tau^{\mathrm{fin}}(\mathbf{0}, v; p) = \sum_{s<\infty} sP_s(p) \tag{2.12}$$

Finally, we introduce the quantity

$$s(n) = (2n)^d\,\pi_n(p_c)\,. \tag{2.13}$$

As we will see, $s(n)$ is the order of magnitude of the size of the largest critical clusters on scale n.

Length scales in the model are naturally expressed in terms of the correlation length $\xi(p)$, defined by the limit

$$1/\xi(p) = -\lim_{|v|_\infty \to \infty} \frac{1}{|v|_\infty} \log \tau^{\mathrm{fin}}(\mathbf{0}, v; p) \tag{2.14}$$

taken with v along a coordinate axis. It is known that $\xi(p) < \infty$ for all $p \neq p_c$ and $\xi(p) \to \infty$ as $p \uparrow p_c$ (see [Gri89], Theorem 5.49 and equation (5.57) for $p < p_c$; for $p > p_c$ the finiteness of the correlation length follows from [GM90]), but there is no proof yet that $\xi(p) \to \infty$ as $p \downarrow p_c$ in dimension $d > 2$.

Alternatively, lengths may be expressed in terms of the finite-size scaling correlation length $L_0(p, \varepsilon)$, introduced in [CCF85] for $p < p_c$ and in [CC87] for $p > p_c$. $L_0(p, \varepsilon)$ is defined in terms of the probablities

$$\begin{aligned} R^{\mathrm{fin}}_{L,M}(p) = Pr_p\{ &\exists \text{ a finite, occupied cluster } \mathcal{C} \text{ containing a bond-crossing of} \\ &[0, L] \times [0, M] \cdots \times [0, M] \text{ in the 1-direction}\}\,. \end{aligned} \tag{2.15}$$

These are finite-cluster analogues of the crossing probability

$$\begin{aligned} R_{L,M}(p) = Pr_p\{ &\exists \text{ a occupied bond-crossing of } [0, L] \times [0, M] \cdots \times [0, M] \\ &\text{in the 1-direction}\}\,, \end{aligned} \tag{2.16}$$

discussed in the introduction.

By [CCGKS89], Theorem 5, and the fact that $\xi(p) < \infty$ for $p \neq p_c$, the crossing probability $R^{\mathrm{fin}}_{L,3L}(p) \to 0$ as $L \to \infty$. For $p \neq p_c$, the finite-size scaling length

$$L_0(p) = L_0(p, \varepsilon) = \min\{L \geq 1 \mid R^{\mathrm{fin}}_{L,3L}(p) \leq \varepsilon\} \tag{2.17}$$

is therefore well defined and finite. In fact, we may use the bounds of Theorem 5 in [CCGKS89] to show that for each $\varepsilon > 0$ there are constants $C_1 = C_1(d, \varepsilon) < \infty$ and $C_2 = C_2(d) < \infty$, such that

$$\frac{L_0(p, \varepsilon) - 1}{\xi(p)} \leq C_1 + C_2 \log\bigl(1 + \xi(p)\bigr) \quad \text{if} \quad p \neq p_c\,. \tag{2.18}$$

For $p < p_c$, $R^{\mathrm{fin}}_{L,3L}(p) = R_{L,3L}(p)$, and the finite-size scaling correlation length $L_0(p)$ can be analyzed by the methods of [ACCFR83], [CC84], [CCF85], [CCFS86] and [Kes87]. It then is straightforward to show that there exists a constant $a(d) > 0$ such that for $\varepsilon < a(d)$, the scaling behavior of $L_0(p, \varepsilon)$ is independent of ε for $p < p_c$, in the sense that $L_0(p, \varepsilon_1)/L_0(p, \varepsilon_2)$ is bounded away from 0 and infinity for any two fixed values $\varepsilon_1, \varepsilon_2 < a(d)$. This scaling behavior is also essentially the same as that of the standard correlation length $\xi(p)$. More specifically, for $0 < \varepsilon < a(d)$, the bound (2.18) can be complemented by the lower bound[1]

$$L_0(p, \varepsilon) \geq C_3 \xi(p) \quad \text{if} \quad p < p_c\,, \tag{2.19}$$

[1]As for the upper bound (2.18), K. Alexander [Ale96] has recently shown that one can take $C_2(d) = 0$ if $d = 2$ and $p < p_c$.

for some constant $C_3 = C_3(d, \varepsilon) > 0$ that can be made arbitrary large by chosing $a(d)$ small enough; this follows for instance, with $C_3 = \frac{1}{2}\log(A(d)/\varepsilon)$, from the fact that $\tau(\mathbf{0}, x; p) \leq 2dR_{|x|,3|x|}$ and (4.4) below. Hereafter we will assume that $\varepsilon < a(d)$; we usually suppress the ε-dependence in our notation.

Above p_c, we expect that the definition (2.17) again coincides, say in the sense of equation (2.18) and (2.19), with the standard correlation length $\xi(p)$ above threshold, but we are actually not able to prove an analogue to (2.19) for $p > p_c$, except for $d = 2$, where one can use a Harris ring construction [Har60] (see also [BCKS97]) in conjunction with the Russo-Seymour-Welsh Lemma ([Rus78], [SW78]). Assuming that $P_\infty(p_c) = 0$, it is known, however, (for any $d \geq 2$) that $L_0(p) \to \infty$ as $p \downarrow p_c$, see [CC87]. Note that our results do not depend on the validity of the bound (2.19) *above* p_c.

A quantity that is very much related to the crossing probability $R_{L,M}$ is the expected number of crossing clusters,

$$N_{L,M}(p) = E_p\{\mathcal{N}_{L,M}\}\,, \tag{2.20}$$

where

$$\begin{aligned}\mathcal{N}_{L,M} := &\text{ number of occupied clusters in } [0,L]\times[0,M]\times\cdots\times[0,M]\\ &\text{ that cross } [0,L]\times[0,M]\times\cdots\times[0,M] \text{ in the 1-direction}\,.\end{aligned} \tag{2.21}$$

It is easy to see (cf. Section 4) that

$$R_{L,M}(p) \leq N_{L,M}(p) \leq \sum_{m=1}^{\infty} [R_{L,M}(p)]^m\,, \tag{2.22}$$

which immediately implies that $N_{L,3L}(p) \to 0$ as $L \to \infty$ if $p < p_c$. For $p > p_c$, on the other hand, $R_{L,3L}(p) \to 1$ as $L \to \infty$, so that $\liminf_{L\to\infty} N_{L,3L}(p) \geq 1$. Uniqueness of the infinite cluster therefore suggests that $N_{L,3L}(p) \to 1$ as $L \to \infty$ for fixed $p > p_c$. This is indeed the case, see Lemma 4.2.

At the critical point, finally, it is expected [Con85], [Arc87] that $N_{L,3L}(p)$ is bounded uniformly in L if d is smaller than the so-called upper critical dimesion d_c, and grows like L^{d-6} above d_c, with logarithmic corrections for $d = d_c$. Very recently, Aizenman [Aiz97] has rigorously verified certain aspects of the expected high-dimensional behavior under a strong but natural assumption on the behavior of the connectivity function above d_c.

We close this section with the definitions of some of the standard power laws which are expected to characterize the scaling behavior of relevant quantities in percolation, noting that the existence of these power laws has not yet been rigorously established in low dimensions. We define $G(n) \approx n^\alpha$ to mean

$$G(n) = g(n)n^\alpha \tag{2.23}$$

where $g(n)$ is a slowly varying function in the sense that for each $\varepsilon > 0$ one can find constants $n_0 = n_0(\varepsilon) < \infty$ and $C = C(\varepsilon) < \infty$ such that for all $m \geq n \geq n_0$

$$\frac{1}{C(\varepsilon)}\Big(\frac{m}{n}\Big)^{-\varepsilon} \leq \frac{g(m)}{g(n)} \leq C(\varepsilon)\Big(\frac{m}{n}\Big)^{\varepsilon}. \tag{2.24}$$

In a similar way, we use $G(p) \approx |p-p_c|^\alpha$ to mean

$$G(p) = g(p)|p-p_c|^\alpha\,, \tag{2.25}$$

where $g(p)$ is again a slowly varying function, this time in the sense that for each $\varepsilon > 0$, there exist constants $b = b(\varepsilon) > 0$ and $C = C(\varepsilon) < \infty$ such that

$$\frac{1}{C(\varepsilon)}\Big(\frac{p'-p_c}{p-p_c}\Big)^{-\varepsilon} \leq \frac{g(p')}{g(p)} \leq C(\varepsilon)\Big(\frac{p'-p_c}{p-p_c}\Big)^{\varepsilon} \tag{2.26}$$

if $0 < |p-p_c| \leq |p'-p_c| \leq b$ and either both p and p' lie below p_c or both p and p' lie above p_c. Note that these requirements are stronger than saying that $\log G(n)/\log n \to \alpha$ and $\log G(p)/\log|p-p_c| \to \alpha$, respectively, but weaker than saying that $g(\cdot)$ is slowly varying in the traditional sense.

At p_c, the power laws of relevance to us are

$$\pi_n(p_c) \approx n^{-1/\rho}\,, \tag{2.27}$$

$$P_{\geq s}(p_c) \approx s^{-1/\delta}\,, \tag{2.28}$$

and

$$\sup_{x:|x|_\infty \geq n} \tau(\mathbf{0}, x; p_c) \approx \frac{1}{n^{d-2+\eta}}\,, \tag{2.29}$$

where we have slightly deviated from the usual definition $\tau(\mathbf{0}, x; p_c) \approx |x|_\infty^{-(d-2+\eta)}$. For the approach to the critical point, the standard power laws are

$$\xi(p) \approx |p-p_c|^{-\nu} \quad p < p_c\,, \tag{2.30}$$

$$\chi(p) \approx |p-p_c|^{-\gamma} \quad p < p_c\,, \tag{2.31}$$

$$\xi(p) \approx |p-p_c|^{-\nu'} \quad p > p_c\,, \tag{2.32}$$

$$\chi(p) \approx |p-p_c|^{-\gamma'} \quad p > p_c\,, \tag{2.33}$$

and

$$P_\infty(p) \approx |p-p_c|^{\beta} \quad p > p_c\,. \tag{2.34}$$

3. Statement of Results

Our main results are the hyperscaling equalities (1.19) and (1.20), which we will prove under certain assumptions, namely the uniform boundedness of critical crossing probabilities and the existence of the exponent ρ. However, before stating our main results in a more precise form, we will state the corresponding hyperscaling *in*equalities, which will be proved without any assumptions.

3.1. Hyperscaling Inequalities.

In order to prove the inequalities, we will use a slight modification of the quantity $s(m)$ introduced in (2.13), namely we will consider

$$\tilde{s}(n) = \frac{1}{\pi_n(p_c)} \sum_{m=0}^{n} |\partial B_m| \pi^2_{\lfloor m/2 \rfloor}(p_c), \tag{3.1}$$

where the boundary ∂B_m of the hypercube B_m is the set of all points in $\mathbb{Z}^d$ that have ℓ_∞ distance m from the origin. Note that

$$\tilde{s}(n) \geq |B_n| \pi_n(p_c) \geq s(n) \tag{3.2}$$

by the monotonicity of $\pi_n(p_c)$ in n.

Proposition 3.1. *For all $d \geq 2$,*

$$\sup_{x : |x|_\infty | \geq 2n} \tau(\mathbf{0}, x; p_c) \leq [\pi_n(p_c)]^2 \tag{3.3}$$

and

$$P_{\geq \tilde{s}(n)}(p_c) \leq 2\pi_n(p_c). \tag{3.4}$$

Corollary. *Assume that the exponents ρ, η, and δ exist. Then*

$$d - 2 + \eta \geq 2/\rho. \tag{3.5}$$

and

$$d\rho \geq \delta + 1 \tag{3.6}$$

Proof. The inequality (3.5) follows immediately from (3.3) and the definitions of the exponents ρ and η. By Proposition 4.3 below, the existence of ρ implies that $1/\rho \leq (d-1)/2$. Therefore, if ρ exists,

$$\sum_{m=0}^{n} |\partial B_m| \pi_m^2(p_c) \approx n^{d-2/\rho} \tag{3.7}$$

and hence

$$\tilde{s}(n) \approx s(n) \approx n^{d-1/\rho}. \tag{3.8}$$

If δ also exists, this in turn implies

$$P_{\geq \tilde{s}(n)}(p_c) \approx \tilde{s}(n)^{-1/\delta} \approx n^{-(d-1/\rho)/\delta}. \tag{3.9}$$

Equation (3.4) therefore implies the inequality (3.6). $\square$

3.2. Hyperscaling in Terms of the Critical Exponents ρ, δ and η.

We begin by stating our main result in its simplest form, namely that uniform boundedness of the critical crossing probabilities implies hyperscaling, provided the exponent ρ exists. First, however, we must precisely define our notion of uniform boundedness. We say that *the critical crossing probabilities are uniformly bounded* if there exists a constant $\widetilde{\varepsilon} > 0$ such that

$$R_{L,3L}(p_c) \leq 1 - \widetilde{\varepsilon} \quad \text{for all} \quad L \geq 1 \,. \tag{3.10}$$

Theorem 3.2. *Assume that the critical exponent ρ exists in the sense*[2] *of* (2.27). *Then uniform boundedness* (3.10) *of the critical crossing probabilities implies that the exponents δ and η exist* (*in the sense of* (2.28) *and* (2.29)), *with*

$$d\rho = \delta + 1 \tag{3.11}$$

and

$$2 - \eta = d\frac{\delta - 1}{\delta + 1} \,. \tag{3.12}$$

Remarks:.

(i) Assuming the usual scaling relations $\gamma + 2\beta = \beta(\delta + 1)$ and $\gamma = \nu(2 - \eta)$, the relation (3.12) is equivalent to the standard hyperscaling relation $d\nu = \gamma + 2\beta$.

(ii) In the course of proving Theorem 3.2, we will in fact prove the stronger statement that uniform boundedness of $N_{L,3L}(p_c)$, and the existence of ρ in the sense of (2.27), imply the existence of δ and η in the sense of (2.28) and (2.29), and imply (3.11) and (3.12). See Remark (ix). If the existence of all three exponents is assumed, we can prove the even stronger statement that hyperscaling is valid as long as $N_{L,3L}(p_c)$ grows more slowly than any power in L, see Theorem 3.3 below.

Theorem 3.3. *Assume that the critical exponents ρ, δ and η exist in the sense of* (2.27), (2.28) *and* (2.29). *If $N_{L,3L}(p_c)$ grows more slowly than any power L (i.e., if for all $\varepsilon > 0$ there exist a constant $C_\varepsilon < \infty$ such that $N_{L,3L}(p_c) \leq C_\varepsilon L^\varepsilon$), then ρ, δ and η obey the hyperscaling relations* (3.11) *and* (3.12).

3.3. An Axiomatic Approach to Hyperscaling.

In order to prove Theorem 3.2, we will first prove a similar theorem in terms of upper and lower bounds on $P_{\geq s}$ and τ, *without* assuming the existence of any exponents. Given these bounds, we will then be able to show that the existence of ρ implies the existence of δ and η, together with the hyperscaling relations (3.11) and (3.12). We will also prove several additional results, which are needed in a companion paper [BCKS97]. The axiomatic form of Theorem 3.2 is stated in this subsection, and the additional results are given in the next subsection. In order to

[2] Recall that the symbol $\approx$ is defined as equality up to a slowly varying function, see (2.23) and (2.24).

state these results, we will need several assumptions. Two of them, Assumptions (I) and (II) below, will be used to prove the axiomatic form of Theorem 3.2; these two assumptions deal with behavior at p_c only. The two additional Assumptions (III) and (IV) below, will be used to prove the further results stated in Section 3.4.

