J. Eur. Math. Soc. 19, 725–814

DOI 10.4171/JEMS/679

Remco van der Hofstad · Asaf Nachmias

Hypercube percolation

Received March 3, 2014

Abstract. We study bond percolation on the Hamming hypercube $\{0, 1\}^m$ around the critical probability p_c. It is known that if $p = p_c(1 + O(2^{-m/3}))$, then with high probability the largest connected component $\mathscr{C}_1$ is of size $\Theta(2^{2m/3})$. Here we show that for any sequence $\varepsilon(m)$ such that $\varepsilon(m) = o(1)$ but $\varepsilon(m) \gg 2^{-m/3}$ percolation on the hypercube at $p_c(1 + \varepsilon(m))$ has

$$|\mathscr{C}_1| = (2 + o(1))\varepsilon(m)2^m \quad \text{and} \quad |\mathscr{C}_2| = o(\varepsilon(m)2^m),$$

with high probability, where $\mathscr{C}_2$ is the second largest component. This resolves a conjecture of Borgs, Chayes, the first author, Slade and Spencer [18].

Keywords. Hypercube, percolation, critical behavior, mean-field results, scaling window, birth of the giant component, non-backtracking random walk, mixing time

Contents

R. van der Hofstad: Department of Mathematics and Computer Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands; e-mail: rhofstad@win.tue.nl

A. Nachmias: Department of Mathematics, University of British Columbia,
121-1984 Mathematics Rd, Vancouver, BC, Canada V6T1Z2; e-mail: asafnach@math.ubc.ca

Mathematics Subject Classification (2010): 60K35, 82B43, 05C80

1. Introduction

Percolation on the Hamming hypercube $\{0, 1\}^m$ is a combinatorial model proposed in 1979 by Erdős and Spencer [24]. The study of its phase transition poses two inherent difficulties. Firstly, its non-trivial geometry makes the combinatorial "subgraph count" techniques unavailable. Secondly, the critical probability where the phase transition occurs is significantly larger than $1/(m-1)$, making the method of stochastic domination by branching processes very limited. Unfortunately, these are the two prominent techniques for obtaining scaling windows in mean-field settings (see e.g. [4, 13, 19, 23, 32, 38, 44, 46, 47, 48]).

In light of the second difficulty, Borgs, Chayes, the first author, Slade and Spencer [16, 17, 18] suggested that the precise location p_c of the phase transition is the unique solution to the equation

$$\mathbb{E}_{p_c}|\mathscr{C}(0)| = \lambda 2^{m/3}, \tag{1.1}$$

where $\mathscr{C}(0)$ is the connected component containing the origin, $|\mathscr{C}(0)|$ denotes its size, and $\lambda \in (0, 1)$ denotes an arbitrary constant. Later it will become clear how λ is chosen. The lace expansion was then employed by the authors to show that at $p = p_c(1 + O(2^{-m/3}))$ the largest connected component $\mathscr{C}_1$ is of size $\Theta(2^{2m/3})$ with high probability —the same asymptotics as in the critical Erdős and Rényi random graph both with respect to the size of the cluster and the width of the scaling window (see Section 1.1 for more details). However, this result does not rule out the possibility that this critical behavior proceeds beyond the $O(2^{-m/3})$ window and does not give an upper bound on the width of the scaling window.

The authors conjectured that the giant component "emerges" just above this window (see [18, Conjecture 3.2]). They were unable to prove this primarily because their combination of lace expansion and sprinkling methodology breaks down for p above the scaling window. In this paper we resolve their conjecture:

Theorem 1.1. *Consider bond percolation on the Hamming hypercube $\{0, 1\}^m$ with $p = p_c(1 + \varepsilon)$, where $p_c = p_c(\lambda)$ with $\lambda \in (0, \infty)$ a fixed constant, and $\varepsilon = \varepsilon(m) = o(1)$ is a positive sequence with $\varepsilon(m) \gg 2^{-m/3}$. Then*

$$\frac{|\mathscr{C}_1|}{2\varepsilon 2^m} \xrightarrow{\mathbf{P}} 1,$$

where $\xrightarrow{\mathbf{P}}$ denotes convergence in probability, and

$$\mathbb{E}|\mathscr{C}(0)| = (4 + o(1))\varepsilon^2 2^m.$$

Furthermore, the second largest component $\mathscr{C}_2$ satisfies

$$\frac{|\mathscr{C}_2|}{\varepsilon 2^m} \xrightarrow{\mathbf{P}} 0.$$

The main novelty of our approach is showing that large percolation clusters behave in some sense like uniform random sets. We use this to deduce that two large clusters tend to "clump" together and form a giant component. This analysis replaces the appeal to the hypercube's isoperimetric inequality which is key in all the previous works on this problem (see further details in Section 1.3). It essentially rules out the possibility that two large percolation clusters are "worst-case" sets, that is, sets which saturate the isoperimetric inequality (e.g., two balls of radius $m/2-\sqrt{m}$ around the two poles of the hypercube). The precise behavior of the non-backtracking random walk on the hypercube plays a key role in proving such statements. Our proof combines this idea with some combinatorial ideas (the "sprinkling" method of [3], see Section 1.3), and ideas originating in statistical physics (Aizenman and Barsky's [1] differential inequalities and variants of the triangle condition). Our proof methods are general and apply to other families of graphs such as various expanders of high degree and high girth, finite tori of dimension growing with the length and products of complete graphs of any dimension (answering a question asked in [32]). We state our most general theorem in Section 1.5 and illustrate its use with some examples.

The problem of establishing a phase transition for the appearance of a component of size order 2^m was solved in the breakthrough work of Ajtai, Komlós and Szemerédi [3]. They proved that when the retention probability of an edge is scaled as $p = c/m$ for a fixed constant $c > 0$, the model exhibits a phase transition: if $c < 1$, then the largest component has size of order m, and if $c > 1$, then the largest component has size linear in 2^m, with high probability.

At about the same time, Bollobás [13] initiated a study of zooming in onto the large scale properties of the phase transition on the Erdős and Rényi [23] random graph $G(n, p)$ (see Section 1.1 below). However, unlike $G(n, p)$, the phase transition in the hypercube does *not* occur around $p = 1/(\deg - 1)$, where deg denotes the degree of the graph. In fact, it was shown by the first author and Slade [35, 36] that p_c of the hypercube $\{0, 1\}^m$ satisfies

$$p_c = \frac{1}{m-1} + \frac{5/2}{m^3} + O(m^{-4}). \tag{1.2}$$

Here and below we write $f(m) = O(g(m))$ if $|f(m)|/|g(m)|$ is uniformly bounded from above by a positive constant, $f(m) = \Theta(g(m))$ if $f(m) = O(g(m))$ and $g(m) = O(f(m))$, and $f(m) = o(g(m))$ if $f(m)/g(m)$ tends to 0 with m. We also say that a sequence of events $(E_m)_{m\geq 1}$ *occurs with high probability* (whp) when $\lim_{m\to\infty} \mathbf{P}(E_m) = 1$.

The first improvement to [3] was obtained by Bollobás, Kohayakawa and Łuczak [15]. They showed that if $p=(1+\varepsilon(m))/(m-1)$ with $\varepsilon(m)=o(1)$ but $\varepsilon(m)\geq 60m^{-1}(\log m)^3$, then $|\mathscr{C}_1| = (2 + o(1))\varepsilon(m)2^m$ whp. In view of (1.2), it is clear that one cannot improve the regime of $\varepsilon(m)$ in their result to more than $\varepsilon(m) \geq m^{-2}$. In [18], the authors show that if $\varepsilon(m) \geq \mathrm{e}^{-cm^{1/3}}$ and $p = p_c(1 + \varepsilon(m))$, then $|\mathscr{C}_1| \geq c\varepsilon(m)2^m$ whp. Note that $\mathrm{e}^{-cm^{1/3}} \gg 2^{-\alpha m}$ for any $\alpha > 0$, so the requirement on $\varepsilon(m)$ of Theorem 1.1 is much weaker. Our result, combined with those in [16, 17, 18], shows that it is sharp and therefore fully identifies the phase transition on the hypercube.

Other models of statistical physics, such as random minimal spanning trees and bootstrap percolation on the hypercube, have been studied before; we refer the reader to [7, 8, 50]. In the remainder of this section we present some of the necessary background and context of the result, briefly describe our techniques (we provide a more detailed overview of the proof in the next section) and present a general theorem that is used to establish scaling windows for percolation on various other graphs studied in the literature.

1.1. The Erdős and Rényi random graph. Recall that $G(n, p)$ is obtained from the complete graph by retaining each edge of the complete graph on n vertices with probability p and erasing it otherwise, independently for all edges. Write $\mathscr{C}_j$ for the jth largest component obtained this way. An inspiring discovery of Erdős and Rényi [23] is that this model exhibits a phase transition when p is scaled like $p = c/n$. When $c < 1$ we have $|\mathscr{C}_1| = \Theta(\log n)$ whp, while $|\mathscr{C}_1| = \Theta(n)$ whp when $c > 1$.

The investigation of the case $c \sim 1$, initiated by Bollobás [13] and continued by Łuczak [44], revealed an intricate picture of the phase transition's nature. See [14] for results up to 1984, and [4, 38, 39, 45] for references to subsequent work. We briefly describe these here.

The critical window. If $p = (1 + an^{-1/3})/n$ for some fixed $a \in \mathbb{R}$, then for any fixed integer $j \geq 1$,

$$(|\mathscr{C}_1|/n^{2/3}, \ldots, |\mathscr{C}_j|/n^{2/3}) \xrightarrow{d} (\chi_1, \ldots, \chi_j),$$

where $(\chi_i)_{i=1}^j$ are random variables supported on $(0, \infty)$, and $\xrightarrow{d}$ denotes convergence in distribution.

The subcritical phase. Let $\varepsilon = \varepsilon(n) = o(1)$ be a non-negative sequence with $\varepsilon \gg n^{-1/3}$ and set $p = (1-\varepsilon)/n$. Then, for any fixed integer $j \geq 1$,

$$\frac{|\mathscr{C}_j|}{2\varepsilon^{-2}\log(\varepsilon^3 n)} \xrightarrow{\mathbf{P}} 1.$$

The supercritical phase. Let $\varepsilon = \varepsilon(n) = o(1)$ be a non-negative sequence with $\varepsilon \gg n^{-1/3}$ and set $p = (1+\varepsilon)/n$. Then

$$\frac{|\mathscr{C}_1|}{2\varepsilon n} \xrightarrow{\mathbf{P}} 1,$$

and, for any fixed integer $j \geq 2$,

$$\frac{|\mathscr{C}_j|}{2\varepsilon^{-2}\log(\varepsilon^3 n)} \xrightarrow{\mathbf{P}} 1.$$

Thus, the prominent qualitative features of this phase transition are:

(1) The emergence of the giant component just above the scaling window. That is, only in the supercritical phase do we find that $|\mathscr{C}_2| \ll |\mathscr{C}_1|$, and that $|\mathscr{C}_1|/n$ increases suddenly but smoothly above the critical value (in mathematical physics jargon, the phase transition is of second order).
(2) Concentration of the size of the largest connected components outside the scaling window and non-concentration inside the window.
(3) Duality: $|\mathscr{C}_2|$ in the supercritical phase has the same asymptotics as $|\mathscr{C}_1|$ in the corresponding subcritical phase.

Theorem 1.1 proves (1) and the concentration in the supercritical regime in (2). Property (3) on the hypercube remains an open problem (see Section 8).

1.2. Random subgraphs of transitive graphs. Let us briefly review the study of percolation on finite transitive graphs presented in [16, 17, 18] (see also [6]). We focus here only on some of the many results obtained in those papers. Let G be a finite transitive graph and write V for the number of vertices of G. Let $p \in [0, 1]$ and write G_p for the random graph obtained from G by retaining each edge with probability p and erasing it with probability $1-p$, independently for all edges. We also write $\mathbf{P}_p$ for this probability measure. We say an edge is *p-open* [*p-closed*] if it was retained [erased]. We say that a path in the graph is p-open if all of its edges are p-open. For two vertices x, y we write $x \leftrightarrow y$ for the event that there exists a p-open path connecting x and y. For an integer $j \geq 1$ we write $\mathscr{C}_j$ for the jth largest component of G_p (breaking ties arbitrarily), and for a vertex v we write $\mathscr{C}(v)$ for the component in G_p containing v.

For two vertices x, y we denote

$$\nabla_p(x, y) = \sum_{u,v} \mathbf{P}_p(x \leftrightarrow u)\mathbf{P}_p(u \leftrightarrow v)\mathbf{P}_p(v \leftrightarrow y). \tag{1.3}$$

The quantity $\nabla_p(x, y)$, known as the *triangle diagram*, was introduced by Aizenman and Newman [2] to study critical percolation on high-dimensional infinite lattices. In that setting, an important feature of an infinite graph G is whether $\nabla_{p_c}(0,0) < \infty$. This condition is often referred to as the *triangle condition*. In high dimensions, Hara and Slade [28] proved the triangle condition. It allows one to deduce that numerous critical exponents attain the same values as they do on an infinite regular tree (see e.g. [9, 2, 40, 42]). See also [31] for a survey of results on high-dimensional percolation, including a discussion of percolation on high-dimensional tori such as the hypercube.

When G is a finite graph, $\nabla_p(0,0)$ is obviously finite; however, there is still a *finite* triangle condition which in turn guarantees that random critical subgraphs of G have the same geometry as random subgraphs of the complete graph on V vertices, where V denotes the number of vertices in G. That is, in the finite setting the role of the infinite regular tree is played by the complete graph. Let us make this heuristic formal.

We always have $V \to \infty$, and $\lambda \in (0,1)$ is a fixed constant. Let $p_c = p_c(\lambda)$ be defined by

$$\mathbb{E}_{p_c(\lambda)}|\mathscr{C}(0)| = \lambda V^{1/3}. \tag{1.4}$$

The *finite triangle condition* is the assumption that $\nabla_{p_c(\lambda)}(x,y) \le \mathbf{1}_{\{x=y\}} + a_0$ for some $a_0 = a_0(\lambda)$ sufficiently small when λ is small. The *strong triangle condition*, defined in [17, (1.26)], is the statement that there exists a constant C such that for all $p \le p_c$,

$$\nabla_p(x,y) \le \mathbf{1}_{\{x=y\}} + C\chi(p)^3/V + \alpha_G, \tag{1.5}$$

where $\alpha_G \to 0$ as $m \to \infty$ and $\chi(p) = \mathbb{E}_p|\mathscr{C}(0)|$ denotes the expected cluster size. Throughout this paper, we will assume that the strong triangle condition holds. In fact, in all examples where the finite triangle condition is proved to hold, actually the strong triangle condition (1.5) is proved. In [17], (1.5) is shown to hold for various graphs: the complete graph, the hypercube and high-dimensional tori $\mathbb{Z}_n^d$. In particular, the next theorem states (1.5) for the hypercube.

Theorem 1.2 ([17]). *Consider percolation on the hypercube $\{0,1\}^m$. Then for any λ there exists a constant $C = C(\lambda) > 0$ such that for any $p \le p_c(\lambda)$ (as defined in (1.4)),*

$$\nabla_p(x,y) \le \mathbf{1}_{\{x=y\}} + C(\chi(p)^3/V + 1/m). \tag{1.6}$$

As we discuss in Remark 4 below Theorem 1.3, our methodology can be used to yield a simple proof of Theorem 1.2 without relying on the lace-expansion methods derived in [17]. The main effort in [16] is to show that under condition (1.5) the phase transition behaves similarly to the one in $G(n,p)$ described in the previous section. The main results obtained in [16] are the following:

The critical window. Let G be a finite transitive graph for which (1.5) holds. Then, for $p = p_c(1 + O(V^{-1/3}))$,

$$\mathbf{P}(A^{-1}V^{2/3} \le |\mathscr{C}_1| \le AV^{2/3}) = 1 - O(A^{-1}).$$

The subcritical phase. Let G be a finite transitive graph for which (1.5) holds. Let $\varepsilon = o(1)$ be a non-negative sequence with $\varepsilon \gg V^{-1/3}$ and set $p = p_c(1-\varepsilon)$. Then, for all fixed $\delta > 0$,

$$\mathbf{P}\big(|\mathscr{C}_1| \le (2+\delta)\varepsilon^{-1}\log(\varepsilon^3 V)\big) = 1 - o(1).$$

The supercritical phase. Let G be a finite transitive graph for which (1.5) holds. Let $\varepsilon = o(1)$ be a non-negative sequence with $\varepsilon \gg V^{-1/3}$ and set $p = p_c(1+\varepsilon)$. Then

$$\mathbf{P}(|\mathscr{C}_1| \geq A\varepsilon V) = O(A^{-1}).$$

Thus, while these results hold in a very general setting, they are incomplete. Most notably, in the supercritical phase there is no matching lower bound on $|\mathscr{C}_1|$. So, a priori, it is possible that $|\mathscr{C}_1|$ is still of order $V^{2/3}$ when $p = p_c(1+\varepsilon)$ for some $\varepsilon \gg V^{-1/3}$ and that the scaling window is in fact much larger than $V^{-1/3}$. It remains an open problem whether (1.5) alone implies that $|\mathscr{C}_1|/(\varepsilon V)$ converges in probability to a constant in the supercritical phase.

As we mentioned before, the particular case of the hypercube was addressed in [18]. There the authors employed some of the result of [16, 17] together with a *sprinkling* argument to provide a lower bound of order $\varepsilon 2^m$ on $|\mathscr{C}_1|$ valid only when $\varepsilon \geq \mathrm{e}^{-cm^{1/3}}$. We will rely on the sprinkling method for the arguments in this paper, so let us briefly expand on it.

1.3. Sprinkling. The sprinkling technique was invented by Ajtai, Komlós and Szemerédi [3] to show that $|\mathscr{C}_1| = \Theta(2^m)$ when $p = (1+\varepsilon)/m$ for fixed $\varepsilon > 0$, and can be described as follows. Fix some small $\theta > 0$ and write $p_1 = (1+(1-\theta)\varepsilon)/m$ and $p_2 \geq \theta\varepsilon/m$ such that $(1-p_1)(1-p_2) = 1-p$. It is clear that G_p is distributed as the union of the edges in two independent copies of G_{p_1} and G_{p_2}. The sprinkling method consists of two steps. The first step is performed in G_{p_1} and uses a branching process comparison argument together with the Azuma–Hoeffding concentration inequality to deduce that whp at least $c_2 2^m$ vertices are contained in connected components of size at least $2^{c_1 m}$ for some small but fixed constants $c_1, c_2 > 0$. In the second step we add the edges of G_{p_2} (these are the "sprinkled" edges) and show that they connect many of the clusters of size at least $2^{c_1 m}$ into a giant cluster of size $\Theta(2^m)$.

Let us give some details on how the last step is done. A key tool here is the *isoperimetric inequality* for the hypercube, stating that two disjoint subsets of the hypercube of size at least $c_2 2^m/3$ have at least $2^m/m^{100}$ disjoint paths of length $C(c_2)\sqrt{m}$ connecting them, for some constant $C(c_2)$. (The m^{100} in the denominator is not sharp, but this is immaterial as long as it is a polynomial in m.) This fact is used in the following way. Write V' for the set of vertices that are contained in a component of size at least $2^{c_1 m}$ in G_{p_1} so that $V' \geq c_2 2^m$. We say that *sprinkling fails* when $|\mathscr{C}_1| \leq c_2 2^m/3$ in $G_{p_1} \cup G_{p_2}$. If sprinkling fails, then we can partition $V' = A \uplus B$ so that both A and B have cardinality at least $c_2 2^m/3$ and *any* path of length at most $C(c_2)\sqrt{m}$ between them has an edge that is p_2-closed. The number of such partitions is at most $2^{2^m/2^{c_1 m}}$. The probability that a path of length k has a p_2-closed edge is $1 - p_2^k$. Applying the isoperimetric inequality and using the fact that the paths guaranteed to exist by it are disjoint so that the edges in them are independent, we find that the probability that sprinkling fails is at most

$$2^{2^m/2^{c_1 m}} \cdot (1 - (\theta\varepsilon/m)^{C(c_2)\sqrt{m}})^{2^m/m^{100}} = \mathrm{e}^{-2^{(1+o(1))m}}, \tag{1.7}$$

which tends to 0.

1.4. Revised sprinkling. The sprinkling argument above is not optimal due to the use of the isoperimetric inequality. It is wasteful because it assumes that large percolation clusters can be "worst-case" sets, that is, sets which saturate the isoperimetric inequality (e.g., two balls of radius $m/2 - \sqrt{m}$ around two vertices at Hamming distance m). However, it is in fact very improbable for percolation clusters to be similar to this kind of worst-case sets. Our approach replaces the use of the isoperimetric inequality by proving statements showing that large percolation clusters are "close" to uniform random sets of similar size, more precisely, they behave like a random collection of connected components of size ε^{-2}. This allows us to deduce that two large clusters share many closed edges with the property that if we open even *one* of them, then the two clusters connect. While previously we had paths of length $\sqrt{m}$ connecting the two clusters, here we will have paths of length precisely 1. The final line of our proof, replacing (1.7), will be

$$2^{2\varepsilon V/(k(m)\varepsilon^{-2})} \cdot (1-\theta\varepsilon/m)^{m\varepsilon^2 V} \le \mathrm{e}^{-\theta\varepsilon^3 V(1+o(1))}, \tag{1.8}$$

where $k(m)$ is some sequence with $k(m) \to \infty$ very slowly. This tends to 0 since $\varepsilon^3 V \to \infty$. Compared with the logic leading to (1.7), this line is rather suggestive. We will see that whp $2\varepsilon V$ vertices are in components of size at least $k(m)\varepsilon^{-2}$, explaining the $2^{2\varepsilon V/(k(m)\varepsilon^{-2})}$ factor in (1.8). The main effort in this paper is to justify the second factor showing that for any partition of these vertices into two sets of size εV, the number of closed edges between them is at least $\varepsilon^2 m V$—the same number of edges one would expect two uniform random sets of size εV to have between them. Therefore, given a partition, the probability that sprinkling fails for it is bounded by $(1-\theta\varepsilon/m)^{m\varepsilon^2 V}$.

1.5. The general theorem. Our methods use relatively simple geometric properties of the hypercube and apply to a larger set of underlying graphs. We present this general setting that the majority of the paper assumes and briefly discuss some other cases for which our main theorem holds aside from the hypercube. We remark that the impatient reader may proceed assuming the underlying graph is always the hypercube $\{0,1\}^m$—we have set up the notation to support this, since the hypercube, in some sense, is our most "difficult" example.

The geometric conditions of our underlying graphs will be stated in terms of non-backtracking random walks. The main advantage of this approach is that these conditions are relatively easy to verify. Let G be a finite transitive graph on V vertices and degree m. Consider the non-backtracking random walk on it (this is just a simple random walk not allowed to traverse back on the edge it just came from; see Section 3.4 for a precise definition). For any vertices x, y, we write $\mathbf{p}^t(x,y)$ for the probability that the walk started at x visits y at time t. For any $\xi > 0$, we write $T_{\mathrm{mix}}(\xi)$ for the ξ*-upper-uniform mixing time* of the walk, that is,

$$T_{\mathrm{mix}}(\xi) = \min\left\{t:\ \max_{x,y} \frac{\mathbf{p}^t(x,y) + \mathbf{p}^{t+1}(x,y)}{2} \le (1+\xi)V^{-1}\right\}.$$

Theorem 1.3. *Let G be a transitive graph on V vertices with degree m and define p_c as in (1.1) with $\lambda = 1/10$. Assume that there exists a sequence $\alpha_G = o(1)$ with $\alpha_G \ge 1/m$ such that if we set $t_{\mathrm{mix}} = T_{\mathrm{mix}}(\alpha_G)$ then:*

(1) *$m \to \infty$ as $V \to \infty$,*
(2) *$[p_c(m-1)]^{t_{\mathrm{mix}}} = 1 + O(\alpha_G)$,*
(3) *for any vertices x, y,*

$$\sum_{u,v} \sum_{\substack{t_1,t_2,t_3=0 \\ t_1+t_2+t_3 \geq 3}}^{t_{\mathrm{mix}}} \mathbf{p}^{t_1}(x,u)\mathbf{p}^{t_2}(u,v)\mathbf{p}^{t_3}(v,y) = O(\alpha_G / \log V).$$

Then

(a) *the finite triangle condition* (1.5) *holds* (*and hence the results in* [16] *described in Section* 1.2 *hold*),
(b) *for any sequence $\varepsilon = \varepsilon(m)$ satisfying $\varepsilon \gg V^{-1/3}$ and $\varepsilon = o(1/t_{\mathrm{mix}})$, bond percolation on G with $p = p_c(1+\varepsilon)$ has*

$$\frac{|\mathscr{C}_1|}{2\varepsilon V} \xrightarrow{\mathbf{P}} 1, \qquad \mathbb{E}|\mathscr{C}(0)| = (4+o(1))\varepsilon^2 V, \qquad \frac{|\mathscr{C}_2|}{\varepsilon V} \xrightarrow{\mathbf{P}} 0.$$

Remark 1. In the case of the hypercube $\{0,1\}^m$ we will take $\alpha_G = m^{-1}\log m$ and verify the conditions of Theorem 1.3. This is done in Section 7. Although the behavior of random walks on the hypercube is well understood, we have not been able to find an estimate on the uniform mixing time of the non-backtracking random walk yielding $T_{\mathrm{mix}}(m^{-1}\log m) = \Theta(m \log m)$ in the literature. To show this we use the recent paper of Fitzner and the first author [25] in which the non-backtracking walk transition matrix on the hypercube is analyzed. We use this result in Lemma 7.1 to verify condition (3), and condition (2) follows directly from (1.2) (though (2) can also be verified by elementary means without using (1.2)—see Remark 4).

Remark 2. Note that part (b) of Theorem 1.3 only applies when $\varepsilon(m) = o(1/t_{\mathrm{mix}})$ and not for any $\varepsilon = \varepsilon(m) = o(1)$. Thus, for a complete proof of Theorem 1.1, we also require a separate argument dealing with the regime $\varepsilon(m) \geq c/t_{\mathrm{mix}}$—in the case of the hypercube and other graphs mentioned in this paper, this is a much easier regime in which previous techniques based on sprinkling and isoperimetric inequalities are effective.

Remark 3. Random walk conditions for percolation on finite graphs were first given by the second author in [47]. The significant difference between the two approaches is that in [47] the condition requires controlling the random walk behavior for a period of time that is as long as the critical cluster diameter, that is, $V^{1/3}$. The outcome is that the results of [47] only apply when $p_c = (1 + O(V^{-1/3}))/(m-1)$, and hence do not apply in the case of the hypercube. Here we are only interested in the behavior of the random walk up to the mixing time, even though typical percolation paths are much longer. The reason for this is that it turns out that it is enough to randomize the beginning of a percolation path in order to conclude that the end point is uniformly distributed (see Section 2.4). Another difference is that the results in [47] only show that $|\mathscr{C}_1| \geq c\varepsilon V$ for some $c > 0$ and do not give the precise asymptotic value of $|\mathscr{C}_1|$ as we do here.

Remark 4. Our approach also enables us to give a simple proof for the fact that the finite triangle condition (Theorem 1.2) holds for the hypercube without using the lace expansion as in [17]; this is done in [34]. Our proof of this fact relies on the estimate $p_c = 1/(m-1) + O(m^{-3})$ (which is much weaker than (1.2) but also much easier to prove) and on the argument presented in Section 2.4. In fact, in the current paper we only rely on this easy estimate for p_c, so our main result, Theorem 1.1, is in fact self-contained and does not rely at any time on results obtained via the lace expansion in [17] (we do use arguments of [16] which rely on the triangle condition). We refer to [34] for a more extensive discussion on this subject.

In many cases, verifying the conditions of Theorem 1.3 is done using known methods from the theory of percolation and random walks (note that condition (2) involves both a random walk and a percolation estimate). We illustrate this in Section 7 in the case of the hypercube (thus proving Theorem 1.1) and for expander families of high degree and high girth (see [37] for an introduction to expanders). This is a class of graphs that contains various examples such as Payley graphs (see e.g. [21]), products K_n^d of complete graphs and many others. Percolation on products of complete graphs was studied in [32, 33, 47] in the cases $d = 2, 3$; our expander theorem allows us to provide a complete description of the phase transition in any fixed dimension d, answering a question posed in [32]. Recall that a sequence of graphs G_n is called an *expander* family if there exists a constant $c > 0$ such that the second largest eigenvalue of the transition matrix of the simple random walk is at most $1 - c$ (the largest eigenvalue is 1). Also, the *girth* of a graph is the length of the shortest cycle. It is a classical fact that on expanders $T_{\rm mix}(V^{-1}) = O(\log V)$, where V is the number of vertices of the graph (see e.g. [5, below (19)]).

Theorem 1.4. *Let G_m be a transitive family of expanders with degree $m \to \infty$ and V vertices. Assume that $m \geq c \log V$ and the girth of G is at least $c \log V/\log(m-1)$ for some fixed $c > 0$. Then the conditions of Theorem 1.3 hold, and hence the conclusions of that theorem hold.*

For products K_n^d of complete graphs, the girth equals 3, $V = n^d$ and $m = d(n-1)$, so that the girth assumption is satisfied for $c \leq 3(1 - o(1))/d$ and n sufficiently large. Theorem 1.3 applies to other examples of graphs, not included in the last theorem, for example, products of complete graphs K_n^d where d may depend on n (as long as $n + d \to \infty$) and finite tori $\mathbb{Z}_n^d$ but only when $d = d(n)$ grows at some rate with n. We omit the details since they are rather similar. We emphasize, however, that there are important examples which our methods are insufficient to solve. Most prominently, there are bounded degree expanders with low girth (the case of girth $\geq (2/3 + \varepsilon) \log_{m-1} n$ was solved in [47]) and finite tori $\mathbb{Z}_n^d$ where d is large but fixed. It seems that new ideas are required to study percolation on these graphs (see Section 8).

1.6. Organization. This paper is organized as follows. In Section 2 we give an overview of our proof, stating the main results upon which the proof is based. In Section 3 we prove several estimates on the number of vertices satisfying various properties, such as having large clusters, or surviving up to great depth. We further prove detailed estimates on connection probabilities. In Section 4 we prove expected volume estimates both in the critical and supercritical regimes. In Section 5 we prove an intrinsic-metric regularity theorem, showing that most vertices that survive long and have a large cluster size have neighborhoods that are sufficiently regular. See Section 2.5 for an explanation on how the estimates of Section 5 are used in the proof. In Section 6 we show that most large clusters have many closed edges between them, which is the main result in our proof. In Section 7 we perform the improved sprinkling argument as indicated in Section 1.4, and complete the proof of Theorem 1.1. In Section 8 we discuss several open problems. We close the paper with an Appendix where we sharpen the arguments in [9] and [17] to obtain the asymptotics of the supercritical cluster tail.

2. Overview of the proof

In this section we give an overview of the key steps in our proof. Throughout the rest of the paper we assume that $\varepsilon = \varepsilon(m)$ is a sequence such that $\varepsilon = o(1)$ but $\varepsilon^3 V \to \infty$.

2.1. Notation and tools. Let G be a transitive graph and recall that G_p is obtained from G by independently retaining each edge with probability p. Recall that $x \leftrightarrow y$ denotes that there exists a p-open path connecting x and y. We write $d_{G_p}(x, y)$ for the length of a shortest p-open path between x, y, and set $d_{G_p}(x, y) = \infty$ if x is not connected to y in G_p. We write $x \overset{r}{\longleftrightarrow} y$ if $d_{G_p}(x, y) \le r$; $x \overset{=r}{\longleftrightarrow} y$ if $d_{G_p}(x, y) = r$; and $x \overset{[a,b]}{\longleftrightarrow} y$ if $d_{G_p}(x, y) \in [a, b]$. Further, we write $x \overset{P[a,b]}{\longleftrightarrow} y$ if there exists an open path of length in $[a, b]$ between x and y (not necessarily a shortest path). The event $\{x \overset{[a,b]}{\longleftrightarrow} y\}$ is not increasing with respect to adding edges, but the event $\{x \overset{P[a,b]}{\longleftrightarrow} y\}$ *is*, which often makes it easier to deal with. Whenever the sign $\leftrightarrow$ appears, it will be clear what p is, and we will drop it from the notation. The *intrinsic metric ball* of radius r around x and its boundary are defined by

$$B_x^G(r) = \{y \colon d_{G_p}(x, y) \le r\}, \qquad \partial B_x^G(r) = \{y \colon d_{G_p}(x, y) = r\}.$$

Note that these are random sets of the graph and not balls in the shortest path metric of the graph G. We often drop the G from the above notation and write $B_x(r)$ when it is clear what the underlying graph G is. We also denote

$$B_x^G([a, b]) = \{y \colon x \overset{[a,b]}{\longleftrightarrow} y\}.$$

Our graphs always contain a marked vertex that we call *the origin* and denote by 0. In the case of the hypercube this is taken to be the all-zero vector. We often drop 0 from the notation and write $B(r)$ for $B_0(r)$ whenever possible.

We now define the intrinsic metric *one-arm* event. This was introduced in [49] to study the mixing time of critical $G(n, p)$ clusters and was very useful in the context of high-dimensional percolation in [40]. Define the event

$$H^G(r) = \{\partial B_0^G(r) \neq \emptyset\}$$

for any integer $r \geq 0$, and

$$\Gamma(r) = \sup_{G' \subset G} \mathbf{P}(H^{G'}(r)), \tag{2.1}$$

where the supremum is over all subgraphs G' of G. The reason for the somewhat unnatural definition of Γ is that the event $\partial B_0^G(r) \neq \emptyset$ is not monotone with respect to addition of edges. Indeed, turning an edge from closed to open may shorten a shortest path, rendering a configuration such that the event $\partial B_0^G(r) \neq \emptyset$ no longer occurs. This non-monotonicity problem arises whenever one conditions on $B_0^G(r)$ and would like to estimate the probability that some $v \in \partial B_0^G(r)$ survives an additional ℓ generations, that is, $\partial B_v^G(\ell) \neq \emptyset$ off $B_0^G(r)$. A priori, the survival probability may be much larger on the subgraph $G \setminus B_0^G(r)$ than on G itself; the next theorem shows this is not the case.

The following theorem studies the survival probability and expected ball sizes at p_c, and is the finite graph analogue of a theorem of Kozma and the second author [40]. The proof is almost identical to the one in [40] and is given explicitly in [30, 41].

Theorem 2.1 (Volume and survival probability [40]). *Let G be a finite transitive graph on V vertices such that the finite triangle condition (1.5) holds, and consider percolation on G at $p = p_c(\lambda)$ with any $\lambda > 0$. Then there exists a constant $C = C(\lambda) > 0$ such that for any $r > 0$,*

(1) $\mathbb{E}|B(r)| \leq Cr$,
(2) $\Gamma(r) \leq C/r$.

We often need to consider percolation performed at different values of p. We write $\mathbf{P}_p$ and $\mathbb{E}_p$ for the probability distribution and the corresponding expectation operator with parameter p when necessary. Furthermore, we sometimes need to consider percolation configurations at different p's on the same probability space. This is a standard procedure called the *simultaneous coupling* and it works as follows. For each edge e of our graph G, we draw an independent uniform random variable $U(e)$ in $[0, 1]$. We say that the edge e receives the value $U(e)$. For any $p \in [0, 1]$, the set of p-open edges is distributed precisely as $\{e : U(e) \leq p\}$. In this way, $G_{p_1} \subset G_{p_2}$ with probability 1 whenever $p_1 \leq p_2$.

2.2. Tails of the supercritical cluster size. We start by describing the tail of the cluster size in the supercritical regime. Note that the following theorem requires only the finite triangle condition, and not the stronger assumptions of Theorem 1.3, and so the restriction $\varepsilon = o(1/t_{\mathrm{mix}})$ is not needed.

Theorem 2.2 (Bounds on the cluster tail). *Let G be a finite transitive graph of degree m on V vertices such that the finite triangle condition (1.5) holds and set $p = p_c(1+\varepsilon)$ where $\varepsilon = o(1)$ and $\varepsilon \gg V^{-1/3}$. Then, for any k satisfying $k \gg \varepsilon^{-2}$,*

$$\mathbf{P}(|\mathscr{C}(0)| \geq k) \leq 2\varepsilon\big[1 + O\big(\varepsilon + (\varepsilon^3 V)^{-1} + (\varepsilon^2 k)^{-1/4} + \alpha_G\big)\big], \tag{2.2}$$

and, for the sequence $k_0 = \varepsilon^{-2}(\varepsilon^3 V)^{\alpha}$ for any $\alpha \in (0, 1/3)$, there exists a $c = c(\alpha) > 0$ such that

$$\mathbf{P}(|\mathscr{C}(0)| \geq k_0) \geq 2\varepsilon\big[1 + O\big(\varepsilon + (\varepsilon^3 V)^{-c} + \alpha_G\big)\big]. \tag{2.3}$$

Theorem 2.2 is reminiscent of the fact that a branching process with Poisson progeny distribution of mean $1+\varepsilon$ has survival probability $2\varepsilon(1+O(\varepsilon))$ when $\varepsilon = o(1)$. Upper and lower bounds of order ε for the cluster tail were proved already in [17] using Barsky and Aizenman's differential inequalities [9]. However, to get the precise constant 2 we need to sharpen these differential inequalities and handle some error terms in them that were neglected in the past. This derivation and the proof of Theorem 2.2 are presented in the Appendix. The proof is entirely self-contained.

