arXiv:2106.06400v2 [math.PR] 14 Jun 2021

On the derivation of mean-field percolation critical exponents from the triangle condition

Tom Hutchcroft

Statslab, DPMMS, University of Cambridge.
Email: t.hutchcroft@maths.cam.ac.uk

June 15, 2021

Abstract. We give a new derivation of mean-field percolation critical behaviour from the triangle condition that is quantitatively much better than previous proofs when the triangle diagram ∇_{p_c} is large. In contrast to earlier methods, our approach continues to yield bounds of reasonable order when the triangle diagram ∇_p is unbounded but diverges slowly as $p \uparrow p_c$, as is expected to occur in percolation on $\mathbb{Z}^d$ at the upper-critical dimension $d = 6$. Indeed, we show in particular that if the triangle diagram diverges polylogarithmically as $p \uparrow p_c$ then mean-field critical behaviour holds to within a polylogarithmic factor. We apply the methods we develop to deduce that for long-range percolation on the hierarchical lattice, mean-field critical behaviour holds to within polylogarithmic factors at the upper-critical dimension.

As part of the proof, we introduce a new method for comparing diagrammatic sums on general transitive graphs that may be of independent interest.

Contents

1 Introduction **2**
1.1 Other exponents via the relative entropy method 4
1.2 About the proof 6
1.3 Applications to the hierarchical lattice 7

2 Derivation of the main differential inequality **8**

3 Comparing diagrams: a sojourn into linear algebra **12**

4 The tail of the volume and the infinite cluster density **23**

5 Applications to the hierarchical lattice **27**

1 Introduction

The triangle condition is a sufficient condition for mean-field critical behaviour in Bernoulli percolation models that was introduced by Aizenman and Newman [4] and proven to hold on high-dimensional Euclidean lattices in the milestone work of Hara and Slade [21]. Let $G = (V, E)$ be a connected, locally finite, transitive graph, let o be a fixed vertex of G, and for each $0 \leq p \leq 1$ and $x, y \in V$ let $T_p(x, y) = \mathbb{P}_p(x \leftrightarrow y)$ be the probability that x and y are connected in Bernoulli-p bond percolation. We think of T_p as a matrix indexed by $V \times V$ and refer to it as the **two-point matrix**. For each $0 \leq p \leq 1$ we define the **triangle diagram** ∇_p to be

$$\nabla_p := \sum_{x,y \in V} T_p(o, x) T_p(x, y) T_p(y, o) = T_p^3(o, o).$$

We say that G satisfies the **triangle condition** if ∇_p is finite when p is equal to the critical probability $p_c = p_c(G)$. Aizenman and Newman [4] and Barsky and Aizenman [7] proved that if G is, say, a Cayley graph satisfying the triangle condition then G has mean-field critical behaviour for percolation in the sense that if K denotes the cluster of the origin then

$$\mathbb{E}_p|K| \asymp |p - p_c|^{-1} \qquad \text{as } p \uparrow p_c, \tag{1.1}$$

$$\mathbb{P}_{p_c}(|K| \geq n) \asymp n^{-1/2} \qquad \text{as } n \uparrow \infty, \text{ and} \tag{1.2}$$

$$\mathbb{P}_p(|K| = \infty) \asymp |p - p_c| \qquad \text{as } p \downarrow p_c, \tag{1.3}$$

where we write $\asymp$ to denote an equality holding to within multiplication by positive constants within the vicinity of the relevant limit point. Note that the *lower bounds* of (1.1) and (1.3) are known to hold on every transitive graph [1, 17]. The triangle condition has now been proven to hold in a variety of high-dimensional settings [11, 13, 19, 23, 24, 27, 28, 42], and further works studying critical behaviour in high-dimensional percolation under the triangle condition include [10, 12, 15, 31, 32, 38]; see [22] for an overview of this extensive literature and [20] for further background on percolation.

A new differential inequality for the susceptibility. The proofs of eqs. (1.1)–(1.3) rely crucially on *differential inequalities*. Let us consider as a paradigmatic example the rate of growth of the *susceptibility* $\mathcal{X}_p := \mathbb{E}_p|K|$ as $p \uparrow p_c$, whose behaviour under the triangle condition was determined by Aizenman and Newman [4]. It is a fundamental fact known as the *sharpness of the phase transition* and due originally to Menshikov [35] and Aizenman and Barsky [1] that $\mathcal{X}_p < \infty$ if and only if $p < p_c$; see also [16, 17, 26] for alternative proofs. Moreover, the susceptibility $\mathcal{X}_p$ is a smooth function of p on $[0, p_c)$ [20, Chapter 6.4] and satisfies the differential inequality

$$\frac{d\mathcal{X}_p}{dp} \leq 2\deg(o)\mathcal{X}_p^2 \tag{1.4}$$

for every $0 < p < p_c$ [22, Equation 4.2.4]. This inequality can be integrated to obtain that the *mean-field lower bound*

$$\mathcal{X}_p \geq \frac{1}{2\deg(o)}|p - p_c|^{-1} \tag{1.5}$$

holds for every $0 \leq p < p_c$. In order to establish mean-field critical behaviour under the triangle condition, the key step is to establish a *lower bound* on the derivative of $\mathcal{X}_p$ of the same order

as the upper bound of (1.4), which can then be integrated to obtain an upper bound on $\mathcal{X}_p$ of the same order as the lower bound of (1.5). To do this, one considers the so-called *open triangle* $\nabla_{p_c}(r) = \inf_{x:d(o,x)\leq r} T^3_{p_c}(o,x)$, which tends to zero as $r \to \infty$ when the triangle condition holds [30], and argues that for each $r \geq 1$ there exists $\varepsilon(r) > 0$ such that

$$\frac{d\mathcal{X}_p}{dp} \geq \varepsilon(r)(1 - \nabla_{p_c}(r))\mathcal{X}_p^2 \tag{1.6}$$

for every $p_c/2 \leq p < p_c$. See also the proof of [20, Theorem 10.68] for a detailed treatment of the case $G = \mathbb{Z}^d$. Taking r to be sufficiently large that $\nabla_{p_c}(r) < 1$ gives a differential inequality of the required form. Unfortunately, the dependence of $\varepsilon(r)$ on r given by these arguments tends to be very bad (e.g. exponentially small in r), so that the implicit constants appearing in eqs. (1.1)–(1.3) tend to be extremely large when the triangle sum ∇_{p_c} is finite but large. The same shortcomings make these methods poorly suited to situations where, say, ∇_{p_c} is infinite but ∇_p diverges slowly as $p \uparrow p_c$, as is expected to happen on $\mathbb{Z}^d$ at the upper-critical dimension $d = 6$.

In this paper we give a new derivation of mean-field critical behaviour from the triangle condition that is arguably simpler than the standard method (in the case of $\mathbb{Z}^d$) and that gives much better quantitative control when the triangle sum is large. We then apply this new method to study long-range percolation on the *hierarchical lattice* at the upper-critical dimension as discussed in detail in Section 1.3. Since we are interested in applying our results to both long-range and nearest-neighbour models, we work with percolation on *weighted graphs*, which we now briefly define. A **weighted graph** $G = (V, E, J)$ is a countable graph (V, E) together with an assignment of non-negative **weights** $\{J_e : e \in E\}$ such that $\sum_{e\in E_v^{\to}} J_e < \infty$ for each $v \in V$, where $E_v^{\to}$ denotes the set of oriented edges of G emanating from the vertex v. (Although all our graphs are unoriented, it is useful to think of each unoriented edge as corresponding to a pair of oriented edges pointing in opposite directions.) Locally finite graphs can be considered as weighted graphs by setting $J_e \equiv 1$. A graph automorphism of (V, E) is a weighted graph automorphism of (V, E, J) if it preserves the weights, and a weighted graph G is said to be **transitive** if any vertex can be mapped to any other vertex by an automorphism. Given a weighted graph $G = (V, E, J)$ and $\beta \geq 0$, we define Bernoulli-β bond percolation on G to be the random subgraph of G in which each edge e is chosen to be either retained or deleted independently at random with retention probability $1 - e^{-\beta J_e}$, and write $\mathbb{P}_\beta = \mathbb{P}_{G,\beta}$ for the law of this random subgraph.

The main contribution of the paper is the following simple explicit strengthening of the differential inequality (1.6). Throughout the paper we will assume without loss of generality that $\sum_{e\in E_o^{\to}} J_e = 1$; multiplying all edge weights by a constant so that this holds is equivalent to multiplying the parameter β by a constant. As above, when G is transitive we fix a vertex o of G, write K for the cluster of o, and write $\mathcal{X}_\beta = \mathbb{E}_\beta|K|$ for the susceptibility, which is a finite-valued, smooth function of β on $[0, \beta_c)$ by sharpness of the phase transition [1, 17, 26].

Theorem 1.1. *Let G be a connected, unimodular, transitive weighted graph that is normalised so that $\sum_{e\in E_o^{\to}} J_e = 1$. Then*

$$\frac{d\mathcal{X}_\beta}{d\beta} \geq \frac{\mathcal{X}_\beta(\mathcal{X}_\beta - \nabla_\beta)}{3\beta^2\nabla_\beta^2} \tag{1.7}$$

for every $0 \leq \beta < \beta_c$.

Here, we recall that a connected, transitive weighted graph $G = (V, E, J)$ is said to be **unimodular** if it satisfies the **mass-transport principle**, meaning that if $F : V \times V \to [0, \infty]$ is diagonally invariant in the sense that $F(\gamma x, \gamma y) = F(x, y)$ for every $x, y \in V$ and $\gamma \in \mathrm{Aut}(G)$ then

$$\sum_{x \in V} F(o, x) = \sum_{x \in V} F(x, o). \tag{1.8}$$

Most transitive weighted graphs arising in applications are unimodular, including all amenable transitive weighted graphs and all Cayley graphs of countable groups [44]. Indeed, if G is a Cayley graph of a countable group Γ then left multiplication by an element of Γ defines an automorphism of G and we trivially verify (1.8) by writing

$$\sum_{\gamma \in \Gamma} F(\mathrm{id}, \gamma) = \sum_{\gamma \in \Gamma} F(\gamma^{-1}\mathrm{id}, \gamma^{-1}\gamma) = \sum_{\gamma \in \Gamma} F(\gamma^{-1}, \mathrm{id}) = \sum_{\gamma \in \Gamma} F(\gamma, \mathrm{id}).$$

(Note that all *nonunimodular* transitive graphs are proven to have mean-field critical behaviour for percolation in [27] via different methods that are exclusive to the nonunimodular case.)

Integrating the differential inequality (1.7) yields the following corollary, which recovers the estimate (1.1) when the triangle condition holds.

Corollary 1.2. *Let G be a connected, unimodular, transitive weighted graph that is normalised so that $\sum_{e \in E_o^{\to}} J_e = 1$. Then*

$$\mathcal{X}_\beta \leq \left(1 + \frac{\beta_c - \beta}{3\beta\beta_c}\right) \left[\int_\beta^{\beta_c} \frac{1}{3\lambda^2 \nabla_\lambda^2} \, \mathrm{d}\lambda\right]^{-1}$$

for every $0 \leq \beta < \beta_c$.

Note that the $(1 + (\beta_c - \beta)/(3\beta\beta_c))$ prefactor appearing here converges to 1 as $\beta \uparrow \beta_c$.

1.1 Other exponents via the relative entropy method

Classically, Barsky and Aizenman [7] proved that (1.2) and (1.3) hold under the triangle condition by proving a further partial differential inequality concerning the *magnetization* $M_{\beta,h} = \mathbb{E}_\beta[1 - e^{-h|K|}] \asymp \sum_{n=1}^{\lceil 1/h \rceil} \mathbb{P}_\beta(|K| \geq n)$. Specifically, they proved that there exists a constant C such that for each $r \geq 1$ there exists $\varepsilon(r) > 0$ such that

$$M_{\beta,h} - h\frac{\partial M_{\beta,h}}{\partial h} \geq \varepsilon(r)\Big(1 - C\nabla_\beta(r)\Big) M_{\beta,h}^2 \frac{\partial M_{\beta,h}}{\partial h} - \frac{1}{\varepsilon(r)} h M_{\beta,h} \frac{\partial M_{\beta,h}}{\partial h} \tag{1.9}$$

for every $\beta_c/2 \leq \beta < \beta_c$ and $h \geq 0$; the upper bound of (1.2) can be deduced from this partial differential inequality via an elementary (if non-obvious) calculation. This method suffers from the same quantitative shortcomings affecting Aizenman and Newman's analysis of the susceptibility and does not provide estimates of reasonable order when the triangle sum is unbounded but diverges slowly as $\beta \uparrow \beta_c$.

Rather than following this strategy, we instead observe that the upper bounds of (1.2) and (1.3) can each be deduced directly from the upper bound of (1.1) using the *relative entropy method*

pioneered in the recent work of Dewan and Muirhead [15], which yields good quantitative control of all the relevant constants. This allows us to completely avoid proving any partial differential inequality of the form (1.9) and obtain versions of (1.2) and (1.3) with good quantitative dependence on the value of the triangle diagram ∇_{p_c}. More specifically, we adapt the methods of Dewan and Muirhead to prove the following general inequality.

Theorem 1.3. *Let $G = (V, E, J)$ be a connected, transitive weighted graph that is normalised so that $\sum_{e \in E_o^{\rightarrow}} J_e = 1$. Then*

$$\mathbb{P}_{\beta_2}(|K| \geq n) \leq 2\mathbb{P}_{\beta_1}(|K| \geq n) + \frac{4}{\beta_1}|\beta_2 - \beta_1|^2 \sum_{k=1}^{n} \mathbb{P}_{\beta_1}(|K| \geq k) \leq \left(\frac{2}{n} + \frac{4}{\beta_1}|\beta_2 - \beta_1|^2\right) \mathcal{X}_{\beta_1}$$

for every $\beta_2 \geq \beta_1 > 0$ and $n \geq 0$.

Taking $\beta_1 = \beta_c - n^{-1/2}$ and $\beta_2 = \beta_c$ in this inequality shows that the upper bound of (1.1) implies the upper bound of (1.2), while taking $\beta_1 = \beta_c - \varepsilon$ and $\beta_2 = \beta_c + \varepsilon$ and $n \uparrow \infty$ shows that the upper bound of (1.1) implies the upper bound of (1.3). (The complementary fact that the upper bound of (1.2) implies the upper bound of (1.1) was established in [26, Theorem 1.1]. One can also deduce the upper bound of (1.3) from the upper bound of (1.2) using the *extrapolation* technique of Aizenman and Fernandez [2] as discussed in [7].)

Remark 1.4. The bounds of Theorem 1.3 are similar to, but quantitatively better than, those appearing in the work of Newman [36, 37], which relies on differential inequalities obtained via large-deviations analysis of the *fluctuation* $(1-p)\#\{\text{open edges in } K\} - p\#\{\text{closed edges touching } K\}$. One can sharpen Newman's analysis by using maximal inequalities instead of large-deviations estimates but the bounds one obtains this way are still not as strong as those of Theorem 1.3. As explained to us by Stephen Muirhead, it appears to be a rather general phenomenon that relative entropy methods give 'non-differential improvements' to estimates based on the analysis of the fluctuation.

The following corollary is illustrative of what can be done with the new methods we introduce. The case $\alpha = 0$ of the corollary recovers the classical fact the triangle condition implies mean-field critical behaviour as proven by Aizenman and Newman [4] and Barsky and Aizenman [7].