As before, our first assumption is the uniform boundedness of $R_{L,3L}$ at p_c. The second replaces the assumption that ρ exists, and can actually be proven from the existence of ρ. The third, which will be used in only one theorem below, Theorem 3.6, is the assumption that uniform boundedness of $R_{L,3L}(p)$ continues to hold for $p > p_c$, as long as $L \leq L_0(p)$. Finally, our fourth assumption is that $\pi_n(p)$ behaves like $\pi_n(p_c)$ as long as $n \leq L_0(p)$. Since $L_0(p)$ depends on ε, see equation (2.17), Assumptions (III) and (IV) implicitly involve the constant ε. We assume that they are true for all nonzero $\varepsilon < \varepsilon_0$, where $\varepsilon_0 = \varepsilon_0(d)$ is a suitable constant. Our assumptions are:

(I) There exists a constant $\widetilde{\varepsilon} > 0$ such that

$$R_{L,3L}(p_c) \leq 1 - \widetilde{\varepsilon} \quad \text{for all} \quad L \geq 1\,. \tag{3.13}$$

(II) There exist constants $D_1 > 0$ and $\rho_1 > \frac{2}{d}$, such that

$$\frac{\pi_m(p_c)}{\pi_n(p_c)} \geq D_1 \Big(\frac{m}{n}\Big)^{-1/\rho_1} \quad \text{for all} \quad m \geq n \geq 1; \tag{3.14}$$

(III) There exists a constant $\widetilde{\varepsilon} > 0$ such that

$$R_{L,3L}(p) \leq 1 - \widetilde{\varepsilon} \quad \text{if} \quad L \leq L_0(p)\,. \tag{3.15}$$

(IV) There exist constants $D_2 > 0$ and $D_3 < \infty$ such that

$$D_2 \leq \frac{\pi_n(p)}{\pi_n(p_c)} \leq D_3 \quad \text{if} \quad n \leq L_0(p); \tag{3.16}$$

Remarks.

(iii) As noted above, Assumption (II) follows from the assumption that the exponent ρ exists in the sense of (2.27). Indeed, by Proposition 4.3 below,

$$\pi_n(p_c) \geq C_1 n^{-\frac{d-1}{2}}\,, \tag{3.17}$$

for some constant $C_1 = C_1(d) > 0$. The existence of the exponent ρ in the sense of equation (2.27) therefore immediately implies the bound

$$1/\rho \leq \frac{d-1}{2}\,. \tag{3.18}$$

Using once more the assumption that ρ exists, we have that for all $\varepsilon > 0$ there is a constant $C_\varepsilon > 0$ such that

$$\frac{\pi_m(p_c)}{\pi_n(p_c)} \geq C_\varepsilon \left(\frac{m}{n}\right)^{\varepsilon - 1/\rho} \qquad \text{for all} \qquad m \geq n\,. \tag{3.19}$$

Equations (3.18) and (3.19) together imply Assumption (II) with, for example, $\rho_1 = \frac{2}{d-1/2}$.
(iv) By the rescaling inequalities of Lemma 4.1 (iii) below, Assumption (III) is equivalent to the (formally weaker) assumption that there exist constants $\bar{\varepsilon} > 0$ and $\sigma_0 > 0$ such that $R_{L,3L}(p) \leq 1 - \bar{\varepsilon}$ if $L \leq \sigma_0 \, L_0(p)$.
(v) By Theorem 3.8 ii) below, Assumption (I) implies that $\pi_n(p_c) \to 0$ as $n \to \infty$, and hence $P_\infty(p_c) = 0$. Note that this implies in particular that $L_0(p) \to \infty$ as $p \downarrow p_c$, see Section 2 above.
(vi) The fact that $L_0(p)$ below p_c is defined in terms of the crossing probabilities $R_{L,3L}(p)$ is used at several points in this paper. By contrast, our results *above* p_c depend only weakly on the precise definition of $L_0(p)$. Namely, our results remain true as long as Assumptions (III) and (IV) hold for our definition of $L_0(p)$ above p_c.
(vii) In many lemmas and propositions below we can replace Assumption (I) by the following weaker assumption:

$$\text{(I*)} \qquad N_{L,3L}(p_c) = E_{p_c}\{\mathcal{N}_{L,3L}\} \text{ is bounded in } L.$$

Lemma 4.2 below shows that this assumption is implied by Assumption (I). We explictly indicate below when (I) can be replaced by (I*).

The theorems in this and the following subsection describe various important properties of the quantities $\pi_n, P_s, P_{\geq s}$ and χ. Throughout, the basic parameter p is bounded away from 0 and 1, that is we restrict p to $\zeta_0 \leq p \leq 1 - \zeta_0$ for some small strictly positive ζ_0. No further mention of ζ_0 will be made. Many constants C_i appear in this paper. These are always *finite and strictly positive*, even when this is not indicated. In different formulae the same symbol C_i may denote different constants. All these constants depend on ε, d, ζ_0 and the constants which appear in Assumptions (I) – (IV). This dependence will not be indicated in our notation.

Theorem 3.4. *Under Assumptions (I) and (II), there are constants C_i, $0 < C_i < \infty$ such that*

$$C_1[\pi_n(p_c)]^2 \leq \sup_{x:|x|_\infty \geq 2n} \tau(\mathbf{0}, x; p_c) \leq [\pi_n(p_c)]^2 \tag{3.20}$$

and

$$C_2 \pi_n(p_c) \leq P_{\geq s(n)}(p_c) \leq C_3 \pi_n(p_c)\,. \tag{3.21}$$

Assumption (I) is not needed for the upper bound in (3.21), and neither Assumption (I) nor Assumption (II) is needed for the upper bound in (3.20).

Note that the upper bound in (3.20) was already stated in (3.3). It is included again here for completeness. The upper bound in (3.21) is reminiscent of (3.4), but equation (3.4) concerns $\tilde{s}(n)$, while the upper bound in (3.21) concerns $s(n)$. Assumption (II) is needed to get from $\tilde{s}(n)$ to $s(n)$.

We claim that Theorem 3.4 implies Theorem 3.2. To see this, we first recall that the existence of ρ in the sense of (2.27) implies Assumption (II), see Remark (iii) above. The assumptions of Theorem 3.2 therefore imply those of Theorem 3.4.

But the bound (3.20), together with the existence of ρ, immediately implies the existence of η, with

$$d - 2 + \eta = \frac{2}{\rho}\,. \tag{3.22}$$

To see that (3.21) and the existence of ρ implies the existence of δ is slightly less obvious, and will be shown in Section 3.5. Here, we will just show that (3.21) and the existence of both ρ and δ implies (3.12). Indeed, assuming the existence of ρ and δ, we get

$$P_{\geq s(n)}(p_c) \approx s(n)^{-1/\delta} \approx n^{-(d-1/\rho)/\delta}, \tag{3.23}$$

which together with (3.21) implies that

$$\delta = d\rho - 1. \tag{3.24}$$

Solving (3.24) for ρ and inserting the result into (3.22), we obtain (3.11) and (3.12). Modulo the proof of the existence of δ, given in Section 3.5 below, the proof of Theorem 3.2 therefore reduces to that of Theorem 3.4.

Remarks.

(viii) In the course of proving Theorem 3.4, we will prove a bound on the finite-size scaling of the largest cluster in a finite box. We use the notation $W_B^{(1)}$ for the size of the largest cluster in B and recall the definition (2.9) of $B_n = B_n(\mathbf{0})$. Under the assumptions of Theorem 3.4, we then prove that for $p = p_c$, the expected size of $W_{B_n}^{(1)}$ is bounded from below and above by (two different) constants times $s(n)$, see Remark (xii) following the proof of Proposition 5.2. Assuming the existence of the exponent ρ in the sense of (2.27), this therefore implies that

$$E_{p_c}\{W_{B_n}^{(1)}\} \approx n^{d_f} \qquad \text{where} \qquad d_f = d - \frac{1}{\rho}\,. \tag{3.25}$$

As in Proposition 3.1, an *upper* bound can be obtained without any assumptions. Namely, we can show that

$$E_{p_c}\{W_{B_n}^{(1)}\} \leq 3\tilde{s}(n)\,, \tag{3.26}$$

see Remark (xii).

(ix) As stated in Theorem 3.4, Assumption (I) is not used in the proof of the upper bounds in (3.20) and (3.21). As can be seen from the proof of Proposition 5.2, it is further true that the lower bound in (3.21) remains valid if Assumption (I) is replaced by (I*). Furthermore, the lower bound in (3.20) remains valid under Assumption (I*) and (3.29) below. Since (3.29) is trivial if ρ is known to exist, it follows that all conclusions of Theorem 3.4, and hence of Theorem 3.2, remain valid under the hypotheses that Assumption (I*) holds and that ρ exists.

3.4. Additional Hyperscaling and Related Results.

In this subsection, we state several results of independent interest, some of which are related to the proof of Theorem 3.4, and others of which are necessary for the proofs in our companion paper [BCKS97] on finite-size scaling. The first gives a scaling relation for the exponent γ.

Theorem 3.5. *Under Assumptions (I), (II) and (IV), there are constants C_i such that*

$$C_1 L_0(p)^d [\pi_{L_0(p)}(p_c)]^2 \leq \chi(p) \leq C_2 L_0(p)^d [\pi_{L_0(p)}(p_c)]^2 \quad \textit{if} \quad p < p_c \,. \tag{3.27}$$

Neither Assumption (I) nor Assumption (IV) is needed for the upper bound in (3.27).

Remarks.

(x) Assume that the bound (3.27) holds and that $L_0(p) \approx \xi(p)$ as $p \uparrow p_c$ (which seems quite reasonable in view of (2.18) and (2.19)). Then the existence of the exponents ρ and ν implies the existence of γ, with $\gamma = (d - 2/\rho)\nu$. Combined with (3.22), this gives the scaling relation

$$\gamma = \nu(2 - \eta) \,. \tag{3.28}$$

(xi) Since Assumptions (I) and (IV) are not used in the proof of the upper bound in (3.27), the following bound is true in all dimensions $d \geq 2$, provided the exponents ρ, ν and γ exist:

$$\gamma \leq (d - 2/\rho)\nu. \,. \tag{3.29}$$

Next, we have a *lower* bound of the form (3.27) for $\chi^{\text{fin}}(p)$ above p_c.

Theorem 3.6. *Under Assumptions (I), (II) and (III), there is a constant $C_1 > 0$ such that*

$$\chi^{\text{fin}}(p) \geq C_1 L_0(p)^d [\pi_{L_0(p)}(p_c)]^2 \quad \textit{if} \quad p > p_c \,. \tag{3.30}$$

The next theorem is a statement relating $P_{\geq s(n)}$ at p_c to $P_{\geq s(n)}$ at $p \neq p_c$, provided $n/L_0(p)$ is small.

Theorem 3.7. *Under Assumptions (I), (II) and (IV), there are constants C_i, $0 < C_i < \infty$, and σ_1, $0 < \sigma_1 \leq 1$, such that*

$$C_1 P_{\geq s(n)}(p_c) \leq P_{\geq s(n)}(p) \leq C_2 P_{\geq s(n)}(p_c) \quad \textit{if} \quad n \leq \sigma_1 L_0(p) \,. \tag{3.31}$$

The last two theorems give upper bounds on $\pi_n(p)$ and $P_{\geq s(n)}$.

Theorem 3.8.
i) There exist constants C_i, $0 < C_i < \infty$, such that

$$\frac{\pi_n(p)}{\pi_{L_0(p)}(p)} \leq C_1 e^{-C_2 n / L_0(p)} \quad \textit{if} \quad p < p_c \quad \textit{and} \quad n \geq L_0(p) \,. \tag{3.32}$$

ii) Under Assumption (I), there exist constants $0 < C_3 < \infty$ and $0 < \rho_2 < \infty$ such that

$$\frac{\pi_m(p_c)}{\pi_n(p_c)} \leq C_3 \Big(\frac{m}{n}\Big)^{-1/\rho_2} \quad \textit{if} \quad m \geq n \,. \tag{3.33}$$

Theorem 3.9. *Under Assumption (II), there exist constants $0 < C_i < \infty$ such that*

$$\frac{P_{\geq xs(L_0(p))}(p)}{\pi_{L_0(p)}(p_c)} \leq C_1 e^{-C_2 x} \quad if \quad p < p_c \quad and \quad x \geq 1\,. \tag{3.34}$$

If, in addition, Assumptions (I) and (IV) are valid, then

$$\frac{P_{\geq xs(L_0(p))}(p)}{P_{\geq s(\sigma_1 L_0(p))}(p)} \leq C_3 e^{-C_2 x} \quad if \quad p < p_c \quad and \quad x \geq 1\,, \tag{3.35}$$

where σ_1 is the constant from Theorem 3.7.

3.5. The Existence of δ.

In this subsection, we prove that (3.33) and the existence of ρ imply the existence of δ. To this end, we introduce for $t \geq 1$, $n_* = n_*(t) := \max\{n \mid s(n) \leq t\}$. This is well defined because $s(n) \to \infty$ as $n \to \infty$, by Assumption (II). By definition we then have

$$s(n_*) \leq t \leq s(n_* + 1). \tag{3.36}$$

Combining this with the monotonicity of $P_{\geq t}$ and the bounds in (3.21) we get

$$C_2 p_c \pi_{n_*}(p_c) \leq C_2 \pi_{n_*+1}(p_c) \leq P_{\geq s(n_*+1)}(p_c) \leq P_{\geq t}(p_c) \leq P_{\geq s(n_*)} \leq C_3 \pi_{n_*}(p_c). \tag{3.37}$$

Assume now that

$$\pi_n(p_c) = g(n) n^{-1/\rho} \tag{3.38}$$

with $g(\cdot)$ satisfying (2.24). Let $\lambda > [d - 1/\rho]^{-1}$. Then choose $\varepsilon > 0$ so small that $\lambda(d - 1/\rho - \varepsilon) \geq 1$ and $C_4 = C_4(\varepsilon)$ so large that

$$\frac{1}{C(\varepsilon)} \Big(\frac{C_4}{2}\Big)^{d-1/\rho-\varepsilon} \geq 2.$$

We claim that then for any $1 \leq t_1 \leq t_2$, it holds that

$$\frac{n_*(t_2)}{n_*(t_1)} \leq C_4 \Big(\frac{t_2}{t_1}\Big)^{\lambda}. \tag{3.39}$$

To see this, write n_i for $n_*(t_i)$ and let $r = (t_2/t_1)^\lambda \geq 1, m \geq C_4 r n_1$. Then

$$\begin{aligned}
s(m) &= m^{d-1/\rho} g(m) \\
&= \left(\frac{m}{n_1+1}\right)^{d-1/\rho} \frac{g(m)}{g(n_1+1)} (n_1+1)^{d-1/\rho} g(n_1+1) \\
&\geq \frac{1}{C(\varepsilon)} \left(\frac{m}{n_1+1}\right)^{d-1/\rho-\varepsilon} s(n_1+1) \\
&\geq \frac{1}{C(\varepsilon)} \left(\frac{C_4 r}{2}\right)^{d-1/\rho-\varepsilon} t_1 \text{ (see (3.36))} \\
&= \frac{1}{C(\varepsilon)} \Big(\frac{C_4}{2}\Big)^{d-1/\rho-\varepsilon} \Big(\frac{t_2}{t_1}\Big)^{\lambda(d-1/\rho-\varepsilon)} t_1 \\
&\geq 2t_2.
\end{aligned}$$

Thus, we see that $s(m) \geq 2t_2$ for all $m \geq C_4 r n_1$, and therefore $n_2 \leq C_4 r n_1$, as claimed.

Next, let $\widetilde{\lambda} < [d-1/\rho]^{-1}$, $\varepsilon > 0$ so small that $\widetilde{\lambda}(d-1/\rho-\varepsilon) \leq 1$, and $C_5 = C_5(\varepsilon)$ so small that

$$C(\varepsilon) C_5^{d-1/\rho+\varepsilon} \leq \frac{1}{2}.$$

Then one proves analogously to (3.39) that for $1 \leq t_1 \leq t_2$,

$$\frac{n_*(t_2)}{n_*(t_1)} \geq C_5 \Big(\frac{t_2}{t_1}\Big)^{\widetilde{\lambda}}. \tag{3.40}$$

It now follows from (3.37)-(3.40) and (2.24) that if we define $h(\cdot)$ by

$$P_{\geq t}(p_c) = t^{-1/\delta} h(t)$$

with ρ given by (3.24), then $h(\cdot)$ satisfies (2.24) with g replaced by h. Thus the exponent δ exists and has the value given by (3.24), as claimed in the discussion following the statement of Theorem 3.4.

3.6. Organization of the Proofs.

As explained after the statement of Theorem 3.4, the proof of Theorem 3.2 follows from Theorem 3.4 and the considerations of the previous subsection. The proofs of the other results can be found in the next four sections.