Let $Z_{\geq k}$ denote the number of vertices with cluster size at least k, i.e.,

$$Z_{\geq k} = \big|\{v : |\mathscr{C}(v)| \geq k\}\big|. \tag{2.4}$$

We use Theorem 2.2 to show that $Z_{\geq k_0}$, with k_0 as in the theorem, is concentrated. This advances us towards the first factor on the left hand side of (1.8).

Lemma 2.3 (Concentration of $Z_{\geq k_0}$). *In the setting of Theorem 2.2, if $m \to \infty$, then*

$$\frac{Z_{\geq k_0}}{2\varepsilon V} \xrightarrow{\mathbf{P}} 1 \quad \textit{and} \quad \mathbb{E}|\mathscr{C}(0)| \leq (4+o(1))\varepsilon^2 V.$$

2.3. Many boundary edges between large clusters. The factor $(1-\theta\varepsilon/m)^{m\varepsilon^2 V}$ in (1.8) suggests that after partitioning the large clusters into two sets of order εV vertices, as we did before, the number of closed edges connecting them is of order $m\varepsilon^2 V$. This is the content of Theorem 2.4 below, which is the main effort of this paper. It is rather intuitive if one believes that large clusters are uniform random sets. Indeed, let v be a vertex in one of the sets of the partition. It has degree m and hence we expect $m\varepsilon$ of these neighbors to belong to the second set of the partition. Summing over all vertices v we obtain of the order $\varepsilon^2 m V$ edges. Making this a precise statement requires some details which we now provide.

We work under the general assumptions of Theorem 1.3. In particular, we are given sequences ε, α_G, both $o(1)$, such that $\varepsilon^3 V \to \infty$ and $\alpha_G \geq 1/m$. Without loss of generality we assume that

$$\alpha_G \geq (\varepsilon^3 V)^{-1/2}, \tag{2.5}$$

otherwise we replace the original α_G by $(\varepsilon(m)^3 V)^{-1/2}$, and note that in both cases $\alpha_G = o(1)$ and $\alpha_G \geq 1/m$.

Let us start by introducing some notation. For vertices x, y and radii j_x, j_y, we define the event

$$\mathscr{A}(x, y, j_x, j_y) = \big\{\partial B_x(j_x) \neq \emptyset,\ \partial B_y(j_y) \neq \emptyset \text{ and } B_x(j_x) \cap B_y(j_y) = \emptyset\big\}. \tag{2.6}$$

Intuitively, if $\mathscr{A}(x, y, j_x, j_y)$ occurs for j_x and j_y sufficiently large, then x and y are both in the giant component. The event $\mathscr{A}(x, y, j_x, j_y)$ plays a central role throughout our paper.

We continue by choosing some parameters. The role of each will become clear later. We set

$$M = M(m) = \log\log\log(\varepsilon^3 V \wedge \alpha_G^{-1} \wedge (\varepsilon t_{\mathrm{mix}})^{-1}), \qquad r = r(m) = M(m)\varepsilon^{-1}. \tag{2.7}$$

Note that $M(m) \to \infty$ in our setting. We choose $r_0 = r_0(m)$ to be

$$r_0 = \frac{\varepsilon^{-1}}{2} \log(\alpha_G \varepsilon^3 V). \tag{2.8}$$

It is important that r_0 is chosen so that $r_0 \gg r$. Additionally, in Corollary 4.6 we prove that $\mathbb{E}|B(j)| = \Theta(\varepsilon^{-1}(1+\varepsilon)^j)$ as long as $j \le \varepsilon^{-1}[\log(\varepsilon^3 V) - 4\log\log(\varepsilon^3 V)]$—the same asymptotics as in a Poisson $1+\varepsilon$ branching process (though in the branching process the estimate is valid for all j). This implies that $\mathbb{E}|B(r_0)| = \Theta\big(\sqrt{\alpha_G \varepsilon V}\big)$, a fact that we use throughout the paper, but not right now.

For vertices x, y we define

$$S_{2r+r_0}(x,y) = \Big|\Big\{(u,u') \in E(G) : \{x \overset{2r+r_0}{\longleftrightarrow} u\} \circ \{y \overset{2r+r_0}{\longleftrightarrow} u'\},$$
$$|B_u(2r+r_0)| \cdot |B_{u'}(2r+r_0)| \le \mathrm{e}^{40M} \varepsilon^{-2} (\mathbb{E}|B(r_0)|)^2\Big\}\Big|.$$

The edges counted in $S_{2r+r_0}(x,y)$ are the ones that will be used in the sprinkling. Informally, a pair (x,y) of vertices is *good* when their clusters are large and $S_{2r+r_0}(x,y)$ is large, so that their clusters have many bonds between them. We make this quantitative in the following definition:

Definition 2.1 ((r, r_0)-good pairs). We say that x, y are (r, r_0)-*good* if:

(1) $\mathscr{A}(x, y, 2r, 2r)$,
(2) $|\mathscr{C}(x)| \ge (\varepsilon^3 V)^{1/4} \varepsilon^{-2}$ and $|\mathscr{C}(y)| \ge (\varepsilon^3 V)^{1/4} \varepsilon^{-2}$, and
(3) $S_{2r+r_0}(x,y) \ge (\log M)^{-1} V^{-1} m \varepsilon^{-2} (\mathbb{E}|B(r_0)|)^2$.

Write P_{r,r_0} for the number of (r, r_0)-good pairs.

Theorem 2.4 (Most large clusters share many boundary edges). *Let G be a graph on V vertices and degree m satisfying the assumptions of Theorem 1.3. Assume that $\varepsilon = \varepsilon(m)$ satisfies*

$$\varepsilon \gg V^{-1/3} \quad \textit{and} \quad \varepsilon = o(1/t_{\mathrm{mix}}), \tag{2.9}$$

as in part (b) *of Theorem 1.3. Take M and $r = M\varepsilon^{-1}$ as in (2.7), and r_0 as in (2.8). Then*

$$\frac{P_{r,r_0}}{(2\varepsilon V)^2} \xrightarrow{\mathbf{P}} 1.$$

In light of Theorem 2.2, we expect that the number of pairs of vertices (x, y) with $|\mathscr{C}(x)| \ge (\varepsilon^3 V)^{1/4}\varepsilon^{-2}$ and $|\mathscr{C}(y)| \ge (\varepsilon^3 V)^{1/4}\varepsilon^{-2}$ is close to $(2\varepsilon V)^2$. Theorem 2.4 shows that the majority of these pairs have clusters that share many edges between them. This allows us to proceed with the sprinkling argument leading to (1.8), and we perform this in Section 7.1 leading to the proof of Theorem 1.3. Since the latter proof assumes only Theorem 2.4, the curious reader may now skip to Section 7.1 to see how this is done.

2.4. Uniform connection bounds and the role of the random walk. We briefly expand here on one of our most useful percolation inequalities and its connection with random walks. In the analysis of the Erdős–Rényi random graph $G(n, p)$, symmetry plays a special role. One instance of this symmetry is that the function $f(x) = \mathbf{P}(0 \leftrightarrow x)$ is constant whenever $x \ne 0$ and its value is precisely $(V-1)^{-1}(\mathbb{E}|\mathscr{C}(0)| - 1)$, and 1 when $x = 0$.

Such a statement clearly does not hold on the hypercube at p_c: the probability that two neighbors are connected is at least m^{-1} (recall (1.2)) and the probability that 0 is connected to one of the vertices in the barycenter of the cube is at most $\sqrt{m}\, 2^{-m}\mathbb{E}|\mathscr{C}(0)|$ by symmetry.

A key observation in our proof is that one can recover this symmetry as long as we require the connecting paths to be longer than the mixing time of the random walk. A precise statement is that percolation at p_c occurs on any graph G satisfying the assumptions of Theorem 1.3,

$$\mathbf{P}(0 \xleftrightarrow{[t_{\mathrm{mix}},\infty)} x) \le (1+o(1))\frac{\mathbb{E}|\mathscr{C}(0)|}{V}, \tag{2.10}$$

where t_{mix} is the uniform mixing time, as defined in Theorem 1.3. This is the content of Lemma 3.13 (or rather it follows by taking $r \to \infty$ in the lemma). In addition to allowing us to estimate difficult sums such as $\nabla_p(0,0)$ in (1.3) (see Section 3.6) and other similar quantities, this estimate also plays a key role in the high level idea of the proof, as we now explain.

2.5. Sketch of proof of Theorem 2.4. The difficulty in Theorem 2.4 is the requirement (3) in Definition 2.1. Indeed, conditioned on survival (that is, on the event $\mathscr{A}(x,y,2r,2r)$), the random variable $S_{2r+r_0}(x,y)$ is not concentrated, and hence it is hard to prove that it is large with high probability. In fact, even the variable $|B(r_0)|$ is not concentrated. This is not surprising: the number of descendants at generation n of a branching process with mean $\mu > 1$ divided by μ^n converges as $n \to \infty$ to a non-trivial random variable. Intuitively, this non-concentration occurs because the first generations of the process have a strong and lasting effect on the future of the population.

In order to counteract this, we condition on the event $\mathscr{A}(x,y,r,r)$ *and* on the entire balls $B_x(r)$ and $B_y(r)$ including all the open and closed edges touching them (during the actual proof we will use some other radii j_x, j_y between r and $2r$, but this is a technical matter). We will prove that given this conditioning the variable $S_{r+r_0}(x,y)$ is concentrated around the value

$$|\partial B_x(r)|\,|\partial B_y(r)|V^{-1}m(\mathbb{E}|B(r_0)|)^2, \tag{2.11}$$

and that $|\partial B_x(r)|\,|\partial B_y(r)| \ge \varepsilon^{-2}$ with high probability, implying that requirement (3) in Definition 2.1 occurs with high probability conditioned on the event above. Our choice of r_0 in (2.8) is made in such a way that the above quantity is large (however, later we will see that r_0 cannot be too large).

Let us elaborate on the estimate (2.11). Assume that $B_x(r) = A$ and $B_y(r) = B$ and write $\mathbf{P}_{A,B}$ and $\mathbb{E}_{A,B}$ for the conditional probability and expectation given $B_x(r) = A$ and $B_y(r) = B$. We have

$$\mathbb{E}_{A,B} S_{r+r_0}(x,y) \approx \sum_{a\in\partial A,\, b\in\partial B}\sum_{(u,u')}\mathbf{P}_{A,B}(\{a \xleftrightarrow{r_0} u\}\circ\{b \xleftrightarrow{r_0} u'\}),$$

where we do not write equality because (a) we have ignored the second condition in the definition of $S_{r+r_0}(x,y)$, on $|B_u(2r+r_0)|\cdot|B_{u'}(2r+r_0)|$; (b) some edges (u,u') may be over-counted in the sum; and (c) we have neglected to count the closed edges (u, u') that connect A and B (that is, occurring in height smaller than r). However, it turns out that all of these contributions are small compared to (2.11). It is a standard matter by now to use the triangle condition in order to deduce that for *most* edges (u, u'),

$$\mathbf{P}_{A,B}(\{a \overset{r_0}{\longleftrightarrow} u\} \circ \{b \overset{r_0}{\longleftrightarrow} u'\}) \approx \mathbf{P}_{A,B}(a \overset{r_0}{\longleftrightarrow} u)\mathbf{P}_{A,B}(b \overset{r_0}{\longleftrightarrow} u'),$$

so to proceed we need a good lower bound on $\mathbf{P}_{A,B}(a \overset{r_0}{\longleftrightarrow} u)$. The uniform connection bounds, that is, Lemma 3.13 or (2.10), immediately imply that $\mathbf{P}(a \overset{r_0}{\longleftrightarrow} u) \geq (1 - o(1))V^{-1}\mathbb{E}|B(r_0)|$ for *most* vertices u (since $\sum_u \mathbf{P}(a \overset{r_0}{\longleftrightarrow} u) = \mathbb{E}|B(r_0)|$). Had we had the same estimate for $\mathbf{P}_{A,B}(a \overset{r_0}{\longleftrightarrow} u)$, the lower bound on the conditional first moment required to prove the estimate (2.11) would follow immediately. However, the probability $\mathbf{P}_{A,B}(a \overset{r_0}{\longleftrightarrow} u)$ may heavily depend on the sets A and B.

To that end, in Section 5 we establish an intrinsic metric regularity theorem, similar in spirit to the extrinsic metric regularity theorem presented in [42]. Roughly, it states that for *most* sets A (more precisely, the weight of sets not having this is $o(\varepsilon)$) for which $B_x(r) = A$ satisfies $\partial B_x(r) \neq \emptyset$, *most* vertices $a \in \partial A$ satisfy

$$\sum_u \mathbf{P}_A(a \overset{r_0}{\longleftrightarrow} u) \geq (1 - o(1))\mathbb{E}|B(r_0)|,$$

where $\mathbf{P}_A$ is the conditional probability given $B_x(r) = A$. Thus, the expected size of the "future" of most vertices on the boundary is not affected by the conditioning on a typical "past".

At this point comes another crucial application of the uniform connection bounds as in (2.10). Indeed, even if the expected "future" of a vertex has the same asymptotics with or without conditioning, we cannot a priori rule out the possibility that this conditional "future" concentrates on a small remote portion of the underlying graph G—this can potentially violate the concentration around the value in (2.11). However, our uniform connection bounds stated in Lemma 3.13 are robust enough to deal with conditioning and immediately imply that $\mathbf{P}_A(a \overset{r_0}{\longleftrightarrow} u) = (1 - o(1))V^{-1}\mathbb{E}|B(r_0)|$ for most a in ∂A and for most vertices u. In other words, not only did the conditioning not influence the size of the "future", it also left its distribution approximately unaltered. These considerations allow us to give a lower bound of (2.11) on the conditional expectation. This and the conditional second moment calculation required to show concentration are performed in Section 6.

3. Preliminaries

In this section we provide some preliminary results that we will use. These involve various expectations and probabilities related to the random variable $|\partial B(r)|$ in Section 3.2 and 3.3, non-backtracking random walks in Section 3.4 and its relation to uniform bounds for connection probabilities in Section 3.5. In Section 3.6 we use these results to prove part (a) of Theorem 1.3. Finally, in Section 3.7 we bound triangle and square diagrams. The results in this section do *not* rely on the assumptions of Theorem 1.3 but sometimes we do assume the finite triangle condition (1.5).

3.1. The "off" method and BK-Reimer inequality. We will frequently handle the events $\partial B(r) \neq \emptyset$ and $x \overset{=t}{\longleftrightarrow} y$. These events are non-monotone with respect to adding edges; indeed, adding an edge may shorten a shortest path and prevent the events from holding. This non-monotonicity is a technical difficulty which unfortunately manifests itself in many of the arguments in this paper. Our main tools to deal with this problem are the BK-Reimer inequality [12, 53] and the notion of events occurring "off" a set of vertices. For the BK-Reimer inequality we use the formulation in [20].

For a subset A of vertices, we say that an event $\mathscr{M}$ occurs *off* A, intuitively, if it occurs in $G_p \setminus A$. Formally, for a percolation configuration ω, we write ω_A for the configuration obtained from ω by turning all the edges touching A to closed. The event "$\mathscr{M}$ occurs off A" is defined to be $\{\omega\colon \omega_A \in \mathscr{M}\}$. We also frequently write $\mathbf{P}_{\mathrm{off}\,A}$ to denote the measure $\mathbf{P}_{\mathrm{off}\,A}(\mathscr{M}) = \mathbf{P}_p(\mathscr{M} \text{ off } A)$. Equivalently, $\mathbf{P}_{\mathrm{off}\,A}$ can be thought of as a percolation measure in which all edges touching A are closed with probability 1 and the rest are distributed independently as before. We often drop p from the notation when it is clear what p is. This framework also allows us to address the case when $A = A(\omega)$ is a *random* set measurable with respect to G_p (the most prominent example is $A = B_0(r)$). In this case, the event $\{\mathscr{M} \text{ occurs off } A(\omega)\}$ is defined to be

$$\{\mathscr{M} \text{ occurs off } A(\omega)\} = \{\omega\colon \omega_{A(\omega)} \in \mathscr{M}\}.$$

Let us review an example occurring frequently in our arguments in which $\mathscr{M}$ is an arbitrary event and $A = B_x(s)$. In this case,

$$\mathbf{P}(\mathscr{M} \text{ off } B_x(s)) = \sum_A \mathbf{P}(B_x(s) = A)\mathbf{P}(\mathscr{M} \text{ off } A), \tag{3.1}$$

where we have used the fact that

$$\mathbf{P}(\mathscr{M} \text{ off } B_x(s) \mid B_x(s) = A) = \mathbf{P}(\mathscr{M} \text{ off } A),$$

since the events do not depend on edges touching A on both sides of the equation, and the marginal of the two distributions on the edges not touching A is the same product measure. In terms of this notation, for a subset A of vertices, we define

$$B_x^G(r; A) = \{y\colon d_{G_p}(x, y) \le r \text{ off } A\}, \quad \partial B_x^G(r; A) = \{y\colon d_{G_p}(x, y) = r \text{ off } A\}$$

to be the intrinsic ball off A, and its boundary. We finally say that $\mathscr{M}$ occurs *only on* A if $\mathscr{M}$ occurs but $\mathscr{M}$ off A does not occur. We frequently rely on the following inclusion:

Claim 3.1. *For any event $\mathscr{M}$ and any subset of vertices $A \subset V(G)$,*

$$\mathscr{M} \setminus \{\mathscr{M} \textit{ only on } A\} \subset \{\mathscr{M} \textit{ off } A\}.$$

Proof. By definition of "$\mathscr{M}$ only on A" the event on the left hand side equals

$$\mathscr{M} \cap \big\{\mathscr{M}^c \cup \{\mathscr{M} \text{ off } A\}\big\}.$$

From this, it is easy to see that this event implies $\mathscr{M}$ off A. □

Remark. Equality does not hold in Claim 3.1 (unless the right hand side is replaced by $\mathscr{M} \cap \{\mathscr{M} \text{ off } A\}$). This can easily be seen by taking a non-monotone event, say $\partial B_x(r) \neq \emptyset$.

The following lemmas are inequalities valid for any graph G and any p.

Lemma 3.2 (Disjoint survival). *For any graph G, $p \in [0,1]$, vertices x, y, z and integers r, s,*

$$\mathbf{P}\big(\partial B_x(r) \neq \emptyset, \partial B_y(s) \neq \emptyset, B_x(r) \cap B_y(s) = \emptyset\big) \\ \leq \mathbf{P}(\partial B_x(r) \neq \emptyset) \max_{A \subset V(G)} \mathbf{P}_{\mathrm{off}\, A}(\partial B_y(s) \neq \emptyset).$$

Proof. We condition on $B_x(r) = A$ such that $\partial B_x(r) \neq \emptyset$ and $\mathbf{P}(B_x(r) = A) > 0$. The left hand side is at most

$$\sum_{A:\, \partial B_x(r) \neq \emptyset} \mathbf{P}(B_x(r) = A)\mathbf{P}(\partial B_y(s) \neq \emptyset \text{ off } A),$$

as in (3.1). The lemma now follows. □

Lemma 3.3. *For any graph G, $p \in [0,1]$, vertices u, v and integers $r, \ell > 0$,*

$$\mathbf{P}(0 \overset{=r}{\longleftrightarrow} u \text{ and } 0 \overset{\ell}{\longleftrightarrow} v) \leq \sum_z \sum_{t=0}^{r} \mathbf{P}(0 \overset{=t}{\longleftrightarrow} z)\mathbf{P}(z \overset{\ell}{\longleftrightarrow} v) \max_{A \subset V(G)} \mathbf{P}_{\mathrm{off}\, A}(z \overset{=r-t}{\longleftrightarrow} u).$$

Proof. We claim that if $0 \overset{=r}{\longleftrightarrow} u$ and $0 \overset{\ell}{\longleftrightarrow} v$, then there exist $z \in V(G)$ and $t \leq r$ such that the following two events occur disjointly:

(a) there exists a shortest open path η of length r between 0 and u such that $\eta(t) = z$;
(b) there exists an open path between z and ℓ of length at most ℓ.

Indeed, if the event occurs, let η be the lexicographical first shortest path of length r between 0 and u, and let γ be an open path of length at most ℓ between 0 and v. We take z to be the last vertex on γ which belongs to η, and choose t such that $\eta(t) = z$. The witness for the first event is the set of open edges of η together with all the closed edges in G_p (the closed edges determine that η is a shortest path), and the second witness is the edges of γ from z to v. These are disjoint witnesses so we may use the BK-Reimer inequality and bound

$$\mathbf{P}(0 \overset{=r}{\longleftrightarrow} u \text{ and } 0 \overset{\ell}{\longleftrightarrow} v) \leq \sum_{z \in V(G),\, t \leq r} \mathbf{P}((a))\mathbf{P}(z \overset{\ell}{\longleftrightarrow} v).$$

To bound $\mathbf{P}((\mathrm{a}))$ we condition on $B_0(t) = A$ such that A satisfies $0 \overset{=t}{\longleftrightarrow} z$, so

$$\mathbf{P}((\mathrm{a})) = \sum_{A:\, 0 \overset{=t}{\longleftrightarrow} z} \mathbf{P}(B_0(t) = A)\mathbf{P}(z \overset{=r-t}{\longleftrightarrow} u \text{ off } A),$$

and the lemma follows. □

3.2. Survival probabilities. In this section, we prove Lemma 2.3 and a few other useful estimates of a similar nature. In the rest of this section we only rely on the finite triangle condition (1.5), Theorem 2.1 and Theorem 2.2 (both of which follow from the triangle condition, as shown in [40] and the Appendix).

Lemma 3.4 (Relating connection probabilities for different p's). *Let $p_1, p_2 \in [0,1]$ satisfy $p_1 \le p_2$ and let $r > 0$ be an integer. Then, for any graph G and vertex v:*

(1) $\mathbf{P}_{p_2}(\partial B_v(r) \neq \emptyset) \le (p_2/p_1)^r \mathbf{P}_{p_1}(\partial B_v(r) \neq \emptyset)$,
(2) $\mathbb{E}_{p_2}|\partial B_v(r)| \le (p_2/p_1)^r \mathbb{E}_{p_1}|\partial B_v(r)|$.

Proof. We recall the standard simultaneous coupling between percolation measure at different p's discussed in Section 2.1. Order all the paths in G of length r starting at v lexicographically. Write $\mathscr{A}$ for the event that $\partial B_v(r) \neq \emptyset$ in G_{p_2} and that the lexicographically first p_2-open shortest path of length r starting at v is in fact p_1-open. We claim that

$$\mathbf{P}(\mathscr{A}) = (p_1/p_2)^r \mathbf{P}(\partial B_v(r) \neq \emptyset \text{ in } G_{p_2}). \tag{3.2}$$

Indeed, conditioned on the edges of G_{p_2}, the value $U(e)$ of each edge in G_{p_2} is distributed uniformly on the interval $[0, p_2]$. Hence, the probability of the first shortest path being p_1-open in this conditioning is precisely $(p_1/p_2)^r$, which proves (3.2). To see (1), note that if the first p_2-open shortest path is p_1-open, then it is a shortest path of length r in G_{p_1}, so that $\mathscr{A}$ implies that $\partial B_v(r) \neq \emptyset$ in G_{p_1}, whence

$$\mathbf{P}_{p_1}(\partial B_v(r) \neq \emptyset) \ge (p_1/p_2)^r \mathbf{P}_{p_2}(\partial B_v(r) \neq \emptyset).$$

The proof of (2) is similar and we omit the details. □

Corollary 3.5 (Correlation length is $1/\varepsilon$). *Let G be a transitive finite graph for which (1.5) holds and set $p = p_c(1+\varepsilon)$. Then, for any subset A of vertices, any vertex v and any integer r:*

(1) $\mathbf{P}_{\text{off } A}(\partial B_v(\varepsilon^{-1}) \neq \emptyset) = O(\varepsilon)$,
(2) $\mathbb{E}|B_v(r; A)| = O(r(1+\varepsilon)^r)$.

Proof. The result is immediate by combining Lemma 3.4 and Theorem 2.1. □

Remark. In Section 4 we will show a sharp estimate replacing (2) in the above corollary.

Lemma 3.6 (Supercritical survival probability). *Let G be a transitive finite graph for which (1.5) holds and set $p = p_c(1+\varepsilon)$. Then, for any $M \to \infty$ and any subset A of vertices,*

$$\mathbf{P}_{\text{off } A}(\partial B(M\varepsilon^{-1}) \neq \emptyset) \le (2 + o(1))\varepsilon,$$

and for any $M \le \log\log(\varepsilon^3 V)$ such that $M \to \infty$,

$$\mathbf{P}(\partial B(M\varepsilon^{-1}) \neq \emptyset) \ge (2 - o(1))\varepsilon.$$

Proof. To prove the upper bound we write

$$\begin{aligned}\mathbf{P}_{\text{off } A}(\partial B(M\varepsilon^{-1}) \neq \emptyset) = {}& \mathbf{P}_{\text{off } A}\big(\partial B(M\varepsilon^{-1}) \neq \emptyset,\ |\mathscr{C}(0)| > \sqrt{M}\,\varepsilon^{-2}\big)\\ &+ \mathbf{P}_{\text{off } A}\big(\partial B(M\varepsilon^{-1}) \neq \emptyset,\ |\mathscr{C}(0)| \le \sqrt{M}\,\varepsilon^{-2}\big).\end{aligned}$$

The first term on the right hand side is at most $(2 + o(1))\varepsilon$ by Theorem 2.2—note that we have used the fact that the event $|\mathscr{C}(0)| \ge k$ is monotone, so $\mathbf{P}_{\text{off } A}(|\mathscr{C}(0)| \ge k) \le$

$\mathbf{P}(|\mathscr{C}(0)| \geq k)$. It remains to show that the second term is $o(\varepsilon)$. Indeed, if this event occurs, then there exists a radius $j \in [M\varepsilon^{-1}/3, 2M\varepsilon^{-1}/3]$ such that

$$0 < \partial B(j; A) \leq 3M^{-1/2}\varepsilon^{-1} \quad \text{and} \quad \partial B(M\varepsilon^{-1}; A) \neq \emptyset.$$

Let J be the first level satisfying this. By Corollary 3.5 and the union bound,

$$\begin{aligned}\mathbf{P}_{\mathrm{off}\,A}\big(\exists y \in \partial B(J; A) \text{ with } \partial B_y(M\varepsilon^{-1}/3; A) \neq \emptyset \text{ off } B(J; A) \;\big|\; B(J; A)\big) \\ \leq C\varepsilon|\partial B(J; A)| = O(M^{-1/2}).\end{aligned}$$

Corollary 3.5 also shows that $\mathbf{P}_{\mathrm{off}\,A}(\partial B(M\varepsilon^{-1}/3) \neq \emptyset) = O(\varepsilon)$, so putting this together gives

$$\mathbf{P}_{\mathrm{off}\,A}\big(\partial B(M\varepsilon^{-1}) \neq \emptyset,\ |\mathscr{C}(x)| \leq \sqrt{M}\,\varepsilon^{-2}\big) = O(M^{-1/2}\varepsilon), \tag{3.3}$$

concluding the proof of the upper bound. For the lower bound, take $k_0 = \varepsilon^{-2}(\varepsilon^3 V)^{\alpha}$ for some fixed $\alpha \in (0, 1/3)$. We have

$$\begin{aligned}\mathbf{P}(\partial B(M\varepsilon^{-1}) \neq \emptyset) &\geq \mathbf{P}\big(\partial B(M\varepsilon^{-1}) \neq \emptyset \text{ and } |\mathscr{C}(0)| \geq k_0\big) \\ &= \mathbf{P}(|\mathscr{C}(0)| \geq k_0) - \mathbf{P}\big(|\mathscr{C}(0)| \geq k_0 \text{ and } \partial B(M\varepsilon^{-1}) = \emptyset\big),\end{aligned}$$

so by Theorem 2.2 it suffices to bound the last term on the right hand side from above. Indeed, by Markov's inequality and Corollary 3.5,

$$\begin{aligned}\mathbf{P}\big(|\mathscr{C}(0)| \geq k_0 \text{ and } \partial B(M\varepsilon^{-1}) = \emptyset\big) &\leq \mathbf{P}(|B(M\varepsilon^{-1})| \geq k_0) \\ &\leq \frac{CM\mathrm{e}^{M}\varepsilon^{-1}}{k_0} = O\big(\varepsilon(\varepsilon^3 V)^{-\alpha}\log(\varepsilon^3 V)\big) = o(\varepsilon),\end{aligned} \tag{3.4}$$

since $M \leq \log\log(\varepsilon^3 V)$. □

We proceed with the preparations towards the proof of Lemma 2.3. For an integer $r > 0$, we write N_r for the random variable

$$N_r = |\{x : \partial B_x(r) \neq \emptyset\}|.$$

We think of $1/\varepsilon$ as the correlation length (see [27]). In other words, if $r \gg 1/\varepsilon$, then the vertices contributing to N_r should be those in the giant component.

Lemma 3.7 (N_r is concentrated). *Let G_m be a sequence of transitive finite graphs with degree m for which* (1.5) *holds and $m \to \infty$. Set $p = p_c(1 + \varepsilon)$. Then, for any $r \gg 1/\varepsilon$ satisfying $r \leq \varepsilon^{-1}\log\log(\varepsilon^3 V)$,*

$$\frac{N_r}{2\varepsilon V} \xrightarrow{\mathbf{P}} 1.$$

Proof. We use a second moment method on N_r. Lemma 3.6 and our assumption on r show that

$$\mathbb{E}N_r = (2+o(1))\varepsilon V.$$

The second moment is

$$\mathbb{E}N_r^2 = \sum_{x,y} \mathbf{P}\big(\partial B_x(r) \neq \emptyset \text{ and } \partial B_y(r) \neq \emptyset\big).$$

We have

$$\mathbf{P}\big(\partial B_x(r) \neq \emptyset \text{ and } \partial B_y(r) \neq \emptyset\big) \leq \mathbf{P}\big(\partial B_x(r) \neq \emptyset \text{ and } \partial B_y(r) \neq \emptyset,\ B_x(r) \cap B_y(r) = \emptyset\big) + \mathbf{P}(x \overset{2r}{\leftrightarrow} y).$$

We sum the first term on the right hand side using Lemmas 3.2 and 3.6 and the second term using Corollary 3.5. We get

$$\mathbb{E}N_r^2 \leq (4+o(1))\varepsilon^2 V^2 + O(Vr(1+\varepsilon)^{2r}) = (1+o(1))[\mathbb{E}N_r]^2,$$

since $Vr(1+\varepsilon)^{2r} = o(\varepsilon^2 V^2)$ by our assumption on r and since $\varepsilon^3 V \to \infty$. The assertion of the lemma now follows by Chebyshev's inequality. □

Proof of Lemma 2.3. Take $M = M(m)$ and r as in (2.7) and write

$$Z_{\geq k_0} = N_r + \big|\{x : \partial B_x(r) = \emptyset,\ |\mathscr{C}(x)| \geq k_0\}\big| - \big|\{x : \partial B_x(r) \neq \emptyset,\ |\mathscr{C}(x)| < k_0\}\big|.$$

By Lemma 3.7, $N_r/(2\varepsilon V) \xrightarrow{\mathbf{P}} 1$, so it suffices to show that the expectation of both remaining terms is $o(\varepsilon V)$. The expectation of the first term is

$$V\mathbf{P}(x : \partial B_x(r) = \emptyset,\ |\mathscr{C}(v)| \geq k_0) = o(\varepsilon V), \tag{3.5}$$

by (3.4). The expectation of the second term now must be $o(\varepsilon V)$ since both N_r and $Z_{\geq k_0}$ have mean $(2+o(1))\varepsilon V$ by Theorem 2.2 and Lemma 3.6. This shows that $Z_{\geq k_0}/(2\varepsilon V) \xrightarrow{\mathbf{P}} 1$ as stated in Lemma 2.3.

To prove the upper bound on $\mathbb{E}|\mathscr{C}(0)|$ we write

$$\mathbb{E}|\mathscr{C}(0)| = \sum_y \mathbf{P}(0 \leftrightarrow y) = \sum_y \mathbf{P}(0 \xleftrightarrow{[0,2r)} y) + \sum_y \mathbf{P}(0 \xleftrightarrow{[2r,\infty)} y).$$

By Corollary 3.5,

$$\sum_y \mathbf{P}(0 \xleftrightarrow{[0,2r)} y) = \mathbb{E}|B(2r)| \leq C\varepsilon^{-1}(\log(\varepsilon^3 V))^3 = o(\varepsilon^2 V), \tag{3.6}$$

since $\varepsilon^3 V \gg 1$. If $0 \xleftrightarrow{[2r,\infty)} y$, then the event

$$\big\{\partial B_0(r) \neq \emptyset,\ \partial B_y(r) \neq \emptyset,\ B_0(r) \cap B_y(r) = \emptyset\big\}$$

occurs. Hence Lemmas 3.2 and 3.6 give $\mathbf{P}(0 \xleftrightarrow{[2r,\infty)} y) \leq (4+o(1))\varepsilon^2$, and summing this over y yields the required upper bound on $\mathbb{E}|\mathscr{C}(0)|$. □

3.3. Disjoint survival probabilities. In this section we show that for most pairs x, y the event $\mathscr{A}(x, y, r, r)$ occurs with probability asymptotic to $4\varepsilon^2$. The point is that r is chosen such that $r \gg \varepsilon^{-1}$, where ε^{-1} is the correlation length, but $r \ll \varepsilon^{-1}\log(\varepsilon^3 V)$, which is the order of the diameter of $\mathscr{C}_1$ (we do not prove the estimate on the diameter here, but it can be obtained using the techniques of this paper).

Lemma 3.8 (Number of pairs surviving disjointly). *Let G_m be a sequence of transitive finite graphs with degree m for which (1.5) holds and $m \to \infty$. Set $p = p_c(1+\varepsilon)$. Then, for any $r \gg \varepsilon^{-1}$ satisfying $r \le \varepsilon^{-1}\log\log(\varepsilon^3 V)$,*

$$\frac{|\{\{x, y\}: \mathscr{A}(x, y, r, r) \text{ occurs}\}|}{(2\varepsilon V)^2} \xrightarrow{\mathbf{P}} 1.$$

Proof. Define

$$N_r^{(2)} = |\{\{x, y\}: \mathscr{A}(x, y, r, r) \text{ occurs}\}|.$$

Then

$$N_r^2 - |\{\{x, y\}: x \overset{2r}{\longleftrightarrow} y\}| \le N_r^{(2)} \le N_r^2,$$

and $\mathbb{E}|\{\{x, y\}: x \overset{2r}{\longleftrightarrow} y\}| = o(\varepsilon^2 V^2)$ as shown in (3.6). The result now follows from Markov's inequality and Lemma 3.7. □

Lemma 3.9 (Most pairs have almost independent disjoint survival probabilities). *Let G_m be a sequence of transitive finite graphs with degree m for which (1.5) holds and $m \to \infty$. Set $p = p_c(1+\varepsilon)$. Then, for any $j_x, j_y \le \varepsilon^{-1}\log\log(\varepsilon^3 V)$ such that $j_x, j_y \gg \varepsilon^{-1}$, there exist at least $(1-o(1))V^2$ pairs of vertices x, y such that*

$$\mathbf{P}(\mathscr{A}(x, y, j_x, j_y)) = (1+o(1))4\varepsilon^2.$$

Proof. The upper bound $\mathbf{P}(\mathscr{A}(x, y, j_x, j_x)) \le (1+o(1))4\varepsilon^2$ follows immediately from Lemmas 3.2 and 3.6 and is valid for all pairs x, y. We turn to showing the corresponding lower bound. First, the inequality $\mathbb{E}[N_r^2] \ge (\mathbb{E}N_r)^2$ can be written as

$$\sum_{x,y} \mathbf{P}\big(\partial B_x(r) \ne \emptyset \text{ and } \partial B_y(r) \ne \emptyset\big) \ge V^2 \mathbf{P}(\partial B(r) \ne \emptyset)^2.$$

We take $r = \varepsilon^{-1}\log\log(\varepsilon^3 V)$. Since $\mathbf{P}(\partial B_x(j) \ne \emptyset)$ is decreasing in j, by Lemma 3.6 and our assumption on j_x and j_y we get

$$\sum_{x,y} \mathbf{P}\big(\partial B_x(j_x) \ne \emptyset \text{ and } \partial B_y(j_y) \ne \emptyset\big) \ge (4-o(1))V^2\varepsilon^2.$$

Secondly, Corollary 3.5 implies that

$$\sum_{x,y} \mathbf{P}(x \overset{2r}{\longleftrightarrow} y) = V\mathbb{E}|B(2r)| \le CVr(1+\varepsilon)^{2r} = o(\varepsilon^2 V^2),$$

by our choice of r and since $\varepsilon^3 V \to \infty$. Since

$$\mathscr{A}(x, y, j_x, j_y) \subseteq \{\partial B_x(j_x) \ne \emptyset,\ \partial B_y(j_y) \ne \emptyset\} \setminus \{x \overset{2r}{\longleftrightarrow} y\},$$

we deduce that

$$\sum_{x,y} \mathbf{P}(\mathscr{A}(x, y, j_x, j_y)) \ge (4 - o(1))\varepsilon^2 V^2,$$

and since the upper bound was valid for all x, y, the lemma follows. □

Lemma 3.10. *Let G_m be a sequence of transitive finite graphs with degree m for which (1.5) holds and $m \to \infty$. Set $p = p_c(1+\varepsilon)$. Then, for any $r \le \varepsilon^{-1} \log\log(\varepsilon^3 V)$ such that $r \gg \varepsilon^{-1}$,*

$$\frac{\left|\{(x, y): \mathscr{A}(x, y, r, r) \text{ and } |\mathscr{C}(x)| \le (\varepsilon^3 V)^{1/4}\varepsilon^{-2}\}\right|}{(\varepsilon V)^2} \xrightarrow{\mathbf{P}} 0.$$

Proof. By Lemma 3.7, the assertion follows from the statement that

$$\left|\{x: \partial B_x(r) \ne \emptyset \text{ and } |\mathscr{C}(x)| \le (\varepsilon^3 V)^{1/4}\varepsilon^{-2}\}\right| / (\varepsilon V) \xrightarrow{\mathbf{P}} 0.$$

To show this, note that

$$\begin{aligned}\mathbf{P}\big(\partial B_x(r) = \emptyset \text{ and } |\mathscr{C}(x)| \ge (\varepsilon^3 V)^{1/4}\varepsilon^{-2}\big) &\le \mathbf{P}\big(|B_x(r)| \ge (\varepsilon^3 V)^{1/4}\varepsilon^{-2}\big)\\ &\le \frac{Cr(1+\varepsilon)^r}{(\varepsilon^3 V)^{1/4}\varepsilon^{-2}} = o(\varepsilon).\end{aligned}$$

Hence, by Theorem 2.2,

$$\mathbf{P}\big(\partial B_x(r) \ne \emptyset \text{ and } |\mathscr{C}(x)| \ge (\varepsilon^3 V)^{1/4}\varepsilon^{-2}\big) \ge (2 - o(1))\varepsilon.$$

Together with Lemma 3.6, this yields

$$\mathbf{P}\big(\partial B_x(r) \ne \emptyset \text{ and } |\mathscr{C}(x)| \le (\varepsilon^3 V)^{1/4}\varepsilon^{-2}\big) = o(\varepsilon),$$

concluding our proof. □

3.4. Using the non-backtracking random walk. In the rest of this section we provide several basic percolation estimates which we use throughout the paper. These include bounds on long and short connection probabilities and bounds on various triangle and square diagrams. It is here that we make crucial use of the geometry of the graph and of the behavior of the random walk on it, namely, the assumptions of Theorem 1.3. We frequently use *non-backtracking* random walk estimates. Such a walk is a simple random walk on a graph that is not allowed to traverse back on an edge it has just walked on. Let us first define it formally.

