EXPANSION OF PERCOLATION CRITICAL POINTS FOR HAMMING GRAPHS

LORENZO FEDERICO, REMCO VAN DER HOFSTAD, FRANK DEN HOLLANDER, AND TIM HULSHOF

Abstract. The Hamming graph $H(d,n)$ is the Cartesian product of d complete graphs on n vertices. Let $m = d(n-1)$ be the degree and $V = n^d$ be the number of vertices of $H(d,n)$. Let $p_c^{(d)}$ be the critical point for bond percolation on $H(d,n)$. We show that, for $d \in \mathbb{N}$ fixed and $n \to \infty$,

$$p_c^{(d)} = \frac{1}{m} + \frac{2d^2-1}{2(d-1)^2}\frac{1}{m^2} + O(m^{-3}) + O(m^{-1}V^{-1/3}),$$

which extends the asymptotics found in [10] by one order. The term $O(m^{-1}V^{-1/3})$ is the width of the critical window. For $d = 4, 5, 6$ we have $m^{-3} = O(m^{-1}V^{-1/3})$, and so the above formula represents the full asymptotic expansion of $p_c^{(d)}$. In [16] we show that this formula is a crucial ingredient in the study of critical bond percolation on $H(d,n)$ for $d = 2, 3, 4$. The proof uses a lace expansion for the upper bound and a novel comparison with a branching random walk for the lower bound. The proof of the lower bound also yields a refined asymptotics for the susceptibility of a subcritical Erdős-Rényi random graph.

MSC 2010. 60K35, 60K37, 82B43.
Keywords and phrases. Hamming graph, percolation, critical point, critical window, lace expansion.

1. Introduction and main result

1.1. **Percolation on the Hamming graph.** The Hamming graph $H(d,n)$ is the Cartesian product of d complete graphs on n vertices (e.g., $H(3,7) = K_7 \times K_7 \times K_7$). Bernoulli bond percolation is the model where, given a graph, each edge is retained independently with the same probability p. In this paper we study the location of the critical point of bond percolation on $H(d,n)$ for the phase transition in the size of the largest connected component when d is fixed and $n \to \infty$.

Formally, we define the Hamming graph $H(d,n)$ for $d, n \in \mathbb{N}$ as the graph with vertex set $\mathcal{V} := \{0, 1, \dots, n-1\}^d$ and edge set

$$\mathcal{E} := \{(v,w) : v, w \in \mathcal{V},\ v_j \neq w_j \text{ for exactly one } j\}. \tag{1.1}$$

Thus, $H(d,n)$ is a transitive graph on $V := n^d$ vertices with degree $m := d(n-1)$. Bernoulli bond percolation is synonymous with the probability space $(\Omega, \mathbb{P}_p)$, where $\Omega := \{0,1\}^{\mathcal{E}}$ and $\mathbb{P}_p$ is the measure such that

$$\mathbb{P}_p(\omega) = \prod_{e \in \mathcal{E}} \big((1-p)\delta_{0,\omega(e)} + p\delta_{1,\omega(e)}\big) \qquad \forall\, \omega \in \Omega, \tag{1.2}$$

Date: September 2, 2018.

where $\delta_{x,y}$ is the Kronecker delta. When $\omega(e) = 1$ we say that the edge e is *open*, when $\omega(e) = 0$ we say that the edge e is *closed*. Given a vertex $x \in \mathcal{V}$, we write $\mathscr{C}(x)$ for the graph whose vertex set consists of all vertices that can be reached from x through a path of open edges, and whose edge set consists of all open edges between these vertices. We call $\mathscr{C}(x)$ the *connected component* of x, or *cluster* of x, and write $|\mathscr{C}(x)|$ for its number of vertices. We write $\mathscr{C}_1$ for the cluster $\mathscr{C}(x)$ with the largest cardinality $|\mathscr{C}(x)|$ (using some tie-breaking rule). Two of the main objects of study in percolation are $|\mathscr{C}(x)|$ and $|\mathscr{C}_1|$, the cardinalities of $\mathscr{C}(x)$ and $\mathscr{C}_1$. For percolation on *infinite graphs* G it is often observed that the *critical point of the percolation phase transition on* G, defined by

$$p_c^G := \inf\{p \in [0,1] \colon \mathbb{P}_p(|\mathscr{C}(x)| = \infty) > 0\}, \tag{1.3}$$

is non-trivial, i.e., $p_c^G \in (0,1)$ (see for example Grimmett [20]) for most infinite graphs (an exception being $\mathbb{Z}^1$). Moreover, Aizenman and Barsky [1] and independently Menshikov [34] proved that on transitive graphs,

$$p_c^G = \sup\{p \in [0,1] \colon \mathbb{E}_p[|\mathscr{C}(x)|] < \infty\}. \tag{1.4}$$

Since we consider percolation on $H(d,n)$ with d, n finite and $\mathbb{P}_p$ is a product measure, any event that is measurable with respect to $\mathbb{P}_p$ has a probability that is a polynomial in p, and therefore is continuous in p: the finite model cannot undergo a non-trivial phase transition in p as described above. Nevertheless, it *does* make sense to study the percolation phase transition on finite graphs in the limit as $n \to \infty$. To see why, let us give a rough sketch of an important related problem: the emergence of the giant component in the *Erdős-Rényi Random Graph* (ERRG).

1.2. **Giant component.** The Erdős-Rényi random graph is the common name for percolation on the complete graph K_n. Erdős and Rényi [14] proved that in the limit as $n \to \infty$, if $p = p(n) < n^{-1}$, then $|\mathscr{C}_1| = \Theta(\log n)$ w.h.p., [1] while if $p > n^{-1}$, then $|\mathscr{C}_1| = \Theta(n)$ w.h.p. Moreover, zooming in on the transition point n^{-1} by choosing $p = (1+\varepsilon_n)n^{-1}$ for a sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ such that $\lim_{n\to\infty} \varepsilon_n = 0$, Bollobás [7] showed that [2]

- ▷ $|\mathscr{C}_1| = \Theta(\varepsilon_n^{-2} \log(\varepsilon_n^3 n))$ w.h.p. when $\varepsilon_n^3 n \to -\infty$ (subcritical),
- ▷ $|\mathscr{C}_1| = \Theta(n^{2/3})$ w.h.p. when $\varepsilon_n n^3 \to a \in \mathbb{R}$ (critical),
- ▷ $|\mathscr{C}_1| = \Theta(\varepsilon_n n)$ w.h.p. when $\varepsilon_n n^3 \to +\infty$ (supercritical).

What this shows is that the size of the largest component undergoes a sharp transition around n^{-1}. As mentioned above, there is no critical point for a finite graph, but the transition occurs in a slice of the parameter space with a width of order $n^{-4/3}$, which is asymptotically vanishing with respect to the center of the window located around n^{-1}. This behaviour inspired the notion of *critical window*: to indicate that the transition of the ERRG occurs around n^{-1} in a range of width $n^{-4/3}$, we use the short-hand notation [3]

$$p_c^{K_n} = n^{-1} + O(n^{-4/3}). \tag{1.5}$$

[1] Given a sequence of random variables $(X_n)_{n\in\mathbb{N}}$, we write $X_n = \Theta(f(n))$ w.h.p. (with high probability) if there exist constants $C \geq c > 0$ such that $\mathbb{P}_p(cf(n) \leq X_n \leq Cf(n)) \to 1$ as $n \to \infty$.

[2] Subsequent results in [3, 33, 36, 40] are much sharper and comprehensive than what is summarized here, and there is an extensive body of literature on the problem.

[3] Given three sequences $(a_n), (b_n), (c_n)$, we write that $a_n = b_n + O(c_n)$ when there exists a constant $K < \infty$ such that $|a_n - b_n| \leq Kc_n$ for all n.

Erdős and Spencer [15] conjectured that if we replace K_n by a more "geometric" graph sequence (their primary candidate was $H(d,2)$, the d-dimensional *hypercube*, with $d \to \infty$), then the critical behaviour should remain largely intact. In fact, it turned out that to a large extent the picture is the same for a large class of graph sequences with "sufficiently weak" geometries.

Of particular interest to us here are the papers by Borgs et al. [9–11], demonstrating that graph sequences satisfying the so-called *triangle condition* (which serves as an indicator of what is meant by sufficiently weak geometry; see e.g. [2, 9–11, 21]) have a phase transition that strongly resembles that of the ERRG, and that both $(H(d,2))_{d\in\mathbb{N}}$ and $(H(d,n))_{n\in\mathbb{N}}$ satisfy the triangle condition. More precisely, consider a sequence $(G_m)_{m\in\mathbb{N}} = (\mathcal{V}_m, \mathcal{E}_m)_{m\in\mathbb{N}}$ of vertex transitive graphs of degree m, and write $V_m := |\mathcal{V}_m|$. Write $x \longleftrightarrow y$ for the event that $y \in \mathscr{C}(x)$, and define the *two-point function* $\tau_p(x-y) := \mathbb{P}_p(x \longleftrightarrow y)$ and the *susceptibility* $\chi(p) := \mathbb{E}_p[|\mathscr{C}(x)|] = \sum_{y\in\mathcal{V}_m} \tau_p(x-y)$ (note that $\chi(p)$ does not depend on x by transitivity and that $\tau_p(x-y)$ depends on the relative difference of x and y only because the graphs under consideration are tori). The triangle condition is satisfied for percolation on (G_m) if for all p such that $\chi(p)^3/V_m \le \beta_0$ for some sufficiently small β_0, and for all $x, y \in \mathcal{V}_m$, we have [4]

$$\nabla_p(x,y) := \sum_{u,v} \tau_p(x-u)\tau_p(u-v)\tau_p(v-y) = \delta_{x,y} + 10\,\frac{\chi(p)^3}{V} + O(m^{-1}). \tag{1.6}$$

Borgs et al. prove that the triangle condition holds for a class of models that includes $(H(d,n))_{n\in\mathbb{N}}$ for any fixed $d \ge 2$ (see [10, Theorem 1.3]). An alternative proof, applying to e.g. Hamming graphs and hypercubes, was given by van der Hofstad and Nachmias [26, 27].

1.3. **Critical window.** Fix some $\theta \in (0,\infty)$ and define $p_c^{G_m}(\theta)$ as the unique solution of the equation

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

Borgs et al. [9, 10] prove that if we consider percolation on a sequence (G_m) that satisfies the triangle condition (1.6) with $p = p_c(\theta)(1+\varepsilon_m)$ and $\varepsilon_m \to 0$, then we see subcritical behaviour when $\varepsilon^3 V \to -\infty$ and critical behaviour when $\varepsilon^3 V \to a \in \mathbb{R}$, just as in the ERRG. Sharper results about mean-field supercritical behaviour of percolation models when $\varepsilon^3 V \to \infty$ were derived later by van der Hofstad and Nachmias [26], who investigate the supercritical phase and thus establish that (1.7) really constitutes the critical window for several high-dimensional tori including the hypercube and Hamming graphs. Moreover, it was shown in [9, Theorem 1.1] that the critical window satisfies

$$p_c(\theta) = m^{-1} + O(m^{-2}) + O(m^{-1}V^{-1/3}), \tag{1.8}$$

and that $p_c(\theta_1) - p_c(\theta_2) = O(m^{-1}V^{-1/3})$ for any $\theta_1, \theta_2 > 0$, i.e., any choice of θ yields *the same* critical window.

Compare (1.8) with the critical window of the ERRG in (1.5), and note that K_n has $O(m^{-1}V^{-1/3}) = O(n^{-4/3})$ because $m = n-1$ and $V = n$. Thus, by that analogy, the second error term above corresponds to the width of the critical window, while the first error term can be viewed as a "correction" in m^{-1} to p_c itself. In this interpretation, (1.8) describes the critical window *asymptotically precisely* for the two-dimensional Hamming graph $H(2,n)$, since

[4] Here and below we will frequently suppress sub- and superscripts when their presence is clear from context. Likewise, we do not always stress that we are considering asymptotic results for sequences.

in this case $m = 2(n-1)$ and $V = n^2$, so that the correction term m^{-2} is vanishingly small compared to $m^{-1}V^{-1/3}$. Moreover, (1.8) is also asymptotically precise for $H(3,n)$ because the two O-terms coincide.

1.4. **Expansion of the critical point.** This brings us to the main result of our paper. We write $p_c^{(d)}(\theta)$ for the critical value of percolation on $H(d,n)$ defined in (1.8), and compute the second term of $p_c^{(d)}(\theta)$ for all $d \geq 2$:

Theorem 1.1 (Critical window for percolation on $H(d,n)$)**.** *For all $\theta \in (0,\infty)$ and all $d \geq 2$,*

$$p_c^{(d)}(\theta) = m^{-1} + \frac{2d^2-1}{2(d-1)^2}\, m^{-2} + O(m^{-3}) + O(m^{-1}V^{-1/3}), \tag{1.9}$$

where the constants in the error terms may depend on θ.

Observe that for $d \geq 4$, the correction term of order m^{-2} is asymptotically larger than the width of the critical window, and that when $d = 4,5,6$ the above expansion is again asymptotically precise, since we have $m^{-3} = O(m^{-1}V^{-1/3})$.

To see the relevance of Theorem 1.1, we compare it with other expansions of p_c in the literature. The van der Hofstad and Slade [29] proved that for percolation on G, with G either the infinite lattice $\mathbb{Z}^d$ with nearest-neighbour edges or the hypercube $H(d,2)$, as $d \to \infty$, p_c^G can be expanded up to three terms as

$$p_c^G = m^{-1} + m^{-2} + \frac{7}{2}m^{-3} + O(m^{-4}), \tag{1.10}$$

where in both cases m denotes the degree of the graph G. Moreover, they [28] also proved that, for any $N \in \mathbb{N}$,

$$p_c^{\mathbb{Z}^d} = \sum_{k=1}^{N} a_k (2d)^{-k} + O((2d)^{-N-1}), \qquad p_c^{H(d,2)}(\theta) = \sum_{k=1}^{N} b_k d^{-k} + O(d^{-N-1}), \tag{1.11}$$

where $(a_k), (b_k)$ are *rational* coefficients. The critical window of the hypercube has width $O(d^{-1}2^{-d/3})$, so we believe that the expansion cannot be asymptotically precise, regardless of the choice of N. Furthermore, it was conjectured that the expansion for $p_c^{\mathbb{Z}^d}$, although it may exist, is *divergent* for all d as $N \to \infty$ (in the sense that the power series $z \mapsto \sum_{k=1}^{\infty} a_k z^k$ has radius of convergence 0). We conjecture that the expansion for the Hamming graph is very different. We believe that for any $d \geq 2$ there exist coefficients $(c_k(d))$ such that

$$p_c^{(d)}(\theta) = m^{-1} + \frac{2d^2-1}{2(d-1)^2}m^{-2} + \sum_{k=3}^{\lfloor d/3 \rfloor} c_k(d) m^{-k} + O(m^{-1}V^{-1/3}), \tag{1.12}$$

i.e., we conjecture that $p_c^{(d)}$ has an asymptotically precise expansion in m^{-1} of order $\lfloor d/3 \rfloor$ for all d. Heydenreich and van der Hofstad state the conjecture in (1.12) as [22, Open Problem 15.4].

Theorem 1.1 in [9] confirms this conjecture for $d = 2,3$, and our current work confirms it for $d = 4,5,6$. The argument of van der Hofstad and Slade [28] establishing (1.11) for the lattice and the hypercube crucially uses the fact that a ball of a radius r restricted to a d' dimensional subspace has the same shape for all $d \geq d'$, so that we can express each coefficient in terms of events that happen on a *fixed* subgraph. Balls in the Hamming graph instead grow very rapidly when n increases. Each coefficient is obtained as a limit and it will be

more involved to prove the existence of this limit. Hence we do not have significant evidence suggesting that all coefficients in (1.12) have to be rational.

We note that the existence of a finite asymptotically precise expansion makes the proof of the critical window of the Hamming graph more challenging than for the hypercube. Roughly speaking, because the critical window of the hypercube is exponentially narrower than any of the expansion terms, we can approximate $p_c^{H(d,2)}$ up to any fixed order by a value p that is in fact *subcritical*, by choosing a negative coefficient for the error term. This allows one to exploit the fact that $\chi(p)$ is poly-logarithmic in V, which simplifies the analysis considerably. In our case, the approximating p will be much closer to $p_c^{(d)}$, and so we need a much more refined analysis. We will explain this in more detail in Section 5.5.