In Section 4, we prove Theorem 3.8, Proposition 3.1 and the upper bounds in Theorem 3.4. The proofs of the latter two are contained in the proof of Proposition 4.5. Most of the remaining statements are proven in Section 5: Proposition 5.2 implies the lower bound in (3.21), and, together with Proposition 4.5, completes the proof of Theorem 3.7, while Proposition 5.3 implies the lower bound in (3.20), thus completing the proof of Theorem 3.4. Theorems 3.5 and 3.6 are just the statements of Proposition 5.4. Theorem 3.3 is proven at the end of Section 5. In Section 6, we give a general moment estimate, which is then used in Section 7 to prove Theorem 3.9.

It is worth noting that Theorem 3.9, the proof of which is rather complicated, is not needed for our results on hyperscaling. We establish Theorem 3.9 in *this* paper because its proof fits in with the others here, and we will need it for our companion paper [BCKS97] on finite-size scaling.

4. Preliminaries

We start with a lemma which follows easily from the methods of [ACCFR83]:

Lemma 4.1. *Let $d \geq 2$. Then there exists a constant $1 \leq C(d) < \infty$ such that*
i) For all p and for all $L \geq 1$

$$R_{L,6L}(p) \leq 1 - \big(1 - R_{L,3L}(p)\big)^{C(d)}. \tag{4.1}$$

ii) For all p and all M and L with $1 \le L \le M \le 2L$

$$R_{M,3M}(p) \le 1 - \big(1 - R_{L,3L}(p)\big)^{C(d)} \le C(d) R_{L,3L}(p)\,. \tag{4.2}$$

iii) For all p and all $L \ge 1$

$$R_{2L,6L}(p) \le \Big[1 - \big(1 - R_{L,3L}(p)\big)^{C(d)}\Big]^2 \le C(d)^2 \big[R_{L,3L}(p)\big]^2\,. \tag{4.3}$$

iv) For $p < p_c$, $\varepsilon < A(d) := C(d)^{-2}$ and $L \ge L_0(p;\varepsilon)$

$$R_{L,3L}(p) \le \Big(\frac{\varepsilon}{A(d)}\Big)^{L/2L_0(p;\varepsilon)}\,. \tag{4.4}$$

v) For all $L \ge 1$

$$R_{L,3L}(p_c) \ge A(d)\,. \tag{4.5}$$

Proof. Let $F_{L,M}$ be the event that in the box $[0,L] \times [0,M] \cdots \times [0,M]$, there is no crossing in the 1-direction. As shown in [ACCFR83], the event $F_{L,6L}$ can be guaranteed by patching together a finite number of translations and rotations of the event $F_{L,3L}$. This gives

$$1 - R_{L,6L}(p) = Pr_p\{F_{L,6L}\} \ge \big(Pr_p\{F_{L,3L}\}\big)^{C(d)} = \big(1 - R_{L,3L}(p)\big)^{C(d)}\,; \tag{4.6}$$

see [ACCFR83] for details. (4.2) follows from (4.1) and the fact that $R_{I,J}(p)$ is increasing in J and decreasing in I. (4.3) follows from (4.1) and the fact that $R_{2L,M}(p) \le \big[R_{L,M}(p)\big]^2$. In order to obtain (4.4), one first uses the bound (4.2) (and the fact that $C(d) \le 1/A(d)$) to obtain

$$R_{L,3L(p)} \le \frac{1}{A(d)} R_{2^k L_0(p;\varepsilon), 2^k 3 L_0(p;\varepsilon)}(p),$$

where k is chosen in such a way that $2^k L_0(p;\varepsilon) \le L < 2^{k+1} L_0(p;\varepsilon)$. Iterating the bound (4.3), one then gets (4.4). Statement v) finally follows from the bound (4.3) and the fact that $\xi(p)$ diverges as $p \uparrow p_c$. Indeed, assume that (4.5) fails for some $L_1 \ge 1$. Then $R_{n,3n}(p_c)$ decays exponentially in n by (4.3) and the argument for (4.4) with L_0 replaced by L_1. As a consequence $\pi_n(p_c)$ decays exponentially in n, which in turn implies exponential decay of $\tau(0,x;p_c)$, in contradiction with the fact that $\xi(p)$ diverges as $p \uparrow p_c$. $\square$

The next lemma gives a relation between the crossing probabilities $R_{L,M}(p)$ and the expected number of crossing clusters $N_{L,M}(p)$.

Lemma 4.2.
i) For all $d \geq 2$ and all p

$$R_{L,M}(p) \leq N_{L,M}(p) \leq \sum_{m=1}^{\infty} [R_{L,M}(p)]^m \,. \tag{4.7}$$

ii) For all $d \geq 2$, as $L \to \infty$,

$$N_{L,3L}(p) \to \begin{cases} 0 & \text{if} \quad p < p_c \\ 1 & \text{if} \quad p > p_c. \end{cases} \tag{4.8}$$

Proof.
i) The lower bound on $N_{L,M}(p)$ is obvious. We are therefore left with the upper bound. Recall the definition (2.21) of $\mathcal{N}_{L,M}$, and define $\widetilde{\mathcal{N}}_{L,M}$ to be the number of disjoint crossings of $[0, L] \times [0, M] \times \cdots \times [0, M]$ in the 1-direction. Then

$$N_{L,M}(p) = E_p\{\mathcal{N}_{L,M}\} \leq E_p\{\widetilde{\mathcal{N}}_{L,M}\} = \sum_{m=1}^{\infty} Pr_p\{\widetilde{\mathcal{N}}_{L,M} \geq m\} \,. \tag{4.9}$$

By the van den Berg-Kesten inequality [BK85],

$$Pr_p\{\widetilde{\mathcal{N}}_{L,M} \geq m\} \leq [Pr_p\{\widetilde{\mathcal{N}}_{L,M} \geq 1\}]^m \,.$$

Since $Pr_p\{\widetilde{\mathcal{N}}_{L,M} \geq 1\} = R_{L,M}(p)$, this implies the upper bound in (4.7).
ii) For $p < p_c$, the bound in (4.8) follows from (4.4) and the upper bound in (4.7). For $p > p_c$, we use that the probability of $\mathcal{N}_{L,3L} \geq 2$ goes to zero exponentially in L if $p > p_c$ by [KZ90], see also [Aiz97]. Since $\mathcal{N}_{L,3L} \leq (3L+1)^{d-1}$, this implies that $\limsup N_{L,3L}(p) \leq 1$ as $L \to \infty$. The lemma now follows from the fact that $R_{L,3L}(p) \to 1$ as $L \to \infty$ for all $p > p_c$ (see (1.1)). $\square$

Proof of Theorem 3.8.
We start with the proof of ii). By Assumption (I), the probability $1 - R_{2n,6n}$ that there is no occupied crossing in the 1-direction of the block

$$[n, 3n] \times [-3n, 3n]^{d-1} \tag{4.10}$$

is at least $\widetilde{\varepsilon}$. The cube B_{3n} is the union of B_n and the block in (4.10) plus $2d - 1$ more blocks congruent to the block in (4.10). Let F_n be the event that none of these $2d$ blocks congruent to (4.10) has an occupied crossing in the short direction. By the Harris–FKG inequality and Assumption (I), the probability (at p_c) of F_n is at least $\widetilde{\varepsilon}^{2d}$. If F_n occurs, then $\partial B(n)$ is not connected to $\partial B(3n)$. Moreover F_n is independent of the event $\{0 \leftrightarrow \partial B(n)\}$. It follows that for all n

$$\pi_{3n}(p_c) \leq \pi_n(p_c)[1 - Pr_{p_c}\{F_n\}] \leq \pi_n(p_c)[1 - \widetilde{\varepsilon}^{2d}]. \tag{4.11}$$

By iteration

$$\pi_{3^j n}(p_c) \le \pi_n(p_c)[1-\widetilde{\varepsilon}^{2d}]^j.$$

Since π_m is decreasing in m, (3.33) follows.

In order to prove i), let us assume for the moment that $n \ge 3L_0(p)$, Then, as in the proof of (4.11),

$$\begin{aligned}\frac{\pi_n(p)}{\pi_{L_0(p)}} &\le 1-[1-R_{n-L_0(p),2n}(p)]^{2d} \le 2dR_{n-L_0(p),2n}(p)\\ &\le 2dR_{n-L_0(p),3(n-L_0(p))}(p)\,,\end{aligned} \tag{4.12}$$

where we used the monotonicity of $R_{L,M}(p)$ in M in the last step. By the rescaling bound (4.4), $R_{n-L_0(p),3(n-L_0(p))}(p)$ decays exponentially in $(n-L_0(p))/L_0(p)$, so that we obtain the existence of a constant $C_2 > 0$ such that

$$\frac{\pi_n(p)}{\pi_{L_0(p)}} \le 2d\, e^{-C_2 n/L_0(p)} \quad \text{if} \quad p < p_c \quad \text{and} \quad n \ge 3L_0(p)\,. \tag{4.13}$$

For $L_0(p) \le n < 3L_0(p)$ we use the monotonicity of $\pi_n(p)$ in n to conclude the proof. □

The next proposition gives the lower bound (3.17).

Proposition 4.3. *For $d \ge 2$ there is a constant $C_1 = C_1(d) > 0$ such that*

$$\pi_n(p_c) \ge C_1\Big(\frac{1}{n}\Big)^{\frac{d-1}{2}}, \quad n \ge 1\,. \tag{4.14}$$

Proof. To prove (4.14) one can simply copy the argument from [BK85], Corollary 3.15. This argument shows that

$$R_{2n,6n}(p_c) \le \sum_{0\le x_2,\dots,x_d\le 6n} [\pi_n(p_c)]^2.$$

(4.14) follows immediately from this and the fact that $R_{L,3L}(p_c)$ is bounded away from 0 (see for instance Theorem 5.1 in [Kes82] or statement v) of Lemma 4.1 above). □

Lemma 4.4. *If Assumption (II) holds, then for $\beta > 1/\rho_1 - 1$ (and a fortiori for $\beta > d/2-1 = (d-2)/2$) there exist constants $C_1 = C(\beta, d)$ and $C_2 = C_2(d)$ such that for all $L \ge 1$*

$$\sum_{m=0}^{L} (m+1)^\beta \pi_m(p_c) \le C_1 L^{\beta+1}\pi_L(p_c), \tag{4.15}$$

and

$$\sum_{m=0}^{L} (m+1)^{d-1} \pi_m^2(p_c) \le C_2 L^d \pi_L^2(p_c), \tag{4.16}$$

Proof. By Assumption (II) and the fact that

$$\frac{\pi_0(p_c)}{\pi_L(p_c)} = \frac{1}{\pi_L(p_c)} \leq O(1)\frac{\pi_1(p_c)}{\pi_L(p_c)}\,,$$

we have

$$\begin{aligned}\sum_{m=0}^{L}(m+1)^{\beta}\pi_m(p_c) &= \pi_L(p_c)\sum_{m=0}^{L}(m+1)^{\beta}\frac{\pi_m(p_c)}{\pi_L(p_c)} \\ &\leq C_3\pi_L(p_c)\sum_{m=0}^{L}(m+1)^{\beta}\left(\frac{L}{m+1}\right)^{1/\rho_1} \\ &= C_3L^{1/\rho_1}\pi_L(p_c)\sum_{m=0}^{L}(m+1)^{\beta-1/\rho_1} \leq C_1\pi_L(p_c)L^{\beta+1}\,,\end{aligned}$$

which proves (4.15). The proof of (4.16) is almost the same (recall that ρ_1 in Assumption (II) is assumed strictly larger than $2/d$). $\square$

The last proposition in this Section gives an upper bound on $P_{\geq s(n)}(p)$, which implies in particular the upper bound in (3.21). The proof of the proposition also gives the proof of the Proposition 3.1 and the upper bound in (3.20).

Proposition 4.5. *Under Assumptions (II) and (IV), there exists a constant $0 < C_3 < \infty$ such that*

$$P_{\geq s(n)}(p) \leq C_3\pi_n(p_c) \quad \textit{if} \quad n \leq L_0(p)\,. \tag{4.17}$$

If $p \leq p_c$, (4.17) remains true for all $n < \infty$, and Assumption (IV) is not needed.

Proof. First, we claim that for $n \leq \frac{1}{2}|x-y|_\infty$ and all p

$$\tau(x,y;p) \leq [\pi_n(p)]^2\,. \tag{4.18}$$

Indeed, for $n \leq \frac{1}{2}|x-y|_\infty$, the event $\{x \leftrightarrow y\}$ is contained in the intersection of the two events $\{x \leftrightarrow \partial B_n(x)\}$ and $\{y \leftrightarrow \partial B_n(y)\}$. Since these two events are independent, and each have the probability $\pi_n(p)$, the bound (4.18) follows. Note that this proves the upper bounds in (3.3) and (3.20) in Proposition 3.1 and Theorem 3.4, respectively.

For any $s > 0$ and any $p \in [0,1]$,

$$Pr_p\{|\mathcal{C}(\mathbf{0})| \geq s \text{ and } \mathcal{C}(\mathbf{0}) \not\subseteq B_n\} \leq Pr_p\{\mathbf{0} \leftrightarrow \partial B_n\} = \pi_n(p) \tag{4.19}$$

and

$$\begin{aligned}
&Pr_p\{|\mathcal{C}(\mathbf{0})| \geq s \text{ and } \mathcal{C}(\mathbf{0}) \subset B_n\} \\
&\qquad \leq Pr_p\{|\mathcal{C}(\mathbf{0}) \cap B_n| \geq s\} \\
&\qquad \leq \frac{1}{s} E_p\{|\mathcal{C}(\mathbf{0}) \cap B_n|\} \\
&\qquad = \frac{1}{s} \sum_{x \in B_n} \tau(\mathbf{0}, x; p) \\
&\qquad \leq \frac{1}{s} \sum_{x \in B_n} \left[\pi_{\lfloor |x|/2 \rfloor}(p)\right]^2 \\
&\qquad = \frac{1}{s} \sum_{m=0}^{n} |\partial B_m| \left[\pi_{\lfloor m/2 \rfloor}(p)\right]^2 ,
\end{aligned}$$

where we have used the bound (4.18) in the second to last step. As a consequence

$$P_{\geq s}(p) \leq \pi_n(p) + \frac{1}{s} \sum_{m=0}^{n} |\partial B_m| \left[\pi_{\lfloor m/2 \rfloor}(p)\right]^2 . \tag{4.20}$$

The inequality (4.20) proves both the remaining statement of Proposition 3.1 and the statement of Proposition 4.5. Indeed, choosing $p = p_c$ and $s = \tilde{s}(n)$, it gives the bound (3.4). If instead we choose $s = s(n)$ and $n \leq L_0(p)$, we can use Assumption (IV) and Lemma 4.4 to obtain (4.17). If $p \leq p_c$, Assumption (IV) can be replaced by the monotonicity of $\pi_m(p)$ in p. $\square$

5. Some Important Scaling Properties

In this section we derive a number of scaling properties of the functions $\pi_n, P_{\geq s}$, τ and χ. To this end, we will first prove a lower bound on the expectation of the largest cluster in a finite box of the form $B_n = B_n(\mathbf{0})$, see (2.9). We need some notation. For a finite box $\Lambda \subset \mathbb{Z}^d$, we denote the connected component of x in $\mathcal{C}(x) \cap \Lambda$ by $\mathcal{C}_\Lambda(x) = \mathcal{C}_\Lambda(x;\omega)$; this is therefore the collection of all points which are connected to x by an occupied path in Λ. $\mathcal{C}_\Lambda^{(1)}, \mathcal{C}_\Lambda^{(2)}, \cdots \mathcal{C}_\Lambda^{(k)}$ denote the occupied clusters in Λ, ordered from largest to smallest size, with lexicographic order between clusters of the same size. $W_\Lambda^{(i)} = |\mathcal{C}_\Lambda^{(i)}|$ denotes the size of the ith largest cluster in Λ. For an arbitrary event A, $I[A]$ denotes the indicator function of the event A, and for an arbitrary subset $\Lambda \subset \mathbb{Z}^d$, $\partial\Lambda$ denotes the set of all points in Λ that have a nearest neighbor $y \in \Lambda^c = \mathbb{Z}^d \setminus \Lambda$.