The *non-backtracking random walk* on an undirected graph $G = (V(G), E(G))$, starting from a vertex $x \in V(G)$, is a Markov chain $\{X_t\}$ with transition matrix $\mathbf{P}^x$ on the state space of *directed* edges

$$\vec{E} = \{(x, y): \{x, y\} \in E(G)\}.$$

If $X_t = (x, y)$, then we write $X_t^{(1)} = x$ and $X_t^{(2)} = y$. Also, for notational convenience, we write

$$\mathbf{P}_{(x,w)}(\cdot) = \mathbf{P}^x(\cdot \mid X_0 = (x, w)) \quad \text{and} \quad \mathbf{p}^t(x, y) = \mathbf{P}^x(X_t^{(2)} = y).$$

The non-backtracking walk starting from a vertex x has initial state given by

$$\mathbf{P}^x(X_0 = (x, y)) = \mathbf{1}_{\{(x,y)\in \vec{E}\}} \frac{1}{\deg(x)},$$

and transition probabilities given by

$$\mathbf{P}_{(u,v)}(X_1 = (v, w)) = \mathbf{1}_{\{(v,w)\in \vec{E},\, w\neq v\}} \frac{1}{\deg(v) - 1},$$

where we write $\deg(x)$ for the degree of x in G. The following lemma will be useful.

Lemma 3.11. *Let G be a regular graph of degree m. Then, for $t \geq 0$,*

$$\frac{m}{m-1} \sum_u \mathbf{p}^1(0, u)\mathbf{p}^t(u, z) = \mathbf{p}^{t+1}(0, z) + \frac{1}{m-1}\mathbf{p}^{t-1}(0, z),$$

with the convention that $\mathbf{p}^{-1}(0, z) = 0$.

Proof. For any neighbour u of 0 we have $\mathbf{p}^1(0, u) = m^{-1}$, hence

$$\sum_u \mathbf{p}^1(0, u)\mathbf{p}^t(u, z) = m^{-1} \sum_{u:\, u\sim 0} \Bigg[\sum_{\omega} \frac{1}{m(m-1)^{t-1}} + \sum_{\omega'} \frac{1}{m(m-1)^{t-1}} \Bigg],$$

where ω runs over the non-backtracking paths of length t from u to z such that the first step is *not* 0, and ω' runs over the same paths such that the first step *is* 0. Similarly

$$\mathbf{p}^{t+1}(0, z) = m^{-1} \sum_{u:\, u\sim 0} \sum_{\omega} \frac{1}{(m-1)^t},$$

where ω runs over the same paths as above. Thus

$$\Bigg[\frac{m}{m-1} \sum_u \mathbf{p}^1(0, u)\mathbf{p}^t(u, z) \Bigg] - \mathbf{p}^{t+1}(0, z) = m^{-1} \sum_{u:\, u\sim 0} \sum_{\omega'} \frac{1}{(m-1)^t}.$$

By changing the order of summation the last sum equals

$$m^{-1} \sum_{\omega''} \sum_{u:\, u\sim 0,\, u\neq \omega''_1} (m-1)^{-t},$$

where ω'' runs over non-backtracking paths of length $t-1$ from 0 to z. For each ω'' the number of summands in the second sum is precisely $m-1$, so that the sum equals

$$\sum_{\omega''} m^{-1}(m-1)^{-(t-1)},$$

which is precisely $(m-1)^{-1}\mathbf{p}^{t-1}(0, z)$ as required. □

3.5. Uniform upper bounds on long connection probabilities. In this section we show that long percolation paths are asymptotically equally likely to end at any vertex in G.

Lemma 3.12. *Let G be a graph satisfying the assumptions in Theorem 1.3 and consider percolation on it with $p \le p_c(1+\varepsilon)$. Then, for any integer $t \ge t_{\mathrm{mix}}$ and any vertex x,*

$$\mathbf{P}_p(0 \overset{=t}{\longleftrightarrow} x) + \mathbf{P}_p(0 \overset{=t+1}{\longleftrightarrow} x) \le \frac{2 + O(\alpha_G + \varepsilon t_{\mathrm{mix}})}{V} \mathbb{E}|\partial B(t - t_{\mathrm{mix}})|.$$

Proof. If $0 \overset{=t}{\longleftrightarrow} x$, then there exists a vertex v and a simple path ω of length t_{mix} from x to v such that the event

$$\{\omega \text{ is open}\} \circ \{v \overset{=t-t_{\mathrm{mix}}}{\longleftrightarrow} x\}$$

occurs. Indeed, consider a shortest path η of length t between 0 to x. Take $v = \eta(t_{\mathrm{mix}})$ and $\omega = \eta[1, t_{\mathrm{mix}}]$. Now, the first witness is the path ω and the witness for $\{v \overset{=t-t_{\mathrm{mix}}}{\longleftrightarrow} x\}$ is the path $\eta[t_{\mathrm{mix}}, t]$ together with all the closed edges in G_p (which determine that $\eta[t_{\mathrm{mix}}, t]$ is a shortest path). These are disjoint witnesses. If $0 \overset{=t+1}{\longleftrightarrow} x$ occurs, then we get the same conclusion with ω of length $t_{\mathrm{mix}} + 1$. We now apply the BK-Reimer inequality and the fact that the probability that ω is open is precisely $p^{|\omega|}$. This yields

$$\mathbf{P}_p(0 \overset{=t}{\longleftrightarrow} x) \le p^{t_{\mathrm{mix}}} \sum_v \sum_{\substack{\omega\colon |\omega| = t_{\mathrm{mix}} \\ \omega[t_{\mathrm{mix}}] = v}} \mathbf{P}_p(v \overset{=t-t_{\mathrm{mix}}}{\longleftrightarrow} x),$$

and

$$\mathbf{P}_p(0 \overset{=t+1}{\longleftrightarrow} x) \le p^{t_{\mathrm{mix}}+1} \sum_v \sum_{\substack{\omega\colon |\omega| = t_{\mathrm{mix}}+1 \\ \omega[t_{\mathrm{mix}}+1] = v}} \mathbf{P}_p(v \overset{=t-t_{\mathrm{mix}}}{\longleftrightarrow} x).$$

We now bound these by relaxing the requirement that ω is simple and only requiring that it is non-backtracking. Since $t_{\mathrm{mix}} = T_{\mathrm{mix}}(\alpha_G)$, we find by definition that

$$\begin{aligned} &\frac{\big|\{\omega\colon |\omega| = t_{\mathrm{mix}}, \omega[t_{\mathrm{mix}}] = v\}\big|}{m(m-1)^{t_{\mathrm{mix}}-1}} + \frac{\big|\{\omega\colon |\omega| = t_{\mathrm{mix}}+1, \omega[t_{\mathrm{mix}}+1] = v\}\big|}{m(m-1)^{t_{\mathrm{mix}}}} \\ &\qquad = \mathbf{p}^{t_{\mathrm{mix}}}(0, v) + \mathbf{p}^{t_{\mathrm{mix}}+1}(0, v) \le \frac{2 + 2\alpha_G}{V}, \end{aligned}$$

where we have enumerated only non-backtracking paths. Using this, condition (2) and summing over v gives

$$\begin{aligned} \mathbf{P}_p(0 \overset{=t}{\longleftrightarrow} x) + \mathbf{P}_p(0 \overset{=t+1}{\longleftrightarrow} x) &\le \frac{2 + O(\alpha_G)}{V} [p(m-1)]^{t_{\mathrm{mix}}} \mathbb{E}|\partial B(t - t_{\mathrm{mix}})| \\ &\le \frac{2 + O(\alpha_G + \varepsilon t_{\mathrm{mix}})}{V} \mathbb{E}|\partial B(t - t_{\mathrm{mix}})|, \end{aligned}$$

concluding our proof. ◻

Lemma 3.13. *Let G be a graph satisfying the assumptions in Theorem 1.3 and consider percolation on it with $p \le p_c(1+\varepsilon)$. Then, for any $r \ge t_{\rm mix}$ and any vertex x,*

$$\mathbf{P}_p(0 \xleftrightarrow{P[t_{\rm mix},r]} x) \le \frac{1 + O(\alpha_G + \varepsilon t_{\rm mix})}{V} \mathbb{E}|B(r)|.$$

Proof. The proof is similar to that of Lemma 3.12. If the event occurs, then there exists a vertex v and a simple path ω of length $t_{\rm mix}$ from 0 to v such that the event

$$\{\omega \text{ is open}\} \circ \{v \overset{r}{\longleftrightarrow} x\}$$

occurs. Hence,

$$\mathbf{P}_p(0 \xleftrightarrow{P[t_{\rm mix},r]} x) \le p^{t_{\rm mix}} \sum_v \sum_{\substack{\omega:\, |\omega| = t_{\rm mix} \\ \omega[t_{\rm mix}] = v}} \mathbf{P}_p(v \overset{r}{\longleftrightarrow} x),$$

by the BK inequality. By the same argument,

$$\mathbf{P}_p(0 \xleftrightarrow{P[t_{\rm mix},r]} x) \le p^{t_{\rm mix}+1} \sum_v \sum_{\substack{\omega:\, |\omega| = t_{\rm mix}+1 \\ \omega[t_{\rm mix}+1] = v}} \mathbf{P}_p(v \overset{r}{\longleftrightarrow} x).$$

The reason we make two such similar estimates is that due to possible periodicity, in each of the estimates the sum over v may be 0 on half the vertices. We now take the average of these two estimates, sum over v to get the $\mathbb{E}|B(r)|$ factor and use the same analysis as in Lemma 3.12 using condition (2). This gives the required assertion of the lemma. □

Lemma 3.14. *Let G be a graph satisfying the assumptions in Theorem 1.3 and consider percolation on it with $p \le p_c(1+\varepsilon)$. Then, for any $r \ge t_{\rm mix}$ and any vertex x,*

$$\mathbf{P}(0 \overset{[r,2r]}{\longleftrightarrow} x) \le \frac{1 + O(\alpha_G + \varepsilon t_{\rm mix})}{V} \mathbb{E}|B([r - t_{\rm mix}, 2r - t_{\rm mix}])|.$$

Proof. This follows by summing the estimate of Lemma 3.12 and using the fact that

$$\mathbb{E}|\partial B(t)| \le p(m-1)\mathbb{E}|\partial B(t-1)| \le (1 + m^{-1} + \varepsilon)\mathbb{E}|\partial B(t-1)|,$$

where the last inequality is due to condition (2) and the first inequality holds since, given $\partial B(t-1)$, $|\partial B(t)|$ is stochastically bounded above by a sum of $|\partial B(t-1)|$ binomial random variables with parameters $m-1$ and p. The lemma follows since $\alpha_G \ge m^{-1}$. □

We close this section with a remark to be used later. We would often like to have these uniform connection bounds off some subsets of vertices. The proofs of the lemmas in this section immediately generalize to such a setting, because the claim "the number of paths from 0 to v of length $t_{\rm mix}$ is at most n" still holds even if we are in $G \setminus A$, for any subset A of vertices. We state the required assertion here and omit the proof:

Lemma 3.15 (Uniform connection bounds off sets). *Consider percolation on $G = \{0,1\}^m$ with $p = p_c(1+\varepsilon)$, and let A be any subset of vertices. Then, for any $r \ge t_{\rm mix}$ and any vertex x,*

$$\mathbf{P}_{\text{off}\,A}(0 \overset{=r}{\longleftrightarrow} x) \le (2 + O(\alpha_G + \varepsilon t_{\rm mix}))V^{-1}\mathbb{E}|\partial B(r - t_{\rm mix}; A)|,$$

$$\mathbf{P}_{\text{off}\,A}(0 \xleftrightarrow{[t_{\rm mix},r]} x) \le (1 + O(\alpha_G + \varepsilon t_{\rm mix}))V^{-1}\mathbb{E}|B(r; A)|.$$

3.6. Proof of part (a) of Theorem 1.3. We demonstrate the use of Lemma 3.13 by showing that the finite triangle condition holds under the assumptions of Theorem 1.3. We begin with an easy calculation.

Claim 3.16. *On any regular graph G of degree m, for vertices x, y,*

$$\sum_{u,v}\sum_{t_1,t_2,t_3:\,t_1+t_2+t_3\in\{0,1,2\}} \mathbf{p}^{t_1}(x,u)\mathbf{p}^{t_2}(u,v)\mathbf{p}^{t_3}(v,y) = \mathbf{1}_{\{x=y\}} + O(m^{-1}),$$
$$\sum_{u,v}\sum_{t_1,t_2,t_3:\,t_1+t_2+t_3\in\{1,2\}} \mathbf{p}^{t_1}(x,u)\mathbf{p}^{t_2}(u,v)\mathbf{p}^{t_3}(v,y) = O(m^{-1}).$$

Proof. We prove both statements simultaneously. The contribution coming from $t_1 + t_2 + t_3 = 0$ is the one we get when $x = u = v = y$, giving $\mathbf{1}_{\{x=y\}}$. The contributions from $t_1 + t_2 + t_3 = 1$ can only come from the cases $u = v = y$ and $(t_1, t_2, t_3) = (1, 0, 0)$, or $u = v = x$ and $(t_1, t_2, t_3) = (0, 0, 1)$, or $u = x$ and $v = y$ and $(t_1, t_2, t_3) = (0, 1, 0)$. These are easily bounded using the fact that $\max_z \mathbf{p}^t(w, z) \le 1/(m-1)$ for any $t \ge 1$. We perform a similar case analysis to bound the contributions from $t_1 + t_2 + t_3 = 2$. If $(t_1, t_2, t_3) = (0, 0, 2)$, then we must have $u = v = x$ and $\mathbf{p}^2(v, y) = O(m^{-1})$; this argument also handles the case where one of the other t_i's is 2. In the case $(t_1, t_2, t_3) = (1, 1, 0)$ we must have $v = y$ and u is a neighbor of both x and y. There are at most m such u's, and for each we have $\mathbf{p}^1(x, u)\mathbf{p}^1(u, y) = O(m^{-2})$. The cases $(t_1, t_2, t_3) \in \{(0, 1, 1), (1, 0, 1)\}$ are handled similarly. □

Proof of part (a) of Theorem 1.3. Let $p \le p_c$. If one of the connections in the sum $\nabla_p(x, y)$ is of length in $[t_{\mathrm{mix}}, \infty)$, say between x and u, then we may estimate

$$\sum_{u,v}\mathbf{P}_p(x \xleftrightarrow{[t_{\mathrm{mix}},\infty)} u)\mathbf{P}_p(u \leftrightarrow v)\mathbf{P}_p(v \leftrightarrow y)$$
$$\le \frac{(1+o(1))\mathbb{E}_p|\mathscr{C}(0)|}{V}\sum_{u,v}\mathbf{P}_p(u \leftrightarrow v)\mathbf{P}_p(v \leftrightarrow y) = \frac{(1+o(1))(\mathbb{E}_p|\mathscr{C}(0)|)^3}{V},$$

where we have used Lemma 3.13 (and taken $r \to \infty$ on both sides of the lemma) for the first inequality. Thus, it remains to deal with short connections,

$$\nabla_p(x, y) \le \sum_{u,v}\mathbf{P}_p(x \overset{t_{\mathrm{mix}}}{\longleftrightarrow} u)\mathbf{P}_p(u \overset{t_{\mathrm{mix}}}{\longleftrightarrow} v)\mathbf{P}_p(v \overset{t_{\mathrm{mix}}}{\longleftrightarrow} y) + O(V^{-1}\chi(p)^3).$$

We write

$$\mathbf{P}_p(x \overset{t_{\mathrm{mix}}}{\longleftrightarrow} u) = \sum_{t_1=0}^{t_{\mathrm{mix}}}\mathbf{P}_p(x \overset{=t_1}{\longleftrightarrow} u),$$

and do the same for all three terms so that

$$\nabla_p(x, y) \le \sum_{u,v}\sum_{t_1,t_2,t_3}^{t_{\mathrm{mix}}}\mathbf{P}_p(x \overset{=t_1}{\longleftrightarrow} u)\mathbf{P}_p(u \overset{=t_2}{\longleftrightarrow} v)\mathbf{P}_p(v \overset{=t_3}{\longleftrightarrow} y) + O(V^{-1}\chi(p)^3). \tag{3.7}$$

We bound

$$\mathbf{P}_p(x \overset{=t_1}{\longleftrightarrow} u) \le m(m-1)^{t_1-1}\mathbf{p}^{t_1}(x, u)p^{t_1},$$

simply because $m(m-1)^{t_1-1}\mathbf{p}^{t_1}(x,u)$ is an upper bound on the number of simple paths of length t_1 starting at x and ending at u. Hence

$$\begin{aligned}
&\nabla_p(x,y)\\
&\quad\le \frac{m^3}{(m-1)^3}\sum_{u,v}\sum_{t_1,t_2,t_3}^{t_{\mathrm{mix}}}[p(m-1)]^{t_1+t_2+t_3}\mathbf{p}^{t_1}(x,u)\mathbf{p}^{t_2}(u,v)\mathbf{p}^{t_3}(v,y)+O(V^{-1}\chi(p)^3).
\end{aligned}$$

Since $p\le p_c$, assumption (2) gives $[p(m-1)]^{t_1+t_2+t_3}=1+O(\alpha_G)$, and condition (3) together with Claim 3.16 yields

$$\nabla_p(x,y)\le \mathbf{1}_{\{x=y\}}+O(V^{-1}\chi(p)^3)+O(m^{-1}+\alpha_G/\log V),$$

concluding the proof. □

3.7. Restricted triangle and square diagrams. In this section, we provide several extensions to the triangle condition (1.5). We will bound the triangle diagram in the supercritical phase (which requires a bound on the length of connections, otherwise the sums blow up) and estimate a square diagram which will be useful in a key second moment calculation in Section 6.

Lemma 3.17 (Short supercritical triangles). *Let G be a graph satisfying the assumptions in Theorem 1.3 and consider percolation on it with $p\le p_c(1+\varepsilon)$. Then*

$$\max_{x,y}\sum_{u,v\colon\{u,v\}\neq\{0,0\}}\mathbf{P}_p(x\overset{t_{\mathrm{mix}}}{\longleftrightarrow}u)\mathbf{P}_p(u\overset{t_{\mathrm{mix}}}{\longleftrightarrow}v)\mathbf{P}_p(v\overset{t_{\mathrm{mix}}}{\longleftrightarrow}y)=O(\alpha_G+\varepsilon t_{\mathrm{mix}}).$$

Proof. This follows immediately from assumptions (2) and (3) of Theorem 1.3 and the usual bound

$$\mathbf{P}_p(x\overset{=s}{\longleftrightarrow}u)\le p^s m(m-1)^{s-1}\mathbf{p}^s(x,u),$$

as before. □

Corollary 3.18 (Long supercritical triangles). *Let G be a graph satisfying the assumptions in Theorem 1.3 and consider percolation on it with $p\le p_c(1+\varepsilon)$. Let $r_1=r_1(m)$, $r_2=r_2(m)$, $r_3=r_3(m)$ be integers that are all at least t_{mix}. Then*

$$\begin{aligned}
&\max_{x,y}\sum_{u,v\colon\{u,v\}\neq\{x,y\}}\mathbf{P}(x\overset{r_1}{\leftrightarrow}u)\mathbf{P}(u\overset{r_2}{\leftrightarrow}v)\mathbf{P}(v\overset{r_3}{\leftrightarrow}y)\\
&\qquad\le O(\alpha_G+\varepsilon t_{\mathrm{mix}})+\frac{3+O(\alpha_G+\varepsilon t_{\mathrm{mix}})}{V}\mathbb{E}|B(r_1)|\,\mathbb{E}|B(r_2)|\,\mathbb{E}|B(r_3)|. \qquad (3.8)
\end{aligned}$$

Proof. We split the sum into two cases. The first case is that at least one of the connection events occurs with a path of length at least t_{mix}. For instance, if $x\xleftrightarrow{P[t_{\mathrm{mix}},r_1]}u$ occurs, then we use Lemma 3.13 to bound, uniformly in x,y,

$$\sum_{u,v} \mathbf{P}(x \xleftrightarrow{P[t_{\rm mix},r_1]} u)\mathbf{P}(u \overset{r_2}{\longleftrightarrow} v)\mathbf{P}(v \overset{r_3}{\longleftrightarrow} y)$$
$$\leq \frac{1+O(\alpha_G+\varepsilon t_{\rm mix})}{V}\mathbb{E}|B(r_1)| \sum_{u,v} \mathbf{P}(u \overset{r_2}{\longleftrightarrow} v)\mathbf{P}(v \overset{r_3}{\longleftrightarrow} y)$$
$$= \frac{1+O(\alpha_G+\varepsilon t_{\rm mix})}{V}\mathbb{E}|B(r_1)|\,\mathbb{E}|B(r_2)|\,\mathbb{E}|B(r_3)|.$$

The second case is when all the connections occur with paths of length at most $t_{\rm mix}$, in which case we use Lemma 3.17 to get an $O(\alpha_G+\varepsilon t_{\rm mix})$ bound. □

Lemma 3.19 (Supercritical square diagram). *Let G be a graph satisfying the assumptions in Theorem 1.3 and consider percolation on it with $p \leq p_c(1+\varepsilon)$. Let $r_1, r_2 \leq t_{\rm mix}$. Then*

$$\sum_{(u_1,u_1'),(u_2,u_2'),z_1,z_2} \mathbf{P}(z_1 \overset{r_1}{\longleftrightarrow} u_1)\mathbf{P}(z_1 \overset{r_1}{\longleftrightarrow} u_2)\mathbf{P}(z_2 \overset{r_2}{\longleftrightarrow} u_1')\mathbf{P}(z_2 \overset{r_2}{\longleftrightarrow} u_2')$$
$$\leq Cm^2(\mathbb{E}|B(r_1)|)^2(\mathbb{E}|B(r_2)|)^2 + CVm^2 t_{\rm mix}\alpha_G.$$

Proof. See Figure 1. If one of the connections is of length at least $t_{\rm mix}$, then we use Lemma 3.13 and the summation simplifies. For instance, if $u_1 \xleftrightarrow{[t_{\rm mix},r_1]} z_1$, then we use Lemma 3.13 and sum over z_1, and subsequently over u_1 and u_2. This gives a bound of

$$\frac{Cm^2(\mathbb{E}|B(r_1)|)^2}{V} \sum_{u_1',u_2',z_2} \mathbf{P}(z_2 \overset{r_2}{\longleftrightarrow} u_1')\mathbf{P}(z_2 \overset{r_2}{\longleftrightarrow} u_2') \leq Cm^2[\mathbb{E}|B(r_1)|^2\mathbb{E}|B(r_2)|]^2,$$

where $C > 1$ is an upper bound on $1+O(\alpha_G+\varepsilon t_{\rm mix})$.

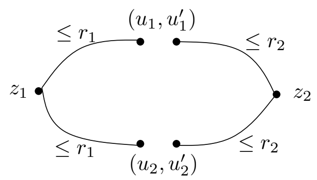

Fig. 1. A square diagram.

It remains to bound the sum

$$\sum_{(u_1,u_1'),(u_2,u_2'),z_1,z_2} \mathbf{P}(z_1 \overset{t_{\rm mix}}{\longleftrightarrow} u_1)\mathbf{P}(z_1 \overset{t_{\rm mix}}{\longleftrightarrow} u_2)\mathbf{P}(z_2 \overset{t_{\rm mix}}{\longleftrightarrow} u_1')\mathbf{P}(z_2 \overset{t_{\rm mix}}{\longleftrightarrow} u_2')$$
$$= V \sum_{(u_1,u_1'),(u_2,u_2'),z_2} \mathbf{P}(0 \overset{t_{\rm mix}}{\longleftrightarrow} u_1)\mathbf{P}(0 \overset{t_{\rm mix}}{\longleftrightarrow} u_2)\mathbf{P}(z_2 \overset{t_{\rm mix}}{\longleftrightarrow} u_1')\mathbf{P}(z_2 \overset{t_{\rm mix}}{\longleftrightarrow} u_2'), \tag{3.9}$$

by transitivity. We write this sum as $V\sum_{u_2'} f(u_2')g(u_2')$, where

$$g(u_2') = \sum_{(u_1,u_1'),z_2} \mathbf{P}(0 \overset{t_{\rm mix}}{\longleftrightarrow} u_1)\mathbf{P}(z_2 \overset{t_{\rm mix}}{\longleftrightarrow} u_1')\mathbf{P}(z_2 \overset{t_{\rm mix}}{\longleftrightarrow} u_2'),$$
$$f(u_2') = \sum_{u_2 \sim u_2'} \mathbf{P}(0 \overset{t_{\rm mix}}{\longleftrightarrow} u_2). \tag{3.10}$$

We then bound

$$
\begin{aligned}
V \sum_{u_2'} f(u_2')g(u_2') &\le V\Big(\sum_{u_2'} f(u_2')\Big)\Big(\max_{u_2'} g(u_2')\Big)\\
&= Vm\mathbb{E}|B(t_{\mathrm{mix}})| \max_x \sum_{(u_1,u_1'),z_2} \mathbf{P}(0 \xleftrightarrow{t_{\mathrm{mix}}} u_1)\mathbf{P}(z_2 \xleftrightarrow{t_{\mathrm{mix}}} u_1')\mathbf{P}(z_2 \xleftrightarrow{t_{\mathrm{mix}}} x).
\end{aligned}
$$

By condition (2) in Theorem 1.3, we can write the above as

$$
\begin{aligned}
&V \sum_{u_2'} f(u_2')g(u_2')\\
&\quad\le CVm^2\mathbb{E}|B(t_{\mathrm{mix}})| \max_x \sum_{u_1,u_1',z_2} \sum_{t_1,t_2,t_3\ge 0}^{t_{\mathrm{mix}}} \mathbf{p}^{t_1}(0,u_1)\mathbf{p}^1(u_1,u_1')\mathbf{p}^{t_2}(u_1',z_2)\mathbf{p}^{t_3}(z_2,x)\\
&\quad\le CVm^2\mathbb{E}|B(t_{\mathrm{mix}})|\\
&\qquad\times \max_x \sum_{u_1,z_2} \sum_{t_1,t_2,t_3\ge 0}^{t_{\mathrm{mix}}} \mathbf{p}^{t_1}(0,u_1)\Big[\mathbf{p}^{t_2+1}(u_1,z_2) + \frac{1}{m-1}\mathbf{p}^{t_2-1}(u_1,z_2)\Big]\mathbf{p}^{t_3}(z_2,x)\\
&\quad\le CVm^2\mathbb{E}|B(t_{\mathrm{mix}})|(\alpha_G + O(1/m)) \le CVm^2\mathbb{E}|B(t_{\mathrm{mix}})|\alpha_G,
\end{aligned}
$$

where we use Lemma 3.11 in the second inequality, and Claim 3.16, condition (2) of Theorem 1.3 and $\alpha_G \ge 1/m$ in the final inequality. Further, $\mathbb{E}|B(t_{\mathrm{mix}})| = O(t_{\mathrm{mix}})$ by Corollary 3.5 and the fact that $t_{\mathrm{mix}} = o(\varepsilon^{-1})$. This concludes our proof. □

4. Volume estimates

In this section, we study the expected volume of intrinsic balls and their boundaries at various radii in both the critical and supercritical phase.

4.1. In the critical regime. Given a subset A of vertices and integer $r \ge 0$ we write

$$
G(r;A) = \mathbb{E}|\partial B(r;A)|, \qquad G(r) = \max_{A\subseteq V(G)} G(r;A).
$$

Theorem 2.1 implies that for "most" r's the value $G(r)$ is bounded above by a constant (more precisely, given any fixed A and R, the number of r's satisfying $1 \le r \le R$ and $\mathbb{E}|\partial B(r;A)| \ge C/\delta$ is at most δR). The following useful result states that $G(r)$ is at most a constant *for all* r. We believe that this estimate should hold only under the assumption of the triangle condition but we are only able to prove it under the stronger assumptions of Theorem 1.3. The proof's strategy is similar to the proof of Theorem 2.1(1) in [40], that is, finding recursive inequalities involving $G(r)$. However, the details are somewhat more involved, and the uniform connection bounds (Lemma 3.15) are used to decouple a sum we could not decouple otherwise.

Theorem 4.1 (Expected boundary size). *Let G be a graph satisfying the assumptions of Theorem 1.3 and consider percolation on it with $p = p_c$. Then there exists a constant $C > 0$ such that for any integer r,*

$$
G(r) \le C.
$$

Proof. Define $F(r) = \mathbb{E}|B(r)|$ and $F(r; A) = \mathbb{E}|B(r; A)|$, so that $F(r; A) \le F(r)$ for all subsets A. Define $G^*(r) = \max_{s \le r} G(s)$, and let $r \ge 2t_{\text{mix}}$ be a maximizer of G^*, that is, $G(r') \le G(r)$ for any $r' < r$. Let $A = A(r)$ be the subset of vertices which maximizes $G(r; A)$, so that $G(r; A) = G(r) = G^*(r)$. We will prove that there exists $c > 0$ such that for any integer $s \ge 0$,

$$F(r+s; A) \ge cG^*(r)F(s; A). \tag{4.1}$$

We begin by bounding

$$F(r+s; A) \ge \sum_v \mathbf{P}_{\text{off } A}(0 \xleftrightarrow{(r,r+s]} v).$$

For a vertex u we define $\mathscr{C}^u(0; A) = \{x : 0 \leftrightarrow x \text{ off } A \cup \{u\}\}$. Now, for any vertex v, if there exists $u \ne v$ such that $0 \overset{=r}{\longleftrightarrow} u$ off A and $u \overset{s}{\longleftrightarrow} v$ off $\mathscr{C}^u(0; A)$, then $0 \xleftrightarrow{(r,r+s]} v$ off A. Furthermore, if such a u exists, then it is unique because otherwise $v \in \mathscr{C}^u(0; A)$. We deduce that

$$F(r+s; A) \ge \sum_{v \ne u} \mathbf{P}_{\text{off } A}\big(0 \overset{=r}{\longleftrightarrow} u \text{ and } \{u \overset{s}{\longleftrightarrow} v \text{ off } \mathscr{C}^u(0; A)\}\big).$$

We now condition on $\mathscr{C}^u(0; A) = H$ for some admissible H (that is, for which the probability of the event $\mathscr{C}^u(0; A) = H$ is positive, and in which $0 \overset{=r}{\longleftrightarrow} u$ occurs). In this conditioning, we also condition on the status of all edges touching H. Note that by definition $A \cap H = \emptyset$. We can write the right hand side of the last inequality as

$$\sum_{v \ne u} \sum_{H : 0 \overset{=r}{\leftrightarrow} u \text{ off } A} \mathbf{P}_{\text{off } A}(\mathscr{C}^u(0; A) = H)\mathbf{P}_{\text{off } A}(u \overset{s}{\longleftrightarrow} v \text{ off } H),$$

in the same way as we derived (3.1). This can be rewritten as

$$\sum_{v \ne u} \sum_{H : 0 \overset{=r}{\leftrightarrow} u \text{ off } A} \mathbf{P}_{\text{off } A}(\mathscr{C}^u(0; A) = H)\big[\mathbf{P}_{\text{off } A}(u \overset{s}{\longleftrightarrow} v) - \mathbf{P}_{\text{off } A}(u \overset{s}{\longleftrightarrow} v \text{ only on } H)\big].$$

We open the parenthesis and find that the first part of this sum is precisely $G^*(r)F(s; A)$ since r and A were maximizers. We need to show that the second part of the sum is of lower order. To that end, note that if $u \overset{s}{\longleftrightarrow} v$ only on H, then there exists $h \in H$ such that $h \ne u$ and $\{u \overset{s}{\longleftrightarrow} h\} \circ \{h \overset{s}{\longleftrightarrow} v\}$. By the BK inequality, we bound the second part of the sum above by

$$\sum_u \sum_{H : 0 \overset{=r}{\leftrightarrow} u \text{ off } A} \sum_{h \in H,\, h \ne u, v} \mathbf{P}_{\text{off } A}(\mathscr{C}^u(0; A) = H)\mathbf{P}_{\text{off } A}(u \overset{s}{\longleftrightarrow} h)\mathbf{P}_{\text{off } A}(h \overset{s}{\longleftrightarrow} v).$$

Summing over v and changing the order of summation shows that the last sum is at most

$$F(s; A) \sum_{u \ne h} \mathbf{P}_{\text{off } A}(0 \overset{=r}{\longleftrightarrow} u, 0 \leftrightarrow h)\mathbf{P}_{\text{off } A}(u \overset{s}{\longleftrightarrow} h).$$

We bound this from above using Lemma 3.3 by

$$F(s;A)\sum_{u\neq h,z}\sum_{t\leq r}\mathbf{P}_{\mathrm{off}\,A}(0\overset{=t}{\longleftrightarrow}z)\\ \times\max_{D\subseteq V(G)}\mathbf{P}(z\overset{=r-t}{\longleftrightarrow}u\text{ off }A\cup D)\mathbf{P}_{\mathrm{off}\,A}(z\leftrightarrow h)\mathbf{P}_{\mathrm{off}\,A}(u\overset{s}{\longleftrightarrow}h). \tag{4.2}$$

We sum this separately for $t\leq r-t_{\mathrm{mix}}$ and $t\in[r-t_{\mathrm{mix}},r]$. For $t\leq r-t_{\mathrm{mix}}$ we bound, for any $D\subset V$,

$$\mathbf{P}(z\overset{=r-t}{\longleftrightarrow}u\text{ off }A\cup D)\leq 3G(r-t-t_{\mathrm{mix}};A\cup D)/V\leq 3G^*(r)/V,$$

where the first inequality is by Lemma 3.15 and the second by definition of $G^*(r)$. Hence, the sum over $t\leq r-t_{\mathrm{mix}}$ in (4.2) is at most

$$\frac{3G^*(r)F(s;A)}{V}\sum_{u,h,z}\mathbf{P}_{\mathrm{off}\,A}(0\overset{r}{\leftrightarrow}z)\mathbf{P}_{\mathrm{off}\,A}(z\leftrightarrow h)\mathbf{P}_{\mathrm{off}\,A}(h\overset{s}{\longleftrightarrow}u)\\ \leq 3G^*(r)F(s;A)(\mathbb{E}|\mathscr{C}(0)|)^3/V=3\lambda^3G^*(r)F(s;A),$$

where the inequality is obtained by summing over u,h and z (in that order), and the equality is due to the definition of p_c in (1.1). Our $\lambda=1/10$ is chosen small enough so that $3\lambda^3\leq 1/2$.