1.5. **Scaling limit of largest cluster sizes.** Besides offering an interesting comparison with other graphs with sufficiently weak geometry, the expansion of $p_c^{(d)}$ also has another motivation. The Hamming graph is an excellent example to investigate the universality class of the ERRG, since it has a non-trivial geometry yet is highly mean field. See [17, 24, 25, 35] for a small sample of the literature from this perspective. A crucial motivation for the present paper is that it serves as a companion paper to [16], where we establish the *scaling limit* of the cluster sizes of the largest clusters within the critical window. More precisely, writing $\mathscr{C}_j$ for the j-th largest cluster, we prove that for any fixed $N \in \mathbb{N}$ and for $d = 2, 3, 4$ the largest critical clusters of Hamming graph percolation satisfy

$$\big(V^{-2/3}|\mathscr{C}_j|\big)_{j\geq 1} \xrightarrow{\mathrm{d}} (X_j)_{j\geq 1} \tag{1.13}$$

for a certain sequence of θ-dependent continuous random variables $(X_i)_{i\in\mathbb{N}}$ supported on $[0,\infty)$. Aldous [3] proved this scaling limit for the ERRG. Since then, many other random graph models have been shown to have the same (or at least a similar) scaling limit. See for instance [5, 6, 32, 38] and the references therein. The above result for the Hamming graph, however, is the first indication that the same scaling occurs for models with an underlying high-dimensional geometry. Moreover, it is the most precise determination to date of the critical behaviour of percolation on a finite transitive graph (other than the ERRG scaling limit of Aldous). The proof of (1.13) and various other results in [16] crucially rely on the asymptotically precise determination of the critical window that we give here.

1.6. **Alternative definition of the critical point.** It is worth noting that a disadvantage of the definition p_c in (1.7) is that it imposes an ad hoc relation between p_c and $V^{1/3}$, which is known not to hold in general and believed to be associated with "high-dimensional" models. In other words, (1.7) is possibly only a valid definition of p_c for percolation models in the universality class of the ERRG. Nachmias and Peres in [37] observed that it would be desirable to have a definition of p_c that applies more generally, and they proposed

$$\tilde{p}_c^G := \operatorname*{argmax}_{p\in(0,1)} \frac{\frac{\mathrm{d}}{\mathrm{d}p}\chi_G(p)}{\chi_G(p)} \tag{1.14}$$

as a definition of the critical point for *any* graph G. Their motivation for this definition is that Russo's formula [41] implies that $p = \tilde{p}_c^G$ is the point where a small change in p has the greatest impact on the relative size of the connected components, i.e., $\chi(p)$ changes most dramatically at $\tilde{p}_c^G$. A serious downside of this definition appears to be that $\tilde{p}_c^G$ may be very difficult to compute. Thus far, the only non-trivial determination of $\tilde{p}_c^G$ is given in recent work

by Janson and Warnke [31]. They determine that, for the ERRG, $|\tilde{p}_c^{K_n} - 1/n| = O(n^{-4/3})$, so $\tilde{p}_c^{K_n}$ is a point inside the critical window (1.5), that $\chi(p)^{-1}\frac{\mathrm{d}}{\mathrm{d}p}\chi(p)$ around $\tilde{p}_c^{K_n}$ describes the critical window (1.5) as well, and that, interestingly, $\tilde{p}_c^{K_n}$ does *not* equal either $1/n$ or $1/(n-1)$. It would be interesting to see whether their methods can be applied to the current setting of percolation on $H(d,n)$.

1.7. **Susceptibility of the subcritical ERRG.** In Section 2 we prove Theorem 1.1, and also derive refined asymptotics for the susceptibility of a subcritical ERRG, its second moment, and its *surplus:* given a connected graph G, let $\mathrm{Sp}(G) := |\mathcal{E}(G)| - |\mathcal{V}(G)| + 1$ denote the number of *surplus edges* in G. Besides being interesting in their own right, these will be crucial for proving the lower bound on $p_c^{(d)}$, because the restriction of critical percolation on $H(d,n)$ to a one-dimensional subspace of $H(d,n)$ is equivalent to a subcritical ERRG. To prove the lower bound of Theorem 1.1 we rely on the following asymptotics, which, to the best of our knowledge, are sharper than results in the literature:

Theorem 1.2 (Second order asymptotics for susceptibility of the subcritical ERRG)**.** *Let $G = G(n,p)$ be the ERRG with $p = \frac{\lambda}{n-1}$ and $0 < \lambda < 1$. Then as $n \to \infty$,*

$$\chi_G(p) = \mathbb{E}_p[|\mathscr{C}(v)|] = \frac{1}{1-\lambda} - \frac{2\lambda^2 - \lambda^4}{2(1-\lambda)^4} n^{-1} + O(n^{-2}), \tag{1.15}$$

$$\mathbb{E}_p[|\mathscr{C}(v)|^2] = \frac{1}{(1-\lambda)^3} + O(n^{-1}), \tag{1.16}$$

$$\mathbb{E}_p[\mathrm{Sp}(\mathscr{C}(v))] = \frac{\lambda^3}{2(1-\lambda)^2} n^{-1} + O(n^{-2}). \tag{1.17}$$

The second-order coefficient computed in (1.15) improves the result by Durrett in [13, Theorem 2.2.1], which states that $\chi_G(p) = (1-\lambda)^{-1} - O(n^{-1})$, while (1.16) provides the matching lower bound to well-known upper bound derived with the usual branching process domination. To achieve the sharper asymptotics we need a new way to encode the usual breadth-first search in the ERRG with the help of a branching random walk. We believe that there exists an infinite polynomial expansion of $\chi_G(p)$ in powers of p for all $p = \frac{\lambda}{n-1}$ with $0 < \lambda < 1$. There is substantial literature related to (1.17), see e.g. the classic book on random graphs by Bollobás [8, Section 5.2] as well as the seminal paper by Janson, Knuth Łuczak and Pittel [30] computing generating functions of components having various cycle structures. As far as we are aware, the second order asymptotics in (1.17) is new.

1.8. **Outline.** We prove Theorem 1.1 by separately proving a lower bound and an upper bound on $p_c^{(d)}$. In Section 2 we prove Theorem 1.2. This theorem is used in Section 3 to prove the lower bound in Theorem 1.1 with the help of an *exploration process* that uses the fact that the restriction of critical bond percolation on $H(d,n)$ to a one-dimensional subspace has the same distribution as a subcritical ERRG. This is used to obtain a sharp enough *branching process* upper bound on the susceptibility. In Section 4 we estimate connection probabilities and estimate bubble, triangle and polygon diagrams. In Section 5 we prove the upper bound in Theorem 1.1 with the help of the *lace expansion*. Perhaps surprisingly, these disparate methods yield compatible bounds, due to the fact that both methods are asymptotically sharp. The lace expansion method may be improved to prove Theorem 1.1, but this would be

more difficult than our current proof and less interesting. We do not see how the exploration process proof could be improved to also prove the upper bound in Theorem 1.1.

2. Susceptibility of the subcritical Erdős-Rényi Random Graph

In this section we prove Theorem 1.2. To give our estimate of the expected size of a subcritical cluster, we couple a breadth-first exploration process of the cluster to a process related to a *Branching Random Walk* (BRW). The breadth-first exploration exploration process is defined in Section 2.1, the branching random walk exploration in 2.2. The proof of the susceptibility asymptoticis is given in Section 2.3.

2.1. Breadth-first/surplus exploration.

We start by defining a version of the breadth-first (BF) exploration. This is a very standard tool in the study of the ERRG (see e.g. [23, Section 5.2.1]). In a nutshell, a breadth-first exploration is a process that, starting from a vertex v, "discovers" its adjacent edges, "activating" the direct neighbours of v in some fixed order, and then explores those vertices, discovering their adjacent edges and activating any unexplored, unactivated neighbours, and so on, always choosing the vertex that was activated the longest time ago as the next vertex to explore from. The BF exploration keeps track of which vertices have been explored (the "dead" set), which vertices have been activated but not explored (the "active" set), and the time at which a vertex was activated or explored. Crucially, the "traditional" BF exploration will only explore a vertex once, so the process terminates once all vertices are explored, and the edges associated with newly activated vertices describe a *subtree* of the component of v, but the process provides little information about the *surplus*, i.e., the discovered edges that do not activate new vertices (also sometimes referred to as the "tree excess" of the graph). For our purposes it is important that we also know about the surplus, so we consider the following modification of the BF exploration:

Definition 2.1 (BF exploration process of a graph)**.** *Given a graph $G = (\mathcal{V}, \mathcal{E})$ and a vertex $v \in \mathcal{V}$ we define the* breadth-first/surplus (BF) exploration process *as the sequence of* dead, active *and* surplus sets $(\mathbf{D}(t), \mathbf{A}(t), \mathbf{Sp}(t))_{t\geq 0}$ *as follows:*

- ▷ ***Initiation.*** *Initiate the exploration with the dead, active and surplus sets at time $t = 0$ as*

$$\mathbf{D}(0) := \varnothing, \qquad \mathbf{A}(0) := \{v\}, \qquad \mathbf{Sp}(0) := \varnothing, \tag{2.1}$$

and at time $t = 1$ as

$$\begin{aligned} \mathbf{D}(1) &:= \{v\}, \\ \mathbf{A}(1) &:= \{w \,:\, \{v, w\} \in \mathcal{E}\}, \\ \mathbf{Sp}(1) &:= \varnothing. \end{aligned} \tag{2.2}$$

- ▷ ***Time $t \geq 2$.*** *Choose the vertex $v^t \in \mathbf{A}(t-1)$ that minimizes $\min\{i : v^t \in \mathbf{A}(i)\}$, breaking ties according to an arbitrary but predetermined rule.*[5] *Update the active,*

[5] An example of such a rule: Fix an order on the vertex set $\mathcal{V}$. If at step $t-1$ we have explored and/or activated a total of k vertices, and we activate ℓ more at step t, then we assign to these ℓ newly explored vertices the labels $k+1$ through $k+\ell$, according to the order on $\mathcal{V}$. At time $s+1$ we explore from the active vertex with the smallest label.

dead and surplus sets as follows:

$$\begin{aligned}\mathbf{D}(t) &:= \mathbf{D}(t-1) \cup \{v^t\},\\ \mathbf{A}(t) &:= (\mathbf{A}(t-1) \setminus \{v^t\}) \cup \{w \notin \mathbf{A}(t-1) \cup \mathbf{D}(t-1) : \{v^t, w\} \in \mathcal{E}\},\\ \mathbf{Sp}(t) &:= \mathbf{Sp}(t-1) \cup \{\{v^t, w\} \in \mathcal{E} : w \in \mathbf{A}(t-1)\}.\end{aligned} \tag{2.3}$$

▷ **Stop.** *Terminate the exploration when* $\mathbf{A}(t) = \varnothing$. *Set* $T = t$.

Note that $\mathbf{D}(t)$ and $\mathbf{A}(t)$ are subsets of $\mathcal{V}$, whereas $\mathbf{Sp}(t)$ is a subset of $\mathcal{E}$. When $\mathbf{A}(t) = \varnothing$, this means that we have completely explored the connected component $\mathscr{C}(v)$ and $T = |\mathbf{D}(T)| = |\mathscr{C}(v)|$. In the BF we find a new edge every time we activate a vertex (except the initial vertex v) or we discover an edge between active vertices. It follows that $|\mathcal{E}(\mathscr{C}(v))| = |\mathbf{D}(T)| - 1 + |\mathbf{Sp}(T)|$. We conclude that that $|\mathbf{Sp}(T)| = |\mathcal{E}(\mathscr{C}(v))| - |\mathbf{D}(T)| + 1 = \mathrm{Sp}(\mathscr{C}(v))$.

2.2. **The branching random walk exploration.** The subtree generated by a "traditional" BF exploration is often studied through a comparison to a branching process (see e.g. [13, 23]). To study our BF exploration, we define a suitable extension, the *branching random walk (BRW) exploration,* in which we randomly embed a branching process in the graph, and keep track of its self-intersections.[6] This is made precise in the following definition:

Definition 2.2 (Branching random walk)**.** *Given an* m*-regular graph* $G_m = (\mathcal{V}, \mathcal{E})$ *and* $p \in [0,1]$*, we define the* p-branching random walk *(*p*-BRW) on* G_m *started at* $v \in \mathcal{V}$ *as the pair* $(\mathcal{T}, \phi_v)$*, where* $\mathcal{T}$ *is a* $\mathsf{Bin}(m,p)$ *Galton-Watson tree, and* ϕ_v *is a random mapping of* $\mathcal{T}$ *into the vertex set* $\mathcal{V}$ *whose law satisfies: (1)* ϕ_v *maps the root* ρ *of* $\mathcal{T}$ *to* v*; (2) given any node* $x \in \mathcal{T}$ *and its set of children* $\mathbf{C}(x) \subset \mathcal{T}$*, the marginal law of* $\phi_v(\mathbf{C}(x))$ *is the same as that of* $|\mathbf{C}(x)|$ distinct *neighbours of* $\phi_v(x)$ *in* G_m *chosen uniformly at random, independently for all* $x \in \mathcal{T}$*. (Here, for a set* $A \subset \mathcal{T}$ *and a mapping* $\phi_v \colon \mathcal{T} \to G$*, we define* $\phi_v(A) = \cup_{a \in A} \phi_v(a)$*, and by convention set* $\phi_v(\varnothing) = \varnothing$*.)*

Next, we define a process that explores a p-BRW and keeps track of any self-intersections. Briefly, the idea is that we explore the p-BRW by exploring the tree $\mathcal{T}$ in a breadth-first fashion from the root upward. If the p-BRW intersects its own trace, then we declare the particle that intersected, and all its offspring, to have become "ghosts". We differentiate between particles that became ghosts through intersecting with active and dead vertices. In Proposition 2.4 below we prove that this exploration process can be coupled to a BF exploration of a percolation cluster:

Definition 2.3 (BRW exploration process)**.** *Given an* m*-regular graph* $G_m = (\mathcal{V}, \mathcal{E})$*, a vertex* $v \in \mathcal{V}$*, and a* p*-BRW* $(\mathcal{T}, \phi_v)$ *on* G_m*, we define the* BRW exploration process $(\mathtt{A}(t), \mathtt{D}(t), \mathtt{P}^{\mathtt{A}}(t), \mathtt{P}^{\mathtt{D}}(t))_{t=0}^{T}$ *as the sequence of* dead, active, active ghost *and* dead ghost *sets as follows:*

▷ ***Initiation.*** *Initiate the exploration with the* dead, active, active ghost *and* dead ghost *sets at time* $t = 0$ *as*

$$\mathtt{D}(0) := \varnothing, \quad \mathtt{A}(0) := \{\rho\}, \quad \mathtt{P}^{\mathtt{A}}(0) = \varnothing, \quad \mathtt{P}^{\mathtt{D}}(0) = \varnothing, \tag{2.4}$$

[6] From now on the term *nodes* will refer to elements of GW trees, while *vertices* will refer to elements of graphs. Moreover, the *progeny* of a node x will indicate the set of vertices whose path to the root ρ passes through x, while the *children* of x are only the vertices for which x is the first vertex encountered on such a path. We write $\mathbf{C}(x)$ for the set of children of x in $\mathcal{T}$.

and at time $t = 1$ as

$$\mathtt{D}(1) := \{\rho\}, \quad \mathtt{A}(1) := \{y \in \mathbf{C}(\rho)\}, \quad \mathtt{P}^{\mathtt{A}}(1) := \varnothing, \quad \mathtt{P}^{\mathtt{D}}(1) := \varnothing. \tag{2.5}$$

▷ ***Time*** *$t \geq 2$. Choose the node $x^t \in \mathtt{A}(t-1)$ that minimizes $\min\{i : x^t \in \mathtt{A}(i)\}$, breaking ties according to an arbitrary but predetermined rule, and update the exploration as follows:*

$$\begin{aligned}
\mathtt{D}(t) &:= \mathtt{D}(t-1) \cup \{x^t\},\\
\mathtt{A}(t) &:= (\mathtt{A}(t-1) \setminus \{x^t\}) \cup \big\{y \in \mathbf{C}(x^t) : \phi_v(y) \notin \phi_v\big(\mathtt{D}(t-1) \cup \mathtt{A}(t-1)\big)\big\},\\
\mathtt{P}^{\mathtt{A}}(t) &:= \mathtt{P}^{\mathtt{A}}(t-1) \cup \big\{y \in \mathbf{C}(x^t) : \phi_v(y) \in \phi_v\big(\mathtt{A}(t-1)\big)\big\},\\
\mathtt{P}^{\mathtt{D}}(t) &:= \mathtt{P}^{\mathtt{D}}(t-1) \cup \big\{y \in \mathbf{C}(x^t) : \phi_v(y) \in \phi_v\big(\mathtt{D}(t-1)\big)\big\}.
\end{aligned} \tag{2.6}$$

▷ ***Stop.*** *If $\mathtt{A}(t) = \varnothing$, then terminate the exploration. Set $T = t$.*

Using the BRW exploration, we define the subgraph $\tilde{\mathscr{C}}(v)$ as the graph traced out by a p-BRW where the particles are killed when they intersect with the active set. More precisely, we let $\tilde{\mathcal{T}}$ be the subtree in $\mathcal{T}$ induced by $\mathtt{D}(T) \cup \mathtt{P}^{\mathtt{A}}(T)$, and define

$$\tilde{\mathscr{C}}(v) := \big(\phi_v(\mathtt{D}(T)), \{\{\phi_v(x), \phi_v(y)\} : \{x, y\} \in \tilde{\mathcal{T}}\}\big). \tag{2.7}$$

Note that, by Definition 2.3, $\phi_v(\mathtt{D}(T) \cup \mathtt{P}^{\mathtt{A}}(T)) = \phi_v(\mathtt{D}(T))$, so $\tilde{\mathscr{C}}(v)$ is indeed a subgraph of $G_m = (\mathcal{V}, \mathcal{E})$.