Corollary 1.5. *Let G be a connected, unimodular, transitive weighted graph and suppose that there exists $\alpha \geq 0$ such that $\nabla_{\beta_c - \varepsilon} \preceq (\log(1/\varepsilon))^\alpha$ as $\varepsilon \downarrow 0$. Then*

$$\mathcal{X}_{\beta_c - \varepsilon} \preceq \left(\log \frac{1}{\varepsilon}\right)^{2\alpha} \cdot \frac{1}{\varepsilon} \qquad \text{as } \varepsilon \downarrow 0, \tag{1.10}$$

$$\mathbb{P}_{\beta_c}\left(|K| \geq n\right) \preceq (\log n)^{2\alpha} \cdot \frac{1}{n^{1/2}} \qquad \text{as } n \uparrow \infty, \text{ and} \tag{1.11}$$

$$\mathbb{P}_{\beta_c + \varepsilon}\left(|K_v| = \infty\right) \preceq \left(\log \frac{1}{\varepsilon}\right)^{2\alpha} \cdot \varepsilon \qquad \text{as } \varepsilon \downarrow 0. \tag{1.12}$$

In particular, mean-field critical behaviour holds to within polylogarithmic factors.

See [3, Proposition 3.1] and [34, Theorem 1.5.4] for related results for the four-dimensional Ising model and self-avoiding walk. The hypothesis $\nabla_{\beta_c - \varepsilon} \preceq (\log(1/\varepsilon))^\alpha$ is expected to hold with $\alpha > 0$ for nearest-neighbour percolation on $\mathbb{Z}^d$ at the upper-critical dimension $d = 6$, but proving this appears to be completely beyond the scope of existing methods.

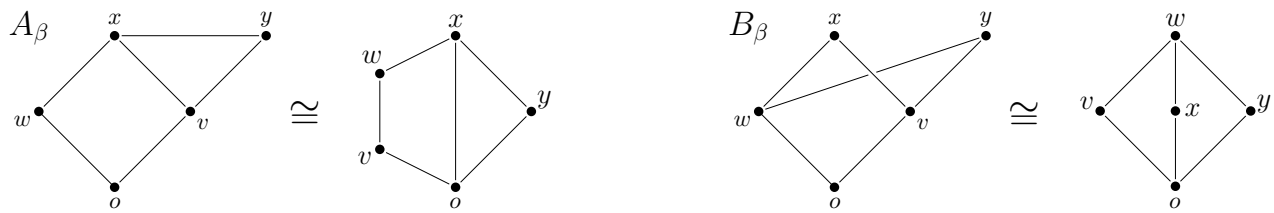

Figure 1: The diagrammatic sums A_β and B_β. In each case, we fix the origin o and sum over every other vertex, with an edge of the diagram representing a copy of the two-point matrix T_β. When G is unimodular we can change which vertex of the diagram is pinned to the origin without changing the value of the corresponding sum.

1.2 About the proof

The differential inequality of Theorem 1.1 will be deduced from a more fundamental estimate involving two diagrammatic sums A_β and B_β that are more complicated than the usual triangle diagram. Let G be a countable, transitive weighted graph, and for each $\beta \geq 0$ consider the diagrammatic sums

$$A_\beta = \sum_{v,w,x,y \in V} T_\beta(o,w)T_\beta(o,v)T_\beta(w,x)T_\beta(v,x)T_\beta(v,y)T_\beta(y,x) \tag{1.13}$$

and

$$B_\beta = \sum_{v,w,x,y \in V} T_\beta(o,w)T_\beta(o,v)T_\beta(w,x)T_\beta(v,x)T_\beta(w,y)T_\beta(v,x), \tag{1.14}$$

both of which belong to $[1,\infty]$. See Figure 1 for graphical representations of these sums. When G is unimodular, the mass-transport principle (1.8) allows us to exchange the roles of o and v to write these sums more succinctly as

$$\begin{aligned} A_\beta &= \sum_{v,w,x,y \in V} T_\beta(o,v)T_\beta(v,w)T_\beta(w,x)T_\beta(o,x)T_\beta(o,y)T_\beta(y,x) \\ &= \sum_{x \in V} T_\beta(o,x)T_\beta^2(o,x)T_\beta^3(o,x) \end{aligned} \tag{1.15}$$

and

$$B_\beta = \sum_{v,w,x,y \in V} T_\beta(o,v)T_\beta(v,w)T_\beta(o,x)T_\beta(x,w)T_\beta(o,y)T_\beta(y,w) = \sum_{w \in V} T_\beta^2(o,w)^3 \tag{1.16}$$

for each $\beta \geq 0$. We will deduce Theorem 1.1 as a corollary of the following two propositions.

Proposition 1.6. *Let $G = (V,E,J)$ be a connected, transitive weighted graph that is normalised so that $\sum_{e \in E_o^\rightarrow} J_e = 1$. Then*

$$\frac{d\mathcal{X}_\beta}{d\beta} \geq \frac{\mathcal{X}_\beta(\mathcal{X}_\beta - \nabla_\beta)}{\beta^2(2A_\beta + B_\beta)} \tag{1.17}$$

for every $0 < \beta < \beta_c$.

Proposition 1.7. *Let $G = (V, E, J)$ be a connected, unimodular, transitive weighted graph. Then*

$$B_\beta \leq A_\beta \leq \nabla_\beta^2$$

for every $\beta \geq 0$.

Note that Proposition 1.6 does *not* require unimodularity.

The inequality $A_\beta \leq \nabla_\beta^2$ of Proposition 1.7 follows easily from the fact that T_β is positive definite [4, Lemma 3.3] and hence that the maximal entries of T_β^3 lie on its diagonal. The inequality $B_\beta \leq A_\beta$ holds for rather more subtle reasons, but is also a consequence of T_β being positive definite. In the case $G = \mathbb{Z}^d$ this inequality admits a relatively straightforward proof by Fourier analysis, which is given at the beginning of Section 3. The general proof of Proposition 1.7 uses some interesting and (in our view) rather obscure facts about positive definite matrices which we review and develop in the remainder of Section 3. We are optimistic that the techniques developed in Section 3 may find further applications to probability theory in the future, perhaps to the problem of implementing the lace expansion [43] in non-Euclidean settings.

1.3 Applications to the hierarchical lattice

We now describe the applications of our results to long-range percolation on the hierarchical lattice. We begin by defining the model. Let $d \geq 1$, $L \geq 2$, and let $\mathbb{T}_L^d = (\mathbb{Z}/L\mathbb{Z})^d$ be the discrete d-dimensional torus of side length L. We define the **hierarchical lattice** $\mathbb{H}_L^d$ to be the countable Abelian group $\bigoplus_{i=1}^\infty \mathbb{T}_L^d = \{x = (x_1, x_2, \ldots) \in (\mathbb{T}_L^d)^\mathbb{N} : x_i = 0 \text{ for all but finitely many } i \geq 0\}$ together with the group-invariant ultrametric

$$\|y - x\| = \begin{cases} 0 & x = y \\ L^{h(x,y)} & x \neq y \end{cases} \qquad \text{where } h(x,y) = \max\{i \geq 1 : x_i \neq y_i\}.$$

A function $J : \mathbb{H}_L^d \to [0, \infty)$ is said to be **radially symmetric** if $J(x)$ can be expressed as a function of $\|x\|$ and is said to be **integrable** if $\sum_{x \in \mathbb{H}_L^d} J(x) < \infty$. For example, writing $\langle x \rangle = 1 \vee \|x\|$ for each $x \in \mathbb{H}_L^d$, the function $J(x) = \langle x \rangle^{-d-\alpha}$ is radially symmetric and integrable whenever α is positive. Given a radially symmetric, integrable function $J : \mathbb{H}_L^d \to [0, \infty)$, we define can define a transitive weighted graph with vertex set $\mathbb{H}_L^d$, edge set $\{\{x, y\} : x, y \in \mathbb{H}_L^d,\ J(x - y) > 0\}$, and weights given by $J(\{x, y\}) = J(x - y)$; percolation on the resulting weighted graph is referred to as *long-range percolation on the hierarchical lattice*. This model has the convenient feature that for each $n \geq 1$ the ultrametric ball $\Lambda_n := \{x \in \mathbb{H}_L^d : \|x\| \leq L^n\}$ induces a weighted subgraph of the hierarchical lattice that is itself transitive.

The principal result of our earlier work [28] states that if $J : \mathbb{H}_L^d \to [0, \infty)$ is an integrable, radially symmetric function satisfying $c\|x\|^{-d-\alpha} \leq J(x) \leq C\|x\|^{-d-\alpha}$ for some positive constants c and C and every $x \in \mathbb{H}_L^d \setminus \{0\}$ then there exist positive constants a and A such that

$$a\langle x - y \rangle^{-d+\alpha} \leq \mathbb{P}_{\beta_c}(x \leftrightarrow y) \leq A\langle x - y \rangle^{-d+\alpha} \tag{1.18}$$

for every $x, y \in \mathbb{H}_L^d$. It follows from this result that the model satisfies the triangle condition if and only if $\alpha < d/3$ [28, Corollary 1.4], and indeed it is proven in [28, Corollary 1.5] that the model does

not have mean-field critical exponents when $\alpha > d/3$.

The techniques we develop in this paper can be used to prove that mean-field critical behaviour holds to within polylogarithmic factors in the upper-critical case $\alpha = d/3$.

Theorem 1.8. *Let $J : \mathbb{H}_L^d \to [0,\infty)$ be a radially symmetric, integrable function, let $0 < \alpha < d$, and suppose that there exist constants c and C such that $c\|x\|^{-d-\alpha} \leq J(x) \leq C\|x\|^{-d-\alpha}$ for every $x \in \mathbb{H}_L^d \setminus \{0\}$. If $\alpha = d/3$ then*

$$\mathcal{X}_{\beta_c-\varepsilon} \preceq \left(\log \frac{1}{\varepsilon}\right)^2 \cdot \frac{1}{\varepsilon} \qquad \text{as } \varepsilon \downarrow 0, \tag{1.19}$$

$$\mathbb{P}_{\beta_c}\left(|K| \geq n\right) \preceq (\log n)^2 \cdot \frac{1}{n^{1/2}} \qquad \text{as } n \uparrow \infty, \text{ and} \tag{1.20}$$

$$\mathbb{P}_{\beta_c+\varepsilon}\left(|K_v| = \infty\right) \preceq \left(\log \frac{1}{\varepsilon}\right)^2 \cdot \varepsilon \qquad \text{as } \varepsilon \downarrow 0. \tag{1.21}$$

In particular, mean-field critical behaviour holds to within polylogarithmic factors.

It is believed under the same hypotheses that, for example, there should exist an exponent $g > 0$ such that $\mathcal{X}_{\beta_c-\varepsilon} \asymp (\log 1/\varepsilon)^g \varepsilon^{-1}$ as $\varepsilon \downarrow 0$. Indeed, for nearest neighbour percolation on $\mathbb{Z}^d$ at the upper-critical dimension $d = 6$, Essam, Gaunt, and Guttmann [18] predicted that this holds with $g = 2/7$. See [8] and references therein for related results for weakly self-avoiding walk and the φ^4 model on hierarchical lattices.

Note that Theorem 1.8 is not an instance of Corollary 1.5 since the results of [28] do not *a priori* give any control of the rate of growth of the near-critical triangle $\nabla_{\beta_c-\varepsilon}$ when $\alpha = d/3$. Related results stating roughly that the exponents take values close to their mean-field values when α is slightly larger than $d/3$ are given in Theorem 5.1.

2 Derivation of the main differential inequality

In this section we prove Proposition 1.6. Our proof uses **Russo's formula** [20, Chapter 2.4], which states in our context that if $G = (V, E, J)$ is a countable weighted graph and $A \subseteq \{0,1\}^E$ is an increasing event depending on at most finitely many edges then

$$\frac{d}{d\beta}\mathbb{P}_\beta(A) = \sum_{e\in E} J_e e^{-\beta J_e}\mathbb{P}_\beta(e \text{ is pivotal for } A) = \sum_{e\in E} J_e \mathbb{P}_\beta(e \text{ is closed pivotal for } A).$$

Here we recall that a set $A \subseteq \{0,1\}^E$ is said to be increasing if $(\omega \in A) \Rightarrow (\omega' \in A)$ for every $\omega, \omega' \in \{0,1\}^E$ such that $\omega'(e) \geq \omega(e)$ for every $e \in E$, and that given a percolation configuration ω and an increasing event A, an edge $e \in E$ is said to be **pivotal** if $\omega \cup \{e\} \in A$ and $\omega \setminus \{e\} \notin A$. More generally [20, Equation 2.28], if A is an increasing event that may depend on infinitely many edges, we have the inequality

$$\left(\frac{d}{d\beta}\right)_+ \mathbb{P}_\beta(A) \geq \sum_{e\in E} J_e e^{-\beta J_e}\mathbb{P}_\beta(e \text{ is pivotal for } A) = \sum_{e\in E} J_e \mathbb{P}_\beta(e \text{ is closed pivotal for } A)$$

where $\left(\frac{d}{d\beta}\right)_+ \mathbb{P}_\beta(A) = \liminf_{\varepsilon\downarrow 0} \frac{1}{\varepsilon}\left(\mathbb{P}_{\beta+\varepsilon}(A) - \mathbb{P}_\beta(A)\right)$ is the **lower-right Dini derivative** of $\mathbb{P}_\beta(A)$. (For brevity we say that an edge is *closed pivotal* for A if it is closed and pivotal for A.)

Let $G = (V, E, J)$ be a countable weighted graph, and let S be a finite set of vertices of G. Let $\partial_E^{\rightarrow} S$ denote the set of oriented edges of G with $e^- \in S$ and $e^+ \notin S$. Given $u, v \in V \setminus S$, we write $\{u \leftrightarrow v \text{ off } S\}$ to mean that there exists an open path connecting u and v that does not visit any vertex of S. For each $\beta \geq 0$ let $\Phi_\beta(S)$ be defined by

$$\Phi_\beta(S) = \sum_{e \in \partial_E^{\rightarrow} S} \sum_{v \in V \setminus S} J_e \mathbb{P}_\beta(e^+ \leftrightarrow v \text{ off } S).$$

Note that an edge e is a closed pivotal for the event $\{o \leftrightarrow v\}$ if and only if there exists a (necessarily unique) orientation of e such that $e^- \in K$, $e^+ \notin K$, and v is connected to e^+ off of K. As such, Russo's formula implies that

$$\left(\frac{d}{d\beta}\right)_+ \mathbb{E}_\beta |K_v| \geq \mathbb{E}_\beta \left[\Phi_\beta(K_v)\right] \tag{2.1}$$

for every $\beta \geq 0$ and $v \in V$.

We begin our analysis with the following key lemma, which does not require transitivity. We write $\mathcal{X}_\beta^{\min} = \inf_{v \in V} \mathbb{E}_\beta |K_v|$ and $\mathcal{X}_\beta^{\max} = \sup_{v \in V} \mathbb{E}_\beta |K_v|$ so that $\mathcal{X}_\beta^{\min} = \mathcal{X}_\beta^{\max} = \mathcal{X}_\beta$ when G is transitive.

Lemma 2.1. *Let G be a countable weighted graph normalized so that $\sup_{v \in V} \sum_{e \in E_v^{\rightarrow}} J_e \leq 1$. Then*

$$\Phi_\beta(S) \geq \frac{\left(\mathcal{X}_\beta^{\min} |S| - \sum_{u,v \in S} T_\beta(u,v)\right)^2}{\beta^2 \mathcal{X}_\beta^{\max} \sum_{u,v,w \in S} T_\beta(u,v) T_\beta(u,w)} \tag{2.2}$$

for every $\beta > 0$ and every finite set of vertices S in G.

The proof of this lemma will apply the *van den Berg and Kesten* (BK) *inequality* [45] and the attendant notion of the *disjoint occurrence* $A \circ B$ of two sets A and B; we refer the unfamiliar reader to [20, Chapter 2.3] for background.