Lemma 5.1. *Under Assumption (II), there exists a strictly positive constant C_1 such that*

$$E_p\{W_{B_{3n}}^{(1)}\} \geq C_1 \frac{s(3n)}{E_{p_c}(\mathcal{N}_{2n,6n})} , \tag{5.1}$$

provided $p \geq p_c$. If Assumption (IV) holds as well, then

$$E_p\{W^{(1)}_{B_{3n}}\} \geq C_1 \frac{s(3n)}{E_p(\mathcal{N}_{2n,6n})} \tag{5.2}$$

for all $p < p_c$ and all $n \leq L_0(p)/3$.

Proof. By monotonicity in p we may assume $p \leq p_c$. Without loss of generality we further assume that $n \geq 2$. Let

$$m = \lfloor \frac{n}{2} \rfloor ,$$

implying that that $n-1 \leq 2m \leq n$. Let B_r be the cube $B_r(\mathbf{0}) = [-r, r]^d$, let A_n be the annulus

$$A_n = \{x \in B_{3n} \mid \max_{i=1,\dots,d} |x_i| \geq n\} ,$$

and let

$$M(n) = \text{number of clusters in } A_n \text{ which connect } \partial B_n \text{ to } \partial B_{3n}.$$

For $m \leq k \leq 3n$, define

$$V(m,k) = \text{number of sites in } B_m \text{ connected by an occupied path to } \partial B_k \,.$$

Assume for a moment that $M(n) < r$ for some integer r. Then the $V(m, 3n)$ vertices in B_m which are connected to ∂B_{3n} must be connected to one of the at most $r-1$ occupied clusters connecting ∂B_n to ∂B_{3n}, so that these vertices decompose into at most $r-1$ clusters in B_{3n}. At least one of these components must have size $\geq V(m, 3n)/(r-1)$ so that

$$W^{(1)}_{B_{3n}} \geq \frac{1}{r-1} V(m,3n) \quad \text{if} \quad M(n) < r \,. \tag{5.3}$$

As a consequence,

$$\begin{aligned} E_p\{W^{(1)}_{B_{3n}}\} &\geq \frac{1}{r-1} E_p\big\{V(m,3n) I[M(n) < r]\big\} \\ &\geq \frac{1}{r-1}\Big[E_p\big\{V(m,3n)\big\} - E_p\big\{V(m,3n) I[M(n) \geq r]\big\}\Big] \end{aligned} \tag{5.4}$$

Using the Assumption (IV) in the case $p < p_c$, we bound

$$\begin{aligned} E_p\{V(m,3n)\} &\geq \sum_{v \in B_m} \widetilde{\pi}_{3n}(p) \geq |B_m| \frac{\pi_{3n}(p)}{2d} \quad \text{(see (2.10))} \\ &\geq |B_m| D_2 \frac{\pi_{3n}(p_c)}{2d} \,. \end{aligned} \tag{5.5}$$

On the other hand, we have

$$\begin{aligned} E_p\big\{V(m,3n)I[M(n)\geq r]\big\} &\leq E_p\{V(m,n)I[M(n)\geq r]\} \\ &= E_p\{V(m,n)\}Pr_p\{M(n)\geq r\} \\ &\leq |B_m|\pi_m(p)\frac{E_p\{M(n)\}}{r} \\ &\leq |B_m|\pi_m(p_c)\frac{E_p\{M(n)\}}{r}\,. \end{aligned} \tag{5.6}$$

Next we claim that

$$E_p\{M(n)\} \leq 2dE_p\{\mathcal{N}_{2n,6n}\}\,. \tag{5.7}$$

Indeed, let $S_n^{(\pm i)} = \{x\in A_n \mid \pm x_i \geq n\}$, and let $\mathcal{N}_{2n,6n}^{(\pm i)}$ be the number of occupied clusters in $S_n^{(\pm i)}$ that cross $S_n^{(\pm i)}$ in the i'th direction, $i=1,\dots,d$. Consider a cluster C_{A_n} in A_n that connects ∂B_n to ∂B_{3n}. Then, by the arguments leading to the proof of (4.11), each such cluster contains a path that crosses at least one of the $2d$ slabs $S_n^{(\pm i)}$, $i=1,\dots,d$. As a consequence,

$$M(n) \leq \sum_{i=1}^{d}\left(\mathcal{N}_{2n,6n}^{(+i)} + \mathcal{N}_{2n,6n}^{(-i)}\right)\,, \tag{5.8}$$

which in turn implies (5.7). Using finally Assumption (II) together with the fact that $3n/m \leq 9$ to bound $\pi_m(p_c)$ by a constant times $\pi_{3n}(p_c)$, we obtain the bound

$$E_p\big\{V(m,3n)I[M(n)\geq r]\big\} \leq C_2|B_m|\pi_{3n}(p_c)\frac{E_p\{\mathcal{N}_{2n,6n}\}}{r} \tag{5.9}$$

where $C_2<\infty$ is a constant that depends only on the dimension d and the constants in Assumption (II).

Combining the bounds (5.4), (5.5), and (5.9), we get

$$E_p\{W_{B_{3n}}^{(1)}\} \geq \frac{1}{r-1}|B_m|\pi_{3n}(p_c)\left[\frac{D_2}{2d} - C_2\frac{E_p\{\mathcal{N}_{2n,6n}\}}{r}\right].$$

Choosing

$$r = 1 + \lfloor \frac{4dC_2}{D_2}E_p\{\mathcal{N}_{2n,6n}\}\rfloor$$

we finally get

$$E_p\{W_{B_{3n}}^{(1)}\} \geq \frac{(D_2)^2}{(4d)^2C_2}|B_m|\pi_{3n}(p_c)\frac{1}{E_p\{\mathcal{N}_{2n,6n}\}}\,. \tag{5.10}$$

Since we took $2m+1\geq n$, and hence $|B_m|\pi_{3n}(p_c)\geq 6^{-d}s(3n)$, (5.10) proves the Lemma. $\square$

We next prove a lower bound on $P_{\geq s(n)}(p)$. Together with Proposition 4.5, this shows that $P_{\geq s(n)}(p)$ is comparable to $\pi_n(p_c)$ for $n\leq$ some multiple of $L_0(p)$, implying in particular Theorem 3.7.

Proposition 5.2. *Under the Assumptions (I), (II) and (IV), there exist constants C_4 and σ_1, with $0 < C_4 < \infty$ and $0 < \sigma_1 \le 1$, such that*

$$P_{\ge s(n)}(p) \ge C_4 \pi_n(p_c) \quad \textit{if} \quad n \le \sigma_1 L_0(p)\,. \tag{5.11}$$

If $p = p_c$, Assumption (IV) is not needed, and Assumption (I) can be replaced by Assumption (I).*

Proof. We prove the proposition under Assumptions (I), (II) and (IV), and leave it to the reader to check that Assumption (IV) is not used in the proof of the result at p_c. The fact that, for $p = p_c$, Assumption (I) can be replaced by (I*) is obvious from the proof below.

In order to prove the lower bound on $P_{\ge s(n)}(p)$, we will need a relation between the distribution of $W_\Lambda^{(1)}$ and $P_{\ge s}$. To this end, we use the fact that, for an arbitrary positive s,

$$\begin{aligned} E_p\{W_\Lambda^{(1)}\} &\le s + E_p\{W_\Lambda^{(1)} I[W_\Lambda^{(1)} \ge s]\} \\ &\le s + \sum_{i\ge 1} E_p\{W_\Lambda^{(i)} I[W_\Lambda^{(i)} \ge s]\} \\ &= s + \sum_{v\in\Lambda} Pr_p\{|\mathcal{C}_\Lambda(v)| \ge s\} \\ &\le s + \sum_{v\in\Lambda} Pr_p\{|\mathcal{C}(v)| \ge s\} \\ &= s + |\Lambda| P_{\ge s}(p)\,. \end{aligned} \tag{5.12}$$

Taking $s = \frac{1}{2}E_p\{W_\Lambda^{(1)}\}$, this gives

$$P_{\ge \frac{1}{2}E_p\{W_\Lambda^{(1)}\}}(p) \ge \frac{1}{|\Lambda|}\sum_{v\in\Lambda} Pr_p\left\{|\mathcal{C}_\Lambda(v)| \ge \frac{1}{2}E_p\{W_\Lambda^{(1)}\}\right\} \ge \frac{1}{2|\Lambda|}E_p\{W_\Lambda^{(1)}\}\,. \tag{5.13}$$

Next, we use Lemma 4.2 and the monotonicity of $R_{L,M}(p)$ in p to conclude that Assumption (I) implies that

$$E_p\{\mathcal{N}_{2n,6n}\} \le C_2 \quad \text{for all} \quad p \le p_c\,. \tag{5.14}$$

Combining the fact that $E_p\{W_{B_n}^{(1)}\}$ is monotone increasing in n with Lemma 5.1, equation (5.14) and the fact that $\pi_n(p_c)$ is monotone decreasing in n, we get the existence of a strictly positive constant $\widetilde{C}_1$ such that

$$E_p\{W_{B_n}^{(1)}\} \ge E_p\{W_{B_{3\lfloor \frac{n}{3}\rfloor}}^{(1)}\} \ge \frac{C_1}{C_2} s(3\lfloor \frac{n}{3}\rfloor) \ge \widetilde{C}_1 s(n) \quad \text{if} \quad n \le L_0(p)\,, \tag{5.15}$$

and hence, see (5.13),

$$P_{\ge\frac{1}{2}\widetilde{C}_1 s(n)}(p) \ge P_{\ge \frac{1}{2}E_p\{W_{B_n}^{(1)}\}}(p) \ge \widetilde{C}_1 \frac{s(n)}{2|B_n|} \ge C_3 \pi_n(p_c) \quad \text{if} \quad n \le L_0(p)\,, \tag{5.16}$$

where C_3 is a strictly positive constant. Now, for given n, let $k = \lceil (\frac{1}{2} D_1 \widetilde{C}_1)^{-2/d} \rceil$ and $m = kn$, where D_1 is the constant of Assumption (II). Then

$$\frac{s(m)}{s(n)} \geq D_1 \left(\frac{m}{n}\right)^{d/2} \geq \frac{2}{\widetilde{C}_1} \tag{5.17}$$

and hence

$$\begin{aligned} P_{\geq s(n)}(p) &\geq P_{\geq \frac{1}{2}\widetilde{C}_1 s(m)}(p) \geq C_3 \pi_m(p_c) \text{ (by (5.16))} \\ &\geq C_3 D_1 \left(\frac{m}{n}\right)^{-d/2} \pi_n(p_c) \geq C_4 \pi_n(p_c), \end{aligned} \tag{5.18}$$

provided $m \leq L_0(p)$. Thus (5.18) and hence (5.11) will hold for $n \leq \sigma_1 L_0(p)$, for some suitable $\sigma_1 > 0$. $\square$

Remark (xii). Combined with Proposition 4.5, Proposition 5.2 proves both Theorem 3.7 and the "hyperscaling" relation (1.19b). In the course of this proof, we have actually shown that for $p = p_c$, the expected size of $W_{B_n}^{(1)}$ is bounded from below and above by (two different) constants times $s(n)$, provided Assumptions (I) and (II) hold. Indeed, combining equations (5.15), (5.12) with $s = s(n)$, and (4.17), we get the existence of constants $0 < \widetilde{C}_i < \infty$ such that

$$\widetilde{C}_1 s(n) \leq E_{p_c}\{W_{B_n}^{(1)}\} \leq \widetilde{C}_2 s(n)\,. \tag{5.19}$$

Bounds of this form and extensions thereof are studied in great detail in [BCKS97]. Note that Assumption (I) is not needed for the upper bound in (5.19), so that the "hyperscaling inequality"

$$d_f \leq d - 1/\rho \tag{5.20}$$

is valid under the sole assumption that ρ exists (see (3.25) for the definition of d_f). Indeed, if we are willing to replace $s(n)$ in (5.19) by $\tilde{s}(n)$, then the upper bound holds without any assumption, see (3.26). To see this, we use (5.12) with $s = \tilde{s}(n)$, followed by (3.4) and (3.2).

Proposition 5.3. *Under Assumptions (I), (II) and (IV), there are constants* $0 < \alpha < 1$ *and* $0 < C_1 < \infty$ *such that*

$$\sum_{x \in B_n(\mathbf{0})} \tau(\mathbf{0}, x; p) \geq C_1 n^d [\pi_n(p_c)]^2 \quad \text{if} \quad n \leq L_0(p) \tag{5.21}$$

and

$$\sum_{x \in B_{\lfloor \alpha n \rfloor}(\mathbf{0})} \tau(\mathbf{0}, x; p) \leq \frac{C_1}{2} n^d [\pi_n(p_c)]^2 \quad \text{if} \quad n \leq L_0(p)\,. \tag{5.22}$$

As in Proposition 5.2, if $p = p_c$*, Assumption (IV) is not needed.*

Corollary. *Under Assumptions (I) and (II) the lower bound in (3.20) holds.*

Proof of Corollary. We take $p = p_c$ and subtract (5.22) from (5.21) to obtain

$$(2n+1)^d \sup_{\lfloor \alpha n \rfloor < |x| \le n} \tau(\mathbf{0}, x; p_c) \ge \sum_{x \in B_n \setminus B_{\lfloor \alpha n \rfloor}} \tau(\mathbf{0}, x; p_c) \ge n^d [\pi_n(p_c)]^2 \frac{C_1}{2}.$$

Using Assumption (II), we get

$$\sup_{x \ge \lfloor \alpha n \rfloor} \tau(\mathbf{0}, x; p_c) \ge \frac{C_1}{2} D_1^2 \alpha^{2/\rho_1} [\pi_{\lfloor \alpha n \rfloor}(p_c)]^2.$$

Now just replace $\lfloor \alpha n \rfloor$ by n. $\square$

Proof of Proposition 5.3. We again prove the proposition under Assumptions (I), (II) and (IV), and leave it to the reader to check that Assumption (IV) is not needed at p_c.

In order to prove (5.21), we may assume without loss of generality that $p \le p_c$. For $B_n = B_n(\mathbf{0})$, and $m \le \sigma_1 n$, where σ_1 is the constant of Proposition 5.2, we then bound

$$\begin{aligned}
\sum_{x \in B_n(\mathbf{0})} \tau(\mathbf{0}, x; p) &= E_p\{|\mathcal{C}(\mathbf{0}) \cap B_n|\} \\
&\ge s(m) Pr_p\{|\mathcal{C}(\mathbf{0}) \cap B_n| \ge s(m)\} \\
&\ge s(m) Pr_p\{|\mathcal{C}(\mathbf{0})| \ge s(m), \mathbf{0} \not\leftrightarrow \partial B_n\} \\
&\ge s(m) \left(Pr_p\{|\mathcal{C}(\mathbf{0})| \ge s(m)\} - \pi_n(p)\right) \\
&\ge s(m) \left(Pr_p\{|\mathcal{C}(\mathbf{0})| \ge s(m)\} - \pi_n(p_c)\right) \\
&= s(m) \left(P_{\ge s(m)}(p) - \pi_n(p_c)\right) \\
&\ge s(m) \left(C_4 \pi_m(p_c) - \pi_n(p_c)\right) ,
\end{aligned} \tag{5.23}$$

where C_4 is the constant of Proposition 5.2. Appealing to Theorem 3.8 ii) and the monotonicity of $\pi_n(p)$ in n, we therefore obtain the existence of a constant $k_1 \ge 1$ with $\frac{1}{\sigma_1} \le k_1 < \infty$, such that

$$\sum_{x \in B_n(\mathbf{0})} \tau(\mathbf{0}, x; p) \ge s(m) \frac{C_4}{2} \pi_m(p_c) \ge \frac{C_4}{2} m^d [\pi_n(p_c)]^2 \tag{5.24}$$

if $mk_1 \le n$ and $n \le L_0(p)$. Choosing $m = \lfloor n/k_1 \rfloor$, we then obtain (5.21).