We now show that $\tilde{\mathscr{C}}(v)$ has the same law as $\mathscr{C}(v)$, the connected component of v in an ERRG, by coupling the BF and BRW explorations:

Proposition 2.4 (Coupling of BF and BRW explorations)**.** *Consider percolation on an m-regular graph G_m with parameter p. Consider the BF exploration on the percolated graph $G_m(p)$ and the p-BRW exploration processes on G_m, both starting from the vertex v (and using the same tie-breaking rule). Then $\mathscr{C}(v)$ with respect to $\mathbb{P}_p$ has the same law as $\tilde{\mathscr{C}}(v)$.*

Proof. We show inductively that we can couple each step of the BRW and of the BF exploration in such a way that $\tilde{\mathscr{C}}(v) = \mathscr{C}(v)$ almost surely. We start by showing that there exists a coupling such that for all $t \geq 0$,

$$\begin{aligned}
\mathbf{D}(t) &= \phi_v(\mathtt{D}(t)),\\
\mathbf{A}(t) &= \phi_v(\mathtt{A}(t)),\\
\mathbf{Sp}(t) &= \bigcup_{s \leq t} \big\{\{\phi_v(x^s), w\} : w \in \phi_v(\mathtt{P}^{\mathtt{A}}(s) \setminus \mathtt{P}^{\mathtt{A}}(s-1))\big\}.
\end{aligned} \tag{2.8}$$

We start with the inductive base. At time $t = 0$, by Definitions 2.1 and 2.3,

$$\begin{aligned}
\mathbf{D}(0) &= \varnothing = \phi_v(\mathtt{D}(0)),\\
\mathbf{A}(0) &= \{v\} = \phi_v(\{\rho\}) = \phi_v(\mathtt{A}(0)),\\
\mathbf{Sp}(0) &= \varnothing = \big\{\{\phi_v(x^0), w\} : w \in \phi_v(\mathtt{P}^{\mathtt{A}}(0))\big\}.
\end{aligned} \tag{2.9}$$

Next, we prove the inductive step: the induction hypothesis is that the relations in (2.8) holds for all $r < t$. We extend the coupling so that it also holds at time t. Our assumption is that we use the same tie-breaking rule for both explorations, so by the induction hypothesis we choose $v^t = \phi_v(x^t)$.

Given $\phi_v(x^t)$, fix a set $\mathcal{U}_k^t = \{u^1, u^2, \dots, u^k\}$ of k neighbours of $\phi_v(x^t)$. By Definition 2.2, the mapping ϕ_v is such that $|\mathbf{C}(x^t)|$ neighbours of $\phi_v(x^t)$ are distinct neighbours chosen uniformly at random, so

$$\begin{aligned}\mathbb{P}\big(\phi_v(\mathbf{C}(x^t)) = \mathcal{U}_k^t\big) &= \mathbb{P}(|\mathbf{C}(x^t)| = k)\,\mathbb{P}\big(\phi_v(\mathbf{C}(x^t)) = \mathcal{U}_k^t \mid |\mathbf{C}(x^t)| = k\big)\\ &= \binom{m}{k} p^k (1-p)^{m-k} \binom{m}{k}^{-1} = p^k(1-p)^{m-k}.\end{aligned} \tag{2.10}$$

Next, consider the BF exploration at time t. Given $(\mathbf{D}(s), \mathbf{A}(s), \mathbf{Sp}(s))_{s=0}^{t-1}$, we can determine v^t. For $t \geq 0$, let $\mathbf{N}(t)$ denote an independent set-valued random variable that contains the vertex w with probability p, independently for all w such that $\{v^t, w\} \in \mathcal{E}$, so that $\mathbb{P}\big(\mathbf{N}(t) = \mathcal{U}_k^t\big) = p^k(1-p)^{m-k}$. For every set $\mathcal{U}_k^t$ of neighbours of v^t we have $\mathbb{P}\big(\mathbf{N}(t) = \mathcal{U}_k^t\big) = \mathbb{P}\big(\phi_v(\mathbf{C}(x^t)) = \mathcal{U}_k^t\big)$, and so there exists a trivial coupling of $\mathbf{N}(t)$ and $\phi_v(\mathbf{C}(x^t))$ such that $\mathbb{P}\big(\mathbf{N}(t) = \phi_v(\mathbf{C}(x^t))\big) = 1$.

Consider an edge $\{v^t, w\}$. Observe that if $w \notin \mathbf{D}(t-1)$, then $\{v^t, w\}$ has not been discovered by the exploration, so it is open in the percolation conditionally independently with probability p, while if $w \in \mathbf{D}(t-1)$, then $\{v^t, w\}$ has been discovered in the BF exploration, so its status can be determined from $(\mathbf{D}(s), \mathbf{A}(s), \mathbf{Sp}(s))_{s=0}^{t-1}$. Let $\mathbf{X}(t)$ denote the vertices that are end-points of edges that are discovered in the t-th step, i.e.,

$$\mathbf{X}(t) := \big\{\{v^t, w\} \,:\, \{v^t, w\} \in \mathscr{C}(v) \setminus \{\{v^t, u\} \,:\, u \in \mathbf{D}(t-1)\}\big\}. \tag{2.11}$$

Note that if $w \in \mathbf{X}(t)$, then either w becomes activated at time t or $w \in \mathbf{A}(t-1)$. By the above observation, we can couple $\mathbf{X}(t)$ to $\mathbf{N}(t)$ such that $\mathbf{X}(t) = \mathbf{N}(t) \setminus \mathbf{D}(t-1)$ almost surely, conditionally on $(\mathbf{D}(s), \mathbf{A}(s), \mathbf{Sp}(s))_{s=0}^{t-1}$.

Consider henceforth the setting in which $\mathbf{X}(t)$, $\mathbf{N}(t)$, $\phi_v(\mathbf{C}(v^t))$, and $(\mathbf{D}(s), \mathbf{A}(s), \mathbf{Sp}(s))_{s=0}^{t-1}$ are *simultaneously* coupled according to the above description. (Since both $\mathbf{N}(t)$ and $\phi_v(\mathbf{C}(x^t))$ are essentially independent p-random subsets, it is easy to make this coupling explicit; we leave those details to the reader.) Using this coupling, the induction hypothesis (2.8), and Definitions 2.1 and 2.3, we derive

$$\mathbf{D}(t) \setminus \mathbf{D}(t-1) = \{v^t\} = \phi(x^t) = \phi_v(\mathtt{D}(t)) \setminus \phi_v(\mathtt{D}(t-1)), \tag{2.12}$$

and

$$\begin{aligned}\mathbf{A}(t) \setminus \mathbf{A}(t-1) &= \mathbf{X}(t) \setminus \mathbf{A}(t-1)\\ &= \mathbf{N}(t) \setminus (\mathbf{D}(t-1) \cup \mathbf{A}(t-1))\\ &= (\phi_v(\mathbf{C}(x^t)) \setminus (\phi_v(\mathtt{D}(t-1)) \cup \phi_v(\mathtt{A}(t-1)))\\ &= \phi_v(\mathtt{A}(t)) \setminus \phi_v(\mathtt{A}(t-1)),\end{aligned} \tag{2.13}$$

and

$$\begin{aligned}\mathbf{Sp}(t) \setminus \mathbf{Sp}(t-1) &= \{\{v^t, w\} \,:\, w \in \mathbf{X}(t) \cap \mathbf{A}(t-1)\}\\ &= \{\{v^t, w\} \,:\, w \in \mathbf{N}(t) \cap \mathbf{A}(t-1)\}\\ &= \{\{\phi_v(x^t), w\} \,:\, w \in \phi_v(\mathbf{C}(x^t)) \cap \phi_v(\mathtt{A}(t-1))\}\\ &= \{\{\phi_v(x^t), w\} \,:\, w \in \phi_v(\mathtt{P}^{\mathtt{A}}(t) \setminus \mathtt{P}^{\mathtt{A}}(t-1))\}.\end{aligned} \tag{2.14}$$

Since $\phi_v(x^t) = v^t$, we obtain that (2.8) holds also at time t almost surely, and thus, by induction, for all $t \in \{0, 1, \dots, T\}$, almost surely.

To conclude the proof, we show that the coupling (2.8) for all $t \in \{0, 1, \dots, T\}$ implies that $\tilde{\mathscr{C}}(v) = \mathscr{C}(v)$ almost surely. Recall the definition of $\tilde{\mathscr{C}}(v)$ in (2.7), and of $\tilde{\mathcal{T}}$ above it. Since $\phi_v(\mathtt{D}(T)) = \mathbf{D}(T)$, it follows directly from (2.8) that the vertex sets of $\tilde{\mathscr{C}}(v)$ and $\mathscr{C}(v)$ coincide. To see that the edge sets coincide, note that, by Definition 2.3, $\tilde{\mathcal{T}}$ contains only edges $\{x, y\} \in \mathcal{T}$ such that $y \in \mathbf{C}(x)$, with $x \in \mathtt{D}(T)$ and $y \in \mathtt{D}(T) \cup \mathtt{P}^{\mathtt{A}}(T)$. Indeed, by the construction of the BRW exploration it is impossible that both $x, y \in \mathtt{P}^{\mathtt{A}}(T)$, since vertices in $\mathtt{P}^{\mathtt{A}}$ are never explored further. Let $s = \min\{t : x \in \mathtt{D}(t)\}$. Then, from the definition of the BRW exploration, $x = x^s$ and $y \in (\mathtt{A}(s) \setminus \mathtt{A}(s-1)) \cup (\mathtt{P}^{\mathtt{A}}(s) \setminus \mathtt{P}^{\mathtt{A}}(s-1))$. We then obtain

$$\mathcal{E}(\tilde{\mathscr{C}}(v)) = \bigcup_{s=0}^{T} \big\{\{\phi_v(x^s), \phi_v(y)\} \,:\, \phi_v(y) \in \phi_v((\mathtt{A}(s) \setminus \mathtt{A}(s-1)) \cup (\mathtt{P}^{\mathtt{A}}(s) \setminus \mathtt{P}^{\mathtt{A}}(s-1))\big\}. \quad (2.15)$$

An application of (2.8) now completes the proof. □

We conclude by deriving some consequences of Proposition 2.4 that will be useful in the proof of Theorem 1.2:

Corollary 2.5. *Consider a BF exploration on an m-regular graph G_m with parameter p, and a p-BRW exploration processes on G_m, both starting from the vertex v. Then*

$$|\mathtt{D}(T)| \stackrel{d}{=} |\mathscr{C}(v)|, \quad (2.16)$$

$$|\mathtt{P}^{\mathtt{A}}(T)| \stackrel{d}{=} \mathrm{Sp}(\mathscr{C}(v)), \quad (2.17)$$

$$\mathbb{E}[|\mathtt{P}^{\mathtt{D}}(T)|] = p\mathbb{E}\Big[\sum_{t=1}^{T} |\mathtt{D}(t-1)|\Big]. \quad (2.18)$$

Proof. Equations (2.16) and (2.17) follow immediately from Proposition 2.4. To prove (2.18) we use (2.6) and (2.10), from which we get

$$\mathbb{E}[|\mathtt{P}^{\mathtt{D}}(t) \setminus \mathtt{P}^{\mathtt{D}}(t-1)| \mid \mathcal{F}(t-1)] = p|\mathtt{D}(t-1)|, \quad (2.19)$$

where $\mathcal{F}(t-1)$ is the σ-algebra generated by the first $t-1$ steps of the BRW exploration. Summing over $t \le T$ we get the claim. □

2.3. **Proof of the susceptibility asymptotics.**

Proof of Theorem 1.2. Consider a BF exploration on $G(n,p)$ and a BRW exploration of a p-BRW on K_n, both with $p = \frac{\lambda}{n-1}$, coupled as in the proof of Proposition 2.4. Corollary 2.5 implies that $\chi_G(p) = \mathbb{E}[|\tilde{\mathscr{C}}(v)|] = \mathbb{E}[|\mathtt{D}(T)|]$. We know that $\mathbb{E}[|\mathcal{T}|] = \frac{1}{1-\lambda}$ (see [39]). From the definition of the BRW exploration we know that $\mathcal{T} \setminus \mathtt{D}(T)$ consists of the nodes $\mathtt{P}^{\mathtt{D}}(T) \cup \mathtt{P}^{\mathtt{A}}(T)$ and their progeny in $\mathcal{T}$. Using (2.16) and (2.17), we can thus write

$$\begin{aligned} \mathbb{E}_p[|\mathscr{C}(v)|] &= \mathbb{E}[|\mathcal{T}|] - \mathbb{E}[|\mathcal{T}|]\,\mathbb{E}[|\mathtt{P}^{\mathtt{A}}(T) \cup \mathtt{P}^{\mathtt{D}}(T)|] \\ &= \mathbb{E}[|\mathcal{T}|] - \mathbb{E}[|\mathcal{T}|]\left(\mathbb{E}_p[\mathrm{Sp}(\mathscr{C}(v))] + p\mathbb{E}\Big[\sum_{t=1}^{T} |\mathtt{D}(t-1)|\Big]\right). \end{aligned} \quad (2.20)$$

Since $|\mathtt{D}(t-1)| = t-1$ for all $t \le T$ and $|\mathtt{D}(T)| \stackrel{d}{=} |\mathscr{C}(v)|$,

$$\mathbb{E}\Big[\sum_{t=1}^{T} |\mathtt{D}(t-1)|\Big] = \tfrac{1}{2}\mathbb{E}[T(T-1)] = \tfrac{1}{2}\mathbb{E}\big[|\mathscr{C}(v)|^2 - |\mathscr{C}(v)|\big]. \tag{2.21}$$

As a result we obtain that

$$\mathbb{E}_p[|\mathscr{C}(v)|] = \mathbb{E}[|\mathcal{T}|]\left(1 - \mathbb{E}_p[\mathrm{Sp}(\mathscr{C}(v))] - \tfrac{1}{2}p\mathbb{E}_p\big[|\mathscr{C}(v)|^2 - |\mathscr{C}(v)|\big]\right). \tag{2.22}$$

Pakes proved in [39, Section 2.2] that

$$\mathrm{Var}(|\mathcal{T}|) = \frac{\lambda}{(1-\lambda)^3} + O(p), \tag{2.23}$$

so that $\mathbb{E}[|\mathcal{T}|^2] = (1-\lambda)^{-3} + O(p)$. Moreover, Durrett shows in [13, Section 2] that $\mathbb{E}_p[|\mathscr{C}(v)|] = (1-\lambda)^{-1} - O(n^{-1})$.