Proof of Lemma 2.1. Fix $S \subseteq V$ and $\beta > 0$. Observe that for each $u \in S$ and $w \notin S$ we have that

$$\{u \leftrightarrow w\} \subseteq \bigcup_{e \in \partial_E^{\rightarrow} S} \{u \leftrightarrow e^-\} \circ \{e \text{ open}\} \circ \{e^+ \leftrightarrow w \text{ off } S\}.$$

Indeed, if γ is a simple open path from u to w and e is the oriented edge that is crossed by γ as it leaves S for the last time, then the pieces of γ before and after it crosses e are disjoint witnesses for $\{u \leftrightarrow e^-\}$ and $\{e^+ \to w \text{ off } S\}$ that are both disjoint from e. Thus, applying a union bound and the BK inequality yields that

$$\begin{aligned} T_\beta(u,w) &\leq \sum_{v \in S} T_\beta(u,v) \sum_{e \in E_v^{\rightarrow}} \mathbb{1}(e^+ \notin S)(1 - e^{-\beta J_e}) \mathbb{P}_\beta(e^+ \leftrightarrow w \text{ off } S) \\ &\leq \beta \sum_{v \in S} T_\beta(u,v) \sum_{e \in E_v^{\rightarrow}} \mathbb{1}(e^+ \notin S) J_e \mathbb{P}_\beta(e^+ \leftrightarrow w \text{ off } S) \end{aligned}$$

for every $u \in S$ and $w \in V \setminus S$, where we used the inequality $1 - e^{-t} \leq t$ in the second line. Summing

over $u \in S$ and $w \in V \setminus S$, we deduce that

$$\sum_{w \in V \setminus S} \sum_{u \in S} T_\beta(u,w) \leq \beta \sum_{w \in V \setminus S} \sum_{u,v \in S} T_\beta(u,v) \sum_{e \in E_v^{\rightarrow}} \mathbb{1}(e^+ \notin S) J_e \mathbb{P}_\beta(e^+ \leftrightarrow w \text{ off } S).$$

Let $f : V \to \mathbb{R}$ be defined by

$$f(v) = \mathbb{1}(v \in S) \sum_{w \in V \setminus S} \sum_{e \in E_v^{\rightarrow}} \mathbb{1}(e^+ \notin S) J_e \mathbb{P}_\beta(e^+ \leftrightarrow w \text{ off } S)$$

so that the above inequality may be rewritten as

$$\sum_{w \in V \setminus S} \sum_{u \in S} T_\beta(u,w) \leq \beta \langle T_\beta \mathbb{1}_S, f \rangle.$$

Applying Cauchy-Schwarz and (a trivial special case of) Hölder's inequality, we obtain that

$$\sum_{w \in V \setminus S} \sum_{u \in S} T_\beta(u,w) \leq \beta \langle T_\beta \mathbb{1}_S, f \rangle \leq \beta \|T_\beta \mathbb{1}_S\|_2 \|f\|_2 \leq \beta \|T_\beta \mathbb{1}_S\|_2 \|f\|_1^{1/2} \|f\|_\infty^{1/2}. \tag{2.3}$$

We have from the definitions that $\|T_\beta \mathbb{1}_S\|_2^2 = \sum_{u,v,w \in S} T_\beta(u,v) T_\beta(u,w)$, $\|f\|_1 = \Phi_\beta(S)$, and

$$\sum_{w \in V \setminus S} \sum_{u \in S} T_\beta(u,w) = \sum_{w \in V} \sum_{u \in S} T_\beta(u,w) - \sum_{u,v \in S} T_\beta(u,w) \geq \mathcal{X}_\beta^{\min} |S| - \sum_{u,v \in S} T_\beta(u,v).$$

Since $\sum_{e \in E_v^{\rightarrow}} J_e \leq 1$ for every $v \in V$ we also have that

$$\|f\|_\infty \leq \sup_{v \in V} \sum_{w \in V} \sum_{e \in E_v^{\rightarrow}} J_e \mathbb{P}_\beta(e^+ \leftrightarrow w) \leq \mathcal{X}_\beta^{\max},$$

and the claim follows by substituting these four estimates into (2.3) and rearranging. □

We now deduce Proposition 1.6 from Lemma 2.1. Before beginning the proof, let us recall that the following trivial consequence of Cauchy-Schwarz inequality: If X and Y are random variables defined on the same probability space such that Y is positive almost surely and $\mathbb{E}[Y]$ is finite then

$$\mathbb{E}\left[\frac{X^2}{Y}\right] \geq \frac{\mathbb{E}[X]^2}{\mathbb{E}[Y]}.$$

Proof of Proposition 1.6. Recall that K denotes the cluster of the origin. We have by Russo's formula and (2.2) that

$$\frac{d\mathcal{X}_\beta}{d\beta} \geq \mathbb{E}_p\left[\Phi_\beta(K)\right] \geq \mathbb{E}_\beta\left[\frac{\left(\mathcal{X}_\beta |K| - \sum_{x,y \in K} T_\beta(x,y)\right)^2}{\beta^2 \mathcal{X}_\beta \sum_{a,b,c \in K} T_\beta(a,b) T_\beta(a,c)}\right]$$

for every $0 < \beta < \beta_c$. (As previously mentioned, $\mathcal{X}_\beta$ is a smooth function of β on $[0, \beta_c)$ [20, Chapter 6.4], so that its usual derivative and lower-right Dini derivative coincide on this interval.) Applying the Cauchy-Schwarz inequality as above with the random variables $X = \mathcal{X}_\beta |K| - \sum_{x,y \in K} T_\beta(x,y)$

and $Y = \sum_{a,b,c \in K} T_\beta(a,b)T_\beta(a,c)$ yields that

$$\frac{d\mathcal{X}_\beta}{d\beta} \geq \frac{1}{\beta^2 \mathcal{X}_\beta} \mathbb{E}_\beta \left[\mathcal{X}_\beta |K| - \sum_{x,y \in K} T_\beta(x,y) \right]^2 \mathbb{E}_\beta \left[\sum_{a,b,c \in K} T_\beta(a,b)T_\beta(a,c) \right]^{-1} \tag{2.4}$$

for every $0 < \beta < \beta_c$. (Note that the random variable $\sum_{a,b,c \in K} T_\beta(a,b)T_\beta(a,c)$ is bounded by $|K|^3$ and is therefore integrable for $\beta < \beta_c$ by sharpness of the phase transition [1, 17, 26].) Thus, to complete the proof it suffices to prove that

$$\mathbb{E}_\beta \left[\sum_{x,y \in K} T_\beta(x,y) \right] \leq \mathcal{X}_\beta \nabla_\beta \quad \text{and} \quad \mathbb{E}_\beta \left[\sum_{a,b,c \in K} T_\beta(a,b)T_\beta(a,c) \right] \leq (2A_\beta + B_\beta)\mathcal{X}_\beta \tag{2.5}$$

for every $0 < \beta < \beta_c$; the claimed estimate (1.17) will then follow by substituting (2.5) into (2.4).

Both inequalities of (2.5) will follow by the same reasoning used to prove the "tree graph inequalities" of Aizenman and Newman [4], an account of which can also be found in [20, Chapter 6.3]. We begin with the first of the two inequalities claimed in (2.5). Let $x, y \in V$, and suppose that x and y are both connected to o. We claim that there must exist a vertex $w \in V$ (possibly equal to one of o, x, or y) such that the event $\{o \leftrightarrow w\} \circ \{w \leftrightarrow x\} \circ \{w \leftrightarrow y\}$ holds. Indeed, if γ_1 is an open simple path from v to x and γ_2 is an open simple path from v to y, then the last vertex of γ_1 visited by γ_2 has this property. Thus, we have by the BK inequality and the union bound that

$$\mathbb{P}_\beta(x, y \in K) \leq \sum_{w \in V} T_\beta(o,w)T_\beta(w,x)T_\beta(w,y) \tag{2.6}$$

for every $v, x, y \in V$. It follows that

$$\begin{aligned} \mathbb{E}_\beta \left[\sum_{x,y \in K} T_\beta(x,y) \right] &= \sum_{x,y \in V} \mathbb{P}_\beta(x, y \in K) T_\beta(x,y) \\ &\leq \sum_{x,y,w \in V} T_\beta(o,w)T_\beta(w,x)T_\beta(w,y)T_\beta(x,y) = \mathcal{X}_\beta \nabla_\beta \end{aligned}$$

as claimed. We now turn to the second inequality of (2.5), whose proof is very similar although the required expressions are larger. Let $a, b, c \in V$. It follows by similar reasoning to the two-vertex case above that if a, b, and c are all connected to o then there exist vertices x and y such that at least one of the events

$$\begin{aligned} &\{o \leftrightarrow x\} \circ \{x \leftrightarrow a\} \circ \{x \leftrightarrow y\} \circ \{y \leftrightarrow b\} \circ \{y \leftrightarrow c\}, \\ &\{o \leftrightarrow x\} \circ \{x \leftrightarrow b\} \circ \{x \leftrightarrow y\} \circ \{y \leftrightarrow a\} \circ \{y \leftrightarrow c\}, \text{ or} \\ &\{o \leftrightarrow x\} \circ \{x \leftrightarrow c\} \circ \{x \leftrightarrow y\} \circ \{y \leftrightarrow a\} \circ \{y \leftrightarrow b\} \end{aligned}$$

holds. See [20, Equation 6.93] and its proof for details. Thus, it follows from a union bound and the

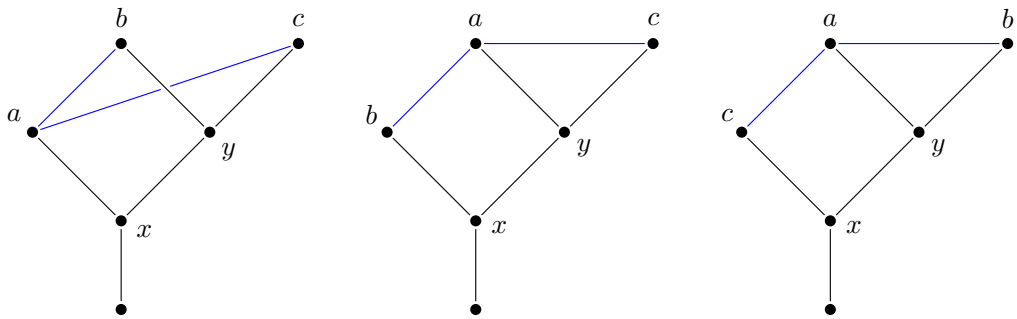

Figure 2: Diagrammatic representations of the three sums contributing to the upper bound on the expectation of $\mathbb{E}_\beta \sum_{a,b,c\in K} T_\beta(a,b)T_\beta(a,c)$. Fixing a, b, and c, weighting each black edge by T_β, and summing over x and y gives the tree-graph bound on the probability that a, b, and c all belong to the cluster of the origin.

BK inequality that

$$\begin{aligned}\mathbb{P}_\beta(a,b,c\in K) \leq \sum_{x,y\in V} & T_\beta(o,x)T_\beta(x,a)T_\beta(x,y)T_\beta(y,b)T_\beta(y,c) \\ &+ \sum_{x,y\in V} T_\beta(o,x)T_\beta(x,b)T_\beta(x,y)T_\beta(y,a)T_\beta(y,c) \\ &+ \sum_{x,y\in V} T_\beta(o,x)T_\beta(x,c)T_\beta(x,y)T_\beta(y,a)T_\beta(y,b) \qquad (2.7)\end{aligned}$$

for every $a,b,c\in V$ and $\beta\geq 0$. Summing over $a,b,c\in V$, we deduce from this together with the definitions of A_β and B_β that

$$\mathbb{E}_\beta\left[\sum_{a,b,c\in K} T_\beta(a,b)T_\beta(a,c)\right] = \sum_{a,b,c\in V} \mathbb{P}_\beta(a,b,c\in K)T_\beta(a,b)T_\beta(a,c) \leq \mathcal{X}_\beta(2A_\beta + B_\beta)$$

for every $0\leq\beta<\beta_c$ as claimed; see Figure 2 for graphical representations of the resulting diagrammatic sums. □

3 Comparing diagrams: a sojourn into linear algebra

In this section we prove Proposition 1.7, which states that $B_\beta \leq A_\beta \leq \nabla_\beta^2$ for every unimodular transitive weighted graph G and every $\beta\geq 0$. The inequality $B_\beta\leq A_\beta$ is by far the more subtle to establish. We begin by proving the proposition in the special case of $G=\mathbb{Z}^d$ using Fourier analysis. Note that a very similar proof may also be implemented on the Hierarchical lattice since it is a locally compact Abelian group and hence also admits a version of the Fourier transform. We will use the alternative expressions for A_β and B_β derived from the mass-transport principle in (1.15) and (1.16) throughout this section.

Proof of Proposition 1.7 in the case $G=\mathbb{Z}^d$ via Fourier analysis. Let $0\leq\beta<\beta_c$ and let $\mathbb{T}^d = [-\pi,\pi]^d$ be the d-dimensional torus. Let $\tau_\beta(x) = T_\beta(0,x)$, which belongs to $\ell^1(\mathbb{Z}^d)$ by sharpness

of the phase transition, and let $\hat{\tau}_\beta : \mathbb{T}^d \to \mathbb{C}$ be its Fourier transform

$$\hat{\tau}_\beta(\theta) = \sum_{x \in \mathbb{Z}^d} \tau_\beta(x) e^{i\theta \cdot x}.$$

The symmetry $\tau_\beta(x) = \tau_\beta(-x)$ ensures that $\hat{\tau}_\beta(\theta) \in \mathbb{R}$ for every $\theta \in \mathbb{T}^d$, while the fact that τ_β is real-valued implies that $\hat{\tau}_\beta(\theta) = \hat{\tau}_\beta(-\theta)$ for every $\theta \in \mathbb{T}^d$. Moreover, it is a lemma due to Aizenman and Newman [4, Lemma 3.3] (see also Lemma 3.1 below) that the matrix T_β is positive semidefinite and hence that $\hat{\tau}_\beta(\theta)$ is non-negative for every $\theta \in \mathbb{T}^d$.

Letting $*$ denote convolution (on either $\mathbb{Z}^d$ or $\mathbb{T}^d$, taken with respect to either counting measure or normalized Lebesgue measure as appropriate) and using that Fourier transforms exchange the roles of convolution and multiplication, we can write the diagrammatic sums A_β and B_β as

$$A_\beta = \sum_{x \in \mathbb{Z}^d} \tau_\beta(x)[\tau_\beta * \tau_\beta](x)[\tau_\beta * \tau_\beta * \tau_\beta](x) = [\hat{\tau}_\beta * \hat{\tau}_\beta^2 * \hat{\tau}_\beta^3](0)$$

and

$$B_\beta = \sum_{x \in \mathbb{Z}^d} [\tau_\beta * \tau_\beta](x)^3 = \hat{\tau}_\beta^2 * \hat{\tau}_\beta^2 * \hat{\tau}_\beta^2(0),$$

where we used that $\sum_{x \in \mathbb{Z}^d} f(x) = \hat{f}(0)$ for every $f \in \ell^1(\mathbb{Z}^d)$. Similarly, we can express ∇_β via either of the two equivalent expressions

$$\nabla_\beta = \sum_{x \in \mathbb{Z}^d} \tau_\beta(x)[\tau_\beta * \tau_\beta](x) = [\hat{\tau}_\beta * \hat{\tau}_\beta^2](0) \quad \text{and} \quad \nabla_\beta = [\tau_\beta * \tau_\beta * \tau_\beta](0) = \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} \hat{\tau}_\beta^3(\theta) \, \mathrm{d}\theta.$$

This Fourier-analytic perspective can easily be used to prove the bound $A_\beta \leq \nabla_\beta^2$. Indeed, since $\hat{\tau}_\beta$ is non-negative we may apply Hölder's inequality to deduce that

$$\begin{aligned} \hat{\tau}_\beta * \hat{\tau}_\beta^2(\theta) &= \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} \hat{\tau}_\beta(\phi) \hat{\tau}_\beta^2(\theta - \phi) \, \mathrm{d}\phi \leq \frac{1}{(2\pi)^d} \left[\int_{\mathbb{T}^d} \hat{\tau}_\beta(\phi)^3 \, \mathrm{d}\phi \right]^{1/3} \left[\int_{\mathbb{T}^d} \hat{\tau}_\beta(\theta - \phi)^3 \, \mathrm{d}\phi \right]^{2/3} \\ &= \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} \hat{\tau}_\beta(\phi)^3 \, \mathrm{d}\phi = \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} \hat{\tau}_\beta(\phi) \hat{\tau}_\beta^2(-\phi) \, \mathrm{d}\phi = \hat{\tau}_\beta * \hat{\tau}_\beta^2(0) \end{aligned}$$

for every $\theta \in \mathbb{T}^d$, where we used that $\hat{\tau}_\beta$ is even in the second equality on the second line, and hence that

$$A_\beta = \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} [\hat{\tau}_\beta * \hat{\tau}_\beta^2](\phi) \hat{\tau}_\beta^3(-\phi) \, \mathrm{d}\phi \leq [\hat{\tau}_\beta * \hat{\tau}_\beta^2](0) \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} \hat{\tau}_\beta^3(-\phi) \, \mathrm{d}\phi = \nabla_\beta^2$$

as claimed, where we used the non-negativity of $\hat{\tau}_\beta$ in the central inequality.

