Random subgraphs of finite graphs: I. The scaling window under the triangle condition

Christian Borgs* Jennifer T. Chayes* Remco van der Hofstad† Gordon Slade‡ Joel Spencer§

May 6, 2003

Abstract

We study random subgraphs of an arbitrary finite connected transitive graph $\mathbb{G}$ obtained by independently deleting edges with probability $1-p$. Let V be the number of vertices in $\mathbb{G}$, and let Ω be their degree. We define the critical threshold $p_c = p_c(\mathbb{G}, \lambda)$ to be the value of p for which the expected cluster size of a fixed vertex attains the value $\lambda V^{1/3}$, where λ is fixed and positive. We show that for any such model, there is a phase transition at p_c analogous to the phase transition for the random graph, provided that a quantity called the triangle diagram is sufficiently small at the threshold p_c. In particular, we show that the largest cluster inside a scaling window of size $|p-p_c| = \Theta(\Omega^{-1}V^{-1/3})$ is of size $\Theta(V^{2/3})$, while below this scaling window, it is much smaller, of order $O(\epsilon^{-2}\log(V\epsilon^3))$, with $\epsilon = \Omega(p_c-p)$. We also obtain an upper bound $O(\Omega(p-p_c)V)$ for the expected size of the largest cluster above the window. In addition, we define and analyze the percolation probability above the window and show that it is of order $\Theta(\Omega(p-p_c))$. Among the models for which the triangle diagram is small enough to allow us to draw these conclusions are the random graph, the n-cube and certain Hamming cubes, as well as the spread-out n-dimensional torus for $n > 6$.

1 Introduction and results

1.1 Background

Random subgraphs of finite graphs are of central interest in modern graph theory. The best known example is the random graph $G(V, p)$. It is defined as the subgraph of the complete graph on V vertices obtained by deleting edges independently with probability $1-p$, and was first studied by

*Microsoft Research, One Microsoft Way, Redmond, WA 98052, USA. `borgs@microsoft.com`, `jchayes@microsoft.com`

†Department of Mathematics and Computer Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands. `rhofstad@win.tue.nl`

‡Department of Mathematics, University of British Columbia, Vancouver, BC V6T 1Z2, Canada. `slade@math.ubc.ca`

§Department of Computer Science, Courant Institute of Mathematical Sciences, New York University, 251 Mercer St., New York, NY 10012, U.S.A. `spencer@cs.nyu.edu`

Erdős and Rényi in 1960 [17]. They showed that when p is scaled as $(1+\epsilon)V^{-1}$, there is a phase transition at $\epsilon = 0$ in the sense that the size of the largest component is $\Theta(\log V)$ for $\epsilon < 0$, $\Theta(V)$ for $\epsilon > 0$, and has the nontrivial behavior $\Theta(V^{2/3})$ for $\epsilon = 0$.

The results of Erdős and Rényi were substantially strengthened by Bollobás [9] and Łuczak [28]. In particular, they showed that the model has a scaling window of width $V^{-1/3}$ in the sense that if $p = (1 + \Lambda_V V^{-1/3})V^{-1}$, then the size of the largest component is $\Theta(V^{2/3})$ whenever Λ_V remains uniformly bounded in V, is $o(V^{2/3})$ whenever $\Lambda_V \to -\infty$, and $\omega(V^{2/3})$ whenever $\Lambda_V \to \infty$.

Considerably less is known for random subgraphs of other finite graphs. An interesting example is the n-cube $\mathbb{Q}_n$, which has vertex set $\{0,1\}^n$ and an edge joining any two vertices that differ in exactly one component. Let $V = 2^n$ denote the number of vertices in $\mathbb{Q}_n$. It is known since the work of Ajtai, Komlós and Szemerédi [5] that for p of the form $p = (1+\epsilon)n^{-1}$, the largest component is of size $O(n)$ when ϵ is fixed and negative, and is of size at least $c2^n$ for some positive $c = c(\epsilon)$ if ϵ is fixed and positive. However, very little is known about the scaling window. The best results available are those of Bollobás, Kohayakawa and Łuczak in [10], who showed the following. We use the standard terminology that a sequence of events E_n occurs *asymptotically almost surely* (a.a.s.) if $\lim_{n\to\infty} \mathbb{P}(E_n) = 0$. In [10], it is shown that that for $p = (n-1)^{-1}(1+\epsilon)$ the size of the largest cluster is at most $O(n\epsilon^{-2})$ if $\epsilon < -e^{-o(n)}$, is a.a.s. $(2\log 2)n\epsilon^{-2}(1+o(1))$ if $\epsilon \leq -(\log n)^2(\log\log n)^{-1}n^{-1/2}$, and is a.a.s. $2\epsilon 2^n(1+o(1))$ if $\epsilon \geq 60n^{-1}(\log n)^3$. Note that the resulting bounds, while much sharper than those established in [5], are still far from establishing the behavior one would expect by analogy with the random graph, namely a window of width $\Theta(V^{-1/3})$ where the largest cluster is of size $\Theta(V^{2/3})$, with different behavior outside the window on either side.

For random subgraphs of finite subsets of $\mathbb{Z}^n$, Borgs, Chayes, Kesten and Spencer [14] systematically developed a relationship between critical exponents and the width of the scaling window. In particular, they determined the size of the largest component inside, below, and above a suitably defined window, under certain scaling and hyperscaling hypotheses (proved in $n = 2$ and conjectured to be valid whenever $n \leq 6$). These results gave the appropriate version of the Erdős and Rényi [17], Bollobás [9] and Łuczak [28] results for random subsets of $\mathbb{Z}^2$.

Very recently, there have been attempts to extend the Erdős and Rényi [17] analysis to more general finite graphs. Frieze, Krivelevich and Martin [18] showed that for random subgraphs of pseudorandom graphs of V vertices, there is a phase transition in which the largest component goes from $\Theta(\log V)$ to $\Theta(V)$. Alon, Benjamini and Stacey [6] use the methods of [5] to study the critical value for the emergence of the $\Theta(V)$ component in random subgraphs of general finite graphs of large girth. Note, however, that in the language of the discussion above, both [18] and [6] consider only ϵ *fixed*; they do not get any results on the scaling window.

In this paper, we study conditions under which random subgraphs of arbitrary finite graphs behave like the random graph $G(V,p)$, both with respect to the critical point and the scaling window. More precisely, let $\mathbb{G}$ be a finite connected *transitive* graph with V vertices of degree Ω. Consider random subgraphs of $\mathbb{G}$ in which edges are deleted independently with probability $1-p$. We show that if a finite version of the so-called triangle diagram is sufficiently small at a suitably defined transition point p_c (see below), then the model behaves like the random graph in the sense that inside a window of width $|p - p_c| = \Theta(\Omega^{-1}V^{-1/3})$ the largest cluster is of order $\Theta(V^{2/3})$ while it is of order $o(V^{2/3})$ below this window. These results are essentially optimal within and below the scaling window. While we do obtain results much stronger than previous results above the scaling window, our bounds in this region are still far from optimal. It is likely that a condition beyond the triangle condition (e.g., an expansion condition) will be necessary to achieve optimal results in this region above the window.

For percolation on *infinite* graphs, the triangle diagram has been recognized as an important quantity since the work of Aizenman and Newman [4] who identified the so-called *triangle condition* as a sufficient condition for mean-field behavior for percolation on $\mathbb{Z}^n$. Here, the term *mean-field behavior* refers to the critical behavior of percolation on a tree, which is well understood. The triangle condition is defined in terms of the *triangle diagram*

$$\nabla_p(x,y) = \sum_{u,v\in\mathbb{V}} \tau_p(x,u)\tau_p(u,v)\tau_p(v,y), \tag{1.1}$$

where the sum goes over the vertices of the underlying graph, and $\tau_p(x,y)$ denotes the probability that x and y are joined by a path of occupied edges (in the random subgraph language, $\tau_p(x,y)$ is the probability that x and y lie in the same component of the random subgraph). On $\mathbb{Z}^n$, the triangle condition is the statement that at the threshold p_c, $\nabla_{p_c}(x,x)$ is finite. The triangle condition was proved on $\mathbb{Z}^n$ by Hara and Slade [22, 23], using the lace expansion, for the nearest-neighbor model for $n \geq 19$ and for a wide class of spread-out (long-range) models for all $n > 6$.

Let $\chi(p)$ denote the expected size of the cluster containing a fixed vertex. Aizenman and Newman used a differential inequality for $\chi(p)$ to show that the triangle condition implies that as $p \nearrow p_c$, the expected cluster size diverges like $(p_c - p)^{-\gamma}$ with $\gamma = 1$. Subsequently, Barsky and Aizenman [7] showed, in particular, that the triangle condition also implies that as $p \searrow p_c$ the percolation probability goes to zero like $(p - p_c)^{\beta}$ with $\beta = 1$. Their proof is based on differential inequalities for the *magnetization*. These inequalities, which were motivated by an earlier inequality of Chayes and Chayes [15, 16], had been used previously by Aizenman and Barsky to prove sharpness of the percolation phase transition on $\mathbb{Z}^n$ [2]. The exponents γ and β are examples of *critical exponents*. For percolation on a tree, the above behavior for the percolation probability and the expected cluster size can be relatively easily established with $\gamma = \beta = 1$.

In order to apply the above methods to prove mean-field behavior for percolation on finite graphs, several hurdles must be overcome. The first is the fact that it is *a priori* unclear how even to define the critical value p_c. Second, the triangle condition must be modified, since $\nabla_p(x,y)$ is always finite on a finite graph. Third, the method of integration of the differential inequalities of [2, 4, 7] requires that at p_c, the expected cluster size diverges, which is again not possible on a finite graph $\mathbb{G}$. All these facts, which we deal with below, require substantial modification and generalizations of the methods and concepts of [2, 4, 7].

In addition to the methods involving differential inequalities, our results are based on a second set of techniques, developed in [14], relating critical exponents and the width of the scaling window. We will apply these methods here to obtain information on the size of the largest cluster from information on the cluster-size distribution.

The results of this paper are valid assuming the triangle condition. For the complete graph $G(V,p)$, we will easily verify the triangle condition below, thereby reproducing some of the known results for the phase transition in the random graph. In [12], we will use the lace expansion to verify the triangle condition for several other examples of finite graphs, including the n-cube and various tori with vertex set $\{0, 1, \ldots, r-1\}^n$. This leads to several new results for these models; see Section 2.2 below.

1.2 The setting

Let $\mathbb{G} = (\mathbb{V}, \mathbb{B})$ be a finite graph. The vertex set $\mathbb{V}$ is any finite set, and the set of bonds (or edges) $\mathbb{B}$ is a subset of the set of all two-element subsets $\{x, y\} \subset \mathbb{V}$. The *degree* of a vertex $x \in \mathbb{V}$ is the number of bonds containing x. A bijective map $\varphi : \mathbb{V} \to \mathbb{V}$ is called a *graph isomorphism* if $\{\varphi(x), \varphi(y)\} \in \mathbb{B}$ whenever $\{x, y\} \in \mathbb{B}$, and $\mathbb{G}$ is called *transitive* if for every pair of vertices $x, y \in \mathbb{V}$ there is a graph-isomorphism φ with $\varphi(x) = y$. Transitive graphs are by definition regular, i.e., each vertex has the same degree.

Let $\mathbb{G}$ be an arbitrary finite, connected, transitive graph with V vertices of degree Ω. We study percolation on $\mathbb{G}$, in which each of the bonds is *occupied* with probability p independently of the other bonds, and *vacant* otherwise. We denote probabilities and expectations in the resulting product measure by $\mathbb{P}_p(\cdot)$ and $\mathbb{E}_p(\cdot)$, respectively.

As usual, we say that x *is connected to* y, written as $x \leftrightarrow y$, when there is a path from x to y consisting of occupied bonds. We define the *connectivity function* $\tau_p(x, y)$ by

$$\tau_p(x, y) = \mathbb{P}_p(x \leftrightarrow y). \tag{1.2}$$

We denote by $C(x)$ the cluster of a vertex x, that is, the set of all vertices in $\mathbb{G}$ which are connected to x, and by $|C(x)|$ the number of vertices in this cluster. Note that the distribution of $|C(x)|$ is invariant under the automorphisms of $\mathbb{G}$, and hence independent of x. Instead of $|C(x)|$, we will therefore often study $|C(0)|$, where 0, the "origin", is an arbitrary fixed vertex in $\mathbb{V}$.

Our main results involve the *cluster size distribution*,

$$P_{\geq k}(p) = \mathbb{P}_p(|C(0)| \geq k), \tag{1.3}$$

the *susceptibility*

$$\chi(p) = \mathbb{E}_p|C(0)|, \tag{1.4}$$

(i.e., the expected size of the cluster of a fixed vertex), and the maximal cluster size

$$|\mathcal{C}_{\max}| = \max\{|C(x)| : x \in \mathbb{G}\}. \tag{1.5}$$

By definition, the function χ is strictly monotone increasing on the interval $[0, 1]$, with $\chi(0) = 1$ and $\chi(1) = V$. Also,

$$\chi(p) = \mathbb{E}_p \sum_{x \in \mathbb{V}} I[x \in C(0)] = \sum_{x \in \mathbb{V}} \tau_p(0, x). \tag{1.6}$$

Recall that for $G(V, p)$ the largest cluster inside the transition window is of order $V^{2/3}$. It is not difficult—in fact, easier—to determine the expected cluster size inside the window, which turns out to be of order $V^{1/3}$. Motivated by this fact, we define the *critical threshold* $p_c = p_c(\mathbb{G}, \lambda)$ of a finite graph $\mathbb{G}$ to be the unique solution to the equation

$$\chi(p_c) = \lambda V^{1/3}, \tag{1.7}$$

with $\lambda > 0$. There is some flexibility in the choice of λ, connected with the fact that the transition takes place within a window and not at a particular value of p. A convenient choice is to take λ to be constant (independent of V). We will always assume that $1 < \lambda V^{1/3} < V$ so that p_c is well defined and $0 < p_c < 1$.

The definition (1.7) is appropriate for graphs that obey mean-field behavior, which we expect only for graphs that are in some sense "high-dimensional." As we discuss in more detail in Section 3.4.2, a different definition of the critical threshold would be appropriate for a graph providing a finite approximation to $\mathbb{Z}^n$ for $n < 6$.

1.3 Main results

In this section, we state our main results, which hold for arbitrary finite connected transitive graphs, provided the triangle diagram (1.1) at p_c is sufficiently small. To be more precise, we will assume that

$$\nabla_{p_c}(x,y) \leq \delta_{x,y} + a_0 \tag{1.8}$$

for a sufficiently small constant a_0, a condition we call the *finite-graph triangle condition*, or more briefly, the *triangle condition*. Although we have not done the necessary computations, the constant a_0 need not be extremely small, and we expect our results to hold for a_0 of the order of $\frac{1}{10}$.

By (1.6), $\sum_{y\in\mathbb{V}} \nabla_p(x,y) = \chi(p)^3$. As a consequence, the triangle condition implies that

$$\lambda^3 \leq a_0 + V^{-1}. \tag{1.9}$$

In other words, small λ is a necessary condition for the triangle condition to hold. It turns out that it also sufficient for many graphs $\mathbb{G}$. For the random graph, this is shown in Section 2.1, and for several other models in [12]; see Section 2.2. Indeed, we will show that for these models,

$$\nabla_p(x,y) = \delta_{x,y} + O(\Omega^{-1}) + O(\chi^3(p)/V) \tag{1.10}$$

whenever $\chi^3(p)/V$ is small enough.

Our results concerning the critical threshold are given in the following theorem. In its statement, we make the abbreviations

$$\epsilon_0 = \frac{1}{\chi(p_c)} = \lambda^{-1}V^{-1/3}, \tag{1.11}$$

and

$$\bar{\nabla}_p = \max_{\{x,y\}\in\mathbb{B}} \nabla_p(x,y). \tag{1.12}$$

Theorem 1.1 (Critical threshold). *For all finite connected transitive graphs $\mathbb{G}$, the following statements hold.*
i) If $\lambda > 0$ and the triangle condition (1.8) holds for some $a_0 < 1$, then

$$1 - \epsilon_0 \leq \Omega p_c \leq \frac{1-\epsilon_0}{1-a_0}. \tag{1.13}$$

ii) Given $0 < \lambda_1 < \lambda_2 < \infty$, let p_i be defined by $\chi(p_i) = \lambda_i V^{1/3}$ $(i = 1, 2)$. If $\bar{\nabla}_{p_2} < 1$, then

$$\frac{\lambda_2-\lambda_1}{\lambda_1\lambda_2}\frac{1}{V^{1/3}} \leq \Omega(p_2-p_1) \leq \frac{1}{1-\bar{\nabla}_{p_2}}\frac{\lambda_2-\lambda_1}{\lambda_1\lambda_2}\frac{1}{V^{1/3}}. \tag{1.14}$$

For example, if $\mathbb{G}$ is the complete graph on n vertices (so that $V = n$ and $\Omega = n-1$) and p_2 is inside the transition window, then p_1 remains within the transition window for any constant $\lambda_1 < \lambda_2$.

Our results concerning the subcritical phase are given in the following theorem.

Theorem 1.2 (Subcritical phase). *There is a (small) constant $b_0 > 0$ such that the following statements hold for all positive λ, all finite connected transitive graphs $\mathbb{G}$ and all p of the form $p = p_c - \Omega^{-1}\epsilon$ with $\epsilon \geq 0$.*
i) If the triangle condition (1.8) *holds for some $a_0 < 1$, then*

$$\frac{1}{\epsilon_0 + \epsilon} \leq \chi(p) \leq \frac{1}{\epsilon_0 + [1 - a_0]\epsilon}. \tag{1.15}$$

ii) If the triangle condition holds for some $a_0 \leq b_0$ and if $\lambda V^{1/3} \geq b_0^{-1}$, then

$$10^{-4}\chi^2(p) \leq \mathbb{E}_p\Big(|\mathcal{C}_{\max}|\Big) \leq 2\chi^2(p)\log(V/\chi^3(p)), \tag{1.16}$$

$$\mathbb{P}_p\Big(|\mathcal{C}_{\max}| \leq 2\chi^2(p)\log(V/\chi^3(p))\Big) \geq 1 - \frac{\sqrt{e}}{[2\log(V/\chi^3(p))]^{3/2}}, \tag{1.17}$$

and, for $\omega \geq 1$,

$$\mathbb{P}_p\Big(|\mathcal{C}_{\max}| \geq \frac{\chi^2(p)}{3600\omega}\Big) \geq \Big(1 + \frac{36\chi^3(p)}{\omega V}\Big)^{-1}. \tag{1.18}$$

Our next theorem states our results inside the scaling window.

Theorem 1.3 (Critical Window). *Let $\lambda > 0$ and $\Lambda < \infty$. Then there are finite positive constants $b_1, \ldots, b_8$ such that the following statements hold for all finite connected transitive graphs $\mathbb{G}$ provided the triangle-condition* (1.8) *holds for some constant $a_0 \leq b_0$ and $\lambda V^{1/3} \geq b_0^{-1}$, with b_0 as in Theorem 1.2. Let $p = p_c + \Omega^{-1}\epsilon$ with $|\epsilon| \leq \Lambda V^{-1/3}$.*
i) If $k \leq b_1 V^{2/3}$, then

$$\frac{b_2}{\sqrt{k}} \leq P_{\geq k}(p) \leq \frac{b_3}{\sqrt{k}}. \tag{1.19}$$

ii)

$$b_4 V^{2/3} \leq \mathbb{E}_p\Big[|\mathcal{C}_{\max}|\Big] \leq b_5 V^{2/3} \tag{1.20}$$

and, if $\omega \geq 1$, then

$$\mathbb{P}_p\Big(\omega^{-1}V^{2/3} \leq |\mathcal{C}_{\max}| \leq \omega V^{2/3}\Big) \geq 1 - \frac{b_6}{\omega}. \tag{1.21}$$

iii)

$$b_7 V^{1/3} \leq \chi(p) \leq b_8 V^{1/3}. \tag{1.22}$$

In the above statements, the constants b_2 and b_3 can be chosen independent of λ and Λ, the constants b_5 and b_8 depend on Λ and not λ, and the constants b_1, b_4, b_6 and b_7 depend on both λ and Λ.

Our results on the supercritical phase are given in the following theorem.

Theorem 1.4 (Supercritical phase). *Let $\lambda > 0$. The following statements hold for all finite connected transitive graphs $\mathbb{G}$ provided the triangle-condition* (1.8) *holds for some constant $a_0 \leq b_0$ and $\lambda V^{1/3} \geq b_0^{-1}$, with b_0 as in Theorem 1.2. Let $p = p_c + \epsilon\Omega^{-1}$ with $\epsilon \geq 0$.*
i)

$$\mathbb{E}_p(|\mathcal{C}_{\max}|) \leq 21\epsilon V + 7V^{2/3}, \tag{1.23}$$

and, for all $\omega > 0$,

$$\mathbb{P}_p\Big(|\mathcal{C}_{\max}| \geq \omega(V^{2/3} + \epsilon V)\Big) \leq \frac{21}{\omega}. \tag{1.24}$$

ii)

$$\chi(p) \leq 81(V^{1/3} + \epsilon^2 V). \tag{1.25}$$

Note that Theorem 1.4 does not give lower bounds on the size of the largest supercritical cluster. We believe that this is not a mere technicality. Indeed, the formation of a giant component in the random graph is closely related to the fact that moderately large clusters have a significant chance to merge into a single, giant component as ϵ is increased beyond the critical value by an amount of order $V^{-1/3}$. This fact involves the geometry of the random graph, and may not be true for arbitrary transitive graphs obeying the triangle condition. It would be interesting to know whether there exists a sequence of transitive graphs $\mathbb{G}_n$ such that the largest cluster above the window is $o(\epsilon V)$, at least if $\epsilon V^{-1/3} \to \infty$ sufficiently slowly. On the other hand, as we explain in more detail in Section 2.2 below, our results apply to the n-cube $\mathbb{Q}_n$, and for $\mathbb{Q}_n$ we prove complementary lower bounds to the upper bounds of Theorem 1.4 in [13]. Our proof of these upper bounds is valid for $\epsilon \geq e^{-cn^{1/3}}$, and not in the full domain $\epsilon \gg V^{-1/3} = 2^{-n/3}$ where we would conjecture that they are valid. The methods of [13] rely heavily on the specific geometry of $\mathbb{Q}_n$ and do not apply at the level of generality of Theorem 1.4.