We will prove that $\mathbb{E}_p[|\mathscr{C}(v)|^2] = (1-\lambda)^{-3} + O(p)$. It follows from Proposition 2.4 that $\mathbb{E}_p[|\mathscr{C}(v)|^2] = \mathbb{E}[|\mathtt{D}(T)|^2] \le \mathbb{E}[|\mathcal{T}|^2]$, which establishes the upper bound. To determine the lower bound, we write

$$\mathbb{E}_p[|\mathscr{C}(v)|^2] = \mathbb{E}[(|\mathcal{T}| - |\mathcal{T} \setminus \mathtt{D}(T)|)^2] \ge \mathbb{E}[|\mathcal{T}|^2] - 2\mathbb{E}[|\mathcal{T}||\mathcal{T} \setminus \mathtt{D}(T)|], \tag{2.24}$$

so it remains to prove that $\mathbb{E}[|\mathcal{T}||\mathcal{T} \setminus \mathtt{D}(T)|] = O(p)$. We write

$$\begin{aligned}\mathbb{E}[|\mathcal{T}||\mathcal{T} \setminus \mathtt{D}(T)|] &= \mathbb{E}\Big[\sum_{x,y\in\mathcal{T}} \mathbb{1}_{\{x\notin\mathtt{D}(T)\}}\Big] \\ &= \sum_{k=1}^{\infty} \mathbb{P}(|\mathcal{T}| = k)\, k\, \mathbb{E}\Big[\sum_{x} \mathbb{1}_{\{x\notin\mathtt{D}(T), x\in\mathcal{T}\}} \big| |\mathcal{T}| = k\Big].\end{aligned} \tag{2.25}$$

Now we bound $\mathbb{E}\big[\mathbb{1}_{\{x\notin\mathtt{D}(T),x\in\mathcal{T}\}}\big||\mathcal{T}| = k\big]$. Suppose that $d_{\mathcal{T}}(\rho, x) = L$, and that η is the path in $\mathcal{T}$ with $\eta(0) = \rho$ and $\eta(L) = x$. From the description of the BRW exploration it follows that $x \notin \mathtt{D}(T)$ if and only if there exists a $t \in \{0, \dots, T\}$ such that $v^t \in \eta$ and $\phi_v(v^t) \in \mathtt{P}^{\mathtt{A}}(t-1) \cup \mathtt{P}^{\mathtt{D}}(t-1)$. Since the mapping of children of a node in the BRW is done uniformly at random, $\phi_v(\eta)$ has the same distribution as a simple random walk path on K_n, conditioned on going from ρ to x in L steps. Conditionally on $|\mathcal{T}| = k$, we have $|\mathtt{A}(t-1) \cup \mathtt{D}(t-1)| \le k$ for all t, so for all t the probability that $\phi_v(v^t) \in \phi_v(\mathtt{A}(t-1) \cup \mathtt{D}(t-1))$ is at most $k/(n-1)$ by the symmetry of K_n. Conditionally on $|\mathcal{T}| = k$, we have $L \le k-1$, and so

$$\mathbb{E}\big[\mathbb{1}_{\{x\notin\mathtt{D}(T),x\in\mathcal{T}\}}\big||\mathcal{T}| = k\big] \le \frac{k(k-1)}{n-1}. \tag{2.26}$$

Inserting this into (2.25) we obtain

$$\mathbb{E}[|\mathcal{T}||\mathcal{T} \setminus \mathtt{D}(T)|] = \sum_{k=1}^{n} \mathbb{P}(|\mathcal{T}| = k)\, k^2\, \frac{k(k-1)}{n-1} \le \frac{p}{\lambda}\mathbb{E}[|\mathcal{T}|^4] = O(n^{-1}). \tag{2.27}$$

Inserting this into (2.24) we get (1.16), and thus it also follows that

$$\mathbb{E}_p[|\mathscr{C}(v)|^2] - \mathbb{E}_p[|\mathscr{C}(v)|] = \frac{1}{(1-\lambda)^3} - \frac{1}{1-\lambda} + O(n^{-1}) = \frac{2\lambda - \lambda^2}{(1-\lambda)^3} + O(n^{-1}). \tag{2.28}$$

Next, we compute $\mathbb{E}_p[\mathrm{Sp}(\mathscr{C}(v))]$. Note that $\mathrm{Sp}(\mathscr{C}(v))$ is bounded from above by the number of vertex-disjoint cycles in $\mathscr{C}(v)$, since removing a surplus edge from a graph destroys at least

one such cycle. Here, given a graph $G = (\mathcal{V}, \mathcal{E})$ and $v_1, \dots, v_k \in \mathcal{V}$, we say that a subgraph L_k is a (vertex-disjoint) *cycle of length* k if L_k has vertex set $\{v^1, \dots, v^k\} \subset \mathcal{V}$ and edge set $\cup_{i=1}^{k} \{v^i, v^{i+1 \bmod k}\}$. We write $\mathcal{L}_k(G)$ for the set of cycles of length k in a graph G. For $G(n,p)$ we may thus bound

$$\mathbb{E}_p[\mathrm{Sp}(\mathscr{C}(v))] \le \sum_{k=3}^{n} \sum_{L_k \in \mathcal{L}_k(K_n)} \mathbb{P}_p\big(L_k \in \mathcal{L}_k(\mathscr{C}(v))\big). \tag{2.29}$$

Note that $|\mathcal{L}_k(K_n)| = \frac{n!}{2k(n-k)!}$. The probability that a given set of k edges is open in $G(n,p)$ is p^k. Moreover, the probability that a vertex w in $G(n,p)$ is connected to a given set of k vertices is at most $k\frac{\mathbb{E}_p[|\mathscr{C}(w)|]}{n} \le \frac{k}{(1-\lambda)n}$ for any w, by the symmetry of K_n. Combining these estimates, and using that $p = \frac{\lambda}{n-1}$, we bound

$$\begin{aligned}
\mathbb{E}_p[\mathrm{Sp}(\mathscr{C}(v))] &\le \sum_{k=3}^{n} \frac{n!}{2k(n-k)!} p^k \frac{k}{(1-\lambda)n}(1 + O(n^{-1})) \\
&= \sum_{k=3}^{n} \frac{(n-1)!}{(n-k)!} \frac{1}{(n-1)^{k-1}} \frac{\lambda^{k-1}}{2(1-\lambda)}\, p\,(1 + O(n^{-1})) \\
&\le \sum_{k=3}^{n} \frac{\lambda^k}{2(1-\lambda)} n^{-1} + O(n^{-2}) = \frac{\lambda^3}{2(1-\lambda)^2} n^{-1} + O(n^{-2}).
\end{aligned} \tag{2.30}$$

This establishes the upper bound in (1.17).

It remains to prove a matching lower bound. Before we start, let us recall a standard tool from percolation theory: the *van den Berg-Kesten (BK) inequality* [4]: We say that an event A is *increasing* with respect to p if $\mathbb{P}_p(A) \ge \mathbb{P}_q(A)$ whenever $p \ge q$. We say that two increasing events A and B *occur disjointly*, and write $A \circ B$, if the occurrence of A and B can be verified by inspecting disjoint sets of edges (which may depend on the percolation configuration). For instance, the event $\{v \leftrightarrow w\}$ is increasing, and $\{v \leftrightarrow w\} \circ \{v' \leftrightarrow w'\}$ implies that there exists a path of open edges between v and w and another path of open edges between v' and w', and that these paths are edge disjoint. The BK-inequality states that $\mathbb{P}_p(A \circ B) \le \mathbb{P}_p(A)\mathbb{P}_p(B)$. See e.g. Grimmett's classic book on percolation [19] for more details.

We use the BK-inequality to prove a lower bound on the expected surplus. Since the removal of a surplus edge must destroy at least one cycle in the graph, we can bound $\mathrm{Sp}(\mathscr{C}(v))$ from below by the number of vertex-disjoint cycles in $\mathscr{C}(v)$ that are edge-disjoint from any other cycle in $\mathscr{C}(v)$. A cycle is edge-disjoint from other cycles if and only if the cycle does not contain a pair of vertices that are connected by a path outside the cycle. Each cycle of length k has $k(k-1)/2$ pairs of vertices, and the probability that any two vertices are connected is at most $\frac{1}{(1-\lambda)n}$ by the symmetry of K_n. Writing $\mathcal{V}(L_k)$ for the vertex set of the cycle L_k and using the same reasoning as in (2.30) as well as inclusion-exclusion, we thus obtain

$$\begin{aligned}
\mathbb{E}_p[\mathrm{Sp}(\mathscr{C}(v))] \ge \sum_{k=3}^{n} \sum_{L_k \in \mathcal{L}_k(K_n)} \Big[&\mathbb{P}_p\big(L_k \in \mathcal{L}_k(\mathscr{C}(v))\big) \\
&- \mathbb{P}_p\Big(\bigcup_{\{x,y\} \subset \mathcal{V}(L_k)} \big\{L_k \in \mathcal{L}_k(\mathscr{C}(v))\big\} \circ \{x \leftrightarrow y\} \Big) \Big]
\end{aligned} \tag{2.31}$$

$$
\begin{aligned}
&\geq \sum_{k=3}^{n} \frac{n!}{(n-k)!} \frac{p^k}{2(1-\lambda)n} \Big(1 - \frac{k(k-1)}{2(1-\lambda)n}\Big) \\
&\geq \sum_{k=3}^{n} \Big(\frac{\lambda^k}{2(1-\lambda)} n^{-1} + O(n^{-2})\Big)\Big(1 - \frac{k(k-1)}{2(1-\lambda)} n^{-1}\Big) \\
&= \frac{\lambda^3}{2(1-\lambda)^2} n^{-1} + O(n^{-2}),
\end{aligned}
$$

where in the first step the union is over all two-element subsets of $\mathcal{V}(L_k)$, in the second step we first use the union bound and then the BK-inequality, and in the third step we use that $p = \lambda/(n-1)$ and $\frac{(n-1)!}{(n-k)!} = (n-1)^{k-1} - O(kn^{k-2})$. This completes the proof of (1.17).

Inserting the bounds (2.28), (2.30), and (2.31) into (2.22), we conclude that

$$
\begin{aligned}
\mathbb{E}_p[|\mathscr{C}(v)|] &= \frac{1}{1-\lambda}\Big(1 - \frac{\lambda^3}{2(1-\lambda)^2} n^{-1} - \tfrac{1}{2} p \frac{2\lambda - \lambda^2}{(1-\lambda)^3} + O(n^{-2})\Big) \\
&= \frac{1}{1-\lambda} - \frac{2\lambda^2 - \lambda^4}{2(1-\lambda)^4} n^{-1} + O(n^{-2}),
\end{aligned}
\tag{2.32}
$$

which proves (1.15) and thus completes the proof of Theorem 1.2. □

3. The lower bound on $p_c^{(d)}(\theta)$ via an exploration process

In this section we use the bound on the susceptibility of the subcritical ERRG to determine a lower bound on $p_c^{(d)}(\theta)$, the critical value of the Hamming graph. We achieve this by bounding $\chi(p)$ from above with the use of an exploration process, and then substituting this bound into $\chi(p_c^{(d)}) = \theta V^{-1/3}$, the equation that defines $p_c^{(d)}$ (recall (1.7)). The exploration process that we use is designed with the geometry of the Hamming graph in mind, so let us start by investigating this geometry further.

Recall that the Hamming graph $H(d,n)$ can be viewed as the $(d-1)$-fold Cartesian product of complete graphs K_n. If we arrange the n^d vertices of $H(d,n)$ on a d-dimensional hypercubic grid in the obvious way, then the edges of $H(d,n)$ are precisely those edges that have both end-points on a line that is parallel to an axis of the grid. This inspires the following definition: Given some $i \in \{1,\dots,d\}$ and a vertex $v = (v_1,\dots,v_d) \in \{0,\dots,n-1\}^d$, we call the subgraph of $H(d,n)$ induced by the set

$$
\big\{(v_1,\dots,v_{i-1},w,v_{i+1},\dots,v_d) \in \mathcal{V} \,:\, w \in \{0,\dots,n-1\}\big\}, \tag{3.1}
$$

the *i-directional line of $H(d,n)$ through v*. When i and v are unimportant we refer to such subgraphs simply as *lines*. We write $\mathscr{C}_i(v)$ for the set of vertices that can be reached by a path of open edges in the i-directional line through v. Note that any line of $H(d,n)$ is isomorphic to K_n, so that $\mathscr{C}_i(v)$ has the same law as an Erdős-Rényi random graph on n vertices with parameter p. Moreover, because this is a graph on n vertices and we choose $p = \frac{1+O(m^{-1})}{d(n-1)}$ with $d \geq 2$ (in accordance with (1.8)), this ERRG is *subcritical*. Writing $\chi_{\text{line}}(p) := \mathbb{E}_p[|\mathscr{C}_i(v)|]$ for the expected size of a connected component *within a line* (i.e., the set over vertices that can be reached from v using only open edges in the line), we get from Theorem 1.2 that

$$
\chi_{\text{line}}\Big(\frac{1+O(m^{-1})}{d(n-1)}\Big) = \frac{1}{1-p(d-1)}\Big(1 - \frac{2d^2-1}{2(d-1)^3} m^{-1} + O(m^{-2})\Big). \tag{3.2}
$$

We use this fact repeatedly below.

We next define the exploration process that allows us to estimate $\chi(p)$. To first order, this estimation simply yields a Galton-Watson branching process. But this is an overestimate, and we can give a (negative) second order correction to it by correcting for the over counting that arises because we ignored loops in the graph. We thus will have to find a bound on the number of loops. An important insight into the structure of percolation on the Hamming graph is that loops are much more likely to occur within lines than outside lines. Our exploration crucially uses this fact: we only subtract the correction for loops within lines, which gives us the desired upper bound.

Roughly speaking, the line-wise exploration process defined below works as follows. We have two sets, the *active* and *dead* sets A and D. We start with a single vertex in the active set. At any given time we move an active vertex to the dead set, and add all the vertices connected to that vertex *through a line* to the active set. Because we want to avoid "feedback loops" in the process, we need to keep track of the line that we have previously explored from. The *parent set* P of ordered pairs of vertices and their parents in the exploration is a technical addition to the process that takes care of this. The process stops when the active set becomes empty.

Definition 3.1. *The* line-wise exploration process $(\mathsf{A}(t), \mathsf{D}(t), \mathsf{P}(t))_{t=0}^{T}$ *on* $H(d,n)$ *started at the vertex* v *is the* T*-step discrete-time process defined as follows:*

▷ ***Initiation.*** *Define the* dead, active *and* parent sets *at time* $t = 0$

$$\mathsf{D}(0) := \varnothing, \qquad \mathsf{A}(0) := \{v\}, \qquad \mathsf{P}(0) := \varnothing.$$

and at time $t = 1$

$$\mathsf{D}(1) := \{v\}, \qquad \mathsf{A}(1) := \bigcup_{i\in[d]} \{w : w \in \mathscr{C}_i(v)\} \setminus \{v\}, \qquad \mathsf{P}(1) := \big\{\{w,v\} \in \mathcal{E} : w \in \mathsf{A}(1)\big\}.$$

▷ ***Time*** $t \geq 2$***.*** *Choose a vertex* v^t *according to an arbitrary but predetermined rule from* $\mathsf{A}(t-1)$*. Let* u *be the vertex such that* $\{u, v^t\} \in \mathsf{P}(t-1)$ *and write* j_t *for the unique direction such that* $v^t_{j_t} \neq u_{j_t}$*. Update*

$$\begin{aligned}
\mathsf{D}(t) &:= \mathsf{D}(t-1) \cup \{v^t\},\\
\mathsf{A}(t) &:= \Big(\mathsf{A}(t-1) \cup \bigcup_{i\in[d]\colon i\neq j_t} \big\{w : w \in \mathscr{C}_i(v^t)\big\}\Big) \setminus \{v^t\},\\
\mathsf{P}(t) &:= \mathsf{P}(t-1) \cup \big\{\{w, v^t\} \in \mathcal{E} : w \in \mathsf{A}(t) \setminus \mathsf{A}(t-1)\big\}.
\end{aligned}$$

(Note that, by the definition of $\mathsf{P}(t)$*, the vertex* w *above is always unique.)*

▷ ***Stop.*** *Terminate the process when* $\mathsf{A}(t) = \varnothing$*. Set* $T = t$*.*

Now we are ready to complete the proof of the lower bound on $p_c^{(d)}(\theta)$ in Theorem 1.1:

Proof of the lower bound in Theorem 1.1. Note that $\mathsf{D}(T) = |\mathscr{C}(v)|$. Moreover, the line-wise exploration process can be naturally coupled to the (modified) Galton-Watson process where the offspring distribution is given by the law of the sum of the sizes of $d-1$ independent ERRG clusters minus one (where the $d-1$ accounts for the fact that the line we are in has already been explored, and the minus one accounts for the fact that the vertex v^t has already been counted). This Galton-Watson process has a modification at the root, where we consider

d independent ERRG cluster sizes, to account for the fact that in the first step we have not yet explored any lines. Let Z_p denote the total progeny of this GW-process. A standard argument tells us that $|\mathscr{C}(v)|$ is stochastically dominated by Z_p, because Z_p "ignores" the loops of $|\mathscr{C}(v)|$ that do not occur within a line. The offspring distribution of this GW-process has mean $\mu = (d-1)(\chi_{\text{line}}(p)-1)$ (except at the root, where it has mean $\mu_\rho = d(\chi_{\text{line}}(p)-1)$), so

$$\mathbb{E}_p[|\mathscr{C}(v)|] \leq \mathbb{E}[Z_p] = 1 + \frac{\mu_\rho}{1-\mu} = \chi_{\text{line}}(p)\,\frac{1}{1-(d-1)(\chi_{\text{line}}(p)-1)}. \tag{3.3}$$

Consequently, the value p_ℓ that solves the equation

$$\theta V^{1/3} = \chi_{\text{line}}(p)\,\frac{1}{1-(d-1)(\chi_{\text{line}}(p_\ell)-1)} \tag{3.4}$$

is a lower bound on $p_c^{(d)}(\theta)$.