It remains to establish the less obvious bound $B_\beta \leq A_\beta$. To this end, we claim that if $f : \mathbb{T}^d \to [0, \infty)$ is an arbitrary non-negative function then

$$f^2 * f^2(\theta) \leq f * f^3(\theta) \tag{3.1}$$

for every $\theta \in \mathbb{T}^d$. Indeed, if $f, g : \mathbb{T}^d \to [0, \infty)$ are any two functions then the AM-GM inequality

implies that

$$\begin{aligned}[f^2 * g^2](\theta) &= \frac{1}{(2\pi)^d}\int_{\mathbb{T}^d} f^2(\phi)g^2(\theta-\phi)\,\mathrm{d}\phi \\ &\leq \frac{1}{(2\pi)^d}\int_{\mathbb{T}^d}\frac{1}{2}\Big[f(\phi)g^3(\theta-\phi)+f^3(\phi)g(\theta-\phi)\Big]\,\mathrm{d}\phi = \frac{1}{2}\Big([f*g^3](\theta)+[f^3*g](\theta)\Big)\end{aligned}$$

for every $\theta \in \mathbb{T}^d$, and the claim follows by taking $f = g$. Applying the inequality (3.1) with $f = \hat{\tau}_\beta$ and using positivity of $\hat{\tau}_\beta^2$ we deduce that $\hat{\tau}_\beta^2 * \hat{\tau}_\beta^2 * \hat{\tau}_\beta^2(\theta) \leq \hat{\tau}_\beta * \hat{\tau}_\beta^3 * \hat{\tau}_\beta^2(\theta)$ for every $\theta \in \mathbb{T}^d$ and consequently that $B_\beta \leq A_\beta$ as claimed. □

The goal of the remainder section is to remove all the Fourier analysis from the above proof so that it may be carried out on an arbitrary unimodular transitive weighted graph. If the reader is only interested in the cases $G = \mathbb{Z}^d$ and $G = \mathbb{H}_L^d$ they may safely skip the rest of this subsection. Let us stress however that we believe that the tools we introduce below are in fact rather elegant and, in some cases, do not appear to have been used in probability theory before.

The general proof of Proposition 1.7 will be based on the theory of infinite positive (semi)definite matrices and the Schur product theorem, with a particularly important role being played by a theorem of Ando [6] concerning the interplay between the *Hadamard product* (i.e., the entrywise product) and the usual matrix product. Since the results we use are spread around several disparate sources and since we expect the content of this section to be highly unfamiliar to most probabilists, we give a thorough, self-contained, and probabilist-friendly treatment of most of the results that we use. A secondary purpose of this self-contained account is to prove infinite-dimensional cases of various results that are usually stated only in the finite-dimensional case. In addition to the paper on Ando [6], the primary sources we used to prepare this section were [9] and [14, Appendix A], both of which are written in a very accessible manner.

Let V be a countable (finite or infinite) set, and denote by $\mathcal{M}(V) = \mathbb{R}^{V\times V}$ the set of all matrices indexed by V. Let $L^0(V)$ be the set of *finitely supported* functions $L^0(V) := \{f \in \mathbb{R}^V : f(v) = 0$ for all but finitely many $v\}$, so that every $T \in \mathcal{M}(V)$ defines a linear map

$$\begin{array}{rcl} L^0(V) & \longrightarrow & \mathbb{R}^V \\ f & \longmapsto & Tf \end{array} \qquad \text{where} \qquad Tf(u) = \sum_{v\in V} T(u,v)f(v) \quad \text{for each } u \in V$$

in the natural way, where the assumption that $f \in L^0(V)$ means that convergence is not an issue.

A matrix $T \in \mathcal{M}(V)$ is said to be **symmetric** if $T(u,v) = T(v,u)$ for every $u, v \in V$. A matrix $T \in \mathcal{M}(V)$ is said to be **positive semidefinite** if it symmetric and satisfies $\langle Tf, f\rangle := \sum_{u,v\in V} f(u)T(u,v)f(v) \geq 0$ for every $f \in L^0(V)$ and is said to be **positive definite** if there exists $\lambda > 0$ such that $\langle Tf, f\rangle \geq \lambda\langle f, f\rangle$ for every $f \in L^0(V)$. (Again, restricting to $f \in L^0(V)$ means that both inner products are trivially well-defined.) Given $S, T \in \mathcal{M}(V)$, we write $S \geq T$ to mean that $S - T$ is positive semidefinite and write $S > T$ to mean that $S - T$ is positive definite.

The relevance of these notions to percolation theory is established by the following lemma of Aizenman and Newman [4, Lemma 3.3]. We include a proof for completeness.

Lemma 3.1 (Aizenman and Newman)**.** *Let G be a countable weighted graph. Then T_β is positive* semidefinite *for every $\beta \geq 0$. Moreover, if G has $\sup_{v\in V}\sum_{e\in E_v^{\to}} J_e < \infty$ then T_β is positive* definite *for every $\beta \geq 0$.*

Proof of Lemma 3.1. The matrix T_β is trivially symmetric. Let $\mathscr{C}$ be the set of clusters of ω, each of which is a set of vertices. Then for each finitely supported $f : V \to \mathbb{R}$ we have that

$$\begin{aligned}\langle T_\beta f, f\rangle &= \sum_{u,v\in V} f(u)f(v)\mathbb{P}_\beta(u \leftrightarrow v) = \mathbb{E}_\beta\left[\sum_{u,v\in V} f(u)f(v)\mathbb{1}(u \leftrightarrow v)\right] \\ &= \mathbb{E}_\beta\left[\sum_{C\in\mathscr{C}}\Big(\sum_{u\in C} f(u)\Big)^2\right] \geq 0, \end{aligned} \tag{3.2}$$

so that T_β is positive semidefinite as required. Similarly, letting $\mathscr{C}_1 \subseteq \mathscr{C}$ be the set of singleton clusters of ω, that is, clusters containing exactly one vertex, we have by (3.2) that

$$\begin{aligned}\langle T_\beta f, f\rangle \geq \mathbb{E}_\beta\left[\sum_{C\in\mathscr{C}_1}\Big(\sum_{u\in C} f(u)\Big)^2\right] &= \sum_{u\in V} f(u)^2\mathbb{P}_\beta(\{u\} \text{ is a singleton cluster}) \\ &= \sum_{u\in V} f(u)^2 \prod_{e\in E_v^{\to}} e^{-\beta J_e} \end{aligned} \tag{3.3}$$

which is easily seen to establish the second claim. □

Remark 3.2. The proof that the two-point matrix is positive semidefinite works for *any* percolation process, that is, any random subgraph of a given graph, and is not specific to Bernoulli percolation. Similarly, the proof of positive definiteness applies to any percolation process in which the probability that a vertex lies in a singleton cluster is bounded away from zero.

Define $\mathcal{M}^2(V) \subseteq \mathcal{M}(V)$ to be the algebra of V-indexed matrices T such that $\langle Tf, Tf\rangle = \sum_{u\in V}(\sum_{v\in V}|T(u,v)f(v)|)^2 < \infty$ for every $f \in L^2(V)$, so that $Tf \in \mathbb{R}^V$ is well-defined and belongs to $L^2(V)$ for every $f \in L^2(V)$. It follows from the principle of uniform boundedness that $T \in \mathcal{M}^2(V)$ if and only if

$$\|T\|_{2\to 2} = \sup\left\{\frac{\langle Tf, Tf\rangle^{1/2}}{\langle f, f\rangle^{1/2}} : f \in L^2(V), \langle f,\, f\rangle > 0\right\} < \infty,$$

where the inner product $\langle Tf, Tf\rangle$ is always well-defined as an element of $[0,\infty]$ since it is a sum of non-negative terms. Thus, if $T \in \mathcal{M}^2(V)$ then the associated linear map $f \mapsto Tf$ from $L^2(V)$ to itself is continuous. Since $L^0(V)$ is dense in $L^2(V)$ it follows that a symmetric matrix $T \in \mathcal{M}^2(V)$ is positive semidefinite if and only if $\langle Tf, f\rangle \geq 0$ for every $f \in L^2(V)$ and that T is positive definite if and only if there exists $\lambda > 0$ such that $\langle Tf, f\rangle \geq \lambda$ for every $f \in L^2(V)$.

It is a consequence of the Cauchy-Schwarz inequality that

$$\|T\|_{2\to 2} \leq \|T\|_{1\to 1}^{1/2}\|T\|_{\infty\to\infty}^{1/2} := \sqrt{\sup_{u\in V}\sum_{v\in V}|T(u,v)|}\sqrt{\sup_{u\in V}\sum_{v\in V}|T(v,u)|}$$

for every $T \in \mathcal{M}(V)$ (this is a special case of the Riesz-Thorin theorem), and when T is symmetric this simplifies to

$$\|T\|_{2\to 2} \leq \|T\|_{1\to 1} := \sup_{u\in V} \sum_{v\in V} |T(u,v)|.$$

In particular, it follows that if G is a transitive weighted graph then the associated two-point matrix T_β satisfies $\|T_\beta\|_{2\to 2} \leq \|T_\beta\|_{1\to 1} = \mathcal{X}_\beta$ for every $\beta \geq 0$ and hence by the sharpness of the phase transition that T_β defines a bounded operator on $L^2(V)$ for every $0 \leq \beta < \beta_c$. (In fact it is possible in some contexts for this boundedness to continue hold for β strictly larger than β_c, see [25] for a thorough discussion.) Moreover, it follows by spectral considerations that if $T \in \mathcal{M}^2(V)$ is positive (semi)definite then so is T^k for every $k \geq 1$, so that if G is an infinite, connected, locally finite, quasi-transitive graph then T_β^k is positive definite for every $0 \leq \beta < \beta_c$ and $k \geq 1$.

Let us now prove the easy part of Proposition 1.7.

Lemma 3.3. *Let G be a connected, unimodular transitive weighted graph. Then $A_\beta \leq \nabla_\beta^2$ for every $0 \leq \beta \leq \beta_c$.*

Proof of Lemma 3.3. By left continuity of T_β, it suffices to prove the claim for every $0 \leq \beta < \beta_c$. Fix one such β. Since G is unimodular, we have by (1.15) that $A_\beta = \sum_{v\in V} T_\beta(o,v)T_\beta^2(o,v)T_\beta^3(o,v)$. Since T_β^3 is positive definite, it has the property that

$$\langle T_\beta^3(\mathbb{1}_u - \mathbb{1}_v), (\mathbb{1}_u - \mathbb{1}_v)\rangle = T_\beta^3(u,u) + T_\beta^3(v,v) - 2T_\beta^3(u,v) \geq 0$$

for every $u, v \in V$ and hence by transitivity that

$$T_\beta^3(u,v) \leq \frac{1}{2}\left[T_\beta^3(u,u) + T_\beta^3(v,v)\right] = T_\beta^3(o,o) = \nabla_\beta$$

for every $u, v \in V$. It follows that

$$A_\beta \leq \nabla_\beta \sum_{v\in V} T_\beta(o,v)T_\beta^2(v,o) = \nabla_\beta T_\beta^3(o,o) = \nabla_\beta^2$$

as claimed. □

It remains to prove that $B_\beta \leq A_\beta$. This proof will rely on various more sophisticated pieces of technology, which we now begin to introduce. Recall that if $S, T \in \mathcal{M}(V)$ are V-indexed matrices, the **Hadamard product** $S \circ T \in \mathcal{M}(V)$ is defined by entrywise multiplication

$$[S \circ T](u,v) := S(u,v)T(u,v) \qquad \text{for every } u, v \in V.$$

Note that if $S, T \in \mathcal{M}^2(V)$ then $S \circ T \in \mathcal{M}^2(V)$ also and indeed satisfies $\|S \circ T\|_{2\to 2} \leq \|S\|_{2\to 2}\|T\|_{2\to 2}$ [9, Equation 2.2]. The Hadamard product may be used to give various alternative expressions for the diagrams A_β and B_β. The most obvious way to do this, which is analogous to what we did in the case $\mathbb{Z}^d$, is to use (1.15) and (1.16) to write

$$A_\beta = \sum_{v\in V}\left[T_\beta \circ T_\beta^2 \circ T_\beta^3\right](o,v) \qquad \textit{and} \qquad B_\beta = \sum_{v\in V}\left[T_\beta^2 \circ T_\beta^2 \circ T_\beta^2\right](o,v).$$

However, we will see that the equivalent expressions

$$A_\beta = \Big[T_\beta \circ T_\beta^3\Big] T_\beta^2(o,o) \qquad \textit{and} \qquad B_\beta = \Big[T_\beta^2 \circ T_\beta^2\Big] T_\beta^2(o,o)$$

are better suited to our needs. We stress that these expressions contain both Hadamard products and ordinary matrix products: for example, T_β^2 denotes the ordinary square of T_β while $[T_\beta \circ T_\beta^3]T_\beta^2$ denotes the ordinary matrix product of $T_\beta \circ T_\beta^3$ with T_β^2.

The key estimate needed to complete the proof of Proposition 1.7 is the following proposition, which can be thought of as a generalisation of the inequality (3.1).

Proposition 3.4. *Let V be a countable set and let $T \in \mathcal{M}^2(V)$ be positive semidefinite. Then*

$$T^2 \circ T^2 \leq T \circ T^3.$$

We will deduce this proposition as a special case of a theorem of Ando [6], which in turn is a (non-obvious) consequence of the Schur product theorem. While all these theorems are usually stated for finite matrices, they are also valid for infinite matrices; indeed, we will see that the infinite cases of these theorems can be immediately reduced to the finite cases.