We close this section with a theorem that gives a more precise bound on the susceptibility below the window, under the assumption that the stronger triangle condition (1.10) holds. We make the constants in (1.10) explicit by assuming that

$$\nabla_p(x,y) \leq \delta_{x,y} + K_1\Omega^{-1} + K_2\frac{\chi^3(p)}{V} \tag{1.26}$$

for some constants K_1, $K_2 < \infty$ and all $p \leq p_c$. Let

$$a = K_1\Omega^{-1} + K_2\lambda^3, \tag{1.27}$$

$$\tilde{K}_2 = K_2/(1-a), \tag{1.28}$$

$$\tilde{a}(\epsilon) = K_1\Omega^{-1} + K_2\lambda^3\frac{\epsilon_0}{\epsilon_0 + (1-a)\epsilon}. \tag{1.29}$$

Theorem 1.5 (Sharpened bounds). *Let* $\lambda > 0$ *and let* $\mathbb{G}$ *be a finite connected transitive graph such that* (1.26) *holds for all* $p \leq p_c$*, with the constant in* (1.27) *obeying* $a < 1$*. Let* $p = p_c - \Omega^{-1}\epsilon$ *with* $\epsilon \geq 0$*, and let* $\tilde{K}_2$ *and* $\tilde{a}(\epsilon)$ *be given by* (1.28)*–*(1.29)*. Then*

$$1 - \epsilon_0 \leq \Omega p_c \leq \frac{1-\epsilon_0}{1 - K_1\Omega^{-1} - \tilde{K}_2\lambda^3\epsilon_0}, \tag{1.30}$$

$$\frac{1}{\epsilon_0 + \epsilon} \leq \chi(p) \leq \frac{1}{\epsilon_0 + [1-\tilde{a}(\epsilon)]\epsilon}. \tag{1.31}$$

The inequality (1.30) implies that $|\Omega p_c - 1| = O(\Omega^{-1} + \lambda^{-1}V^{-1/3})$. The significance of (1.31) is most apparent if we consider a sequence of graphs with $\lambda > 0$ fixed, $V \to \infty$ and $\Omega \to \infty$, for ϵ such that $\epsilon/\epsilon_0 \to \infty$. In this limit, (1.31) implies that

$$\chi(p) = \frac{1}{\epsilon}[1 + o(1)]. \tag{1.32}$$

We will apply Theorem 1.5 to the random graph in Section 2.1.

1.4 General sequences of finite graphs

To illustrate our theorems, it is instructive to consider a sequence of finite connected transitive graphs $\mathbb{G}_n = (\mathbb{V}_n, \mathbb{B}_n)$ with $|\mathbb{V}_n| \to \infty$. We will say that such a sequence obeys the *finite-graph triangle condition* if there exist a $\lambda > 0$ such that the condition (1.8) holds for all n, with a constant a_0 that is at most as large as the constant b_0 in Theorem 1.2.

Consider thus a sequence of finite connected transitive graphs $\mathbb{G}_n$ satisfying the finite-graph triangle condition. Consider also a sequence of probabilities of the form

$$p_n = p_c + \Lambda_n \Omega^{-1} V_n^{-1/3}. \tag{1.33}$$

Motivated by the random graph (and our theorems) we say that the sequence p_n is *inside the window*, if $\limsup_{n\to\infty} |\Lambda_n| < \infty$, *below the window* if $\Lambda_n \to -\infty$, and *above the window* if $\Lambda_n \to \infty$ as $n \to \infty$. In order to avoid dealing with higher order corrections in $\epsilon_n = \Lambda_n V_n^{-1/3}$, we assume here that $\epsilon_n \to 0$.

Consider first a sequence *below the window*, i.e., assume that $\Lambda_n \to -\infty$ as $n \to \infty$. The first statement of Theorem 1.2 then implies that

$$\chi(p_n) = \Theta(\epsilon_n^{-\gamma}) \tag{1.34}$$

with $\gamma = 1$, while the second implies that

$$\Theta(\Lambda_n^{-2} V_n^{2/3}) \le \mathbb{E}_{p_n}\Big(|\mathcal{C}_{\max}|\Big) \le \Theta(\Lambda_n^{-2}\, V_n^{2/3} \log \Lambda_n), \tag{1.35}$$

and

$$\Theta(\Lambda_n^{-2} V_n^{2/3}) \le |\mathcal{C}_{\max}| \le \Theta(\Lambda_n^{-2}\, V^{2/3} \log \Lambda_n) \quad \text{a.a.s. as } n \to \infty. \tag{1.36}$$

Note that this implies, in particular, that below the window, $|\mathcal{C}_{\max}| = o(V_n^{2/3})$ a.a.s. as $n \to \infty$.

Next, consider a sequence p_n *inside the window*, i.e., a sequence of the form (1.33) with $\limsup |\Lambda_n| < \infty$. Theorem 1.3 then implies that

$$\chi(p_n) = \Theta(V_n^{1/3}), \tag{1.37}$$

$$\mathbb{E}_p\Big[|\mathcal{C}_{\max}|\Big] = \Theta(V_n^{2/3}), \tag{1.38}$$

with the probability of the event

$$\omega(n)^{-1} \le \frac{|\mathcal{C}_{\max}|}{\mathbb{E}_p\Big[|\mathcal{C}_{\max}|\Big]} \le \omega(n) \tag{1.39}$$

going to one whenever $\omega(n) \to \infty$ as $n \to \infty$.

Let us finally consider a sequence p_n *above the window*, i.e., a sequence of the form (1.33) with $\Lambda_n \to \infty$. Theorem 1.4 i) then implies that the expected size of the largest cluster is $O(\epsilon_n V_n)$, and Theorem 1.4 ii) shows that $\chi(p_n) = O(\epsilon_n^2 V_n)$.

1.5 The percolation probability and magnetization

It is a major result for percolation on $\mathbb{Z}^n$ that the value of p for which $\chi(p)$ becomes infinite is the same as the value of p where the percolation probability, or order parameter, $\mathbb{P}_p(|C(0)| = \infty)$, becomes positive [2, 29]. In the present setting, since the graph is finite, there can be no infinite cluster and the definition of the order parameter needs to be adapted. A natural definition of the *finite-size order parameter* is the ratio of the expected maximal cluster size to the volume V:

$$\theta(p) = \frac{\mathbb{E}_p(|\mathcal{C}_{\max}|)}{V}. \tag{1.40}$$

However, we are unable to prove a good lower bound on (1.40) in the supercritical regime, and we therefore consider an alternative definition in terms of the cluster size distribution $P_{\geq k}(p)$. Parameterizing p as $p = p_c + \epsilon\Omega^{-1}$, we define the percolation probability by

$$\theta_\alpha(p) = \mathbb{P}_p(|C(x)| \geq N_\alpha) = P_{\geq N_\alpha}(p), \tag{1.41}$$

where

$$N_\alpha = \frac{1}{\epsilon^2}\Big(\epsilon V^{1/3}\Big)^\alpha. \tag{1.42}$$

Here α is a constant with $0 < \alpha < 1$. This definition is motivated by the known behavior of the random graph. Above the window (corresponding to $\epsilon V^{1/3} \to \infty$), it is known that a.a.s., the largest component has size $|\mathcal{C}_{\max}| = 2\epsilon V[1+o(1)]$, while the second largest has size $2\epsilon^{-2}\log(\epsilon^3 V)(1+o(1))$. For the random graph above the window, the cutoff N_α in (1.41) is therefore much larger than the second largest, and much smaller than the largest cluster. As a consequence, the ratio of $\theta_\alpha(p)$ and $\theta(p)$ goes to one when considered on the random graph above threshold. (The above reasoning actually suggests the wider range $0 < \alpha < 3$ for α, but for technical reasons we require $0 < \alpha < 1$.) Our results for $\theta_\alpha(p)$ are stated in the following theorem.

Theorem 1.6 (The percolation probability). *Let $\lambda > 0$ and $0 < \alpha < 1$. Then there are finite positive constants b_9, b_{10}, b_{11}, b_{12} such that the following statements hold for all finite connected transitive graphs $\mathbb{G}$ provided the triangle-condition (1.8) holds for some constant $a_0 \leq b_0$ and $\lambda V^{1/3} \geq b_0^{-1}$, with b_0 as in Theorem 1.2. Let $p = p_c + \epsilon\Omega^{-1}$.*
i)

$$b_{10}\epsilon \leq \theta_\alpha(p) \leq 27\epsilon, \tag{1.43}$$

where the lower bound holds when $b_9 V^{-1/3} \leq \epsilon \leq 1$ and the upper bound holds when $\epsilon V^{1/3} \geq 1$.
ii) If $\max\{b_{12}V^{-1/3}, V^{-\eta}\} \leq \epsilon \leq 1$, where $\eta = \frac{1}{3}\frac{3-2\alpha}{5-2\alpha}$, then

$$\mathbb{P}_p\Big(|\mathcal{C}_{\max}| \leq [1 + (\epsilon V^\eta)^{-1}]\theta_\alpha(p)V\Big) \geq 1 - \frac{b_{11}}{(\epsilon V^\eta)^{3-2\alpha}}. \tag{1.44}$$

In the above statements, the constants b_9, b_{10}, b_{11} and b_{12} depend on both α and λ.

Theorem 1.6 i) is analogous to results proved for percolation on $\mathbb{Z}^n$ (assuming high n for the upper bound) in [2, 7, 22]. Theorem 1.6 ii) shows that it is unlikely that the largest supercritical cluster is larger than $\theta_\alpha(p)V$, at least for ϵ not too small. As we will describe in more detail in Section 2.2 below, it is shown in [13] that when $\mathbb{G}$ is the n-cube, it is possible also to prove a lower bound on $|\mathcal{C}_{\max}|$, so that $|\mathcal{C}_{\max}|$ is of the same order of magnitude as $\theta_\alpha(p)V$, at least when ϵ is not too small. The fact that $\theta_\alpha(p)$ can be used in this way serves as further justification for the definition (1.41). In [14], a similar approach was used for $\mathbb{Z}^n$ in low dimensions.

Our analysis of $\theta_\alpha(p)$, and more generally our analysis of the cluster size distribution $P_{\geq k}(p)$, is primarily based on an analysis of the magnetization. Let $P_k(p)$ be the probability that the $|C(0)| = k$. The *magnetization* $M(p, \gamma)$ is defined by

$$M(p,\gamma) = 1 - \sum_{k=1}^{V} (1-\gamma)^k P_k(p). \tag{1.45}$$

Thus $M(p, \gamma)$ is essentially the generating function for the sequence $P_k(p)$, and $M(p, 0) = 0$ for all p. Estimates on $M(p, \gamma)$ for small γ can be converted into estimates on $P_k(p)$ for large k, via an analysis reminiscent of a Tauberian theorem. The name "magnetization" is used because $M(p, \gamma)$ is analogous the the magnetization in spin systems, and the variable $h \geq 0$ defined by $\gamma = 1 - e^{-h}$ plays the role of an external magnetic field in that context. Our main results for the magnetization are summarized in the following theorem.

Theorem 1.7 (The magnetization). *Assume that a_0 is sufficiently small, and let $0 \leq \gamma \leq 1$.*
i) If $p \leq p_c$ then

$$\frac{1}{3} \min\{\sqrt{\gamma}, \gamma\chi(p)\} \leq M(p,\gamma) \leq \min\{\sqrt{12\gamma},\, \gamma\chi(p)\}. \tag{1.46}$$

ii) If $p = p_c + \Omega^{-1}\epsilon$ and $\epsilon \geq 0$ then

$$M(p,\gamma) \leq \sqrt{12\gamma} + 13\epsilon. \tag{1.47}$$

Let $0 \leq \alpha < 1$ and $\rho > 0$. There is a positive $c = c(\alpha, \lambda)$ and $b_{13} = b_{13}(\alpha, \lambda, \rho)$ such that if $b_{13}V^{-1/3} \leq \epsilon \leq 1$ then

$$M(p, \rho N_\alpha^{-1}) \geq c\epsilon \min\{1, \rho^{1/(2-\alpha)}\}. \tag{1.48}$$

1.6 Guide to the paper

In Section 2, we discuss several examples where our general results can be applied. In Section 3, we indicate some of the main ideas that enter into the proofs of our main results.

The following table indicates where the various theorems are proved. The notation [u.b.] refers to the upper bounds on $|\mathcal{C}_{\max}|$ and [l.b.] to the lower bounds.

Theorem	1.1	1.2 i), ii) [u.b]	1.2 ii) [l.b]	1.3 i)	1.3 ii-iii)	1.4 i)
Section	4	4	7	6	3	6
Theorem	1.4 ii)	1.5	1.6 i)	1.6 ii)	1.7	
Section	8	4	6	9	5	

There is no dependence on Section 4 in Sections 5–9. The bounds on the magnetization proved in Section 5 are crucial for Sections 6–9. Section 7 depends on Section 6, which in turn depends on Section 5. Sections 8 and 9 each depend on Section 5 and on no other section. Sections 7, 8 and 9 are mutually independent. Three differential inequalities, needed in Sections 4, 5, and 8, are proved in Appendix A.

2 Examples

2.1 The random graph

In this section, we illustrate both the finite-graph triangle condition and our results when $\mathbb{G}$ is the random graph on n vertices. In the notation of the last section, we thus consider the graph $\mathbb{G} = K_n$, the complete graph on n vertices, with $V = n$ vertices of degree $\Omega = n - 1$.

2.1.1 The triangle condition for the random graph

For the random graph, the triangle diagram can be explicitly and easily calculated in terms of the expected cluster size $\chi(p)$, as follows. Due to the high degree of symmetry of the complete graph, the two-point function takes on only the two distinct values $\tau_p(x,x) = 1$ and $\tau_p(x,y) = \tau$ (say) for $x \neq y$, so that $\tau_p(x,y) = \delta_{x,y} + \tau(1 - \delta_{x,y})$. The triangle diagram (1.1) is therefore given by

$$\nabla_p(x,y) = \begin{cases} 1 + 3(n-1)\tau^2 + (n-1)(n-2)\tau^3 & \text{if } x = y, \\ 3\tau + 3(n-2)\tau^2 + [1 + (n-1)(n-2)]\tau^3 & \text{if } x \neq y. \end{cases} \tag{2.1}$$

Also, by (1.6), $\tau = (n-1)^{-1}(\chi(p) - 1)$. Since $\chi(p) \leq n$, this implies that $\tau \leq n^{-1}\chi(p)$. It is then straightforward to see that

$$\nabla_p(x,y) \leq \delta_{x,y} + \frac{\chi^3(p)}{n}\left[1 + 3\chi^{-1}(p) + 3\chi^{-2}(p)\right] \leq \delta_{x,y} + 7n^{-1}\chi^3(p). \tag{2.2}$$

Recalling that by definition,

$$\chi(p_c) = \lambda n^{1/3}, \tag{2.3}$$

we have thus obtained the triangle condition (1.8) with $a_0 = 7\lambda^3$. In addition, (1.26) holds with $K_1 = 0$ and $K_2 = 7$.

2.1.2 The phase transition for the random graph

Having verified the triangle condition, we can now apply the results of Section 1.3 provided we take λ to be a sufficiently small constant. Starting with Theorem 1.5, since $\Omega = n-1 = n(1+O(n^{-1}))$, (1.30) implies that

$$p_c = \frac{1}{n}(1 + O(n^{-1/3})). \tag{2.4}$$

While we cannot expect that $p_c = 1/n$ (in fact, (1.30) implies that $p_c < 1/n$ if λ is small enough), it differs from the traditional value by only a small amount, small enough to keep it inside the scaling window. Thus our definition of p_c is quite sensible for the random graph.

In Theorems 1.2–1.5, we have used the parameter $\epsilon = \Omega(p - p_c)$. For the random graph, we will use the scaling $p = p_c(1 + \Lambda_n n^{-1/3})$, which corresponds to $\epsilon = np_c\Lambda_n n^{-1/3}$. Then up to constants, ϵ is equivalent to

$$\epsilon_n = \Lambda_n n^{-1/3}. \tag{2.5}$$

Note that if $\Lambda_n \to -\infty$ then for $K_1 = 0$ we have $\tilde{a}(\epsilon) = \Theta(|\Lambda_n^{-1}|)$, and (1.31) implies the simpler statement

$$\chi(p) = \frac{1}{|\epsilon_n|}(1 + O(\Lambda_n^{-1})), \tag{2.6}$$

as claimed below in (2.7).

The conclusions of Theorems 1.2–1.6 for the random graph are summarized in the following theorem.

Theorem 2.1. *Let $p = p_n = p_c(1 + \Lambda_n n^{-1/3})$ with $p_c = p_c(n, \lambda)$ defined by (2.3). There exists a constant λ_0 such that the following statements are true for all fixed, strictly positive $\lambda \le \lambda_0$, with the constants implicit in our $O(\cdot)$ and $\Theta(\cdot)$ possibly depending on λ.*
i) (Subcritical phase). If $\Lambda_n \to -\infty$ as $n \to \infty$ then

$$\chi(p) = \frac{n^{1/3}}{|\Lambda_n|}(1 + O(\Lambda_n^{-1})), \tag{2.7}$$

$$n^{2/3}\Theta(\Lambda_n^{-2}) \le \mathbb{E}_p\Big(|\mathcal{C}_{\max}|\Big) \le 6n^{2/3}\,\frac{\log|\Lambda_n|}{\Lambda_n^2}(1 + O(\Lambda_n^{-1})), \tag{2.8}$$

with

$$n^{2/3}\Theta(\Lambda_n^{-2}) \le |\mathcal{C}_{\max}| \le 6n^{2/3}\,\frac{\log|\Lambda_n|}{\Lambda_n^2}(1 + O(\Lambda_n^{-1})) \quad \text{a.a.s. as } n \to \infty. \tag{2.9}$$

ii) (Critical window). If $\Lambda = \limsup|\Lambda_n| < \infty$ as $n \to \infty$ then

$$\chi(p) = \Theta(n^{1/3}), \qquad \mathbb{E}_p\Big[|\mathcal{C}_{\max}|\Big] = \Theta(n^{2/3}), \tag{2.10}$$

with

$$\omega(n)^{-1} \le \frac{|\mathcal{C}_{\max}|}{\mathbb{E}_p\Big[|\mathcal{C}_{\max}|\Big]} \le \omega(n) \quad \text{a.a.s. as } n \to \infty \tag{2.11}$$

whenever $\omega(n) \to \infty$ as $n \to \infty$. If $kn^{-2/3}$ is small enough (depending on Λ), then

$$P_{\ge k}(p) = \Theta(k^{-1/2}). \tag{2.12}$$

iii) (Supercritical phase). Let $0 < \alpha < 1$. If $\Lambda_n \to \infty$ as $n \to \infty$ and $\epsilon_n = \Lambda_n n^{-1/3} \to 0$ then

$$\chi(p) = O(n^{1/3}\Lambda_n^2), \qquad \mathbb{E}_p(|\mathcal{C}_{\max}|) = O(\epsilon_n n), \tag{2.13}$$

and

$$\theta_\alpha(p) = \Theta(\epsilon_n). \tag{2.14}$$

If $\Lambda_n \to \infty$ at least as fast as n^η, where $\eta = \frac{1}{3}\frac{3-2\alpha}{5-2\alpha}$, then

$$\mathbb{P}_p\Big(|\mathcal{C}_{\max}| \le \Big(1 + \frac{1}{\epsilon n^\eta}\Big)\theta_\alpha(p)n\Big) \ge 1 - O\Big(\frac{1}{(\epsilon n^\eta)^{3-2\alpha}}\Big). \tag{2.15}$$

It is interesting to compare Theorem 2.1 with previously known results for the phase transition in the random graph. Since $p_c = n^{-1} + O(n^{-4/3})$, if we change our parametrization to $p = n^{-1} + \Lambda n^{-4/3}$ then we effectively change Λ by a constant. This affects the constants in the critical window and has an asymptotically negligible effect in the subcritical and supercritical phases. This new parametrization is the standard parametrization used (with λ instead of Λ) in much of the random graph literature. We will refer to the book of Janson, Łuczak and Rucinski [27], where references to the original literature can be found. Results in [27] are expressed in terms of the variable s, where $\frac{n}{2}+s$ gives the number of occupied edges, and our formulas can be compared to theirs by setting $s=\frac{\Lambda}{2}n^{2/3}$.

In the subcritical phase, we show that the largest component has size between $c_1 n^{2/3}\Lambda^{-2}$ and $c_2 n^{2/3}\Lambda^{-2}\log|\Lambda|$, while [27, Theorem 5.6] gives, in particular, that the largest component is asymptotically of size $6n^{2/3}\Lambda^{-2}\log|\Lambda|$. The constant 6 in the upper bounds in (2.8)–(2.9) is therefore sharp.