We insert (3.2) in (3.4), to obtain

$$\theta V^{1/3} = \chi_{\text{line}}(p)\,\frac{1}{1-(d-1)\Big(\dfrac{1}{1-p_\ell(n-1)}\Big(1-\dfrac{2d^2-1}{2(d-1)^3}m^{-1}+O(m^{-2})\Big)-1\Big)}. \tag{3.5}$$

Solving this with respect to p_ℓ, we find

$$p_\ell = m^{-1} + \frac{2d^2-1}{2(d-1)^2}m^{-2} + O(m^{-1}V^{-1/3} + m^{-3}), \tag{3.6}$$

which gives us the desired lower bound on $p_c^{(d)}(\theta)$. □

We have found a lower bound on $p_c^{(d)}(\theta)$ by rather explicitly using the product structure of the Hamming graph to find a good branching process domination. To find an upper bound on $p_c^{(d)}(\theta)$ with the same method would be much more involved, since then we would need to thin the GW tree further to take into account the loops outside lines as well. Instead, we apply the lace expansion, which is a less direct, but much more robust method.

4. Bounds on connection probabilities

In Section 5 we show how the lace expansion can be used to express an upper bound on $p_c^{(d)}(\theta)$ in terms of products and sums of connection probabilities $\mathbb{P}_p(x \leftrightarrow y)$. In this section, we collect some preliminaries that will be used throughout the analysis in Section 5. This section is organized as follows. In Section 4.1 we derive estimates on connection probabilities. In Section 4.2 we estimate bubble, triangle and polygon diagrams.

4.1. Connection probabilities.

Given an event A for percolation on $H(d,n)$, we define the event $\{A \text{ within one line}\}$ to be the subset of all configurations $\omega \in A$ such that it can be verified that $\omega \in A$ by fixing a line of $H(d,n)$ and then only inspecting the status of the edges of ω within that line. Likewise, we define $\{A \text{ through multiple lines}\} := A \setminus \{A \text{ within one line}\}$. We define $\{v \overset{\leq r}{\longleftrightarrow} w\}$ to be the event that $w \in \mathscr{C}(v)$ and there exists a path of length at most r. We further write $\{v \overset{>r}{\longleftrightarrow} w\}$ (respectively, $\{v \overset{=r}{\longleftrightarrow} w\}$) for the event that $w \in \mathscr{C}(v)$ and there is a (not necessarily shortest) simple path of open edges from v to w containing more than (respectively, exactly) r edges. Not requiring minimality deviates a bit from the common use of this notation, but for all the purposes of the present paper it makes

little difference and often simplifies proofs. This is because the event $\{v \xleftrightarrow{=r} w\}$ is increasing, whereas the event that the graph distance equals r is not.

We start with a proposition about the probability of two points being connected by a path that is longer than the mixing time of $H(d,n)$. Given a graph G, we define the t-step *non-backtracking random walk* (NBW) on G started at x as the uniform measure on paths $(X_1,\dots,X_t)$ such that $X_1 = x$ and $X_i \neq X_{i-2}$ for all $i \in \{3,\dots,t\}$ (i.e., the path never backtracks). For two vertices x, y of G, we write $\mathbf{p}^t_{\mathrm{NBW}}(x,y)$ for the probability that a t-step non-backtracking random walk started at x ends at y. Given a connected aperiodic graph G and $\alpha \in (0,1)$, we define the *uniform non-backtracking mixing time* as

$$t_{\mathrm{mix}}(G;\alpha) := \min\Big\{t\colon \max_{x,y} \mathbf{p}^t_{\mathrm{NBW}}(x,y) \le (1+\alpha)V^{-1}\Big\}. \tag{4.1}$$

In the remainder of this paper we will use

$$t_{\mathrm{mix}} := t_{\mathrm{mix}}(H(d,n); n^{-1}) = O(\log n), \tag{4.2}$$

so $\alpha = n^{-1}$. The above bound is proved by Fitzner and van der Hofstad [18]. For this choice of α, the following proposition is a direct consequence of [26, Theorem 1.4 and Lemma 3.14]:

Proposition 4.1 (A uniform connection bound)**.** *Consider percolation on $H(d,n)$ with $d \ge 2$ and $p \le p_c^{(d)}$. Then*

$$\mathbb{P}_p\big(x \xleftrightarrow{>t_{\mathrm{mix}}} y\big) \le \frac{\chi(p)}{V}(1+O(m^{-1})). \tag{4.3}$$

Heuristically this proposition can be explained with the idea that percolation paths in sufficiently high-dimensional graphs at criticality look like random walk paths, so if the path is longer than the mixing time, then the connection probabilities become uniform over the graph. For Hamming graphs, there is little difference between non-backtracking walk and simple random walk, so in many of our bounds we use simple random walk instead.

We proceed with a useful bound on the two-point function:

Proposition 4.2 (Asymptotics for the two-point function on the Hamming graph)**.** *Consider percolation on $H(d,n)$ with $d \ge 2$ and $p = m^{-1}(1+O(m^{-1}))$ and $p \le p_c^{(d)}$. For any $v, w \in \mathcal{V}$,*

$$\tau_p(v-w) = \delta_{v,w} + \frac{d}{d-1}\frac{1}{m}\mathbb{1}_{\{v,w\}\in\mathcal{E}} + \frac{\chi(p)}{V}(1+O(m^{-1})) + O(m^{-(d(v,w)\vee 2)}), \tag{4.4}$$

where $d(v,w)$ is the graph distance between v and w on $H(d,n)$.

Proof. If $v = w$ then $\tau_p(v-w) = \mathbb{P}_p(v \leftrightarrow w) = 1$ by definition. This gives rise to the factor $\delta_{v,w}$ above. Let us henceforth assume that $v \neq w$. We divide the event $\{v \leftrightarrow w\}$ into three disjoint events as follows:

- $A = \big\{v \xleftrightarrow{>t_{\mathrm{mix}}} w\big\}$,
- $B = \big\{v \xleftrightarrow{\le t_{\mathrm{mix}}} w$ through different lines$\big\}$,
- $C = \big\{v \xleftrightarrow{\le t_{\mathrm{mix}}} w$ within one line$\big\}$.

We bound their contributions separately.

By Proposition 4.1 we have

$$\mathbb{P}_p(A) = \frac{\chi(p)}{V}(1+O(m^{-1})). \tag{4.5}$$

For the bound on $\mathbb{P}_p(B)$ we distinguish three different cases: $d(v,w)=1$, $1<d(v,w)<d$, $d(v,w)=d$.

Case $1<d(v,w)<d$**:** Write $\mathbf{p}^t(v,w)$ for the probability that a simple random walk started at v is at w after t steps, and let $\mathcal{P}_k(v,w)$ denote the set of all simple paths of length k from v to w in $H(d,n)$. There are at most $m^k\mathbf{p}^k(u,v)$ such paths, and so

$$\begin{aligned}\mathbb{P}_p(B) &\leq \sum_{k=d(v,w)}^{t_{\text{mix}}} \sum_{\eta\in\mathcal{P}_k(v,w)} \mathbb{P}_p(\eta \text{ is open}) \leq \sum_{k=d(v,w)}^{t_{\text{mix}}} p^k m^k \mathbf{p}^k(v,w)\\ &= (1+o(1)) \sum_{k=d(v,w)}^{t_{\text{mix}}} \mathbf{p}^k(v,w),\end{aligned} \tag{4.6}$$

where for the last bound we use that $(mp)^k = 1+o(1)$ for all $k\leq t_{\text{mix}}$ by (1.8) and (4.2). Define the set

$$\mathcal{H}(v,w) := \big\{u\in\mathcal{V} : u_i = v_i \text{ for all } i\in\{1,\dots,d\} \text{ such that } v_i = w_i\big\}. \tag{4.7}$$

We can view $\mathcal{H}(v,w)$ as the "lowest-dimensional hyperplane" that contains both v and w. We write

$$\mathbf{p}^k(v,w) =: \mathbf{p}^k_{\mathcal{H}(v,w)}(v,w) + \mathbf{p}^k_{\neg\mathcal{H}(v,w)}(v,w), \tag{4.8}$$

where $\mathbf{p}^k_{\mathcal{H}(v,w)}(v,w)$ is the probability to go from v to w in k steps without leaving $\mathcal{H}(v,w)$.

If the walker started from v is to reach w, then it will need to take a step in each direction such that $v_i\neq w_i$. At each step of a walk on the Hamming graph the probability that the walk stays in $\mathcal{H}(v,w)$ and gets closer to $w=(w_1,\dots,w_d)$ is at most $\frac{d(v,w)}{m}$, since it has to move in one of at most $d(v,w)$ directions, say direction j, exactly to the unique neighbour that has jth coordinate w_j. There are at most $k^{d(v,w)}$ orders in which the distance-decreasing steps can occur among k steps. Therefore

$$\begin{aligned}\sum_{k=d(v,w)}^{t_{\text{mix}}} \mathbf{p}^k_{\mathcal{H}(v,w)}(v,w) &\leq \sum_{k=d(v,w)}^{t_{\text{mix}}} k^{d(v,w)} \left(\frac{d(v,w)}{m}\right)^{d(v,w)} \left(\frac{d(v,w)}{d}\right)^{k-d(v,w)}\\ &\leq \left(\frac{d}{m}\right)^{d(v,w)} \sum_{k=d(v,w)}^{t_{\text{mix}}} k^{d(v,w)} \left(\frac{d(v,w)}{d}\right)^{k} = O(m^{-d(v,w)}).\end{aligned} \tag{4.9}$$

If the path from v to w leaves $\mathcal{H}(v,w)$, then it will have to take at least $d(v,w)+1$ steps in the direction of w. Since $k\leq t_{\text{mix}}$, there are at most t_{mix} places along the walk where these steps can occur, so we can bound

$$\begin{aligned}\sum_{k=d(v,w)}^{t_{\text{mix}}} \mathbf{p}^k_{\neg\mathcal{H}(v,w)}(v,w) &\leq t_{\text{mix}}\cdot {t_{\text{mix}}}^{d(v,w)+1} \left(\frac{d}{m}\right)^{d(v,w)+1}\\ &= O\big(m^{-d(v,w)-1} {t_{\text{mix}}}^{d(v,w)+2}\big).\end{aligned} \tag{4.10}$$

Case $d(v,w)=d$: The walk needs to take at least d steps in the direction of w. Hence, using the same argument as for (4.10), we obtain

$$\sum_{k=d}^{t_{\mathrm{mix}}}(mp)^k\mathbf{p}^k(v,w)=O(m^{-d}{t_{\mathrm{mix}}}^{d+1}). \tag{4.11}$$

Case $d(v,w)=1$: Given a random walk starting from v, we write Z for the (random) first vertex on the walk such that $d(Z,w)=2$ and T for the number of steps it took to go from v to Z. A standard path-counting argument gives an upper bound on $\mathbb{P}_p(B)$ for the case $d(v,w)=1$ as follows:

$$\mathbb{P}_p(B)\le\sum_{t=1}^{t_{\mathrm{mix}}}\sum_{z\in\mathcal{V}}\mathbb{P}_p(Z=z,T=t)\sum_{k=1}^{t_{\mathrm{mix}}-t}(mp)^{k+t}\mathbf{p}^k(z,w). \tag{4.12}$$

By definition, the vertex Z is unique and $\mathbb{P}_p(Z=z)\neq 0$ only if $d(z,w)=2$, and, by the transitivity of $H(d,n)$, for every k we know that $\mathbf{p}^k(z,w)=\mathbf{p}^k(y,w)$ if $d(z,w)=d(y,w)$. Thus

$$\mathbb{P}_p(B)\le\sum_{t=1}^{t_{\mathrm{mix}}}\mathbb{P}_p(T=t)\sum_{k=1}^{t_{\mathrm{mix}}}(mp)^{t_{\mathrm{mix}}}\mathbf{p}^k(z,w)=O(m^{-(2\vee d(v,w))}+m^{-d}{t_{\mathrm{mix}}}^{d+1}), \tag{4.13}$$

where we have used the assumption that $p\le m^{-1}(1+O(m^{-1}))$ and (4.2) for the second bound. We further assumed $p\ge m^{-1}$ for the first bound. Note that this can be done without loss of generality.

It remains to bound $\mathbb{P}_p(C)$. We consider the cases $d(v,w)=1$ and $d(v,w)\ge 2$. By definition, $\mathbb{P}_p(C)=0$ if $d(v,w)\ge 2$, which takes care of the latter case. If $d(v,w)=1$, then the probability to connect within that line is simply the two-point function of an ERRG with n vertices and edge probability $p=\frac{1+O(m^{-1})}{d(n-1)}$, which is

$$\frac{\chi_G(p)-1}{n-1}=\frac{1+O(m^{-1})}{(d-1)(n-1)}, \tag{4.14}$$

for every pair of distinct vertices, due to the symmetry of K_n. We therefore obtain

$$\mathbb{P}_p(C)=\frac{d}{d-1}\frac{1+O(m^{-1})}{m}\mathbb{1}_{\{v,w\}\in\mathcal{E}}. \tag{4.15}$$

Adding $\delta_{v,w}$ and the three bounds in (4.5), (4.13) and (4.15) for $\mathbb{P}_p(A)$, $\mathbb{P}_p(B)$ and $\mathbb{P}_p(C)$, respectively, completes the proof. □

4.2. **Bubble, triangle and polygon diagrams.** The final estimates of this section involve the so-called bubble, triangle, and polygon diagrams. As was already alluded to in the introduction, these diagrams, the triangle diagram in particular, are very important quantities in the study of high-dimensional percolation. We start with their definition.