Theorem 3.5 (Schur product theorem)**.** *Let V be countable, and let $S, T \in \mathcal{M}(V)$. If S and T are both positive semidefinite, then $S \circ T$ is positive semidefinite.*

Proof of Theorem 3.5. Recall that if V is a countable set and $T \in \mathcal{M}(V)$ the matrices $T|_W \in \mathcal{M}(W)$ where $W \subseteq V$ and $T|_W(u,v) = T(u,v)$ for every $u, v \in W$ are referred to as the **principal submatrices** of T. By the definitions, a matrix $T \in \mathcal{M}(V)$ is positive definite if and only if all its finite dimensional principal submatrices are. Thus, since the Hadamard product commutes with taking principal submatrices in the sense that $S|_W \circ T|_W = (S \circ T)|_W$ for every $S, T \in \mathcal{M}(V)$ and $W \subseteq V$, it suffices to consider the case that V is finite. That is, the countably infinite case of the Schur product theorem immediately reduces to the finite case, which is classical; see https://en.wikipedia.org/wiki/Schur_product_theorem for three distinct proofs. (A further proof works by observing that $S \circ T$ is a principal submatrix of the tensor product $S \otimes T$, which in turn is easily seen to be positive definite as a direct consequence of the relevant definitions.) □

Before stating Ando's theorem on Hadamard products, we recall that for every positive semidefinite matrix $T \in \mathcal{M}^2(V)$ there exists a unique positive semidefinite matrix $T^{1/2} \in \mathcal{M}^2(V)$, known as the **principal square root** of T, such that $(T^{1/2})^2 = T$. (Be careful to note that there may exist other matrices S for which $S^2 = T$ or $SS^\perp = T$, but any such matrix other than $T^{1/2}$ will not be positive semidefinite.) Moreover, the principal square root $T^{1/2}$ of T commutes with T under multiplication in the sense that $T^{1/2}T = TT^{1/2}$. Indeed, the matrix $T^{1/2}T$ is also equal to the principal square root of T^3 and is denoted simply by $T^{3/2}$. Similarly, one may define for each $\alpha \geq 0$ a symmetric, positive semidefinite matrix T^α such that $(T^\alpha)_{\alpha \geq 0}$ is a semigroup in the sense that T^α depends continuously on α and $T^\alpha T^\beta = T^\beta T^\alpha = T^{\alpha+\beta}$ for every $\alpha, \beta \geq 0$. All of these claims are immediate consequences of the spectral theorem for bounded self-adjoint operators, see e.g. [41, Chapter VII] for further background. Similar spectral considerations imply that every positive definite matrix $T \in \mathcal{M}^2(V)$ has a well-defined inverse $T^{-1} \in \mathcal{M}^2(V)$ with $\|T^{-1}\|_{2\to 2}^{-1} = \sup\{\lambda : \langle Tf, f\rangle \geq \lambda \langle f, f\rangle$ for every $f \in L^0(V)\}$.

We are now ready to state the aforementioned theorem of Ando, which is a special case of [6, Theorem 12]. While Ando stated his theorem for finite-dimensional matrices, the proof extends easily to the infinite-dimensional case. We remark that the paper of Ando contains a huge number of further related inequalities.

Theorem 3.6 (Ando)**.** *Let V be a countable set, and let $S, T \in \mathcal{M}^2(V)$ be commuting, positive semidefinite matrices. Then*

$$(ST)^{1/2} \circ (ST)^{1/2} \leq S \circ T.$$

As discussed above, we include a full proof of this theorem with the dual aims of making it accessible to probabilists and verifying the infinite-dimensional case. Before starting the proof, let us note that the theorem immediately implies Proposition 3.4.

Proof of Proposition 3.4 given Theorem 3.6. Let $T \in \mathcal{M}^2(V)$ be positive semidefinite. Applying Theorem 3.6 to the commuting, positive semidefinite matrices T and T^3 yields that $T^2 \circ T^2 \leq T \circ T^3$ as claimed. □

The proof of Theorem 3.6 will proceed by applying the Schur product theorem to various judiciously defined *block matrices*. This methodology, which was pioneered by Ando, also leads to many further interesting inequalities for the Hadamard product, see [9] and references therein. Let V be a countable set and consider the disjoint union $V \sqcup V = V \times \{1, 2\}$. Given $A, B, C, D \in \mathcal{M}(V)$, we define the **block matrix**

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \in \mathcal{M}(V \sqcup V) \qquad \text{by} \qquad \begin{bmatrix} A & B \\ C & D \end{bmatrix} \big((u,i),(v,j)\big) = \begin{cases} A(u,v) & i=1,\, j=1 \\ C(u,v) & i=2,\, j=1 \\ B(u,v) & i=1,\, j=2 \\ D(u,v) & i=2,\, j=2 \end{cases}$$

so that if $f, g \in L^0(V)$ and we define

$$\begin{bmatrix} f \\ g \end{bmatrix} \in L^0(V \sqcup V) \quad \text{by} \quad \begin{bmatrix} f \\ g \end{bmatrix}(v,i) = \begin{cases} f(v) & i=1 \\ g(v) & i=2 \end{cases} \qquad \text{then} \qquad \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} f \\ g \end{bmatrix} = \begin{bmatrix} Af + Bg \\ Cf + Dg \end{bmatrix}.$$

Similarly, if $A, B, C, D, E, F, G, H \in \mathcal{M}^2(V)$ then we can compute the products of the associated block matrices in the usual way

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} E & F \\ G & H \end{bmatrix} = \begin{bmatrix} AE + BG & AF + BH \\ CE + DG & CG + DH \end{bmatrix} \in \mathcal{M}^2(V).$$

For each $T \in \mathcal{M}(V)$, we define $T^\perp$ to be the transpose matrix $T^\perp(u,v) = T(v,u)$. The following lemma is an infinite-dimensional version of [9, Theorem 1.3.3].

Lemma 3.7. *Let V be a countable set, and let $S, T, X \in \mathcal{M}^2(V)$ be such that S is positive semidefinite and T is positive definite. Then*

$$\begin{bmatrix} S & X \\ X^\perp & T \end{bmatrix} \geq 0 \qquad \textit{if and only if} \qquad S \geq XT^{-1}X^\perp.$$

Before proving this lemma, let us note that if $S, T \in \mathcal{M}^2(V)$ are *congruent* in the sense that there exists a matrix $C \in \mathcal{M}^2(V)$ with inverse $C^{-1} \in \mathcal{M}^2(V)$ such that $S = CTC^{\perp}$, then S is positive semidefinite if and only if T is. Indeed, first note that congruence is an equivalence relation since if $S = CTC^{\perp}$ then $T = C^{-1}S(C^{-1})^{\perp}$. Congruence is also trivially seen to preserve symmetry. To conclude, note that if T is positive semidefinite then we have that $\langle Sf, f\rangle = \langle CTC^{\perp}f, f\rangle = \langle T(C^{\perp}f), (C^{\perp}f)\rangle \geq 0$ for every $f \in L^2(V)$ so that S is positive semidefinite as claimed.

Proof of Lemma 3.7. Since T is positive definite, it is invertible with inverse $T^{-1} \in \mathcal{M}^2(V)$. Write I and O for the identity and zero matrices indexed by V respectively. The block matrix we are interested in is congruent to

$$\begin{aligned}\begin{bmatrix} I & -XT^{-1} \\ O & I \end{bmatrix} \begin{bmatrix} S & X \\ X^{\perp} & T \end{bmatrix} \begin{bmatrix} I & -XT^{-1} \\ O & I \end{bmatrix}^{\perp} &= \begin{bmatrix} I & -XT^{-1} \\ O & I \end{bmatrix} \begin{bmatrix} S & X \\ X^{\perp} & T \end{bmatrix} \begin{bmatrix} I & O \\ -T^{-1}X^{\perp} & I \end{bmatrix} \\ &= \begin{bmatrix} S - XT^{-1}X^{\perp} & O \\ O & T \end{bmatrix}, \end{aligned} \tag{3.4}$$

where we note that this is indeed a congruence since

$$\begin{bmatrix} I & -XT^{-1} \\ O & I \end{bmatrix} \quad \text{is invertible with inverse} \quad \begin{bmatrix} I & -XT^{-1} \\ O & I \end{bmatrix}^{-1} = \begin{bmatrix} I & XT^{-1} \\ O & I \end{bmatrix}.$$

The matrix on the right hand side of (3.4) is positive semidefinite if and only if

$$\left\langle \begin{bmatrix} S - XT^{-1}X^{\perp} & O \\ O & T \end{bmatrix} \begin{bmatrix} f \\ g \end{bmatrix}, \begin{bmatrix} f \\ g \end{bmatrix} \right\rangle = \left\langle (S - XT^{-1}X^{\perp})f, f \right\rangle + \langle Tg, g\rangle \geq 0,$$

for every $f, g \in L^0(V)$, and since T is positive definite this is the case if and only if $S - XT^{-1}X^{\perp}$ is positive semidefinite as claimed. □

We will also use the following easy fact. Note that the implication here is only in one direction.

Lemma 3.8. *Let V be countable and let $S, T \in \mathcal{M}(V)$ be symmetric matrices.*

1. *If the block matrix $\begin{bmatrix} S & T \\ T & S \end{bmatrix}$ is positive semidefinite then $S - T$ is positive semidefinite.*
2. *If the block matrix $\begin{bmatrix} S & T \\ T & S \end{bmatrix}$ is positive definite then $S - T$ is positive definite.*

Proof of Lemma 3.8. We can compute that

$$\left\langle \begin{bmatrix} S & T \\ T & S \end{bmatrix} \begin{bmatrix} f \\ -f \end{bmatrix}, \begin{bmatrix} f \\ -f \end{bmatrix} \right\rangle = 2\left\langle (S - T)f, f \right\rangle$$

for every $f \in L^0(V)$, and the claim follows from the definitions. □

We are now ready to prove Theorem 3.6.

Proof of Theorem 3.6. The assumption that S and T commute implies that $(ST)^{1/2} = (TS)^{1/2} = ((ST)^{1/2})^{\perp}$. In particular, we have that $S = (ST)^{1/2}T^{-1}(ST)^{1/2}$ and $T = (ST)^{1/2}S^{-1}(ST)^{1/2}$, and we deduce from Lemma 3.7 that the block matrices

$$\begin{bmatrix} S & (ST)^{1/2} \\ (ST)^{1/2} & T \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} T & (ST)^{1/2} \\ (ST)^{1/2} & S \end{bmatrix}$$

are both positive semidefinite. Applying the Schur product theorem we deduce that

$$\begin{bmatrix} S \circ T & (ST)^{1/2} \circ (ST)^{1/2} \\ (ST)^{1/2} \circ (ST)^{1/2} & S \circ T \end{bmatrix}$$

is positive semidefinite also, so that the claim follows from Lemma 3.8. □

The trace on the group von Neumann algebra. We now wish to apply Ando's theorem to prove Proposition 1.7. To do this, we will employ the notion of the *trace* on the *von Neumann algebra* associated to the action of $\mathrm{Aut}(G)$ on G. We assure the unfamiliar reader that, despite the intimidating name, this is in fact a very simple notion.

Given a connected, transitive weighted graph $G = (V, E, J)$ with automorphism group $\mathrm{Aut}(G)$, we define the von Neumann algebra $\mathcal{A}(G)$ to be the set of matrices $T \in \mathcal{M}^2(V)$ that are diagonally invariant under the action of $\mathrm{Aut}(G)$ on V in the sense that $T(\gamma u, \gamma v) = T(u, v)$ for every $u, v \in V$ and $\gamma \in \mathrm{Aut}(G)$. Note that $\mathcal{A}(G)$ is indeed a von Neumann algebra in the sense that it is a weak*-closed subspace of $\mathcal{M}^2(V) \cong \mathcal{B}(L^2(V))$ that is closed under multiplication and adjunction. It follows trivially from the definitions that $\mathcal{A}(G)$ is also closed under Hadamard products. We will also need the slightly less obvious fact that $\mathcal{A}(G)$ is closed under taking principal square roots of its positive semidefinite elements, which we now prove.

Lemma 3.9. *Let $G = (V, E, J)$ be a connected, transitive weighted graph. If $T \in \mathcal{A}(G)$ is positive semidefinite then its principal square root $T^{1/2}$ also belongs to $\mathcal{A}(G)$.*

Proof of Lemma 3.9. It suffices to prove that $T^{1/2}$ is diagonally invariant, i.e., that $T^{1/2}(\gamma u, \gamma v) = T^{1/2}(u, v)$ for every $\gamma \in \mathrm{Aut}(G)$ and $u, v \in V$. Fix $\gamma \in \mathrm{Aut}(G)$ and let $S \in \mathcal{M}(V)$ be defined by $S(u, v) = T^{1/2}(\gamma u, \gamma v)$ for each $u, v \in V$. Since $\gamma \in \mathrm{Aut}(G)$ was arbitrary, it suffices to prove that S is equal to $T^{1/2}$. Since the principal square root of T is unique, it suffices to prove that S is positive semidefinite and satisfies $S^2 = T$. We begin by proving that S is bounded and positive semidefinite, noting that is is clearly symmetric. For each $f \in L^0(V)$ let $\gamma f \in L^0(V)$ be given by $\gamma f(u) = f(\gamma^{-1}u)$. Since γ is a bijection we have that that

$$Sf(u) = \sum_{v \in V} S(u, v)f(v) = \sum_{v \in V} T^{1/2}(\gamma u, \gamma v)f(v) = \sum_{v \in V} T^{1/2}(\gamma u, v)f(\gamma^{-1}v) = [T^{1/2}(\gamma f)](\gamma u)$$

for every $f \in L^0(V)$ and $u \in V$ and hence that

$$\|Sf\| = \|T^{1/2}(\gamma f)\| \leq \|T\|_{2\to 2}^2 \|\gamma f\| = \|T\|_{2\to 2}^2 \|f\|$$

and

$$\langle Sf, f\rangle = \sum_{u\in V}\sum_{v\in V} T^{1/2}(\gamma u, v) f(u) f(\gamma^{-1} v) = \sum_{u\in V}\sum_{v\in V} T^{1/2}(u, v) f(\gamma^{-1} u) f(\gamma^{-1} v) = \langle T^{1/2}(\gamma f), \gamma f\rangle$$

for every $f \in L^0(V)$. Since $T^{1/2}$ is bounded and positive semidefinite it follows that S is bounded and positive semidefinite also. We now verify that $S^2 = T$ by computing that

$$\begin{aligned} S^2(u, v) &= \sum_{w\in V} S(u, w) S(w, v) = \sum_{w\in V} T^{1/2}(\gamma u, \gamma w) T^{1/2}(\gamma w, \gamma v) \\ &= \sum_{w\in V} T^{1/2}(\gamma u, w) T^{1/2}(w, \gamma v) = T(\gamma u, \gamma v) = T(u, v) \end{aligned}$$

for every $u, v \in V$. We deduce that $S = T^{1/2}$ by uniqueness of the principal square root of T. □

We now define the *trace* on $\mathcal{A}(G)$ and establish its basic properties.

Proposition 3.10 (The trace on the von Neumann algebra $\mathcal{A}(G)$)**.** *Let $G = (V, E, J)$ be a* unimodular*, connected, transitive weighted graph. Define* $\mathrm{Tr} : \mathcal{A}(G) \to \mathbb{R}$ *by*

$$\mathrm{Tr}(T) = T(o, o) \qquad \textit{for every } T \in \mathcal{A}(G).$$

Then Tr *is a* ***tracial state*** *on $\mathcal{A}(G)$. This means that the following conditions hold:*

1. $\mathrm{Tr} : \mathcal{A}(G) \to \mathbb{R}$ *is linear.*
2. $\mathrm{Tr}(I) = 1$*, where $I(u, v) = \mathbb{1}(u = v)$ is the identity matrix.*
3. $\mathrm{Tr}(TT^{\perp}) \geq 0$ *for every $T \in \mathcal{A}(G)$, with equality if and only if $T = 0$.*
4. $\mathrm{Tr}(ST) = \mathrm{Tr}(TS)$ *for every $S, T \in \mathcal{A}(G)$.*

See e.g. [5, §5] and [33, Chapter 10.8] for previous uses of similar notions in probabilistic contexts.