In the critical window, we show that the largest component has size $\Theta(n^{2/3})$, while [27, Theorem 5.20] gives, in particular, that the largest component has size $X_1' n^{2/3}$ where X_1' is a random variable with a nontrivial distribution over $(0,\infty)$.

In the supercritical phase, let $\epsilon=\Lambda n^{-1/3}$, so that $p=\frac{1}{n}(1+\epsilon)$. We show that the largest component has size $O(\epsilon n)$. As mentioned below the statement of Theorem 1.4, we have no lower bound on the largest subcritical cluster in our general setting. In [27, Theorem 5.12], the largest component is shown asymptotically to have size $2\epsilon n$ (their $\overline{s}$ is asymptotic to s when $\epsilon\to 0$). Moreover, [27, Theorem 5.7] yields that the rth largest component (for any fixed $r\geq 2$) has size asymptotic to $6n^{2/3}\Lambda^{-2}\log\Lambda$. We are unable to get any reasonable upper bounds on the size of the second largest component.

Although our results are not state-of-the-art for the random graph, it is nevertheless striking that they follow from a general theory that makes no calculation specific to the random graph apart from the simple verification of the finite-graph triangle condition in Section 2.1.1. More importantly, our theorems apply much more generally, to models such as the n-cube and finite tori in $\mathbb{Z}^n$ for $n>6$, where they imply strong new results.

2.2 The n-cube and several tori

In [12], we use the lace expansion to prove quite generally that for finite graphs that are tori the triangle condition for percolation is implied by a certain triangle condition for simple random walk on the graph. As we show in [12], the latter is easily verified for the following graphs with vertex set $\{0,1,\ldots,r-1\}^n$:

1. The narrow torus: an edge joins vertices that differ by 1 in exactly one component, with the periodic boundary condition that 0 and $r-1$ differ by 1, for $r\geq 2$ fixed and $n\to\infty$. For $r=2$, this is the n-cube.

2. The Hamming torus: an edge joins vertices that differ in exactly one component, again with the periodic boundary condition, for $r\geq 2$ fixed and $n\to\infty$.

3. The wide torus in high dimensions: the same edge set as in (i) but now n is large and fixed and we study the limit $r\to\infty$ to approximate $\mathbb{Z}^n$.

4. The wide spread-out torus in dimensions $n>6$: an edge joins vertices $x=(x_1,\ldots,x_n)$ and $y=(y_1,\ldots,y_n)$ if $\max_{i=1,\ldots,n}|x_i-y_i|\leq L$ (with periodic boundary conditions) with $n>6$ fixed, L large and fixed, in the limit $r\to\infty$ to approximate range-L percolation on $\mathbb{Z}^n$.

Our conclusions thus apply to the percolation phase transition for each of the above examples. The above examples are all high-dimensional graphs. We do not expect the triangle condition to hold for low-dimensional graphs, and in particular do not expect the triangle condition to hold for the wide tori in dimensions $n \leq 6$. Nor do we expect the conclusions of our theorems to hold in low dimensions.

Combined with [12], our results show that the phase transition for percolation on the n-cube $\mathbb{Q}_n$ shares several features with the phase transition for the random graph. In particular, it follows from the triangle condition for $\mathbb{Q}_n$ proved in [12] and Theorem 1.5 that $p_c(\mathbb{Q}_n, \lambda) = n^{-1} + O(n^{-2})$, for any sufficiently small choice of λ. In [26], this series is substantially extended. In [13], we use the lower bound on the percolation probability of (1.43) to prove a lower bound on the largest supercritical cluster for the n-cube. This leads to a substantial improvement of some of the results of [5, 10].

Our results for the wide tori in high dimensions show that in a window of width $r^{-n/3}$ centered at $p_c = p_c(r, n)$, the largest cluster has size $\Theta(r^{2n/3})$. It is interesting to compare this with a previous result for $\mathbb{Z}^n$. For $p = p_c(\mathbb{Z}^n)$, consider the restriction of percolation configurations to a large box of side r, under the *bulk* boundary condition in which the clusters in the box are defined to be the intersection of the box with clusters in the infinite lattice (and thus clusters in the box need not be connected within the box). How large is the largest cluster in the box, as $r \to \infty$? The combined results of Aizenman [1] and Hara, van der Hofstad and Slade [21] show that for spread-out models with $n > 6$ the largest cluster has size of order r^4, and there are order r^{n-6} clusters of this size. For the nearest-neighbor model in dimensions $n \gg 6$, the same results follow from the combined results of [1] and Hara [20]. The size r^4 for the largest critical cluster size is different than the $r^{2n/3}$ that we prove for $p = p_c(r, n)$ under the *periodic* boundary condition of the torus. Aizenman [1] had raised the question whether a change from bulk to periodic boundary conditions would change the r^4 to $r^{2n/3}$. It would be interesting to attempt to extend our results, to show that $p_c(\mathbb{Z}^n)$ lies inside the critical window centered at $p_c(r, n)$ for large r, thereby providing an affirmative answer to Aizenman's question.

3 Overview of the proofs

3.1 Differential inequality for the susceptibility

The results for the critical threshold and the subcritical susceptibility, stated in Theorems 1.1, 1.2 i) and 1.5 are all derived from the differential inequality

$$[1 - \bar{\nabla}_p]\Omega \leq -\frac{d\chi^{-1}(p)}{dp} \leq \Omega, \tag{3.1}$$

with $\bar{\nabla}_p$ defined by (1.12). This differential inequality was proved by Aizenman and Newman [4] with infinite graphs in mind, but its proof applies also to finite transitive graphs. We recall the proof of (3.1) in Appendix A.1. The triangle condition is used to bound the left side of (3.1) from below. In Section 4.1, we will show that integration of (3.1) leads directly to proofs of Theorems 1.1, 1.2 i) and 1.5.

3.2 Differential inequalities for the magnetization

Aizenman and Barsky [2] used differential inequalities for the magnetization to prove sharpness of the phase transition for percolation on $\mathbb{Z}^n$. In [7], they derived a complementary differential inequality, assuming the triangle condition, which implied that on $\mathbb{Z}^n$ the magnetization and percolation probability behave asymptotically as $M(p_c, \gamma) = \Theta(\sqrt{\gamma})$ and $\mathbb{P}_p(|C(0)| = \infty) = \Theta(p - p_c)$. In Section 5, we recall the statement of the differential inequalities of [2], and in Appendix A.2 we derive a variant of the complementary differential inequality of [7]. In Section 5, we show how to integrate the differential inequalities to obtain the bounds on the magnetization stated in Theorem 1.7. In performing the integration, care is required to deal with the finite size effects.

The bounds on the magnetization proved in Theorem 5 lie at the heart of our method. They play a crucial role in all of Sections 6–9 and in the proofs of Theorems 1.2 ii)–1.6.

3.3 The cluster size distribution

The magnetization is a generating function for the sequence $P_k(p)$, and its behavior for small γ is closely related to the behavior of $P_{\geq k}(p)$ for large k. This is made precise in Section 6, where Theorem 1.3 i) and related bounds on $P_{\geq k}(p)$ are obtained from the bounds on the magnetization proven in Section 5. The upper bounds on the magnetization easily lead to upper bounds on the cluster size distribution for all $p \in [0, 1]$. The lower bounds are more difficult. We will need matching upper and lower bounds on $M(p, \gamma)$ to obtain good lower bounds on $P_{\geq k}(p)$, and, in the supercritical phase, our lower bounds on $M(p, \gamma)$ are in the restricted form given in (1.48), with γ proportional to N_α^{-1}. Our bounds on $P_{\geq k}(p)$ then lead to a proof of the bounds on $\theta_\alpha(p) = P_{\geq N_\alpha}(p)$ stated in Theorem 1.6 i).

3.4 The scale of the largest cluster

3.4.1 The random variable $Z_{\geq k}$

Given $k > 0$, let

$$Z_{\geq k} = \sum_{x \in \mathbb{V}} I[|C(x)| \geq k] \tag{3.2}$$

denote the number of vertices that lie in clusters of size k or larger. Then

$$\mathbb{E}_p(Z_{\geq k}) = V P_{\geq k}(p). \tag{3.3}$$

By definition, $|\mathcal{C}_{\max}| \geq k$ if and only if $Z_{\geq k} \geq k$, and hence, by the Markov inequality,

$$\mathbb{P}_p\Big(|\mathcal{C}_{\max}| \geq k\Big) \leq \frac{V P_{\geq k}(p)}{k}. \tag{3.4}$$

and

$$\mathbb{E}_p\Big(|\mathcal{C}_{\max}|\Big) \leq k + \mathbb{E}_p\Big(Z_{\geq k}\Big) = k + V P_{\geq k}(p). \tag{3.5}$$

In addition,

$$|\mathcal{C}_{\max}| = \max\{k : Z_{\geq k} \geq k\}, \tag{3.6}$$

and hence the random variables $\{Z_{\geq k}\}_{k \geq 1}$ provide a characterization of $|\mathcal{C}_{\max}|$.

3.4.2 A useful heuristic

The identity (3.6) suggests that if the distribution of $Z_{\geq k}$ is sufficiently concentrated about its mean, then it should be the case that

$$\mathbb{E}_p(|\mathcal{C}_{\max}|) = \Theta\Big(\max\{k : \mathbb{E}_p(Z_{\geq k}) \geq k\}\Big). \tag{3.7}$$

Define $k_0 = k_0(p)$ to be the solution of the equation

$$k_0 = \mathbb{E}_p(Z_{\geq k_0}) = VP_{\geq k_0}(p). \tag{3.8}$$

Then we are led to expect that

$$\mathbb{E}_p(|\mathcal{C}_{\max}|) = \Theta(k_0(p)). \tag{3.9}$$

Under certain conditions, this heuristic was made rigorous in [14] to analyze percolation on finite subsets of $\mathbb{Z}^n$, $n \leq 6$, and it underlies our approach to obtaining bounds on $|\mathcal{C}_{\max}|$ from bounds on the cluster size distribution $P_{\geq k}(p)$. As a reality check, we note that for the random graph it is not difficult to verify that as $|\epsilon| \to 0$,

$$k_0(p) = \begin{cases} 2\epsilon^{-2}\log(\epsilon^3 V)(1+o(1)) & \text{below the window,} \\ \Theta(V^{2/3}) & \text{inside the window,} \\ 2\epsilon V(1+o(1)) & \text{above the window.} \end{cases} \tag{3.10}$$

To leading order, this is precisely the size of the largest cluster of the random graph, confirming (3.9). Since we are working in settings where random graph scaling should apply, (3.10) also serves as a guide for our more general transitive graphs.

In particular, as noted in [14], if at the critical threshold we have

$$P_{\geq k}(p_c) = \Theta(k^{-1/\delta}), \tag{3.11}$$

then $k_0 = \Theta(V^{\delta/(\delta+1)})$ and (3.9) predicts that $\mathbb{E}_{p_c}(|\mathcal{C}_{\max}|) = \Theta(V^{\delta/(\delta+1)})$. This provides a connection between the critical exponent δ and the size of the largest cluster at criticality. If we assume that $\chi(p_c)$ is well approximated by $\mathbb{E}_{p_c}(|\mathcal{C}_{\max}|)\mathbb{P}_{p_c}(0 \in \mathcal{C}_{\max}) \approx V^{\delta/(\delta+1)}V^{-1+\delta/(\delta+1)}$, it also suggests that the correct definition of the critical threshold, in general, is that value of p for which $\chi(p) = V^{(\delta-1)/(\delta+1)}$. Again, a constant factor λ could be introduced on the right side without significant effect. For a critical branching process, it is the case that $\delta = 2$. For percolation on $\mathbb{Z}^n$ with n sufficiently large, it was proved in [25] that $\delta = 2$ in the sense that $P_k(p_c) = ck^{-3/2}(1 + k^{-a})$ for some $a, c > 0$. On the other hand, it is believed that δ is strictly greater than 2 below the upper critical dimension $n = 6$. Thus we expect that the results of Section 2.2 do not extend to wide tori for $n < 6$, and that our definition of p_c also requires modification in this case, namely in (1.7), the exponent $1/3$ should be replaced by $(\delta - 1)/(\delta + 1)$.

We have in mind a high-dimensional graph $\mathbb{G}$ for which cycles are of limited importance. Since each vertex has Ω neighbors, criticality corresponds to $p\Omega \approx 1$, or $p_c \approx \Omega^{-1}$. According to the above, the value $\delta = 2$ gives the familiar value $V^{2/3}$ for the largest critical cluster. How near to p_c can we expect this behavior to hold, i.e., how wide is the critical window? Let us consider $p < p_c$, which is easier. If $p = p_c - \Omega^{-1}\epsilon$, we expect that a birth process with survival rate $1 - \epsilon$ gives a good approximation, so that

$$P_{\geq k}(p) \approx \frac{\text{const}}{\sqrt{k}} e^{-k\epsilon^2/2}. \tag{3.12}$$

The exponential is unimportant as long as $\epsilon V^{1/3} \leq O(1)$, leading to $P_{\geq k}(p) \approx k^{-1/2}$ and thus $|\mathcal{C}_{\max}| \approx V^{2/3}$. This suggests that the system behaves critically when $\epsilon = O(V^{-1/3})$.

3.4.3 Our method of proof

Our proofs of bounds on $|\mathcal{C}_{\max}|$ proceed as follows. For $p \leq p_c$, we obtain an upper bound on $|\mathcal{C}_{\max}|$ by applying the upper bound

$$P_{\geq k}(p) \leq \sqrt{\frac{e}{k}} e^{-k/(2\chi^2)}, \tag{3.13}$$

which is valid for $k \geq \chi^2(p)$. The bound (3.13) is proved in [4, Proposition 5.1] and [19, (6.77)] (the proofs apply directly to any finite transitive graph). We use (3.13) in conjunction with (3.4)–(3.5), choosing k in accordance with the subcritical case in (3.10). The details are carried out in Section 4.2. For a lower bound on $|\mathcal{C}_{\max}|$, we prove a variance estimate for $Z_{\geq k}$ and use this in conjunction with the second moment method. The details are carried out in Section 7.

Inside the critical window, our bounds on $|\mathcal{C}_{\max}|$ follow directly from monotonicity and the subcritical and supercritical bounds. This is discussed in Section 3.6.

In the supercritical phase, the bounds on $|\mathcal{C}_{\max}|$ of Theorem 1.4 i) follow directly from our upper bounds on $P_{\geq k}$, and are derived in Section 6. To prove the upper bound on $|\mathcal{C}_{\max}|$ stated in Theorem 1.6 ii), we prove another variance estimate for $Z_{\geq k}$. This estimate allows us to bound the probability that $Z_{\geq N_\alpha}$ differs from its expectation $V\theta_\alpha(p)$ by more than a small multiple of $V\theta_\alpha(p)$. The variance of $Z_{\geq N_\alpha}$ is ultimately estimated in terms of the magnetization, and the details are carried out in Section 9. The restriction $\epsilon \geq V^{-\eta}$ in (1.44) (with $\eta \in (\frac{1}{9}, \frac{3}{15})$ for $\alpha \in (0,1)$) means that this upper bound on $|\mathcal{C}_{\max}|$ has not yet been proven for all p above the window.

3.5 The supercritical susceptibility

The magnetization has a useful and standard probabilistic interpretation. We define i.i.d. vertex variables taking the value "green" and "not green" by declaring that each $x \in \mathbb{V}$ is green with probability $\gamma \in [0,1]$. The vertex variables are independent of the bond variables. Let $\mathcal{G}$ denote the random set of green vertices. Then, by definition,

$$M(p,\gamma) = \sum_{k=1}^{V} [1-(1-\gamma)^k] \mathbb{P}_p(|C(0)| = k) = \mathbb{P}_{p,\gamma}(0 \leftrightarrow \mathcal{G}), \tag{3.14}$$

where $\{0 \leftrightarrow \mathcal{G}\}$ denotes the event that $0 \leftrightarrow x$ for some $x \in \mathcal{G}$. Let

$$\chi(p,\gamma) = (1-\gamma)\frac{\partial}{\partial\gamma} M(p,\gamma) = \sum_{k=0}^{V} k(1-\gamma)^k \mathbb{P}_p(|C(0)| = k) = \mathbb{E}_{p,\gamma}\Big(|C(0)| I(0 \not\leftrightarrow \mathcal{G})\Big) \tag{3.15}$$

and

$$\chi_\perp(p,\gamma) = \sum_{k=0}^{V} k[1-(1-\gamma)^k] \mathbb{P}_p(|C(0)| = k) = \mathbb{E}_{p,\gamma}\Big(|C(0)| I(0 \leftrightarrow \mathcal{G})\Big). \tag{3.16}$$

The proof of Theorem 1.4 ii) is based on the decomposition

$$\chi(p) = \chi(p,\gamma) + \chi_\perp(p,\gamma), \tag{3.17}$$

which is valid for all $\gamma \in [0,1]$.

It follows from (1.47)–(1.48) (with $\alpha = 0$ in the latter) that $M(p,\epsilon^2) = \Theta(\epsilon)$ above the window. For the random graph the largest cluster above the window has size of order ϵV, so the origin is in the largest cluster with probability of order ϵ. Thus the probability that the origin is connected to the green set $\mathcal{G}$ and the probability that the origin is in the largest cluster should both be $\Theta(\epsilon)$, when we choose $\gamma = \epsilon^2$. Thus we regard the green set $\mathcal{G}$ as playing the role of a kind of ersatz giant cluster, when $\gamma = \epsilon^2$. From this perspective, $\chi(p,\epsilon^2)$ corresponds to the expected cluster size omitting the giant cluster, whereas $\chi_\perp(p,\epsilon^2)$ corresponds to the expected cluster size of a vertex that is in the largest cluster. Thus we might expect to prove that for $p \geq p_c$, $\chi(p,\epsilon^2)$ is bounded above by $O(\epsilon^{-1})$ while $\chi_\perp(p,\epsilon^2)$ is bounded above by $O(\epsilon^2 V)$. An upper bound on $\chi(p,\gamma)$ will follow easily from our bounds on the magnetization. To obtain a bound of the form $O(\epsilon^2 V)$ for $\chi_\perp(p,\epsilon^2)$, we will make use of the random variable

$$Z_{\mathcal{G}} = \sum_{x \in \mathbb{V}} I(x \leftrightarrow \mathcal{G}), \tag{3.18}$$

which counts the number of vertices in clusters containing at least one green vertex. This will require a differential inequality for the expectation of $Z_{\mathcal{G}}^2$, which is proved in Appendix A.3.

3.6 Proof of Theorem 1.3 ii-iii)

Finally, we show that the bounds of Theorem 1.3 ii-iii) for the critical window follow from the bounds of Theorems 1.2 and 1.4 for the subcritical and supercritical phases.

Proof of (1.22). By the monotonicity of $\chi(p)$ in p, the lower bound follows from the lower bound of (1.15) (with $p = p_c - \Lambda\Omega^{-1}V^{-1/3}$) and the upper bound follows from the upper bound of (1.25) (with $p = p_c + \Lambda\Omega^{-1}V^{-1/3}$). □

Proof of (1.20). The upper bound follows from monotonicity of $\mathbb{E}_p[|\mathcal{C}_{\max}|]$ in p and the upper bound (1.23) (with $p = p_c + \Lambda\Omega^{-1}V^{-1/3}$). The lower bound follows from the lower bounds of (1.16) and (1.22). □

Proof of (1.21). It follows from the upper bound of (1.20) and Markov's inequality that

$$\mathbb{P}_p\Big(|\mathcal{C}_{\max}| \geq \omega V^{2/3}\Big) \leq \frac{b_5}{\omega} \tag{3.19}$$

for all $\omega > 0$. For the complementary bound, we bound $\mathbb{P}_p(|\mathcal{C}_{\max}| \geq \omega^{-1}V^{2/3})$ below by its value at $p = p_c - \Lambda V^{-1/3}$ and apply (1.18) in conjunction with (1.22). □

4 The subcritical phase

In Section 4.1, we apply a differential inequality for $\chi(p)$ due to Aizenman and Newman [4] to show that the triangle condition (1.8) implies the bounds (1.13), (1.14) and (1.15), and that the stronger triangle condition (1.26) implies the bounds (1.30) and (1.31). In Section 4.2, we apply the bound (3.13) on the cluster size distribution, also due to [4], to prove the upper bounds of (1.16)–(1.17).