The bound (5.22) is immediate from (4.18), Assumption (IV), Lemma 4.4 and Assumption (II). $\square$

Proposition 5.4. *Under Assumption (II), there exists a constant $C_1 < \infty$ such that*

$$\chi(p) \le C_1 L_0(p)^d [\pi_{L_0(p)}(p_c)]^2 \qquad \text{if} \qquad p < p_c\,. \tag{5.25}$$

Under Assumptions (I), (II) and (IV), there exists a constant $C_2 > 0$ such that

$$\chi(p) \ge C_2 L_0(p)^d [\pi_{L_0(p)}(p_c)]^2 \qquad \text{if} \qquad p < p_c. \tag{5.26}$$

Under Assumptions (I), (II) and (III), there exists a constant $C_3 > 0$ such that

$$\chi^{\text{fin}}(p) \ge C_3 L_0(p)^d [\pi_{L_0(p)}(p_c)]^2 \qquad \text{if} \qquad p > p_c. \tag{5.27}$$

Proof. First let $p < p_c$. Rewriting $\chi(p)$ as

$$\chi(p) = \sum_{x \in \mathbb{Z}^d} \tau(\mathbf{0}, x; p)\,, \tag{5.28}$$

we use (4.18) and (3.32) to estimate

$$\sum_{\substack{x \in \mathbb{Z}^d: \\ |x| \ge 2L_0(p)}} \tau(\mathbf{0}, x; p) \le C_3 L_0(p)^d [\pi_{L_0(p)}(p)]^2 \le C_3 L_0(p)^d [\pi_{L_0(p)}(p_c)]^2\,, \tag{5.29}$$

and the bounds (4.18) and (4.16) to estimate

$$\sum_{\substack{x \in \mathbb{Z}^d: \\ |x| \le 2L_0(p)}} \tau(\mathbf{0}, x; p) \le \sum_{\substack{x \in \mathbb{Z}^d: \\ |x| \le 2L_0(p)}} \tau(\mathbf{0}, x; p_c) \le C_4 L_0(p)^d [\pi_{L_0(p)}(p_c)]^2\,. \tag{5.30}$$

Combining (5.28), (5.29) and (5.30) we get (5.25). With σ_1 as in Proposition 5.2, the bound (5.26), on the other hand, follows from

$$\chi(p) \ge s(\lfloor \sigma_1 L_0(p) \rfloor) P_{\ge s(\lfloor \sigma_1 L_0(p) \rfloor)}(p) \ge C_6 s(\lfloor \sigma_1 L_0(p) \rfloor) \pi_{\lfloor \sigma_1 L_0(p) \rfloor}(p_c) \text{ (see (5.11))}, \tag{5.31}$$

and the fact that π_n is decreasing in n.

Now take $p > p_c$. Analogously to (5.31) we have for any n and $C_7 \ge 0$

$$\begin{aligned} \chi^{\text{fin}}(p) &\ge E_p\Big\{ |\mathcal{C}_{B_n}(\mathbf{0})| I[\partial B_n \not\leftrightarrow \partial B_{3n}] \Big\} \\ &= E_p\Big\{ |\mathcal{C}_{B_n}(\mathbf{0})| \Big\} Pr_p\{\partial B_n \not\leftrightarrow \partial B_{3n}\} \\ &\ge C_7 s(L_0(p)) Pr_p\{ |\mathcal{C}_{B_n}(\mathbf{0})| \ge C_7 s(L_0(p))\} Pr_p\{\partial B_n \not\leftrightarrow \partial B_{3n}\}\,. \end{aligned} \tag{5.32}$$

Using Assumption (III), in the same way as we used Assumption (I) in the proof of (4.11), we now have for $n \le \frac{1}{2} L_0(p)$

$$Pr_p\{\partial B_n \not\leftrightarrow \partial B_{3n}\} \ge \widetilde{\varepsilon}^{2d}. \tag{5.33}$$

Finally, as we basically saw already in (5.13), (5.15) and (5.16), under Assumptions (I) and (II),

$$C_3\pi_r(p_c) \le \frac{1}{|B_r|}\sum_{v\in B_r} P_{p_c}\Big\{|\mathcal{C}_{B_r}(v)| \ge \tfrac{1}{2}\widetilde{C}_1 s(r)\Big\}$$
$$\le \sup_{v\in B_r} P_p\Big\{|\mathcal{C}_{B_r}(v)| \ge \tfrac{1}{2}\widetilde{C}_1 s(r)\Big\}, \quad r \ge 1 .$$

Therefore, if we take $\widetilde{C}_1$ and C_3 as in (5.15) and (5.16), then for some $w_0 = w_0(r) \in B_r$, we have

$$Pr_p\Big\{|\mathcal{C}_{B_r}(w_0)| \ge \tfrac{1}{2}\widetilde{C}_1 s(r)\Big\} \ge C_3\pi_r(p_c)\,. \tag{5.34}$$

Finally, we take $r = \lfloor n/2\rfloor$ for $n = \lfloor \frac{1}{2}L_0(p)\rfloor$. Then for $w_0 \in B_r$, $B_r - w_0 \subset B_{2r} \subset B_n$, so that $|\mathcal{C}_{B_n}(\mathbf{0})|$ is stochastically larger than $|\mathcal{C}_{B_r}(w_0) - w_0| = |\mathcal{C}_{B_r}(w_0)|$. If we now take $C_7 > 0$ so small that $C_7 s(n) \le \frac{1}{2}\widetilde{C}_1 s(r)$, then we find

$$Pr_p\Big\{|\mathcal{C}_{B_n}(\mathbf{0})| \ge C_7 s(n)\Big\} \ge Pr_p\Big\{|\mathcal{C}_{B_n}(\mathbf{0})| \ge \tfrac{1}{2}\widetilde{C}_1 s(r)\Big\}$$
$$\ge Pr_p\Big\{|\mathcal{C}_{B_r}(w_0)| \ge \tfrac{1}{2}\widetilde{C}_1 s(r)\Big\} \ge C_3\pi_r(p_c) \ge C_3\pi_{L_0(p)}(p_c). \tag{5.35}$$

The inequality (5.27) follows by combining (5.32), (5.33) and (5.35). □

We close this section with the

Proof of Theorem 3.3.

Assume that ρ exists. By Proposition 4.3, this implies Assumption (II) and the bound $1/\rho \le \frac{d-1}{2}$ — see Remark (iii) in Section 3. Assuming furthermore that $E_{p_c}(\mathcal{N}_{L,3L})$ grows more slowly than any power of L, the bound (5.1) and the existence of ρ then imply that for each $\varepsilon > 0$ there exists a constant $C_1(\varepsilon)$ such that

$$E_{p_c}\{W^{(1)}_{B_{3n}}\} \ge 2C_1(\varepsilon) n^{d-1/\rho-\varepsilon}\,. \tag{5.36}$$

Note that the bound $1/\rho \le \frac{d-1}{2}$ implies that for all sufficiently small ε the right hand side of (5.36) is monotone increasing in n. Combining (5.36) with the bound (5.13), we get the existence of a constant $\widetilde{C}_1(\varepsilon) > 0$ such that

$$P_{\ge C_1(\varepsilon)n^{d-1/\rho-\varepsilon}}(p_c) \ge C_1(\varepsilon)\frac{n^{d-1/\rho-\varepsilon}}{(6n+1)^d} \ge \widetilde{C}_1(\varepsilon) n^{-1/\rho-\varepsilon}\,. \tag{5.37}$$

Assuming finally the existence of δ in the sense of (2.28), this shows that for all $\widetilde{\varepsilon} > 0$

$$\frac{1}{\delta}\left(d - \frac{1}{\rho}\right) \le \frac{1}{\rho} + \widetilde{\varepsilon}, \tag{5.38}$$

which implies $\delta \ge d\rho - 1$. Combined with the bound (3.5), this proves the hyperscaling relation (3.11).

In order to prove (3.12), we first note that, by (5.23), for all $n, m \geq 1$,

$$\sum_{x \in B_n(\mathbf{0})} \tau(\mathbf{0}, x; p_c) \geq s(m) \left(P_{\geq s(m)}(p_c) - \pi_n(p_c) \right) . \tag{5.39}$$

It is now easy to see that for all $\widetilde{\varepsilon} > 0$ there is a constant $C_3(\widetilde{\varepsilon}) < \infty$ such that

$$\sum_{x \in B_n(\mathbf{0})} \tau(\mathbf{0}, x; p_c) \geq C_3(\widetilde{\varepsilon}) n^{d-2/\rho-\widetilde{\varepsilon}} . \tag{5.40}$$

Indeed, this follows immediately from (5.39), the existence of ρ and δ, and the just proven relation $d\rho = \delta + 1$, by choosing $m = n^{1-\varepsilon}$, with $\varepsilon = \varepsilon(\widetilde{\varepsilon})$ sufficiently small. The existence of η, on the other hand, implies the existence of a constant $C_4(\widetilde{\varepsilon}) < \infty$ such that

$$\tau(\mathbf{0}, x; p_c) \leq C_4(\widetilde{\varepsilon}) |x|^{2-d-\eta+\widetilde{\varepsilon}} . \tag{5.41}$$

We claim that (5.40) and (5.41) imply that $d - 2 + \eta \leq 2/\rho$. Indeed, assume that $d - 2 + \eta > 2/\rho$. Then $d - 2 + \eta \geq 2/\rho + 3\widetilde{\varepsilon}$ for all sufficiently small $\widetilde{\varepsilon}$. As a consequence,

$$\tau(\mathbf{0}, x; p_c) \leq C_4(\widetilde{\varepsilon}) |x|^{-2/\rho-2\widetilde{\varepsilon}} . \tag{5.42}$$

Since $2/\rho \leq d-1$, this implies that for all sufficiently small $\widetilde{\varepsilon}$ there exists a constant $\widetilde{C}_4 = \widetilde{C}_4(d, \varepsilon) < \infty$ such that

$$\sum_{x \in B_n(\mathbf{0})} \tau(\mathbf{0}, x; p_c) \leq \widetilde{C}_4 n^{d-2/\rho-2\widetilde{\varepsilon}} , \tag{5.43}$$

in contradiction with (5.40). We thus have shown that $d - 2 + \eta \leq 2/\rho$. Combined with (3.5), this completes the proof of (3.12). $\square$

6. A General Moment Estimate

In this section we prove a fundamental moment estimate and an exponential tail estimate for cluster sizes. In order to prove these estimates, we first bound the moments of the number of vertices in a large cube Λ which are connected to the boundary of a cube which is twice as large. We then show how the tail and moments of the largest cluster in Λ can be bounded in terms of such quantities. For $d = 2$ a faster way to obtain such estimates was given in [Ngu88], but his method does not seem usable when $d > 2$. While our results can be easily stated in terms of the cubes B_n introduced in Section 2, it turns out to be more conveniant to express them in terms of the cubes

$$\Lambda_n = \{ x \in \mathbb{Z}^d \mid -n \leq x_i < n \text{ for } i = 1, \dots, n \} . \tag{6.1}$$

Lemma 6.1. *Define*

$$V(L) := \textit{number of sites in } \Lambda_L \textit{ connected to } \partial\Lambda_{2L}. \tag{6.2}$$

Under Assumption (II) there are constants C_i such that for all integers $k \geq 1$,

$$E_p\big\{V^k(L)\big\} \leq C_1 k! \big(C_2 L^d \pi_L(p_c)\big)^k \quad \textit{if} \quad p \leq p_c \quad \textit{and} \quad L \geq 1\,. \tag{6.3}$$

Consequently, for $0 \leq t < [C_2 L^d \pi_L(p_c)]^{-1}$,

$$E_p\big\{\exp(tV(L))\big\} \leq C_1[1 - tC_2 L^d \pi_L(p_c)]^{-1} \quad \textit{if} \quad p \leq p_c \quad \textit{and} \quad L \geq 1\,. \tag{6.4}$$

When Assumptions (II) and (IV) hold, then (6.3) and (6.4) remain valid for $p > p_c$ and $L \leq L_0(p)$.

Proof. We write Λ for Λ_L and $\widetilde{\Lambda}$ for Λ_{2L}. Now

$$E_p\big\{V^k(L)\big\} = \sum_{v_1,\dots,v_k \in \Lambda} Pr_p\{v_i \leftrightarrow \partial\widetilde{\Lambda}, 1 \leq i \leq k\}. \tag{6.5}$$

Fix $v_1, \dots, v_k \in \Lambda$ and define

$$d(i,j) = |v_i - v_j|_\infty \text{ if } 1 \leq i, j \leq k.$$

Also define for $1 \leq i \leq k$

$$n(i) = \min\{\lfloor \frac{1}{4} d(i,j) \rfloor : 1 \leq j \leq k, j \neq i\}.$$

This $n(i)$ is essentially $1/4$ times the distance from v_i to the nearest point. We then define the cubes

$$G(i) = \prod_{r=1}^{d} [v_{i,r} - n(i), v_{i,r} + n(i)] = B_{n(i)}(v_i) \text{ (see (2.9))},$$

where $v_{i,r}$ denotes the r-th coordinate of v_i. If the points $v_1, \dots, v_k$ are pairwise distinct, these cubes are disjoint, because for some r

$$|v_{i,r} - v_{j,r}| = |v_i - v_j|_\infty \geq 4[n(i) \vee n(j)] > n(i) + n(j).$$

Also, for any j

$$4n(i) \leq d(i,j) \leq 2L, \tag{6.6}$$

so that

$$G(i) \subset \widetilde{\Lambda} \setminus \partial\widetilde{\Lambda}, 1 \leq i \leq k.$$

Consequently, if $v_i \leftrightarrow \partial\widetilde{\Lambda}$, then v_i is connected to a point outside $G(i)$, and therefore $v_i \leftrightarrow \partial G(i)$. It follows that

$$\begin{aligned} &Pr_p\{v_i \leftrightarrow \partial\widetilde{\Lambda}, 1 \le i \le k\} \le Pr_p\{v_i \leftrightarrow \partial G(i), 1 \le i \le k\} \\ &= \prod_{i=1}^k Pr_p\{v_i \leftrightarrow \partial G(i)\} \text{ (because the } G(i) \text{ are disjoint)} \\ &= \prod_{i=1}^k \pi_{n(i)}(p) \le C_3^k \prod_{i=1}^k \pi_{n(i)}(p_c). \end{aligned} \tag{6.7}$$

provided the points $v_1, \dots, v_k$ are pairwise distinct. Here C_3 is some finite constant which may be taken equal to 1 when $p \le p_c$ (by obvious monotonicity in p), and may be taken equal to D_3 when $p > p_c, L \le L_0(p)$, by Assumption (IV) (recall (6.6)). Note that (6.7) remains true if some of the v_i's coincide, since in this case the corresponding $n(i)$'s are zero, so that $\pi_{n(i)}(p) = 1 = \pi_{n(i)}(p)^{s_i}$ if the point v_i appears $s_i \ge 2$ times.

In order to sum over the points $v_1, \dots, v_k$, we assign to each set of points $v_1, \dots, v_k$ certain labeled, rooted trees $T_1, \dots, T_\tau$. We then first sum over all $v_1, \dots, v_k$ keeping the set of trees $T_1, \dots, T_\tau$ fixed, and then sum over the trees $T_1, \dots, T_\tau$. To this end, we inductively choose subsets I_r of $\{1, \dots, k\}$, and labeled rooted trees T_r on I_r. The vertices of T_r, will be denoted by $i_j^{(r)}$, and $i_1^{(r)}$ will be the root. The edges or bonds of T_r will be denoted by $b_j^{(r)}$. These trees will be chosen such that the following properties hold:

i) If ℓ is a child of j (so that (j, ℓ) is a bond in T_r), then $n(\ell) = \lfloor \frac{1}{4} d(\ell, j) \rfloor$.

ii) There exists a child $i_2^{(r)} \in I_r$ of $i_1^{(r)}$ in T_r such that $n(i_2^{(r)}) = n(i_1^{(r)})$.

iii) $\{1, \dots, k\}$ is the disjoint union of $I_1, \dots, I_\tau$.