The proof of this proposition will use the *signed* mass-transport principle, which states that if $G = (V, E, J)$ is a unimodular transitive weighted graph and $F : V \times V \to \mathbb{R}$ is diagonally invariant in the sense that $F(\gamma u, \gamma v) = F(u, v)$ for every $u, v \in V$ and $\gamma \in \mathrm{Aut}(G)$ then $\sum_{v\in V} F(o, v) = \sum_{v\in V} F(v, o)$ provided that F satisfies the integrability condition

$$\sum_{v\in V} |F(o, v)| = \sum_{v\in V} |F(v, o)| < \infty. \tag{3.5}$$

This can be verified by applying the usual mass-transport principle (1.8) to the positive and negative parts of F separately.

Proof of Proposition 3.10. The first two items are trivial. For the third, simply note that

$$\mathrm{Tr}(TT^{\perp}) = \sum_{v\in V} T(o, v) T^{\perp}(v, o) = \sum_{v\in V} T(o, v)^2$$

for every $T \in \mathcal{A}(G)$, from which the claim follows trivially. Finally, for the fourth item, we apply the mass-transport principle to the diagonally invariant function $F(u,v) = S(u,v)T(v,u)$, which satisfies the integrability hypothesis (3.5) since $S, T \in \mathcal{M}^2(V)$, to deduce that

$$\mathrm{Tr}(ST) = \sum_{v \in V} S(o,v)T(v,o) = \sum_{v \in V} S(v,o)T(o,v) = \mathrm{Tr}(TS).$$

as claimed. □

The following is an adaptation to our setting of an inequality due to Fejér [14, Theorem A.8].

Lemma 3.11. *Let $G = (V, E, J)$ be a unimodular, connected, transitive weighted graph. If $S, T \in \mathcal{A}(G)$ are positive semidefinite then $\mathrm{Tr}(ST) \geq 0$.*

It follows in particular from this theorem that if two matrices $S, T \in \mathcal{A}(G)$ satisfy $S \leq T$ then $\mathrm{Tr}(SX) \leq \mathrm{Tr}(TX)$ for every positive semidefinite matrix $X \in \mathcal{A}(G)$.

Proof of Lemma 3.11. We have by Item 4. of Proposition 3.10 that

$$\mathrm{Tr}(ST) = \mathrm{Tr}(S^{1/2}S^{1/2}T^{1/2}T^{1/2}) = \mathrm{Tr}(S^{1/2}T^{1/2}T^{1/2}S^{1/2}).$$

Since $(S^{1/2}T^{1/2})^{\perp} = (T^{1/2})^{\perp}(S^{1/2})^{\perp} = T^{1/2}S^{1/2}$, it follows from item 3 of Proposition 3.10 that $\mathrm{Tr}(ST) \geq 0$ as claimed. □

We are finally ready to prove that $B_\beta \leq A_\beta$ and conclude the proof of Proposition 1.7.

Proof of Proposition 1.7. Let $G = (V, E, J)$ be a unimodular, connected, transitive weighted graph and let $0 \leq \beta < \beta_c$. We have by Lemma 3.3 that $A_\beta \leq \nabla_\beta^2$, so that it suffices to prove that $B_\beta \leq A_\beta$. The two-point matrix T_β is positive definite by Lemma 3.1 and belongs to $\mathcal{M}^2(V)$ and hence to $\mathcal{A}(G)$ by sharpness of the phase transition as discussed above. It follows from Proposition 3.4 that $T_\beta^2 \circ T_\beta^2 \leq T_\beta \circ T_\beta^3$ and hence by Lemma 3.11 that

$$B_\beta = [T_\beta^2 \circ T_\beta^2]T_\beta^2(o,o) = \mathrm{Tr}\Big([T_\beta^2 \circ T_\beta^2]T_\beta^2\Big) \leq \mathrm{Tr}\Big([T_\beta \circ T_\beta^3]T_\beta^2\Big) = [T_\beta \circ T_\beta^3]T_\beta^2(o,o) = A_\beta$$

as claimed. Together with Proposition 1.6 this also concludes the proof of Theorem 1.1. □

Now that we have completed the proof of Theorem 1.1, let us note how it implies Corollary 1.2.

Proof of Corollary 1.2. Let $0 \leq \beta < \beta_c$. Since $\mathcal{X}_{\beta_c} = \infty$ we have that

$$\frac{1}{\mathcal{X}_\beta} = \frac{1}{\mathcal{X}_\beta} - \frac{1}{\mathcal{X}_{\beta_c}} = -\int_\beta^{\beta_c} \frac{d}{d\lambda}\frac{1}{\mathcal{X}_\lambda}\,\mathrm{d}\lambda = \int_\beta^{\beta_c} \frac{1}{\mathcal{X}_\lambda^2}\frac{d\mathcal{X}_\lambda}{d\lambda}\,\mathrm{d}\lambda.$$

It follows from Theorem 1.1 that

$$\frac{1}{\mathcal{X}_\beta} \geq \int_\beta^{\beta_c} \frac{1}{3\lambda^2\nabla_\lambda^2}\,\mathrm{d}\lambda - \int_\beta^{\beta_c} \frac{1}{3\lambda^2\nabla_\lambda\mathcal{X}_\lambda}\,\mathrm{d}\lambda \geq \int_\beta^{\beta_c} \frac{1}{3\lambda^2\nabla_\lambda^2}\,\mathrm{d}\lambda - \frac{1}{\mathcal{X}_\beta}\int_\beta^{\beta_c} \frac{1}{3\lambda^2}\,\mathrm{d}\lambda,$$

where we used that $\mathcal{X}_\lambda$ is increasing in λ and that $\nabla_\lambda \geq 1$ for every $\lambda \geq 0$ in the second inequality. The claim follows by rearranging this inequality. □

4 The tail of the volume and the infinite cluster density

In this section we show that mean-field behaviour of the susceptibility implies mean-field behaviour of the critical volume distribution and the infinite cluster density, then use the resulting estimates together with Theorem 1.1 to prove Corollary 1.5.

We begin by stating and proving a generalization to weighted graphs of an inequality established in the very recent work of Dewan and Muirhead [15, Proposition 2.10], from which we will deduce Theorem 1.3 as a corollary. The statement of this inequality will involve the notion of *decision trees*, which have recently come to play an important role in mathematical physics following the breakthrough work of Duminil-Copin, Raoufi, and Tassion [16].

Let E be a countable set and let $E^* = \bigcup_{n\in\mathbb{N}\cup\{\infty\}} E^n$ be the set of finite or infinite sequences in E, which is equipped with the product topology and associated Borel σ-algebra. A **decision tree** is a function $T : \{0,1\}^E \to E^*$ such that $T_1(\omega) = e_1$ for some fixed $e_1 \in E$ and for each $n \geq 2$ there exists a function $S_n : (E \times \{0,1\})^{n-1} \to E \cup \varnothing$ such that either $S_n\left[\left(T_i, \omega(T_i)\right)_{i=1}^{n-1}\right] = \varnothing$ in which case $T(\omega) = (T_1(\omega), \ldots, T_{n-1}(\omega))$ (i.e., the decision tree halts) or else

$$T_n(\omega) = S_n\left[\left(T_i, \omega(T_i)\right)_{i=1}^{n-1}\right].$$

That is, T is a deterministic procedure for querying the values of the configuration $\omega \in \{0,1\}^E$ that starts by querying the value of some fixed edge e_1 and at each subsequent step chooses either to halt or to query the value of some other edge as a function of the values it has already observed.

Given a decision tree T and $\omega \in \{0,1\}^E$, we write $\tau(\omega)$ for the length of the sequence $T(\omega)$. Given a Borel subset A of $\{0,1\}^E$, we say that a decision tree T **Borel-computes** A if there exists a Borel subset $\mathscr{A}$ of $\{0,1\}^*$ such that $\omega \in A$ if and only if $(T_i(A))_{i=1}^{\tau(\omega)} \in \mathscr{A}$. Given a decision tree T, a measure μ on $\{0,1\}^E$, and $e \in E$, we define the **revealment probability**

$$\mathrm{Rev}(\mu, T, e) = \mu(\{\omega : \exists n \geq 1 \text{ such that } T_n(\omega) = e\}).$$

Given a countable weighted graph $G = (V, E, J)$, $\beta \geq 0$, and a decision tree $T : \{0,1\}^E \to E^*$, we write $\mathrm{Rev}_\beta(T, e) = \mathrm{Rev}(\mathbb{P}_\beta, T, e)$.

Theorem 4.1 (Generalised Dewan-Muirhead)**.** *Let $G = (V, E, J)$ be a countable weighted graph, let $A \subseteq \{0,1\}^E$ be a Borel set, and let T be a decision tree that Borel-computes the event A. Then*

$$|\mathbb{P}_{\beta_1}(A) - \mathbb{P}_{\beta_2}(A)|^2 \leq \frac{|\beta_1 - \beta_2|^2}{\min\{\beta_1, \beta_2\}} \max\{\mathbb{P}_{\beta_1}(A), \mathbb{P}_{\beta_2}(A)\} \sum_{e\in E} J_e \,\mathrm{Rev}_{\beta_1}(T, e) \tag{4.1}$$

for every $\beta_1, \beta_2 \geq 0$.

This inequality plays an interesting complementary role to the *OSSS inequality* of O'Donnel, Saks, Schramm, and Servedio [39], which roughly states that if A is *increasing* then the existence of a decision tree computing A with low *maximum* revealment implies that the logarithmic derivative of $\mathbb{P}_\beta(A)$ is *large*.

Remark 4.2. The original Dewan-Muirhead inequality required A to depend on at most finitely many edges. This assumption is not appropriate in the long-range case, and we have introduced the notion

of a decision tree Borel-computing an event to circumvent this issue. In other parts of the literature, one often defines a decision tree T to compute an event A if A belongs to the *completion* of the σ-algebra generated by T; considering infinite-volume cases of Theorem 4.1 in which $\mathbb{P}_{\beta_1}(A) = 1$ and $\mathbb{P}_{\beta_2}(A) = 0$ – in which case A is trivially in the completion of the σ-algebra generated by *any* decision tree – shows that this notion of computation cannot be used in the statement of Theorem 4.1.

The proof of Theorem 4.1 will rely on the notion of *relative entropy*. Given two probably measures μ and ν on a common measurable space Ω, recall that the **relative entropy** (a.k.a. **Kullback-Liebler divergence**) from ν to μ is defined to be

$$D_{\mathrm{KL}}(\mu||\nu) = \int \log\left(\frac{d\mu}{d\nu}\right) d\mu(x)$$

if μ is absolutely continuous with respect to ν, and is defined to be infinite otherwise. Given random variables X and Y taking values in the same space (but not necessarily defined on the same probability space), we write $D_{\mathrm{KL}}(X||Y)$ for the relative entropy from the law of Y to the law of X. The relevance of this quantity to Theorem 4.1 arises from a generalization of Pinsker's inequality established in [15, Lemma 2.12], which states that if μ and ν are two probability measures on a common measurable space Ω then

$$|\mu(A) - \nu(A)|^2 \leq 2D_{\mathrm{KL}}(\mu||\nu)\max\{\mu(A), \nu(A)\} \tag{4.2}$$

for every event $A \subseteq \Omega$. Indeed, the expression $\frac{|\beta_1-\beta_2|^2}{2\min\{\beta_1,\beta_2\}}\sum_{e\in E} J_e \operatorname{Rev}_{\beta_1}(T,e)$ appearing in Theorem 4.1 will arise as an upper bound on the relative entropy of two appropriately defined random variables. We will also use the **chain rule** for the relative entropy, which states that if (X_1, X_2) and (Y_1, Y_2) are random variables taking values in the same product space $\Omega_1 \times \Omega_2$ then

$$D_{\mathrm{KL}}\Big((X_1, X_2)||(Y_1, Y_2)\Big) = D_{\mathrm{KL}}(X_1||Y_1) + \mathbb{E}_{x\sim X_1}\left[D_{\mathrm{KL}}\Big((X_2|X_1 = x)\ \big|\big|\ (Y_2|Y_1 = x)\Big)\right], \tag{4.3}$$

where we write $(X_2|X_1 = x)$ for the conditional distribution of X_2 given $X_1 = x$ and write $\mathbb{E}_{x\sim X_1}$ for an expectation taken over a random x with the law of X_1.

We now begin to work towards the proof of Theorem 4.1. We will use the following estimate on the relative entropy of Bernoulli random variables, which we express in a form that is convenient for our applications to percolation on weighted graphs.

Lemma 4.3. *For each $p \in [0, 1]$ let $\mathrm{Ber}(p)$ denote the law of a Bernoulli-p random variable. The estimate*

$$D_{\mathrm{KL}}\Big(\mathrm{Ber}\big(1 - e^{-a}\big)\Big|\Big|\mathrm{Ber}\big(1 - e^{-b}\big)\Big) = D_{\mathrm{KL}}\Big(\mathrm{Ber}\big(e^{-a}\big)\Big|\Big|\mathrm{Ber}\big(e^{-b}\big)\Big) \leq \frac{|a-b|^2}{2\min\{a,b\}} \tag{4.4}$$

holds for every $a, b > 0$.

Proof of Lemma 4.3. We have by definition that

$$D_{\mathrm{KL}}\Big(\mathrm{Ber}(p)\Big|\Big|\mathrm{Ber}(q)\Big) = p\log\frac{p}{q} + (1-p)\log\frac{1-p}{1-q} = \int_q^p \frac{p-s}{s(1-s)}\,\mathrm{d}s$$

for every $0 < p, q < 1$. First assume that $a < b$. Taking $p = e^{-a}$ and $q = e^{-b}$ and using the

substitution $s = e^{-t}$ yields that

$$D_{\mathrm{KL}}\Big(\mathrm{Ber}\big(e^{-a}\big)\Big|\Big|\,\mathrm{Ber}\big(e^{-b}\big)\Big) = \int_a^b \frac{e^{-a}-e^{-t}}{1-e^{-t}}\,\mathrm{d}t \leq \frac{e^{-a}}{1-e^{-a}}\int_0^{b-a}(1-e^{-t})\,\mathrm{d}t.$$

Using that $1-e^{-t} \leq t$ and $e^{-t}/(1-e^{-t}) \leq t$ for $t \geq 0$, we deduce that

$$D_{\mathrm{KL}}\Big(\mathrm{Ber}\big(e^{-a}\big)\Big|\Big|\,\mathrm{Ber}\big(e^{-b}\big)\Big) \leq \frac{e^{-a}(b-a)^2}{2(1-e^{-a})} \leq \frac{|a-b|^2}{2\min\{a,b\}}$$

as claimed. Now assume that $b < a$. Taking $p = e^{-a}$ and $q = e^{-b}$ and using the substitution $s = e^{-t}$ yields that

$$D_{\mathrm{KL}}\Big(\mathrm{Ber}\big(e^{-a}\big)\Big|\Big|\,\mathrm{Ber}\big(e^{-b}\big)\Big) = \int_b^a \frac{e^{-t}-e^{-a}}{1-e^{-t}}\,\mathrm{d}t \leq \frac{e^{-a}}{1-e^{-b}}\int_0^{a-b}(e^{t}-1)\,\mathrm{d}t.$$

Now, using the inequality $e^{-x}(e^x-1) \leq x$ we deduce that $e^{-a}(e^t-1) \leq e^{-b}t$ for every $0 \leq t \leq a-b$ and hence that

$$D_{\mathrm{KL}}\Big(\mathrm{Ber}\big(e^{-a}\big)\Big|\Big|\,\mathrm{Ber}\big(e^{-b}\big)\Big) \leq \frac{e^{-b}}{1-e^{-b}}\int_0^{a-b} t\,\mathrm{d}t \leq \frac{e^{-b}(b-a)^2}{2(1-e^{-b})} \leq \frac{|a-b|^2}{2\min\{a,b\}}$$

as before. □

We are now ready to prove Theorem 4.1.