We now bound the sum in (4.2) for $t\in[r-t_{\mathrm{mix}},r]$. We first bound

$$\mathbf{P}_{\mathrm{off}\,A}(0\overset{=t}{\longleftrightarrow}z)\leq 3G^*(r)/V,$$

as we did before using Lemma 3.15, and pull that term out of the sum. This gives an upper bound of

$$\frac{3G^*(r)F(s;A)}{V}\sum_{u\neq h,z}\sum_{s_1=0}^{t_{\mathrm{mix}}}\max_{D\subseteq V(G)}\mathbf{P}(z\overset{=s_1}{\longleftrightarrow}u\text{ off }A\cup D)\mathbf{P}_{\mathrm{off}\,A}(z\leftrightarrow h)\mathbf{P}_{\mathrm{off}\,A}(u\overset{s}{\longleftrightarrow}h).$$

We would like to sum the first term in the sum over s_1 and get a contribution of $\mathbf{P}(z\overset{t_{\mathrm{mix}}}{\longleftrightarrow}u)$. We cannot do that, however, because the maximizing set D may depend on s_1 so these are not necessarily disjoint events. Instead we estimate, for all $D\subseteq V(G)$,

$$\mathbf{P}(z\overset{=s_1}{\longleftrightarrow}u\text{ off }A\cup D)\leq m(m-1)^{s_1}p_c^{s_1}\mathbf{p}^{s_1}(z,u)\leq(1+o(1))\mathbf{p}^{s_1}(z,u),$$

where the first inequality is since $m(m-1)^{s_1}\mathbf{p}^{s_1}(z,u)$ bounds the number of simple paths of length s_1 connecting z to u, and the second inequality is due to condition (2) of Theorem 1.3. Now, if one of the connections $z\leftrightarrow h$ or $u\overset{s}{\longleftrightarrow}h$ is in fact a connection of length at least t_{mix}, we use Lemma 3.13 to simplify the sum. For instance, if the connection is $z\overset{[t_{\mathrm{mix}},\infty)}{\longleftrightarrow}h$, then we bound the probability of this by $2V^{-1}\mathbb{E}|\mathscr{C}(0)|$ and the sum simplifies to

$$\frac{4G^*(r)F(s;A)\mathbb{E}|\mathscr{C}(0)|}{V^2}\sum_{u\neq h,z}\sum_{s_1=0}^{t_{\mathrm{mix}}}\mathbf{p}^{s_1}(z,u)\mathbf{P}(u\overset{s}{\longleftrightarrow}h);$$

we then sum over h, u, z to s_1 and get a contribution of

$$\frac{4G^*(r)F(s;A)t_{\mathrm{mix}}(\mathbb{E}|\mathscr{C}(0)|)^2}{V} = o(G^*(r)F(s;A)),$$

since $(\mathbb{E}|\mathscr{C}(0)|)^2 = O(V^{2/3})$ and $t_{\mathrm{mix}} = o(\varepsilon^{-1}) = o(V^{1/3})$. It remains to bound

$$\frac{3G^*(r)F(s;A)}{V}\sum_{u\neq h,z}\sum_{s_1=0}^{t_{\mathrm{mix}}}\mathbf{p}^{s_1}(z,u)\mathbf{P}_{\mathrm{off}\,A}(z \overset{t_{\mathrm{mix}}}{\longleftrightarrow} h)\mathbf{P}_{\mathrm{off}\,A}(u \overset{t_{\mathrm{mix}}}{\longleftrightarrow} h),$$

by

$$\frac{CG^*(r)F(s;A)}{V}\sum_{u\neq h,z}\sum_{s_1=0,\, s_2=1,\, s_3=0}^{t_{\mathrm{mix}}}\mathbf{p}^{s_1}(z,u)\mathbf{p}^{s_2}(u,h)\mathbf{p}^{s_3}(h,z) = o(1)\cdot G^*(r)F(s;A),$$

where we have used Claim 3.16 and condition (3) of Theorem 1.3. This concludes the proof of (4.1).

We now prove the main result assuming (4.1). First, for any $r \le 2t_{\mathrm{mix}}$ the number of non-backtracking paths emanating from 0 is at most $m(m-1)^{r-1}$, and hence, for any A,

$$G(r;A) \le m(m-1)^{r-1}p_c^r = 1 + o(1),$$

by condition (2) of Theorem 1.3. It remains to consider the case $r \ge 2t_{\mathrm{mix}}$. Assume for contradiction that there exists some $r \ge 2t_{\mathrm{mix}}$ such that r is the maximizer in the definition $G^*(r)$, and $G^*(r) \ge 2/c$ where c is the constant from (4.1). Fix such an r and let $A = A(r)$ be the maximizing set as in (4.1). Now, setting $s = r$ in (4.1) gives

$$F(2r;A) \ge cG^*(r)F(r;A) \ge 2F(r;A).$$

Setting $s = 2r$ in (4.1) gives

$$F(3r;A) \ge cG^*(r)F(2r;A) \ge 4F(r;A),$$

and so by induction, for any k,

$$F(kr;A) \ge 2^{k-1}F(r;A).$$

We have reached a contradiction, since on the right hand side we have a quantity going to ∞ with k (note that A cannot contain 0, otherwise it will not be maximizing, so $F(r;A) \ge 1$), and on the left hand side our quantity is bounded by V. □

We now wish to obtain the reverse inequality to Theorem 4.1, that is, a lower bound on $\mathbb{E}|\partial B(r)|$. Of course, this cannot hold for all r, but it turns out to hold as long as $r \ll V^{1/3}$. This is the correct upper bound on r because the diameter of critical clusters is of order $V^{1/3}$ (see [49]).

Lemma 4.2 (Lower bound on critical expected ball). *Let G be any transitive finite graph and set $p = p_c$ where p_c is defined in (1.4). Then there exists a constant $\xi > 0$ such that for all $r \le \xi V^{1/3}$,*

$$\mathbb{E}|B(r)| \ge r/4.$$

Proof. For convenience write $c = 1/4$. Assume for contradiction that $\mathbb{E}|B(r)| \le cr$. Given this assumption, we will prove by induction that for any integer $k \ge 0$,

$$\mathbb{E}\big|B\big([r(1+k/2), r(1+(k+1)/2)]\big)\big| \le 2^{k+1}c^{k+2}r. \tag{4.3}$$

For $k = 0$, since $\mathbb{E}|B(r)| \le cr$ there exists $r' \in [r/2, r]$ such that $\mathbb{E}|\partial B(r')| \le 2c$, so

$$\begin{aligned}\mathbb{E}|B([r, 3r/2])| &= \sum_A \mathbf{P}(B(r') = A)\mathbb{E}[|B([r, 3r/2])| \mid B(r') = A] \\ &= \sum_A \mathbf{P}(B(r') = A)\mathbb{E}\Big[\sum_{a \in \partial A} \mathbb{E}|B_a(3r/2 - r'; A)|\Big] \\ &\le cr\mathbb{E}|\partial B(r')| \le 2c^2 r,\end{aligned}$$

where the inequality follows since $\mathbb{E}|B_a(3r/2 - r'; A)| \le \mathbb{E}|B(r)| \le cr$ for any A by monotonicity. Assume now that (4.3) holds for some k. Then there exists $r' \in [r(1+k/2), r(1+(k+1)/2)]$ such that $\mathbb{E}|\partial B(r')| \le 2^{k+2}c^{k+2}$. By conditioning on $B(r') = A$ as before we get

$$\mathbb{E}\big|B\big([r(1+(k+1)/2), r(1+(k+2)/2)]\big)\big| \le cr \cdot 2^{k+2}c^{k+2},$$

concluding the proof of (4.3).

Now, since $c < 1/2$ it is clear that the sum over k of (4.3) is at most Cr, contradicting the fact that $\mathbb{E}_{p_c}|\mathscr{C}(0)| = \lambda V^{1/3}$ by our definition of p_c in (1.4). Note that the constant ξ may depend on λ. □

Lemma 4.3 (Lower bound on expected boundary size). *Let G be a transitive finite graph for which* (1.5) *holds and set $p = p_c$. Then there exist constants $c, \xi > 0$ such that for any $r \le \xi V^{1/3}$,*

$$\mathbb{E}|\partial B(r)| \ge c.$$

Proof. By Lemma 4.2 and Theorem 2.1 we know that $\mathbb{E}|B([2r, Cr])| \ge r$ for some large fixed $C > 0$. Also,

$$\mathbb{E}|B(Cr)|^2 \le \sum_{x,y} \mathbf{P}(0 \overset{Cr}{\longleftrightarrow} x,\ 0 \overset{Cr}{\longleftrightarrow} y) \le \sum_{x,y,z} \mathbf{P}(0 \overset{Cr}{\longleftrightarrow} z)\mathbf{P}(z \overset{Cr}{\longleftrightarrow} x)\mathbf{P}(z \overset{Cr}{\longleftrightarrow} y) \le Cr^3,$$

by Theorem 2.1. From the inequality

$$\mathbf{P}(X > a) \ge (\mathbb{E}X - a)^2/\mathbb{E}X^2 \tag{4.4}$$

valid for any non-negative random variable X and $a < \mathbb{E}X$, we obtain, with $a = 0$,

$$\mathbf{P}(\partial B(2r) \neq \emptyset) \ge c/r$$

for some $c > 0$. Furthermore, given $B(r)$, each vertex of $\partial B(r)$ has probability at most Cr^{-1} of reaching $\partial B(2r)$ by Theorem 2.1. Hence, for any $\zeta = \zeta(m) > 0$,

$$\begin{aligned}&\mathbf{P}\big(\partial B(2r) \neq \emptyset \text{ and } |\partial B(r)| \le \zeta r\big) \\ &\quad = \mathbf{P}\big(\partial B(2r) \neq \emptyset \,\big|\, 0 < |\partial B(r)| \le \zeta r\big)\mathbf{P}(0 < |\partial B(r)| \le \zeta r) \le C^2\zeta/r.\end{aligned}$$

We now have

$$\begin{aligned}\mathbf{P}(|\partial B(r)| \geq \zeta r) &\geq \mathbf{P}(\partial B(2r) \neq \emptyset) - \mathbf{P}\big(\partial B(2r) \neq \emptyset \text{ and } |\partial B(r)| \leq \zeta r\big)\\ &\geq c/r - C^2\zeta/r,\end{aligned}$$

and the lemma follows by choosing $\zeta > 0$ small enough. □

4.2. In the supercritical regime. In this section, we extend the volume estimates to the supercritical regime. The following is an immediate corollary of Theorem 4.1:

Lemma 4.4 (Upper bound on supercritical volume). *Let G be a graph satisfying the assumptions in Theorem 1.3 and consider percolation on it with $p \leq p_c(1+\varepsilon)$. Then, for any r and any $A \subset V(G)$,*

$$\mathbb{E}|\partial B(t; A)| \leq C(1+\varepsilon)^t \quad \textit{and} \quad \mathbb{E}|B(r; A)| \leq C\varepsilon^{-1}(1+\varepsilon)^r.$$

Proof. The first assertion is immediate from Theorem 4.1 and Lemma 3.4. The second assertion follows by summing the first over $t \leq r$. □

The corresponding lower bound is more complicated to obtain, and as before, can only hold up to some value of r. In conjunction with Lemma 4.4, it identifies $\mathbb{E}|B(r)| = \Theta(\varepsilon^{-1}(1+\varepsilon)^r)$ for appropriate r's.

Theorem 4.5 (Lower bound on supercritical volume). *Let G be a graph satisfying the assumptions in Theorem 1.3 and consider percolation on it with $p = p_c(1+\varepsilon)$. Then, for any r satisfying*

$$\mathbb{E}|B(r)| \leq \frac{\varepsilon^2 V}{(\log(\varepsilon^3 V))^4}, \tag{4.5}$$

the following bound holds:

$$\mathbb{E}|B(r)| \geq c\varepsilon^{-1}(1+\varepsilon)^r.$$

Proof. First, we may assume that

$$r \leq \varepsilon^{-1}\log(\varepsilon^3 V), \tag{4.6}$$

since otherwise the assumption of the lemma cannot hold together with the conclusion. Recall now the simultaneous coupling (described at the end of Section 2.1) between percolation at $p_1 = p_c$ and $p_2 = p_c(1+\varepsilon)$. Let

$$\mathscr{A}_\ell(x) = \{0 \overset{=\ell}{\longleftrightarrow} x \text{ in } G_{p_2}\},$$

and given a simple path η of length ℓ between 0 and x, write

$$\begin{aligned}&\mathscr{A}_\ell(x,\eta)\\ &\quad= \{0 \overset{=\ell}{\longleftrightarrow} x \text{ in } G_{p_2} \text{ and } \eta \text{ is the lexicographically first } p_2\text{-open path between 0 and } x\},\end{aligned}$$

so that $\mathscr{A}_\ell(x) = \bigcup_\eta \mathscr{A}_\ell(x,\eta)$. Write $\mathscr{B}_\ell(x,\eta)$ for the event that the edges of η are in fact p_1-open (not just p_2). We have

$$\mathscr{A}_\ell(x,\eta) \cap \mathscr{B}_\ell(x,\eta) \subseteq \{0 \overset{=\ell}{\longleftrightarrow} x \text{ in } G_{p_1}\},$$

so

$$\bigcup_{\ell\in[r-\varepsilon^{-1},r],\,\eta} \mathscr{A}_\ell(x,\eta)\cap\mathscr{B}_\ell(x,\eta) \subseteq \{0 \xleftrightarrow{[r-\varepsilon^{-1},r]} x \text{ in } G_{p_1}\}. \tag{4.7}$$

We will show that

$$\sum_x \mathbf{P}\Big(0 \xleftrightarrow{[r-\varepsilon^{-1},r]} x \text{ in } G_{p_1} \setminus \bigcup_{\ell\in[r-\varepsilon^{-1},r],\,\eta} \mathscr{A}_\ell(x,\eta)\cap\mathscr{B}_\ell(x,\eta)\Big) = o(\varepsilon^{-1}), \tag{4.8}$$

and first complete the proof subject to (4.8). Since $\{\mathscr{A}_\ell(x,\eta)\cap\mathscr{B}_\ell(x,\eta)\}_{\ell,\eta}$ are disjoint events, (4.7) and (4.8) show that

$$\sum_{x,\,\ell\in[r-\varepsilon^{-1},r],\,\eta} \mathbf{P}(\mathscr{A}_\ell(x,\eta)\cap\mathscr{B}_\ell(x,\eta)) \geq \mathbb{E}_{p_1}|B([r-\varepsilon^{-1},r])| - o(\varepsilon^{-1}) \geq c\varepsilon^{-1},$$

where the last inequality uses Lemma 4.3 and the fact that $r \ll V^{1/3}$ by (4.6) and $\varepsilon \gg V^{-1/3}$. From this the required result follows since

$$\mathbf{P}(\mathscr{B}_\ell(x,\eta) \mid \mathscr{A}_\ell(x,\eta)) = (1+\varepsilon)^{-\ell},$$

which implies that

$$\begin{aligned}
\mathbb{E}_{p_2}|B(r)| &\geq \sum_{x,\,\ell\in[r-\varepsilon^{-1},r],\,\eta} \mathbf{P}(\mathscr{A}_\ell(x,\eta)) \\
&= \sum_{x,\,\ell\in[r-\varepsilon^{-1},r],\,\eta} \mathbf{P}(\mathscr{A}_\ell(x,\eta))\mathbf{P}\big(\mathscr{B}_\ell(x,\eta) \mid \mathscr{A}_\ell(x,\eta)\big)(1+\varepsilon)^\ell \\
&\geq (1+\varepsilon)^{r-\varepsilon^{-1}} \sum_{x,\,\ell\in[r-\varepsilon^{-1},r],\,\eta} \mathbf{P}(\mathscr{A}_\ell(x,\eta)\cap\mathscr{B}_\ell(x,\eta)) \geq (1+\varepsilon)^{r-\varepsilon^{-1}} c\varepsilon^{-1}.
\end{aligned}$$

Thus, our main effort is to show (4.8) under the restriction of (4.5) and (4.6). Fix x and assume that the event

$$\{0 \xleftrightarrow{[r-\varepsilon^{-1},r]} x \text{ in } G_{p_1}\} \setminus \bigcup_{\ell\in[r-\varepsilon^{-1},r],\,\eta} \mathscr{A}_\ell(x,\eta)\cap\mathscr{B}_\ell(x,\eta) \tag{4.9}$$

occurs. In words, this event means that either the shortest p_2-open path is shorter than the shortest p_1-path, or they have the same length but the lexicographically first shortest p_2-path contains an edge having value in $[p_1, p_2]$. This implies that the p_2-path *shortcuts* the p_1-path. Formally, given vertices u, v and integers $\ell \in [r-\varepsilon^{-1}, r]$, $k \in [0, \ell]$, $t \in [2, \ell]$ with $k + t \leq \ell$ write $\mathscr{T}(u, v, x, \ell, k, t)$ for the event that there exist paths $\eta_1, \eta_2, \eta_3, \gamma$ in the graph such that

(1) η_1 is a shortest p_1-open path of length k connecting 0 to u,
(2) η_2 is a shortest p_1-open path of length t connecting u to v,
(3) η_3 is a shortest p_1-open path of length $\ell - t - k$ connecting v to x,
(4) $B_x(\ell - k - t) \cap B_0(k) = \emptyset$ in G_{p_1},
(5) γ is a p_2-open path of length at most t connecting u to v and one of the edges of γ receives value in $[p_1, p_2]$, and
(6) η_1, η_2, η_3 and γ are disjoint paths

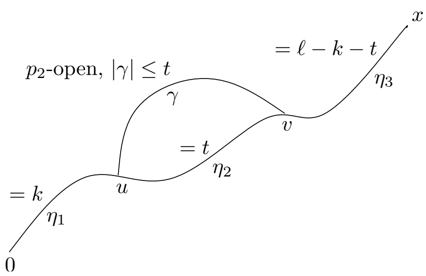

Fig. 2. The event $\mathscr{T}(u, v, x, \ell, k, t)$—an excursion.

(see Figure 2). The event (4.9) implies that $\mathscr{T}(u, v, x, \ell, k, t)$ occurs for some u, v, ℓ, k, t satisfying the conditions above. Our treatment of the case $t \ge t_{\text{mix}}$ is easier than the case $t \le t_{\text{mix}}$, so let us perform the former first. When $t \ge t_{\text{mix}}$ we forget about condition (4) and the special edge with value $[p_1, p_2]$ in (5) and take a union over ℓ, k and $t \in [t_{\text{mix}}, r]$ of the event $\mathscr{T}(u, v, x, \ell, k, t)$. This union implies the existence of vertices u, v such that the following events occur disjointly:

- $0 \overset{r}{\longleftrightarrow} u$ in G_{p_1},
- $u \xleftrightarrow{P[t_{\text{mix}}, r]} v$ in G_{p_1},
- $v \overset{r}{\longleftrightarrow} x$ in G_{p_1},
- $u \overset{r}{\longleftrightarrow} v$ in G_{p_2}.

Indeed, the witnesses to these (monotone) events are the paths $\eta_1, \eta_2, \eta_3, \gamma$. We now wish to use the BK inequality; however, as the astute reader may have already noticed, our witnesses are not stated in an i.i.d. product measure. Let us expand briefly on how we may still use the BK inequality. We may consider our simultaneous coupling measure to be an i.i.d. product measure by putting on each edge a countable infinite sequence of independent random bits receiving 0 with probability 1/2 and 1 otherwise such that this sequence encodes the uniform [0, 1] random variable attached to each edge. In this setting, a witness for an edge being p-open is the sequence of bits attached to the edge, and similarly for the edge being p-closed. Similarly, we define events of the form "E_1 in G_{p_1} occurs disjointly from E_2 in G_{p_2}". With this definition of witnesses we may use the BK inequality here to bound the probability of the union above and sum over x (as in (4.8)). This gives an upper bound of

$$\sum_{u,v,x} \mathbf{P}_{p_1}(0 \overset{r}{\longleftrightarrow} u)\mathbf{P}_{p_1}(u \xleftrightarrow{P[t_{\text{mix}},r]} v)\mathbf{P}_{p_1}(v \overset{r}{\longleftrightarrow} x)\mathbf{P}_{p_2}(u \overset{r}{\longleftrightarrow} v).$$

We sum over x and get a factor of r by Theorem 2.1. We bound $\mathbf{P}_{p_1}(u \xleftrightarrow{P[t_{\text{mix}},r]} v) \le CrV^{-1}$ by Lemma 3.13 and Theorem 2.1. We then sum over v and get a factor of $\mathbb{E}_{p_2}|B(r)|$, and over u to get another factor of r. Altogether this gives an upper bound of

$$Cr^3\mathbb{E}_{p_2}|B(r)|/V = O\big(\varepsilon^{-1}(\log(\varepsilon^3 V))^{-1}\big) = o(\varepsilon^{-1}),$$

by (4.5).

We now treat the case $t \in [2, t_{\mathrm{mix}}]$. We claim that the event $\mathscr{T}(u, v, x, \ell, k, t)$ implies that there exist disjoint paths η_2, γ between u and v such that $|\eta_2| = t$ and $|\gamma| \le t$ and the intersection of the following events occurs:

(a) η_2 is p_1-open,
(b) γ is p_2-open, and one of its edges receives value in $[p_1, p_2]$,
(c) $0 \overset{=k}{\longleftrightarrow} u$ off $\eta_2 \cup \gamma$ and $v \overset{=\ell-k-t}{\longleftrightarrow} x$ off $\eta_2 \cup \gamma \cup B_0(k)$ in G_{p_1}.

Indeed, let $\eta_1, \eta_2, \eta_3, \gamma$ be the disjoint paths guaranteed to exist in the definition of $\mathscr{T}(u, v, x, \ell, k, t)$. The paths η_2 and γ show that both (a) and (b) indeed occur (note that we have relaxed the requirement that η_2 is a shortest p_1-open path). To see that (c) occurs, first observe that for any vertices z, y and integer $\ell \ge 0$,

$$\{z \overset{=\ell}{\longleftrightarrow} y \text{ off } A\} = \bigcup_{\beta\colon |\beta|=\ell,\, \beta\cap A=\emptyset} \Big(\{\beta \text{ is open}\} \cap \bigcap_{\beta'\colon |\beta'|<\ell,\, \beta'\cap A=\emptyset} \{\beta' \text{ has a closed edge}\}\Big),$$

where β, β' are simple paths in G and we slightly abuse notation and write $\beta \cap A = \emptyset$ to denote that the edges of β are disjoint from the edges touching A. To see that (c) holds we note that the event $\mathscr{T}(u, v, x, \ell, k, t)$ implies that η_1 is of length k between 0 and u, is disjoint from $\eta_2 \cup \gamma$, is p_1-open, and any shorter path between 0 and u has a p_1-closed edge in it; in particular, $0 \overset{=k}{\longleftrightarrow} u$ off $\eta_2 \cup \gamma$ occurs in G_{p_1}. Similarly, η_3 is of length $\ell-k-t$ between v and x, is disjoint from $\eta_2 \cup \gamma \cup B_0(k)$, is p_1-open, and any shorter path between v and x has a p_1-closed edge in it; in particular $v \overset{=\ell-k-t}{\longleftrightarrow} x$ off $\eta_2 \cup \gamma \cup B_0(k)$ occurs in G_{p_1}.

Now, the events (a), (b), (c) are independent since they are measurable with respect to disjoint sets of edges (the edges of η_2, γ and all the rest). The probability of their intersection is hence

$$p_1^{|\eta_2|} p_2^{|\gamma|} [1 - (p_1/p_2)^{|\gamma|}] \mathbf{P}_{p_1}\big(0 \overset{=k}{\longleftrightarrow} u \text{ off } \eta_2 \cup \gamma \text{ and } v \overset{=\ell-k-t}{\longleftrightarrow} x \text{ off } \eta_2 \cup \gamma \cup B_0(k)\big),$$

where the factor $1 - (p_1/p_2)^{|\gamma|}$ is the probability that one edge of γ has value in $[p_1, p_2]$ conditioned on all edges being p_2-open. We compute the probability on the right hand side as usual by conditioning on $B_0(k)$, which gives

$$p_1^{|\eta_2|} p_2^{|\gamma|} [1 - (p_1/p_2)^{|\gamma|}] \sum_{A\colon 0 \overset{=k}{\longleftrightarrow} u} \mathbf{P}_{p_1}(B_0(k) = A) \mathbf{P}_{p_1}(v \overset{=\ell-k-t}{\longleftrightarrow} x \text{ off } A \cup \eta_2 \cup \gamma).$$

We now start summing all this over $u, v, x, \ell, k, t, \eta_2, \gamma$. We first sum over x the last probability, which gives a constant factor by Theorem 4.1. The sum over A gives a term of $\mathbf{P}_{p_1}(0 \overset{=k}{\longleftrightarrow} u \text{ off } \eta_2 \cup \gamma)$ which we sum over $k \in [0, r]$ and bound this by $\mathbf{P}_{p_1}(0 \overset{r}{\longleftrightarrow} u)$. Furthermore, the number of possible η_2's is at most $m(m-1)^t \mathbf{p}^t(u, v)$, and if $|\gamma| = s \le t$, then the number of such γ's is at most $m(m-1)^{s-1} \mathbf{p}^s(u, v)$. We also bound

$1-(p_1/p_2)^s \le Cs\varepsilon$. All this gives

$$\sum_{u,v,x,\ell,k,\,t\in[2,t_{\mathrm{mix}}]} \mathbf{P}(\mathscr{T}(u,v,x,\ell,k,t))$$
$$\le C\varepsilon \sum_{\substack{u,v,\ell \\ t\in[2,t_{\mathrm{mix}}],\,s\in[1,t]}} (m-1)^{s+t} p_1^t p_2^s s \mathbf{p}^t(u,v)\mathbf{p}^s(u,v)\mathbf{P}_{p_1}(0 \stackrel{r}{\leftrightarrow} u).$$

From (2) of Theorem 1.3 and $t_{\mathrm{mix}} = o(\varepsilon^{-1})$, we get $(m-1)^{s+t} p_1^t p_2^s = 1 + o(\alpha_G)$, so we may bound the above sum by

$$C\varepsilon \sum_{u,\ell} \mathbf{P}_{p_1}(0 \stackrel{r}{\leftrightarrow} u) \sum_{v,\,t\in[2,t_{\mathrm{mix}}],\,s\in[1,t]} s\mathbf{p}^t(u,v)\mathbf{p}^s(u,v).$$

The sum over $\ell \in [r-\varepsilon^{-1}, r]$ gives a factor of ε^{-1}, and since G is transitive, the second sum over v, t, s does not depend on u. Hence we may sum over u separately using Theorem 2.1 to get a bound of

$$Cr \sum_{v,\,t\in[2,t_{\mathrm{mix}}],\,s\in[1,t]} s\mathbf{p}^t(u,v)\mathbf{p}^s(u,v).$$

For each $s \ge 1$ and $s_1 \in \{1,\dots,s\}$, we can bound

$$\mathbf{p}^s(0,v) \le \frac{m}{m-1}\sum_w \mathbf{p}^{s_1}(0,w)\mathbf{p}^{s-s_1}(w,v),$$

because the number of non-backtracking paths of length s from 0 to v is at most the sum over w of the number of non-backtracking paths of length s_1 from 0 to w times the number of non-backtracking paths of length $s-s_1$ from w to v (the factor $m/(m-1)$ comes from properly normalizing these numbers). As a result,

$$\sum_{v,\,t\in[2,t_{\mathrm{mix}}],\,s\in[1,t]} s\mathbf{p}^t(0,v)\mathbf{p}^s(0,v)$$
$$\le \frac{m}{m-1}\sum_{v,w}\sum_{t\in[2,t_{\mathrm{mix}}],\,s_1\in[1,t],\,s_2\le s_1} \mathbf{p}^t(0,v)\mathbf{p}^{s_1}(0,w)\mathbf{p}^{s_2}(w,v) \le \frac{C\alpha_G}{\log V},$$

by (3) in Theorem 1.3 and the fact that $t+s_1+s_2 \ge 3$. Altogether we get

$$\sum_{u,v,x,\ell,k,\,t\in[2,t_{\mathrm{mix}}]} \mathbf{P}(\mathscr{T}(u,v,x,\ell,k,t)) \le \frac{Cr\alpha_G}{\log V} = o(\varepsilon^{-1}),$$

by our assumption (4.6) and since $\alpha_G = o(1)$. This finishes the proof of (4.8) and concludes the proof of the theorem. □

The following are easy corollaries:

Corollary 4.6. *Let G be a graph satisfying the assumptions in Theorem 1.3 and consider percolation on it with $p = p_c(1+\varepsilon)$. Then, for any r satisfying*

$$r \le \varepsilon^{-1}[\log(\varepsilon^3 V) - 4\log\log(\varepsilon^3 V)],$$

the following bound holds:

$$\mathbb{E}|B(r)| = \Theta(\varepsilon^{-1}(1+\varepsilon)^r).$$

In particular, for r_0 defined in (2.8),

$$\mathbb{E}|B(r_0)| = \Theta(\sqrt{\alpha_G \varepsilon V}).$$

Proof. The upper bound follows from Lemma 4.4, and the lower bound from Theorem 4.5. □

Lemma 4.7. *Let G be a graph satisfying the assumptions in Theorem 1.3 and consider percolation on it with $p \le p_c(1+\varepsilon)$. Let r be an integer satisfying the assumptions of Theorem 4.5. Then*

$$\mathbb{E}|B(r)|^2 \le C\varepsilon^{-1}(\mathbb{E}|B(r)|)^2.$$

Proof. If $0 \overset{r}{\longleftrightarrow} x$ and $0 \overset{r}{\longleftrightarrow} y$, then there exists a vertex z and an integer $t \le r$ such that the event

$$\{0 \overset{=t}{\longleftrightarrow} z\} \circ \{z \overset{r-t}{\longleftrightarrow} x\} \circ \{z \overset{r-t}{\longleftrightarrow} y\}$$

holds. Apply BK-Reimer and sum over x, y and then z to bound

$$\mathbb{E}|B(r)|^2 \le \sum_{t=1}^{r} \mathbb{E}|\partial B(t)|\,\mathbb{E}|B(r-t)|\,\mathbb{E}|B(r-t)|.$$

We apply Lemma 4.4 and Theorem 4.1 to obtain

$$\mathbb{E}|B(r)|^2 \le C\varepsilon^{-3}(1+\varepsilon)^{2r},$$

and Theorem 4.5 gives the required claim. □

5. An intrinsic metric regularity theorem

For an increasing event E and a vertex a, we say that a is *pivotal for E* whenever E occurs but does not occur in the modified configuration in which we close all the edges touching a. We write $\mathrm{Piv}(E)$ for the set of pivotal vertices for the event E. For vertices a, x, radii r_1, j_x and $A \subset V(G)$, we define

$$G_{r_1,j_x}(a,x;A) = \mathbb{E}\Big[|\{u : a \xleftrightarrow{P[2t_{\mathrm{mix}},r_1]} u \text{ off } A \setminus \{a\} \text{ and } a \in \mathrm{Piv}(\{x \xleftrightarrow{j_x+r_1} u\})\}| \;\Big|\; B_x(j_x) = A\Big].$$

Definition 5.1 (Regenerative and fit vertices).

(a) Given vertices x, a, radii r_1, $j_x \ge t_{\mathrm{mix}}$ and a real number $\beta > 0$, we say that a is (β, j_x, r_1)-*regenerative* if
- $x \overset{=j_x}{\longleftrightarrow} a$, and
- $G_{r_1,j_x}(a,x;B_x(j_x)) \ge (1-\beta)\mathbb{E}|B(r_1)|$,

and note that this event is determined by the status of the edges touching $B_x(j_x)$.

We say that a is (β, j_x, r_1)-*non-regenerative* if $x \overset{=j_x}{\longleftrightarrow} a$ but a is not (β, j_x, r_1)-regenerative.

(b) Given an additional real number $\delta > 0$, we say that x is $(\delta, \beta, j_x, r_1)$-*fit* if

- $\partial B_x(j_x) \neq \emptyset$, and
- the number of (β, j_x, r_1)-non-regenerative vertices is at most $\delta\varepsilon^{-1}$.

It will also be convenient to combine our error terms. For this, we define

$$\omega = \omega(m) = \alpha_G^{1/2} + \varepsilon t_{\text{mix}}, \tag{5.1}$$

so that $\omega = o(1)$. Our goal in this section is to prove that if $\partial B_x(j_x) \neq \emptyset$, then x is fit with high probability. This is the *intrinsic metric* regularity theorem discussed in Section 2.5.

Theorem 5.1 (Intrinsic regularity). *Let G be a graph satisfying the assumptions of Theorem 1.3. Let $p = p_c(1+\varepsilon)$, $r = r(m) = M/\varepsilon$ where $M = M(m)$ is defined in (2.7), and $r_1 \in [\varepsilon^{-1}, r_0]$, where r_0 is defined in (2.8). For any $\delta, \beta \in (0,1)$ there exist at least $(1 - O(\omega^{1/4}))r$ radii $j_x \in [r, 2r]$ such that*

$$\mathbf{P}(x \text{ is } (\delta, \beta, j_x, r_1)\text{-fit}) \geq \big(1 - O(\delta^{-1}\beta^{-2}\mathrm{e}^{2M}\omega^{1/4})\big)\mathbf{P}(\partial B_x(j_x) \neq \emptyset).$$

We start by proving some preparatory lemmas:

Lemma 5.2. *Assume the setting of Theorem 5.1. Then*

$$\sum_{j_x=r}^{2r}\sum_{a,u}\mathbf{P}\big(x \overset{=j_x}{\longleftrightarrow} a,\ a \overset{P[2t_{\text{mix}},r_1]}{\longleftrightarrow} u \text{ off } B_x(j_x)\setminus\{a\}\big) \geq (1-O(\omega))\mathbb{E}|B([r,2r])|\,\mathbb{E}|B(r_1)|.$$

Proof. We condition on $B_x(j_x) = A$ for any admissible A (that is, A for which the event $x \overset{=j_x}{\longleftrightarrow} a$ occurs and $\mathbf{P}(B_x(j_x) = A) > 0$). Then

$$\mathbf{P}\big(a \overset{P[2t_{\text{mix}},r_1]}{\longleftrightarrow} u \text{ off } B_x(j_x)\setminus\{a\} \bigm| B_x(j_x) = A\big) = \mathbf{P}(a \overset{P[2t_{\text{mix}},r_1]}{\longleftrightarrow} u \text{ off } A\setminus\{a\}),$$

and

$$\mathbf{P}(a \overset{P[2t_{\text{mix}},r_1]}{\longleftrightarrow} u \text{ off } A\setminus\{a\}) = \mathbf{P}(a \overset{P[2t_{\text{mix}},r_1]}{\longleftrightarrow} u) - \mathbf{P}(a \overset{P[2t_{\text{mix}},r_1]}{\longleftrightarrow} u \text{ only on } A\setminus\{a\}).$$

Summing over the first term gives

$$\begin{aligned}\sum_{a,u,\ j_x\in[r,2r]}\mathbf{P}(x \overset{=j_x}{\longleftrightarrow} a)\mathbf{P}(a \overset{P[2t_{\text{mix}},r_1]}{\longleftrightarrow} u) &\geq \mathbb{E}|B([r,2r])|\big(\mathbb{E}|B(r_1)| - \mathbb{E}|B(2t_{\text{mix}})|\big)\\ &= (1-O(\omega))\mathbb{E}|B(r_1)|\,\mathbb{E}|B([r,2r])|,\end{aligned}$$

since $\mathbb{E}|B(2t_{\text{mix}})| \leq Ct_{\text{mix}}$ by Corollary 3.5 and $\mathbb{E}|B(r_1)| \geq c\varepsilon^{-1}$ by Theorem 4.5 and since $r_1 \geq \varepsilon^{-1}$. It remains to bound the sum

$$\sum_{a,u,\ j_x\in[r,2r]}\sum_{A}\mathbf{P}(B_x(j_x) = A)\mathbf{P}(a \overset{P[2t_{\text{mix}},r_1]}{\longleftrightarrow} u \text{ only on } A\setminus\{a\}).$$

As usual, if $a \xleftrightarrow{P[2t_{\mathrm{mix}},r_1]} u$ only on $A \setminus \{a\}$ occurs, then there exists $z \in A$ such that $\{a \overset{r_1}{\leftrightarrow} z\} \circ \{z \overset{r_1}{\leftrightarrow} u\}$. The BK inequality now gives

$$\sum_{a,u,\ j_x\in[r,2r]} \sum_{A} \mathbf{P}(B_x(j_x) = A) \sum_{z\in A\setminus\{a\}} \mathbf{P}(a \overset{r_1}{\leftrightarrow} z)\mathbf{P}(z \overset{r_1}{\leftrightarrow} u). \tag{5.2}$$

We sum over u and extract a factor of $\mathbb{E}|B(r_1)|$. We then change the order of summation, so the sum simplifies to

$$\mathbb{E}|B(r_1)| \sum_{a,\ z\neq a,\ j_x\in[r,2r]} \mathbf{P}(x \xleftrightarrow{=j_x} a,\ x \overset{j_x}{\leftrightarrow} z)\mathbf{P}(a \overset{r_1}{\leftrightarrow} z).$$

We sum over j_x (noting that the events $x \xleftrightarrow{=j_x} a,\ x \overset{j_x}{\leftrightarrow} z$ are disjoint as j_x varies) and bound this sum by

$$\mathbb{E}|B(r_1)| \sum_{a,\ z\neq a} \mathbf{P}(x \overset{2r}{\leftrightarrow} a,\ x \overset{2r}{\leftrightarrow} z)\mathbf{P}(a \overset{r_1}{\leftrightarrow} z).$$

As usual, if $x \overset{2r}{\leftrightarrow} a$ and $x \overset{2r}{\leftrightarrow} z$, then there exists z' such that the event

$$\{x \overset{2r}{\leftrightarrow} z'\} \circ \{z' \overset{2r}{\leftrightarrow} z\} \circ \{z' \overset{2r}{\leftrightarrow} a\}$$

occurs. By the BK inequality we bound the above sum by

$$\mathbb{E}|B(r_1)| \sum_{a,z',\ z\neq a} \mathbf{P}(x \overset{2r}{\leftrightarrow} z')\mathbf{P}(z' \overset{2r}{\leftrightarrow} z)\mathbf{P}(z' \overset{2r}{\leftrightarrow} a)\mathbf{P}(a \overset{r_1}{\leftrightarrow} z).$$

We may now sum over a and $z \neq a$ using Corollary 3.18 and then sum over z' to deduce that this is bounded by

$$C\mathbb{E}|B(r_1)|\,\mathbb{E}|B(2r)|\big[\omega + \mathbb{E}|B(r_1)|(\mathbb{E}|B(2r)|)^2/V\big].$$

This concludes our proof since the second term in brackets is of order at most $\alpha_G^{1/2}\mathrm{e}^{4M}(\varepsilon^3 V)^{-1/2} \leq \alpha_G^{1/2}$ by the upper bound on r_1, our choice of r and M in (2.7), and Corollary 4.6. □

Lemma 5.3. *Assume the setting of Theorem 5.1. There exists a $C > 0$ such that*

$$\sum_{j_x=r}^{2r} \sum_{a,u} \mathbf{P}\big(x \xleftrightarrow{=j_x} a,\ a \xleftrightarrow{P[2t_{\mathrm{mix}},r_1]} u \text{ off } B_x(j_x) \setminus \{a\}, a \in \mathrm{Piv}(\{x \xleftrightarrow{j_x+r_1} u\})\big)$$
$$\geq (1 - O(\omega))\mathbb{E}|B([r, 2r])|\,\mathbb{E}|B(r_1)|.$$

Proof. Fix $j_x \in [r, 2r]$. We rely on Lemma 5.2, and bound the difference in the probabilities appearing in Lemma 5.2 and the one above. If the event

$$\big\{x \xleftrightarrow{=j_x} a,\ a \xleftrightarrow{P[2t_{\mathrm{mix}},r_1]} u \text{ off } B_x(j_x) \setminus \{a\}\big\} \tag{5.3}$$

occurs but $a \notin \mathrm{Piv}(\{x \xleftrightarrow{j_x+r_1} u\})$, then there exist z_1, z_2 and $t \leq j_x$ and paths $\eta_1, \eta_2, \gamma_1, \gamma_2, \gamma_3$ such that

(a) γ_1 is an open path of length at most r_1 connecting a to z_2,
(b) γ_2 is an open path of length at most r_1 connecting z_2 to u,
(c) γ_3 is an open path of length at most r_1+j_x connecting z_1 to z_2,
(d) η_1 is a shortest open path of length precisely t connecting x to z_1,
(e) η_2 is a shortest open path of length precisely $j_x - t$ connecting z_1 to a,
(f) $\gamma_1, \gamma_2, \gamma_3, \eta_1, \eta_2$ are disjoint

(see Figure 3). Indeed, assume that a is not pivotal for $x \xleftrightarrow{j_x+r_1} u$ and (5.3) holds. Let η be the lexicographically first shortest open path of length j_x between x and a, and γ a disjoint open path of length in $[2t_{\text{mix}}, r_1]$ between a and u off $B_x(j_x) \setminus \{a\}$, which exists since (5.3) holds. Since a is not pivotal, we learn that there exist another open path β between x and u of length at most $j_x + r_1$ that does *not* visit a. Hence, β goes "around" a, or in formal words, there exist vertices z_1 and z_2 on β appearing on it in that order such that $z_1 \in \eta$ and $z_2 \in \gamma$ and the part of β between z_1 and z_2 is disjoint from $\eta \cup \gamma$. We take $t < j_x$ such that $\eta(t) = z_1$ and set $\eta_1 = \eta[0, t]$ and $\eta_2 = \eta[t, j_x]$. We let γ_3 be the section of β between z_1 and z_2, and γ_1, γ_2 be the sections of γ from a to z_2 and from z_2 to u, respectively.