Given an integer $i\ge 2$ and vertices $v,x_1,\dots,x_{i-1},w\in\mathcal{V}$, we define the *$i$-gon diagrams*

$$C_i^{(0)}(v,x_1,\dots,x_{i-1},w):=\mathbb{P}_p(v\leftrightarrow x_1)\cdots\mathbb{P}_p(x_{i-1}\leftrightarrow w), \tag{4.16}$$

$$C_i^{(1)}(u,x_1,\dots,x_{i-1},w):=\sum_{v\colon\{u,v\}\in\mathcal{E}}pC_i^{(0)}(v,x_1,\dots,x_{i-1},w), \tag{4.17}$$

$$C_i^{\le k}(v,x_1,\dots,x_{i-1},w) := \sum_{k_1+\cdots+k_i\le k} \mathbb{P}_p(v \overset{=k_1}{\longleftrightarrow} x_1) \times \cdots \times \mathbb{P}_p(x_{i-1} \overset{=k_i}{\longleftrightarrow} w), \tag{4.18}$$

$$C_i^{>k}(v,x_1,\dots,x_{i-1},w) := \sum_{k_1+\cdots+k_i>k} \mathbb{P}_p(v \overset{=k_1}{\longleftrightarrow} x_1) \times \cdots \times \mathbb{P}_p(x_{i-1} \overset{=k_i}{\longleftrightarrow} w), \tag{4.19}$$

where in the case $i=2$ we mean $C_2^{(0)}(v,x_1,y)$, etc. Recall (1.6) and observe that

$$\sum_{x,y\in\mathcal{V}} C_3^{(0)}(v,x,y,w) = \nabla_p(v,w) \le \delta_{v,w} + 10\,\frac{\chi(p)^3}{V} + O(m^{-1}), \tag{4.20}$$

and recall that by the definition of $p_c(\theta)$ in (1.7), we have $\chi(p_c(\theta))^3 = \theta^3 V$. Recall, moreover, the following useful bound by Borgs et al. in [10, Proposition 1.2 and (5.106)] (adapted to our setting): For all $v,w\in\mathcal{V}$, $p = m^{-1}(1+O(m^{-1}+V^{-1/3}))$, and all $p \le p_c^{(d)}(\theta)$ with θ sufficiently small such that $\theta^3 \le \beta_0$ for β_0 such that (1.6) holds,

$$\sum_{x_1,x_2} C_3^{(1)}(v,x_1,x_2,w) = 3\,\frac{\chi(p)^3}{V} + O(m^{-1}). \tag{4.21}$$

We proceed with a bound on "short" polygons:

Lemma 4.3. *Consider percolation on $H(d,n)$ with $p = m^{-1}(1+O(m^{-1}))$. Then for all $v,w\in\mathcal{V}$ and each integer $i\ge 2$,*

$$\sum_{x_1,\dots,x_{i-1}\in\mathcal{V}} C_i^{\le t_{\mathrm{mix}}}(v,x_1,\dots,x_{i-1},w) = \delta_{v,w} + O(m^{-d(v,w)\vee 1} + m^{-d} {t_{\mathrm{mix}}}^{d+i}). \tag{4.22}$$

Proof. We use the same path counting argument as in the proof of Proposition 4.2. Instead of fixing $x_1,\dots,x_{i-1}$ and summing over all possible probabilities, we fix a random walk path of length $k\le t_{\mathrm{mix}}$ between v and w and count the possible ways in which that path could appear in the above sum. There are at most $(k+1)^{i-1}$ possible ways to mark the path with the vertices $x_1,\dots,x_{i-1}$. We thus bound

$$\sum_{x_1,\dots,x_{i-1}\in\mathcal{V}} C_i^{\le t_{\mathrm{mix}}}(v,x_1,\dots,x_{i-1},w) \le \sum_{k=0}^{t_{\mathrm{mix}}} (k+1)^{i-1}(mp)^k \mathbf{p}^k(v,w). \tag{4.23}$$

Compare this bound with (4.6). The only difference is the factor $(k+1)^{i-1}$.

Case $d(v,w)<d$: Recall the definition of $\mathcal{H}(v,w)$ in (4.7). We again consider the contributions from walks that stay within $\mathcal{H}(v,w)$ and those that do not separately, starting with the contribution from the walks that remain within $\mathcal{H}(v,w)$.

If $v=w$ then $\mathcal{H}(v,w)=\{v\}$, so the only contribution comes from the trivial path $v=x_1=\cdots=x_{i-1}=w$, which gives the term $\delta_{v,w}$. If $1<d(v,w)<d$, then to go from v to w the walk needs to take at least one step in each direction j such that $v_j\neq w_j$ and move exactly to a neighbour with jth coordinate w_j. Using a similar argument as in (4.9), we obtain

$$\sum_{k=d(v,w)}^{t_{\mathrm{mix}}} (k+1)^{i-1}(mp)^k \mathbf{p}^k_{\mathcal{H}(v,w)}(v,w) = O(m^{-d(v,w)}). \tag{4.24}$$

Note, in particular, that the extra factor $(k+1)^{i-1}$ compared to (4.9) does not affect the convergence of the sum over k: it only changes the constant in the $O(m^{-d(v,w)})$ term.

Next, consider the contribution of walks that leave $\mathcal{H}(v,w)$. If the walk leaves $\mathcal{H}(v,w)$ along the path from v to w, then it needs to take at least $d(v,w)+1$ distance-decreasing steps. By the same argument as (4.10), we thus bound

$$\begin{aligned}\sum_{k=d(v,w)}^{t_{\mathrm{mix}}} (k+1)^{i-1}(mp)^k \mathbf{p}^k_{\neg\mathcal{H}(v,w)}(v,w) &\le (1+o(1)){t_{\mathrm{mix}}}^{d(v,w)+2+i-1}\left(\frac{d}{m}\right)^{-d(v,w)-1} \\ &= O({t_{\mathrm{mix}}}^{d(v,w)+i+1}m^{-d(v,w)-1}) = O(m^{-d(v,w)}).\end{aligned} \tag{4.25}$$

Case $d(v,w)=d$: The walk needs to take at least d steps towards w to reach it. By the same argument as for the previous bound, we have

$$\sum_{k=d(v,w)}^{t_{\mathrm{mix}}} (k+1)^{i-1}(mp)^k \mathbf{p}^k(v,w) = O(m^{-d}{t_{\mathrm{mix}}}^{d+i}). \tag{4.26}$$

Summing these bounds we get the claim. □

The next lemma combines the above estimates in a convenient form (and is especially useful when $i=2$).

Lemma 4.4. *Consider percolation on $H(d,n)$ with $p \le p_c^{(d)}(\theta)$. For all $i \ge 2$, $j \in \{0,1\}$, and $v,w \in \mathcal{V}$,*

$$\begin{aligned}\sum_{x_1,\dots,x_{i-1}\in\mathcal{V}} C_i^{(j)}(v,x_1,\dots,x_{i-1},w) &= \delta_{v,w}\delta_{j,0} + \theta^i V^{i/3-1}(1+O(m^{-1})) \\ &\quad + O(m^{-d(v,w)\vee 1} + m^{-d}{t_{\mathrm{mix}}}^{d+i}).\end{aligned} \tag{4.27}$$

Proof. For $j=0$, the lemma follows after combining Proposition 4.1 and Lemma 4.3 with (1.7). For $j=1$ we further observe that

$$\sum_{u:\{v,u\}\in\mathcal{E}} p = mp = 1 + O(m^{-1}). \tag{4.28}$$

Moreover, the factor $\delta_{v,w}$ does not arise, because this is due to the trivial path $v = x_1 = \cdots = x_{i-1} = w$, which by Proposition 4.2 now contributes

$$C_i^{(1)}(v,v,\dots,v,v) = p\sum_{u:\{v,u\}\in\mathcal{E}} C_3^{(0)}(u,v,\dots,v,v) = p\sum_{u:\{u,v\}\in\mathcal{E}} \tau_p(v-u) = O(m^{-1}). \tag{4.29}$$

□

Lastly, we derive an improved bound on the triangle diagram in the case where the two intermediate points of the triangle are neighbours in $H(d,n)$ and in the case where one intermediate point of the triangle is constrained to be a neighbour of a fixed auxiliary point:

Lemma 4.5. *Consider percolation on $H(d,n)$ with $p \le p_c^{(d)}(\theta)$. For all $v,w \in \mathcal{V}$ and $j \in \{0,1\}$,*

$$\sum_{x,y:\{x,y\}\in\mathcal{E}} C_3^{(j)}(v,x,y,w) = O(m^{-1} + V^{-1/3}) \tag{4.30}$$

and

$$\sup_z \sum_{x,y:\{y,z\}\in\mathcal{E}} C_3^{(j)}(v,x,y,w) = O(m^{-1} + V^{-1/3}). \tag{4.31}$$

Proof. We prove (4.30) for the case $j = 0$. The proof of (4.30) for the case where $j = 1$ and of (4.31) are almost identical, so we leave them to the reader.

Consider first the contribution due the cases where all the connections are due to short paths (i.e., shorter than t_{mix} in total). By Lemma 4.3, this is bounded by $O(m^{-1})$, even without using the constraint that $\{x, y\} \in \mathcal{E}$. Next, consider the contribution to the left-hand side of (4.30) from the case where the path from y to w is longer than t_{mix}, i.e.,

$$\sum_x \mathbb{P}_p(v \leftrightarrow x) \sum_{y:\{x,y\}\in\mathcal{E}} \mathbb{P}_p(x \leftrightarrow y)\, \mathbb{P}_p(y \xleftrightarrow{>t_{\text{mix}}} w). \tag{4.32}$$

Applying Proposition 4.1 to the last term, Proposition 4.2 to the middle term, and summing over x and y, we obtain the upper bound

$$\chi(p)\, mO(m^{-1})\, O\Big(\frac{\chi(p)}{V}\Big) = O(V^{-1/3}), \tag{4.33}$$

where the final bound is due to (1.7). The contribution due to the case where the path from v to x is longer than t_{mix} is the same by symmetry.

To bound the contributions due to the case where the path from x to y is longer than t_{mix}, we consider the cases $y = w$ and $y \neq w$ separately. The contribution due to a long path between x and y and $y = w$ is given by

$$\sum_x \mathbb{P}_p(v \leftrightarrow x)\, \mathbb{P}_p(x \xleftrightarrow{>t_{\text{mix}}} w) = \chi(p)\, O\Big(\frac{\chi(p)}{V}\Big) = O(V^{-1/3}), \tag{4.34}$$

where the bound follows from Proposition 4.1 and (1.7). The contribution due to a long path between x and y and $y \neq w$ is given by

$$\sum_x \mathbb{P}_p(v \leftrightarrow x) \sum_{\substack{y:\{x,y\}\in\mathcal{E},\\ y\neq w}} \mathbb{P}_p(x \xleftrightarrow{>t_{\text{mix}}} y)\, \mathbb{P}_p(y \leftrightarrow w) = \chi(p)\, mO\Big(\frac{\chi(p)}{V}\Big)\, O(m^{-1}) = O(V^{-1/3}), \tag{4.35}$$

where the bound again follows from Propositions 4.1 and 4.2, and (1.7).

Adding the bound $O(m^{-1})$ due to short paths and the bounds (4.33)–(4.35), we complete the proof of (4.30) for the case $j = 0$. □

5. The upper bound on $p_c(\theta)$ via the lace expansion

In Section 5.1 we recall the background of the lace-expansion technique and state a proposition in which we estimate lace-expansion coefficients. In Section 5.2 we use this proposition to prove the upper bound on $p_c(\theta)$. The proof of the proposition is given in Sections 5.3–5.5.

5.1. Background.

The lace expansion is a method originated by Brydges and Spencer to study self-avoiding walk [12], and was first applied to percolation by Hara and Slade [21]. Hara and Slade's method gives an expansion for the percolation two-point function $\tau_p(x)$. The version of the lace expansion that we use here was derived by Borgs et al. [10], where it is proved for the Hamming graph, among others, that for any $p \in [0, 1]$,

$$\tau_p(x) = \delta_{0,x} + m(J_p * \tau_p)(x) + m(\Pi_p * J_p * \tau_p)(x) + \Pi_p(x), \tag{5.1}$$

where $(f * g)(x) = \sum_{y \in \mathcal{V}} f(y)g(x-y)$ denotes the convolution between f and g, $J_p(x-y) := \mathbb{P}_p(\{x,y\} \text{ is open})$, and $\Pi_p(x)$ is the so-called *irreducible two-point function* or *lace-expansion coefficient*. The lace expansion further determines that $\Pi_p(x)$ is given by the alternating series

$$\Pi_p(x) = \sum_{N=0}^{\infty} (-1)^N \Pi_p^{(N)}(x), \tag{5.2}$$

and the *lace-expansion coefficients* $\Pi_p^{(N)}$ have a well-defined structure that will play an important role in the determination of the upper bound we derive here. Define the *discrete Fourier transform* of a function $f : \mathcal{V} \to \mathbb{R}$ as $\hat{f}(k) = \sum_{x \in \mathcal{V}} \mathrm{e}^{ik \cdot x} f(x)$ for $k \in \mathbb{R}^d$. Taking the Fourier transform of (5.1), we obtain

$$\hat{\tau}_p(k) = \frac{1 + \hat{\Pi}_p(k)}{1 - m\hat{J}_p(k)(1 + \hat{\Pi}_p(k))}. \tag{5.3}$$

Observe that $\hat{\tau}_p(0) = \sum_x \tau_p(x) = \chi(p)$ and $\hat{J}_p(0) = mp$, so setting $k \equiv 0$ and $p \equiv p_c^{(d)}(\theta)$, and applying (1.7), we obtain

$$mp_c(\theta) = \frac{1}{1 + \hat{\Pi}_{p_c}(0)} + \theta^{-1} V^{-1/3}. \tag{5.4}$$

(We will henceforth only consider $\hat{\Pi}_p(k)$ at $k = 0$ and therefore will not write the argument anymore.) Combining (5.2) and (5.4) with the following proposition allows us to determine the upper bound on $p_c^{(d)}(\theta)$ for θ sufficiently small:

Proposition 5.1 (Bounds on the lace-expansion coefficients)**.** *Consider percolation on $H(d,n)$ with $d \geq 3$. Let θ be such that $3\theta^3(1 + 10\,\theta^3) < 1$ and $\theta^3 \leq \beta_0$ for β_0 such that (1.6) holds, and let p be such that $p = m^{-1}(1 + O(m^{-1} + V^{-1/3}))$ and $p \leq p_c^{(d)}(\theta)$. Then*

$$\hat{\Pi}_p^{(0)} \geq \frac{2d-1}{2(d-1)^2} m^{-1} + O(m^{-2} + V^{-1/3}), \tag{5.5}$$

$$\hat{\Pi}_p^{(1)} \leq \frac{d^2 + d - 1}{(d-1)^2} m^{-1} + O(m^{-2} + V^{-1/3}), \tag{5.6}$$

$$\sum_{N=2}^{\infty} \hat{\Pi}_p^{(N)} = O(m^{-2} + V^{-1/3}). \tag{5.7}$$

Note that we do not consider $d = 2$ in the above proposition. This is of no consequence, as our main theorem is already proved for $d = 2, 3$ in [9]. The proofs below do apply to the case $d = 2$, but the bounds become slightly less sharp (because Lemma 4.4 is less sharp when $d = 2$.)

5.2. **Proof of the upper bound on the critical point.** Before starting with the proof of Proposition 5.1, which constitutes the bulk of what remains of this paper, let us complete the proof of the upper bound in Theorem 1.1:

Proof of the upper bound in Theorem 1.1 subject to Proposition 5.1. As was mentioned above, Theorem 1.1 is already proved for $d = 2, 3$ in [9], so let $d \geq 4$. Suppose first that θ is such that

$3\theta^3(1+10\,\theta^3)<1$ and $\theta^3\le\beta_0$ for β_0 such that (1.6) holds. Using (5.2) and Proposition 5.1, we bound

$$\begin{aligned}\hat\Pi_{p_c^{(d)}} &= \hat\Pi^{(0)}_{p_c^{(d)}} - \hat\Pi^{(1)}_{p_c^{(d)}} + O(m^{-2}+V^{-1/3})\\ &\ge \frac{2d-1}{2(d-1)^2}m^{-1} - \frac{d^2+d-1}{(d-1)^2}m^{-1} + O(m^{-2}+V^{-1/3})\\ &\ge -\frac{2d^2-1}{2(d-1)^2}m^{-1} + O(m^{-2}+V^{-1/3}).\end{aligned} \tag{5.8}$$

Applying this bound to (5.4), we get

$$\begin{aligned}p_c^{(d)}(\theta) &\le \frac{m^{-1}}{1-\frac{2d^2-1}{2(d-1)^2}m^{-1}+O(m^{-2}+V^{-1/3})} + \theta^{-1}m^{-1}V^{-1/3}\\ &= m^{-1} + \frac{2d^2-1}{2(d-1)^2}m^{-2} + O(m^{-1}V^{-1/3}+m^{-3}),\end{aligned} \tag{5.9}$$

which gives the upper bound for θ such that $3\theta^3(1+10\,\theta^3)<1$ and $\theta^3\le\beta_0$.

We have determined $p_c(\theta)$ for some $\theta\in(0,\infty)$. Borgs et al. [9, Theorem 1.1] show that $p_c(\theta')=p_c(\theta)(1+O(V^{-1/3}))$ for every $\theta'\in(0,\infty)$, and so we obtain the desired upper bound for arbitrary θ, which completes the proof of the upper bound. □

The remainder of this section is devoted to the proof of Proposition 5.1, which is divided into three further subsections.

5.3. **Analysis of $\hat\Pi_p^{(0)}$: proof of (5.5).** The lace-expansion method yields (see [10, Section 3.2])

$$\hat\Pi_p^{(0)} := \sum_{x\neq 0}\mathbb{P}_p(0\Longleftrightarrow x), \tag{5.10}$$

where $v\Longleftrightarrow w$ denotes the event that there exist two edge-disjoint paths between v and w. In this case we say that v and w are *doubly connected.* This is equivalent to the event that there exists an edge-disjoint cycle of open edges containing both v and w.