Proof of Theorem 4.1. For notational convenience, for each $\omega \in \{0,1\}^E$ we will set $T_i(\omega) = \varnothing$ for $i \geq \tau(\omega)$ so that $T_i(\omega)$ is well-defined for every $\omega \in \{0,1\}^E$ and $i \geq 1$. Let $\omega_1, \omega_2 \in \{0,1\}^E$ be samples of Bernoulli percolation on G with parameters β_1 and β_2 respectively, and let $T_1^1 = T_1(\omega_1), T_2^1 = T_2(\omega_1), \ldots, T_{\tau_1}^1 = T_{\tau(\omega_1)}(\omega_1)$ and $T_2^2 = T_2(\omega_2), T_2^1 = T_2(\omega_2), \ldots, T_{\tau_2}^2 = T_{\tau(\omega_2)}(\omega_2)$ be the edges of G that are revealed when applying the decision tree T to ω_1 and ω_2 respectively. For each $i \geq 1$, let $X_i = \omega_1(T_i^1)$ if $i \leq \tau_1$ and $X_i = \varnothing$ otherwise, so that $X = (X_1, X_2, \ldots)$ is an infinite $\{0,1,\varnothing\}$-valued sequence, and similarly define $Y_i = \omega_2(T_i^2)$ if if $i \leq \tau_2$ and $Y_i = \varnothing$ otherwise. For each $i \geq 1$, let $X^i = (X_1, \ldots, X_i)$ and $Y^i = (Y_1, \ldots, Y_i)$. For each $k \geq 1$, we have by the chain rule that

$$D_{\mathrm{KL}}\big(X^{k+1}||Y^{k+1}\big) = D_{\mathrm{KL}}\big(X^k||Y^k\big) + \mathbb{E}_{x\sim X^k}\Big[D_{\mathrm{KL}}\Big(\big(X_{k+1} \mid X^k = x\big)\Big|\Big|\big(Y_{k+1} \mid Y^k = x\big)\Big)\Big]. \quad (4.5)$$

Since T is a decision tree, for each $x \in \{0,1,\varnothing\}^k$ we either have that $X_{k+1} = \varnothing$ and $Y_{k+1} = \varnothing$ on the events $X^k = x$ and $Y^k = x$ or else that there exists an edge $T_{k+1}(x)$ such that $X_{k+1} = \omega_1(T_{k+1}(x))$ and $Y_{k+1} = \omega_2(T_{k+1}(x))$ on the events $X^k = x$ and $Y^k = x$. Letting $J_{k+1}(x)$ be the weight of the edge $T_{k+1}(x)$, it follows that

$$\begin{aligned} &D_{\mathrm{KL}}\Big(\big(X_{k+1} \mid X^k = x\big)\Big|\Big|\big(Y_{k+1} \mid Y^k = x\big)\Big) \\ &\qquad = \mathbb{1}(T_{k+1}(x) \neq \varnothing) D_{\mathrm{KL}}\Big(\mathrm{Ber}\big(1-e^{-\beta_1 J_{k+1}(x)}\big)\Big|\Big|\,\mathrm{Ber}\big(1-e^{-\beta_2 J_{k+1}(x)}\big)\Big) \\ &\qquad \leq \mathbb{1}(T_{k+1}(x) \neq \varnothing)\frac{|\beta_1-\beta_2|^2}{2\min\{\beta_1,\beta_2\}} J_{k+1}(x). \end{aligned}$$

for every $x \in \{0,1,\varnothing\}^k$ such that $X^k = x$ with positive probability. Taking expectations and letting J_i be the weight of the edge T_i^1 for each $1 \le i \le \tau$, it follows by induction on k that

$$D_{\mathrm{KL}}(X^k||Y^k) \le \frac{|\beta_1 - \beta_2|}{2\min\{\beta_1,\beta_2\}} \mathbb{E}_{\beta_1} \sum_{i=1}^{k\wedge\tau} J_i \tag{4.6}$$

for every $k \ge 1$. Taking the limit as $k \uparrow \infty$ (which is valid since the relative entropy $D_{\mathrm{KL}}(\mu||\nu)$ is lower semicontinuous in (μ,ν) with respect to the weak topology on probability measures [40, Theorem 1]), we deduce that

$$D_{\mathrm{KL}}(X||Y) \le \frac{|\beta_1 - \beta_2|}{2\min\{\beta_1,\beta_2\}} \mathbb{E}_{\beta_1} \sum_{i=1}^{\tau} J_i = \frac{|\beta_1 - \beta_2|^2}{2\min\{\beta_1,\beta_2\}} \sum_{e\in E} J_e \cdot \mathrm{Rev}_{\beta_1}(T,e). \tag{4.7}$$

Now, since T Borel-computes the event A, there exists a Borel-measurable subset $\mathscr{A}$ of $\{0,1,\varnothing\}^{\mathbb{N}}$ such that ω_1 belongs to A if and only if X belongs to $\mathscr{A}$ and ω_2 belongs to A if and only if Y belongs to $\mathscr{A}$. The claim now follows by applying the generalised Pinsker inequality (4.2) to the laws of X and Y and the event $\mathscr{A}$. □

Proof of Theorem 1.3. Fix $0 < \beta_1 < \beta_2$ and $n \ge 0$. We apply Theorem 4.1 taking $A = \{|K| \ge n\}$ and taking T to be a decision tree that explores the cluster of the origin one edge at a time, stopping if and when it first finds n vertices in K; see e.g. [26, Page 14] for a formal definition. This decision tree can only query edges with at least one endpoint in K, and the set of edges queried in T all have endpoints in some set of size at most n. Using the assumption that $\sum_{e\in E_o^{\to}} J_e = 1$, it follows that

$$\sum_{e\in E} J_e \cdot \mathrm{Rev}_{\beta_1}(T,e) \le \mathbb{E}_{\beta_1}\left[|K| \wedge n\right] = \sum_{k=1}^{n} \mathbb{P}_{\beta_1}(|K| \ge k)$$

and hence by Theorem 4.1 that

$$|\mathbb{P}_{\beta_2}(|K| \ge n) - \mathbb{P}_{\beta_1}(|K| \ge n)|^2 \le \frac{1}{\beta_1}|\beta_2 - \beta_1|^2 \mathbb{P}_{\beta_2}(|K| \ge n) \sum_{k=1}^{n} \mathbb{P}_{\beta_1}(|K| \ge k).$$

It follows immediately that at least one of the estimates

$$\mathbb{P}_{\beta_2}(|K| \ge n) \le 2\mathbb{P}_{\beta_1}(|K| \ge n) \qquad \text{or} \qquad \mathbb{P}_{\beta_2}(|K| \ge n) \le \frac{4}{\beta_1}|\beta_2 - \beta_1|^2 \sum_{k=1}^{n} \mathbb{P}_{\beta_1}(|K| \ge k)$$

must hold, implying the first inequality. The second inequality follows by Markov's inequality. □

We deduce Corollary 1.5 from Theorem 1.1 and Theorem 1.3.

Proof of Corollary 1.5. We may assume without loss of generality that $\sum_{e\in E_o^{\to}} J_e = 1$. Let $\alpha \ge 0$, $C < \infty$, and δ be such that $\nabla_{\beta-\varepsilon} \le C(\log(1/\varepsilon))^\alpha$ for every $0 \le \varepsilon \le \delta$. We will write $\asymp$, $\preceq$, and $\succeq$ for equalities and inequalities holding to within positive multiplicative constants depending only on

α, C, and δ. We have by Corollary 1.2 that

$$\mathcal{X}_\beta \preceq \left[\int_\beta^{\beta_c} \left(\log \frac{1}{\beta_c - \lambda}\right)^{-2\alpha} \mathrm{d}\lambda\right]^{-1} \preceq |\beta - \beta_c|^{-1} \left(\log \frac{1}{|\beta - \beta_c|}\right)^{2\alpha} \tag{4.8}$$

for every $\beta_c - \delta \leq \beta < \beta_c$. Substituting (4.8) into Theorem 1.3 with $\beta_1 = \beta_c - n^{-1/2}$ and $\beta_2 = \beta_c$ yields that

$$\mathbb{P}_{\beta_c}(|K| \geq n) \preceq n^{-1} \mathcal{X}_{\beta_c - n^{-1/2}} \preceq n^{-1/2} (\log n)^{2\alpha}$$

for every $n \geq 1$. Similarly, substituting (4.8) into Theorem 1.3 with $\beta_1 = \beta_c - \varepsilon$ and $\beta_2 = \beta_c + \varepsilon$ yields that

$$\mathbb{P}_{\beta_c+\varepsilon}(|K| \geq n) \preceq \left(n^{-1} + \varepsilon^2\right) \varepsilon^{-1} \left(\log \frac{1}{\varepsilon}\right)^{2\alpha}$$

for every sufficiently small $\varepsilon > 0$ and every $n \geq 0$, so that the claimed inequality (1.12) follows by taking $n \uparrow \infty$. This completes the proof. □

5 Applications to the hierarchical lattice

In this section we prove our results concerning long-range percolation on the hierarchical lattice. In addition to Theorem 1.8, we will also prove the following related theorem showing that the critical exponents describing the model are close to their mean-field values when α is close to $d/3$.

Theorem 5.1. *Let $J : \mathbb{H}_L^d \to [0, \infty)$ be a radially symmetric, integrable function, let $0 < \alpha < d$, and suppose that there exist constants c and C such that $c\|x\|^{-d-\alpha} \leq J(x) \leq C\|x\|^{-d-\alpha}$ for every $x \in \mathbb{H}_L^d \setminus \{0\}$. If $d/3 < \alpha < 4d/11$ then*

$$\mathcal{X}_{\beta_c - \varepsilon} \preceq \varepsilon^{-\alpha/(2d-5\alpha)} \qquad \text{as } \varepsilon \downarrow 0, \tag{5.1}$$

$$\mathbb{P}_{\beta_c}\left(|K| \geq n\right) \preceq n^{-(4d-11\alpha)/(4d-10\alpha)} \qquad \text{as } n \uparrow \infty, \text{ and} \tag{5.2}$$

$$\mathbb{P}_{\beta_c+\varepsilon}\left(|K_v| = \infty\right) \preceq \varepsilon^{(4d-11\alpha)/(2d-5\alpha)} \qquad \text{as } \varepsilon \downarrow 0. \tag{5.3}$$

The requirement that $\alpha < 4d/11$ is used only to ensure that the resulting bounds on $\mathbb{P}_{\beta_c}(|K| \geq n)$ and $\mathbb{P}_{\beta_c+\varepsilon}(|K| = \infty)$ are non-trivial; we expect this threshold at $4d/11$ to be a feature of the proof that does not correspond to any true change in behaviour of the model. Indeed, power-law upper bounds on the same quantities have been proven for all $0 < \alpha < d$ via completely different methods in our recent work [29]. (That paper does *not* establish that these exponents approach their mean-field values as $\alpha \downarrow d/3$, or indeed that mean-field behaviour holds for $\alpha < d/3$.)

For the remainder of the section we will fix $d \geq 1$, $L \geq 2$, $0 < \alpha < d$ and a radially symmetric function $J : \mathbb{H}_L^d \to [0, \infty)$ satisfying $c\langle x\rangle^{-d-\alpha} \leq J(x) \leq C\langle x\rangle^{-d-\alpha}$ for some positive constants c and C and every $x \in \mathbb{H}_L^d \setminus \{0\}$. We will assume without loss of generality that $\sum_{x \in \mathbb{H}_L^d} J(x) = 1$ and write $\asymp$, $\preceq$, and $\succeq$ for equalities and inequalities holding to within positive multiplicative constants depending only on d, L, α, c and C.

Recall that for each $n \geq 1$ we write $\Lambda_n = \{x \in \mathbb{H}^d : \langle x\rangle \leq L^n\}$ for the ultrametric ball of radius L^n containing the origin, noting that the restriction to Λ_n of the weighted graph defined in terms of $\mathbb{H}_L^d$ and J is itself transitive. Let K_n denote the cluster of the origin in Λ_n, i.e., the set of vertices

in Λ_n that are connected to the origin by an open path contained in Λ_n. For each $\beta \geq 0$ and $n \geq 0$ we let $T_{n,\beta}$ denote the two-point matrix on Λ_n and define $\mathcal{X}_{n,\beta} := \mathbb{E}_\beta |K_n| = \sum_{x \in \Lambda_n} T_{n,\beta}(0,x)$ and $\nabla_{n,\beta} := \sum_{x,y \in \Lambda_n} T_{n,\beta}(0,x) T_{n,\beta}(x,y) T_{n,\beta}(y,0)$. It is proven in [28, Corollary 1.2] that

$$\mathcal{X}_{n,\beta_c} \asymp L^{\alpha n} \tag{5.4}$$

for every $n \geq 0$, while Theorem 1.1 yields that

$$\frac{d\mathcal{X}_{n,\beta}}{d\beta} \geq \frac{\mathcal{X}_{n,\beta}(\mathcal{X}_{n,\beta} - \nabla_{n,\beta})}{3\beta^2 \nabla_{n,\beta}^2} \tag{5.5}$$

for every $\beta \geq 0$ and $n \geq 0$.

We will apply the differential inequality (5.5) to study the growth of the *correlation length* $\xi(\beta)$ as $\beta \uparrow \beta_c$. Before defining the correlation length, we first develop some general theory that will motivate the definition. Let $G = (V, E, J)$ be a countable weighted graph. Following Duminil-Copin and Tassion [17], for each $v \in V$, $\beta \geq 0$, and finite subset $S \subseteq V$ we consider the quantity

$$\phi_\beta(S, v) := \sum_{e \in \partial^{\to} S} \Big(1 - e^{-\beta J_e}\Big) \mathbb{P}_\beta(v \overset{S}{\leftrightarrow} e^-),$$

where we write $\{x \overset{S}{\leftrightarrow} y\}$ to mean that x and y are connected by an open path all of whose vertices belong to S and recall that $\partial^{\to} S$ denotes the set of oriented edges with $e^- \in S$ and $e^+ \notin S$. When G is infinite and transitive, Duminil-Copin and Tassion [17] proved that the critical parameter β_c admits the alternative characterisation

$$\beta_c = \inf\Big\{\beta \geq 0 : \phi_\beta(S, o) \geq 1 \text{ for every finite } S \ni o\Big\}, \tag{5.6}$$

where o is an arbitrary fixed root vertex. This characterisation is closely related to the following inequality, which slightly sharpens the analysis of [17, Theorem 1.1, Item 2].

Lemma 5.2. *Let $G = (V, E, J)$ be a countable weighted graph and let $S \subseteq \Lambda$ be finite subsets of V. Then*

$$\sum_{x \in A} \mathbb{P}_\beta(v \overset{\Lambda}{\leftrightarrow} x) \leq \sum_{x \in S} \mathbb{P}_\beta(v \overset{S}{\leftrightarrow} x) + \phi_\beta(S, v) \cdot \sup_{u \in A} \sum_{x \in A} \mathbb{P}_\beta(u \overset{\Lambda}{\leftrightarrow} x)$$

for every $\beta \geq 0$ and $v \in S$.