To obtain these trees we slightly vary the construction in [Kes86] pp. 389-390. Assume that $I_1, \dots, I_{r-1}$ and $T_1, \dots, T_{r-1}$ have already been chosen and set $J_{r-1} = \cup_{1 \le s \le r-1} I_s \quad (J_0 = \emptyset)$. Assume that these sets have been chosen so that $I_1, \dots, I_{r-1}$ are disjoint and such that

$$\text{for all } i \notin J_{r-1}, \text{ and } j \in J_{r-1}, \quad \lfloor \frac{1}{4} d(i,j) \rfloor > n(i). \tag{6.8}$$

For $r = 1$ these properties are vacuous. In the first step, we choose $i_1^{(r)}, i_2^{(r)} \notin J_{r-1}$ so that

$$n(i_1^{(r)}) = \min\{n(i) : i \notin J_{r-1}\}. \tag{6.9}$$

and

$$n(i_2^{(r)}) = \lfloor \frac{1}{4} d(i_1^{(r)}, i_2^{(r)}) \rfloor = n(i_1^{(r)}). \tag{6.10}$$

Such a choice is always possible, since by definition $n(i_1^{(r)}) = \lfloor \frac{1}{4} d(i_1^{(r)}, j) \rfloor$ for some $j \in \{1, \dots, k\}$, which, by (6.8), necessarily must lie in $\{1, \dots, k\} \setminus J_{r-1}$. We also set $b_1^{(r)} = (i_1^{(r)}, i_2^{(r)})$.

Assume next that the vertices $i_1^{(r)}, \dots, i_t^{(r)}$ have been chosen so that they are distinct and are all outside J_{r-1}, and that the bonds $b_1^{(r)}, \dots, b_{t-1}^{(r)}$ have been chosen so that they form a tree on $i_1^{(r)}, \dots, i_t^{(r)}$ obeying the condition i) above. We then check whether there exists a pair (ℓ, j) so that

$$\ell \notin J_{r-1} \cup \{i_1^{(r)}, \dots, i_t^{(r)}\},\ j \in \{i_1^{(r)}, \dots, i_t^{(r)}\} \tag{6.11}$$

and

$$n(\ell) = \lfloor \frac{1}{4} |v_\ell - v_j|_\infty \rfloor. \tag{6.12}$$

If no such pair exists then we stop and take T_r as the tree on

$$I_r = \{i_1^{(r)}, \dots, i_t^{(r)}\}$$

which has bonds $b_1^{(r)}, \dots, b_{t-1}^{(r)}$. If there is a pair (ℓ, j) which satisfies (6.11) and (6.12), then we set $i_{t+1}^{(r)} = \ell$ and add the bond $b_t^{(r)} = (j, i_{t+1}^{(r)})$. In other words, ℓ is added as a child of j. We then repeat this process with t replaced by $t+1$ and search for $i_{t+2}^{(r)}$ and $b_{t+1}^{(r)}$, and so on, until for some u no further (ℓ, j) satisfying (6.11) and (6.12) with t replaced by u can be found. Then we take $I_r = \{i_1^{(r)}, \dots, i_u^{(r)}\}$ (here u depends on r but we usually do not indicate this dependence explicitly). The fact that there does not exist a further ℓ which satisfies (6.11) and (6.12) means that (6.8) now also holds with $r-1$ replaced by r. Also, by construction (see (6.11)), I_r is disjoint from $J_{r-1} = \cup_1^{r-1} I_s$, and I_r and T_r obey conditions i) and ii) above. After I_r has been chosen we go on to choose I_{r+1} and T_{r+1} etc., until the whole index set $\{1, \dots, k\}$ has been exhausted. Let us assume that that happens with I_τ, so that

$$J_\tau = \cup_{s=1}^{\tau} I_s = \{1, \dots, k\},$$

which is just condition iii).

To estimate (6.5) we must next sum the right hand side of (6.7) over all possible choices for $v_1, \dots, v_k$. Rewriting

$$\prod_{i=1}^{k} \pi_{n(i)}(p_c) = \prod_{r=1}^{\tau} \prod_{s} \pi_{n(i_s^{(r)})}(p_c), \tag{6.13}$$

we first sum $\prod_s \pi(n(i_s^{(r)}))$ over all choices of $v(i_s^{(r)}), 1 \le s \le u$, while $I_r = \{i_1^{(r)}, \dots, i_u^{(r)}\}$ and T_r are held fixed. (For typographical convenience we sometimes write $\pi(n)$ instead of $\pi_n(p_c)$ and $v(i)$ instead of v_i). To do so, let us bound in how many ways $v(i_1^{(r)}), \dots, v(i_u^{(r)})$ can be chosen when we keep T_r and $n(i_1^{(r)}), \dots, n(i_u^{(r)})$ fixed. For $v(i_1^{(r)})$ we merely use the restriction that $v(i_1^{(r)}) \in \Lambda$. This allows at most $(2L)^d$ possible choices for $v(i_1^{(r)})$. Since all other points $\ell \in I_r$ have a parent $j \in T_r$, which by i) implies that $\lfloor \frac{1}{4} |v(\ell) - v(j)|_\infty \rfloor = n(\ell)$, the points $v(\ell) \in \Lambda$ for $\ell \neq i_1^{(r)}$ have to be chosen at distance $4n(\ell) + \theta$ for some $0 \le \theta \le 3$ from the point $v(j)$, where j is the parent of ℓ in T_r. This allows at most $C_4[n(\ell)+1]^{d-1}$ choices for $v(\ell)$. Altogether, for fixed I_r, T_r and $n(i_j^{(r)})$, there are at most

$$(2L)^d \prod_{s=2}^{u} \{C_4[n(i_s^{(r)})+1]^{d-1}\} \tag{6.14}$$

choices for $v(i_1^{(r)}),\dots,v(i_u^{(r)})$.

We now let also the $n(i_s^{(r)})$ vary in accordance with the conditions i) through iii), keeping merely the set I_r and the tree T_r with root $i_1^{(r)}$ fixed. Recalling in particular the condition ii), we then bound the sum of $\prod_{s=1}^{u} \pi(n(i_s^{(r)}))$ over all possible choices of $v(i_1^{(r)}),\dots,v(i_u^{(r)})$ by

$$\begin{aligned}
(2L)^d C_4^{u-1} &\sum_{\substack{m(1),m(2),\dots,m(u)\\ \exists i\neq 1: m(i)=m(1)}} \pi_{m(1)}(p_c) \prod_{s=2}^{u} [(m(s)+1)^{d-1}\pi_{m(s)}(p_c)] \\
&\le (2L)^d C_4^{u-1}(u-1) \sum_{m(1)=0}^{2L} (m(1)+1)^{d-1}\pi^2_{m(1)}(p_c) \\
&\qquad\times \prod_{s=3}^{u} \sum_{m(s)=0}^{2L} (m(s)+1)^{d-1}\pi_{m(s)}(p_c) \\
&\le (2L)^d (2C_4)^{u-1}(C_5L)^{d(u-1)}\pi_L^u(p_c) \text{ (by Lemma 4.4)} \\
&\le C_6(2C_4C_5^d)^{u-1}(L^d\pi_L(p_c))^u \le [C_7L^d\pi_L(p_c)]^u.
\end{aligned} \tag{6.15}$$

By Cayley's formula, the number of labeled trees T_r on I_r is equal to u^{u-2} (see Theorem 2.1 in [Moo70]). Furthermore, there are u different choices for the root $i_1^{(r)} \in I_r$, and $u \cdot u^{u-2} \le u!e^u$. The bound (6.15) therefore shows that the sum of the right hand side of (6.7) over all choices of $v_1,\dots,v_k$ consistent with a given collection $I_1,\dots,I_\tau$ of sizes $u_1,\dots,u_\tau$ with

$$\sum_1^{\tau} u_r = k, \tag{6.16}$$

is at most

$$[C_7eL^d\pi_L(p_c)]^k \prod_{r=1}^{\tau} u_r!\,. \tag{6.17}$$

Finally, to prove (6.3), we must sum this estimate over all possible partitions of $\{1,\dots,k\}$ into disjoint nonempty sets $I_1,...,I_\tau$. To this end, we first note that the sum of $\prod_{r=1}^{\tau} u_r!$ over partitions into (unordered) sets is equal to the number of partitions of $\{1,\dots,k\}$ into (ordered) sequences $I_1,...,I_\tau$. The number of such partitions, in turn, is bounded by $2^{k-1}k!$. Indeed, any such partition can be obtained by first ordering $\{1, \dots, k\}$ in one of $k!$ possible ways, say as $\{i_{\sigma(1)}, \dots, i_{\sigma(k)}\}$, and then putting τ "separation marks" between some of the pairs $i_{\sigma(j)}, i_{\sigma(j+1)}$ of successive indices (with τ some integer ≥ 0). The ordered set I_s will then be the sequence of integers $i_{\sigma(j)}$ between the $(s-1)$-th and the s-th separation mark. Since there are at most 2^{k-1} ways to choose the locations of the separation marks, this proves our claim. It then follows that the sum of (6.7) over all possible choices of $v_1, \dots, v_k$ is at most

$$2^{k-1} k! [C_7 e L^d \pi_L(p_c)]^k,$$

which proves (6.3). (6.4) is immediate from (6.3). $\square$

When trying to bound the distribution of $W^{(1)}_{\Lambda_n}$ we may assume that n is a power of 2, since we can always replace n by the smallest power of 2 which exceeds n; this can only increase $W^{(1)}_{\Lambda_n}$. For the time being we therefore take

$$n = 2^k,$$

and for some ℓ with $0 \leq \ell \leq k$ we subdivide Λ_n into the $(2^{k-\ell})^d$ disjoint subcubes

$$D(\mathbf{j}) = D_\ell(\mathbf{j}) := \prod_{i=1}^d [j_i 2^{\ell+1}, (j_i+1)2^{\ell+1}), \quad -2^{k-\ell-1} \leq j_i < 2^{k-\ell-1},\ 1 \leq i \leq d. \tag{6.18}$$

Here $\mathbf{j} = (j_1, \dots, j_d)$. Each of these subcubes is congruent to Λ_{2^ℓ}. Now for a cluster $\mathcal{C}$ in Λ_n define

$$U_\ell = U_\ell(\mathcal{C}) = \{\mathbf{j} : \mathcal{C} \cap D_\ell(\mathbf{j}) \neq \emptyset\} \tag{6.19}$$

(this is the collection of (indices of) the blocks $D_\ell(\mathbf{j})$ which contain a point of $\mathcal{C}$). Also define

$$\operatorname{diam}(\mathcal{C}) = \max_{v,w \in \mathcal{C}} |v - w|_\infty. \tag{6.20}$$

Now if

$$\operatorname{diam}(\mathcal{C}) > 2^{\ell+2}, \tag{6.21}$$

then any $v \in \mathcal{C} \cap D_\ell(\mathbf{j})$ is connected in $\mathcal{C} \subset \Lambda_n$ to some vertex outside

$$\prod_{i=1}^d [j_i 2^{\ell+1} - 2^\ell, (j_i+1)2^{\ell+1} + 2^\ell)$$

and hence also to

$$\partial \prod_{i=1}^d [j_i 2^{\ell+1} - 2^\ell, (j_i+1)2^{\ell+1} + 2^\ell).$$

Consequently, if

$$V(2^\ell, \mathbf{j}) := \text{number of vertices in } D_\ell(\mathbf{j}) \text{ connected to}$$
$$\partial \prod_{i=1}^d [j_i 2^{\ell+1} - 2^\ell, (j_i+1)2^{\ell+1} + 2^\ell), \tag{6.22}$$

then (6.21) implies

$$|\mathcal{C} \cap D_\ell(\mathbf{j})| \le V(2^\ell, \mathbf{j}).$$

and

$$|\mathcal{C}| \le \sum_{\mathbf{j} \in U_\ell(\mathcal{C})} V(2^\ell, \mathbf{j}). \tag{6.23}$$

Note that each of the random variables $V(2^\ell, \mathbf{j})$ has the same distribution as the $V(2^\ell)$ of (6.2). Moreover a collection of the random variables $V(2^\ell, \mathbf{j}), \mathbf{j} \in \Gamma$, is independent if $|\mathbf{j}' - \mathbf{j}''|_\infty > 1$ for all $\mathbf{j}', \mathbf{j}'' \in \Gamma, \mathbf{j}' \neq \mathbf{j}''$. In particular, if Γ is any subset of $\mathbb{Z}^d$, and if we set for $\eta_i = 0, 1$ for $1 \le i \le d$,

$$\Gamma(\eta) = \{\mathbf{j} \in \Gamma : j_i \equiv \eta_i (\text{mod } 2)\}, \tag{6.24}$$

then the random variables

$$\{V(2^\ell, \mathbf{j}) : \mathbf{j} \in \Gamma(\eta)\}$$

are i.i.d. (for each choice of η). This independence property quickly leads to the following lemma, which, roughly speaking, gives an exponential bound for the tail of $[s(2^\ell)|U_\ell(\mathcal{C})|]^{-1}|\mathcal{C}|$.

Lemma 6.2. *If Assumption (II) holds and $p \le p_c$, or if Assumptions (II) and (IV) hold, $p > p_c$ and $1 \le 2^\ell \le L_0(p)$, then there exist constants C_i such that*

$$\begin{aligned} &Pr_p\Big\{\exists \mathcal{C} \subset \Lambda_n \textit{ which satisfies } (6.21) \textit{ and } |\mathcal{C}| \ge xs(2^\ell) \textit{ but } |U_\ell(\mathcal{C})| \le r\Big\} \\ &\le C_1 \left(\frac{n}{2^\ell}\right)^d C_2^r e^{-C_3 x}, \; x \ge 0, r \ge 1, n \ge 1. \end{aligned} \tag{6.25}$$

Proof. If $n \le 2^\ell$, then any $\mathcal{C} \subset \Lambda_n$ has $\text{diam}(\mathcal{C}) \le 2n \le 2^{\ell+1}$ and the probability in the left hand side of (6.25) is zero. We may therefore assume that $n \ge 2^\ell$. In fact we may, and shall restrict ourselves to $n = 2^k$ for some $k \ge \ell$.

Note now that if $\mathcal{C}$ is a cluster in Λ_n, then $U_\ell(\mathcal{C})$ is a connected subset of

$$\mathbb{Z}^d \cap \prod_{i=1}^d [-2^{k-\ell-1}, 2^{k-\ell-1}).$$

Therefore, by (6.23), the probability in the left hand side of (6.25) is at most

$$\sum_\Gamma Pr_p\{\sum_{\mathbf{j} \in \Gamma} V(2^\ell, \mathbf{j}) \ge xs(2^\ell)\}, \tag{6.26}$$

where the sum is over all connected sets $\Gamma \subset \prod[-2^{k-\ell-1}, 2^{k-\ell-1})$ with $|\Gamma| \le r$. Now, with $\Gamma(\eta)$ as in (6.24), Γ is the disjoint union of 2^d sets $\Gamma(\eta)$, so that for fixed

Γ and all $0 \le t < [C_2 s(2^\ell)]^{-1}$,

$$\begin{aligned}
&Pr_p\{\sum_{\mathbf{j}\in\Gamma} V(2^\ell, \mathbf{j})\} \ge xs(2^\ell)\} \\
&\le \sum_\eta Pr_p\{\sum_{\mathbf{j}\in\Gamma(\eta)} V(2^\ell,\mathbf{j}) \ge 2^{-d} xs(2^\ell)\} \\
&\le \sum_\eta \exp[-t2^{-d}xs(2^\ell)] E_p \exp\left[t \sum_{\mathbf{j}\in\Gamma(\eta)} V(2^\ell,\mathbf{j})\right] \\
&\le 2^d \exp[-t2^{-d}xs(2^\ell)] \left[\frac{C_1}{1-tC_2 s(2^\ell)}\right]^{|\Gamma|} \text{ (by (6.4)).} \qquad (6.27)
\end{aligned}$$

We now take $t = [2C_2 s(2^\ell)]^{-1}$. Then the right hand side of (6.27) is at most

$$2^d \exp[-t2^{-d}xs(2^\ell)][2C_1]^{|\Gamma|} \le 2^d \exp[-2^{-d-1}x/C_2][2C_1]^r.$$

We substitute this estimate into (6.26). We further use the fact that the number of connected subsets of $\mathbb{Z}^d$ of size $\le r$ which contain a given vertex is at most C_3^r for some constant $C_3 = C_3(d)$ (see for instance [Kes82], equation (5.22)). Thus the number of permissible choices for Γ is at most

$$2^{d(k-\ell)} C_3^r,$$

and the probability in the left hand side of (6.25) is at most

$$2^{d(k-\ell+1)} C_3^r \exp[-2^{-d-1}x/C_2][2C_1]^r,$$

so that (6.25) follows. $\square$

We are ready to prove the principal result of this section.