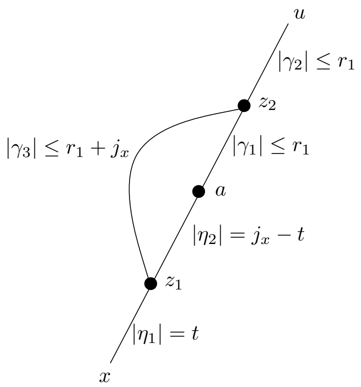

Fig. 3. a is not pivotal for the event $x \xleftrightarrow{j_x+r_1} u$.

For all $j_x, t \in [r, 2r]$ these events (that is, the existence of z_1, z_2 and of the disjoint paths) are disjoint, since η_1 and η_2 are required to be shortest open paths. The union of these events over j_x, t implies that there exist z_1, z_2 such that

$$\{x \xleftrightarrow{2r} z_1\} \circ \{z_1 \xleftrightarrow{2r} a\} \circ \{a \xleftrightarrow{r_1} z_2\} \circ \{z_1 \xleftrightarrow{r_1+2r} z_2\} \circ \{z_2 \xleftrightarrow{r_1} u\},$$

since we can just take $\eta_1, \eta_2, \gamma_1, \gamma_2, \gamma_3$ as our disjoint witnesses. Using the BK inequality we bound the required sum from above by

$$\sum_{\substack{a,u,z_1,z_2 \\ z_1 \neq z_2,\, z_1 \neq a,\, z_2 \neq a}} \mathbf{P}(x \xleftrightarrow{2r} z_1)\mathbf{P}(z_1 \xleftrightarrow{2r} a)\mathbf{P}(a \xleftrightarrow{r_1} z_2)\mathbf{P}(z_1 \xleftrightarrow{r_1+2r} z_2)\mathbf{P}(z_2 \xleftrightarrow{r_1} u).$$

Summing first over u extracts a factor of $\mathbb{E}|B(r_1)|$; we then sum over a and z_2 using Corollary 3.18, and lastly we sum over z_1. This gives a bound of

$$C\mathbb{E}|B(r_1)|\, \mathbb{E}|B(2r)| \big[\omega + \mathbb{E}|B(r_1)|\, \mathbb{E}|B(r_1 + 2r)|\, \mathbb{E}|B(2r)|/V\big].$$

We apply Lemma 5.2 to conclude the proof since the second term in brackets is of order at most $\alpha_G e^{4M} \le \alpha_G^{1/2}$ by the upper bound on r_1, our choice of r and M in (2.7), and Corollary 4.6. Also note that $\mathbb{E}|B([r,2r])| \ge \mathbb{E}|B(2r)|/2$ by Corollary 4.6 and our choice of r and M. $\square$

Lemma 5.4. *Assume the setting of Theorem 5.1. For any vertices x, a,*

$$\sum_{j_x=r}^{2r}\sum_u \mathbf{P}\big(x \overset{=j_x}{\longleftrightarrow} a,\ a \overset{P[2t_{\mathrm{mix}},r_1]}{\longleftrightarrow} u \textit{ off } B_x(j_x)\setminus\{a\}\big) \le (1+O(\omega))V^{-1}\mathbb{E}|B(r_1)|\,\mathbb{E}|B([r,2r])|.$$

Proof. The event $\{x \overset{=j_x}{\longleftrightarrow} a,\ a \overset{P[2t_{\mathrm{mix}},r_1]}{\longleftrightarrow} u \text{ off } B_x(j_x)\setminus\{a\}\}$ implies that

$$\{x \overset{=j_x}{\longleftrightarrow} a\} \circ \{a \overset{r_1}{\leftrightarrow} u\},$$

the second witness is the open edges of an open path of length in $[2t_{\mathrm{mix}}, r_1]$ off $B_x(j_x)\setminus\{a\}$, and the first witness is the lexicographically first shortest open path of length j_x between x and a together with all the closed edges of the graph. The BK-Reimer inequality gives

$$\mathbf{P}\big(x \overset{=j_x}{\longleftrightarrow} a,\ a \overset{P[2t_{\mathrm{mix}},r_1]}{\longleftrightarrow} u \text{ off } B_x(j_x)\setminus\{a\}\big) \le \mathbf{P}(x \overset{=j_x}{\longleftrightarrow} a)\mathbf{P}(a \overset{r_1}{\leftrightarrow} u).$$

We sum over u and $j_x \in [r,2r]$ to find that the sum is bounded by

$$\mathbb{E}|B(r_1)|\mathbf{P}(x \overset{[r,2r]}{\longleftrightarrow} a).$$

Lemma 3.14 gives

$$\mathbf{P}(x \overset{[r,2r]}{\longleftrightarrow} a) \le (1+O(\omega))V^{-1}\mathbb{E}|B([r-t_{\mathrm{mix}}, 2r-t_{\mathrm{mix}}])|.$$

We have

$$\begin{aligned}\mathbb{E}|B([r-t_{\mathrm{mix}}, 2r-t_{\mathrm{mix}}])| &\le \mathbb{E}|B([r,2r])| + \mathbb{E}|B([r-t_{\mathrm{mix}}, r])|\\ &\le (1+O(\varepsilon t_{\mathrm{mix}}))\mathbb{E}|B([r,2r])|\end{aligned}$$

since $\mathbb{E}|B([r-t_{\mathrm{mix}}, r])| \le Ct_{\mathrm{mix}}(1+\varepsilon)^r$ by Theorem 4.1 and Corollary 3.5 and since $\mathbb{E}|B([r,2r])| \ge c\varepsilon^{-1}(1+\varepsilon)^{2r}$ by Corollary 4.6 (we use the assumption $r \gg \varepsilon^{-1}$). Hence

$$\mathbf{P}(x \overset{[r,2r]}{\longleftrightarrow} a) \le (1+O(\omega))V^{-1}\mathbb{E}|B([r,2r])|, \tag{5.4}$$

concluding our proof. $\square$

Proof of Theorem 5.1. By combining Lemmas 5.3 and 5.4 we deduce that for any x there exist at least $(1-O(\omega^{1/2}))V$ vertices a such that

$$\sum_{j_x=r}^{2r}\sum_u \mathbf{P}\big(x \overset{=j_x}{\longleftrightarrow} a,\ a \overset{P[2t_{\mathrm{mix}},r_1]}{\longleftrightarrow} u \text{ off } B_x(j_x)\setminus\{a\}, a \in \mathrm{Piv}(\{x \overset{j_x+r_1}{\longleftrightarrow} u\})\big) = (1+O(\omega^{1/2}))V^{-1}\mathbb{E}|B(r_1)|\,\mathbb{E}|B([r,2r])|.$$

Write $\widetilde{G}$ for the variable

$$\widetilde{G} = \sum_{j_x=r}^{2r} G_{j_x,r_1}(a,x;B_x(j_x))\mathbf{1}_{\{x \overset{=j_x}{\longleftrightarrow} a\}}.$$

Note that $\widetilde{G}$ is a random variable that is measurable with respect to $B_x(2r)$ (that is, it is determined by the status of the edges touching $B_x(2r)$) and that it equals 0 unless $x \xleftrightarrow{[r,2r]} a$. Furthermore, only one of the summands can be nonzero because the events in the indicators are disjoint. Our previous approximate equality can be rewritten as

$$\mathbb{E}\widetilde{G} = (1 + O(\omega^{1/2}))V^{-1}\mathbb{E}|B(r_1)|\,\mathbb{E}|B([r,2r])|.$$

Hence, for at least $(1 - O(\omega^{1/2}))V$ vertices a,

$$\mathbb{E}[\widetilde{G} \mid x \xleftrightarrow{[r,2r]} a] \geq (1 - O(\omega^{1/2}))\mathbb{E}|B(r_1)|, \tag{5.5}$$

by Lemma 3.14. This gives the conditional first moment estimate. The second moment calculation is somewhat easier. We have

$$\mathbb{E}\widetilde{G}^2 = \sum_{j_x=r}^{2r} \sum_{u_1,u_2} \mathbb{E}\Big[\mathbf{P}\big(a \xleftrightarrow{P[2t_{\mathrm{mix}},r_1]} u_1 \text{ off } B_x(j_x) \setminus \{a\} \bigm| B_x(j_x)\big)$$
$$\times \mathbf{P}\big(a \xleftrightarrow{P[2t_{\mathrm{mix}},r_1]} u_2 \text{ off } B_x(j_x) \setminus \{a\} \bigm| B_x(j_x)\big)\mathbf{1}_{\{x \overset{=j_x}{\longleftrightarrow} a\}}\Big].$$

We bound, almost surely in $B_x(j_x)$ and for $i = 1, 2$,

$$\mathbf{P}\big(a \xleftrightarrow{P[2t_{\mathrm{mix}},r_1]} u_i \text{ off } B_x(j_x) \setminus \{a\} \bigm| B_x(j_x)\big) \leq \mathbf{P}(a \overset{r_1}{\longleftrightarrow} u_i),$$

and sum over u_1 and u_2 to get

$$\mathbb{E}\widetilde{G}^2\mathbf{1}_{\{x \xleftrightarrow{[r,2r]} a\}} \leq [\mathbb{E}|B(r_1)|]^2\mathbf{P}(x \xleftrightarrow{[r,2r]} a),$$

so that

$$\mathbb{E}[\widetilde{G}^2 \mid x \xleftrightarrow{[r,2r]} a] \leq [\mathbb{E}|B(r_1)|]^2.$$

Combining this with (5.5), we obtain

$$\mathrm{Var}\big(\widetilde{G} \mid x \xleftrightarrow{[r,2r]} a\big) = O\big([\mathbb{E}|B(r_1)|]^2\omega^{1/2}\big).$$

By Chebyshev's inequality, for any $\beta > 0$,

$$\mathbf{P}\big(\widetilde{G} \leq (1-\beta)\mathbb{E}|B(r_1)| \bigm| x \xleftrightarrow{[r,2r]} a\big) = O(\beta^{-2}\omega^{1/2}).$$

Recall that this holds for at least $(1 - O(\omega^{1/2}))V$ vertices a. Call these vertices *valid*. We have

$$\sum_{a \text{ valid}} \mathbf{P}\big(x \xleftrightarrow{[r,2r]} a,\ \widetilde{G} \leq (1-\beta)\mathbb{E}|B(r_1)|\big) = O\big(\mathbb{E}|B([r,2r])|\beta^{-2}\omega^{1/2}\big),$$

by our previous estimate. Also, since there are at most $O(\omega^{1/2}V)$ invalid a's, we apply (5.4) to bound the sum over all a by

$$\sum_{a} \mathbf{P}\big(x \xleftrightarrow{[r,2r]} a,\ \widetilde{G} \leq (1-\beta)\mathbb{E}|B(r_1)|\big) = O\big(\mathbb{E}|B([r,2r])|\beta^{-2}\omega^{1/2}\big).$$

Returning to our original notation, we rewrite this as

$$\sum_{j_x=r}^{2r} \sum_a \mathbf{P}\big(x \overset{=j_x}{\longleftrightarrow} a,\ G_{j_x,r_1}(a,x;\,B_x(j_x)) \le (1-\beta)\mathbb{E}|B(r_1)|\big)$$
$$= O\big(\mathbb{E}|B([r,2r])|\beta^{-2}\omega^{1/2}\big).$$

Hence, there are at least $(1-O(\omega^{1/4}))r$ radii $j_x \in [r,2r]$ such that

$$\sum_a \mathbf{P}\big(x \overset{=j_x}{\longleftrightarrow} a,\ G_{j_x,r_1}(a,x;\,B_x(j_x)) \le (1-\beta)\mathbb{E}|B(r_1)|\big)$$
$$= O\big(\mathbb{E}|B([r,2r])|r^{-1}\beta^{-2}\omega^{1/4}\big) = O(\mathrm{e}^{2M}\beta^{-2}\omega^{1/4}),$$

where the last inequality is by Lemma 4.4. Given such j_x, write $X(j_x)$ for the random variable

$$X(j_x) = \big|\big\{a\colon x \overset{=j_x}{\longleftrightarrow} a,\ G_{j_x,r_1}(a,x;\,B_x(j_x)) \le (1-\beta)\mathbb{E}|B(r_1)|\big\}\big|,$$

so that $\mathbb{E}X(j_x) \le C\mathrm{e}^{2M}\beta^{-2}\omega^{1/4}$. The variable $X(j_x)$ equals the number of (β, j_x, r_1)-non-regenerative vertices. By Markov's inequality we deduce that for any $\delta > 0$,

$$\mathbf{P}(X(j_x) \ge \delta\varepsilon^{-1}) = O(\varepsilon\delta^{-1}\beta^{-2}\mathrm{e}^{2M}\omega^{1/4}),$$

and we conclude by Lemma 3.6 that at least $(1-O(\omega^{1/4}))r$ radii $j_x \in [r,2r]$ satisfy

$$\mathbf{P}\big(\partial B_x(j_x) \ne \emptyset \text{ and } X(j_x) \le \delta\varepsilon^{-1}\big) \ge \big(1 - O(\delta^{-1}\beta^{-2}\mathrm{e}^{2M}\omega^{1/4})\big)\mathbf{P}(\partial B_x(j_x) \ne \emptyset),$$

as required. □

6. Large clusters are close

In this section, we prove Theorem 2.4 which shows that many closed edges exist between most large clusters. This section involves all our notation from the previous sections and in particular the parameters $V, m, t_{\mathrm{mix}}, \varepsilon(m), \alpha_G$. We define $\beta, k, \ell, \zeta, \delta$ as

$$\beta = (\log M)^{-2}, \quad k = \frac{M}{\log M}, \quad \ell = (\log M)^{1/4}, \quad \zeta = (\log M)^{-1/8}, \quad \delta = \zeta/2. \tag{6.1}$$

For notational convenience we also denote

$$\{a \overset{P[2t_{\mathrm{mix}},r_0],x}{\longleftrightarrow} u\} = \{a \overset{P[2t_{\mathrm{mix}},r_0]}{\longleftrightarrow} u\} \cap \{a \in \mathrm{Piv}(x \overset{j_x+r_0}{\longleftrightarrow} u)\},$$
$$\{b \overset{P[2t_{\mathrm{mix}},r_0],y}{\longleftrightarrow} u'\} = \{b \overset{P[2t_{\mathrm{mix}},r_0]}{\longleftrightarrow} u'\} \cap \{b \in \mathrm{Piv}(y \overset{j_y+r_0}{\longleftrightarrow} u')\}.$$

Let $S_{j_x,j_y,r_0}(x,y)$ be the random variable counting the number of directed edges (u,u') such that there exist vertices a, b with

- $\mathscr{A}(x,y,j_x,j_y)$,
- $x \overset{=j_x}{\longleftrightarrow} a$ and $y \overset{=j_y}{\longleftrightarrow} b$,

- $a \xleftrightarrow{P[2t_{\mathrm{mix}}, r_0], x} u$,
- $b \xleftrightarrow{P[2t_{\mathrm{mix}}, r_0], y} u'$ off $B_x(j_x + r_0)$.

Further define

$$\widehat{S}_{2r,2r,r_0}(x,y) = \Big|\Big\{(u,u')\colon \{x \xleftrightarrow{2r+r_0} u\} \circ \{y \xleftrightarrow{2r+r_0} u'\},\\ |B_u(2r+r_0)| \cdot |B_{u'}(2r+r_0)| \geq \mathrm{e}^{40M} \varepsilon^{-2} (\mathbb{E}|B(r_0)|)^2\Big\}\Big|.$$

We will use the fact that for any $j_x, j_y \in \{r, \ldots, 2r\}^2$,

$$S_{2r+r_0}(x,y) \geq S_{j_x,j_y,r_0}(x,y) - \widehat{S}_{2r,2r,r_0}(x,y), \tag{6.2}$$

where $S_{2r+r_0}(x,y)$ is the random variable defined prior to Theorem 2.4. Finally, write $\mathscr{A}(x,y,j_x,j_y,r_0,\beta,k)$ for the intersection of the events

(1) $\mathscr{A}(x,y,j_x,j_y)$,
(2) $|B_x(j_x)| \leq \varepsilon^{-2}(1+\varepsilon)^{3r}$ and $|B_y(j_y)| \leq \varepsilon^{-2}(1+\varepsilon)^{3r}$ and $|\partial B_x(j_x)| \leq \varepsilon^{-1}(1+\varepsilon)^{3r}$,
(3) $|\partial B_x(j_x)| \geq \mathrm{e}^{k/4}\varepsilon^{-1}$ and $|\partial B_y(j_y)| \geq \mathrm{e}^{k/4}\varepsilon^{-1}$,
(4) x is $(1,\beta,j_x,r_0)$-fit and y is $(1,\beta,j_y,r_0)$-fit,
(5)
$$\mathbb{E}\big[S_{j_x,j_y,r_0}(x,y)\mathbf{1}_{\{\partial B_x(j_x) \overset{2r_0}{\leftrightarrow} \partial B_y(j_y)\}} \mathrel{\big|} B_x(j_x), B_y(j_y)\big] \leq V^{-1} m \varepsilon^{-2} (\mathbb{E}|B(r_0)|)^2 \alpha_G^{1/2}.$$

This event is measurable with respect to $B_x(j_x), B_y(j_y)$. The following three statements will prove Theorem 2.4:

Lemma 6.1. *Assume the setting of Theorem 2.4. Then*

$$\mathbb{E}\big|\big\{(x,y)\colon \mathscr{A}(x,y,2r,2r) \text{ and } \widehat{S}_{2r,2r,r_0}(x,y) \geq \beta^{1/2} V^{-1} m \varepsilon^{-2} (\mathbb{E}|B(r_0)|)^2\big\}\big| = o(\varepsilon^2 V^2).$$

Theorem 6.2. *Assume the setting of Theorem 2.4. Then there exist radii $j_1, \ldots, j_\ell \in [r, 2r]$ such that for at least $(1-o(1))V^2$ pairs (x,y),*

$$\mathbf{P}\Big(\mathscr{A}(x,y,2r,2r) \text{ and } \bigcap_{j_x,j_y \in \{j_1,\ldots,j_\ell\}^2} \mathscr{A}(x,y,j_x,j_y,r_0,\beta,k)^c\Big) = o(\varepsilon^2).$$

Theorem 6.3. *Assume the setting of Theorem 2.4 and let (x,y) be a pair of vertices. Then, for any radii $j_x, j_y \in \{r, \ldots, 2r\}^2$,*

$$\mathbf{P}\big(S_{j_x,j_y,r_0}(x,y) \leq 2\beta^{1/2} V^{-1} m \varepsilon^{-2} (\mathbb{E}|B(r_0)|)^2 \text{ and } \mathscr{A}(x,y,j_x,j_y,r_0,\beta,k)\big)\\ = O(\beta^{1/2}\varepsilon^2).$$

Proof of Theorem 2.4 subject to Lemma 6.1 and Theorems 6.2–6.3. Lemma 3.8 shows that

$$\frac{|\{(x,y)\colon \mathscr{A}(x,y,2r,2r)\}|}{4\varepsilon^2 V^2} \xrightarrow{\mathbf{P}} 1.$$

Thus, it suffices to prove that

$$\frac{|\{(x,y)\colon \mathscr{A}(x,y,2r,2r) \text{ and } x,y \text{ are not } (r,r_0)\text{-good}\}|}{\varepsilon^2 V^2} \xrightarrow{\mathbf{P}} 0.$$

Lemma 3.10 shows that

$$\frac{|\{(x,y)\colon \mathscr{A}(x,y,2r,2r) \text{ and } |\mathscr{C}(x)| \le (\varepsilon^3 V)^{1/4}\varepsilon^{-2} \text{ or } |\mathscr{C}(y)| \le (\varepsilon^3 V)^{1/4}\varepsilon^{-2}\}|}{\varepsilon^2 V^2} \xrightarrow{\mathbf{P}} 0.$$

It remains to handle requirement (3) in the definition of (r,r_0)-good. Let $j_1,\dots,j_\ell$ be the radii guaranteed to exist by Theorem 6.2 and let (x,y) be a pair of vertices for which the assertion of Theorem 6.2 holds. Theorem 6.2 asserts that the number of such pairs is $(1-o(1))V^2$, so the sum of $\mathbf{P}(\mathscr{A}(x,y,2r,2r))$ over pairs not counted is $o(\varepsilon^2 V^2)$. Write $(J(x),J(y))$ for the lexicographically first pair $(j_x,j_y) \in \{j_1,\dots,j_\ell\}^2$ for which the event $\mathscr{A}(x,y,j_x,j_y,r_0,\beta,k)$ occurs, or set $J(x)=J(y)=\infty$ if no such j_x,j_y exist. Then for at least $(1-o(1))V^2$ pairs (x,y),

$$\mathbf{P}\big(\mathscr{A}(x,y,2r,2r),\ J(x)=\infty,\ J(y)=\infty\big) = o(\varepsilon^2).$$

Theorem 6.3 together with the union bound implies that for any such pair (x,y),

$$\sum_{j_x,j_y\in\{j_1,\dots,j_\ell\}^2} \mathbf{P}\big(S_{j_x,j_y,r_0}(x,y) \le 2\beta^{1/2}V^{-1}m\varepsilon^{-2}(\mathbb{E}|B(r_0)|)^2,\ J(x)=j_x,\ J(y)=j_y\big) \\ = O(\beta^{1/2}\ell^2\varepsilon^2),$$

which is $o(\varepsilon^2)$ by our choice of ℓ and β in (6.1). By these last two statements we deduce that

$$\mathbb{E}\big|\big\{(x,y)\colon \mathscr{A}(x,y,2r,2r) \text{ and } \forall j_x,j_y\ \ S_{j_x,j_y,r_0}(x,y) \le 2\beta^{1/2}V^{-1}m\varepsilon^{-2}(\mathbb{E}|B(r_0)|)^2\big\}\big| \\ = o(\varepsilon^2 V^2).$$

This together with (6.2) and Lemma 6.1 implies that

$$\mathbb{E}\big|\big\{(x,y)\colon \mathscr{A}(x,y,2r,2r) \text{ and } S_{2r+r_0}(x,y) \le \beta^{1/2}V^{-1}m\varepsilon^{-2}(\mathbb{E}|B(r_0)|)^2\big\}\big| = o(\varepsilon^2 V^2),$$

concluding our proof since $\beta^{1/2} = (\log M)^{-1}$. □

6.1. Proof of Lemma 6.1: Bounding the error $\widehat{S}_{2r,2r,r_0}$. We begin by providing some useful estimates.

Lemma 6.4. *Assume the setting of Theorem 2.4 and let $p = p_c(1+\varepsilon)$. There exists $C>0$ such that for any positive integer n:*

(1) $\displaystyle\sum_{x,y,(u,u')} \mathbf{P}(\{u \overset{n}{\leftrightarrow} x\}\circ\{u' \overset{n}{\leftrightarrow} y\} \text{ and } u \overset{2n}{\leftrightarrow} u') \le C[m\varepsilon^{-5}(1+\varepsilon)^{4n} + \alpha_G V m\varepsilon^{-2}(1+\varepsilon)^{2n}].$

(2) $\displaystyle\sum_{x,y,(u,u')} \mathbf{P}(\{u \overset{n}{\leftrightarrow} x\}\circ\{u' \overset{n}{\leftrightarrow} y\} \text{ and } x \overset{2n}{\leftrightarrow} y) \le C[m\varepsilon^{-5}(1+\varepsilon)^{4n} + \alpha_G V m\varepsilon^{-2}(1+\varepsilon)^{2n}].$

Proof. We begin by showing (1). If $\{u \overset{n}{\longleftrightarrow} x\} \circ \{u' \overset{n}{\longleftrightarrow} y\}$ and $u \overset{2n}{\longleftrightarrow} u'$, then there exist vertices z_1, z_2 and integers $t_1, t_2 \le n$ such that the event

$$\{u \overset{=t_1}{\longleftrightarrow} z_1,\ u' \overset{=t_2}{\longleftrightarrow} z_2,\ d_{G_p}(u,u') \ge t_1+t_2\} \circ \{z_1 \overset{2n-t_1-t_2}{\longleftrightarrow} z_2\} \circ \{x \overset{n-t_1}{\longleftrightarrow} z_1\} \circ \{y \overset{n-t_2}{\longleftrightarrow} z_2\}$$

or the event

$$\{u \overset{=t_1}{\longleftrightarrow} z_1,\ u' \overset{=t_2}{\longleftrightarrow} z_2,\ d_{G_p}(u,u') \ge t_1+t_2\} \circ \{z_1 \overset{2n-t_1-t_2}{\longleftrightarrow} z_2\} \circ \{x \overset{n-t_1}{\longleftrightarrow} z_2\} \circ \{y \overset{n-t_2}{\longleftrightarrow} z_1\}$$

occurs (see Figure 4(a)). Indeed, let η be the shortest open path of length at most $2n$ between u and u' and let $\gamma_{x,u}, \gamma_{y,u'}$ be two disjoint paths of length at most n connecting x to u and y to u', respectively. We let z_1, z_2 be the first vertices of $\gamma_{x,u}$ and $\gamma_{y,u'}$ which belong to η. There are two possible orderings of z_1, z_2 on η, (u, z_1, z_2, u') or (u, z_2, z_1, u'), which give the two possible events. Assume the ordering on η is (u, z_1, z_2, u') (the two orderings give rise to identical contributions to the sum in (1)); let t_1, t_2 be the distances on η between u and z_1 and between z_2 and u', respectively; and write η_1, η_2 for the corresponding sections of η, and η_3 for the section of η between z_1 and z_2. The paths γ_1 and γ_2 are the sections of $\gamma_{x,u}$ and $\gamma_{y,u'}$ from x to z_1 and from y to z_2, respectively. The witness for the first event is η_1, η_2 together with all the closed edges of G_p (the closed edges determine that η_1, η_2 are indeed shortest open paths, and that $d_{G_p}(u,u') \ge t_1+t_2$); for the second, third and fourth events, the witnesses are just η_3, γ_1 and γ_2, respectively.

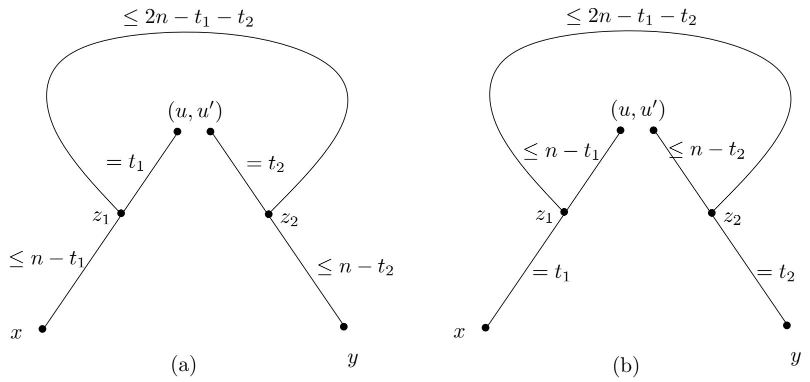

Fig. 4. The edge (u, u') is counted in the first and second sum of Lemma 6.4.

We now apply the BK-Reimer inequality and bound the sum in (1) by

$$2 \sum_{x,y,z_1,z_2,(u,u'),t_1\le n,t_2\le n} \mathbf{P}\big(u \overset{=t_1}{\longleftrightarrow} z_1,\ u' \overset{=t_2}{\longleftrightarrow} z_2,\ d_{G_p}(u,u') \ge t_1+t_2\big)\mathbf{P}(z_1 \overset{2n-t_1-t_2}{\longleftrightarrow} z_2)$$
$$\times \mathbf{P}(x \overset{n-t_1}{\longleftrightarrow} z_1)\mathbf{P}(y \overset{n-t_2}{\longleftrightarrow} z_2).$$

We first sum over x, y and get a factor of $C\varepsilon^{-2}(1+\varepsilon)^{2n-t_1-t_2}$ by Lemma 4.4. The event $\{u \overset{=t_1}{\longleftrightarrow} z_1, u' \overset{=t_2}{\longleftrightarrow} z_2, d_{G_p}(u,u') \ge t_1+t_2\}$ implies that $u \overset{=t_1}{\longleftrightarrow} z_1$ and $u' \overset{=t_2}{\longleftrightarrow} z_2$

off $B_u(t_1)$, hence we may bound its probability by

$$\sum_{A:\, u \overset{=t_1}{\longleftrightarrow} z_1} \mathbf{P}(B_u(t_1) = A)\mathbf{P}(u' \overset{=t_2}{\longleftrightarrow} z_2 \text{ off } A),$$

and so we get an upper bound of

$$C\varepsilon^{-2} \sum_{z_1,z_2,(u,u'),\, t_1 \le n,\, t_2 \le n} (1+\varepsilon)^{2n-t_1-t_2}\mathbf{P}(u \overset{=t_1}{\longleftrightarrow} z_1) \times \max_A \mathbf{P}(u' \overset{=t_2}{\longleftrightarrow} z_2 \text{ off } A)\mathbf{P}(z_1 \overset{2n-t_1-t_2}{\longleftrightarrow} z_2). \tag{6.3}$$

We bound this in two parts. If $t_2 \ge t_{\text{mix}}$, then we use Lemma 3.15 together with Lemma 4.4 to bound, uniformly in A, $\mathbf{P}(u' \overset{=t_2}{\longleftrightarrow} z_2 \text{ off } A) \le CV^{-1}(1+\varepsilon)^{t_2}$. We then sum over z_2 and z_1 in that order using Lemma 4.4 and extract a Vm factor from summing over (u,u'). If $t_2 \le t_{\text{mix}}$ and $t_1 \ge t_{\text{mix}}$, then we use Lemma 3.13 together with Lemma 4.4 to bound $\mathbf{P}(u \overset{=t_1}{\longleftrightarrow} z_1) \le CV^{-1}(1+\varepsilon)^{t_1}$. Further, we use condition (2) of Theorem 1.3 and $\varepsilon = o(1/t_{\text{mix}})$ to bound, uniformly in A, $\mathbf{P}(u' \overset{=t_2}{\longleftrightarrow} z_2 \text{ off } A) \le C\mathbf{p}^{t_2}(u', z_2)$. We then sum over z_1 and z_2 in that order using Lemma 4.4 and extract a Vm factor from summing over (u,u'). All this gives an upper bound of

$$Cm\varepsilon^{-3}(1+\varepsilon)^{4n} \sum_{t_1,t_2 \le n} (1+\varepsilon)^{-t_1-t_2} \le Cm\varepsilon^{-5}(1+\varepsilon)^{4n},$$

as required. We next sum (6.3) over $t_1, t_2 \le t_{\text{mix}}$. We first relax $(1+\varepsilon)^{2n-t_1-t_2} \le (1+\varepsilon)^{2n}$ and $\mathbf{P}(z_1 \overset{2n-t_1-t_2}{\longleftrightarrow} z_2) \le \mathbf{P}(z_1 \overset{2n}{\longleftrightarrow} z_2)$, and then sum over t_1, t_2 to get an upper bound of

$$C\varepsilon^{-2}(1+\varepsilon)^{2n} \sum_{z_1,z_2,(u,u')} \mathbf{P}(u \overset{t_{\text{mix}}}{\longleftrightarrow} z_1)\mathbf{P}(u' \overset{t_{\text{mix}}}{\longleftrightarrow} z_2)\mathbf{P}(z_1 \overset{2n}{\longleftrightarrow} z_2).$$

We now sum over z_1, z_2 using Corollary 3.18 and Lemma 4.4. We conclude that this is bounded by

$$CVm\varepsilon^{-2}(1+\varepsilon)^{2n}[t_{\text{mix}}^2\varepsilon^{-1}(1+\varepsilon)^{2n}/V + \alpha_G] \le C[m\varepsilon^{-5}(1+\varepsilon)^{4n} + \alpha_G Vm\varepsilon^{-2}(1+\varepsilon)^{2n}],$$

since $t_{\text{mix}} \le \varepsilon^{-1}$, as required.