Proof of Lemma 5.2. Fix $\beta \geq 0$ and $v \in S$. For each $x \in \Lambda$, considering the edge crossed by an open path from v to x in Λ as it leaves S for the first time yields that

$$\{v \overset{\Lambda}{\leftrightarrow} x\} \setminus \{v \overset{S}{\leftrightarrow} x\} \subseteq \bigcup_{e \in \partial^{\to} S} \{v \overset{S}{\leftrightarrow} e^-\} \circ \{e \text{ open}\} \circ \{e^+ \overset{\Lambda}{\leftrightarrow} x\}.$$

Thus, it follows by the BK inequality and a union bound that

$$\mathbb{P}_\beta(v \overset{\Lambda}{\leftrightarrow} x) \leq \mathbb{P}(v \overset{S}{\leftrightarrow} x) + \sum_{e \in \partial^{\to} S} \mathbb{P}_\beta(v \overset{S}{\leftrightarrow} e^-)(1 - e^{-\beta J_e}) \mathbb{P}_\beta(e^+ \overset{\Lambda}{\leftrightarrow} x)$$

for every $x \in \Lambda$. The claim follows by summing over $x \in \Lambda$. $\square$

We now return to our setting of the hierarchical lattice as above. For each $n \geq 0$ we define

$$\beta_n = \sup\left\{\beta \geq 0 : \phi_\beta(\Lambda_n, 0) \leq \frac{1}{2}\right\},$$

which satisfies $\beta_n \leq \beta_c$ since $\phi_{\beta_c}(\Lambda_n, 0) \geq 1$ for every $n \geq 1$ by (5.6). It follows from Lemma 5.2 and transitivity of Λ_n that

$$\mathcal{X}_{\beta,N} \leq \mathcal{X}_{\beta,n} + \phi_\beta(\Lambda_n, 0) \cdot \mathcal{X}_{\beta,N}$$

for each $N \geq n \geq 0$ and $\beta \geq 0$, and hence that

$$\mathcal{X}_\beta = \lim_{N\to\infty} \mathcal{X}_{\beta,N} \leq 2\mathcal{X}_{\beta,n} \leq 2\mathcal{X}_{\beta_c,n} \tag{5.7}$$

for every $n \geq 0$ and $0 \leq \beta \leq \beta_n$. We claim moreover that

$$\mathcal{X}_{\beta,n} \asymp \mathcal{X}_{\beta_c,n} \asymp L^{\alpha n} \tag{5.8}$$

for every $n \geq 0$ and $\beta_n \leq \beta \leq \beta_c$. Indeed, we have by the definitions that

$$\phi_\beta(\Lambda_n, 0) = \sum_{y \in \mathbb{H}_L^d \setminus \Lambda_n} \left(1 - e^{-\beta J(y)}\right) \cdot \mathbb{E}_\beta|K_n| \asymp \beta L^{-\alpha n} \cdot \mathbb{E}_\beta|K_n|$$

for every $0 \leq \beta \leq \beta_c$ and $n \geq 0$, and the claimed estimate (5.8) follows from (5.4) and the fact that $\phi_\beta(\Lambda_n, 0) \geq 1/2$ for $\beta \geq \beta_n$. Putting together (5.7) and (1.18) yields that

$$\mathcal{X}_\beta \asymp L^{\alpha n} \qquad \text{for every } n \geq 0 \text{ and } \beta_n \leq \beta \leq \beta_{n+1}. \tag{5.9}$$

It follows in particular that $\beta_n \uparrow \beta_c$ as $n \to \infty$.

For each $0 \leq \beta < \beta_c$ define the **correlation length** $\xi(\beta)$ by

$$\xi(\beta) = L^{n(\beta)} \qquad \text{where} \qquad n(\beta) = \inf\{n \geq 0 : \beta \leq \beta_n\}. \tag{5.10}$$

The estimates (5.7) and (5.8) together justify referring to this quantity as the correlation length: Roughly speaking, the estimate (5.8) shows that the two-point function is comparable to the critical two-point function for x with $\langle x \rangle \leq \xi(\beta)$, while (5.7) shows that scales above the correlation length do not contribute significantly to the susceptibility.

Theorems 1.8 and 5.1 will both follow easily from (5.9) together with the following proposition.

Proposition 5.3. *Let $J : \mathbb{H}_L^d \to [0, \infty)$ be a radially symmetric, integrable function, let $0 < \alpha < d$, and suppose that there exist constants c and C such that $c\|x\|^{-d-\alpha} \leq J(x) \leq C\|x\|^{-d-\alpha}$ for every $x \in \mathbb{H}_L^d \setminus \{0\}$. If $\alpha < 2d/5$ then*

$$\xi(\beta) \preceq \begin{cases} |\beta - \beta_c|^{-1/\alpha} & \alpha < d/3 \\ |\beta - \beta_c|^{-1/\alpha} \left(\log \frac{1}{|\beta - \beta_c|}\right)^{2/\alpha} & \alpha = d/3 \\ |\beta - \beta_c|^{-1/(2d-5\alpha)} & \alpha > d/3 \end{cases}$$

for every $0 \leq \beta < \beta_c$.

Remark 5.4. It follows form (5.9) together with the mean-field lower bound $\mathcal{X}_\beta \succeq |\beta - \beta_c|^{-1}$ that the correlation length always satisfies the mean-field lower bound $\xi(\beta) \succeq (\beta_c - \beta)^{-1/\alpha}$ for every $0 \leq \beta < \beta_c$ no matter the value of $0 < \alpha < d$.

Proof of Proposition 5.3. It suffices to prove that there exists a constant n_0 such the estimate

$$|\beta_n - \beta_c| \preceq \begin{cases} L^{-\alpha n} & \alpha < d/3 \\ (n+1)^2 L^{-\alpha n} & \alpha = d/3 \\ L^{-(2d-5\alpha)n} & \alpha > d/3. \end{cases}$$

holds for every $n \geq n_0$. It follows from (5.8) that there exist positive constants c_1 and C_1 such that $c_1 L^{\alpha n} \leq \mathcal{X}_{\beta,n} \leq C_1 L^{\alpha n}$ for every $n \geq 0$ and $\beta_n \leq \beta \leq \beta_c$. On the other hand, it is an immediate consequence of (5.4) as shown in [28, Equation 2.19] that

$$1 \leq \nabla_{\beta,n} \leq \sum_{x,y \in \Lambda_n} T_{\beta_c}(0,x) T_{\beta_c}(x,y) T_{\beta_c}(y,0) \preceq \begin{cases} 1 & \alpha < d/3 \\ n+1 & \alpha = d/3 \\ L^{(3\alpha - d)n} & \alpha > d/3 \end{cases} \tag{5.11}$$

for every $n \geq 0$ and $0 \leq \beta \leq \beta_c$. Since $\alpha < 2d/5 \leq d/2$ we deduce that $\nabla_{\beta_c,n} = o(\mathcal{X}_{\beta_n,n})$ as $n \to \infty$ and hence by (5.5) that there exist positive constants c_2, c_3 and n_0 such that

$$\frac{d\mathcal{X}_{\beta,n}}{d\beta} \geq c_2 \frac{\mathcal{X}_{\beta,n}^2}{\nabla_\beta^2} \geq c_3 \cdot \begin{cases} L^{2\alpha n} & \alpha < d/3 \\ \frac{1}{(n+1)^2} L^{2\alpha n} & \alpha = d/3 \\ L^{(2d-4\alpha)n} & \alpha > d/3. \end{cases}$$

for every $n \geq n_0$ and every $\beta_n \leq \beta \leq \beta_c$. Integrating this inequality between β_n and β_c yields that

$$\mathcal{X}_{\beta_c,n} \geq c_3 |\beta_c - \beta_n| \cdot \begin{cases} L^{2\alpha n} & \alpha < d/3 \\ \frac{1}{(n+1)^2} L^{2\alpha n} & \alpha = d/3 \\ L^{(2d-4\alpha)n} & \alpha > d/3 \end{cases}$$

and the claim follows by rearranging and applying (5.4) to estimate the left hand side. □

Proof of Theorems 1.8 and 5.1. It follows immediately from eq. (5.9) and Proposition 5.3 that

$$\mathcal{X}_\beta \preceq \xi(\beta)^\alpha \preceq \begin{cases} |\beta - \beta_c|^{-1} & \alpha < d/3 \\ |\beta - \beta_c|^{-1} \left(\log \frac{1}{|\beta - \beta_c|} \right)^2 & \alpha = d/3 \\ |\beta - \beta_c|^{-\alpha/(2d-5\alpha)} & \alpha > d/3 \end{cases}$$

for every $0 < \beta < \beta_c$. The remaining claims follow easily from Theorem 1.3 as in the proof of Corollary 1.5. □

Remark 5.5. The tree-graph estimates (2.6) and (2.7) on the three- and four-point functions used to prove (2.5) are not sharp outside of the mean-field regime, and one can improve the exponent estimates of Theorem 5.1 by instead using the techniques of [28] to directly bound the critical four-point function in the hierarchical case. The improvements we were able to obtain via this method were rather modest and we have chosen not to pursue this further here.

Acknowledgements

The author was supported in part by ERC starting grant 804166 (SPRS). We thank Vivek Dewan, Emmanuel Michta, Stephen Muirhead, and Gordon Slade for helpful comments on a previous version of the manuscript.

References

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

[2] M. Aizenman and R. Fernández. On the critical behavior of the magnetization in high-dimensional Ising models. *J. Statist. Phys.*, 44(3-4):393–454, 1986.

[3] M. Aizenman and R. Graham. On the renormalized coupling constant and the susceptibility in $\varphi 44$ field theory and the ising model in four dimensions. *Nuclear Physics B*, 225(2):261–288, 1983.

[4] M. Aizenman and C. M. Newman. Tree graph inequalities and critical behavior in percolation models. *J. Statist. Phys.*, 36(1-2):107–143, 1984.

[5] D. Aldous and R. Lyons. Processes on unimodular random networks. *Electron. J. Probab.*, 12:no. 54, 1454–1508, 2007.

[6] T. Ando. Concavity of certain maps on positive definite matrices and applications to Hadamard products. *Linear Algebra Appl.*, 26:203–241, 1979.

[7] D. J. Barsky and M. Aizenman. Percolation critical exponents under the triangle condition. *Ann. Probab.*, 19(4):1520–1536, 1991.

[8] R. Bauerschmidt, D. C. Brydges, and G. Slade. *Introduction to a renormalisation group method*, volume 2242 of *Lecture Notes in Mathematics*. Springer, Singapore, 2019.

[9] R. Bhatia. *Positive definite matrices*. Princeton Series in Applied Mathematics. Princeton University Press, Princeton, NJ, 2007.

[10] C. Borgs, J. T. Chayes, R. van der Hofstad, G. Slade, and J. Spencer. Random subgraphs of finite graphs. I. The scaling window under the triangle condition. *Random Structures Algorithms*, 27(2):137–184, 2005.

[11] 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. *Ann. Probab.*, 33(5):1886–1944, 2005.

[12] S. Chatterjee and J. Hanson. Restricted percolation critical exponents in high dimensions. *arXiv preprint arXiv:1810.03750*, 2018.

[13] L.-C. Chen and A. Sakai. Critical two-point functions for long-range statistical-mechanical models in high dimensions. *Ann. Probab.*, 43(2):639–681, 2015.

[14] E. de Klerk. *Aspects of semidefinite programming*, volume 65 of *Applied Optimization*. Kluwer Academic Publishers, Dordrecht, 2002. Interior point algorithms and selected applications.

[15] V. Dewan and S. Muirhead. Upper bounds on the one-arm exponent for dependent percolation models. *arXiv preprint arXiv:2102.12123*, 2021.

[16] H. Duminil-Copin, A. Raoufi, and V. Tassion. Sharp phase transition for the random-cluster and Potts models via decision trees. *Ann. of Math. (2)*, 189(1):75–99, 2019.

[17] H. Duminil-Copin and V. Tassion. A new proof of the sharpness of the phase transition for Bernoulli percolation and the Ising model. *Comm. Math. Phys.*, 343(2):725–745, 2016.

[18] I. Essam, D. Gaunt, and A. Guttmann. Percolation theory at the critical dimension. *Journal of Physics A: Mathematical and General*, 11(10):1983, 1978.

[19] R. Fitzner and R. van der Hofstad. Mean-field behavior for nearest-neighbor percolation in $d > 10$. *Electron. J. Probab.*, 22:Paper No. 43, 65, 2017.

[20] G. Grimmett. *Percolation*, volume 321 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 1999.

[21] T. Hara and G. Slade. Mean-field critical behaviour for percolation in high dimensions. *Comm. Math. Phys.*, 128(2):333–391, 1990.

[22] M. Heydenreich and R. van der Hofstad. *Progress in high-dimensional percolation and random graphs*. CRM Short Courses. Springer, Cham; Centre de Recherches Mathématiques, Montreal, QC, 2017.

[23] M. Heydenreich, R. van der Hofstad, and A. Sakai. Mean-field behavior for long- and finite range Ising model, percolation and self-avoiding walk. *J. Stat. Phys.*, 132(6):1001–1049, 2008.

[24] T. Hutchcroft. Percolation on hyperbolic graphs. *Geometric and Functional Analysis*, 29(3):766–810, Jun 2019.

[25] T. Hutchcroft. The L^2 boundedness condition in nonamenable percolation. *Electron. J. Probab.*, 25:Paper No. 127, 27, 2020.

[26] T. Hutchcroft. New critical exponent inequalities for percolation and the random cluster model. *Probability and Mathematical Physics*, 1(1):147–165, 2020.

[27] T. Hutchcroft. Nonuniqueness and mean-field criticality for percolation on nonunimodular transitive graphs. *Journal of the American Mathematical Society*, 33(4):1101–1165, 2020.

[28] T. Hutchcroft. The critical two-point function for long-range percolation on the hierarchical lattice. *arXiv preprint arXiv:2103.17013*, 2021.

[29] T. Hutchcroft. Power-law bounds for critical long-range percolation below the upper-critical dimension. *Probability Theory and Related Fields*, 2021. To appear. Available at http://arxiv.org/abs/2008.11197.

[30] G. Kozma. The triangle and the open triangle. *Ann. Inst. Henri Poincaré Probab. Stat.*, 47(1):75–79, 2011.

[31] G. Kozma and A. Nachmias. The Alexander-Orbach conjecture holds in high dimensions. *Invent. Math.*, 178(3):635–654, 2009.

[32] G. Kozma and A. Nachmias. Arm exponents in high dimensional percolation. *J. Amer. Math. Soc.*, 24(2):375–409, 2011.

[33] R. Lyons and Y. Peres. *Probability on Trees and Networks*, volume 42 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, New York, 2016. Available at http://pages.iu.edu/~rdlyons/.

[34] N. Madras and G. Slade. *The self-avoiding walk*. Modern Birkhäuser Classics. Birkhäuser/Springer, New York, 2013. Reprint of the 1993 original.

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

[36] C. M. Newman. Some critical exponent inequalities for percolation. *J. Statist. Phys.*, 45(3-4):359–368, 1986.

[37] C. M. Newman. Another critical exponent inequality for percolation: $\beta \geq 2/\delta$. In *Proceedings of the symposium on statistical mechanics of phase transitions—mathematical and physical aspects (Trebon, 1986)*, volume 47, pages 695–699, 1987.

[38] B. G. Nguyen. Gap exponents for percolation processes with triangle condition. *J. Statist. Phys.*, 49(1-2):235–243, 1987.

[39] R. O'Donnell, M. Saks, O. Schramm, and R. A. Servedio. Every decision tree has an influential variable. In *46th Annual IEEE Symposium on Foundations of Computer Science (FOCS'05)*, pages 31–39. IEEE, 2005.

[40] E. C. Posner. Random coding strategies for minimum entropy. *IEEE Trans. Inform. Theory*, IT-21:388–391, 1975.

[41] M. Reed and B. Simon. *Methods of modern mathematical physics. I.* Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York, second edition, 1980. Functional analysis.

[42] R. H. Schonmann. Multiplicity of phase transitions and mean-field criticality on highly non-amenable graphs. *Comm. Math. Phys.*, 219(2):271–322, 2001.

[43] G. Slade. *The lace expansion and its applications*, volume 1879 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2006. Lectures from the 34th Summer School on Probability Theory held in Saint-Flour, July 6–24, 2004, Edited and with a foreword by Jean Picard.

[44] P. M. Soardi and W. Woess. Amenability, unimodularity, and the spectral radius of random walks on infinite graphs. *Math. Z.*, 205(3):471–486, 1990.

[45] J. van den Berg and H. Kesten. Inequalities with applications to percolation and reliability. *J. Appl. Probab.*, 22(3):556–569, 1985.