4.1 The subcritical susceptibility and critical threshold

Recall from (1.12) that $\bar{\nabla}_p = \max_{\{x,y\}\in\mathbb{B}} \nabla_p(x,y)$. In Appendix A.1, we prove the differential inequality

$$[1-\bar{\nabla}_p]\Omega \leq -\frac{d\chi^{-1}}{dp} \leq \Omega, \tag{4.1}$$

which is valid for all $p \in (0,1)$. The differential inequality and its proof are due to Aizenman and Newman [4]. Integration of (4.1) over the interval $[p_1, p_2]$, together with monotonicity of $\bar{\nabla}_p$ in p, gives

$$[1-\bar{\nabla}_{p_2}]\Omega(p_2-p_1) \leq \chi^{-1}(p_1) - \chi^{-1}(p_2) \leq \Omega(p_2-p_1). \tag{4.2}$$

Proof of (1.13) *assuming* (1.8). We set $p_1 = 0$ and $p_2 = p_c$ in (4.2) and note that $\chi(0) = 1$ and $\chi(p_c)^{-1} = \epsilon_0$, to obtain (1.13). □

Proof of (1.14). This follows from (4.2) with p_i defined by $\chi(p_i) = \lambda_i V^{1/3}$. □

Proof of (1.15). This follows from (4.2) with $p_1 = p$ and $p_2 = p_c$. □

Proof of (1.31). The lower bound has been proved already in (1.15). For the upper bound, we first observe that the stronger triangle condition (1.26) implies (1.8) with $a_0 = K_1\Omega^{-1} + K_2\lambda^3$. For $p \leq p_c$, we may therefore use the upper bound of (1.15) to see that

$$\nabla_p(x,y) \leq K_1\Omega^{-1} + K_2\frac{1}{V}\Big(\frac{1}{\epsilon_0 + (1-a_0)\epsilon}\Big)^3 \tag{4.3}$$

for $x \neq y$. We now integrate the lower bound of (4.1) over the interval $[p, p_c]$, using (4.3) to bound the triangle diagram. This gives

$$\begin{aligned}
\chi^{-1}(p) - \epsilon_0 &\geq \int_0^{\epsilon} d\tilde{\epsilon}\Big[1 - K_1\Omega^{-1} - K_2\frac{1}{V}\Big(\frac{1}{\epsilon_0 + (1-a_0)\tilde{\epsilon}}\Big)^3\Big] \\
&= \epsilon(1-K_1\Omega^{-1}) - \frac{K_2}{2V(1-a_0)}\Big[\frac{1}{\epsilon_0{}^2} - \Big(\frac{1}{\epsilon_0 + (1-a_0)\epsilon}\Big)^2\Big] \\
&= \epsilon(1-K_1\Omega^{-1}) - \frac{K_2}{2\epsilon_0{}^2V(1-a_0)}\Big[1 - \frac{1}{1+(1-a_0)\frac{\epsilon}{\epsilon_0}}\Big]\Big[1 + \frac{1}{1+(1-a_0)\frac{\epsilon}{\epsilon_0}}\Big] \\
&\geq \epsilon(1-K_1\Omega^{-1}) - \frac{K_2}{\epsilon_0{}^2V(1-a_0)}\Big[1 - \frac{1}{1+(1-a_0)\frac{\epsilon}{\epsilon_0}}\Big] \\
&= \epsilon\Big(1 - K_1\Omega^{-1} - K_2\lambda^3\frac{1}{1+(1-a_0)\epsilon/\epsilon_0}\Big).
\end{aligned} \tag{4.4}$$

The upper bound in (1.31) is equivalent to (4.4). □

Proof of (1.30). The lower bound of (1.30) was proved already in (1.13). The upper bound follows from (1.31) with $p = 0$, using the lower bound of (1.13) to bound $\epsilon = \Omega p_c$ in (1.29). □

4.2 Upper bound on the largest subcritical cluster

Proof of the upper bound of (1.16) *and of* (1.17). We will prove that

$$\mathbb{E}_p\Big(|\mathcal{C}_{\max}|\Big) \leq 2\chi^2(p)\log(V/\chi^3(p)) \tag{4.5}$$

and

$$\mathbb{P}_p\Big(|\mathcal{C}_{\max}| \le 2\chi^2(p)\log(V/\chi^3(p))\Big) \ge 1 - \frac{\sqrt{e}}{[2\log(V/\chi^3(p)]^{3/2}}, \tag{4.6}$$

if $\chi(p) \le e^{-2}V^{1/3}$. The desired bounds follow immediately from (4.5) and (4.6), provided $\lambda = V^{-1/3}\chi(p_c) \le e^{-2}$. However, it follows from (1.9) and our assumptions $a_0 \le b_0$ and $V^{-1/3} \le \lambda b_0$ that $\lambda^3 \le b_0 + \lambda^3 b_0^3$, which gives $\lambda^3 \le b_0(1-b_0^3)^{-1} \le e^{-2}$.

To prove (4.6), we let $A = \log(V/\chi^3(p))$ and $k = 2A\chi^2(p)$. By assumption, $A \ge 6 \ge 1/2$, and hence $k \ge \chi^2(p)$. We can therefore apply (3.4) and (3.13) to obtain

$$\mathbb{P}_p\Big(|\mathcal{C}_{\max}| \ge 2A\chi^2(p)\Big) \le \frac{VP_{\ge 2A\chi^2(p)}}{2A\chi^2(p)} \le \frac{V\sqrt{e}}{(2A)^{3/2}\chi^3(p)}e^{-A} = \frac{\sqrt{e}}{(2A)^{3/2}}, \tag{4.7}$$

which is the desired bound (4.6).

To prove (4.5), we set $k = 2(A-1)\chi^2(p)$. Combining (3.5) and (3.13) leads to

$$\begin{aligned}\mathbb{E}_p\Big(|\mathcal{C}_{\max}|\Big) &\le 2(A-1)\chi^2(p)\Big[1 + \frac{V\sqrt{e}}{(2(A-1))^{3/2}\chi^3(p)}e^{-A+1}\Big] \\ &= 2Y\chi^2(p)\log(V/\chi^3(p))\end{aligned} \tag{4.8}$$

with

$$Y = \Big(1 - \frac{1}{A}\Big)\Big[1 + \Big(\frac{e}{2(A-1)}\Big)^{3/2}\Big]. \tag{4.9}$$

To complete the proof, it suffices to show that

$$\Big(\frac{e}{2(A-1)}\Big)^{3/2} \le \frac{1}{A}, \tag{4.10}$$

since this implies that $Y \le 1$. To prove (4.10), we use the monotonicity of the function $x \mapsto (x-1)^3/x^2$ and the fact that $A \ge 6$ to conclude that

$$\frac{8(A-1)^3}{A^2} \ge \frac{1000}{36} \ge e^3. \tag{4.11}$$

□

5 The magnetization

In this section, we prove Theorem 1.7. This theorem provides upper and lower bounds on the magnetization, which is defined by

$$M(p,\gamma) = \sum_{k=1}^{V}[1-(1-\gamma)^k]\mathbb{P}_p(|C(0)| = k). \tag{5.1}$$

For fixed p, the function $M(p,\cdot)$ is strictly increasing, with $M(p,0) = 0$ and $M(p,1) = 1$. We denote the inverse function by $\gamma(m)$, so that $M(p,\gamma(m)) = m$ for all $m \in [0,1]$. In addition, for $\gamma \in (0,1)$, $M(p,\gamma)$ is strictly increasing in p. Finally, recalling (3.15), we note that $\partial M/\partial\gamma = (1-\gamma)^{-1}\chi$ is monotone decreasing in γ. Since $M(p,0) = 0$ this implies that

$$\frac{\gamma}{1-\gamma}\chi(p,\gamma) \le M(p,\gamma) \le \gamma\chi(p,0). \tag{5.2}$$

5.1 Bounds on the magnetization

We formulate our results in the general setting of a connected transitive graph $\mathbb{G}$ with V vertices and degree Ω, not necessarily obeying the triangle condition (1.8). Instead, we will assume that one or several of the following conditions hold:

$$p_c \leq a_1, \tag{5.3}$$
$$\Omega p_c \leq 1 + a_2, \tag{5.4}$$
$$\Omega p_c \geq 1 - a_3, \tag{5.5}$$

and last but not least, the triangle condition (1.8) itself. The constants a_0, a_1, a_2 and a_3 in the following statements refer to these assumptions, and when a constant is not mentioned in a theorem, the corresponding assumption is not used.

Note that when we do assume the triangle condition, then the assumptions (5.3)–(5.5) all follow, provided V is large enough. To see this, we note that for any bond $\{x, y\} \in \mathbb{B}$, we have $p \leq \tau_p(x, y) \leq \nabla_p(x, y)$ (just take $u = v = x$ in (1.1)), and hence

$$p_c \leq a_0 \tag{5.6}$$

whenever the triangle condition (1.8) holds. In addition, (5.4)–(5.5) follow from (1.13). Therefore, in particular, the constants a_1, a_2 and a_3 can be made as small as desired by assuming that a_0 and $\epsilon_0 = \lambda^{-1}V^{-1/3}$ are sufficiently small (as assumed in the theorems in Section 1.3).

The following propositions and corollaries immediately imply Theorem 1.7. The first pair gives lower bounds on the magnetization, and the second pair gives upper bounds. For Corollary 5.2, we recall that $N_\alpha = \epsilon^{-2}(\epsilon V^{1/3})^\alpha$ was defined in (1.42).

Proposition 5.1. *(i) Let $0 < p < 1$ and $0 < \gamma < 1$, and let $K = 1 + \frac{\Omega p}{1-p}$. Then*

$$M(p, \gamma) \geq \frac{1}{2K}\left[\sqrt{4K\gamma + \chi^{-2}(p)} - \chi^{-1}(p)\right], \tag{5.7}$$

so that in particular

$$M(p, \gamma) \geq \frac{\sqrt{4K+1} - 1}{2K} \min\{\sqrt{\gamma}, \gamma\chi(p)\}. \tag{5.8}$$

(ii) If $0 < p_0 \leq p < 1$, $0 < \gamma_0 \leq \gamma < 1$ and $0 < \tilde{\alpha} < 1$, then

$$M(p, \gamma) \geq \min\left\{\left(\frac{\gamma}{\gamma_0}\right)^{\tilde{\alpha}} M(p_0, \gamma_0), \frac{p_0}{p}M(p_0, \gamma_0) + (1 - \tilde{\alpha})\frac{p - p_0}{p}\right\}. \tag{5.9}$$

Corollary 5.2. *Assume that a_1 and a_2 are sufficiently small.*
i) If $0 \leq \gamma \leq 1$ and $p \leq p_c$, then

$$M(p, \gamma) \geq \frac{1}{3}\min\{\sqrt{\gamma}, \gamma\chi(p)\}. \tag{5.10}$$

ii) Let $0 \leq \alpha < 1$, $\tilde{\alpha} = (2 - \alpha)^{-1}$, $\rho > 0$, and $p = p_c + \Omega^{-1}\epsilon$. Let $b_{13} = \lambda^{-1-\alpha\tilde{\alpha}}\rho^{-\tilde{\alpha}}$. If $b_{13}V^{-1/3} \leq \epsilon \leq 1$ then

$$M(p, \rho N_\alpha^{-1}) \geq \frac{\epsilon}{3}\min\{(1 - \tilde{\alpha}), \rho^{\tilde{\alpha}}\lambda^{\alpha\tilde{\alpha}}\}. \tag{5.11}$$

Lemma 5.3. *If a_0 and a_3 are sufficiently small, $p \le p_c$ and $0 \le \gamma \le 1$, then*

$$M(p,\gamma) \le \min\{\sqrt{12\gamma},\, \gamma\chi(p)\}. \tag{5.12}$$

Proposition 5.4. *If a_0 and a_3 are sufficiently small, $p = p_c + \Omega^{-1}\epsilon \ge p_c$ and $0 \le \gamma \le 1$, then*

$$M(p,\gamma) \le \sqrt{12\gamma} + 13\epsilon. \tag{5.13}$$

Proof of Theorem 1.7. This is an immediate consequence of Corollary 5.2, Lemma 5.3 and Proposition 5.4. □

Note that for $p \le p_c$ the lower bound (5.10) and the upper bound (5.12) differ only by a constant, for all γ. For $p \ge p_c$, our results are much weaker: If we specialize to γ proportional to N_α^{-1}, and assume that $\epsilon V^{1/3}$ is large enough (in particular, this implies that $N_\alpha^{-1} \le \epsilon^2$), then our lower and upper bounds (5.11) and (5.13) match.

Our bounds on the magnetization are proved using the three differential inequalities stated in the next lemma.

Lemma 5.5. *If $0 < p < 1$ and $0 < \gamma < 1$, then*

$$(1-p)\frac{\partial M}{\partial p} \le \Omega(1-\gamma)M\frac{\partial M}{\partial \gamma}, \tag{5.14}$$

$$M \le \gamma\frac{\partial M}{\partial \gamma} + M^2 + pM\frac{\partial M}{\partial p}, \tag{5.15}$$

and

$$M \ge \left[\binom{\Omega}{2}p^2(1-p)^{\Omega-2}(1-\nabla_p^{\max})^3 - p - \nabla_p^{\max}\right]p\Omega(1-\gamma)M^2\frac{\partial M}{\partial \gamma}, \tag{5.16}$$

where $\nabla_p^{\max} = \max_{x,y\in\mathbb{V}} \nabla_p(x,y)$.

The differential inequalities (5.14)–(5.15) were derived and used by Aizenman and Barsky [2] to prove sharpness of the percolation phase transition on $\mathbb{Z}^n$, and will be used to prove our lower bounds on $M(p,\gamma)$. The derivations in [2] extend without difficulty to an arbitrary transitive graph. The differential inequality (5.16), which is a variant of an inequality derived by Barsky and Aizenman [7], will be used to prove our upper bounds on $M(p,\gamma)$. We give a proof of (5.16) in Appendix A.2.

5.2 Lower bounds on the magnetization

In this section, we prove Proposition 5.1 and Corollary 5.2, using the first two differential inequalities of Lemma 5.5.

Proof of Proposition 5.1. (i) We fix $p \in (0,1)$, and drop the p dependence from the notation. Inserting (5.14) into (5.15), and using $\tilde{K} = \frac{\Omega p}{1-p}$ and $1-\gamma \le 1$, we get

$$M \le \gamma\frac{dM}{d\gamma} + M^2 + \tilde{K}M^2\frac{dM}{d\gamma}. \tag{5.17}$$

Since $M > 0$ as long as $\gamma > 0$, we get

$$\frac{1}{M}\frac{d\gamma}{dM} - \frac{1}{M^2}\gamma \le \tilde{K} + \frac{d\gamma}{dM}, \tag{5.18}$$

where we are using the fact that M has a well-defined inverse function. Therefore,

$$\frac{d}{dM}\left(\frac{\gamma}{M}\right) \le \tilde{K} + \frac{d\gamma}{dM}. \tag{5.19}$$

Next we integrate (5.19) and use that $\gamma(0) = 0$ and $\lim_{M\to 0}\frac{\gamma(M)}{M} = \gamma'(0) = 1/M'(0) = \chi^{-1}(p)$ to get

$$\frac{\gamma}{M} \le \chi^{-1} + \tilde{K}M + \gamma \tag{5.20}$$

where we used the shorthand χ^{-1} for $\chi^{-1}(p)$. Observing that $1-(1-\gamma)^k \ge 1-(1-\gamma) = \gamma$, we see from (5.1) that $\gamma \le M$, which simplifies (5.20) to

$$\frac{\gamma}{M} \le \chi^{-1} + KM, \tag{5.21}$$

where $K = \tilde{K} + 1$. Multiplying by M/K and completing the square on the right side, we thus obtain

$$\frac{\gamma}{K} + \left[\frac{\chi^{-1}}{2K}\right]^2 \le \left[M + \frac{\chi^{-1}}{2K}\right]^2. \tag{5.22}$$

Since $M \ge 0$, this implies that

$$M \ge \sqrt{\frac{\gamma}{K} + \left[\frac{\chi^{-1}}{2K}\right]^2} - \frac{\chi^{-1}}{2K}. \tag{5.23}$$

This completes the proof of (5.7).

To prove (5.8), let us first assume that $\gamma \ge \chi^{-2}(p)$. By (5.7) and the fact that the function $f(x) = \frac{1}{\sqrt{x}}(\sqrt{x+\chi^{-2}} - \chi^{-1})$ is increasing, we conclude that

$$\begin{aligned} M(p,\gamma) &\ge \sqrt{\frac{\gamma}{K}}f(4K\gamma) \\ &\ge \sqrt{\frac{\gamma}{K}}f(4K\chi^{-2}(p)) = \frac{\sqrt{4K+1}-1}{2K}\sqrt{\gamma} \\ &= \frac{\sqrt{4K+1}-1}{2K}\min\{\sqrt{\gamma}, \gamma\chi(p)\}. \end{aligned} \tag{5.24}$$

On the other hand, if $\gamma \le \chi^{-2}(p)$, we use the fact that the function $g(x) = \frac{1}{x}(\sqrt{x+\chi^{-2}} - \chi^{-1})$ is decreasing, together with the bound $M(p,\gamma) \ge 2\gamma g(4K\gamma)$ of (5.7), to arrive at the same conclusion. This completes the proof of (i).

(ii) The result is immediate if $\gamma_0 = \gamma$ or $p_0 = p$ so we assume that $\gamma_0 < \gamma$ and $p_0 < p$. We rewrite (5.15) as

$$0 \le \frac{1}{M}\frac{\partial M}{\partial \gamma} + \frac{1}{\gamma}\frac{\partial}{\partial p}(pM - p), \tag{5.25}$$

and then integrate (5.25) over the rectangle $[\gamma_0, \gamma] \times [p_0, p]$. This yields

$$0 \le \int_{p_0}^{p} d\tilde{p} \log\left(\frac{M(\tilde{p},\gamma)}{M(\tilde{p},\gamma_0)}\right) + \int_{\gamma_0}^{\gamma} d\tilde{\gamma}\frac{1}{\tilde{\gamma}}\left(pM(p,\tilde{\gamma}) - p_0 M(p_0,\tilde{\gamma}) - (p-p_0)\right). \tag{5.26}$$

Since

$$0 \le M(p_0,\gamma_0) \le M(\tilde{p},\tilde{\gamma}) \le M(p,\gamma) \tag{5.27}$$

whenever $(\tilde{p},\tilde{\gamma}) \in [\gamma_0,\gamma] \times [p_0,p]$, it follows that

$$0 \le (p-p_0)\log\left(\frac{M(p,\gamma)}{M(p_0,\gamma_0)}\right) + \log\left(\frac{\gamma}{\gamma_0}\right)\left(pM(p,\gamma) - p_0 M(p_0,\gamma_0) - (p-p_0)\right). \tag{5.28}$$

Dividing by $\log(\gamma/\gamma_0)$, we conclude that

$$M(p,\gamma) \ge \frac{p_0}{p}M(p_0,\gamma_0) + \frac{p-p_0}{p}\left[1 - \frac{\log\{M(p,\gamma)/M(p_0,\gamma_0)\}}{\log\{\gamma/\gamma_0\}}\right]. \tag{5.29}$$

If $M(p,\gamma)/M(p_0,\gamma_0) \le (\gamma/\gamma_0)^{\tilde{\alpha}}$, then (5.29) gives

$$M(p,\gamma) \ge \frac{p_0}{p}M(p_0,\gamma_0) + \frac{p-p_0}{p}\Big[1-\tilde{\alpha}\Big]. \tag{5.30}$$

If, on the other hand, $M(p,\gamma)/M(p_0,\gamma_0) \ge (\gamma/\gamma_0)^{\tilde{\alpha}}$, then it is trivially the case that

$$M(p,\gamma) \ge M(p_0,\gamma_0)(\gamma/\gamma_0)^{\tilde{\alpha}}. \tag{5.31}$$

Therefore, as desired,

$$M(p,\gamma) \ge \min\left\{\left(\frac{\gamma}{\gamma_0}\right)^{\tilde{\alpha}} M(p_0,\gamma_0), \frac{p_0}{p}M(p_0,\gamma_0) + (1-\tilde{\alpha})\frac{p-p_0}{p}\right\}. \tag{5.32}$$

□

Proof of Corollary 5.2. (i) The function $K = K(p) = 1 + \Omega p/(1-p)$ is increasing in p, so $K(p) \le K(p_c)$ for $p \le p_c$. Since the function $(\sqrt{4K+1}-1)/2K$ is decreasing in K, for a lower bound we can replace K by $K(p_c)$ in (5.8). Since $K(p_c) \to 2$ as a_1 and a_2 go to zero, (5.10) then follows.

(ii) We apply Proposition 5.1(ii), whose conclusion is repeated above in (5.32), with $p_0 = p_c$, $\gamma = \rho N_\alpha^{-1}$ and $\gamma_0 = {\epsilon_0}^2 = \chi^{-2}(p_c)$. The requirement $\gamma \ge \gamma_0$ for (5.32) is equivalent to our hypothesis that $\epsilon \ge b_{13} V^{-1/3}$. It suffices to show that

$$\left(\frac{\gamma}{\gamma_0}\right)^{\tilde{\alpha}} M(p_c,\gamma_0) \ge \frac{\epsilon}{3}\rho^{\tilde{\alpha}}\lambda^{\alpha\tilde{\alpha}} \tag{5.33}$$

and

$$\frac{p_c}{p}M(p_c,\gamma_0) + (1-\tilde{\alpha})\frac{p-p_c}{p} \ge \frac{\epsilon}{3}(1-\tilde{\alpha}). \tag{5.34}$$

For (5.33), we use (5.7) and the observation in the proof of part (i) to see that

$$M(p_c, \gamma_0) \geq \frac{\chi^{-1}(p_c)}{2K}\left[\sqrt{4K+1} - 1\right] \geq \frac{1}{3}\epsilon_0 \tag{5.35}$$

if a_1 and a_2 are sufficiently small. Since $(\gamma/\gamma_0)^{\tilde{\alpha}} = \rho^{\tilde{\alpha}}\lambda^{\alpha\tilde{\alpha}}(\epsilon/\epsilon_0)$, (5.33) follows. For (5.34), we bound the first term on the left side below by zero, and note that

$$\frac{p-p_c}{p} = \frac{\epsilon}{\epsilon + \Omega p_c} \geq \frac{\epsilon}{3}, \tag{5.36}$$

since $\epsilon \leq 1$ by assumption and $\Omega p_c \leq 2$ if a_2 is small enough. □

5.3 Upper bounds on the magnetization

We now prove Lemma 5.3 and Proposition 5.4. Lemma 5.3 is proved by integration of the differential inequality (5.15), assuming the triangle condition. We then use the extrapolation principle of [2, 3, 7] to convert the upper bound on $M(p_c, \gamma)$ to an upper bound valid for $p > p_c$. This is perhaps surprising, since M is an increasing function of p. However, it is also increasing in γ, and we will see that it is possible to use the differential inequality (5.14) to compensate for an increase in p with a decrease in γ.