To prove (2) we proceed in a very similar fashion. If $\{u \overset{n}{\leftrightarrow} x\} \circ \{u' \overset{n}{\leftrightarrow} y\}$ and $x \overset{2n}{\leftrightarrow} y$ then there exist vertices z_1, z_2 and $t_1, t_2 \le n$ such that the event

$$\{x \overset{=t_1}{\longleftrightarrow} z_1,\, y \overset{=t_2}{\longleftrightarrow} z_2,\, d_{G_p}(x,y) \ge t_1+t_2\} \circ \{z_1 \overset{2n-t_1-t_2}{\longleftrightarrow} z_2\} \circ \{u \overset{n-t_1}{\longleftrightarrow} z_1\} \circ \{u' \overset{n-t_2}{\longleftrightarrow} z_2\}$$

or the event

$$\{x \overset{=t_1}{\longleftrightarrow} z_2,\, y \overset{=t_2}{\longleftrightarrow} z_1,\, d_{G_p}(x,y) \ge t_1+t_2\} \circ \{z_1 \overset{2n-t_1-t_2}{\longleftrightarrow} z_2\} \circ \{u \overset{n-t_1}{\longleftrightarrow} z_1\} \circ \{u' \overset{n-t_2}{\longleftrightarrow} z_2\}$$

occurs, by the same reasoning as before (see Figure 4(b)). Let us handle the first case only (the second leads to an identical contribution). We appeal to the BK-Reimer inequality and as before we condition on $B_x(t_1)$ and bound

$$\mathbf{P}\big(x \overset{=t_1}{\longleftrightarrow} z_1,\ y \overset{=t_2}{\longleftrightarrow} z_2,\ d_{G_p}(x,y) \ge t_1+t_2\big) \le \sum_{A:\, x \overset{=t_1}{\leftrightarrow} z_1} \mathbf{P}(B_x(t_1)=A)\mathbf{P}(y \overset{=t_2}{\longleftrightarrow} z_2 \text{ off } A).$$

We sum over y then x using Lemma 4.4, which gives a bound of

$$C \sum_{z_1,z_2,(u,u'),t_1,t_2\le n} (1+\varepsilon)^{t_1+t_2}\mathbf{P}(u \overset{n-t_1}{\longleftrightarrow} z_1)\mathbf{P}(z_1 \overset{2n-t_1-t_2}{\longleftrightarrow} z_2)\mathbf{P}(u' \overset{n-t_2}{\longleftrightarrow} z_2).$$

An appeal to Corollary 3.18 and Lemma 4.4 to sum over z_1, z_2 gives a bound of

$$CVm \sum_{t_1,t_2\le n} (1+\varepsilon)^{t_1+t_2}[\varepsilon^{-3}(1+\varepsilon)^{4n-2(t_1+t_2)}/V + \alpha_G] \le C[m\varepsilon^{-5}(1+\varepsilon)^{4n} + \alpha_G Vm\varepsilon^{-2}(1+\varepsilon)^{2n}],$$

where the last inequality is a direct calculation. □

Proof of Lemma 6.1. For convenience set $n = 2r + r_0$. By Markov's inequality, the expectation we need to bound is at most

$$2\beta^{-1/2}Vm^{-1}\varepsilon^2(\mathbb{E}|B(r_0)|)^{-2} \times \sum_{x,y,(u,u')} \mathbf{P}\big(\{x \overset{n}{\leftrightarrow} u\} \circ \{y \overset{n}{\leftrightarrow} u'\},\ |B_u(n)| \ge \mathrm{e}^{20M}\varepsilon^{-1}\mathbb{E}|B(r_0)|\big). \qquad (6.4)$$

We split the sum into

$$S_1 = \sum_{x,y,(u,u')} \mathbf{P}\big(\{x \overset{n}{\leftrightarrow} u\} \circ \{y \overset{n}{\leftrightarrow} u'\},\ |B_u(n)| \ge \mathrm{e}^{20M}\varepsilon^{-1}\mathbb{E}|B(r_0)| \text{ and } B_u(n)\cap B_{u'}(n) = \emptyset\big)$$

and

$$S_2 = \sum_{x,y,(u,u')} \mathbf{P}(\{x \overset{n}{\leftrightarrow} u\} \circ \{y \overset{n}{\leftrightarrow} u'\},\ u \overset{2n}{\leftrightarrow} u').$$

We bound S_1 using the BK inequality,

$$S_1 \le \sum_{x,y,(u,u')} \mathbf{P}\big(x \overset{n}{\leftrightarrow} u,\ |B_u(n)| \ge \mathrm{e}^{20M}\varepsilon^{-1}\mathbb{E}|B(r_0)|\big)\mathbf{P}(y \overset{n}{\leftrightarrow} u').$$

Summing over y and x and then over (u,u') shows that this is at most

$$Vm\mathbb{E}|B(n)| \cdot \mathbb{E}|B(n)|\mathbf{1}_{\{|B(n)|\ge \mathrm{e}^{20M}\varepsilon^{-1}\mathbb{E}|B(r_0)|\}}|.$$

We use the Cauchy–Schwarz inequality to bound

$$\mathbb{E}|B(n)|\mathbf{1}_{\{|B(n)|\ge \mathrm{e}^{20M}\varepsilon^{-1}\mathbb{E}|B(r_0)|\}}| \le \big[\mathbb{E}|B(n)|^2\mathbf{P}(|B(n)| \ge \mathrm{e}^{20M}\varepsilon^{-1}\mathbb{E}|B(r_0)|)\big]^{1/2}.$$

We bound this using Lemma 4.7 and the Markov inequality by:

$$\mathbb{E}|B(n)|\mathbf{1}_{\{|B(n)|\geq e^{20M}\varepsilon^{-1}\mathbb{E}|B(r_0)|\}}| \leq Ce^{-10M}(\mathbb{E}|B(n)|)^{3/2}(\mathbb{E}|B(r_0)|)^{-1/2},$$

and conclude that

$$S_1 \leq Ce^{-10M}Vm(\mathbb{E}|B(n)|)^{5/2}(\mathbb{E}|B(r_0)|)^{-1/2}.$$

We bound S_2 using Lemma 6.4(1) :

$$S_2 \leq C[m\varepsilon^{-5}(1+\varepsilon)^{4n} + \alpha_G Vm\varepsilon^{-2}(1+\varepsilon)^{2n}].$$

We put these two back into (6.4) and infer that we can bound this sum by

$$\frac{CV^2\varepsilon^2(\mathbb{E}|B(n)|)^{5/2}}{\beta^{1/2}e^{10M}(\mathbb{E}|B(r_0)|)^{5/2}} + \frac{CV\varepsilon^{-3}(1+\varepsilon)^{4n}}{\beta^{1/2}(\mathbb{E}|B(r_0)|)^2} + \frac{Cm\alpha_G V^2(1+\varepsilon)^{2n}}{\beta^{1/2}m(\mathbb{E}|B(r_0)|)^2} = o(\varepsilon^2V^2),$$

by our choice of r_0 in (2.8), $n = r_0 + 2r$, $r = M/\varepsilon$, $\beta = (\log M)^{-2}$ and using Corollary 4.6. □

6.2. Proof of Theorem 6.2: Finding good radii. Recall the choice of parameters in (6.1).

Lemma 6.5. *For any radius $r \geq \varepsilon^{-1}$ and any $\zeta > 0$,*

$$\mathbf{P}\big(|\partial B(r)| > 0 \text{ and } \exists j \in [\varepsilon^{-1}, r-\varepsilon^{-1}] \text{ with } |\partial B(j)| \leq \zeta\varepsilon^{-1}\big) \leq O(\zeta\varepsilon).$$

Proof. Assume that the event holds, and let J be the first radius j with $j \in [\varepsilon^{-1}, r-\varepsilon^{-1}]$ which has $|\partial B(j)| \leq \zeta\varepsilon^{-1}$. Conditioned on J and $B(J)$, for $|\partial B(r)| > 0$ to occur, one of the vertices on the boundary of $B(J)$ needs to reach level r. Since $r - j \geq \varepsilon^{-1}$, Corollary 3.5 and the union bound show that the probability of this is at most $C\zeta$. As the probability of $|\partial B(j)| > 0$ is at most $C\varepsilon$ by Corollary 3.5, this concludes the proof. □

In the lemma below, we write $\mathbf{P}_A(\cdot) = \mathbf{P}(\cdot \text{ off } A \mid B_x(j_x) = A)$ and let $\mathbb{E}_A$ be the corresponding expectation.

Lemma 6.6. *There exists $c > 0$ such that for any radius $j_x \in [r, 2r]$ the following statement holds. Let the set A be such that x is $(\delta, \beta, j_x, k\varepsilon^{-1})$-fit and $|\partial B_x(j_x)| \geq \zeta\varepsilon^{-1}$ when $B_x(j_x) = A$. Then*

$$\mathbf{P}_A\big(|\partial B_x(j_x + k\varepsilon^{-1}/2)| \geq \varepsilon^{-1}e^{k/4}\big) \geq c\zeta.$$

Proof. We perform a second moment argument on $|B_x([j_x + k\varepsilon^{-1}/2, j_x + k\varepsilon^{-1}]|$ rather than on the required random variable. Since x is $(\delta, \beta, j_x, k\varepsilon^{-1})$-fit,

$$\mathbb{E}_A|B_x([j_x, j_x + k\varepsilon^{-1}])| \geq (|\partial A| - \delta\varepsilon^{-1})(1-\beta)\mathbb{E}|B(k\varepsilon^{-1})|.$$

Furthermore,

$$\mathbb{E}_A|B_x([j_x, j_x + k\varepsilon^{-1}/2])| \leq |\partial A|\,\mathbb{E}|B(k\varepsilon^{-1}/2)|,$$

by monotonicity. Since $|\partial A| \geq 2\delta\varepsilon^{-1}$ by our choice of ζ and δ, and $\beta = o(1)$ (recall (6.1)), Corollary 4.6 now gives a lower bound on the first moment,

$$\mathbb{E}_A|B_x([j_x + k\varepsilon^{-1}/2, j_x + k\varepsilon^{-1}])| \geq \tfrac{1}{4}|\partial A|\,\mathbb{E}|B(k\varepsilon^{-1})|.$$

To calculate the second moment, if u, v are counted in $|B([j_x, j_x + k\varepsilon^{-1}])|$, then either there exist two vertices a_1, a_2 in ∂A such that

$$\{a_1 \overset{k\varepsilon^{-1}}{\longleftrightarrow} u \text{ off } A\} \circ \{a_2 \overset{k\varepsilon^{-1}}{\longleftrightarrow} v \text{ off } A\},$$

or there exists $a \in \partial A$, a vertex z and $t \leq k\varepsilon^{-1}$ such that

$$\{a \overset{=t}{\longleftrightarrow} z \text{ off } A\} \circ \{z \overset{k\varepsilon^{-1}-t}{\longleftrightarrow} u \text{ off } A\} \circ \{z \overset{k\varepsilon^{-1}-t}{\longleftrightarrow} v \text{ off } A\}.$$

We apply the BK-Reimer inequality and sum over u, v to get

$$\begin{aligned}\mathbb{E}_A|B_x([j_x, j_x + k\varepsilon^{-1}])|^2 &\leq |\partial A|^2(\mathbb{E}|B(k\varepsilon^{-1})|)^2 \\ &\quad + \sum_{a\in\partial A,\, z,\, t\leq k\varepsilon^{-1}} \mathbf{P}_A(a \overset{=t}{\longleftrightarrow} z \text{ off } A)(\mathbb{E}|B(k\varepsilon^{-1} - t)|)^2.\end{aligned}$$

We first sum over z using Lemma 4.4, then appeal again to Corollary 4.6 to get

$$\mathbb{E}_A|B_x([j_x, j_x + k\varepsilon^{-1}])|^2 \leq C(\mathbb{E}|B(k\varepsilon^{-1})|)^2[|\partial A|^2 + |\partial A|\varepsilon^{-1}].$$

By (4.4),

$$\mathbf{P}_A\big(|B_x([j_x + k\varepsilon^{-1}/2, j_x + k\varepsilon^{-1}])| \geq \tfrac{1}{8}|\partial A|\,\mathbb{E}|B(k\varepsilon^{-1})|\big) \geq \frac{c|\partial A|^2}{|\partial A|^2 + |\partial A|\varepsilon^{-1}} \geq c\zeta,$$

where the last inequality holds since $|\partial A| \geq \zeta\varepsilon^{-1}$. By Theorem 4.5, we can write this as

$$\mathbf{P}_A\big(|B_x([j_x + k\varepsilon^{-1}/2, j_x + k\varepsilon^{-1}])| \geq c\zeta\varepsilon^{-2}\mathrm{e}^k\big) \geq c\zeta \tag{6.5}$$

for some constant $c > 0$. Now, if $|B_x([j_x + k\varepsilon^{-1}/2, j_x + k\varepsilon^{-1}])| \geq c\zeta\varepsilon^{-2}\mathrm{e}^k$ and $|\partial B_x(j_x + k\varepsilon^{-1}/2)| \leq \varepsilon^{-1}\mathrm{e}^{k/4}$ occurs, then

$$|\partial B_x(j_x + k\varepsilon^{-1}/2)| \leq \varepsilon^{-1}\mathrm{e}^{k/4} \quad\text{and}\quad \sum_{v\in\partial B_x(j_x+k\varepsilon^{-1}/2)} |B_v(k\varepsilon^{-1}/2; A)| \geq c\zeta\varepsilon^{-2}\mathrm{e}^k$$

both occur. By the Markov inequality and Lemma 4.4, the probability of this event is at most

$$\frac{\varepsilon^{-2}\mathrm{e}^{3k/4}}{c\zeta\varepsilon^{-2}\mathrm{e}^k} = O(\zeta^{-1}\mathrm{e}^{-k/4}) = o(\zeta),$$

by our choice of ζ and k in (6.1). Putting this together with (6.5) yields the assertion of the lemma. □

Lemma 6.7 (Finding good radii). *There exist radii $k_1, \dots, k_\ell$ in $[r, 2r]$ such that*

$$k_{i+1} - k_i \geq k\varepsilon^{-1} \quad \textit{for all } i = 1, \dots, \ell,$$

and

$$\mathbf{P}(x \textit{ is } (\delta, \beta, k_i, k\varepsilon^{-1})\textit{-fit}) = (1 + O(\omega^{1/5}))\mathbf{P}(\partial B_x(k_i) \neq \emptyset),$$
$$\mathbf{P}(x \textit{ is } (1, \beta, k_i, r_0)\textit{-fit}) = (1 + O(\omega^{1/5}))\mathbf{P}(\partial B_x(k_i) \neq \emptyset),$$
$$\mathbf{P}(x \textit{ is } (\delta, \beta, k_i + k\varepsilon^{-1}/2, r_0)\textit{-fit}) = (1 + O(\omega^{1/5}))\mathbf{P}(\partial B_x(k_i + k\varepsilon^{-1}/2) \neq \emptyset).$$

Proof. This is the only place where we use Theorem 5.1. Indeed, say a radius $j \in [r, 2r]$ is *good* if it satisfies the three assertions of the statement with k_i replaced by j. Three appeals to Theorem 5.1 show that at least $(1 - o(1))r$ radii $j \in [r, 2r]$ are good by our choice of δ and β. Now, since $\ell k = o(M)$ and $r = M\varepsilon^{-1}$, it is immediate that there exist ℓ good radii which are $k\varepsilon^{-1}$-separated from each other. □

Lemma 6.8. *For at least $(1 - o(1))V^2$ pairs (x, y) and for any $j_x, j_y \in [r, 2r]$,*

$$\mathbb{E}\big[S_{j_x, j_y, r_0} \mathbf{1}_{\{x \overset{2r_0+4r}{\leftrightarrow} y\}}\big] \leq V^{-1} m(\mathbb{E}|B(r_0)|)^2 \alpha_G^{3/4}.$$

Proof. Lemma 6.4(2) with $n = r_0 + 2r$ and a straightforward calculation with Theorem 4.5 and our choice of parameters shows that

$$\sum_{x,y} \mathbb{E}\big[S_{j_x, j_y, r_0} \mathbf{1}_{\{x \overset{2r_0+j_x+j_y}{\leftrightarrow} y\}}\big] \leq CVm(\mathbb{E}|B(r_0)|)^2 [\alpha_G \mathrm{e}^{8M} + \alpha_G \mathrm{e}^{4M}],$$

which gives the result since $C\alpha_G \mathrm{e}^{8M} \leq \alpha_G^{3/4}$ by our choice of M in (2.7). □

Proof of Theorem 6.2. Recall the requirements (1)–(5) in the definition of $\mathscr{A}(x, y, j_x, j_y, r_0, \beta, k)$ (p. 771). We apply Lemma 6.7 and let $k_1, \dots, k_\ell$ be the corresponding radii. We prove the theorem with radii $\{j_1, \dots, j_\ell\}$ defined by

$$j_i = k_i + k\varepsilon^{-1}/2 \quad \text{for } i = 1, \dots, \ell,$$

and assume x, y are such that the assertion of Lemma 6.8 holds. We will prove that for such x, y,

$$\mathbf{P}\Big(\mathscr{A}(x, y, 2r, 2r) \text{ and } \bigcap_{j_x, j_y \in \{j_1, \dots, j_\ell\}} \{(q) \text{ does not hold for } j_x, j_y\}\Big) = o(\varepsilon^2) \tag{6.6}$$

for $q \in \{1, 2, 3, 4, 5\}$. We do this in the order (1), (2), (4), (5) and (3). Since $\mathscr{A}(x, y, 2r, 2r) \subseteq \mathscr{A}(x, y, j_x, j_y)$ when $j_x, j_y \leq 2r$, (6.6) holds trivially for $q = 1$ and all $x, y, j_x, j_y \leq 2r$.

For any $j_x \in \{j_1, \dots, j_\ell\}$,

$$\mathbf{P}\big(\mathscr{A}(x, y, j_x, j_y) \text{ and } |B_x(j_x)| \geq \varepsilon^{-2}(1+\varepsilon)^{3r}\big) \leq C\varepsilon^2(1+\varepsilon)^{-r} = O(\mathrm{e}^{-M}\varepsilon^2),$$

by the Markov inequality, Lemma 4.4, the BK-Reimer inequality and Corollary 3.5. This implies that

$$\mathbf{P}\big(\mathscr{A}(x, y, 2r, 2r) \text{ and } \exists j_x \in \{j_1, \dots, j_\ell\} \text{ such that } |B_x(j_x)| \geq \varepsilon^{-2}(1+\varepsilon)^{3r}\big) = o(\varepsilon^2),$$

since $\ell = o(\mathrm{e}^M)$. Similarly,

$$\mathbf{P}\big(\mathscr{A}(x, y, j_x, j_y) \text{ and } |\partial B_x(j_x)| \geq \varepsilon^{-1}(1+\varepsilon)^{3r}\big) \leq C\varepsilon^2(1+\varepsilon)^{-r} = O(\mathrm{e}^{-M}\varepsilon^2),$$

leading to the same bound. This proves (6.6) for $q = 2$.

Next, we wish to show that for any $j_x \in \{j_1, \dots, j_\ell\}$,

$$\mathbf{P}\big(\mathscr{A}(x, y, j_x, j_y) \text{ and } x \text{ is not } (1, \beta, j_x, r_0)\text{-fit}\big) = O(\varepsilon^2\omega^{1/5}). \tag{6.7}$$

It is tempting to use the BK-Reimer inequality here; however, we cannot claim that the event that in (6.7) implies that $\partial B_y(j_y) \neq \emptyset$ occurs disjointly from the event that x is not $(1, \beta, j_x, r_0)$-fit, since they are both non-monotone events and the corresponding witnesses may share closed edges. Instead, we condition on $B_x(j_x) = A$ to get

$$\begin{aligned}&\mathbf{P}\big(\mathscr{A}(x, y, j_x, j_y) \text{ and } x \text{ is not } (1, \beta, j_x, r_0)\text{-fit}\big)\\ &\qquad = \sum_{A:\, x \text{ is not } (1,\beta,j_x,r_0)\text{-fit}} \mathbf{P}(B_x(j_x) = A)\mathbf{P}\big(\partial B_y(j_y) \neq \emptyset \text{ off } A\big),\end{aligned}$$

since $(1, \beta, j_x, r_0)$-fit is determined by the status of the edges touching $B_x(j_x)$. We use Corollary 3.5 to bound $\mathbf{P}(\partial B_y(j_y) \neq \emptyset \text{ off } A) = O(\varepsilon)$ and

$$\begin{aligned}&\mathbf{P}\big(\partial B_x(j_x) \neq \emptyset \text{ and } x \text{ is not } (1, \beta, j_x, r_0)\text{-fit}\big)\\ &\qquad = \mathbf{P}(\partial B_x(j_x) \neq \emptyset) - \mathbf{P}\big(\partial B_x(j_x) \neq \emptyset \text{ and } x \text{ is } (1, \beta, j_x, r_0)\text{-fit}\big)\\ &\qquad \leq \mathbf{P}(\partial B_x(j_x) \neq \emptyset)O(\omega^{1/5}),\end{aligned}$$

by our choice of radii in Lemma 6.7, so Corollary 3.5 gives (6.7). Therefore,

$$\begin{aligned}&\mathbf{P}\big(\mathscr{A}(x, y, 2r, 2r) \text{ and } \exists j_x \in \{j_1, \dots, j_\ell\} \text{ such that } x \text{ is not } (1, \beta, j_x, r_0)\text{-fit}\big)\\ &\qquad = O(\ell\varepsilon^2\omega^{1/5}) = o(\varepsilon^2),\end{aligned}$$

by our choice of ℓ in (6.1), of M in (2.7), and of ω in (5.1). This proves (6.6) for $q = 4$.

Similarly, by Lemma 6.8 and Markov's inequality, for any j_x, j_y,

$$\mathbf{P}(\mathscr{A}(x, y, j_x, j_y) \text{ and (5) does not hold}) \leq C\varepsilon^2\alpha_G^{1/4}.$$

The union bound implies that

$$\begin{aligned}&\mathbf{P}\big(\mathscr{A}(x, y, 2r, 2r) \text{ and } \exists j_x, j_y \in \{j_1, \dots, j_\ell\} \text{ such that (5) does not hold}\big)\\ &\qquad = O(\ell^2\alpha_G^{1/4}\varepsilon^2) = o(\varepsilon^2),\end{aligned}$$

by our choice of ℓ. Therefore, (6.6) holds for $q = 5$.

Thus, it remains to prove (6.6) for $q = 3$. This is a difficult requirement; we only prove that one of the radii in $\{j_1, \dots, j_\ell\}$ satisfies it (in fact, all radii do, but that is harder to prove, and we refrain from doing so). As in the proof of (6.6) for $q = 4$, using Corollary 3.5 it is enough to show that

$$\mathbf{P}\big(\partial B_x(2r) \neq \emptyset \text{ and } |\partial B_x(j_x)| \leq \mathrm{e}^{k/4}\varepsilon^{-1}\ \forall j_x \in \{j_1, \dots, j_\ell\}\big) = o(\varepsilon). \tag{6.8}$$

For $i \in \{1, \dots, \ell\}$, we write $\mathscr{A}_i$ for the event that x is $(\delta, \beta, k_i, k\varepsilon^{-1})$-fit; $\mathscr{B}_i$ for the event that $|\partial B_x(j_i)| \le \varepsilon^{-1} \mathrm{e}^{k/4}$; and $\mathscr{D}_i$ for the event

$$\mathscr{D}_i = \big\{ |\partial B_x(k_t)| \ge \zeta \varepsilon^{-1} \ \forall t \in \{1, \dots, i\} \big\},$$

so that

$$\mathbf{P}(\partial B_x(2r) \ne \emptyset \text{ and } |\partial B_x(j_x)|) \le \mathbf{P}(\mathscr{D}_\ell \cap \mathscr{B}_1 \cap \dots \cap \mathscr{B}_\ell) + \mathbf{P}(\{\partial B_x(2r) \ne \emptyset\} \cap \mathscr{D}_\ell^c).$$

Then we can split

$$\mathbf{P}(\mathscr{D}_\ell \cap \mathscr{B}_1 \cap \dots \cap \mathscr{B}_\ell) \le \mathbf{P}(\mathscr{D}_\ell \cap \mathscr{A}_\ell^c) + \mathbf{P}(\mathscr{D}_\ell \cap \mathscr{B}_1 \cap \dots \cap \mathscr{B}_\ell \cap \mathscr{A}_\ell).$$

By our choice of k_i in Lemma 6.7 and Corollary 3.5 we have $\mathbf{P}(\mathscr{D}_\ell \cap \mathscr{A}_\ell^c) \le \varepsilon \omega^{1/5}$, so

$$\begin{aligned}
&\mathbf{P}(\mathscr{D}_\ell \cap \mathscr{B}_1 \cap \dots \cap \mathscr{B}_\ell) \\
&\quad \le \varepsilon \omega^{1/5} + \mathbf{P}(\mathscr{B}_\ell \mid \mathscr{D}_\ell \cap \mathscr{B}_1 \cap \dots \cap \mathscr{B}_{\ell-1} \cap \mathscr{A}_\ell) \mathbf{P}(\mathscr{D}_{\ell-1} \cap \mathscr{B}_1 \cap \dots \cap \mathscr{B}_{\ell-1}).
\end{aligned}$$

Thus, by Lemma 6.6,

$$\mathbf{P}(\mathscr{D}_\ell \cap \mathscr{B}_1 \cap \dots \cap \mathscr{B}_\ell) \le \varepsilon \omega^{1/5} + (1 - c\zeta) \mathbf{P}(\mathscr{D}_{\ell-1} \cap \mathscr{B}_1 \cap \dots \cap \mathscr{B}_{\ell-1}).$$

By iterating we obtain

$$\mathbf{P}(\mathscr{D}_\ell \cap \mathscr{B}_1 \cap \dots \cap \mathscr{B}_\ell) \le \varepsilon \ell \omega^{1/5} + C\varepsilon (1 - c\zeta)^\ell = o(\varepsilon),$$

since $\ell \omega^{1/5} = o(1)$ and $\zeta^{-1} = o(\ell)$ (recall (6.1)), and $\mathbf{P}(\mathscr{D}_1) \le C\varepsilon$ by Corollary 3.5. Lastly, Lemma 6.5 shows that $\mathbf{P}(\{\partial B_x(2r) \ne \emptyset\} \cap \mathscr{D}_\ell^c) = o(\varepsilon)$, proving (6.8) and thus concluding the proof of (6.6) for $q = 3$ and the proof of Theorem 6.2. □

6.3. Proof of Theorem 6.3: Conditional second moment. We now set the stage for the proof of Theorem 6.3. We perform this by a conditional second moment argument on $S_{j_x, j_y, r_0}(x, y)$. We will be conditioning on $B_x(j_x) = A$ and $B_y(j_y) = B$ where A and B are such that the event $\mathscr{A}(x, y, j_x, j_y, r_0, \beta, k)$ holds. We abuse notation, as before, and treat A, B as sets of vertices but our conditioning is on the status of all edges touching $B_x(j_x - 1)$ and $B_y(j_y - 1)$. Thus, while A and B are disjoint sets of vertices, they may be sharing closed edges. With this in mind, we generalize the notation just before Lemma 6.6, and write $\mathbf{P}_A$, $\mathbf{P}_B$ and $\mathbf{P}_{A,B}$ for the measures

$$\begin{aligned}
\mathbf{P}_A(\cdot) &= \mathbf{P}(\cdot \text{ off } A \mid B_x(j_x) = A), \\
\mathbf{P}_B(\cdot) &= \mathbf{P}(\cdot \text{ off } B \mid B_y(j_y) = B), \\
\mathbf{P}_{A,B}(\cdot) &= \mathbf{P}(\cdot \text{ off } A \cup B \mid B_x(j_x) = A,\ B_y(j_y) = B).
\end{aligned}$$

We start by proving five preparatory lemmas.

Lemma 6.9. *Assume that A and B are such that x, y are $(1, \beta, j_x, r_0)$-fit and $(1, \beta, j_y, r_0)$-fit, respectively. Then*

$$\sum_{a,b\in\partial A\times\partial B}\sum_{(u,u')}\mathbf{P}_A(a \xleftrightarrow{P[2t_{\mathrm{mix}},r_0],x} u)\mathbf{P}_B(b \xleftrightarrow{P[2t_{\mathrm{mix}},r_0],y} u')$$
$$\geq (1-8\beta^{1/2})V^{-1}(\mathbb{E}|B(r_0)|)^2 m(|\partial A|-\varepsilon^{-1})(|\partial B|-\varepsilon^{-1}).$$

Proof. Let $a \in \partial A$ be a (β, j_x, r_0)-regenerative vertex. Then, by definition,

$$\sum_u \mathbf{P}_A(a \xleftrightarrow{P[2t_{\mathrm{mix}},r_0],x} u) \geq (1-\beta)\mathbb{E}|B(r_0)|.$$

Denote

$$U = \big\{u : \mathbf{P}_A(a \xleftrightarrow{P[2t_{\mathrm{mix}},r_0],x} u) \leq (1-\beta^{1/2})V^{-1}\mathbb{E}|B(r_0)|\big\},$$

and recall that Lemma 3.15 guarantees that

$$\mathbf{P}_A(a \xleftrightarrow{P[2t_{\mathrm{mix}},r_0],x} u) \leq (1+o(\beta))\mathbb{E}|B(r_0)|/V,$$

by our choice of β in (6.1), so that

$$(1-\beta)\mathbb{E}|B(r_0)| \leq \sum_u \mathbf{P}_A(a \xleftrightarrow{P[2t_{\mathrm{mix}},r_0],x} u)$$
$$\leq |U|(1-\beta^{1/2})V^{-1}\mathbb{E}|B(r_0)| + (V-|U|)(1+o(\beta))V^{-1}\mathbb{E}|B(r_0)|,$$

and we deduce that $|U| \leq 2\beta^{1/2}V$. In other words, for at least $(1-2\beta^{1/2})V$ vertices u,

$$\mathbf{P}_A(a \xleftrightarrow{P[2t_{\mathrm{mix}},r_0],x} u) \geq (1-\beta^{1/2})\mathbb{E}|B(r_0)|/V.$$

Similarly, for any $b \in \partial B$ which is (β, j_y, r_0)-regenerative, there are at least $(1-2\beta^{1/2})V$ vertices u such that

$$\mathbf{P}_B(b \xleftrightarrow{P[2t_{\mathrm{mix}},r_0],y} u) \geq (1-\beta^{1/2})\mathbb{E}|B(r_0)|/V.$$

Thus, for such a and b, at least $(1-4\beta^{1/2})V$ vertices u satisfy both inequalities. Write D for this set of vertices, so that $|D^c| \leq 4\beta^{1/2}V$. Since the degree of each vertex is m, the number of edges having at least one side in D^c is at most $4\beta^{1/2}Vm$. Thus, at least $(1-8\beta^{1/2})Vm$ directed edges (u,u') are such that u and u' both satisfy the above inequalities. Hence

$$\sum_{(u,u')}\mathbf{P}_A(a \xleftrightarrow{P[2t_{\mathrm{mix}},r_0],x} u)\mathbf{P}_B(b \xleftrightarrow{P[2t_{\mathrm{mix}},r_0],y} u') \geq (1-8\beta^{1/2})(\mathbb{E}|B(r_0)|)^2 m/V.$$

Since x is $(1, \beta, j_x, r_0)$-fit and y is $(1, \beta, j_y, r_0)$-fit the number of such pairs a, b is at least $(|\partial A|-\varepsilon^{-1})(|\partial B|-\varepsilon^{-1})$, and the lemma follows. □

Lemma 6.10. *The following bounds hold:*

$$\sum_{\substack{(a,b)\in\partial A\times\partial B\\ (u,u'),\, z_1,\, t_1\in[t_{\rm mix},r_0]}} \mathbf{P}_{A,B}(a \xleftrightarrow{P[2t_{\rm mix},r_0]} u,\, a \xleftrightarrow{r_0} z_1)\mathbf{P}_B(b \xleftrightarrow{=t_1} z_1)\mathbf{P}_B(z_1 \xleftrightarrow{r_0-t_1} u')$$
$$\le \alpha_G^{1/2}|\partial A|\,|\partial B|V^{-1}m(\mathbb{E}|B(r_0)|)^2,$$

$$\sum_{\substack{(a,b)\in\partial A\times\partial B\\ (u,u'),\, z_1,\, t_1\in[t_{\rm mix},r_0]}} \mathbf{P}_{A,B}(a \xleftrightarrow{P[2t_{\rm mix},r_0]} u,\, a \xleftrightarrow{r_0} z_1)\mathbf{P}_B(b \xleftrightarrow{r_0-t_1} z_1)\mathbf{P}_B(z_1 \xleftrightarrow{=t_1} u')$$
$$\le \alpha_G^{1/2}|\partial A|\,|\partial B|V^{-1}m(\mathbb{E}|B(r_0)|)^2.$$

Proof. The proof of the second assertion is identical to the first, so we only prove the first. If $a \xleftrightarrow{P[2t_{\rm mix},r_0]} u$ and $a \xleftrightarrow{r_0} z_1$, then either there exist z_2 and $t_2\in[t_{\rm mix},r_0]$ such that

$$\{a \xleftrightarrow{=t_2} z_2\}\circ\{z_2 \xleftrightarrow{r_0-t_2} u\}\circ\{z_2 \xleftrightarrow{r_0-t_2} z_1\},$$

or there exists z_2 such that

$$\{a \xleftrightarrow{t_{\rm mix}} z_2\}\circ\{z_2 \xleftrightarrow{P[t_{\rm mix},r_0]} u\}\circ\{z_2 \xleftrightarrow{P[t_{\rm mix},r_0]} z_1\},$$

or there exists z_2 such that

$$\{a \xleftrightarrow{P[t_{\rm mix},r_0]} z_2\}\circ\{z_2 \xleftrightarrow{t_{\rm mix}} u\}\circ\{z_2 \xleftrightarrow{P[t_{\rm mix},r_0]} z_1\}.$$

To see this, let η be the lexicographically first shortest open path between a and z_1 with $|\eta|\le r_0$, and let γ be an open path between a and u such that $|\gamma|\in[2t_{\rm mix},r_0]$. Let z_2 be the last vertex (according to the ordering induced by γ) on γ belonging to η (that is, the part of γ after z_2 is disjoint from η, and the part of η after z_2 is disjoint from γ). Let t_2 be the distance between a and z_2 along η. If $t_2\ge t_{\rm mix}$, then the first event occurs: the first witness is the first t_2 open edges of η together with all the closed edges in the graph, the second witness is the set of open edges of γ between z_2 to u (note that there are no more than r_0-t_2 edges since the part of γ between a to z_2 is of length at least t_2), and the third witness is the set of open edges of η between z_2 and u. If $t_2\le t_{\rm mix}$ and the part of γ between z_2 and u is longer than $t_{\rm mix}$, then the second event occurs by similar reasoning. Finally, if $t_2\le t_{\rm mix}$ occurs and the part of γ between z_2 and u is of length at most $t_{\rm mix}$, then the part of γ between a and z_2 is longer than $t_{\rm mix}$ and the third event occurs.

This leads to three different terms; the BK-Reimer inequality shows that the required sum is at most $S^{(a)}+S^{(b)}+S^{(c)}$, where

$$S^{(a)} = \sum_{\substack{(a,b)\in\partial A\times\partial B,\,(u,u'),\\ z_1,z_2,\,t_1,t_2\in[t_{\rm mix},r_0]}} \mathbf{P}_{A,B}(a \xleftrightarrow{=t_2} z_2)\mathbf{P}_{A,B}(z_2 \xleftrightarrow{r_0-t_2} u)\mathbf{P}_{A,B}(z_2 \xleftrightarrow{r_0-t_2} z_1)$$
$$\times\,\mathbf{P}_B(b \xleftrightarrow{=t_1} z_1)\mathbf{P}_B(z_1 \xleftrightarrow{r_0-t_1} u'),$$

$$S^{(b)} = \sum_{\substack{(a,b)\in\partial A\times\partial B,\,(u,u'),\\ z_1,z_2,\,t_1\in[t_{\rm mix},r_0]}} \mathbf{P}_{A,B}(a \xleftrightarrow{t_{\rm mix}} z_2)\mathbf{P}_{A,B}(z_2 \xleftrightarrow{P[t_{\rm mix},r_0]} u)\mathbf{P}_{A,B}(z_2 \xleftrightarrow{P[t_{\rm mix},r_0]} z_1)$$
$$\times\,\mathbf{P}_B(b \xleftrightarrow{=t_1} z_1)\mathbf{P}_B(z_1 \xleftrightarrow{r_0-t_1} u'),$$

$$S^{(c)} = \sum_{\substack{(a,b)\in\partial A\times\partial B,\,(u,u'),\\ z_1,z_2,\,t_1\in[t_{\rm mix},r_0]}} \mathbf{P}_{A,B}(a \xleftrightarrow{P[t_{\rm mix},r_0]} z_2)\mathbf{P}_{A,B}(z_2 \xleftrightarrow{t_{\rm mix}} u)\mathbf{P}_{A,B}(z_2 \xleftrightarrow{P[t_{\rm mix},r_0]} z_1) \times \mathbf{P}_B(b \xleftrightarrow{=t_1} z_1)\mathbf{P}_B(z_1 \xleftrightarrow{r_0-t_1} u').$$

We use Lemma 3.15 together with Lemma 4.4 to bound the factors $\mathbf{P}_{A,B}(a \xleftrightarrow{=t_2} z_2)$ and $\mathbf{P}_B(b \xleftrightarrow{=t_1} z_1)$ in $S^{(a)}$ by $CV^{-1}(1+\varepsilon)^{t_2}$ and $CV^{-1}(1+\varepsilon)^{t_1}$, respectively. This gives

$$S^{(a)} \le CV^{-2} \sum_{\substack{(a,b)\in\partial A\times\partial B,\,(u,u'),\\ t_1,t_2\in[t_{\rm mix},r_0]}} (1+\varepsilon)^{t_1+t_2} \times \sum_{z_1,z_2} \mathbf{P}_{A,B}(z_2 \xleftrightarrow{r_0-t_2} u)\mathbf{P}_{A,B}(z_2 \xleftrightarrow{r_0-t_2} z_1)\mathbf{P}_{A,B}(z_1 \xleftrightarrow{r_0-t_1} u').$$

We sum over z_1, z_2 using Corollary 3.18 together with Lemma 4.4 to get

$$\begin{aligned} S^{(a)} &\le CV^{-2} \sum_{\substack{(a,b)\in\partial A\times\partial B,\,(u,u'),\\ t_1,t_2\in[t_{\rm mix},r_0]}} (1+\varepsilon)^{t_1+t_2}[\alpha_G + \varepsilon^{-3}(1+\varepsilon)^{3r_0-t_1-2t_2}/V] \\ &\le C\alpha_G m|\partial A|\,|\partial B|V^{-1}\varepsilon^{-2}(1+\varepsilon)^{2r_0} + C|\partial A|\,|\partial B|V^{-2}m\varepsilon^{-4}(1+\varepsilon)^{3r_0}r_0, \end{aligned}$$

where the last inequality is an immediate calculation. By Theorem 4.5 the first term is at most

$$C\alpha_G|\partial A|\,|\partial B|V^{-1}m(\mathbb{E}|B(r_0)|)^2,$$

and by our choice of r_0 in (2.8), the second term is at most

$$\frac{\alpha_G^{1/2}\log(\varepsilon^3 V)}{\sqrt{\varepsilon^3 V}}|\partial A|\,|\partial B|V^{-1}m(\mathbb{E}|B(r_0)|)^2.$$

This gives an upper bound on $S^{(a)}$ fitting the error in the assertion of the lemma.