Proposition 6.3.
i) Under Assumption (II) there exist constants C_i such that for all $p \le p_c$, $x \ge 0$ and $n \le L_0(p)$,

$$Pr_p\left\{W^{(1)}_{\Lambda_n} \ge xs(n)\right\} \le C_1 e^{-C_2 x}. \qquad (6.28)$$

If Assumptions (II) and (IV) hold, then (6.28) remains valid for $p > p_c$ (and $n \le L_0(p)$, $x \ge 0$) as well.

ii) Under Assumption (II) there exist constants C_i such that for all $p < p_c, n \ge L_0(p)$ and $x \ge 0$,

$$Pr_p\left\{W^{(1)}_{\Lambda_n} \ge xs(L_0(p))\right\} \le C_1 \left(\frac{n}{L_0(p)}\right)^d e^{-C_2 x}. \qquad (6.29)$$

In particular, for $y > d/C_2$ and any sequence of densities p_n with $L_0(p_n)/n \to 0$,

$$Pr_{p_n}\left\{W^{(1)}_{\Lambda_n} \ge ys(L_0(p_n)) \log\left(\frac{n}{L_0(p_n)}\right)\right\} \to 0 \qquad (6.30)$$

as $n \to \infty$.

iii) Under Assumptions (II) and (IV) there exist constants C_i *such that for all* $p > p_c, n \geq L_0(p)$ *and* $x \geq 0$,

$$Pr_p\left\{W^{(1)}_{\Lambda_n} \geq xs(L_0(p))\right\} \leq C_1\left(\frac{n}{L_0(p)}\right)^d \exp\Big[-C_2 x + C_3\left(\frac{n}{L_0(p)}\right)^d\Big]. \tag{6.31}$$

Proof. There is nothing to prove if $x < 1$, so that we shall assume that $x \geq 1$. In a similar way, we may assume without loss of generality that $xs(n \wedge L_0(p))$ is large enough to guaranty that $|\mathcal{C}| \geq xs(n \wedge L_0(p))$ implies that $\mathrm{diam}(\mathcal{C}) > 2^{\ell+2}$ for $\ell = 0$. Therefore, by a decomposition according to the diameter of the largest cluster in Λ_n,

$$\begin{aligned} &Pr_p\left\{W^{(1)}_{\Lambda_n} \geq xs(n \wedge L_0(p))\right\} \\ &\leq \sum_{\ell \geq 0: 2^\ell \leq L_0(p)} Pr_p\{\exists \text{ cluster } \mathcal{C} \text{ in } \Lambda_n \text{ with} \\ &\qquad 2^{\ell+2} < \mathrm{diam}(\mathcal{C}) \leq 2^{\ell+3} \text{ and } |\mathcal{C}| \geq xs(n \wedge L_0(p))\} \\ &+ Pr_p\{\exists \text{ cluster } \mathcal{C} \text{ in } \Lambda_n \text{ with} \\ &\qquad \mathrm{diam}(\mathcal{C}) > 4L_0(p) \text{ and } |\mathcal{C}| \geq xs(n \wedge L_0(p))\}. \end{aligned} \tag{6.32}$$

We first estimate the sum in the right hand side. If

$$2^{\ell+2} < \mathrm{diam}(\mathcal{C}) \leq 2^{\ell+3}$$

and $\mathcal{C} \cap D_\ell(\mathbf{j}) \neq \emptyset$ for some $\mathbf{j}$, then $\mathcal{C}$ is contained in

$$\bigcup_{|\mathbf{p}-\mathbf{j}|_\infty \leq 4} D_\ell(\mathbf{p}),$$

and by (6.23)

$$|\mathcal{C}| \leq \sum_{|\mathbf{p}-\mathbf{j}|_\infty \leq 4} V(2^\ell, \mathbf{p}).$$

In particular, $|U_\ell| \leq 9^d$ and by Lemma 6.2 with $r = 9^d$,

$$\begin{aligned} &Pr_p\{\exists \text{ cluster } \mathcal{C} \text{ in } \Lambda_n \text{ with } 2^{\ell+2} < \mathrm{diam}(\mathcal{C}) \leq 2^{\ell+3} \text{ and } |\mathcal{C}| \geq xs(n \wedge L_0(p))\} \\ &\leq C_2\left(\frac{n}{2^\ell}\right)^d \exp\Big[-C_3 x\frac{s(n \wedge L_0(p))}{s(2^\ell)}\Big]. \end{aligned} \tag{6.33}$$

Now let us first consider the case of part i). Then $n \leq L_0(p)$ and as in the beginning of the proof of Lemma 6.2 we can restrict the sum in the right hand side of (6.32) to ℓ with $2^\ell \le n$, and also the last term in (6.32) vanishes. Therefore the right hand side of (6.32) is at most

$$\sum_{2^\ell \le n} Pr_p\{\exists \text{ cluster } \mathcal{C} \text{ in } \Lambda_n \text{ with}$$
$$2^{\ell+2} < \operatorname{diam}(\mathcal{C}) \le 2^{\ell+3} \text{ and } |\mathcal{C}| \ge xs(n)\}$$
$$\le \sum_{2^\ell \le n} C_2 \left(\frac{n}{2^\ell}\right)^d \exp\Big[-C_3 x \frac{s(n)}{s(2^\ell)}\Big]. \tag{6.34}$$

By Assumption (II),

$$\frac{s(n)}{s(2^\ell)} \ge D_1 \left(\frac{n}{2^\ell}\right)^{d/2}, \tag{6.35}$$

so that the right hand side of (6.32) is bounded by

$$\sum_{2^\ell \le n} C_2 \left(\frac{n}{2^\ell}\right)^d \exp\Big[-C_4 x \left(\frac{n}{2^\ell}\right)^{d/2}\Big] \le C_5 e^{-C_6 x}.$$

This proves part (i).

For parts (ii) and (iii) we have $n \wedge L_0(p) = L_0(p)$. Making the obvious changes in (6.34) we find that the sum in the right hand side of (6.32) is bounded by

$$\sum_{2^\ell \le L_0(p)} C_2 \left(\frac{n}{2^\ell}\right)^d \exp\Big[-C_4 x \left(\frac{L_0(p)}{2^\ell}\right)^{d/2}\Big]$$
$$\le C_5 \left(\frac{n}{L_0(p)}\right)^d \exp[-C_6 x]. \tag{6.36}$$

This estimate holds for any p. However, the last term in the right hand side of (6.32) has to be treated somewhat differently in the cases $p \le p_c$ and $p > p_c$. Let us first consider the latter case, that is, the case of part (iii). We now define ℓ_0 by

$$2^{\ell_0+2} \le L_0(p) < 2^{\ell_0+3}. \tag{6.37}$$

Then any cluster $\mathcal{C}$ with $\operatorname{diam}(\mathcal{C}) > L_0(p)$ satisfies (6.21) with ℓ replaced by ℓ_0. If also $\mathcal{C} \subset \Lambda_n$, with $n = 2^k$, then by definition,

$$U_{\ell_0}(\mathcal{C}) \subset [-n2^{-\ell_0-1}, n2^{-\ell_0-1})$$

and hence

$$|U_{\ell_0}(\mathcal{C})| \le \left(n2^{-\ell_0}\right)^d.$$

We therefore can bound the last term in (6.32) by (6.25) with ℓ_0 for ℓ and

$$r = \left(n2^{-\ell_0}\right)^d.$$

(6.31) follows easily from (6.32), (6.36) and this last bound.

We now turn to the estimate of the last term in (6.32) in the case of part (ii), that is when $p < p_c$. In that case we again take ℓ_0 as in (6.37), and for a constant C_7 to be determined below, bound the last term in (6.32) by

$$\begin{aligned} &Pr_p\{\exists \text{ cluster } \mathcal{C} \text{ in } \Lambda_n \text{ with } |U_{\ell_0}(\mathcal{C})| > C_7 x\} \\ &\quad + Pr_p\{\exists \text{ cluster } \mathcal{C} \text{ in } \Lambda_n \text{ with } \operatorname{diam}(\mathcal{C}) > 2^{\ell_0+2}, \\ &\qquad\qquad |U_{\ell_0}(\mathcal{C})| \le C_7 x, \text{and } |\mathcal{C}| \ge xs(L_0(p)\}. \end{aligned} \tag{6.38}$$

If ε is sufficiently small, then there exist constants $C_8 = C_8(\varepsilon, d)$ and $C_9 = C_9(\varepsilon, d)$ so that the first probability in (6.38) is at most

$$C_8 \left(\frac{n}{L_0(p)}\right)^d \exp(-C_9 C_7 x). \tag{6.39}$$

This is proven by a renormalized block argument which is given in detail in the proof of Theorem 5.1 in [Kes82]. It is based on the fact that if a cluster $\mathcal{C}$ contains a point in $D_{\ell_0}(\mathbf{j})$ as well as a point outside

$$\prod_{i=1}^d [(j_i - 4)2^{\ell_0+1}, (j_i + 8)2^{\ell_0+1}),$$

then $\mathcal{C}$ contains a crossing in the short direction of one $2d$ blocks congruent to

$$[0, 3L_0(p)] \times \cdots [0, 3L_0(p)] \times [0, L_0(p)] \times [0, 3L_0(p)] \times \cdots [0, 3L_0(p)] \tag{6.40}$$

which surround the cube

$$\prod_{i=1}^d [j_i 2^{\ell_0+1}, j_i 2^{\ell_0+1} + L_0(p)].$$

This is so because

$$\prod_{i=1}^d [j_i 2^{\ell_0+1} - L_0(p), j_i 2^{\ell_0+1} + 2L_0(p)] \setminus \prod_{i=1}^d (j_i 2^{\ell_0+1}, j_i 2^{\ell_0+1} + L_0(p))$$

is the union of $2d$ such blocks,

$$D_{\ell_0}(\mathbf{j}) \subset \prod_{i=1}^d [j_i 2^{\ell_0+1}, j_i 2^{\ell_0+1} + L_0(p)],$$

and

$$\prod_{i=1}^d [j_i 2^{\ell_0+1} - L_0(p), j_i 2^{\ell_0+1} + 2L_0(p)] \subset \prod_{i=1}^d [(j_i - 4)2^{\ell_0+1}, (j_i + 8)2^{\ell_0+1}).$$

By the definition (2.17) and the fact that $R^{\rm fin}_{L,3L}(p) = R_{L,3L}(p)$ if $p < p_c$ the probability that there exists an occupied crossing in the short direction of a block (6.40) is

$$R_{L_0(p),3L_0(p)} \le \varepsilon. \tag{6.41}$$

One can now use (6.41) and a Peierls argument to obtain (6.39). The factor

$$\left(\frac{n}{L_0(p)}\right)^d$$

in (6.39) arises because there are $(n2^{-\ell_0})^d$ blocks $D_{\ell_0}(\mathbf{j})$ in Λ_n which $\mathcal{C}$ can intersect.

Once one has (6.39), one uses (6.25) to estimate the second probability in (6.38). Together one finds that the last term in (6.32) is at most

$$C_{10}\left(\frac{n}{L_0(p)}\right)^d \left[C_8 \exp(-C_9C_7x) + C_2^{C_7x}e^{-C_3x}\right]. \tag{6.42}$$

If we choose

$$C_7 = \frac{C_3}{C_9 + \log C_2},$$

then this bound is at most $C_{11}[n/L_0(p)]^d \exp(-C_{12}x)$ with

$$C_{12} = \frac{C_9C_3}{C_9 + \log C_2}.$$

(6.29) is immediate from (6.32), (6.36) and (6.42). $\square$

Remark (xiii):. The estimates (6.33) and (6.35) also show that under Assumption (II) we have for $p \le p_c, x \ge 0, 0 < y \le 1, 1 \vee 4/y \le n \le L_0(p)$,

$$\begin{aligned} Pr_p\{\exists \text{ cluster } \mathcal{C} \subset \Lambda_n \text{ with } \operatorname{diam}(\mathcal{C}) \le yn \text{ but } |\mathcal{C}| \ge xs(n)\} \\ \le C_1y^{-d}\exp[-C_3xy^{-d/2}]. \end{aligned} \tag{6.43}$$

Indeed, (6.33) holds for $p \le p_c$ under Assumption (II) only, since it relies only on Lemma 6.2. Note also that we may assume $xy^{-d/2} \ge 1$, because otherwise (6.43) is trivial for large enough C_1. Therefore, by (6.33) and (6.35)

$$\begin{aligned} &Pr_p\{\exists \text{ cluster } \mathcal{C} \subset \Lambda_n \text{ with } \operatorname{diam}(\mathcal{C}) \le yn \text{ but } |\mathcal{C}| \ge xs(n)\} \\ &\le \sum_{\ell:2^\ell \le yn} Pr_p\{\exists \text{ cluster } \mathcal{C} \subset \Lambda_n \text{ with } 2^{\ell+2} < \operatorname{diam}(\mathcal{C}) \le 2^{\ell+3} \text{ but } |\mathcal{C}| \ge xs(n)\} \\ &\le \sum_{\ell:2^\ell \le yn} C_2\Big(\frac{n}{2^\ell}\Big)^d \exp\Big[-C_3x\frac{s(n)}{s(2^\ell)}\Big] \qquad\qquad \text{(recall } n \le L_0(p)) \\ &\le \sum_{\ell:2^\ell \le yn} C_2\Big(\frac{n}{2^\ell}\Big)^d \exp\Big[-C_3xD_1\Big(\frac{n}{2^\ell}\Big)^{d/2}\Big] \\ &\le \sum_{j\ge 0} C_2\Big(\frac{2}{y}\Big)^d 2^{jd}\exp[-C_4xy^{-d/2}2^{jd/2}] \\ &\le C_1y^{-d}\exp[-C_5xy^{-d/2}]. \end{aligned}$$

7. Exponential Decay of $P_{\geq s}(p)$

In this section we prove Theorem 3.9. We note that (3.35) is an immediate consequence of (3.34) and the bound (5.11) from Proposition 5.2. We therefore only need to prove (3.34) and we turn to this now.