Proof of Lemma 5.3. We first note that $M(p, \gamma) \leq \gamma\chi(p)$ for all p and γ, by (5.2). Since $M(\cdot, \gamma)$ is increasing, it suffices to prove that

$$M(p_c, \gamma) \leq \sqrt{12\gamma}. \tag{5.37}$$

We assume that the triangle condition is satisfied and that (5.5) holds for a sufficiently small constant a_3. Under these conditions, $1 - a_3 \leq \Omega p_c \leq (1-a_0)^{-1}$, $p_c \leq a_0$ and $\Omega = \Omega p_c / p_c \geq (1-a_3)/a_0$, so (5.16) implies that

$$M(p_c, \gamma) \geq \frac{1}{2e}\Big[1 - O(a_0 \vee a_3)\Big](1-\gamma)\frac{dM(p_c,\gamma)}{d\gamma}M(p_c,\gamma)^2. \tag{5.38}$$

For $p = p_c$, this gives

$$\frac{1}{2}\frac{dM^2}{d\gamma} \leq \frac{2e}{1-\gamma}\Big[1 - O(a_0 \vee a_3)\Big]. \tag{5.39}$$

We integrate (5.39) over the interval $[0, \gamma]$, using $M(p_c, 0) = 0$, to see that

$$M^2(p_c, \gamma) \leq \frac{4e\gamma}{1-\gamma}\Big[1 - O(a_0 \vee a_3)\Big]. \tag{5.40}$$

For $\gamma \in [0, \frac{1}{12}]$, this implies (5.37), provided a_0 and a_3 are sufficiently small. Finally, we note that we can remove the restriction $\gamma \in [0, \frac{1}{12}]$, since trivially, $M(p, \gamma) \leq 1 \leq \sqrt{12\gamma}$ if $\gamma \geq \frac{1}{12}$. □

Proof of Proposition 5.4. Following [7], we apply the extrapolation principle used in [2], to extend (5.37) to (5.13). The extrapolation principle is explained in [3]. In our setting, the finite size effect will need to be taken into account. We find it most convenient to use the variable h rather than γ, and define $\tilde{M}(p, h) = M(p, 1 - e^{-h})$, for $h \geq 0$.

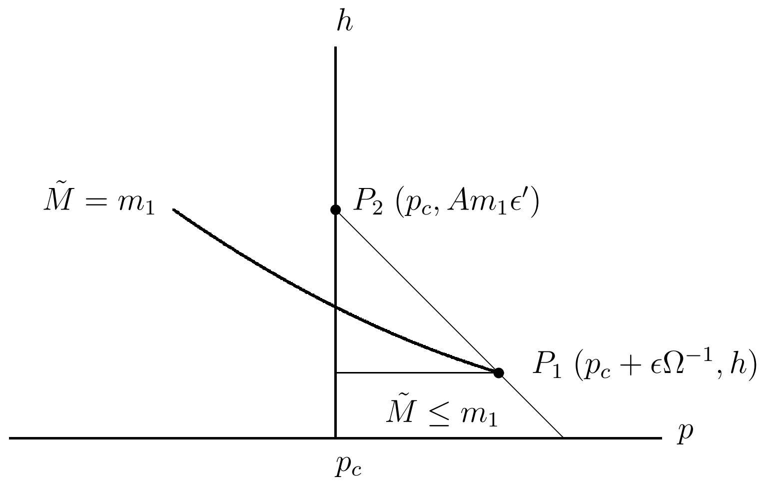

Figure 1: The extrapolation geometry.

Assuming that $\epsilon \le 1$, the differential inequality (5.14) implies that

$$\frac{\partial \tilde{M}}{\partial p} \le A\Omega \tilde{M} \frac{\partial \tilde{M}}{\partial h} \tag{5.41}$$

where $A = (1 - p_c - \Omega^{-1})^{-1} = 1 + O(a_0) + O(a_3)$. For fixed $m \in [0,1]$ and fixed $p \in (0,1)$, we can solve the equation $\tilde{M}(p,h) = m$ for $h = h(p)$, so that $\tilde{M}(p, h(p)) = m$. Differentiation of this identity with respect to p gives

$$\frac{\partial \tilde{M}}{\partial p} + \frac{\partial \tilde{M}}{\partial h} \left. \frac{\partial h}{\partial p} \right|_{\tilde{M}=m} = 0. \tag{5.42}$$

Therefore,

$$0 \le - \left. \frac{\partial h}{\partial p} \right|_{\tilde{M}=m} = \frac{\frac{\partial \tilde{M}}{\partial p}}{\frac{\partial \tilde{M}}{\partial h}} \le A\Omega m. \tag{5.43}$$

The upper and lower bounds of (5.43) imply that a contour line $\tilde{M} = m_1$ in the (p,h)-plane (with p-axis horizontal and h-axis vertical) passing through a point $P_1 = (p_1, h_1)$ is such that $\tilde{M}(P) \le m_1$ for all points P in the first quadrant that lie on or below the lines of slope 0 and $-A\Omega m_1$ through P_1; see Figure 1. In addition, if $P_2 = (p_2, h_2)$ is on the line through P_1 with slope $-A\Omega m_1$, with $p_2 < p_1$, and if we set $m_2 = \tilde{M}(P_2)$, then $m_2 \ge m_1$. For if to the contrary $m_2 < m_1$, then P_1 would lie below the line through P_2 of slope $-A\Omega m_2$, which would imply that $m_1 \le m_2$, a contradiction. We will use the fact that $m_2 \ge m_1$ below.

Fix h, and fix $\epsilon > 0$. Let $P_1 = (p_c + \epsilon\Omega^{-1}, h)$, and define $m_1 = m_1(\epsilon) = \tilde{M}(P_1)$. Let

$$\epsilon' = \epsilon + \frac{h}{Am_1}, \tag{5.44}$$

and define $P_2 = (p_c, Am_1\epsilon')$ and $m_2 = \tilde{M}(P_2)$. The points P_1 and P_2 are collinear on the line through P_1 with slope $-A\Omega m_1$. Therefore, as observed above, $m_1 \le m_2$. Applying (5.37) gives

$$\begin{aligned}\tilde{M}(p_c + \epsilon\Omega^{-1}, h) &= m_1 \le m_2 = \tilde{M}(p_c, Am_1\epsilon') \le \sqrt{12}(1 - e^{-Am_1\epsilon'})^{1/2} \\ &= \sqrt{12}\Big(1 - e^{-Am_1\epsilon} + e^{-Am_1\epsilon}[1 - e^{-h}]\Big)^{1/2} \\ &\le \sqrt{12}(Am_1\epsilon + \gamma)^{1/2}, \end{aligned} \tag{5.45}$$

with $\gamma = 1 - e^{-h}$. The inequality

$$m_1^2 \le 12(Am_1\epsilon + \gamma) \tag{5.46}$$

has roots

$$m^{\pm} = 6A\epsilon \pm \sqrt{12\gamma + (6A\epsilon)^2}. \tag{5.47}$$

The root m^+ is positive and m^- is negative. Thus we have

$$M(p_c + \epsilon\Omega^{-1}, \gamma) = m_1 \le m^+ \le 6A\epsilon + \sqrt{12\gamma + (6A\epsilon)^2} \le 12A\epsilon + \sqrt{12\gamma}. \tag{5.48}$$

This completes the proof of (5.13), since we can choose A arbitrarily close to 1 by choosing a_0 and a_3 sufficiently small. □

6 The cluster size distribution

In this section, we prove Theorems 1.3 i), 1.4 i) and 1.6 i). The magnetization

$$M(p,\gamma) = \sum_{k=1}^{V} [1 - (1-\gamma)^k] \mathbb{P}_p(|C(0)| = k) \tag{6.1}$$

is a generating function for $\mathbb{P}_p(|C(0)| = k)$. In the spirit of a Tauberian theorem, we will use the bounds on $M(p,\gamma)$ established in Section 5 to obtain bounds on $P_{\ge s}(p)$. We recall the upper bound

$$M(p,\gamma) \le \sqrt{12\gamma} + 13\epsilon, \tag{6.2}$$

proved in (5.13) for all $p \ge p_c$ provided a_0 and a_3 are sufficiently small, and the lower bound

$$M(p,\gamma) \ge \frac{1}{3}\min\{\sqrt{\gamma}, \gamma\chi(p)\}, \tag{6.3}$$

proved in (5.10) for all $p \le p_c$ provided a_1 and a_2 are sufficiently small. The discussion of Section 5.1 shows that the constants a_1, a_2 and a_3 can be made arbitrarily small if $a_0 \le b_0$, $\epsilon_0 \le b_0$ and b_0 is chosen small enough.

The cluster size distribution and magnetization are related by the following lemma.

Lemma 6.1. *Let $p \in [0,1]$, $k > 0$ and $0 \le \gamma, \tilde{\gamma} \le 1$. Then*

$$P_{\ge k}(p) \le \frac{e}{e-1} M(p, k^{-1}), \tag{6.4}$$

$$P_{\ge k}(p) \ge M(p,\gamma) - \frac{\gamma}{\tilde{\gamma}} e^{\tilde{\gamma}k} M(p,\tilde{\gamma}). \tag{6.5}$$

Proof. The bound (6.4) follows immediately from the definition of M and the fact that $1-e^{-1} \le 1-(1-k^{-1})^{\ell}$ whenever $\ell \ge k$.

To prove (6.5), we note that $[1-(1-\gamma)^{\ell}] \le \ell\gamma$. Also, $\ell\tilde{\gamma} \le e^{\ell\tilde{\gamma}} - 1 = e^{\ell\tilde{\gamma}}(1-e^{-\ell\tilde{\gamma}})$, which combined with $e^{-\tilde{\gamma}} \ge 1-\tilde{\gamma}$ gives $\ell\tilde{\gamma} \le e^{\ell\tilde{\gamma}}(1-(1-\tilde{\gamma})^{\ell})$. Therefore,

$$\begin{aligned}
M(p,\gamma) &= \sum_{\ell=1}^{V}(1-(1-\gamma)^{\ell})\,\mathbb{P}_p(|C(0)|=\ell) \\
&\le \gamma\sum_{\ell\le k}\ell\,\mathbb{P}_p(|C(0)|=\ell) + \sum_{\ell\ge k}\mathbb{P}_p(|C(0)|=\ell) \\
&\le \frac{\gamma}{\tilde{\gamma}}\,e^{\tilde{\gamma}k}M(p,\tilde{\gamma}) + P_{\ge k}(p). \qquad (6.6)
\end{aligned}$$

□

We will use (6.2)–(6.3) and Lemma 6.1 to prove the bounds in the following lemma.

Lemma 6.2. *There is a constant b_0 such that the following statements hold provided $\epsilon_0 \le b_0$ and the triangle condition (1.8) is valid with $a_0 \le b_0$.*
i) If $p = p_c + \Omega^{-1}\epsilon \ge p_c$ then

$$P_{\ge k}(p) \le 21\epsilon + 6k^{-1/2}. \qquad (6.7)$$

ii) If $p \le p_c$ then

$$\frac{1}{360}k^{-1/2} \le P_{\ge k}(p) \le 6k^{-1/2}. \qquad (6.8)$$

provided $k \le \frac{1}{3600}\chi(p)^2$ for the lower bound (this assumption is not needed for the upper bound).
iii) If $p = p_c + \epsilon\Omega^{-1}$ and $k \le [100(|\epsilon|+\epsilon_0)]^{-2}$ then

$$\frac{1}{360}k^{-1/2} \le P_{\ge k}(p) \le 6k^{-1/2}. \qquad (6.9)$$

Proof. (i) Inserting (6.2) into (6.4) gives (6.7).
(ii) For the upper bound in (6.8), we take $p \le p_c$ and note that $P_k(p) \le P_k(p_c) \le 6k^{-1/2}$, using monotonicity in the first step and (6.7) in the second. For the lower bound, we apply (6.5) with $\tilde{\gamma} = 1/k$. Since

$$M(p,k^{-1}) \le M(p_c,k^{-1}) \le \sqrt{12/k} \qquad (6.10)$$

by (6.2), (6.5) implies that

$$P_{\ge k}(p) \ge M(p,\gamma) - \sqrt{12k}\gamma e. \qquad (6.11)$$

If $\gamma \ge \chi^{-2}(p)$, then (6.3) implies $M(p,\gamma) \ge \frac{1}{3}\sqrt{\gamma}$, and hence

$$P_{\ge k}(p) \ge \frac{1}{3}\sqrt{\gamma} - \sqrt{12k}\gamma e \ge \frac{1}{3}\sqrt{\gamma}\Big(1-30\sqrt{\gamma k}\Big). \qquad (6.12)$$

The choice $\gamma = \frac{1}{60^2 k}$ gives the lower bound of (6.8).
(iii) To prove the upper bound of (6.9), we note that $k \le [100(|\epsilon|+\epsilon_0)]^{-2}$ implies $|\epsilon| \le \frac{1}{100}k^{-1/2}$. It then follows from (6.4) and (6.2) that

$$P_{\ge k}(p) \le P_{\ge k}(p_c + 0.01k^{-1/2}\Omega^{-1}) \le \frac{e}{e-1}\left[\sqrt{12}+0.13\right]k^{-1/2}, \qquad (6.13)$$

which gives the desired bound. For the lower bound, we note that $P_{\geq k}(p) \geq P_{\geq k}(p_c - |\epsilon|\Omega^{-1})$ and that the condition on k in (6.9) implies the condition on k in (6.8) for $p_c - |\epsilon|\Omega^{-1}$, by the lower bound on the susceptibility in (1.15). Therefore (6.9) follows from (6.8). □

Proof of Theorem 1.3 i). Lemma 6.2 iii) immediately implies the statement of Theorem 1.3 i) with $b_1 = [100(\Lambda + \lambda^{-1})]^{-2}$, $b_2 = 1/360$ and $b_3 = 6$. □

Proof of Theorem 1.4 i). We set $k = V^{2/3}$ in (3.5) and apply (6.7) to obtain, as required,

$$\mathbb{E}_p\Big(|\mathcal{C}_{\max}|\Big) \leq 21\epsilon V + 7V^{2/3}. \tag{6.14}$$

The bound (1.24) then follows from Markov's inequality. □

Recall that the percolation probability θ_α is defined, for $p > p_c$ and $0 < \alpha < 1$, by

$$\theta_\alpha(p) = P_{\geq N_\alpha}(p), \tag{6.15}$$

with $N_\alpha = \epsilon^{-2}(\epsilon V^{1/3})^\alpha$. We now prove the bound (1.43) on $\theta_\alpha(p)$ of Theorem 1.6 i).

Proof of Theorem 1.6 i). (*Upper bound on $\theta_\alpha(p)$.*) If $\alpha > 0$ and $\epsilon V^{1/3} \geq 1$, then $N_\alpha \geq \epsilon^{-2}$ and the upper bound of (1.43) follows from (6.7).

(*Lower bound on $\theta_\alpha(p)$.*) We use (6.5) with $k = N_\alpha$, $\tilde{\gamma} = N_\alpha^{-1}$, and $\gamma = \rho N_\alpha^{-1}$ (with $\rho > 0$ to be chosen below) to obtain

$$\theta_\alpha(p) \geq M(p, \rho N_\alpha^{-1}) - \rho e M(p, N_\alpha^{-1}). \tag{6.16}$$

Let $\tilde{\alpha} = (2-\alpha)^{-1}$. Let $b_9 = \lambda^{-1-\alpha\tilde{\alpha}}\rho^{-\tilde{\alpha}}$, and assume that $\epsilon \geq b_9 V^{-1/3}$. By Corollary 5.2 ii),

$$M(p, \rho N_\alpha^{-1}) \geq \frac{\epsilon}{3} \min\{(1-\tilde{\alpha}), \rho^{\tilde{\alpha}}\lambda^{\alpha\tilde{\alpha}}\}. \tag{6.17}$$

Assuming that $N_\alpha \geq \epsilon^{-2}$, which follows if we also assume $b_9 \geq 1$, it follows from Proposition 5.4 that

$$\rho e M(p, N_\alpha^{-1}) \leq \rho e(\sqrt{12} + 13)\epsilon. \tag{6.18}$$

Therefore,

$$\theta_\alpha(p) \geq \epsilon\Big(\frac{1}{3}\min\{(1-\tilde{\alpha}), \rho^{\tilde{\alpha}}\lambda^{\alpha\tilde{\alpha}}\} - \rho e(\sqrt{12}+13)\Big). \tag{6.19}$$

Since $\alpha < 1$ implies $\tilde{\alpha} < 1$, we can make the ratio of the first to the second term as large as we want by choosing ρ sufficiently small depending on α and λ. This gives the lower bound of (1.43), with b_9 and b_{10} depending on α and λ. □

7 Lower bound on the largest subcritical cluster

In this section, we complete the proof of Theorem 1.2 ii) by proving the lower bound of (1.16), and (1.18). Given $s > 0$, let $Z_{\geq s}$ be the number of vertices that lie in clusters of size s or larger, as defined in (3.2). We will use the bound on the variance of $Z_{\geq s}$ given in the following lemma.

Lemma 7.1. *Let $s > 0$ and $p \in (0,1)$. Then*

$$\mathrm{Var}_p\,[Z_{\geq s}] \leq V\chi(p)\,. \tag{7.1}$$

Proof. Let

$$\chi_{\geq s}(p) = \mathbb{E}_p\left[|C(0)|I[|C(0)| \geq s\right]. \tag{7.2}$$

We will prove that

$$\mathbb{E}_p\Big(Z_{\geq s}^2\Big) \leq \left(\mathbb{E}_p\left[Z_{\geq s}\right]\right)^2 + V\chi_{\geq s}(p)\Big(1 - P_{\geq s}(p)\Big), \tag{7.3}$$

which implies (7.1).

We start by rewriting the expectation of $Z_{\geq s}^2$ as

$$\mathbb{E}_p\left[Z_{\geq s}^2\right] = \sum_{x,y\in\mathbb{V}} \sum_{\substack{S:x\in S,\\ |S|\geq s}} \sum_{\substack{T:y\in T,\\ |T|\geq s}} \mathbb{P}_p\Big(C(x) = S, C(y) = T\Big). \tag{7.4}$$

Next, we observe that $C(x)$ and $C(y)$ must be identical if they are not disjoint. As a consequence, the sum decomposes into two terms: the term

$$\sum_{x,y\in\mathbb{V}} \sum_{\substack{S:x,y\in S,\\ |S|\geq s}} \mathbb{P}_p\Big(C(x) = S\Big) = \sum_{x\in\mathbb{V}} \sum_{\substack{S:x\in S,\\ |S|\geq s}} |S|\mathbb{P}_p\Big(C(x) = S\Big) = V\chi_{\geq s}(p) \tag{7.5}$$

and the term

$$\begin{aligned} &\sum_{x\in\mathbb{V}} \sum_{\substack{S:x\in S,\\ |S|\geq s}} \sum_{y\in\mathbb{V}\setminus S} \sum_{\substack{T:y\in T,\\ |T|\geq s}} \mathbb{P}_p\Big(C(x) = S, C(y) = T\Big) \\ &\quad= \sum_{x\in\mathbb{V}} \sum_{\substack{S:x\in S,\\ |S|\geq s}} \mathbb{P}_p\Big(C(x) = S\Big) \sum_{y\in\mathbb{V}\setminus S} \mathbb{P}_p\Big(|C(y)| \geq s \mid C(x) = S\Big). \end{aligned} \tag{7.6}$$

Denoting the set of all edges which either join two points in S or join a point in S to a point in $\mathbb{V}\setminus S$ by $B_+(S)$, we now rewrite the conditional probability as

$$\mathbb{P}_p\Big(|C(y)| \geq s \mid C(x) = S\Big) = \mathbb{P}_p\Big(|C(y)| \geq s \mid \text{all edges in } B_+(S) \text{ are vacant}\Big). \tag{7.7}$$

By the FKG inequality, (7.7) is bounded above by the unconditioned probability $\mathbb{P}_p\Big(|C(y)| \geq s\Big)$. Therefore, (7.6) is bounded by

$$\begin{aligned} &\sum_{x\in\mathbb{V}} \sum_{\substack{S:x\in S,\\ |S|\geq s}} \mathbb{P}_p\Big(C(x) = S\Big) \sum_{y\in\mathbb{V}\setminus S} \mathbb{P}_p\Big(|C(y)| \geq s\Big) \\ &\quad= V \sum_{\substack{S:0\in S,\\ |S|\geq s}} \Big(V - |S|\Big)\mathbb{P}_p\Big(C(0) = S\Big)\mathbb{P}_p\Big(|C(0)| \geq s\Big) \\ &\quad= \left(\mathbb{E}_p\left[Z_{\geq s}\right]\right)^2 - V\chi_{\geq s}(p)P_{\geq s}(p). \end{aligned} \tag{7.8}$$