To estimate $S^{(b)}$ we use Lemma 3.13 to bound $\mathbf{P}_{A,B}(z_2 \xleftrightarrow{P[t_{\rm mix},r_0]} u)$ and $\mathbf{P}_{A,B}(z_2 \xleftrightarrow{P[t_{\rm mix},r_0]} z_1)$. This gives

$$S^{(b)} \le C\frac{(\mathbb{E}|B(r_0)|)^2}{V^2} \sum_{\substack{(a,b)\in\partial A\times\partial B,\,(u,u'),\\ z_1,z_2,\,t_1\in[t_{\rm mix},r_0]}} \mathbf{P}_{A,B}(a \xleftrightarrow{t_{\rm mix}} z_2)\mathbf{P}_B(b \xleftrightarrow{=t_1} z_1)\mathbf{P}_B(z_1 \xleftrightarrow{r_0-t_1} u').$$

We now sum over (u,u'), z_1, z_2 and t_1 using Lemma 4.4 to get

$$\begin{aligned} S^{(b)} &\le C|\partial A|\,|\partial B|V^{-2}m(\mathbb{E}|B(r_0)|)^2 r_0 t_{\rm mix}\varepsilon^{-1}(1+\varepsilon)^{r_0} \\ &\le \frac{Ct_{\rm mix}\varepsilon\alpha_G^{1/2}}{\sqrt{\varepsilon^3 V}}C|\partial A|\,|\partial B|V^{-1}m(\mathbb{E}|B(r_0)|)^2, \end{aligned}$$

by Theorem 4.5 and (2.8). Since $t_{\rm mix} = o(\varepsilon^{-1})$, this fits within the error estimate in the assertion of the lemma. The estimate on $S^{(c)}$ is performed in precisely the same way as for $S^{(b)}$, and gives the same error estimate, concluding the proof. □

Lemma 6.11. *The following bounds hold:*

$$\sum_{\substack{(a,b)\in\partial A\times\partial B,\, a'\in\partial A\\ (u,u'),\, z_1,\, t_1\in[t_{\mathrm{mix}},r_0]}} \mathbf{P}_{A,B}(a \xleftrightarrow{P[2t_{\mathrm{mix}},r_0]} u)\mathbf{P}_{A,B}(a' \overset{r_0}{\leftrightarrow} z_1)\mathbf{P}_B(b \overset{=t_1}{\longleftrightarrow} z_1)\mathbf{P}_B(z_1 \xleftrightarrow{r_0-t_1} u') \leq \alpha_G^{1/2}\varepsilon|\partial A|^2|\partial B|V^{-1}m(\mathbb{E}|B(r_0)|)^2$$

and

$$\sum_{\substack{(a,b)\in\partial A\times\partial B,\, a'\in\partial A\\ (u,u'),\, z_1,\, t_1\in[t_{\mathrm{mix}},r_0]}} \mathbf{P}_{A,B}(a \xleftrightarrow{P[2t_{\mathrm{mix}},r_0]} u)\mathbf{P}_{A,B}(a' \overset{r_0}{\leftrightarrow} z_1)\mathbf{P}_B(b \xleftrightarrow{r_0-t_1} z_1)\mathbf{P}_B(z_1 \overset{=t_1}{\longleftrightarrow} u') \leq \alpha_G^{1/2}\varepsilon|\partial A|^2|\partial B|V^{-1}m(\mathbb{E}|B(r_0)|)^2.$$

Proof. The proof of the second assertion is identical to the first, so we only prove the first. We use Lemma 3.15 to bound $\mathbf{P}_{A,B}(a \xleftrightarrow{P[2t_{\mathrm{mix}},r_0]} u) \leq CV^{-1}\mathbb{E}|B(r_0)|$ and $\mathbf{P}_B(b \overset{=t_1}{\longleftrightarrow} z_1) \leq CV^{-1}(1+\varepsilon)^{t_1}$. We then sum $\mathbf{P}_B(z_1 \xleftrightarrow{r_0-t_1} u')$ over (u,u') to obtain a factor of $Cm\varepsilon^{-1}(1+\varepsilon)^{r_0-t_1}$ by Lemma 4.4. We now sum $\mathbf{P}_{A,B}(a' \overset{r_0}{\leftrightarrow} z_1)$ over z_1 and get another $\mathbb{E}|B(r_0)|$ factor and now sum all this over a, b, a', t_1. This gives a contribution of

$$C|\partial A|^2|\partial B|V^{-2}mr_0(\mathbb{E}|B(r_0)|)^3,$$

by Theorem 4.5. This is at most

$$\frac{C\varepsilon|\partial A|\alpha_G^{1/2}\log(\varepsilon^3 V)}{\sqrt{\varepsilon^3 V}}|\partial A|\,|\partial B|V^{-1}m(\mathbb{E}|B(r_0)|)^2,$$

by our choice of r_0 in (2.8) and Lemma 4.4, concluding our proof. □

Lemma 6.12. *The following bounds hold:*

$$\sum_{\substack{a,b\in\partial A\times\partial B\\ (u,u'),\, z_1\in B,\, t_1\in[t_{\mathrm{mix}},r_0]}} \mathbf{P}_A(a \overset{=t_1}{\longleftrightarrow} z_1)\mathbf{P}_A(z_1 \xleftrightarrow{r_0-t_1} u)\mathbf{P}_B(b \xleftrightarrow{P[2t_{\mathrm{mix}},r_0]} u') \leq C|\partial A|\,|\partial B|\,|B|V^{-2}mr_0(\mathbb{E}|B(r_0)|)^2$$

and

$$\sum_{\substack{a,b\in\partial A\times\partial B\\ (u,u'),\, z_1\in B,\, t_1\in[t_{\mathrm{mix}},r_0]}} \mathbf{P}_A(a \xleftrightarrow{r_0-t_1} z_1)\mathbf{P}_A(z_1 \overset{=t_1}{\longleftrightarrow} u)\mathbf{P}_B(b \xleftrightarrow{P[2t_{\mathrm{mix}},r_0]} u') \leq C|\partial A|\,|\partial B|\,|B|V^{-2}mr_0(\mathbb{E}|B(r_0)|)^2.$$

Proof. The proof of the second assertion is identical to the first so we only prove the first. We use Lemma 3.13 to bound

$$\mathbf{P}_B(b \xleftrightarrow{P[2t_{\mathrm{mix}},r_0]} u') \leq (1+o(1))V^{-1}\mathbb{E}|B(r_0)|,$$

and, as before, we use Lemma 3.12 together with Lemma 4.4 to bound

$$\mathbf{P}_A(a \overset{=t_1}{\longleftrightarrow} z_1) \leq CV^{-1}(1+\varepsilon)^{t_1},$$

sum over u' such that $(u,u') \in E(G)$, and finally use Lemma 4.4 to bound

$$\sum_u \mathbf{P}_A(z_1 \xleftrightarrow{r_0-t_1} u) \leq C\varepsilon^{-1}(1+\varepsilon)^{r_0-t_1}.$$

Altogether, after summing over $a \in \partial A$, $b \in \partial B$, $z_1 \in B$, $t_1 \le r_0$, this gives the bound of

$$C|\partial A|\,|\partial B|\,|B|V^{-2}mr_0\varepsilon^{-1}(1+\varepsilon)^{r_0}\mathbb{E}|B(r_0)| \le C|\partial A|\,|\partial B|\,|B|V^{-2}mr_0(\mathbb{E}|B(r_0)|)^2,$$

where we have used Theorem 4.5. □

Lemma 6.13. *For any $\delta, \beta > 0$,*

$$\mathbb{E}_{A,B}S_{j_x,j_y,r_0}(x,y) \ge (1-8\beta^{1/2})V^{-1}m(\mathbb{E}|B(r_0)|)^2(|\partial A|-\varepsilon^{-1})(|\partial B|-\varepsilon^{-1}) - \mathrm{Err},$$

where

$$\mathrm{Err} \le C|\partial A|\,|\partial B|V^{-1}m(\mathbb{E}|B(r_0)|)^2[r_0(|A|+|B|)V^{-1} + \alpha_G^{1/2}(1+\varepsilon|\partial A|)].$$

Proof. We have

$$\begin{aligned}&\mathbb{E}_{A,B}S_{j_x,j_y,r_0}(x,y)\\ &\quad= \sum_{(a,b)\in\partial A\times\partial B}\sum_{(u,u')}\mathbf{P}_{A,B}\big(\{a \xleftrightarrow{P[2t_{\mathrm{mix}},r_0],x} u\} \text{ and } \{b \xleftrightarrow{P[2t_{\mathrm{mix}},r_0],y} u' \text{ off } B_x(j_x+r_0)\}\big),\end{aligned} \tag{6.9}$$

because the additional requirement that a and b are pivotal in the definitions of $\{a \xleftrightarrow{P[2t_{\mathrm{mix}},r_0],x} u\}$ and $\{b \xleftrightarrow{P[2t_{\mathrm{mix}},r_0],y} u'\}$ implies that they are unique in $\partial A \times \partial B$, so no pair (a,b) is overcounted in the sum. We define $B_{\partial A}(r_0; A\cup B) = \bigcup_{a'\in\partial A} B_{a'}(r_0; A\cup B)$. We condition on $B_{\partial A}(r_0; A\cup B) = H$ for an admissible H (that is, any H that has positive probability and $a \xleftrightarrow{P[2t_{\mathrm{mix}},r_0],x} u$ off B occurs in it). Each summand in (6.9) equals

$$\sum_H \mathbf{P}_{A,B}\big(B_{\partial A}(r_0; A\cup B) = H\big)\mathbf{P}_{A,B}\big(b \xleftrightarrow{P[2t_{\mathrm{mix}},r_0],y} u' \text{ off } H \;\big|\; B_{\partial A}(r_0; A\cup B) = H\big),$$

and we have

$$\mathbf{P}_{A,B}\big(b \xleftrightarrow{P[2t_{\mathrm{mix}},r_0],y} u' \text{ off } H \;\big|\; B_{\partial A}(r_0; A\cup B) = H\big) = \mathbf{P}_B(b \xleftrightarrow{P[2t_{\mathrm{mix}},r_0],y} u' \text{ off } A\cup H),$$

because on both sides the status of the edges touching $A\cup H$ cannot change the occurrence of the event. This gives

$$\begin{aligned}&\mathbb{E}_{A,B}S_{j_x,j_y,r_0}(x,y)\\ &\quad= \sum_{(a,b)\in\partial A\times\partial B}\sum_{(u,u')}\sum_H \mathbf{P}_{A,B}(B_{\partial A}(r_0; A\cup B) = H)\mathbf{P}_B(b \xleftrightarrow{P[2t_{\mathrm{mix}},r_0],y} u' \text{ off } A\cup H).\end{aligned}$$

Now, by Claim 3.1,

$$\begin{aligned}&\mathbf{P}_B(b \xleftrightarrow{P[2t_{\mathrm{mix}},r_0],y} u' \text{ off } A\cup H)\\ &\qquad\qquad \ge \mathbf{P}_B(b \xleftrightarrow{P[2t_{\mathrm{mix}},r_0],y} u') - \mathbf{P}_B(b \xleftrightarrow{P[2t_{\mathrm{mix}},r_0]} u' \text{ only on } A\cup H),\end{aligned}$$

where in the last term we have dropped the requirement that y is pivotal (which only increases the probability). Hence by summing on H we get

$$\mathbb{E}_{A,B} S_{j_x,j_y,r_0}(x,y) \geq \sum_{(a,b)\in\partial A\times\partial B} \sum_{(u,u')} \mathbf{P}_{A,B}(a \xleftrightarrow{P[2t_{\rm mix},r_0],x} u)\mathbf{P}_B(b \xleftrightarrow{P[2t_{\rm mix},r_0],y} u') - S_2,$$

where

$$S_2 = \sum_{(a,b)\in\partial A\times\partial B} \sum_{(u,u')} \sum_{H} \mathbf{P}_{A,B}(B_{\partial A}(r_0; A\cup B)=H)\mathbf{P}_B(b \xleftrightarrow{P[2t_{\rm mix},r_0]} u' \text{ only on } A\cup H). \tag{6.10}$$

As before, $\mathbf{P}_{A,B}(a \xleftrightarrow{P[2t_{\rm mix},r_0],x} u) = \mathbf{P}_A(a \xleftrightarrow{P[2t_{\rm mix},r_0],x} u \text{ off } B)$, since the status of the edges touching $A\cup B$ in both sides does not matter. Claim 3.1 again gives

$$\mathbf{P}_A(a \xleftrightarrow{P[2t_{\rm mix},r_0],x} u \text{ off } B) \geq \mathbf{P}_A(a \xleftrightarrow{P[2t_{\rm mix},r_0],x} u) - \mathbf{P}_A(a \xleftrightarrow{P[2t_{\rm mix},r_0]} u \text{ only on } B),$$

and so we may further expand $\mathbb{E}_{A,B} S_{j_x,j_y,r_0}(x,y) \geq S_1 - S_2 - S_3$ with

$$S_1 = \sum_{(a,b)\in\partial A\times\partial B} \sum_{(u,u')} \mathbf{P}_A(a \xleftrightarrow{P[2t_{\rm mix},r_0],x} u)\mathbf{P}_B(b \xleftrightarrow{P[2t_{\rm mix},r_0],y} u'),$$

$$S_3 = \sum_{(a,b)\in\partial A\times\partial B} \sum_{(u,u')} \mathbf{P}_A(a \xleftrightarrow{P[2t_{\rm mix},r_0]} u \text{ only on } B)\mathbf{P}_B(b \xleftrightarrow{P[2t_{\rm mix},r_0]} u'),$$

and S_2 is defined in (6.10). Lemma 6.9 gives the required lower bound on S_1, which yields the positive contribution in the assertion of this lemma. We now bound S_2 and S_3 from above, starting from S_3. If $a \xleftrightarrow{P[2t_{\rm mix},r_0]} u$ only on B, then either $a \xleftrightarrow{2t_{\rm mix}} u$ or there exist $z_1 \in B$ and $t_1 \in [t_{\rm mix}, r_0]$ such that

$$\{a \xleftrightarrow{=t_1} z_1\} \circ \{z_1 \xleftrightarrow{r_0-t_1} u\} \quad \text{or} \quad \{a \xleftrightarrow{r_0-t_1} z_1\} \circ \{z_1 \xleftrightarrow{=t_1} u\}.$$

Indeed, let γ be the lexicographically first shortest path between a and u. If $|\gamma| \leq 2t_{\rm mix}$, then $a \xleftrightarrow{2t_{\rm mix}} u$, otherwise $|\gamma| \in [2t_{\rm mix}, r_0]$ and we take z_1 to be the first vertex in B visited by γ, and t is such that $\gamma(t) = z_1$. If $t \geq t_{\rm mix}$, then we set $t_1 = t$ and otherwise $t_1 = |\gamma| - t$. In any case $t_1 \in [t_{\rm mix}, r_0]$. When $t \geq t_{\rm mix}$, the witness for $a \xleftrightarrow{=t_1} z_1$ is the set of open edges of the path $\gamma[0,t]$ together with all the closed edges of the graph, and the witness for $z_1 \xleftrightarrow{r_0-t_1} u$ are the open edges of $\gamma[t,|\gamma|]$. The case $t \leq t_{\rm mix}$ is handled similarly. We get

$$\begin{aligned} S_3 \leq &\sum_{\substack{(a,b)\in\partial A\times\partial B\\(u,u')}} \mathbf{P}_A(a \xleftrightarrow{2t_{\rm mix}} u)\mathbf{P}_B(b \xleftrightarrow{P[2t_{\rm mix},r_0]} u') \\ &+ \sum_{\substack{(a,b)\in\partial A\times\partial B\\(u,u'),\, z_1\in B,\, t_1\in[t_{\rm mix},r_0]}} \mathbf{P}_A(a \xleftrightarrow{=t_1} z_1)\mathbf{P}_A(z_1 \xleftrightarrow{r_0-t_1} u)\mathbf{P}_B(b \xleftrightarrow{P[2t_{\rm mix},r_0]} u') \\ &+ \sum_{\substack{(a,b)\in\partial A\times\partial B\\(u,u'),\, z_1\in B,\, t_1\in[t_{\rm mix},r_0]}} \mathbf{P}_A(a \xleftrightarrow{r_0-t_1} z_1)\mathbf{P}_A(z_1 \xleftrightarrow{=t_1} u)\mathbf{P}_B(b \xleftrightarrow{P[2t_{\rm mix},r_0]} u'). \end{aligned}$$

For the first term we bound $\mathbf{P}_B(b \xleftrightarrow{P[2t_{\text{mix}},r_0]} u') \leq CV^{-1}\mathbb{E}|B(r_0)|$ by Lemma 3.13 and sum over everything to get a contribution bounded by

$$\begin{aligned} C|\partial A|\,|\partial B| m V^{-1}\mathbb{E}|B(r_0)| t_{\text{mix}} &\leq C|\partial A|\,|\partial B| V^{-1} m(\mathbb{E}|B(r_0)|)^2 [t_{\text{mix}}(\mathbb{E}|B(r_0)|)^{-1}] \\ &\leq C\alpha_G^{1/2}|\partial A|\,|\partial B| V^{-1} m(\mathbb{E}|B(r_0)|)^2, \end{aligned} \tag{6.11}$$

by our choice of r_0 in (2.8), our assumptions $\alpha_G \geq (\varepsilon^3 V)^{-1/2}$ in (2.5) and $t_{\text{mix}} = o(\varepsilon^{-1})$, and Corollary 4.6. This fits in the second term of Err in the assertion of the lemma. We bound the second and third terms using Lemma 6.12, giving an upper bound of

$$C|\partial A|\,|\partial B|\,|B| V^{-2} m r_0 (\mathbb{E}|B(r_0)|)^2,$$

which fits in the first term of Err in the assertion of the lemma.

We proceed to bound S_2 in (6.10) from above. As before, if $b \xleftrightarrow{P[2t_{\text{mix}},r_0]} u'$ only on $H \cup A$, then either $b \xleftrightarrow{2t_{\text{mix}}} u'$ or there exists $z_1 \in H \cup A$ and $t_1 \in [t_{\text{mix}}, r_0]$ such that

$$\{b \xleftrightarrow{=t_1} z_1\} \circ \{z_1 \xleftrightarrow{r_0 - t_1} u'\} \quad \text{or} \quad \{b \xleftrightarrow{r_0 - t_1} z_1\} \circ \{z_1 \xleftrightarrow{=t_1} u'\}. \tag{6.12}$$

The case where $b \xleftrightarrow{2t_{\text{mix}}} u'$ is handled as before and gives a contribution of $C|\partial A|\,|\partial B| m V^{-1}\mathbb{E}|B(r_0)| t_{\text{mix}}$ which by (6.11) again fits in the second term of Err. To handle the other cases, let us first sum the contribution to S_2 due to (6.12) over $z_1 \in H$. We use the BK-Reimer inequality and change the order of summation to bound this contribution to S_2 by

$$\begin{aligned} &\sum_{\substack{(a,b)\in\partial A\times\partial B\\ (u,u'),z_1,\,t_1\in[t_{\text{mix}},r_0]}} \mathbf{P}_{A,B}\big(a \xleftrightarrow{P[2t_{\text{mix}},r_0]} u,\ \exists a' \in \partial A \text{ such that } a' \xleftrightarrow{r_0} z_1\big)\mathbf{P}_B(b \xleftrightarrow{=t_1} z_1) \\ &\qquad\qquad \times \mathbf{P}_B(z_1 \xleftrightarrow{r_0 - t_1} u') \\ &+ \sum_{\substack{(a,b)\in\partial A\times\partial B\\ (u,u'),z_1,\,t_1\leq r_0}} \mathbf{P}_{A,B}\big(a \xleftrightarrow{P[2t_{\text{mix}},r_0]} u,\ \exists a' \in \partial A \text{ such that } a' \xleftrightarrow{r_0} z_1\big)\mathbf{P}_B(b \xleftrightarrow{r_0 - t_1} z_1) \\ &\qquad\qquad \times \mathbf{P}_B(z_1 \xleftrightarrow{=t_1} u'). \end{aligned}$$

Now, if $a \xleftrightarrow{P[2t_{\text{mix}},r_0]} u$ and there exists $a' \in \partial A$ with $a' \xleftrightarrow{r_0} z_1$, then either $a \xleftrightarrow{P[2t_{\text{mix}},r_0]} u$ and $a \xleftrightarrow{r_0} z_1$, or there exists $a' \in \partial A$ such that $\{a \xleftrightarrow{P[2t_{\text{mix}},r_0]} u\} \circ \{a' \xleftrightarrow{r_0} z_1\}$. Hence we may bound this from above by $(I) + (II)$ where

$$\begin{aligned} (I) = &\sum_{\substack{(a,b)\in\partial A\times\partial B\\ (u,u'),z_1,\,t_1\in[t_{\text{mix}},r_0]}} \mathbf{P}_{A,B}\big(a \xleftrightarrow{P[2t_{\text{mix}},r_0]} u, a \xleftrightarrow{r_0} z_1\big)\mathbf{P}_B(b \xleftrightarrow{=t_1} z_1)\mathbf{P}_B(z_1 \xleftrightarrow{r_0 - t_1} u') \\ &+ \sum_{\substack{(a,b)\in\partial A\times\partial B\\ (u,u'),z_1,\,t_1\leq r_0}} \mathbf{P}_{A,B}\big(a \xleftrightarrow{P[2t_{\text{mix}},r_0]} u, a \xleftrightarrow{r_0} z_1\big)\mathbf{P}_B(b \xleftrightarrow{r_0 - t_1} z_1)\mathbf{P}_B(z_1 \xleftrightarrow{=t_1} u'), \end{aligned}$$

and

$$(II) = \sum_{\substack{(a,b)\in\partial A\times\partial B,\, a'\in\partial A\\ (u,u'),z_1,\, t_1\in[t_{\mathrm{mix}},r_0]}} \mathbf{P}_{A,B}(a \xleftrightarrow{P[2t_{\mathrm{mix}},r_0]} u)\mathbf{P}_{A,B}(a' \overset{r_0}{\longleftrightarrow} z_1)\mathbf{P}_B(b \overset{=t_1}{\longleftrightarrow} z_1)\mathbf{P}_B(z_1 \xleftrightarrow{r_0-t_1} u')$$
$$+ \sum_{\substack{(a,b)\in\partial A\times\partial B,\, a'\in\partial A\\ (u,u'),z_1,\, t_1\le r_0}} \mathbf{P}_{A,B}(a \xleftrightarrow{P[2t_{\mathrm{mix}},r_0]} u)\mathbf{P}_{A,B}(a' \overset{r_0}{\longleftrightarrow} z_1)\mathbf{P}_B(b \xleftrightarrow{r_0-t_1} z_1)\mathbf{P}_B(z_1 \overset{=t_1}{\longleftrightarrow} u').$$

Lemma 6.10 readily gives $(I) \le \alpha_G^{1/2}|\partial A|\,|\partial B|V^{-1}m(\mathbb{E}|B(r_0)|)^2$, which fits into the second term of Err. Lemma 6.11 implies that $(II) \le \alpha_G^{1/2}\varepsilon|\partial A|^2|\partial B|V^{-1}m(\mathbb{E}|B(r_0)|)^2$, which fits in the second term of Err. We sum the contribution to S_2 due to (6.12) over $z_1 \in A$ and bound it from above by

$$\sum_{\substack{(a,b)\in\partial A\times\partial B\\ (u,u'),\, z_1\in A,\, t_1\in[t_{\mathrm{mix}},r_0]}} \mathbf{P}_{A,B}(a \xleftrightarrow{P[2t_{\mathrm{mix}},r_0]} u)\mathbf{P}_B(b \overset{=t_1}{\longleftrightarrow} z_1)\mathbf{P}_B(z_1 \xleftrightarrow{r_0-t_1} u')$$
$$\le C|\partial A|\,|\partial B|\,|A|V^{-2}mr_0(\mathbb{E}|B(r_0)|)^2,$$

by an appeal to Lemma 6.12. This fits in the first term of Err and concludes our proof. □

Lemma 6.14. *The following bound holds:*

$$\mathbb{E}_{A,B}S_{j_x,j_y,r_0}(x,y)^2\mathbf{1}_{\{\partial A \overset{2r_0}{\nleftrightarrow} \partial B\}} \le Q_1 + Q_2 + Q_3,$$

where

$$Q_1 = \big(1 + O(\alpha_G + \varepsilon t_{\mathrm{mix}})\big)V^{-2}m^2(\mathbb{E}|B(r_0)|)^4|\partial A|^2|\partial B|^2,$$
$$Q_2 = CV^{-2}m^2\varepsilon^{-1}(\mathbb{E}|B(r_0)|)^4|\partial A|\,|\partial B|(|\partial A| + |\partial B|),$$
$$Q_3 = CV^{-2}m^2\varepsilon^{-2}(\mathbb{E}|B(r_0)|)^4|\partial A|\,|\partial B|.$$

Proof. Assume that (u_1,u_1') and (u_2,u_2') are two edges, and let a, a_1, a_2 be vertices in ∂A and b, b_1, b_2 vertices in ∂B. Define

$$\begin{aligned}\mathscr{T}(u_1,u_2,a_1,a_2) &= \{a_1 \xleftrightarrow{P[2t_{\mathrm{mix}},r_0]} u_1\} \circ \{a_2 \xleftrightarrow{P[2t_{\mathrm{mix}},r_0]} u_2\},\\ \mathscr{T}(u_1,u_2,a) &= \{a \xleftrightarrow{P[2t_{\mathrm{mix}},r_0]} u_1\} \cap \{a \xleftrightarrow{P[2t_{\mathrm{mix}},r_0]} u_2\}.\end{aligned} \tag{6.13}$$

We define $\mathscr{T}(u_1',u_2',b_1,b_2)$ and $\mathscr{T}(u_1',u_2',b)$ in a similar fashion.

Now, if (u_1,u_1') and (u_2,u_2') are counted in $S_{j_x,j_y,r_0}(x,y)^2\mathbf{1}_{\{\partial A \overset{2r_0}{\nleftrightarrow} \partial B\}}$, then one of the following events must occur off $A \cup B$:

(1) There exist a_1, a_2, b_1, b_2 such that $\mathscr{T}(u_1,u_2,a_1,a_2) \circ \mathscr{T}(u_1',u_2',b_1,b_2)$ occurs.
(2) There exist a_1, a_2, b such that $\mathscr{T}(u_1,u_2,a_1,a_2) \circ \mathscr{T}(u_1',u_2',b)$ occurs, or the symmetric case $\mathscr{T}(u_1,u_2,a) \circ \mathscr{T}(u_1',u_2',b_1,b_2)$ occurs.
(3) There exist a, b such that $\mathscr{T}(u_1,u_2,a) \circ \mathscr{T}(u_1',u_2',b)$ occurs.

(See Figure 5.) Observe that the disjoint occurrence of the events is implied since $\partial A \overset{2r_0}{\nleftrightarrow} \partial B$. We now sum the probability of these events over $(u_1,u_1'), (u_2,u_2')$; this gives

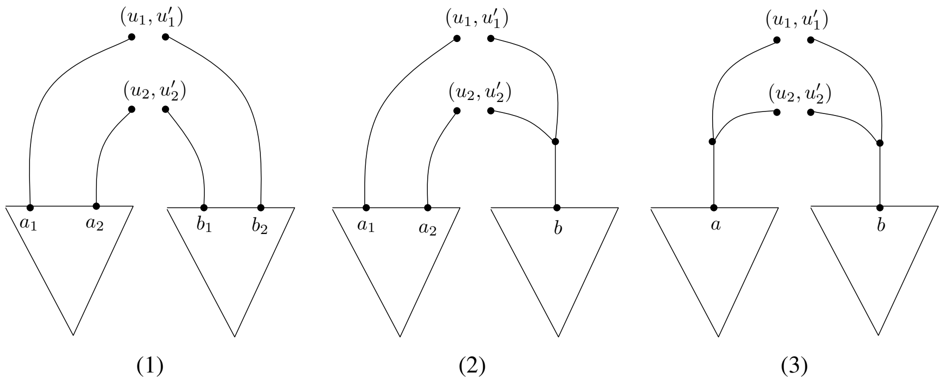

Fig. 5. The three contributions to the second moment of $S_{j_x,j_y,r_0}(x, y)$. The main contribution comes from (1).

three terms which we will bound by Q_1, Q_2 and Q_3, respectively. By Lemma 3.13 and the BK inequality,

$$\mathbf{P}_{A,B}(\mathscr{T}(u_1, u_2, a_1, a_2)) \leq \big(1 + O(\alpha_G + \varepsilon t_{\text{mix}})\big)(\mathbb{E}|B(r_0)|)^2/V^2, \tag{6.14}$$

whence

$$\sum_{\substack{a_1,a_2,b_1,b_2\\(u_1,u_1'),(u_2,u_2')}} \mathbf{P}_{A,B}(\mathscr{T}(u_1, u_2, a_1, a_2) \circ \mathscr{T}(u_1', u_2', b_1, b_2)) \\ \leq \big(1 + O(\alpha_G + \varepsilon t_{\text{mix}})\big)V^{-2}m^2(\mathbb{E}|B(r_0)|)^4|\partial A|^2|\partial B|^2,$$

which equals Q_1. To bound the probability of (2), if $\mathscr{T}(u_1', u_2', b)$ occurs, then as before either there exists a vertex z_1 and $t_1 \in [t_{\text{mix}}, r_0]$ such that

$$\{b \overset{=t_1}{\longleftrightarrow} z_1\} \circ \{z_1 \overset{r_0-t_1}{\longleftrightarrow} u_1'\} \circ \{z_1 \overset{r_0-t_1}{\longleftrightarrow} u_2'\},$$

or there exists z_1 such that

$$\{b \overset{t_{\text{mix}}}{\longleftrightarrow} z_1\} \circ \{z_1 \xleftrightarrow{P[t_{\text{mix}},r_0]} u_1'\} \circ \{z_1 \xleftrightarrow{P[t_{\text{mix}},r_0]} u_2'\},$$

or there exists z_1 such that

$$\{b \xleftrightarrow{P[t_{\text{mix}},r_0]} z_1\} \circ \{z_1 \xleftrightarrow{P[t_{\text{mix}},r_0]} u_1'\} \circ \{z_1 \overset{t_{\text{mix}}}{\longleftrightarrow} u_2'\}.$$

Hence, the BK-Reimer inequality gives

$$\begin{aligned}
&\sum_{u_1',u_2'} \mathbf{P}_{A,B}(\mathscr{T}(u_1', u_2', b)) \\
&\quad \leq \sum_{\substack{u_1',u_2',z_1\\t_1\in[t_{\text{mix}},r_0]}} \mathbf{P}_{A,B}(b \overset{=t_1}{\longleftrightarrow} z_1)\mathbf{P}_{A,B}(z_1 \overset{r_0-t_1}{\longleftrightarrow} u_1')\mathbf{P}_{A,B}(z_1 \overset{r_0-t_1}{\longleftrightarrow} u_2') \\
&\qquad + \sum_{u_1',u_2',z_1} \mathbf{P}_{A,B}(b \overset{t_{\text{mix}}}{\longleftrightarrow} z_1)\mathbf{P}_{A,B}(z_1 \xleftrightarrow{P[t_{\text{mix}},r_0]} u_1')\mathbf{P}_{A,B}(z_1 \xleftrightarrow{P[t_{\text{mix}},r_0]} u_2') \\
&\qquad + \sum_{u_1',u_2',z_1} \mathbf{P}_{A,B}(b \xleftrightarrow{P[t_{\text{mix}},r_0]} z_1)\mathbf{P}_{A,B}(z_1 \xleftrightarrow{P[t_{\text{mix}},r_0]} u_1')\mathbf{P}_{A,B}(z_1 \overset{t_{\text{mix}}}{\longleftrightarrow} u_2').
\end{aligned}$$

We estimate the first sum by summing on u_2', u_1' then on z_1, t_1 using Lemma 4.4 to get a bound of

$$C\varepsilon^{-3}(1+\varepsilon)^{2r_0} \le C\varepsilon^{-1}(\mathbb{E}|B(r_0)|)^2,$$

by Theorem 4.5. The second and third sums are bounded by $Ct_{\mathrm{mix}}(\mathbb{E}|B(r_0)|)^2$, which is of lower order since $\varepsilon t_{\mathrm{mix}} = o(1)$ by (2.9). We use the BK inequality and (6.14) to bound the contribution due to the first event in (2) by

$$\sum_{\substack{a_1,a_2,b,(u_1,u_1')\\(u_2,u_2')}} \mathbf{P}_{A,B}(\mathscr{T}(u_1,u_2,a_1,a_2))\mathbf{P}_{A,B}(\mathscr{T}(u_1',u_2',b)) \le CV^{-2}m^2\varepsilon^{-1}(\mathbb{E}|B(r_0)|)^4|\partial A|^2|\partial B|.$$

The symmetric case in (2) obeys the same bound with the roles of $|\partial A|$ and $|\partial B|$ reversed. This contribution equals Q_2. To bound the contribution due to (3), we note that

$$\sum_{a,b,(u_1,u_1'),(u_2,u_2')} \mathbf{P}_{A,B}(\mathscr{T}(u_1,u_2,a))\mathbf{P}_{A,B}(\mathscr{T}(u_1',u_2',b))$$

is bounded using the BK-Reimer inequality by the three sums

$$\sum_{\substack{a,b,(u_1,u_1'),(u_2,u_2')\\ z_1,z_2,t_1,t_2\in[t_{\mathrm{mix}},r_0]}} \mathbf{P}_{A,B}(a \overset{=t_1}{\longleftrightarrow} z_1)\mathbf{P}_{A,B}(z_1 \overset{r_0-t_1}{\longleftrightarrow} u_1)\mathbf{P}_{A,B}(z_1 \overset{r_0-t_1}{\longleftrightarrow} u_2)\mathbf{P}_{A,B}(b \overset{=t_2}{\longleftrightarrow} z_2) \mathbf{P}_{A,B}(z_2 \overset{r_0-t_2}{\longleftrightarrow} u_1')\mathbf{P}_{A,B}(z_2 \overset{r_0-t_2}{\longleftrightarrow} u_2'),$$

$$\sum_{\substack{a,b,(u_1,u_1'),(u_2,u_2')\\ z_1,z_2,t_1\in[t_{\mathrm{mix}},r_0]}} \mathbf{P}_{A,B}(a \overset{=t_1}{\longleftrightarrow} z_1)\mathbf{P}_{A,B}(z_1 \overset{r_0-t_1}{\longleftrightarrow} u_1)\mathbf{P}_{A,B}(z_1 \overset{r_0-t_1}{\longleftrightarrow} u_2)\mathbf{P}_{A,B}(b \overset{t_{\mathrm{mix}}}{\longleftrightarrow} z_2) \mathbf{P}_{A,B}(z_2 \overset{P[t_{\mathrm{mix}},r_0]}{\longleftrightarrow} u_1')\mathbf{P}_{A,B}(z_2 \overset{P[t_{\mathrm{mix}},r_0]}{\longleftrightarrow} u_2'),$$

$$\sum_{\substack{a,b,(u_1,u_1')\\(u_2,u_2'),z_1,z_2}} \mathbf{P}_{A,B}(a \overset{t_{\mathrm{mix}}}{\longleftrightarrow} z_1)\mathbf{P}_{A,B}(z_1 \overset{P[t_{\mathrm{mix}},r_0]}{\longleftrightarrow} u_1)\mathbf{P}_{A,B}(z_1 \overset{P[t_{\mathrm{mix}},r_0]}{\longleftrightarrow} u_2) \mathbf{P}_{A,B}(b \overset{t_{\mathrm{mix}}}{\longleftrightarrow} z_2)\mathbf{P}_{A,B}(z_2 \overset{P[t_{\mathrm{mix}},r_0]}{\longleftrightarrow} u_1')\mathbf{P}_{A,B}(z_2 \overset{P[t_{\mathrm{mix}},r_0]}{\longleftrightarrow} u_2').$$