Given a graph G and a vertex x in G, write $\mathcal{L}_k(G;x)$ for the set of vertex-disjoint cycles of length k in G that contain x. Note that $|\mathcal{L}_k(K_n;0)|=\frac{(n-1)!}{2(n-k)!}$, because we must choose the $k-1$ other vertices for the cycle from $n-1$ possible choices, and although their order matters, the direction of the cycle does not, which explains the factor $\frac12$. By only considering vertex-disjoint cycles in $\hat\Pi_p^{(0)}$ that are contained within a line that intersects 0, we can apply a similar argument to that of (2.31) to bound

$$\begin{aligned}\hat\Pi_p^{(0)} &\ge \sum_{x\neq0}\mathbb{P}_p(0\Longleftrightarrow x \text{ within one line})\\ &\ge d\sum_{k=3}^{n}(k-1)\sum_{L_k\in\mathcal{L}_k(K_n;0)}\Big[\mathbb{P}_p(L_k\in\mathcal{L}_k(\mathscr{C}(0);0))\\ &\qquad -\mathbb{P}_p\Big(\bigcup_{\{x,y\}\subset\mathcal{V}(L_k)}\{L_k\in\mathcal{L}_k(\mathscr{C}(0);0)\}\circ\{x\leftrightarrow y\}\Big)\Big]\end{aligned} \tag{5.11}$$

$$\geq d\sum_{k=3}^{n}(k-1)\frac{(n-1)!}{2(n-k)!}p^k\Big(1-\tfrac{1}{2}k(k-1)\mathbb{P}_p(1\in\mathscr{C}(0))\Big),$$

where in the second inequality $\mathscr{C}(0)$ is to be viewed as the cluster of vertex 0 in $G(n,p)$, and we use that every cycle of length k that passes through 0 passes through $(k-1)$ other vertices, and in the third inequality $\mathbb{P}_p(1\in\mathscr{C}(0))$ denotes the probability that vertices 0 and 1 of K_n are in the same cluster in $G(n,p)$. Use that $p=m^{-1}(1+O(m^{-1}+V^{-1/3}))$ by assumption, that $\frac{(n-1)!}{(n-k)!}=(n-1)^{k-1}-O(kn^{k-2})$, and that $\mathbb{P}_p(1\in\mathscr{C}(0))=\frac{d}{(d-1)n}(1+O(m^{-1}+V^{-1/3}))$ by Theorem 1.2, to bound (after a short computation)

$$\hat{\Pi}_p^{(0)}=\frac{2d-1}{2(d-1)^2}m^{-1}(1+O(m^{-1}+V^{-1/3})). \tag{5.12}$$

□

5.4. **Analysis of $\hat{\Pi}_p^{(1)}$: proof of (5.6).** We start with a few definitions:

Definition 5.2 (Connections through a subset and pivotal edges)**.**

- ▷ *We say that two vertices x,y are* connected through *a set $\mathcal{W}\subset\mathcal{V}$, and write $x\overset{\mathcal{W}}{\longleftrightarrow}y$, if $x\leftrightarrow y$ occurs and all the paths connecting the two vertices in the percolation configuration have at least one vertex in $\mathcal{W}$.*
- ▷ *Given a vertex x and an edge e, we define the set $\tilde{\mathscr{C}}^e(x)$ as the percolation cluster of x in the (possibly modified) percolation configuration where the edge e is set to closed.*
- ▷ *Given two vertices x,y and a directed edge (u,v), we say that (u,v) is* pivotal *for $x\leftrightarrow y$ if $x\leftrightarrow y$ occurs on the (possibly modified) configuration where $\{u,v\}$ is set to open, and $x\leftrightarrow u$ and $v\leftrightarrow y$ occur, but $x\leftrightarrow y$ does not occur on the (possibly modified) configuration where $\{u,v\}$ is set to closed. Note that the direction of the edge is important and that $\{(u,v)$ is pivotal for $x\leftrightarrow y\}\neq\{(v,u)$ is pivotal for $x\leftrightarrow y\}$.*

We define the event

$$E'(v,x;\mathcal{W}):=\{v\overset{\mathcal{W}}{\longleftrightarrow}x\}\cap\{\nexists\text{ pivotal }(u',v')\text{ for }v\leftrightarrow x\ \text{ such that }v\overset{\mathcal{W}}{\longleftrightarrow}u'\}. \tag{5.13}$$

This event is a central object of the percolation lace expansion (see e.g. [21, (1.36)]). We can now give the definition of $\hat{\Pi}_p^{(1)}$ from [10]:

$$\hat{\Pi}_p^{(1)}:=\sum_x p\sum_{u,v:\{u,v\}\in\mathcal{E}}\mathbb{E}_0[\mathbb{1}_{\{0\Leftrightarrow u\}}\mathbb{P}_1(E'(v,x;\tilde{\mathscr{C}}_0^{(u,v)}(0))]. \tag{5.14}$$

The subscripts 0 and 1 indicate that the events happen on two *distinct* percolation configurations on the Hamming graph, and that these percolation configurations are "nested" in such a way that $\tilde{\mathscr{C}}_0^{(u,v)}$ is a set-valued random variable with respect to $\mathbb{E}_0$, but for any fixed realization of $\tilde{\mathscr{C}}_0^{(u,v)}$ it is viewed as a deterministic set with respect to $\mathbb{P}_1$ (see [10] for a more in-depth discussion of this construction). Our proof strategy will be to first identify and bound the main contributions to $\hat{\Pi}_p^{(1)}$. For this we can apply the BK-inequality to all the events in the formulation of $\hat{\Pi}_p^{(1)}$ and replace the complicated probabilities that appear there by products of two-point functions.

The analysis of $\hat{\Pi}_p^{(1)}$ proceeds by reduction of the complicated event to a collection of disjointly occurring two-point events, followed by a repeated application of the BK-inequality. Such bounds are standard, and are known in the literature as *diagrammatic bounds.*

Inspecting (5.13) and (5.14), we see that on the percolation measure $\mathbb{P}_1$ we need that v is connected to x through $\tilde{\mathscr{C}}_0^{(u,v)}(0)$, that any vertex $z \in \tilde{\mathscr{C}}_0^{(u,v)}(0)$ on the path from v to x must be doubly connected to x (otherwise there would exist a pivotal (u', v')), and that on a path from z to x there must be a vertex y that is connected to v. Moreover, on the percolation measure $\mathbb{P}_0$, the vertex z is connected to 0, and 0 is doubly connected to u, so there must exist exist a vertex t on a path from 0 to u such that t is connected to z. See Figure 1.

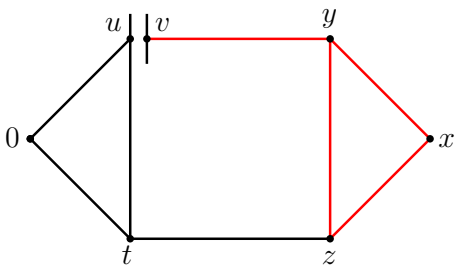

Figure 1. Diagrammatic description of the events contributing to $\hat{\Pi}_p^{(1)}$. Black lines refer to connections that occur in level 0 of percolation, red lines to connections that occur in level 1. The directed edge (u, v) is represented by two vertical dashes.

The main contribution will come from the simple case when $0 = u = t$ and $x = y = z$. To get a sharp bound on this term, it is necessary to use that (5.13) implies that the connection from 0 to x on the $\mathbb{P}_0$-percolation configuration is not allowed to use the edge $\{0, v\}$. We will therefore bound this contribution by

$$M := p \sum_x \sum_{v:\{0,v\}\in\mathcal{E}} \mathbb{P}_p(0 \leftrightarrow x \text{ without using } \{0,v\})\mathbb{P}_p(v \leftrightarrow x). \tag{5.15}$$

Following the same derivation as in [10, Section 4.1], but isolating the main contribution, we can use the BK-inequality to derive the following upper bound from (5.14):

$$\hat{\Pi}_p^{(1)} \le M + \sum_{u,t,z,y,x} C_3^{(0)}(0,u,t,0)C_3^{(1)}(u,y,z,t)C_2^{(0)}(y,x,z)(1-\delta_{0,u}\delta_{0,t}\delta_{x,y}\delta_{y,z}). \tag{5.16}$$

Our bound on (5.16) will rely heavily on the kind of path-counting methods that we applied in the previous sections.

We proceed by bounding the main term M. We claim that

$$\begin{aligned} M &\le p \sum_{v:\ \{0,v\}\in\mathcal{E}} \Big(\big(p + \mathbb{P}_p(v \overset{\ge 2}{\longleftrightarrow} 0)\big) + \mathbb{P}_p(0 \overset{\ge 2}{\longleftrightarrow} x) + \sum_{\substack{x: x\neq v,\\ x\neq 0}} \mathbb{P}(0 \leftrightarrow x)\mathbb{P}_p(v \leftrightarrow x) \Big) \\ &\le p \sum_{v:\ \{0,v\}\in\mathcal{E}} \Big(p + \sum_x C_2^{\ge 2}(0,x,v) \Big). \end{aligned} \tag{5.17}$$

The first term in the first inequality is due to the case $x = 0$, the second term to $x = v$, and the third term to the remaining cases. Note that if we had not restricted the connection from

0 to x to occur without using $\{0,v\}$, then the second term would have had an additional p, so the same argument would have given the upper bound $p\sum_v \big(2p+\sum_x C_2^{\geq 2}(0,x,v)\big)$, which, as it will turn out, is not sharp enough.

We use the (by now) familiar path-counting estimates to bound the right-hand side of (5.17) by

$$\begin{aligned} &p\sum_{v\colon \{0,v\}\in\mathcal{E}}\Big(p+\sum_x C_2^{\geq 2}(0,x,v)\Big) \qquad (5.18)\\ &\leq p^2m + p\sum_{v\colon \{0,v\}\in\mathcal{E}}\Big(\sum_{k=2}^{\infty}(k+1)\mathbf{p}^k_{\mathcal{H}(0,v)}(0,v)+\sum_{k=3}^{t_{\mathrm{mix}}}(k+1)\mathbf{p}^k_{\neg\mathcal{H}(0,v)}(0,v)+\sum_x C_2^{>t_{\mathrm{mix}}}(0,x,v)\Big)\\ &=: p^2m + M_1 + M_2 + M_3, \end{aligned}$$

where the factors $(k+1)$ in the second and third term on the right-hand side are due to interchanging the sum over k with the sum over x. The term M_1 is the contribution from walks that are constrained to remain within one line:

$$\begin{aligned} M_1 &\leq pm\sum_{k=2}^{\infty}(k+1)p^k(n-1)^{k-1}\\ &= p\sum_{k=2}^{\infty}(k+1)d^{-k+1}(1+O(m^{-1}+V^{-1/3}))\\ &= \frac{3d-2}{(d-1)^2}m^{-1}(1+O(m^{-1}+V^{-1/3})). \end{aligned} \qquad (5.19)$$

Using that $d(v,w)\geq 1$ and using an estimate similar to (4.12), we further bound

$$M_2 \leq p\sum_{v:\{0,v\}\in\mathcal{E}}\sum_{k=3}^{t_{\mathrm{mix}}}(mp)^k k(k+1)(n-1)^{-1} = O(m^{-2}). \qquad (5.20)$$

Next, we use Proposition 4.1 and (1.7) to bound

$$\begin{aligned} M_3 &= p\sum_{v:\{0,v\}\in\mathcal{E}}\sum_x C_2^{>t_{\mathrm{mix}}}(0,x,v) = pm\frac{\chi(p)^2}{V}(1+O(m^{-1}))\\ &\leq \theta^2 V^{-1/3}(1+O((m^{-1}+V^{-1/3})). \end{aligned} \qquad (5.21)$$

Inserting the bounds for M_1, M_2 and M_3 into (5.18), we conclude that

$$M \leq p^2m + \frac{3d-2}{(d-1)^2}m^{-1} + O(m^{-2}+V^{-1/3}) = \frac{d^2+d-1}{(d-1)^2}m^{-1} + O(m^{-2}+V^{-1/3}), \qquad (5.22)$$

and so it remains to show that the other terms in (5.16) are error terms.

We split the remaining contributions on the right-hand side of (5.16) according to the relative locations of y and z as follows: $y=z$, $d(y,z)=1$, $d(y,z)\geq 2$.

Case $d(y,z)\geq 2$**:** Apply Lemma 4.4 twice and (4.21) once to obtain

$$\sum_{t,u,y,z} C_3^{(0)}(0,u,t,0)C_3^{(1)}(u,y,z,t)\sup_{z'\colon d(y,z')\geq 2}\sum_x C_2^{(0)}(y,x,z') = O(m^{-2}+V^{-1/3}). \qquad (5.23)$$

Case $d(y,z)=1$: Apply Lemma 4.4 twice and Lemma 4.5 once to obtain

$$\sum_{t,u} C_3^{(0)}(0,u,t,0) \sup_{t'} \Big(\sum_{y,z:\{y,z\}\in\mathcal{E}} C_3^{(1)}(u,y,z,t') \Big) \sup_{z':d(y,z')=1} \Big(\sum_{x} C_2^{(0)}(y,x,z') \Big) = O(m^{-2}+V^{-1/3}). \tag{5.24}$$

Case $y=z$: Split again, according to whether $d(t,u)\geq 2$, $d(t,u)=1$, $t=u$. To bound the case $y=z$ and $d(t,u)\geq 2$, apply Lemma 4.4 three times to obtain

$$\sum_{t,u} C_3^{(0)}(0,u,t,0) \sup_{t':d(t',u)\geq 2} \Big(\sum_{y} C_2^{(1)}(u,y,t') \Big) \sum_{x} C_2^{(0)}(y,x,y) = O(m^{-2}+V^{-1/3}). \tag{5.25}$$

Similarly, to bound the case $y=z$ and $d(t,u)=1$, apply Lemmaa 4.4 twice and Lemma 4.5 once to obtain

$$\sum_{t,u:\{t,u\}\in\mathcal{E}} C_3^{(0)}(0,u,t,0) \sup_{t':d(t',u)=1} \Big(\sum_{y} C_2^{(1)}(u,y,t') \Big) \sum_{x} C_2^{(0)}(y,x,y) \;=\; O(m^{-2}+V^{-1/3}). \tag{5.26}$$

All that remains is the case $y=z$ and $t=u$. Recall that in M we have already accounted for the term where $t=u=0$ and $x=y=z$. Applying Proposition 4.2 with the constraint $t=u\neq 0$ to the first sum in (5.16) and Lemma 4.4 to the second and third sum, we obtain

$$\sum_{t\neq 0} C_2^{(0)}(0,t,0) \sum_{y} C_2^{(1)}(t,y,t) \sum_{x} C_2^{(0)}(y,x,y) = O(m^{-2}+V^{-1/3}). \tag{5.27}$$

The case $t=u=0$ and $y=z\neq x$ are analogous and give the same bound.

We have thus shown that all the remaining terms in (5.16) are of order $O(m^{-2}+V^{-1/3})$, which, combined with the bound (5.22) on the main term M, completes the proof of (5.6). □

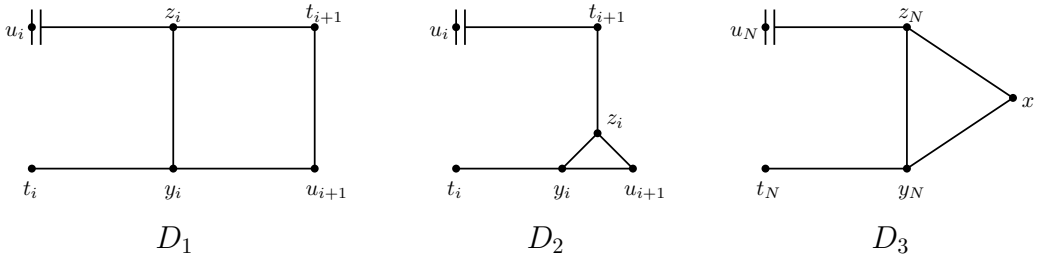

Figure 2. Diagrammatic descriptions of D_1, D_2, and D_3.