The combination of (7.5) and (7.8) proves (7.3) and hence (7.1). □

Proof of (1.18). Let $p \leq p_c$ and $\omega \geq 1$. Assume that $\epsilon_0 \leq b_0$ and that (1.8) holds for some $a_0 \leq b_0$ with b_0 as in Lemma 6.2. We must show that

$$\mathbb{P}_p\Big(|\mathcal{C}_{\max}| \geq \frac{\chi^2(p)}{3600\omega}\Big) \geq \Big(1 + \frac{36\chi^3(p)}{\omega V}\Big)^{-1}. \tag{7.9}$$

By definition, $|\mathcal{C}_{\max}| \geq s$ if and only if $Z_{\geq s} > 0$. By the Cauchy–Schwarz inequality,

$$\mathbb{E}_p\Big[Z_{\geq s}\Big] = \mathbb{E}_p\Big[Z_{\geq s} I[Z_{\geq s} > 0]\Big] \leq \sqrt{\mathbb{E}_p\Big[Z_{\geq s}^2\Big]\mathbb{P}_p\Big(Z_{\geq s} > 0\Big)} \tag{7.10}$$

and thus

$$\mathbb{P}_p\Big(|\mathcal{C}_{\max}| \geq s\Big) = \mathbb{P}_p\Big(Z_{\geq s} > 0\Big) \geq \frac{\Big(\mathbb{E}_p\Big[Z_{\geq s}\Big]\Big)^2}{\mathbb{E}_p\Big[Z_{\geq s}^2\Big]} = (1+x)^{-1} \tag{7.11}$$

where

$$x = \frac{\mathrm{Var}_p\Big[Z_{\geq s}\Big]}{\Big(\mathbb{E}_p\Big[Z_{\geq s}\Big]\Big)^2}. \tag{7.12}$$

By Lemma 7.1, the variance of $Z_{\geq s}$ is bounded by $V\chi(p)$. Combined with (3.3), this gives $x \leq \chi(p)V^{-1}[P_{\geq s}(p)]^{-2}$ and thus

$$\mathbb{P}_p\Big(|\mathcal{C}_{\max}| \geq s\Big) \geq \Big(1 + \frac{\chi(p)}{V[P_{\geq s}(p)]^2}\Big)^{-1}. \tag{7.13}$$

To complete the proof, we note that (7.9) is trivial if $\omega \geq \chi^2(p)/3600$. For $\omega \leq \chi^2(p)/3600$, we chose $s = \chi^2(p)/3600\omega$ and use (6.8) to bound $P_{\geq s}(p)$ from below by $\frac{\sqrt{\omega}}{6\chi(p)}$. This gives (7.9). $\square$

Proof of the lower bound of (1.16). We recall from (1.9) that $\chi^3(p)/V \leq \chi^3(p_c)/V = \lambda^3 \leq a_0 + V^{-1}$. Therefore, by (7.9), we can choose a_0 and V^{-1} sufficiently small that, say,

$$\mathbb{E}_p\Big(|\mathcal{C}_{\max}|\Big) \geq \frac{\chi^2(p)}{3600}\mathbb{P}_p\Big(|\mathcal{C}_{\max}| \geq \frac{\chi^2(p)}{3600}\Big) \geq 10^{-4}\chi^2(p). \tag{7.14}$$

This gives the lower bound of (1.16). $\square$

8 Upper bound on the supercritical susceptibility

In this section, we prove the bound (1.25) of Theorem 1.4 ii), by showing that if a_0 and a_3 are sufficiently small and $p = p_c + \Omega^{-1}\epsilon$ with $0 \leq \epsilon \leq 1$, then

$$\chi(p) \leq 81V^{1/3} + (81\epsilon)^2 V. \tag{8.1}$$

The proof of (8.1) is based on the decomposition

$$\chi(p) = \chi(p,\gamma) + \chi_\perp(p,\gamma) \tag{8.2}$$

discussed in Section 3.5, where

$$\chi(p,\gamma) = \mathbb{E}_{p,\gamma}\Big(|C(0)| I(0 \not\leftrightarrow \mathcal{G})\Big) = \sum_{k=0}^{V} k(1-\gamma)^k \mathbb{P}_p(|C(0)| = k) \tag{8.3}$$

and

$$\chi_\perp(p,\gamma) = \mathbb{E}_{p,\gamma}\Big(|C(0)| I(0 \leftrightarrow \mathcal{G})\Big) = \sum_{k=0}^{V} k[1-(1-\gamma)^k]\mathbb{P}_p(|C(0)| = k). \tag{8.4}$$

For an upper bound on $\chi(p,\gamma)$, it follows from Proposition 5.4 and the lower bound of (5.2) that

$$\chi(p,\gamma) \le \sqrt{\frac{12}{\gamma}} + \frac{13\epsilon}{\gamma}. \tag{8.5}$$

whenever a_0 and a_3 are sufficiently small and $p = p_c + \Omega^{-1}\epsilon \ge p_c$. This gives a bound $O(\epsilon^{-1})$, if we choose γ proportional to ϵ^2. To obtain a bound of the form $O(\epsilon^2 V)$ for $\chi_\perp(p,\epsilon^2)$, we will make use of the random variable $Z_{\mathcal{G}} = \sum_{x\in\mathbb{V}} I(x \leftrightarrow \mathcal{G})$. As a first step, we prove the following two lemmas, which give bounds on $\chi_\perp(p,\gamma)$ that are valid for all p and γ.

Lemma 8.1. *Let $0 \le p \le 1$ and $0 \le \gamma \le 1$. Then*

$$\chi_\perp(p,\gamma) \le \frac{1}{V}\mathbb{E}_{p,\gamma}(Z_{\mathcal{G}}^2) \le VM^2(p,\gamma) + \chi_\perp(p,\gamma), \tag{8.6}$$

Proof. Under the condition that $0 \leftrightarrow \mathcal{G}$, $|C(0)|$ can be bounded by the number of vertices that are connected to a green vertex, so that

$$|C(0)|I(0 \leftrightarrow \mathcal{G}) \le \sum_{x\in\mathbb{V}} I(x \leftrightarrow \mathcal{G}) I(0 \leftrightarrow \mathcal{G}). \tag{8.7}$$

Combined with transitivity and the definition of $Z_{\mathcal{G}}$, this implies the lower bound in (8.6).

To prove the upper bound, we decompose the expectation of $Z_{\mathcal{G}}^2$ as

$$\begin{aligned}\mathbb{E}_{p,\gamma}(Z_{\mathcal{G}}^2) = &\sum_{x,y\in\mathbb{V}} \mathbb{E}_{p,\gamma}[I(x \leftrightarrow \mathcal{G})I(y \leftrightarrow \mathcal{G})I(x \not\leftrightarrow y)] \\ &+ \sum_{x,y\in\mathbb{V}} \mathbb{E}_{p,\gamma}[I(x \leftrightarrow \mathcal{G})I(y \leftrightarrow \mathcal{G})I(x \leftrightarrow y)]. \end{aligned} \tag{8.8}$$

As an upper bound, the three events in the first term can be replaced by $\{x \leftrightarrow \mathcal{G}\} \circ \{y \leftrightarrow \mathcal{G}\}$. We then use the BK inequality (with respect to the joint bond/vertex measure) to bound the first term by

$$\sum_{x,y\in\mathbb{V}} \mathbb{E}_{p,\gamma}[I(x \leftrightarrow \mathcal{G})]\mathbb{E}_{p,\gamma}[I(y \leftrightarrow \mathcal{G})] = V^2M^2(p,\gamma). \tag{8.9}$$

Since the second term can be rewritten as

$$\sum_{x,y\in\mathbb{V}} \mathbb{E}_{p,\gamma}[I(x \leftrightarrow \mathcal{G})I(x \leftrightarrow y)] = \sum_{x\in\mathbb{V}} \mathbb{E}_{p,\gamma}[|C(x)|I(x \leftrightarrow \mathcal{G})] = V\chi_\perp(p,\gamma), \tag{8.10}$$

this proves the upper bound in (8.6). ∎

Lemma 8.2. *Let $0 \le p \le 1$ and $0 \le \gamma \le 1$. Then*

$$\frac{\gamma}{1-\gamma}\mathbb{E}_{p,\gamma}\Big[|C(0)|^2 I(0 \not\leftrightarrow \mathcal{G})\Big] \le \chi_\perp(p,\gamma) \le \gamma\chi^3(p). \tag{8.11}$$

Proof. We first note that

$$\frac{1}{1-\gamma}\mathbb{E}_{p,\gamma}\Big[|C(0)|^2 I(0 \not\leftrightarrow \mathcal{G})\Big] = \sum_{k=1}^{V}(1-\gamma)^{k-1}k^2P_k(p) = \frac{\partial\chi_\perp(p,\gamma)}{\partial\gamma} \tag{8.12}$$

is monotone decreasing in γ. Using this fact and the observation that $\chi_\perp(p,0)=0$, integration over $[0,\gamma]$ gives

$$\gamma\frac{1}{1-\gamma}\mathbb{E}_{p,\gamma}\Big[|C(0)|^2 I(0 \not\leftrightarrow \mathcal{G})\Big] \le \chi_\perp(p,\gamma) \le \gamma\bigg[\frac{1}{1-\gamma}\mathbb{E}_{p,\gamma}\Big[|C(0)|^2 I(0 \not\leftrightarrow \mathcal{G})\Big]\bigg]_{\gamma=0} \tag{8.13}$$

The right hand side is simply $\gamma\mathbb{E}_p[|C(0)|^2]$, which is less than $\gamma\chi(p)^3$ by the tree-graph inequalities [4]. ☐

In Appendix A.3, we use Lemmas 8.1–8.2 to derive the differential inequality

$$\frac{\partial}{\partial p}\mathbb{E}_{p,\gamma}\Big[Z_{\mathcal{G}}^2\Big] \le \frac{3\Omega}{1-p}\frac{1-\gamma}{\gamma}M(p,\gamma)\mathbb{E}_{p,\gamma}\Big[Z_{\mathcal{G}}^2\Big], \tag{8.14}$$

for any $0 \le p \le 1$ and $0 \le \gamma \le 1$. We use this to prove the following lemma, which is the final ingredient needed for the proof of (8.1).

Lemma 8.3. *If a_0 and a_3 are sufficiently small, $p = p_c + \Omega^{-1}\epsilon \ge p_c$ and $0 \le \gamma \le 1$, then*

$$\chi_\perp(p,\gamma) \le 13\gamma V \exp\left\{\frac{3(1-\gamma)}{1-p}\left(\sqrt{12\frac{\epsilon^2}{\gamma}}+13\frac{\epsilon^2}{\gamma}\right)\right\}. \tag{8.15}$$

Proof. We divide (8.14) by the expectation on its right side and integrate over the interval $[p_c,p]$. Since $M(p,\gamma)/(1-p)$ is monotone increasing in p, the right side (after the above division) is bounded by its value at the upper limit p of integration. This leads to

$$\mathbb{E}_{p,\gamma}\Big[Z_{\mathcal{G}}^2\Big] \le \mathbb{E}_{p_c,\gamma}\Big[Z_{\mathcal{G}}^2\Big]\exp\left\{\frac{3\epsilon}{1-p}\frac{1-\gamma}{\gamma}M(p,\gamma)\right\}. \tag{8.16}$$

We will show that

$$\mathbb{E}_{p_c,\gamma}(Z_{\mathcal{G}}^2) \le 13\gamma V^2. \tag{8.17}$$

With (8.6), (8.16) and Proposition 5.4, this gives the desired estimate. To prove (8.17), we combine the bounds of Lemmas 8.1 and 8.2 with Proposition 5.4, to get

$$\mathbb{E}_{p_c,\gamma}(Z_{\mathcal{G}}^2) \le 12\gamma V^2 + \gamma\chi^3(p_c)V = 12\gamma V^2 + \gamma\lambda^3 V^2. \tag{8.18}$$

It suffices to show that $\lambda \le 1$. If $V=1$, this bound is trivial, so let us assume that $V \ge 2$. But in this case we may use the bound (1.9) to conclude that $\lambda \le 1$ whenever $a_0 \le 1/2$. ☐

Proof of Theorem 1.4 ii). It follows from (8.2), (8.5) and (8.15) that

$$\chi(p) \le \frac{1}{\epsilon}\left(\sqrt{12\frac{\epsilon^2}{\gamma}}+13\frac{\epsilon^2}{\gamma}\right) + 13\gamma V\exp\left\{\frac{3}{1-p}\left(\sqrt{12\frac{\epsilon^2}{\gamma}}+13\frac{\epsilon^2}{\gamma}\right)\right\}. \tag{8.19}$$

To estimate the factor $1-p$, we use the bound (5.6) and the definition (5.5) of a_3 to see that

$$1-p = 1-p_c-\Omega^{-1}\epsilon \ge 1-a_0-\frac{a_0}{1-a_3} \tag{8.20}$$

whenever $0 \le \epsilon \le 1$. Setting $\gamma = 52\epsilon^2$, and assuming that a_0 and a_3 are chosen small enough to ensure that $\frac{1}{1-p}(\sqrt{3/13} + 1/4) \le 3/4$, we then get

$$\chi(p) \le \frac{3}{4\epsilon} + (26\epsilon)^2 V e^{9/4} \le \frac{3}{4\epsilon} + (81\epsilon)^2 V. \tag{8.21}$$

Let $\epsilon' = \frac{1}{81}V^{-1/3}$. We distinguish the cases $\epsilon < \epsilon'$ and $\epsilon \ge \epsilon'$. In the first case, we use monotonicity and (8.21) to obtain

$$\chi(p) \le \chi(p_c + \Omega^{-1}\epsilon') \le \frac{3}{4}81V^{1/3} + V^{1/3} \le 81V^{1/3}. \tag{8.22}$$

If $\epsilon \ge \epsilon'$, we use $\epsilon^{-1} \le 81V^{1/3}$ to obtain

$$\chi(p) \le \frac{3}{4}81V^{1/3} + (81\epsilon)^2 V \le 81V^{1/3} + (81\epsilon)^2 V. \tag{8.23}$$

The combination of these two estimates gives (8.1). □

9 Upper bound on the largest supercritical cluster

In this section, we prove the upper bound on the largest supercritical cluster stated in Theorem 1.6 ii). For $p \le p_c$, we used the variance bound Lemma 7.1. For $p \ge p_c$, we will use the following alternate bound on the variance of $Z_{\ge s}$. For its statement, we define

$$\chi_{<s}(p) = \mathbb{E}_p\left[|C(0)|I[|C(0)| < s\right]. \tag{9.1}$$

Lemma 9.1. *Let $s > 0$ and $p \in (0,1)$. Then*

$$\mathrm{Var}_p\left[Z_{\ge s}\right] \le (1 + p\,\Omega s)V\chi_{<s}(p) \le 4s(1 + p\,\Omega s)VM(p, s^{-1}). \tag{9.2}$$

Proof. We define the random variable $Z_{<s} = V - Z_{\ge s} = \sum_{v\in\mathbb{V}} I[|C(v)| < s]$ and express the variance as

$$\mathrm{Var}_p[Z_{\ge s}] = \mathrm{Var}_p[Z_{<s}] = \sum_{v,w,S,T}\Big[\mathbb{P}_p[C(v) = S, C(w) = T] - \mathbb{P}_p[C(v) = S]\mathbb{P}_p[C(w) = T]\Big], \tag{9.3}$$

where the sum is over connected sets S, T with $|S| < s$, $|T| < s$ and $v \in S$, $w \in T$. Let $\mathrm{dist}(\cdot,\cdot)$ denoting the graph distance on $\mathbb{G}$. If $\mathrm{dist}(S,T) > 1$, the above events are independent and so the difference is zero. If $\mathrm{dist}(S,T) = 0$, then $S \cap T \ne 0$, and the first probability is zero unless $S = T$. The corresponding contribution from the first term is just

$$\sum_{\substack{S:|S|<s,\\ v,w\in S}} \mathbb{P}_p[C(v) = S] = V\chi_{<s}(p), \tag{9.4}$$

implying that the contribution from both terms can be bounded by (9.4).

We are left with the contribution of the terms (v, w, S, T) with $\mathrm{dist}(S,T) = 1$. We need some notation. Given a connected set S, let $E(S)$ be the set of all edges with both endpoint in S, and

∂S be the set of edges with exactly one endpoint in S. Finally, let $A(S)$ be the event that $E(S)$ contains a set E of occupied edges such that the graph (S, E) is connected. With this notation, the event $C(v) = S$ is just the intersection of the event $A(S)$ with the event that all edges in ∂S are vacant. Note also, that the two events are independent, so that $\mathbb{P}_p(C(v) = S)$ is the product $\mathbb{P}_p[A(S)](1-p)^{|\partial S|}$.

If $\mathrm{dist}(S, T) = 1$, the events $A(S)$, $A(T)$ and the event that the edges in $\partial S \cup \partial T$ are vacant are independent. For these (S, T), the difference in (9.3) can thus be rewritten as

$$\begin{aligned}\mathbb{P}_p[C(v) = S, &C(w) = T] - \mathbb{P}_p[C(v) = S]\mathbb{P}_p[C(w) = T] \\ &= \mathbb{P}_p[A(S)]\mathbb{P}_p[A(T)]\Big((1-p)^{|\partial S \cup \partial T|} - (1-p)^{|\partial S|}(1-p)^{|\partial T|}\Big) \\ &= \mathbb{P}_p[C(v) = S, C(w) = T]\Big(1 - (1-p)^{|\partial S \cap \partial T|}\Big). \end{aligned} \tag{9.5}$$

To continue, we use the inequality $1 - (1-p)^k \leq pk$ to obtain

$$1 - (1-p)^{-|\partial S \cap \partial T|} \leq p|\partial S \cap \partial T| = p \sum_{x,y:\mathrm{dist}(x,y)=1} I[x \in S]I[y \in T]. \tag{9.6}$$

Combining (9.6) with the identity (9.5), we now bound the contribution to (9.3) due to the terms (v, w, S, T) with $\mathrm{dist}(S, T) = 1$ by

$$\begin{aligned} p &\sum_{\substack{v,w,x,y \\ \mathrm{dist}(x,y)=1}} \mathbb{P}_p\Big(|C(v)| < s, x \in C(v), |C(w)| < s, y \in C(w), C(x) \neq C(y)\Big) \\ &= p \sum_{\substack{x,y,v,w \\ \mathrm{dist}(x,y)=1}} \mathbb{P}_p\Big(|C(x)| < s, v \in C(x), |C(y)| < s, w \in C(y), C(x) \neq C(y)\Big) \\ &= p \sum_{\substack{x,y \\ \mathrm{dist}(x,y)=1}} \mathbb{E}_p\Big(|C(x)|\, I[|C(x)| < s] |C(y)|\, I[|C(y)| < s, x \not\leftrightarrow y]\Big) \\ &\leq ps \sum_{\substack{x,y \\ \mathrm{dist}(x,y)=1}} \mathbb{E}_p\Big(|C(x)|\, I[|C(x)| < s]\Big) \\ &= ps\Omega V \chi_{<s}(p). \end{aligned} \tag{9.7}$$

Combining this term with (9.4), we obtain the first bound of (9.2).

For the second bound of (9.2), it suffices to show that

$$\chi_{<s}(p) \leq 4sM(p, s^{-1}). \tag{9.8}$$

For $s \geq 2$, we bound $\chi(p, s^{-1})$ in (3.15) from below by restricting the sum over k to $k < s$. We then use $(1 - s^{-1})^k \geq (1 - s^{-1})^s \geq 1/4$ to conclude that $\chi_{<s}(p) \leq 4\chi(p, s^{-1})$. Combined with (5.2), this gives (9.8) for $s \geq 2$. If $s \leq 1$, the left side of (9.8) is zero and the bound is trivial. Finally, if $1 < s < 2$ then the left side of (9.8) is 1, whereas it follows from the fact that $M(p, \gamma) \geq \mathbb{P}_{p,\gamma}(0 \in \mathcal{G}) = \gamma$ that the right side of (9.8) is at least 4. $\square$

Proof of Theorem 1.6 ii). Let $p = p_c + \Omega^{-1}\epsilon$. It suffices to prove that under the hypotheses of Theorem 1.6 there are constants b_{11}, b_{12} such that

$$\mathbb{P}_p\Big(|\mathcal{C}_{\max}| \geq [1 + (\epsilon V^\eta)^{-1}]V\theta_\alpha(p)\Big) \leq \frac{b_{11}}{(\epsilon V^\eta)^{3-2\alpha}} \tag{9.9}$$

if $\max\{b_{12}V^{-1/3}, V^{-\eta}\} \leq \epsilon \leq 1$. (The proof actually applies for $b_{12}V^{-1/3} \leq \epsilon \leq 1$, but the result is not meaningful unless $\epsilon \geq V^{-\eta}$.) To prove (9.9), we first note that $V\theta_\alpha(p) \geq N_\alpha$ if $b_{10}[\epsilon V^{1/3}]^{3-\alpha} \geq 1$, by (1.43). To satisfy $b_{10}[\epsilon V^{1/3}]^{3-\alpha} \geq 1$, we will take $b_{12} \geq b_{10}^{-1/(3-\alpha)}$. Thus we have

$$\begin{aligned}\mathbb{P}_p\Big(|\mathcal{C}_{\max}| \geq [1+(\epsilon V^\eta)^{-1}]V\theta_\alpha(p)\Big) &= \mathbb{P}_p\Big(Z_{\geq N_\alpha} \geq |\mathcal{C}_{\max}| \geq [1+(\epsilon V^\eta)^{-1}]V\theta_\alpha(p)\Big)\\ &\leq \mathbb{P}_p\Big(|Z_{\geq N_\alpha} - V\theta_\alpha(p)| \geq (\epsilon V^\eta)^{-1}V\theta_\alpha(p)\Big).\end{aligned} \tag{9.10}$$

It therefore suffices to show that if $\max\{b_{12}V^{-1/3}, V^{-\eta}\} \leq \epsilon \leq 1$ then

$$\mathbb{P}_p\Big(|Z_{\geq N_\alpha} - V\theta_\alpha(p)| \geq (\epsilon V^\eta)^{-1}V\theta_\alpha(p)\Big) \leq \frac{b_{11}}{(\epsilon V^\eta)^{3-2\alpha}}. \tag{9.11}$$

By the variance estimate (9.2) and the triangle condition, $\mathrm{Var}[Z_{\geq s}] \leq 12s^2VM(p,s^{-1})$. Since $M(p,\gamma) \leq \sqrt{12\gamma} + 13\epsilon$ by (5.13), it follows from the fact that $N_\alpha^{-1} \leq \epsilon^2$ that

$$\mathrm{Var}[Z_{\geq N_\alpha}] \leq 200\epsilon N_\alpha^2 V = 200\frac{1}{\epsilon^{3-2\alpha}}V^{1+2\alpha/3}. \tag{9.12}$$

Therefore, by Chebyshev's inequality and (1.43),

$$\mathbb{P}\Big(|Z_{\geq N_\alpha} - V\theta_\alpha| \geq (\epsilon V^\eta)^{-1}V\theta_\alpha\Big) \leq (\epsilon V^\eta)^2\frac{\mathrm{Var}[Z_{\geq N_\alpha}]}{V^2\theta_\alpha^2} \leq \frac{200}{b_{10}^2}\frac{1}{\epsilon^{3-2\alpha}V^{1-2\eta-2\alpha/3}}. \tag{9.13}$$

Since $\eta = \frac{3-2\alpha}{15-6\alpha}$, the important factor on the right side is equal to $(\epsilon V^\eta)^{-(3-2\alpha)}$. This gives (9.11) and completes the proof. ∎

A Appendix: Derivation of differential inequalities

A.1 Differential inequality for the susceptibility

In this section, we prove (3.1), which is restated here as Proposition A.1. We follow the original proof of Aizenman and Newman [4], with a minor extension for the lower bound to deal with an arbitrary transitive graph $\mathbb{G}$. The proof also provides an instructive preliminary to the proof of (5.16) in Appendix A.2.