Clearly, for any $0 < \alpha \leq 1$,

$$P_{\geq xs(L_0(p))}(p) \leq Pr_p\Big\{|\mathcal{C}(\mathbf{0}) \cap \Lambda_{\lfloor \alpha L_0(p) \rfloor}| \geq \frac{x}{2} s(L_0(p))\Big\} + Pr_p\Big\{|\mathcal{C}(\mathbf{0}) \cap \Lambda^c_{\lfloor \alpha L_0(p) \rfloor}| \geq \frac{x}{2} s(L_0(p))\Big\}. \tag{7.1}$$

In order to bound the first term, we introduce the "rings" or annuli

$$R_\ell := B_{2^{\ell+1}} \setminus B_{2^\ell} = [-2^{\ell+1}, 2^{\ell+1}]^d \setminus [-2^\ell, 2^\ell]^d.$$

Then, $|\mathcal{C}(\mathbf{0}) \cap \Lambda_{\lfloor \alpha L_0(p) \rfloor}| \geq \frac{x}{2} s(L_0(p))$ implies

$$\frac{x}{2} s(L_0(p)) \leq |\mathcal{C}(\mathbf{0}) \cap \Lambda_{\lfloor \alpha L_0(p) \rfloor}| \leq 5^d + \sum_{\ell : 2 \leq 2^\ell \leq \alpha L_0(p)} |\mathcal{C}(\mathbf{0}) \cap R_\ell|\,. \tag{7.2}$$

On the other hand, define k by $2^k \leq L_0(p) < 2^{k+1}$ and let

$$\beta = 2^{d+2} C_2 \frac{\log 2}{\rho_1 D_1} + 1, \tag{7.3}$$

where C_2 is the same constant as in (6.4). Then, by virtue of (5.17),

$$\begin{aligned} \beta & \sum_{\ell : 2 \leq 2^\ell \leq \alpha L_0(p)} (k - \ell) s(2^\ell) \\ & \leq \beta \frac{s(\lfloor \alpha L_0(p) \rfloor)}{D_1} \sum_{\ell \leq k+1-\log(1/\alpha)/\log 2} (k - \ell) \left(\frac{2^\ell}{\lfloor \alpha L_0(p) \rfloor} \right)^{d/2} \\ & \leq \beta C_6 \frac{s(\lfloor \alpha L_0(p) \rfloor)}{D_1} \left[1 + \frac{\log(1/\alpha)}{\log 2} \right], \end{aligned} \tag{7.4}$$

where C_6 is a constant which depends only on the dimension d. Choosing α sufficiently small, $n_0 = n_0(\alpha)$ sufficiently large, and p in such a way that $L_0(p) \geq n_0$, we therefore get

$$\beta \sum_{\ell : 2 \leq 2^\ell \leq \alpha L_0(p)} (k - \ell) s(2^\ell) < s(L_0(p)) - 2(5^d) \tag{7.5}$$

and hence (for $x \geq 1$)

$$\frac{\beta x}{2} \sum_{\ell : 2 \leq 2^\ell \leq \alpha L_0(p)} (k - \ell) s(2^\ell) < \frac{x}{2} s(L_0(p)) - 5^d. \tag{7.6}$$

We fix $0 < \alpha < 1/2$ and n_0 so that this holds for all p with $L_0(p) \geq n_0$. It then follows from (7.2) and (7.6) that if $L_0(p) \geq n_0$, there must be a smallest ℓ with

$$2 \leq 2^\ell \leq \alpha L_0(p) \qquad \text{and} \qquad |\mathcal{C}(\mathbf{0}) \cap R_\ell| > \frac{\beta x}{2}(k-\ell)s(2^\ell).$$

Therefore,

$$\begin{aligned} &Pr_p\Big\{|\mathcal{C}(\mathbf{0}) \cap \Lambda_{\lfloor \alpha L_0(p) \rfloor}| \geq \frac{x}{2}s(L_0(p))\Big\} \\ &\qquad \leq \sum_{\ell : 2 \leq 2^\ell \leq \alpha L_0(p)} Pr_p\Big\{|\mathcal{C}(\mathbf{0}) \cap R_\ell| > \frac{\beta x}{2}(k-\ell)s(2^\ell)\Big\}. \end{aligned} \tag{7.7}$$

But $|\mathcal{C}(\mathbf{0}) \cap R_\ell| > \frac{\beta x}{2}(k-\ell)s(2^\ell)$ implies $|\mathcal{C}(\mathbf{0}) \cap R_\ell| > 0$ and hence

$$\mathbf{0} \leftrightarrow \partial B(2^{\ell-1}), \tag{7.8}$$

and, in the notation of (6.18) and (6.22),

$$\begin{aligned} &\frac{\beta x}{2}(k-\ell)s(2^\ell) < |\mathcal{C}(\mathbf{0}) \cap R_\ell| \\ &\qquad \leq \text{ number of vertices in } R_\ell \text{ connected to } \partial B_{2^{\ell-1}} \\ &\qquad \leq \sum_{D_{\ell-1}(\mathbf{j}) \cap R_\ell \neq \emptyset} [\text{ number of vertices in } D_{\ell-1}(\mathbf{j}) \text{ connected to } \partial B_{2^{\ell-1}}] \\ &\qquad \leq \sum_{D_{\ell-1}(\mathbf{j}) \cap R_\ell \neq \emptyset} V(2^{\ell-1}, \mathbf{j}). \end{aligned} \tag{7.9}$$

The event in (7.8) and the sum in the right hand side of (7.9) are independent. Therefore

$$\begin{aligned} &Pr_p\Big\{|\mathcal{C}(\mathbf{0}) \cap R_\ell| > \frac{\beta x}{2}(k-\ell)s(2^\ell)\Big\} \\ &\qquad \leq Pr_p\{\mathbf{0} \leftrightarrow \partial B_{2^{\ell-1}}\} Pr_p\Big\{\sum_{D_{\ell-1}(\mathbf{j}) \cap R_\ell \neq \emptyset} V(2^{\ell-1}, \mathbf{j}) > \frac{\beta x}{2}(k-\ell)s(2^\ell)\Big\}. \end{aligned} \tag{7.10}$$

The number of $\mathbf{j}$ for which $D_{\ell-1}(\mathbf{j}) \cap R_\ell \neq \emptyset$ is bounded by some constant C_7, uniformly in ℓ. As in (6.27) we therefore obtain

$$\begin{aligned} &Pr_p\Big\{\sum_{D_{\ell-1}(\mathbf{j}) \cap R_\ell \neq \emptyset} V(2^{\ell-1}, \mathbf{j}) > \frac{\beta x}{2}(k-\ell)s(2^\ell)\Big\} \\ &\qquad \leq 2^d \exp\big[-t2^{-d}\frac{\beta x}{2}(k-\ell)s(2^\ell)\big]\left[\frac{C_1}{1 - tC_2 s(2^{\ell-1})}\right]^{C_7}, \quad t \geq 0. \end{aligned}$$

We take $t = [2C_2 s(2^{\ell-1})]^{-1}$ to obtain

$$Pr_p\Big\{ \sum_{D_{\ell-1}(\mathbf{j})\cap R_\ell \neq \emptyset} V(2^{\ell-1}, \mathbf{j}) > \frac{\beta x}{2}(k-\ell)s(2^\ell)\Big\} \leq C_8 \exp\big[-\frac{\beta D_1 x}{2^{d+2}C_2}(k-\ell)\big], \tag{7.11}$$

for all $\ell \geq 1$, because $s(2^\ell)/s(2^{\ell-1}) \geq D_1$, by virtue of Assumption (II). Substitution of (7.11) into (7.10) and use of the monotonicity of $\pi_n(p)$ in p to bound $\pi_{2^{\ell-1}}(p)$ in terms of $\pi_{2^{\ell-1}}(p_c)$ gives for $\ell \geq 1$

$$\begin{aligned} Pr_p\Big\{|\mathcal{C}(\mathbf{0})\cap R_\ell| > \frac{\beta x}{2}(k-\ell)s(2^\ell)\Big\} &\leq C_8 \pi_{2^{\ell-1}}(p_c)\exp\big[-\frac{\beta D_1 x}{2^{d+2}C_2}(k-\ell)\big] \\ &\leq \frac{C_8}{D_1}\pi_{L_0(p)}(p_c)\left(\frac{L_0(p)}{2^{\ell-1}}\right)^{1/\rho_1}\exp\big[-\frac{\beta D_1 x}{2^{d+2}C_2}(k-\ell)\big]. \end{aligned} \tag{7.12}$$

In view of (7.7), the definitions of β and k, and the fact that $\alpha < 1/2, x \geq 1$, this gives

$$\begin{aligned} &Pr_p\Big\{|\mathcal{C}(\mathbf{0})\cap \Lambda_{\lfloor \alpha L_0(p)\rfloor}| \geq \frac{x}{2}s(L_0(p))\Big\} \\ &\quad\leq \frac{C_8}{D_1}\pi_{L_0(p)}(p_c)\sum_{\ell:2\leq 2^\ell \leq \alpha L_0(p)}\left(\frac{L_0(p)}{2^{\ell-1}}\right)^{1/\rho_1}\exp\big[-\frac{\beta D_1 x}{2^{d+2}C_2}(k-\ell)\big] \\ &\quad\leq C_9 \pi_{L_0(p)}(p_c)\exp[-C_{10}x]. \end{aligned} \tag{7.13}$$

Next, define ℓ_1 by $2^{\ell_1} \leq \lfloor \alpha L_0(p)/4\rfloor < 2^{\ell_1+1}$. (Thus $|k-\ell_1-\log(4/\alpha)/\log 2| \leq 2$.) Then, as in (7.8), (7.9), the event $|\mathcal{C}(\mathbf{0})\cap\Lambda^c_{\lfloor\alpha L_0(p)\rfloor}| \geq \frac{x}{2}s(L_0(p))$ is contained in the intersection of

$$\{\mathbf{0} \leftrightarrow \partial B_{\lfloor \alpha L_0(p)/4\rfloor}\}$$

and

$$\sum_{D_{\ell_1}(\mathbf{j})\cap\mathcal{C}(\mathbf{0})\cap\Lambda^c_{\lfloor\alpha L_0(p)\rfloor}\neq\emptyset} V(2^{\ell_1}, \mathbf{j}) \geq \frac{x}{2}s(L_0(p)). \tag{7.14}$$

Analogously to (6.19) we now define

$$\widetilde{U}_\ell = \widetilde{U}_{\ell,\alpha L_0(p)} = \{\mathbf{j} : D_\ell(\mathbf{j}) \leftrightarrow \partial\Lambda_{\lfloor\alpha L_0(p)/4\rfloor},\, D_\ell(\mathbf{j})\cap\Lambda^c_{\lfloor\alpha L_0(p)\rfloor} \neq \emptyset\}.$$

Then

$$\sum_{D_{\ell_1}(\mathbf{j})\cap\mathcal{C}(\mathbf{0})\cap\Lambda^c_{\lfloor\alpha L_0(p)\rfloor}\neq\emptyset} V(2^{\ell_1}, \mathbf{j}) \leq \sum_{\mathbf{j}\in\widetilde{U}_{\ell_1}} V(2^{\ell_1}, \mathbf{j}),$$

so that for any $C_7 > 0$

$$
\begin{aligned}
&Pr_p\Big\{|\mathcal{C}(\mathbf{0}) \cap \Lambda^c_{\lfloor \alpha L_0(p)\rfloor}| \geq \frac{x}{2} s(L_0(p))\Big\} \\
&\leq Pr_p\Big\{\mathbf{0} \leftrightarrow \partial B_{\lfloor \alpha L_0(p)/4\rfloor}, \sum_{\mathbf{j}\in \widetilde{U}_{\ell_1}} V(2^{\ell_1}, \mathbf{j}) \geq \frac{x}{2} s(L_0(p))\Big\} \\
&= \pi_{\lfloor \alpha L_0(p)/4\rfloor}(p) Pr_p\Big\{\sum_{\mathbf{j}\in \widetilde{U}_{\ell_1}} V(2^{\ell_1}, \mathbf{j}) \geq \frac{x}{2} s(L_0(p))\Big\} \\
&\leq \pi_{\lfloor \alpha L_0(p)/4\rfloor}(p) Pr_p\{|\widetilde{U}_{\ell_1}| \geq C_7 x\} \\
&\qquad\qquad + \pi_{\lfloor \alpha L_0(p)/4\rfloor}(p) \sup_{\Gamma} Pr_p\Big\{\sum_{\mathbf{j}\in \Gamma} V(2^{\ell_1}, \mathbf{j}) \geq \frac{x}{2} s(L_0(p))\Big\}.
\end{aligned}
\tag{7.15}
$$

Here Γ runs over all subsets of $\mathbb{Z}^d$ with $|\Gamma| < C_7 x$. Exactly the same method as used to estimate (6.38) when $p < p_c$ can now be used to show that

$$
\begin{aligned}
&Pr_p\{|\widetilde{U}_{\ell_1}| \geq C_7 x\} + \sup_{\Gamma} Pr_p\Big\{\sum_{\mathbf{j}\in \Gamma} V(2^{\ell_1}, \mathbf{j}) \geq \frac{x}{2} s(L_0(p))\Big\} \\
&\leq C_4 \exp[-C_5 x].
\end{aligned}
\tag{7.16}
$$

For $L_0(p) \geq n_0$, the lemma now follows from (7.13), (7.15) and (7.16). On the other hand, for $L_0(p) \leq n_0$, p is bounded away from p_c, and the proof of the lemma just reduces to the proof of the well known fact (see [Gri89], Ch. 3 for references) that for $p < p_c$ the size distribution $P_{\geq s}(p)$ decays exponentially in s with some strictly positive decay constant (rather than a decay constant proportional to $1/s(L_0(p))$). □

References

[Aiz97] M. Aizenman, *On the number of incipient spanning clusters*, to appear in Nucl. Phys. B[FS] (1997).

[Ale96] K. Alexander, *Private communication* (1996).

[ACCFR83] M. Aizenman, J. T. Chayes, L. Chayes, J. Fröhlich and L. Russo, *On a sharp transition from area law to perimeter law in a system of random surfaces*, Commun. Math. Phys. **92** (1983), 19–69.

[Arc87] L. de Arcangelis, *Multiplicity of the infinite clusters in percolation above six dimensions*, J. Phys. A: Math. Gen. **20** (1987), 3057–3061.

[AS92] A. Aharony and D. Stauffer, *Introduction to Percolation Theory*, Taylor and Francis, London, 1992.

[BCKS97] C. Borgs, J. T. Chayes, H. Kesten and J. Spencer, *Birth of the infinite cluster: Finite-size scaling in percolation. Preprint* (1997).

[BK85] J. van den Berg and H. Kesten, *Inequalities with applications to percolation and reliability*, J. Appl. Probab. **22** (1985), 556–569.

[BW95] L. Berlyand and J. Wehr, *The probability distribution of the percolation threshold in a large system*, J. Phys. A: Math. Gen. **28** (1995), 7127–7133.

[BW97] L. Berlyand and J. Wehr, *Nongaussian limiting behavior of the percolation threshold in a large system*, to appear in Commun. Math. Phys. (1997).

[CC84] J. T. Chayes and L. Chayes, *Percolation and random media*, in Random Systems and Gauge Theories, Les Houches, Session XLIII (K. Osterwalder and R. Stora, eds.), Elsevier, Amsterdam, 1984, pp. 1001–1142.

[CC86] J. T. Chayes and L. Chayes, *Bulk transport properties and critical exponent inequalities for random resistor and flow networks*, Commun. Math. Phys. **105** (1986), 133–152.

[CC87] J. T. Chayes and L. Chayes, *On the upper critical dimension in Bernoulli percolation*, Commun. Math. Phys. **113** (1987), 27–48.

[CCF85] J. T. Chayes, L. Chayes and J. Fröhlich, *The low-temerature behavior of disordered magnets*, Commun. Math. Phys. **100** (1985), 399–437.

[CCFS86] J. T. Chayes, L. Chayes, D. Fisher and T. Spencer, *Finite-size scaling and correlation lengths for disordered systems*, Phys. Rev. Lett. **57** (1986), 2999–3002.

[CCGKS89] J. T. Chayes, L. Chayes, G. R. Grimmett, H. Kesten and R. H. Schonmann, *The correlation length for the high-density phase of Bernoulli percolation*, Ann. Probab. **17** (1989), 1277–1302.

[Con85] A. Coniglio, *Shapes, surfaces and interfaces in percolation clusters*, in Proc. Les Houches Conference on Physics of Finely Divided Matter (M. Daoud and N. Boccara, eds.), Springer-Verlag, 1985, pp. 84–109.

[Gri89] G. Grimmett, *Percolation*, Springer-Verlag, New York, 1989.

[GM90] G. R. Grimmett and J. M. Marstrand, *The supercritical phase of percolation is well behaved*, Proc. R. Soc. Lond. **A 430** (1990), 439–457.

[Har60] T. E. Harris, *A lower bound for the critical probability in a certain percolation process*, Proc. Cambridge Philos. Soc. **56** (1960), 13–20.

[Kes80] H. Kesten, *On the time constant and path length of first passage percolation*, Advances in Applied Probability **12** (1980), 848–863.

[Kes82] H. Kesten, *Percolation Theory for Mathematicians*, Birkhäuser, Boston, 1982.

[Kes86] H. Kesten, *The incipient infinite cluster in two-dimensional percolation*, Probab. Th. Rel. Fields **73** (1986), 369–394.

[Kes87] H. Kesten, *Scaling relations for 2D-percolation*, Commun. Math. Phys. **109** (1987), 109–156.

[KZ90] H. Kesten and Y. Zhang, *The probability of a large finite cluster in supercritical Bernoulli percolation*, Ann. Probab. **18** (1990), 537–555.

[Moo70] J. W. Moon, *Counting labelled trees*, Canadian Math. Monographs No 1, Can. Math. Congress, 1970.

[Ngu88] B. Nguyen, *Typical cluster size for two-dimensional percolation processes*, J. Stat. Phys. **50** (1988), 715–726.

[Pis96] A. Pisztora, *Surface order large deviations for Ising, Potts and percolation models.*, Probab. Theory Related Fields **104** (1996), 427–466.

[Rus78] L. Russo, *A note on percolation*, Z. Wahrsch.verw. Geb. **43** (1978), 39–48.

[SW78] P. D. Seymour and D. J. A. Welsh, *Percolation probabilities on the square lattice*, Ann. Discrete Math. 3 (1978), 227–245.

[Sim80] B. Simon, *Correlation inequalities and the decay of correlations in ferromagnets*, Commun. Math. Phys. **77** (1980), 111–126.