5.5. **Analysis of $\hat{\Pi}_p^{(N)}$ for $N\geq 2$: proof of (5.7).** To investigate $\hat{\Pi}_p^{(N)}$ for $N\geq 2$, we again bound the events in terms of products of two-point functions. We define the quantities

$$D_1(t_i,u_i,y_i,z_i,t_{i+1},u_{i+1}) := C_3^{(1)}(t_i,y_i,z_i,u_i) C_3^{(0)}(y_i,t_{i+1},u_{i+1},z_i), \tag{5.28}$$

$$D_2(t_i,u_i,y_i,z_i,t_{i+1},u_{i+1}) := C_5^{(1)}(t_i,y_i,u_{i+1},z_i,t_{i+1},u_i) \mathbb{P}_p(y_i \leftrightarrow z_i), \tag{5.29}$$

$$D_3(t_N,u_N,y_N,z_N,x) := C_3^{(1)}(t_N,y_N,z_N,u_N) C_2^{(0)}(y_N,x,z_N). \tag{5.30}$$

See Figure 2. We write the bounds on $\hat{\Pi}_p^{(N)}$ from [10, Section 4] in the current notation:

$$\hat{\Pi}_p^{(N)} \le \sum_{t_1,\dots,t_N} \sum_{u_1,\dots,u_N} \sum_{y_1,\dots,y_N} \sum_{z_1,\dots,z_N} \sum_{x} C_3^{(0)}(0,t_1,u_1,0) \times \prod_{i=2}^{N} [D_1(t_i,u_i,y_i,z_i,t_{i+1},u_{i+1}) + D_2(t_i,u_i,y_i,z_i,t_{i+1},u_{i+1})] D_3(t_N,u_N,y_N,z_N,x). \tag{5.31}$$

Our strategy for bounding (5.31) will be to first show that the "tails" of the diagrams, i.e.,

$$F_1 := \sup_a \sum_{b,c,f,g,h,k,x} D_1(0,a,b,c,f,g) D_3(f,g,h,k,x), \tag{5.32}$$

$$F_2 := \sup_a \sum_{b,c,f,g,h,k,x} D_2(0,a,b,c,f,g) D_3(f,g,h,k,x), \tag{5.33}$$

(see Figure 3) are of order $O(m^{-2}+V^{-1/3})$ because of the bounds derived in Section 4, and then to bound what remains of (5.31) by $O(\theta^{N-2})$ with the help of repeated applications of the bounds in (1.6) and (4.21). Summing $\hat{\Pi}_p^{(N)}$ over N, we will thus also get $O(m^{-2}+V^{-1/3})$ when θ is sufficiently small.

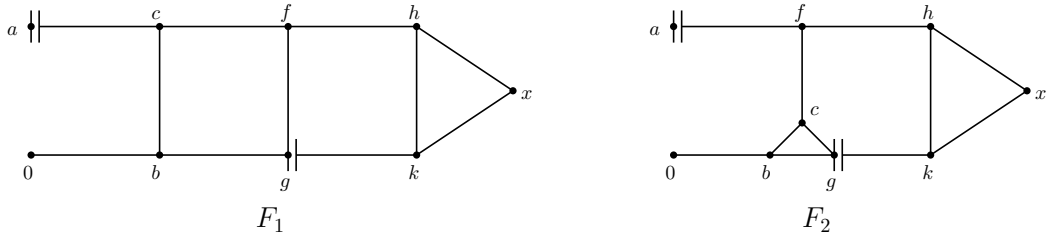

Figure 3. Diagrammatic description of the summands in F_1, F_2.

To bound F_1, we treat the cases $h=k$, $d(h,k)=1$, $d(h,k)\ge 2$ separately.

Case $d(h,k)\ge 2$**:** Apply Lemma 4.4 to the second sum in D_3, and apply (1.6) and (4.21) to the remaining three triangle diagrams. This, combined with (1.7), yields the bound $O(m^{-2}+V^{-1/3})$.

Case $d(h,k)=1$**:** Apply Lemma 4.5 to the first sum in D_3, Lemma 4.4 to the second sum in D_3, and (1.6) and (4.21) to the sums in D_1, to again obtain the bound $O(m^{-2}+V^{-1/3})$.

Case $h=k$**:** Apply Lemma 4.4 to the second sum in D_3, to obtain the bound

$$\sup_a \sum_{b,c,f,g,h} C_3^{(1)}(0,b,c,a) C_3^{(0)}(b,g,h,c) C_2^{(1)}(f,g,h) O(1). \tag{5.34}$$

Compare this with the bound (5.16) on $\hat{\Pi}_p^{(1)}$. The main difference is that the placement of the extra open edges is different (i.e., where the zeros and ones are in the superscripts). We can analyze (5.34) essentially in the same way as (5.16). The most important difference is that here the term due to $0=a=b=c$ and $f=g=h$ does *not* give main contribution, because $j=1$ in the first term here, and so by Lemma 4.4 the term in (5.34) that corresponds to M

in (5.16) is of order $O(m^{-2}+V^{-1/3})$ here. Following the same steps as in the bound on (5.16) above, we conclude that all other terms are also of order $O(m^{-2}+V^{-1/3})$, and hence,

$$F_1 = O(m^{-2}+V^{-1/3}). \tag{5.35}$$

To bound F_2 we again treat the cases $h=k$, $d(h,k)=1$, $d(h,k)\geq 2$ separately.

Cases $d(h,k)=1$ and $d(h,k)\geq 2$: The sums in D_3 can be bounded in the same way as above, to yield a factor $O(m^{-2}+V^{-1/3})$. In [10, (4.47)] it is proved that

$$\sup_a \sum_{b,c,f,g} D_2(0,a,b,c,f,g) \leq (1+10\,\theta^3+O(m^{-1}))(3\theta^3+O(m^{-1})), \tag{5.36}$$

so the bound on D_2 is $O(1)$. Consequently, these cases contribute $O(m^{-2}+V^{-1/3})$ to F_2, as required.

Case $h=k$: The final case is more subtle. We consider the contributions from $f=g$, $d(f,g)=1$, $d(f,g)\geq 2$ separately.

If $h=k$ and $d(f,g)\geq 2$, then we may apply Lemma 4.4 to the first sum in D_3 for a factor $O(m^{-2}+V^{-1/3})$, and to the second sum in D_3 for a factor $O(1)$. Further applying (5.36) to the D_2 term in F_2, we find that this case contributes $O(m^{-2}+V^{-1/3})$.

The contribution due to $h=k$ and $f=g$ is given by

$$\sup_a \sum_{b,c,f,h,x} C_5^{(1)}(0,b,f,c,f,a)\mathbb{P}_p(b\leftrightarrow c)C_2^{(1)}(f,h,f)C_2^{(0)}(h,x,h). \tag{5.37}$$

By Lemma 4.4, the third term is bounded by $O(m^{-1}+V^{-1/3})$ and the fourth term is bounded by $O(1)$. To bound the two remaining terms, we write

$$\begin{aligned}
&\sup_a \sum_{b,c,f} C_5^{(1)}(0,b,f,c,f,a)\,\mathbb{P}_p(b\leftrightarrow c) \\
&\quad = \sup_a \sum_{b,c,f} \mathbb{P}_p(0\leftrightarrow b)\,C_3^{(0)}(b,c,f,b)\,\mathbb{P}_p(c\leftrightarrow f) \sum_{v:\{a,v\}\in\mathcal{E}} p\,\mathbb{P}_p(f\leftrightarrow v) \\
&\quad = \sup_{a'} \sum_{b} C_2^{(1)}(0,b,a') \sum_{c,f} C_3^{(0)}(b,c,f,b)\,\mathbb{P}_p(c\leftrightarrow f),
\end{aligned} \tag{5.38}$$

where for the second equality we use translation invariance of the two-point function to shift the vertex a to $a'=a-f+b$ and f to b in the term $\sum_{v:\{a,v\}\in\mathcal{E}} p\,\mathbb{P}_p(f\leftrightarrow v)$. Using that $\mathbb{P}_p(c\leftrightarrow f)\leq 1$ and applying Lemma 4.4 twice, we find that the remaining terms are bounded by $O(m^{-1}+V^{-1/3})$. Combined with the bounds on the other two terms, we thus conclude that the case where $h=k$ and $f=g$ contributes $O(m^{-2}+V^{-1/3})$.

It remains to bound the case $h=k$ and $d(f,g)=1$. We apply Lemmas 4.4 and 4.5 to bound the two sums in D_3 by $O(m^{-1}+V^{-1/3})$ and write the remaining D_2 terms as

$$\begin{aligned}
&\sup_a \sum_{b,c,f} \sum_{g:\{f,g\}\in\mathcal{E}} C_5^{(1)}(0,b,g,c,f,a)\,\mathbb{P}_p(b\leftrightarrow c) \\
&\qquad\qquad \leq \sup_{a'} \sum_{b,f'} C_3^{(1)}(0,b,f',a') \sup_f \sum_{c,g:\{f,g\}\in\mathcal{E}} C_3^{(0)}(b,c,g,b),
\end{aligned} \tag{5.39}$$

where for the bound we use the translation invariance of the two-point function again to shift a to $a' = a - c + b$, f to $f' = f - c + b$ and c to b, and we take the supremum over f in the second sum for an upper bound. Apply Lemma 4.5 to the second sum and Lemma 4.4 to the first, to bound this factor by $O(m^{-1} + V^{-1/3})$ also.

We have thus bounded all contributions to F_2, and conclude that

$$F_2 = O(m^{-2} + V^{-1/3}). \tag{5.40}$$

The remaining terms in (5.31) are $O(\theta^{N-2})$. This is proved by Borgs et al. [10]. More precisely, combining [10, (4.42)] with (1.6), (1.7), (4.21), (5.35) and (5.40), we may conclude that for all $N \geq 2$,

$$\hat{\Pi}_p^{(N)} = (1 + 10\theta^3)\big(3\theta^3(1 + 10\,\theta^3)\big)^{N-2} O(m^{-2} + V^{-1/3}). \tag{5.41}$$

Hence, the sum $\sum_{N=2}^{\infty} \hat{\Pi}_p^{(N)}$ is of order $O(m^{-2} + V^{-1/3})$ when $3\theta^3(1 + 10\,\theta^3) < 1$, which is the constraint on θ that we assumed. □

Acknowledgments

The work in this paper is supported by the Netherlands Organisation for Scientific Research (NWO) through Gravitation-grant NETWORKS-024.002.003. RvdH is also supported by NWO through VICI-grant 639.033.806, and FdH by the European Research Council (ERC) through Advanced Grant VARIS-267356.

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 C. M. Newman. Tree graph inequalities and critical behavior in percolation models. *J. Statist. Phys.*, **36**(1-2):107–143, (1984).
[3] D. Aldous. Brownian excursions, critical random graphs and the multiplicative coalescent. *Ann. Probab.*, **25**(2):812–854, (1997).
[4] J. v. d. Berg and H. Kesten. Inequalities with applications to percolation and reliability. *J. Appl. Prob.*, **22**:556–569, (1985).
[5] S. Bhamidi, R. v. d. Hofstad, and J. S. H. v. Leeuwaarden. Scaling limits for critical inhomogeneous random graphs with finite third moments. *Electron. J. Probab.*, **15**:no. 54, 1682–1703, (2010).
[6] S. Bhamidi, R. v. d. Hofstad, and J. S. H. v. Leeuwaarden. Novel scaling limits for critical inhomogeneous random graphs. *Ann. Probab.*, **40**(6):2299–2361, (2012).
[7] B. Bollobás. The evolution of random graphs. *Trans. Amer. Math. Soc.*, **286**(1):257–274, (1984).
[8] B. Bollobás. *Random graphs*, volume 73 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, (2001).
[9] C. Borgs, J. Chayes, R. v. d. 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).
[10] C. Borgs, J. Chayes, R. v. d. 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).
[11] C. Borgs, J. Chayes, R. v. d. Hofstad, G. Slade, and J. Spencer. Random subgraphs of finite graphs. III. The phase transition for the n-cube. *Combinatorica*, **26**(4):395–410, (2006).
[12] D. Brydges and T. Spencer. Self-avoiding walk in 5 or more dimensions. *Comm. Math. Phys.*, **97**(1-2):125–148, (1985).
[13] R. Durrett. *Random graph dynamics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, (2007).
[14] P. Erdős and A. Rényi. On random graphs. I. *Publ. Math. Debrecen*, **6**:290–297, (1959).
[15] P. Erdős and J. Spencer. Evolution of the n-cube. *Comput. Math. Appl.*, **5**(1):33–39, (1979).

[16] L. Federico, R. van der Hofstad, F. den Hollander, and T. Hulshof. The scaling limit for critical percolation on the Hamming graph. *In preparation*, (2016).
[17] L. Federico, R. van der Hofstad, and T. Hulshof. Connectivity threshold for random subgraphs of the Hamming graph. *Electron. Commun. Probab.*, **21**:Paper No. 27, 8, (2016).
[18] R. Fitzner and R. van der Hofstad. Non-backtracking random walk. *J. Stat. Phys.*, **150**(2):264–284, (2013).
[19] G. Grimmett. *Percolation*. Springer, Berlin, 2nd edition, (1999).
[20] G. Grimmett. Percolation and disordered systems. In *Percolation theory at Saint-Flour*, Probab. St.-Flour, pages 141–303. Springer, Heidelberg, (2012).
[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. v. d. Hofstad. Progress in high-dimensional percolation and random graphs. To appear in the *CRM Short Courses Series* with Springer.
[23] R. v. d. Hofstad. *Random graphs and complex networks*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, (2017).
[24] R. v. d. Hofstad and M. J. Luczak. Random subgraphs of the 2D Hamming graph: the supercritical phase. *Probab. Theory Related Fields*, **147**(1-2):1–41, (2010).
[25] R. v. d. Hofstad, M. J. Luczak, and J. Spencer. The second largest component in the supercritical 2D Hamming graph. *Random Structures Algorithms*, **36**(1):80–89, (2010).
[26] R. v. d. Hofstad and A. Nachmias. Hypercube percolation. *arXiv:1201.3953*, to appear in *J. Eur. Math. Soc.*, (2012).
[27] R. v. d. Hofstad and A. Nachmias. Unlacing hypercube percolation: a survey. *Metrika*, **77**(1):23–50, (2014).
[28] R. v. d. Hofstad and G. Slade. Asymptotic expansions in n^{-1} for percolation critical values on the n-cube and $\mathbb{Z}^n$. *Random Structures Algorithms*, **27**(3):331–357, (2005).
[29] R. v. d. Hofstad and G. Slade. Expansion in n^{-1} for percolation critical values on the n-cube and $\mathbb{Z}^n$: the first three terms. *Combin. Probab. Comput.*, **15**(5):695–713, (2006).
[30] S. Janson, D. Knuth, T. Łuczak, and B. Pittel. The birth of the giant component. *Random Structures Algorithms*, **4**(3):231–358, (1993). With an introduction by the editors.
[31] S. Janson and L. Warnke. On the critical probability in percolation. *arXiv:1611.08549*, (2016).
[32] A. Joseph. The component sizes of a critical random graph with given degree sequence. *Ann. Appl. Probab.*, **24**(6):2560–2594, (2014).
[33] T. Łuczak, B. Pittel, and J. C. Wierman. The structure of a random graph at the point of the phase transition. *Trans. Amer. Math. Soc.*, **341**(2):721–748, (1994).
[34] M. V. Menshikov. Coincidence of critical points in percolation problems. *Soviet Mathematics, Doklady*, **33**:856–859, (1986).
[35] P. Miłoś and B. Şengül. Existence of a phase transition of the interchange process on the Hamming graph. *arXiv:1605.03548*, (2016).
[36] A. Nachmias and Y. Peres. Component sizes of the random graph outside the scaling window. *ALEA Lat. Am. J. Probab. Math. Stat.*, **3**:133–142, (2007).
[37] A. Nachmias and Y. Peres. Critical random graphs: diameter and mixing time. *Ann. Probab.*, **36**(4):1267–1286, (2008).
[38] A. Nachmias and Y. Peres. Critical percolation on random regular graphs. *Random Structures Algorithms*, **36**(2):111–148, (2010).
[39] A. G. Pakes. Some limit theorems for the total progeny of a branching process. *Advances in Appl. Probability*, **3**:176–192, (1971).
[40] B. Pittel. On the largest component of the random graph at a nearcritical stage. *J. Combin. Theory Ser. B*, **82**(2):237–269, (2001).
[41] L. Russo. On the critical percolation probabilities. *Z. Wahrsch. Verw. Gebiete*, **56**(2):229–237, (1981).

E-mail address: `l.federico@tue.nl`

E-mail address: `r.w.v.d.hofstad@tue.nl`

E-mail address: `denholla@math.leidenuniv.nl`

E-mail address: `w.j.t.hulshof@tue.nl`