Proposition A.1. *For all $p \in (0,1)$,*

$$[1-\bar{\nabla}_p]\Omega\chi(p)^2 \leq \frac{d\chi(p)}{dp} \leq \Omega\chi(p)^2. \tag{A.1}$$

Recall that $E \circ F$ denotes the event that E and F occur disjointly. Given a bond configuration, we say that a bond is *pivotal* for $x \leftrightarrow y$ if $x \leftrightarrow y$ in the possibly modified configuration in which the bond is made occupied, whereas x is not connected to y in the possibly modified configuration in which the bond is made vacant.

Proof of the upper bound in (A.1). By Russo's formula (see [19, Theorem (2.25)]),

$$\frac{d}{dp}\tau_p(x,y) = \sum_{\{u,v\}\in\mathbb{B}} \mathbb{P}_p(\{u,v\} \text{ is pivotal for } x \leftrightarrow y). \tag{A.2}$$

Therefore, by the BK inequality,

$$\frac{d}{dp}\tau_p(x,y) \le \sum_{(u,v)} \mathbb{P}_p(\{x \leftrightarrow u\} \circ \{v \leftrightarrow y\}) \le \sum_{(u,v)} \tau_p(x,u)\tau_p(v,y), \tag{A.3}$$

where the sum over (u,v) is a sum over *directed* bonds. We then perform the sums over y, v, u (in that order) and use transitivity to obtain the desired upper bound. □

For the lower bound of (A.1), we will use the following definition and lemmas. In the first lemma, we use transitivity to give an alternate representation for $\sum_{v:\{0,v\}\in\mathbb{B}} \nabla_p(0,v)$. This is related to an issue raised by Schonmann [30] (see also [31]), who pointed out that the use of differential inequalities plus the triangle condition to prove mean-field behavior on general *infinite* transitive graphs can be accomplished under the additional assumption that the graph is unimodular, and that it is an open problem to determine whether the assumption of unimodularity is essential. Finite transitive graphs are always unimodular, so the issue raised in [30] is less relevant for our purposes. In any case, we will bypass the issue altogether by applying the following lemma. For its statement, we define

$$T_1(z) = \sum_{(u,v)} \sum_{y\in\mathbb{V}} \tau_p(z,u)\tau_p(z,y)\tau_p(y,v). \tag{A.4}$$

The equality (A.5) of Lemma A.2 will be applied only in (A.19) and (A.48).

Lemma A.2. *For each $u, z \in \mathbb{V}$,*

$$T_1(z) = \sum_{v:\{u,v\}\in\mathbb{B}} \nabla_p(u,v) \le \Omega\bar{\nabla}_p. \tag{A.5}$$

Proof. The inequality follows from the definition of $\bar{\nabla}_p$ in (1.12). To prove the equality, let

$$T_2(u) = \sum_{v:\{u,v\}\in\mathbb{B}} \nabla_p(u,v) = \sum_{v:\{u,v\}\in\mathbb{B}} \sum_{y,z\in\mathbb{V}} \tau_p(u,z)\tau_p(z,y)\tau_p(y,v). \tag{A.6}$$

We first prove that $T_2(u)$ is independent of u; a similar proof applies for $T_1(z)$. By transitivity, there is a graph automorphism $\varphi = \varphi_u$ such that $\varphi(u) = 0$, where 0 is a fixed vertex. Since $\tau_p(x,y) = \tau_p(\varphi(x),\varphi(y))$,

$$T_2(u) = \sum_{v:\{u,v\}\in\mathbb{B}} \sum_{y,z\in\mathbb{V}} \tau_p(0,\varphi(z))\tau_p(\varphi(z),\varphi(y))\tau_p(\varphi(y),\varphi(v)). \tag{A.7}$$

Since φ is an automorphism, $\{u,v\}\in\mathbb{B}$ if and only if $\{\varphi(u),\varphi(v)\} = \{0,\varphi(v)\}\in\mathbb{B}$. Similarly, as x runs over all vertices, so does $\varphi(x)$. Relabelling $\varphi(y)$ to y, $\varphi(v)$ to v, and $\varphi(z)$ to z, we thus get

$$T_2(u) = \sum_{v:\{0,v\}\in\mathbb{B}} \sum_{y,z\in\mathbb{V}} \tau_p(0,z)\tau_p(z,y)\tau_p(y,v) = T_2(0). \tag{A.8}$$

Since $T_i(x)$ $(i = 1,2)$ is independent of $x \in \mathbb{V}$, it is equal to the average of its sum over $x \in \mathbb{V}$. Since $\sum_{z\in\mathbb{V}} T_1(z) = \sum_{u\in\mathbb{V}} T_2(u)$, this implies that $T_1(z) = T_2(u)$ for all $u, z \in \mathbb{V}$, which proves the equality in (A.5). □

Definition A.3. (a) Given a bond configuration, and $A \subset \mathbb{V}$, we say x and y are *connected in* A, if there is an occupied path from x to y having all its endpoints in A, or if $x = y \in A$. We define a restricted two-point function by

$$\tau^A(x, y) = \mathbb{P}(x \text{ and } y \text{ are connected in } \mathbb{V}\backslash A). \tag{A.9}$$

(b) Given a bond configuration, and $A \subset \mathbb{V}$, we say x and y are *connected through* A, if $x \leftrightarrow y$ and every occupied path connecting x to y has at least one bond with an endpoint in A. This event is written as $x \overset{A}{\leftrightarrow} y$.
(c) Given a bond configuration, and a bond b, we define $\tilde{C}^b(x)$ to be the set of vertices connected to x, in the new configuration obtained by setting b to be vacant.
(d) Given an event E, we define the event $\{E \text{ occurs on } \tilde{C}^{(u,v)}(x)\}$ to be the set of configurations such that E occurs on the modified configuration in which every bond that does not have an endpoint in $\tilde{C}^{(u,v)}(x)$ is made vacant. We say that $\{E \text{ occurs in } \mathbb{V}\backslash\tilde{C}^{(u,v)}(x)\}$ if E occurs on the modified configuration in which every bond that does not have both endpoints in $\mathbb{V}\backslash\tilde{C}^{(u,v)}(x)$ is made vacant.

Lemma A.4. *Fix $p \in [0, 1]$. Given a bond (u, v), a vertex w and events E, F,*

$$\begin{aligned}\mathbb{E}_p \Big(I[E \text{ occurs on } \tilde{C}^{(u,v)}(w) \;\&\; (u, v) \text{ is occupied } \;\&\; F \text{ occurs in } \mathbb{V}\backslash\tilde{C}^{(u,v)}(w)]\Big) \\ = p\mathbb{E}_p \Big(I[E \text{ occurs on } \tilde{C}^{(u,v)}(w)]\, \mathbb{E}_p \Big(I[F \text{ occurs in } \mathbb{V}\backslash\tilde{C}^{(u,v)}(w)]\Big)\Big).\end{aligned} \tag{A.10}$$

The identity (A.10) *is also valid if the event* $\{(u, v)$ *is occupied*$\}$ *is removed from the left side and* p *is removed from the right side.*

The above lemma is present in [4] in an implicit form. Its elementary proof can be found in [12, Lemma 3.2]. In the nested expectation on the right side of (A.10), the set $\tilde{C}^{(u,v)}(w)$ is a random set with respect to the outer expectation, but it is deterministic with respect to the inner expectation. The inner expectation on the right side effectively introduces a second percolation model on a second lattice, which is coupled to the original percolation model via the set $\tilde{C}^{(u,v)}(w)$.
Proof of the lower bound in (A.1). By definition,

$$\begin{aligned}&\{(u, v) \text{ is pivotal for } 0 \leftrightarrow x\} \\ &\quad = \{0 \leftrightarrow u \text{ occurs on } \tilde{C}^{(u,v)}(0)\} \cap \{v \leftrightarrow x \text{ occurs in } \mathbb{V}\backslash\tilde{C}^{(u,v)}(0)\}.\end{aligned} \tag{A.11}$$

By Lemma A.4 and (A.11),

$$\begin{aligned}\mathbb{P}_p((u, v) \text{ is pivotal for } 0 \leftrightarrow x) &= \mathbb{E}\Big(I[0 \leftrightarrow u \text{ occurs on } \tilde{C}^{(u,v)}(0)]\, \tau^{\tilde{C}^{(u,v)}(0)}(v, x)\Big) \\ &= \mathbb{E}\Big(I[0 \leftrightarrow u]\, \tau^{\tilde{C}^{(u,v)}(0)}(v, x)\Big).\end{aligned} \tag{A.12}$$

In the second equality of (A.12), we dropped the condition "occurs on $\tilde{C}^{(u,v)}(0)$," because of the fact that $\tau^{\tilde{C}^{(u,v)}(0)}(v, x) = 0$ on the event $\{0 \leftrightarrow u\} \setminus \{0 \leftrightarrow u \text{ occurs on } \tilde{C}^{(u,v)}(0)\}$. The identity (A.12) can be rewritten as

$$\mathbb{P}_p((u, v) \text{ is pivotal for } 0 \leftrightarrow x) = \tau_p(0, u)\tau_p(v, x) - \mathbb{E}\Big(I[0 \leftrightarrow u]\mathbb{P}_p(v \xleftrightarrow{\tilde{C}^{(u,v)}(0)} x)\Big). \tag{A.13}$$

By the BK inequality, for $A \subset \mathbb{V}$ we have

$$\begin{aligned}\mathbb{P}_p(v \xleftrightarrow{A} x) &= \mathbb{P}_p\Big(\bigcup_{y\in A} \{v \leftrightarrow y\} \circ \{y \leftrightarrow x\}\Big) \\ &\leq \sum_{y \in \mathbb{V}} I[y \in A] \mathbb{P}_p(\{v \leftrightarrow y\} \circ \{y \leftrightarrow x\}) \\ &\leq \sum_{y \in \mathbb{V}} I[y \in A] \tau_p(v,y)\tau_p(y,x). \end{aligned} \tag{A.14}$$

Therefore, for $A = \tilde{C}^{(u,v)}(0) \subset C(0)$, we have

$$\mathbb{P}_p(v \xleftrightarrow{\tilde{C}^{(u,v)}(0)} x) \leq \sum_{y\in\mathbb{V}} I[y \in C(0)]\tau_p(v,y)\tau_p(y,x). \tag{A.15}$$

Substitution yields

$$\mathbb{P}_p((u,v) \text{ is pivotal for } 0 \leftrightarrow x) \geq \tau_p(0,u)\tau_p(v,x) - \sum_{y\in\mathbb{V}} \mathbb{P}_p(0 \leftrightarrow u \leftrightarrow y)\tau_p(v,y)\tau_p(y,x). \tag{A.16}$$

The tree-graph bound [4] (which is an elementary consequence of the BK inequality) implies that

$$\mathbb{P}_p(0 \leftrightarrow u, 0 \leftrightarrow y) \leq \sum_{z\in\mathbb{V}} \tau_p(0,z)\tau_p(z,y)\tau_p(z,u). \tag{A.17}$$

Therefore,

$$\begin{aligned}\mathbb{P}_p((u,v) \text{ is pivotal for } 0 \leftrightarrow x) \geq{}& \tau_p(0,u)\tau_p(v,x) \\ &- \sum_{y,z\in\mathbb{V}} \tau_p(0,z)\tau_p(z,y)\tau_p(z,u)\tau_p(y,v)\tau_p(y,x).\end{aligned} \tag{A.18}$$

Recalling (A.2), and performing the sums over u, v, x, leads to

$$\begin{aligned}\frac{d\chi(p)}{dp} &\geq \Omega\chi(p)^2 - \chi(p) \sum_{z\in\mathbb{V}} \tau_p(0,z) \sum_{(u,v)} \sum_{y\in\mathbb{V}} \tau_p(z,y)\tau_p(z,u)\tau_p(y,v) \\ &= \Omega\chi(p)^2 - \chi(p)\sum_{z\in\mathbb{V}} \tau_p(0,z) T_1(z) \\ &\geq \Omega\chi(p)^2[1-\bar{\nabla}_p], \end{aligned} \tag{A.19}$$

by Lemma A.2. □

A.2 Differential inequality for the magnetization

In this section, we prove the differential inequality (5.16). Our method of proof for (5.16) is related to, but simpler than, the method used in [7] to prove an analogous statement for percolation on $\mathbb{Z}^n$. See [24, Section 3] for related results for $\mathbb{Z}^n$ which are stronger but more difficult to prove. We restate (5.16) as (A.20) in the following lemma. Note that, by (3.15), the factor $(1-\gamma)\partial M/\partial\gamma$ on the right side of (5.16) can be replaced by $\chi(p,\gamma)$.

Lemma A.5. *If* $0 < p < 1$ *and* $0 < \gamma < 1$*, then*

$$M(p,\gamma) \geq \left[\binom{\Omega}{2} p^2(1-p)^{\Omega-2}(1-\nabla_p^{\max})^3 - p - \nabla_p^{\max}\right] p\Omega(1-\gamma)M^2(p,\gamma)\frac{\partial M(p,\gamma)}{\partial\gamma}, \tag{A.20}$$

where $\nabla_p^{\max} = \sup_{x,y\in\mathbb{V}} \nabla_p(x,y)$.

Proof. Recall the use of the "green" set $\mathcal{G}$ discussed in Section 3.5. Let $\{v \Leftrightarrow \mathcal{G}\}$ denote the event that there exist $x, y \in \mathcal{G}$, with $x \neq y$, such that there are disjoint connections $v \leftrightarrow x$ and $v \leftrightarrow y$. Let $F_{(u,v)}$ denote the event that the bond (u,v) is occupied and pivotal for the connection from 0 to $\mathcal{G}$, with $\{v \Leftrightarrow \mathcal{G}\}$. Let $F = \cup_{(u,v)} F_{(u,v)}$, and note that the union is disjoint. Since $0 \leftrightarrow \mathcal{G}$ when F occurs, $M = \mathbb{P}(0 \leftrightarrow \mathcal{G}) \geq \mathbb{P}(F)$, and it suffices to prove that $\mathbb{P}(F)$ is bounded below by the right side of (A.20).

For $x, y \in \mathbb{V}$, we define a "green-free" analogue of the two-point function by

$$\tau_{p,\gamma}(x,y) = \mathbb{P}_{p,\gamma}(x \leftrightarrow y, x \not\leftrightarrow \mathcal{G}), \tag{A.21}$$

so that

$$\chi(p,\gamma) = \sum_{x\in\mathbb{V}} \tau_{p,\gamma}(0,x) \tag{A.22}$$

and $\chi(p,0) = \chi(p)$. Given a subset $A \subset \mathbb{V}$, we define

$$\tau_{p,\gamma}^{A}(x,y) = \mathbb{P}_{p,\gamma}((x \leftrightarrow y, x \not\leftrightarrow \mathcal{G}) \text{ in } \mathbb{V}\backslash A). \tag{A.23}$$

When $\gamma \neq 0$, we extend the definition of "occurs in" and "occurs on" in Definition A.3 as in [24, Definition 2.2]. In particular, we now say that E occurs on A if, given a configuration, E occurs on the new configuration obtained by setting all bonds not touching A to be vacant and all vertices not in A to be not green. By definition of $F_{(u,v)}$, it can be seen by conditioning on the set $\tilde{C}^{(u,v)}(v)$ that

$$\mathbb{P}_{p,\gamma}(F) = p\sum_{(u,v)} \mathbb{P}_{p,\gamma}\left[(0 \leftrightarrow u \;\&\; 0 \not\leftrightarrow \mathcal{G}) \text{ in } \mathbb{V}\backslash\tilde{C}^{(u,v)}(v), \;\; v \Leftrightarrow \mathcal{G} \text{ on } \tilde{C}^{(u,v)}(v)\right]. \tag{A.24}$$

It then follows from [24, Lemma 2.4] (a straightforward extension of Lemma A.4 to allow for the presence of a magnetic field) that

$$\mathbb{P}_{p,\gamma}(F) = p\sum_{(u,v)} \mathbb{E}_{p,\gamma}\left[\tau_{p,\gamma}^{\tilde{C}^{(u,v)}(v)}(0,u) I[v \Leftrightarrow \mathcal{G} \text{ on } \tilde{C}^{(u,v)}(v)]\right]. \tag{A.25}$$

We use the identities

$$\tau_{p,\gamma}^{\tilde{C}^{(u,v)}(v)}(0,u) = \tau_{p,\gamma}(0,u) - \left(\tau_{p,\gamma}(0,u) - \tau_{p,\gamma}^{\tilde{C}^{(u,v)}(v)}(0,u)\right) \tag{A.26}$$

and

$$I[v \Leftrightarrow \mathcal{G} \text{ on } \tilde{C}^{(u,v)}(v)] = I[v \Leftrightarrow \mathcal{G}] - \left(I[v \Leftrightarrow \mathcal{G}] - I[v \Leftrightarrow \mathcal{G} \text{ on } \tilde{C}^{(u,v)}(v)]\right). \tag{A.27}$$

Recalling (A.22), it follows that

$$\begin{aligned} P_{p,\gamma}(F) = \; & p\Omega\chi(p,\gamma)\mathbb{P}_{p,\gamma}(0 \Leftrightarrow \mathcal{G}) - p\sum_{(u,v)} \tau_{p,\gamma}(0,u)\mathbb{E}_{p,\gamma}\left[I[v \Leftrightarrow \mathcal{G}] - I[v \Leftrightarrow \mathcal{G} \text{ on } \tilde{C}^{(u,v)}(v)]\right] \\ & - p\sum_{(u,v)} \mathbb{E}_{p,\gamma}\left[\left(\tau_{p,\gamma}(0,u) - \tau_{p,\gamma}^{\tilde{C}^{\{u,v\}}(v)}(0,u)\right) I[v \Leftrightarrow \mathcal{G} \text{ on } \tilde{C}^{(u,v)}(v)]\right]. \end{aligned} \tag{A.28}$$

We write (A.28) as $X_1 - X_2 - X_3$, bound X_1 from below, and bound X_2 and X_3 from above.

Lower bound on X_1. We will prove that

$$\mathbb{P}_{p,\gamma}(0 \Leftrightarrow \mathcal{G}) \geq \binom{\Omega}{2} p^2(1-p)^{\Omega-2} M^2(p,\gamma)(1-\nabla_p^{\max})^3, \tag{A.29}$$

which implies that

$$X_1 \geq p\Omega\chi(p,\gamma)\binom{\Omega}{2} p^2(1-p)^{\Omega-2} M^2(p,\gamma)(1-\nabla_p^{\max})^3. \tag{A.30}$$

To begin, we note that the event $\{0 \Leftrightarrow \mathcal{G}\}$ contains the event $\cup_{e,f} E_{e,f}$, where the union is over unordered pairs of neighbors e, f of the origin, the union is disjoint, and the event $E_{e,f}$ is defined as follows. Let $E_{e,f}$ be the event that the bonds $(0,e)$ and $(0,f)$ are occupied, all other bonds incident on 0 are vacant, and that in the reduced graph $\mathbb{G}^- = (\mathbb{V}^-, \mathbb{B}^-)$ obtained by deleting the origin and each of the Ω bonds incident on 0 from $\mathbb{G}$ the following three events occur: $e \leftrightarrow \mathcal{G}$, $f \leftrightarrow \mathcal{G}$, and $C(e) \cap C(f) = \varnothing$. Let $\mathbb{P}_{p,\gamma}^-$ denote the joint bond/vertex measure on $\mathbb{G}^-$. Then

$$\begin{aligned}\mathbb{P}_{p,\gamma}(0 \Leftrightarrow \mathcal{G}) &\geq \mathbb{P}_{p,\gamma}\Big(\cup_{\{e,f\}} E_{e,f}\Big) = \sum_{\{e,f\}} \mathbb{P}_{p,\gamma}(E_{e,f}) \\ &= p^2(1-p)^{\Omega-2} \sum_{\{e,f\}} \mathbb{P}_{p,\gamma}^-(e \leftrightarrow \mathcal{G},\ \ f \leftrightarrow \mathcal{G},\ \ C(e) \cap C(f) = \varnothing).\end{aligned} \tag{A.31}$$

Let W denote the event whose probability appears on the right side of (A.31). Conditioning on the set $C(e) = A \subset \mathbb{V}^-$, we see that

$$\mathbb{P}_{p,\gamma}^-(W) = \sum_{A:A\ni e} \mathbb{P}_{p,\gamma}^-(C(e) = A,\ \ e \leftrightarrow \mathcal{G},\ \ f \leftrightarrow \mathcal{G},\ \ C(e) \cap C(f) = \varnothing). \tag{A.32}$$

This can be rewritten as

$$\begin{aligned}\mathbb{P}_{p,\gamma}^-(W) &= \sum_{A:A\ni e} \mathbb{P}_{p,\gamma}^-((C(e) = A,\ \ e \leftrightarrow \mathcal{G}) \text{ on } A,\ \ f \leftrightarrow \mathcal{G} \text{ in } \mathbb{V}^- \setminus A) \\ &= \sum_{A:A\ni e} \mathbb{P}_{p,\gamma}^-(C(e) = A,\ \ e \leftrightarrow \mathcal{G})\, \mathbb{P}_{p,\gamma}^-(f \leftrightarrow \mathcal{G} \text{ in } \mathbb{V}^- \setminus A).\end{aligned} \tag{A.33}$$

Let $M^-(x) = \mathbb{P}_{p,\gamma}^-(x \leftrightarrow \mathcal{G})$, for $x \in \mathbb{V}^-$. Then, by the BK inequality and the fact that the two-point function on $\mathbb{G}^-$ is bounded above by the two-point function on $\mathbb{G}$,

$$\mathbb{P}_{p,\gamma}^-(f \leftrightarrow \mathcal{G} \text{ in } \mathbb{V}^- \setminus A) = M^-(e) - \mathbb{P}_{p,\gamma}^-(f \overset{A}{\leftrightarrow} \mathcal{G}) \geq M^-(e) - \sum_{y\in A} \tau_{p,0}(f,y) M^-(y). \tag{A.34}$$

By definition and the BK inequality,

$$M^-(x) = M(p,\gamma) - \mathbb{P}_{p,\gamma}(x \xleftrightarrow{\{0\}} \mathcal{G}) \geq M(p,\gamma)(1 - \tau_{p,0}(0,x)) \geq M(p,\gamma)(1-\nabla_p^{\max}). \tag{A.35}$$

In the above, we also used $\tau_{p,0}(0,x) \leq \nabla_p(0,x)$, which follows from (1.1) (with $u = v = y = 0$).

It follows from (A.33)–(A.35) that

$$\mathbb{P}^{-}_{p,\gamma}(W) \geq M(p,\gamma)(1-\nabla_p^{\max}) \sum_{A: A\ni e} \mathbb{P}^{-}_{p,\gamma}(C(e)=A,\ \ e \leftrightarrow \mathcal{G})\Big[1 - \sum_{y\in A} \tau_{p,0}(f,y)\Big]$$
$$= M(p,\gamma)(1-\nabla_p^{\max})\Big[M^{-}(e) - \sum_{y\in \mathbb{V}^{-}} \tau_{p,0}(f,y)\mathbb{P}^{-}_{p,\gamma}(e\leftrightarrow y,\ e\leftrightarrow \mathcal{G})\Big]. \tag{A.36}$$

By the BK inequality,

$$\mathbb{P}^{-}_{p,\gamma}(e\leftrightarrow y,\ e\leftrightarrow \mathcal{G}) \leq \sum_{w\in\mathbb{V}^{-}} \tau_{p,0}(e,w)\tau_{p,0}(w,y)M^{-}(w), \tag{A.37}$$

and hence, by (A.35)–(A.36),

$$\mathbb{P}^{-}_{p,\gamma}(W) \geq M(p,\gamma)(1-\nabla_p^{\max})\Big[M^{-}(e) - \sum_{y,w\in\mathbb{V}^{-}} \tau_{p,0}(f,y)\tau_{p,0}(e,w)\tau_{p,0}(w,y)M^{-}(w)\Big]$$
$$\geq M^2(p,\gamma)(1-\nabla_p^{\max})^3. \tag{A.38}$$

This completes the proof of (A.29), and hence of (A.30).

Upper bound on X_2. This is the easiest term. By definition,

$$X_2 = p\sum_{(u,v)} \tau_{p,\gamma}(0,u)\mathbb{E}_{p,\gamma}\left[I[v\Leftrightarrow\mathcal{G}] - I[v\Leftrightarrow\mathcal{G} \text{ on } \tilde{C}^{(u,v)}(v)]\right]. \tag{A.39}$$

For the difference of indicators to be nonzero, the double connection from v to $\mathcal{G}$ must be realized via the bond $\{u,v\}$, which therefore must be occupied. The difference of indicators is therefore bounded above by the indicator that the events $\{v\leftrightarrow\mathcal{G}\}$, $\{u\leftrightarrow\mathcal{G}\}$ and $\{\{u,v\}$ occupied$\}$ occur disjointly. Thus we have

$$\mathbb{E}_{p,\gamma}\left[I[v\Leftrightarrow\mathcal{G}] - I[v\Leftrightarrow\mathcal{G} \text{ on } \tilde{C}^{(u,v)}(v)]\right] \leq pM^2(p,\gamma), \tag{A.40}$$

and hence

$$X_2 \leq p^2\Omega M^2(p,\gamma)\chi(p,\gamma). \tag{A.41}$$

Upper bound on X_3. By definition,

$$X_3 = p\sum_{(u,v)} \mathbb{E}_{p,\gamma}\left[\left(\tau_{p,\gamma}(0,u) - \tau_{p,\gamma}^{\tilde{C}^{(u,v)}(v)}(0,u)\right) I[v\Leftrightarrow\mathcal{G} \text{ on } \tilde{C}^{(u,v)}(v)]\right]. \tag{A.42}$$

The difference of two-point functions is the expectation of

$$\begin{aligned} &I[0\leftrightarrow u, 0\not\leftrightarrow\ \mathcal{G}] - I[0\leftrightarrow u \text{ in } \mathbb{V}\backslash\tilde{C}^{(u,v)}(v), 0\not\leftrightarrow\ \mathcal{G}]\\ &+ I[0\leftrightarrow u \text{ in } \mathbb{V}\backslash\tilde{C}^{(u,v)}(v), 0\not\leftrightarrow\ \mathcal{G}] - I[(0\leftrightarrow u, 0\not\leftrightarrow\ \mathcal{G}) \text{ in } \mathbb{V}\backslash\tilde{C}^{(u,v)}(v)]\\ &\leq I[0 \xleftrightarrow{\tilde{C}^{(u,v)}(v)} u, 0\not\leftrightarrow\ \mathcal{G}], \end{aligned} \tag{A.43}$$

since the second line is non-positive and the first line equals the third line. Since the indicator in (A.42) is bounded above by $I[v\Leftrightarrow\mathcal{G}]$, it follows that

$$X_3 \leq p\sum_{(u,v)} \mathbb{E}_{p,\gamma}\left[\mathbb{P}_{p,\gamma}[0 \xleftrightarrow{\tilde{C}^{(u,v)}(v)} u, 0\not\leftrightarrow\ \mathcal{G}]\ I[v\Leftrightarrow\mathcal{G}]\right]. \tag{A.44}$$

By [24, Lemma 4.3] (which is proved by conditioning on $\mathcal{G}$),

$$\mathbb{P}_{p,\gamma}[0 \overset{A}{\Leftrightarrow} u, 0 \not\leftrightarrow \mathcal{G}] \leq \sum_{y \in \mathbb{V}} \tau_{p,\gamma}(0,y)\tau_{p,0}(y,u)I[y \in A]. \tag{A.45}$$

The important point in (A.45) is that the condition $0 \not\leftrightarrow \mathcal{G}$ on the left side is retained in the factor $\tau_{p,\gamma}(0,y)$ on the right side (but not in $\tau_{p,0}(y,u)$). With (A.44), this gives

$$X_3 \leq p \sum_{(u,v)} \sum_{y \in \mathbb{V}} \tau_{p,\gamma}(0,y)\tau_{p,0}(y,u)\mathbb{E}_{p,\gamma}\left[I[v \Leftrightarrow \mathcal{G}]I[y \in \tilde{C}^{(u,v)}(v)]\right]. \tag{A.46}$$

Since

$$I[v \Leftrightarrow \mathcal{G}]I[y \in \tilde{C}^{(u,v)}(v)] \leq I[\{v \leftrightarrow w \leftrightarrow y, v \leftrightarrow \mathcal{G}\} \circ \{v \leftrightarrow \mathcal{G}\}], \tag{A.47}$$

a further application of BK gives

$$\begin{aligned} X_3 &\leq p \sum_{y \in \mathbb{V}} \tau_{p,\gamma}(0,y) \sum_{(u,v)} \tau_{p,0}(y,u) \sum_{w \in \mathbb{V}} \tau_{p,0}(v,w)\tau_{p,0}(y,w)M^2(p,\gamma) \\ &= pM^2(p,\gamma)\chi(p,\gamma)T_1(0) \leq pM^2(p,\gamma)\chi(p,\gamma)\Omega\bar{\nabla}_p, \end{aligned} \tag{A.48}$$

using Lemma A.2 in the last step.

The combination of (A.30), (A.41) and (A.48) completes the proof of (A.20). □

A.3 The differential inequality (8.14)

Let $0 \leq p \leq 1$, $0 \leq \gamma \leq 1$, and let $Z_{\mathcal{G}}$ denote the number of vertices that are connected to a green vertex. The differential inequality (8.14) states that

$$\frac{\partial}{\partial p}\mathbb{E}_{p,\gamma}\big[Z_{\mathcal{G}}^2\big] \leq \frac{3\Omega}{1-p}\frac{1-\gamma}{\gamma}M(p,\gamma)\mathbb{E}_{p,\gamma}\big[Z_{\mathcal{G}}^2\big]. \tag{A.49}$$

Proof of (A.49). Let $A_{x,y}$ be the event that $x \leftrightarrow \mathcal{G}$ and $y \leftrightarrow \mathcal{G}$. Then

$$\mathbb{E}_{p,\gamma}\big[Z_{\mathcal{G}}^2\big] = \sum_{x,y \in \mathbb{V}} \mathbb{P}_{p,\gamma}(A_{x,y}), \tag{A.50}$$

and hence, by Russo's formula,

$$\begin{aligned} \frac{\partial}{\partial p}\mathbb{E}_{p,\gamma}\big[Z_{\mathcal{G}}^2\big] &= \sum_{x,y \in \mathbb{V}} \sum_{\{u,v\} \in \mathbb{B}} \mathbb{P}_{p,\gamma}(\{u,v\} \text{ is pivotal for } A_{x,y}) & \text{(A.51)} \\ &= \frac{1}{1-p} \sum_{x,y \in \mathbb{V}} \sum_{\{u,v\} \in \mathbb{B}} \mathbb{P}_{p,\gamma}(\{u,v\} \text{ is vacant and pivotal for } A_{x,y}). & \text{(A.52)} \end{aligned}$$

If $\{u,v\}$ is vacant and pivotal for $A_{x,y}$, then exactly one of the two endpoints of the edge $\{u,v\}$ is connected to a green vertex. Moreover, if one of the two endpoints of $\{u,v\}$ is connected to a green vertex, and the other is not, then the edge $\{u,v\}$ is automatically vacant. As a consequence,

$$\frac{\partial}{\partial p}\mathbb{E}_{p,\gamma}\big[Z_{\mathcal{G}}^2\big] = \frac{1}{1-p} \sum_{x,y \in \mathbb{V}} \sum_{(u,v)} \mathbb{P}_{p,\gamma}(\{\{u,v\} \text{ is pivotal for } A_{x,y}\} \cap \{u \leftrightarrow \mathcal{G}\} \cap \{v \not\leftrightarrow \mathcal{G}\}), \tag{A.53}$$

where the sum over (u, v) is a sum over *directed* edges. To analyze the probability in (A.53), we distinguish two cases: either exactly one of the two vertices x and y is connected to a green vertex, or neither of them is connected to a green vertex. It is not possible that both are connected to $\mathcal{G}$, because we are in a situation where $\{u, v\}$ is vacant, and it cannot then also be pivotal for $A_{x,y}$.

Let us first estimate the contribution due to the event that neither x nor y is connected to a green vertex. A moment's reflection shows that this contribution can be rewritten as

$$\mathbb{P}_{p,\gamma}(\{u \leftrightarrow \mathcal{G}\} \cap \{x \leftrightarrow y \leftrightarrow v \not\leftrightarrow \mathcal{G}\}). \tag{A.54}$$

We will estimate (A.54) by applying the BK inequality, as generalized by van den Berg and Fiebig [8] to cover intersections of increasing and decreasing events, to the joint distribution $\mathbb{P}_{p,\gamma}$ (alternatively, the decoupling inequalities of [11] could be applied). With respect to $\mathbb{P}_{p,\gamma}$, the event $\{u \leftrightarrow \mathcal{G}\}$ is increasing, whereas the event $\{x \leftrightarrow y \leftrightarrow v \not\leftrightarrow \mathcal{G}\}$ is the intersection of an increasing and a decreasing event. In addition, these events must occur disjointly. Therefore, by the BK inequality, (A.54) is bounded by

$$\mathbb{P}_{p,\gamma}(u \leftrightarrow \mathcal{G})\mathbb{P}_{p,\gamma}(x \leftrightarrow y \leftrightarrow v \not\leftrightarrow \mathcal{G}). \tag{A.55}$$

Consider now the contribution from the event that x is connected to a green vertex, while y is not. This contribution can be rewritten as

$$\mathbb{P}_{p,\gamma}(\{u \leftrightarrow \mathcal{G}\} \cap \{x \leftrightarrow \mathcal{G}\} \cap \{y \leftrightarrow v \not\leftrightarrow \mathcal{G}\}), \tag{A.56}$$

which we bound by

$$\mathbb{P}_{p,\gamma}(\{u \leftrightarrow \mathcal{G}\} \cap \{x \leftrightarrow \mathcal{G}\})\mathbb{P}_{p,\gamma}(y \leftrightarrow v \not\leftrightarrow \mathcal{G}). \tag{A.57}$$

Interchanging the role of x and y, we obtain a similar bound on the contribution of the term with $y \leftrightarrow \mathcal{G}$ and $x \not\leftrightarrow \mathcal{G}$.

Inserting these three bounds into (A.53), and recalling (A.22), we get

$$\begin{aligned}
\frac{\partial}{\partial p}\mathbb{E}_{p,\gamma}\Big[Z_{\mathcal{G}}^2\Big] &\leq \frac{1}{1-p}\sum_{x,y\in\mathbb{V}}\sum_{(u,v)} \mathbb{P}_{p,\gamma}(u \leftrightarrow \mathcal{G})\mathbb{P}_{p,\gamma}(x \leftrightarrow y \leftrightarrow v \not\leftrightarrow \mathcal{G}) \\
&\quad + \frac{2}{1-p}\sum_{x,y\in\mathbb{V}}\sum_{(u,v)} \mathbb{P}_{p,\gamma}(\{u \leftrightarrow \mathcal{G}\} \cap \{x \leftrightarrow \mathcal{G}\})\mathbb{P}_{p,\gamma}(y \leftrightarrow v \not\leftrightarrow \mathcal{G}) \\
&= \frac{V\Omega}{1-p}M(p,\gamma)\mathbb{E}_{p,\gamma}[|C(0)|^2 I(0 \not\leftrightarrow \mathcal{G})] + \frac{2\Omega}{1-p}\chi(p,\gamma)\mathbb{E}_{p,\gamma}\Big[Z_{\mathcal{G}}^2\Big].
\end{aligned} \tag{A.58}$$

To complete the proof of (A.49), we estimate (A.58) by using Lemmas 8.2 and 8.1 for the first term, and using the lower bound of (5.2) for the second term. □

Acknowledgments

This work began during a conversation at afternoon tea, while RvdH, GS and JS were visiting Microsoft Research. The authors thank Benny Sudakov for bringing the question of the critical point of the n-cube to our attention, and for telling us about the recent papers [6] and [18]. The work of GS was supported in part by NSERC of Canada. The work of RvdH was carried out in part at the University of British Columbia and in part at Delft University of Technology.

References

[1] M. Aizenman. On the number of incipient spanning clusters. *Nucl. Phys. B [FS]*, **485**:551–582, (1997).

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

[3] M. Aizenman and R. Fernández. On the critical behavior of the magnetization in high dimensional Ising models. *J. Stat. Phys.*, **44**:393–454, (1986).

[4] M. Aizenman and C.M. Newman. Tree graph inequalities and critical behavior in percolation models. *J. Stat. Phys.*, **36**:107–143, (1984).

[5] M. Ajtai, J. Komlós, and E. Szemerédi. Largest random component of a k-cube. *Combinatorica*, **2**:1–7, (1982).

[6] N. Alon, I. Benjamini, and A Stacey. Percolation on finite graphs and isoperimetric inequalities. Preprint, (2002).

[7] D.J. Barsky and M. Aizenman. Percolation critical exponents under the triangle condition. *Ann. Probab.*, **19**:1520–1536, (1991).

[8] J. van den Berg and U. Fiebig. Inequalities with applications to percolation and reliability. *Ann. Probab.*, **15**:354–374, (1987).

[9] B. Bollobás. The evolution of random graphs. *Trans. Amer. Math. Soc.*, **286**:257–274, (1984).

[10] B. Bollobás, Y. Kohayakawa, and Łuczak. The evolution of random subgraphs of the cube. *Random Struct. Alg.*, **3**:55–90, (1992).

[11] C. Borgs and J.T. Chayes. The covariance matrix of the Potts model: a random cluster analysis. *J. Stat. Phys.*, **82**:1235–1297, (1996).

[12] C. Borgs, J.T. Chayes, R. van der Hofstad, G. Slade, and J. Spencer. Random subgraphs of finite graphs: II. The lace expansion and the triangle condition. Preprint, (2003).

[13] C. Borgs, J.T. Chayes, R. van der Hofstad, G. Slade, and J. Spencer. Random subgraphs of finite graphs: III. The phase transition for the n-cube. Preprint, (2003).

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

[15] J.T. Chayes and L. Chayes. Inequality for the infinite-cluster density in Bernoulli percolation. *Phys. Rev. Lett.*, **56**:1619–1622, (1986).

[16] J.T. Chayes and L. Chayes. The mean field bound for the order parameter of Bernoulli percolation. In H. Kesten, editor, *Percolation Theory and Ergodic Theory of Infinite Particle Systems*. Springer, New York, (1987).

[17] P. Erdős and A. Rényi. On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, **5**:17–61, (1960).

[18] A. Frieze, M. Krivelevich, and R. Martin. Emergence of a giant component in random subgraphs of pseudo-random graphs. Preprint, (2002).

[19] G. Grimmett. *Percolation.* Springer, Berlin, 2nd edition, (1999).

[20] T. Hara. Critical two-point functions for nearest-neighbour high-dimensional self-avoiding walk and percolation. In preparation.

[21] T. Hara, R. van der Hofstad, and G. Slade. Critical two-point functions and the lace expansion for spread-out high-dimensional percolation and related models. *Ann. Probab.*, **31**:349–408, (2003).

[22] T. Hara and G. Slade. Mean-field critical behaviour for percolation in high dimensions. *Commun. Math. Phys.*, **128**:333–391, (1990).

[23] T. Hara and G. Slade. Mean-field behaviour and the lace expansion. In G. Grimmett, editor, *Probability and Phase Transition*, Dordrecht, (1994). Kluwer.

[24] T. Hara and G. Slade. The scaling limit of the incipient infinite cluster in high-dimensional percolation. I. Critical exponents. *J. Stat. Phys.*, **99**:1075–1168, (2000).

[25] T. Hara and G. Slade. The scaling limit of the incipient infinite cluster in high-dimensional percolation. II. Integrated super-Brownian excursion. *J. Math. Phys.*, **41**:1244–1293, (2000).

[26] R. van der Hofstad and G. Slade. Asymptotic expansions for percolation critical values on Z^d and the n-cube. In preparation.

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

[28] T. Łuczak. Component behavior near the critical point of the random graph process. *Random Structures Algorithms*, **1**:287–310, (1990).

[29] M.V. Menshikov. Coincidence of critical points in percolation problems. *Soviet Mathematics, Doklady*, **33**:856–859, (1986).

[30] R.H. Schonmann. Multiplicity of phase transitions and mean-field criticality on highly non-amenable graphs. *Commun. Math. Phys.*, **219**:271–322, (2001).

[31] R.H. Schonmann. Mean-field criticality for percolation on planar non-amenable graphs. *Commun. Math. Phys.*, **225**:453–463, (2002).