To bound the first sum, we use Lemmas 3.12 and 4.4 to bound $\mathbf{P}_{A,B}(a \overset{=t_1}{\longleftrightarrow} z_1) \le CV^{-1}(1+\varepsilon)^{t_1}$ and $\mathbf{P}_{A,B}(b \overset{=t_2}{\longleftrightarrow} z_2) \le CV^{-1}(1+\varepsilon)^{t_2}$. We then use Lemmas 3.19 and 4.4 to sum over $z_1, z_2, (u_1,u_1'), (u_2,u_2')$. This gives us an upper bound of

$$CV^{-2}|\partial A|\,|\partial B|m^2\varepsilon^{-6}(1+\varepsilon)^{4r_0} + C|\partial A|\,|\partial B|V^{-1}m^2 t_{\mathrm{mix}}\alpha_G \sum_{t_1,t_2\in[t_{\mathrm{mix}},r_0]} (1+\varepsilon)^{t_1+t_2} \le CV^{-2}|\partial A|\,|\partial B|m^2\varepsilon^{-2}(\mathbb{E}|B(r_0)|)^4,$$

where the last inequality is due to Theorem 4.5 and our choice of r_0 in (2.8). This is contained in Q_3. To bound the second sum, we use Lemma 3.15 to bound each of the last two terms by $CV^{-1}\mathbb{E}|B(r_0)|$. We then sum over (u_1,u_1') and (u_2,u_2') using Lemma 4.4. Next we sum over z_1, z_2 using Lemma 4.4 and finally over a, b, t_1 to conclude that this sum is at most

$$C|\partial A|\,|\partial B|V^{-2}m^2(\mathbb{E}|B(r_0)|)^4\varepsilon^{-1}t_{\mathrm{mix}},$$

which is contained in Q_3 since $\varepsilon t_{\mathrm{mix}} = o(1)$. For the third sum we use Lemma 3.15 four times, and then sum over everything to get a bound of

$$C|\partial A|\,|\partial B|V^{-2}m^2(\mathbb{E}|B(r_0)|)^4 t_{\mathrm{mix}}^2,$$

which is also contained in Q_3, concluding our proof. □

Proof of Theorem 6.3. We condition on the events $B_x(j_x) = A$ and $B_y(j_y) = B$ such that the event $\mathscr{A}(x, y, j_x, j_y, r_0, \beta, k)$ holds. By requirement (2) of $\mathscr{A}(x, y, j_x, j_y, r_0, \beta, k)$ (p. 771) and our choice of parameters,

$$r_0(|A| + |B|)V^{-1} \le V^{-1}\varepsilon^{-3}\log(\varepsilon^3 V)\mathrm{e}^{3M} \le (\log(\varepsilon^3 V))^{-1},$$

and

$$\alpha_G^{1/2}\varepsilon|\partial A| \le \mathrm{e}^{3M}\alpha_G^{1/2} \le \alpha_G^{1/4}.$$

Hence the error term in Lemma 6.13 is at most

$$\mathrm{Err} \le C[(\log(\varepsilon^3 V))^{-1} + \alpha_G^{1/4}]|\partial A|\,|\partial B|V^{-1}m(\mathbb{E}|B(r_0)|)^2.$$

Lemma 6.13 together with requirement (5) in the definition of $\mathscr{A}(x, y, j_x, j_y, r_0, \beta, k)$ and our choice of β in (6.1) (in particular, $\beta \ll \alpha_G^{1/4} \wedge (\log(\varepsilon^3 V))^{-1}$ by (2.7)) give

$$\mathbb{E}_{A,B}\big[S_{j_x,j_y,r_0}(x, y)\mathbf{1}_{\{\partial A \overset{2r_0}{\nleftrightarrow} \partial B\}}\big] \ge (1 - C\beta^{1/2})V^{-1}m(\mathbb{E}|B(r_0)|)^2|\partial A|\,|\partial B|.$$

Since $|\partial A|$ and $|\partial B|$ are at least $\mathrm{e}^{k/4}\varepsilon^{-1}$, we have

$$\varepsilon^{-1}|\partial A|^2|\partial B| + \varepsilon^{-1}|\partial A|^2|\partial B| + \varepsilon^{-2}|\partial A|\,|\partial B| \le C\mathrm{e}^{-k/4}|\partial A|^2|\partial B|^2,$$

hence, by Lemma 6.14 and our choice of parameters,

$$\mathbb{E}_{A,B}\big[S_{j_x,j_y,r_0}(x, y)^2\mathbf{1}_{\{\partial A \overset{2r_0}{\nleftrightarrow} \partial B\}}\big] \le (1 + O(\mathrm{e}^{-k/4}))V^{-2}m^2(\mathbb{E}|B(r_0)|)^4|\partial A|^2|\partial B|^2.$$

We conclude that

$$\mathbf{P}_{A,B}\big(S_{j_x,j_y,r_0}(x, y) \ge 2\beta^{1/2}V^{-1}m(\mathbb{E}|B(r_0)|)^2|\partial A|\,|\partial B|\big) \ge 1 - O(\beta^{1/2}),$$

where we have used the fact that $\mathrm{e}^{-k/4} = o(\beta)$ and (4.4). This concludes our proof since $|\partial A|$ and $|\partial B|$ are at least ε^{-1}. □

7. Proofs of main theorems

7.1. Proof of Theorem 1.3. In Section 2.4 we already proved Theorem 1.3(a), so we may assume that the finite triangle condition (1.5) holds and focus on part (b). Since $|\mathscr{C}_1| \le k_0 + Z_{\ge k_0}$ where k_0 is from Theorem 2.2, Lemma 2.3 immediately gives $|\mathscr{C}_1| \le (2 + o(1))\varepsilon V$ whp, showing the required upper bound on $|\mathscr{C}_1|$—note that this argument only uses the finite triangle condition, hence it is valid for any $\varepsilon(m)$ satisfying $\varepsilon(m) \gg V^{-1/3}$ and $\varepsilon(m) = o(1)$. For the lower bound we will additionally assume, as part (b) requires, that $\varepsilon(m) = o(1/t_{\mathrm{mix}})$, and show that

$$\mathbf{P}_p\big(|\mathscr{C}_1| \geq (2 - o(1))\varepsilon V\big) = 1 - o(1). \tag{7.1}$$

This establishes (b) of Theorem 1.3. Recall that $p = p_c(1+\varepsilon)$ is our percolation probability, let $\theta > 0$ be an arbitrary small constant and let p_2, p_1 satisfy

$$p_2 = \theta\varepsilon/m, \qquad p_c(1+\varepsilon) = p_1 + (1-p_1)p_2,$$

so that $p_c(1 + (1-\theta)\varepsilon) \leq p_1 \leq p_c(1+\varepsilon)$ since $p_c \geq 1/m$. Denote by G_{p_1} and G_{p_2} two independent percolation instances of G with parameters p_1 and p_2, respectively. The sprinkling procedure relies on the fact that G_p is distributed as $G_{p_1} \cup G_{p_2}$. We first apply Theorem 2.4 to G_{p_1} and deduce that for M, r defined in (2.7) and r_0 defined in (2.8),

$$\mathbf{P}_{p_1}\big(P_{r,r_0} \geq (1-3\theta)4\varepsilon^2 V^2\big) \geq 1 - o(1). \tag{7.2}$$

Now we wish to show that if we "sprinkle" this configuration in G_{p_1}, that is, we add independent p_2-open edges, then most of these vertices join together to form one cluster of size roughly $2\varepsilon V$. To make this formal, given G_{p_1}, we construct an auxiliary simple graph H with vertex set

$$V(H) = \{x \in G_{p_1} : |\mathscr{C}(x)| \geq (\varepsilon^3 V)^{1/4}\varepsilon^{-2}\}$$

and edge set

$$E(H) = \{(x,y) \in V(H)^2 : x, y \text{ are } (r, r_0)\text{-good}\}.$$

Thus, using Lemma 2.3 with $k_0 = \varepsilon^{-2}(\varepsilon^3 V)^{1/4}$ and (7.2), with probability at least $1 - o(1)$,

$$|V(H)| = (2 + o(1))\varepsilon V, \qquad |E(H)| \geq (1-3\theta)4\varepsilon^2 V^2. \tag{7.3}$$

Denote $v = |V(H)|$ and write $x_1, \ldots, x_v$ for the vertices in G_{p_1} corresponding to those of H. Given G_{p_1} for which the event in (7.3) occurs, we will show that whp in $G_{p_1} \cup G_{p_2}$ there is no way to partition the set of vertices into $M_1 \uplus M_2 = \{x_1, \ldots, x_v\}$ with $|M_1| \geq 3\theta v$ and $|M_2| \geq 3\theta v$ so that there is no open path in $G_{p_1} \cup G_{p_2}$ connecting a vertex in M_1 to a vertex in M_2. This implies that whp the largest connected component in $G_{p_1} \cup G_{p_2}$ is of size at least $(1-3\theta)v$.

To show this, we first claim that the number of such partitions is at most $2^{3(\varepsilon^3 V)^{3/4}}$ since $|\mathscr{C}(x_i)| \geq (\varepsilon^3 V)^{1/4}\varepsilon^{-2}$. Secondly, given such a partition, we claim that the number of edges (u, u') such that $u \in M_1$ and $u' \in M_2$ (note that, by definition, these edges must be p_1-closed) is at least $\mathrm{e}^{-40M}(\log M)^{-1}\theta\varepsilon^2 Vm$. To see this, we consider the set of edges in H for which both endpoints of the edge lie in either M_1 or M_2 (more precisely, the vertices of H corresponding to M_1 and M_2). This number is at most

$$|M_1|^2 + |M_2|^2 \leq (3\theta v)^2 + (1-3\theta)^2 v^2 \leq (1-5\theta)v^2,$$

where we have used the fact that $\theta > 0$ is a small enough constant, $|M_1| + |M_2| = v$ and both $|M_1|$ and $|M_2|$ are in $[3\theta v, (1-3\theta)v]$. By (7.3), the number of edges in H such that one end is in M_1 and the other in M_2 is at least $\theta\varepsilon^2 V^2$. In other words, there are at least $\theta\varepsilon^2 V^2$ pairs $(x, y) \in M_1 \times M_2$ such that $S_{2r+r_0}(x, y) \geq (\log M)^{-1} V^{-1} m\varepsilon^{-2}(\mathbb{E}|B(r_0)|)^2$. Note that is a large number due to our condition (2.8). In total, we counted at least $\theta\varepsilon^2 V^2 \cdot (\log M)^{-1} V^{-1} m\varepsilon^{-2}(\mathbb{E}|B(r_0)|)^2$ edges (u, u') and no edge is counted more than $|B_u(2r+r_0)|\cdot|B_{u'}(2r+r_0)|$ times, which is at most $\mathrm{e}^{40M}\varepsilon^{-2}(\mathbb{E}|B(r_0)|)^2$ by the definition of $S_{2r+r_0}(x, y)$, and the second claim follows.

Hence, if $|\mathscr{C}_1| \leq (1-3\theta)v$, then there exists such a partition in which all of the above edges (u, u') are p_2-closed. By the two claims above, the probability of this is at most

$$2^{3(\varepsilon^3 V)^{3/4}}(1-p_2)^{\mathrm{e}^{-40M}(\log M)^{-1}\theta m\varepsilon^2 V} \leq 2^{3(\varepsilon^3 V)^{3/4}}\mathrm{e}^{-\mathrm{e}^{-40M}(\log M)^{-1}\theta^2\varepsilon^3 V} = o(1),$$

since $p_2 = \theta\varepsilon/m$ and by our choice of parameters in (2.7) and (2.9). This concludes the proof of (7.1) since $\theta > 0$ was arbitrary, and establishes the required estimate on $|\mathscr{C}_1|$ of Theorem 1.3(b).

We now use (7.1) to show the required bounds on $\mathbb{E}|\mathscr{C}(0)|$ and $|\mathscr{C}_2|$. The upper bound $\mathbb{E}|\mathscr{C}(0)| \leq (4+o(1))\varepsilon^2 V$ is stated in Lemma 2.3, and the lower bound follows immediately from our estimate on $\mathscr{C}_1$. Indeed, write $\mathscr{C}_j$ for the jth largest component. Then

$$\mathbb{E}|\mathscr{C}(0)| = V^{-1}\sum_{v\in V(G)}\mathbb{E}|\mathscr{C}(v)| = V^{-1}\sum_{j\geq 1}\mathbb{E}|\mathscr{C}_j|^2 \geq V^{-1}\mathbb{E}|\mathscr{C}_1|^2 \geq (4-o(1))\varepsilon^2 V,$$

where the first equality is by transitivity, the second equality is because each component $\mathscr{C}_j$ is counted $|\mathscr{C}_j|$ times in the sum on the left, and the last inequality is due to (7.1). Furthermore, by this inequality and Lemma 2.3, we deduce that

$$\sum_{j\geq 2}\mathbb{E}|\mathscr{C}_j|^2 = o(\varepsilon^2 V^2),$$

and hence $|\mathscr{C}_2| = o(\varepsilon V)$ whp. This concludes the proof of Theorem 1.3. □

7.2. Proof of Theorem 1.1. In this section we restrict our attention to the hypercube and prove Theorem 1.1. We begin by showing that t_{mix}, defined in Theorem 1.3 with $\alpha_G = m^{-1}\log m$, satisfies $t_{\mathrm{mix}} = O(m\log m)$: see Lemma 7.1. The proof of Theorem 1.1 is then split into two cases. In the first case we assume that $\varepsilon(m) \leq 1/m^2$ so that $\varepsilon = o(1/t_{\mathrm{mix}})$ and appeal to Theorem 1.3. In the second case we perform the classical sprinkling argument for the case $\varepsilon \geq 1/m^2$, as done in [18].

Lemma 7.1 (NBW estimates). *On the hypercube* $\{0, 1\}^m$*,*

$$T_{\mathrm{mix}}(m^{-1}\log m) = O(m\log m),$$

and for any integer $L \geq 1$*,*

$$\sup_{x,y}\sum_{u,v}\sum_{\substack{t_1,t_2,t_3=0\\ t_1+t_2+t_3\geq 3}}^{L}\mathbf{p}^{t_1}(x,u)\mathbf{p}^{t_2}(u,v)\mathbf{p}^{t_3}(v,y) \leq O(1/m^2) + O(L^3/V). \tag{7.4}$$

Proof. We make use of the results in [25], as we explain now. The estimate $T_{\mathrm{mix}}(m^{-1}\log m) = O(m\log m)$ is [25, Theorem 3.5]. We next explain how to prove (7.4), which will give condition (3) in Theorem 1.3 for $L = Am\log m$ and an appropriate $A > 0$.

Let $D\colon \{0,1\}^m \to [0,1]$ be the simple random walk transition probability on the hypercube, that is, $D(v) = 1/m$ whenever v is a neighbor of the all-zero vector. Our proof of (7.4) relies on Fourier theory. For convenience, we take the Fourier dual of $\{0,1\}^m$ to be $\{0,1\}^m$. Then the Fourier transform $\hat{f}$ of $f\colon \{0,1\}^m \to \mathbb{R}$ is given by

$$\hat{f}(k) = \sum_{x\in\{0,1\}^m} (-1)^{x\cdot k} f(x), \tag{7.5}$$

with inverse Fourier transform

$$f(x) = \frac{1}{V} \sum_{k\in\{0,1\}^m} (-1)^{x\cdot k} \hat{f}(k). \tag{7.6}$$

For the hypercube, $\hat{D}(k)$ takes the appealingly simple form

$$\hat{D}(k) = 1 - 2a(k)/m, \tag{7.7}$$

where $a(k)$ is the number of non-zero coordinates of k.

In [25, Theorem 3.5] it is proved that for $m \geq 2$ and $t \geq 1$, with $\hat{\mathbf{p}}^t(k)$ denoting the Fourier transform of $x \mapsto \mathbf{p}^t(0,x)$,

$$|\hat{\mathbf{p}}^t(k)| \leq \max\bigl(|\hat{D}(k)|, 1/\sqrt{m-1}\bigr)^{t-1}, \tag{7.8}$$

and $\hat{\mathbf{p}}^0(k) = 1$. This gives us all the necessary bounds to prove the NBW triangle condition (7.4).

Denote the sum in (7.4) by S. The contribution to S of $t_1+t_2+t_3 = 3$ equals $O(1/m^2)$. Thus, we are left to bound the contribution due to t_1, t_2, t_3 with $t_1 + t_2 + t_3 \geq 4$. For any $t \geq 1$,

$$\mathbf{p}^t(x,y) \leq \frac{m}{m-1}(D * \mathbf{p}^{t-1})(x,y), \tag{7.9}$$

where, for $f, g\colon \{0,1\}^m \to \mathbb{R}$, we define the convolution $f * g$ by

$$(f * g)(x) = \sum_{y\in\{0,1\}^m} f(y)g(x-y). \tag{7.10}$$

Therefore,

$$S \leq C/m^2 + 3^4\left(\frac{m}{m-1}\right)^4 \sup_{x,y} \sum_{\substack{s_1,s_2,s_3=0\\ s_1+s_2+s_3\geq 0}}^{L} (D^{*4} * \mathbf{p}^{s_1} * \mathbf{p}^{s_2} * \mathbf{p}^{s_3})(x,y), \tag{7.11}$$

where 3^4 is an upper bound on the number of ways we can add 4 to the coordinates of (s_1, s_2, s_3) to get (t_1, t_2, t_3) with $t_1 + t_2 + t + 3 \geq 4$. The above can be bounded in terms of Fourier transforms as

$$\begin{aligned} S &\leq C/m^2 + \frac{C}{V} \sup_{x,y} \sum_{k\in\{0,1\}^m} (-1)^{k\cdot(y-x)} \sum_{\substack{s_1,s_2,s_3=0\\ s_1+s_2+s_3\geq 0}}^{L} \hat{D}(k)^4 \hat{\mathbf{p}}^{s_1}(k)\hat{\mathbf{p}}^{s_2}(k)\hat{\mathbf{p}}^{s_3}(k) \\ &\leq C/m^2 + \frac{C}{V} \sum_{k\in\{0,1\}^m} \sum_{\substack{s_1,s_2,s_3=0\\ s_1+s_2+s_3\geq 0}}^{L} \hat{D}(k)^4 |\hat{\mathbf{p}}^{s_1}(k)|\,|\hat{\mathbf{p}}^{s_2}(k)|\,|\hat{\mathbf{p}}^{s_3}(k)|. \end{aligned} \tag{7.12}$$

The contribution of $k = 0$ equals L^3/V since $\hat{D}(0) = \hat{\mathbf{p}}^t(0) = 1$, and the contribution due to $k = 1$ (where 1 denotes the all-1 vector) obeys the same bound. It is not hard to adapt the proof of [17, Proposition 1.2] to show that the sum over $k \neq 0, 1$ is $O(1/m^2)$. We give the details of this computation now.

Writing $x_+ = \max(x, 0)$ for $x \in \mathbb{R}$, and noting that there are at most two values of s for which $(s-1)_+ = t$, we obtain

$$\begin{aligned}
S &\le C/m^2 + 2L^3/V \\
&\quad + \frac{C}{V} \sum_{k \in \{0,1\}^m :\, k \neq 0,1} \sum_{s_1,s_2,s_3=0}^{L} \hat{D}(k)^4 \max\big(|\hat{D}(k)|, 1/\sqrt{m-1}\big)^{(s_1-1)_+ + (s_2-1)_+ + (s_3-1)_+} \\
&\le C/m^2 + 2L^3/V \\
&\quad + \frac{C2^3}{V} \sum_{k \in \{0,1\}^m :\, k \neq 0,1} \hat{D}(k)^4 \sum_{s_1,s_2,s_3=0}^{\infty} \max\big(|\hat{D}(k)|, 1/\sqrt{m-1}\big)^{s_1+s_2+s_3} \\
&= C/m^2 + 2L^3/V + \frac{C2^3}{V} \sum_{k \in \{0,1\}^m :\, k \neq 0,1} \frac{\hat{D}(k)^4}{\big[1 - \max\big(|\hat{D}(k)|, 1/\sqrt{m-1}\big)\big]^3}.
\end{aligned} \tag{7.13}$$

We bound

$$\begin{aligned}
&\frac{1}{V} \sum_{k \in \{0,1\}^m :\, k \neq 0,1} \frac{\hat{D}(k)^4}{\big[1 - \max\big(|\hat{D}(k)|, 1/\sqrt{m-1}\big)\big]^3} \\
&\qquad \le \frac{1}{V} \sum_{k \in \{0,1\}^m :\, k \neq 0,1} \hat{D}(k)^4 \left[\frac{1}{[1 - |\hat{D}(k)|]^3} + \frac{1}{\big[1 - 1/\sqrt{m-1}\big]^3} \right].
\end{aligned} \tag{7.14}$$

We next use the fact that $V^{-1} \sum_{k \in \{0,1\}^m} \hat{D}(k)^4$ is the probability that a four-step simple random walk on the hypercube returns to its starting point, which is $O(1/m^2)$. Alternatively, and more useful for the proof that follows, we can write

$$\begin{aligned}
\frac{1}{V} \sum_{k \in \{0,1\}^m} \hat{D}(k)^4 &= 2^{-m} \sum_{j=0}^{m} \binom{m}{j} (1 - 2j/m)^4 \\
&= m^{-4} \mathbb{E}[(2X - m)^4] = O(1/m^2),
\end{aligned} \tag{7.15}$$

where X has a binomial distribution with parameters $1/2$ and m, and we use the fact that $\mathbb{E}[(2X - m)^4] = O(m^2)$. We use similar ideas to deal with the contribution involving $[1 - |\hat{D}(k)|]^{-3}$, which we rewrite as

$$\frac{1}{V} \sum_{k \in \{0,1\}^m :\, k \neq 0,1} \frac{\hat{D}(k)^4}{[1 - |\hat{D}(k)|]^3} = 2^{-m} \sum_{j=1}^{m-1} \binom{m}{j} \frac{(1 - 2j/m)^4}{[(2j/m) \wedge (2 - 2j/m)]^3}. \tag{7.16}$$

The sum $2^{-m} \sum_{j \notin [m/4, 3m/4]} \binom{m}{j}$ is exponentially small in m by either Stirling's formula or large deviation bounds on the binomial distribution with parameters m and $1/2$. When

$j \in [m/4, 3m/4]$, we can bound $1/[(2j/m) \wedge (2-2j/m)]^3 \le 8$ to estimate the above sum by $O(1/m^2)$ in the same way as in (7.15). Together with (7.13), this completes the proof of (7.4). □

Proof of Theorem 1.1. We start by proving the theorem in the case $\varepsilon(m) \le 1/m^2$. We take $\alpha_G = m^{-1}\log m$. Lemma 7.1 shows that $t_{\text{mix}} = O(m \log m)$ and condition (3) of Theorem 1.3 holds. Condition (2) of Theorem 1.3 holds by (1.2). Condition (1) is fulfilled automatically, so in this case Theorem 1.1 follows from Theorem 1.3.

We now handle the case $\varepsilon \ge 1/m^2$ and $\varepsilon = o(1)$. We start by proving (7.1) in this case. In [18], it is proven that $|\mathscr{C}_1| \ge c\varepsilon V$ whp in this case, and the argument used there is based on isoperimetry together with Lemma 2.3 and suffices to prove the required $2\varepsilon V$ estimate in our setting as well, as we now show.

Let $\theta > 0$ be a small arbitrary constant. As before, fix the sprinkling probability $p_2 = \theta\varepsilon/m$ and take p_1 such that $p = p_c(1+\varepsilon) = p_1 + (1-p_1)p_2$ so that $p_1 = p_c(1+(1-\theta+o(1))\varepsilon)$. By Lemma 2.3, whp in G_{p_1},

$$2(1-2\theta)\varepsilon V \le Z_{\ge k_0} \le 2(1+\theta)\varepsilon V$$

for $k_0 = \varepsilon^{-2}(\varepsilon^3 V)^{1/4}$. As a result, there are at most $2(1+\theta)\varepsilon V/k_0 = 2(1+\theta)(\varepsilon^3 V)^{3/4}$ clusters of size at least k_0. Denote these clusters by $(\mathscr{D}_i)_{i\in I}$, so that $|I| \le 2(1+\theta)(\varepsilon^3 V)^{3/4}$.

As before, we now perform sprinkling and add the edges of G_{p_2}. We bound the probability that after sprinkling there is a partition of the clusters $(\mathscr{D}_i)_{i\in I}$ into two sets S, T both containing at least $\theta\varepsilon V$ vertices such that there is no path in G_{p_2} connecting them. If there is no such partition, then the largest component in $G_{p_1} \cup G_{p_2}$ has size at least $(2-3\theta)\varepsilon V$ and we conclude the proof. We follow [18, proof of Proposition 2.5].

Since $|I| \le 2(1+\theta)(\varepsilon^3 V)^{3/4}$, the number of such partitions is at most $2^{2(1+\theta)(\varepsilon^3 V)^{3/4}}$. We estimate the probability that given such a partition there is no p_2-open path connecting them. By [18, Lemma 2.4], whenever $\Delta \ge 1$ satisfies

$$e^{-\Delta^2/2m} \le \theta\varepsilon/2, \tag{7.17}$$

there is a collection of at least $\frac{1}{2}\theta\varepsilon m^{-2\Delta} V$ edge disjoint paths connecting S and T, each of length at most Δ. This is where the isoperimetric inequality on the hypercube is being used. Note that Δ needs to be large, in fact, we set $\Delta = m^{2/3}$ and use the fact that $\varepsilon \ge m^{-2}$ so that (7.17) holds. The probability that a path of length Δ has a p_2-closed edge in it is $1 - p_2^\Delta$. Since the paths are disjoint, these events are independent, and we find that the probability that they all have a p_2-closed edge in them is at most

$$(1-p_2^\Delta)^{\frac{1}{2}\theta\varepsilon m^{-2\Delta} V} \le e^{-cp_2^\Delta \theta\varepsilon m^{-2\Delta} V} = e^{-c\theta^\Delta \varepsilon^\Delta m^{-3\Delta} V}. \tag{7.18}$$

Thus, the total probability that sprinkling fails is at most

$$2^{2(1+\theta)(\varepsilon^3 V)^{3/4}} e^{-c\theta^\Delta\varepsilon^\Delta m^{-3\Delta} V} = e^{-c2^{(1-o(1))m}}, \tag{7.19}$$

since $\varepsilon \ge m^{-2}$ (in fact, this argument works as long as $\varepsilon \ge e^{-cm^{1/3}}$). The proof of (7.1) when $\varepsilon \gg V^{-1/3}$ and $\varepsilon = o(1)$ is now completed.

The remaining estimations of $|\mathscr{C}_1|$, $\mathbb{E}|\mathscr{C}(0)|$ and $|\mathscr{C}_2|$ only rely on (7.1) and Lemma 2.3 and are performed exactly as in the conclusion of the proof of Theorem 1.3. This completes the proof of Theorem 1.1. □

7.3. Proof of Theorem 1.4. The expansion and girth assumptions of the theorem allow us to deduce some crude yet sufficient bounds on $\mathbf{p}^t(\cdot,\cdot)$, namely, that there exists some constant $q>0$ such that

$$\mathbf{p}^t(0,0) \le \begin{cases} V^{-q}, & t \le C\log V, \\ CV^{-1}, & t \ge C\log V, \end{cases} \qquad \mathbf{p}^t(x,y) \le \begin{cases} (m-1)^{-t}, & t \le (c\log_{m-1} V)/2, \\ V^{-q}, & t \ge (c\log_{m-1} V)/2. \end{cases}$$

Indeed, the second bound on $\mathbf{p}^t(0,0)$ comes from the classical fact that $T_{\mathrm{mix}}(CV^{-1}) = O(\log V)$ (see e.g. [5, below (19)]). The first bound on $\mathbf{p}^t(0,0)$ comes from the girth assumption. Indeed, the graph induced on the vertices of graph distance at most $\lfloor g/2 \rfloor$, where g is the girth, is a tree. Hence, in order for the walker to return to 0 at time t, he must be at distance $t-\lfloor g/2\rfloor$ from 0 and then take the unique path of length $\lfloor g/2\rfloor$ to 0 so that q can be taken to be any number smaller than $c/2$. The bounds on $\mathbf{p}^t(x,y)$ are proved similarly.

We take $\alpha_G = C(\log V)^{-1}$ (which is at least $1/m$ by our assumption that $m \ge c\log V$) and prove that conditions (2) and (3) of Theorem 1.3 hold. Note that $t_{\mathrm{mix}} = O(\log V)$. To prove condition (2) we show that percolation with $p = (m-1)^{-1}(1+\alpha_G/\log V)$ has $\mathbb{E}_p|\mathscr{C}(0)| \gg V^{1/3}$ so that $p_c \le p$. To get this lower bound on $\mathbb{E}_p|\mathscr{C}(0)|$ in this regime of p one could use a classical sprinkling argument. However, it is quicker to apply [47, Theorem 4] and verify that

$$\varepsilon^{-1} r \sum_{t=1}^{2r} [(1+\varepsilon)^{t\wedge r} - 1]\mathbf{p}^t(0,0) = o(1), \tag{7.20}$$

where $\varepsilon = \alpha_G/\log V$ and $r = \varepsilon^{-1}[\log(\varepsilon^3 V) - 3\log\log(\varepsilon^3 V)]$. Theorem 4 of [47] then yields $\mathbf{P}(|\mathscr{C}_1| \ge b\varepsilon V/(\log(\varepsilon^3 V))^3) = 1-o(1)$ for some $b>0$, which immediately gives a lower bound on $\mathbb{E}|\mathscr{C}(0)|$ since

$$\mathbb{E}|\mathscr{C}(0)| \ge V^{-1}\mathbb{E}|\mathscr{C}_1|^2 \ge (1+o(1))b^2\varepsilon^2 V/(\log(\varepsilon^3 V))^6 \gg V^{1/3},$$

by our choice of ε. We use our bounds on $\mathbf{p}^t(0,0)$ above and sum (7.20) separately for $t \le C\log V$ and $t \ge C\log V$. For $t \le C\log V$ we have $(1+\varepsilon)^t - 1 = O(\varepsilon t)$ and use our first bound $\mathbf{p}^t(0,0) \le V^{-q}$ to get

$$\varepsilon^{-1} r \sum_{t=1}^{C\log V} [(1+\varepsilon)^{t\wedge r} - 1]\mathbf{p}^t(0,0) \le r \sum_{t=1}^{C\log V} tV^{-q} = o(1).$$

When $t \ge C\log V$ we obtain

$$(1+\varepsilon)^{t\wedge r} - 1 \le (1+\varepsilon)^r = \varepsilon^3 V(\log(\varepsilon^3 V))^{-3} = O(\varepsilon^3 V(\log V)^{-3}),$$

by our choice of ε. We use our second bound $\mathbf{p}^t(0,0) \le CV^{-1}$ to deduce that

$$\varepsilon^{-1} r \sum_{t=C\log V}^{2r} [(1+\varepsilon)^{t\wedge r} - 1]\mathbf{p}^t(0,0) = O(r^2\varepsilon^2(\log V)^{-3}) = o(1),$$

since $r \le C(\log V)^2$. This concludes the verification of condition (2) of Theorem 1.3.

To verify condition (3) we need to prove the bound

$$\sum_{u,v} \sum_{t_1,t_2,t_3:\, t_1+t_2+t_3\geq 3}^{C\log V} \mathbf{p}^{t_1}(x,u)\mathbf{p}^{t_2}(u,v)\mathbf{p}^{t_3}(v,y) = O((\log V)^{-2}). \tag{7.21}$$

We first handle the special case of $(t_1,t_2,t_3) = (1,1,1)$. An immediate calculation with Lemma 3.11 shows that (on any regular graph of degree m)

$$\sum_{u,v} \mathbf{p}^1(x,u)\mathbf{p}^1(u,v)\mathbf{p}^1(v,y) = O(1/m^2).$$

In all other cases of (t_1,t_2,t_3) we use our bound on $\mathbf{p}^{t_i}(x,y)$ for $i \in 1,2,3$ such that t_i is the largest of t_1,t_2,t_3 (which must be at least 2). We pull this bound out of the sum, and sum the other two terms over u and v to get a multiplicative contribution of precisely 1. The sum over (t_1,t_2,t_3) such that $3 \leq t_1+t_2+t_3 < 15$ is bounded by $C(\log V)^{-2}$, since the number of such triplets is bounded, and each contributes at most $C(\log V)^{-2}$ because one of the t_i's is at least 2, so that our bounds on $\mathbf{p}^t(x,y)$ guarantee that for this t_i we have $\mathbf{p}^{t_i}(\cdot,\cdot) \leq O(1/m^2) \leq O(1/(\log V)^2)$ by the assumption that $m \geq c\log V$. Similarly, the sum over triplets (t_1,t_2,t_3) such that $t_1+t_2+t_3 \geq 15$ and $t_i \leq t_{\mathrm{mix}}$ is also bounded by $C(\log V)^{-2}$ since the number of such triplets is at most $C(\log V)^3$, and each contributes at most $C(\log V)^{-5}$ because at least one of the t_i's is at least 5 and for this t_i we have $\mathbf{p}^{t_i}(\cdot,\cdot) \leq C(\log V)^{-5}$, again by our assumption that $m \geq c\log V$. This completes the verification of conditions (2) and (3) of Theorem 1.3 and concludes the proof. □

8. Open problems

(1) In this paper we prove a law of large numbers for $|\mathscr{C}_1|$ above the critical window for percolation on the hypercube. Show that $|\mathscr{C}_1|$ satisfies a central limit theorem in this regime. In $G(n,p)$, this and much more was established by Pittel and Wormald [52].

(2) Show that $|\mathscr{C}_2| = (2+o(1))\varepsilon^{-2}\log(\varepsilon^3 2^m)$ when $p = p_c(1+\varepsilon)$ and that $|\mathscr{C}_1| = (2+o(1))\varepsilon^{-2}\log(\varepsilon^3 2^m)$ when $p = p_c(1-\varepsilon)$ for $\varepsilon \gg V^{-1/3}$ and $\varepsilon = o(1)$. This is the content of [18, Conjectures 3.1 and 3.3]. In [15] this is proved for $\varepsilon \geq 60(\log n)^3/n$ in the supercritical regime, and for $\varepsilon \geq (\log n)^2/(n^{1/2}\log\log n)$ in the subcritical regime. In $G(n,p)$, these results are proved in [52] and [39, Theorem 5.6].

(3) Show that $(|\mathscr{C}_j|2^{-2m/3})_{j\geq 1}$ converges in distribution when $p = p_c(1+t2^{-m/3})$ and $t \in \mathbb{R}$ is fixed and identify the limit distribution. Up to a time change, this should be the limiting distribution of $(|\mathscr{C}_j|n^{-2/3})_{j\geq 1}$ in $G(n,p)$ with $p = (1+tn^{-1/3})/n$ identified by Aldous [4].

(4) Consider percolation on the nearest-neighbor torus $\mathbb{Z}_n^d$ where d is a large fixed constant and $n \to \infty$ with $p = p_c(1+\varepsilon)$ such that $\varepsilon \gg n^{-d/3}$ and $\varepsilon = o(1)$. Show that $|\mathscr{C}_1|/(\varepsilon n^d)$ converges to a constant. Does this constant equal the limit as $\varepsilon \downarrow 0$ of $\varepsilon^{-1}\theta_{\mathbb{Z}^d}(p_c(1+\varepsilon))$? Here $\theta_{\mathbb{Z}^d}(p)$ denotes the probability that the cluster of the origin is infinite at p-bond percolation on the infinite lattice $\mathbb{Z}^d$. The techniques of this paper

are not sufficient to show this, mainly because condition (2) of Theorem 1.3 does not hold in $\mathbb{Z}_n^d$ (in fact, it is easy to see that $p_c - (2d-1)^{-1} \geq c > 0$ for some positive constant $c = c(d)$—this is always the case when our underlying transitive graph has constant degree and short cycles). The critical regime of this graph is well understood thanks to [16, 17, 29, 30].

(5) Show that the finite triangle condition (1.5) holds on any family of expander graphs.

(6) Let $\delta > 0$ be a fixed constant and consider the giant component $\mathscr{C}_1$ obtained by performing percolation on the hypercube with $p = (1+\delta)/m$. Show that whp the mixing time of the simple random walk on $\mathscr{C}_1$ is polynomial in m. Is this mixing time of order m^2? This is what one expects by the analogous question on $G(n, p)$ (see [10, 26]). Further analogy with the near-critical $G(n, p)$ (see [22]) suggests that whp the mixing time on $\mathscr{C}_1$ when $p = p_c(1+\varepsilon)$ with the usual condition that $\varepsilon \gg 2^{-m/3}$ and $\varepsilon = o(1)$ is of order $\varepsilon^{-3}\log(\varepsilon^3 2^m)$.

Appendix. Asymptotics of the supercritical cluster tail

Our goal in this section is to prove Theorem 2.2. In [16], Theorem 2.2 is proved without the precise constant 2. Here we sharpen this proof to get this constant. We assume that G is a general transitive graph having degree m and volume V satisfying the finite triangle condition (1.5). In order to stay close to the notation in [16], we define

$$\nabla_p^{\max} = \sup_{x \neq y} \nabla_p(x, y)$$

and

$$\tau_p(x) = \mathbf{P}_p(0 \leftrightarrow x). \tag{A.1}$$

Proposition A.1 (Upper bound on the cluster tail). *Let G be a finite transitive graph of degree m on V vertices such that the finite triangle condition* (1.5) *holds, and set $p = p_c(1+\varepsilon)$ where $\varepsilon = o(1)$ and $\varepsilon \gg V^{-1/3}$. Then, for every $k = k_\varepsilon$ satisfying $k_\varepsilon \geq \varepsilon^{-2}$,*

$$\mathbf{P}_p(|\mathscr{C}(0)| \geq k) \leq 2\varepsilon\big[1 + O\big(\varepsilon + (\varepsilon^3 V)^{-1} + (\varepsilon^2 k)^{-1/4} + \alpha_G\big)\big]. \tag{A.2}$$

Proposition A.2 (Lower bound on the cluster tail). *Let G be a finite transitive graph of degree m on V vertices such that* (1.5) *holds, and set $p = p_c(1+\varepsilon)$ where $\varepsilon = o(1)$ and $\varepsilon \gg V^{-1/3}$. Then, for every $\alpha \in (0, 1/3)$, there exists a $c = c(\alpha) > 0$ such that*

$$\mathbf{P}_p\big(|\mathscr{C}(0)| \geq \varepsilon^{-2}(\varepsilon^3 V)^{\alpha}\big) \geq 2\varepsilon\big[1 + O\big(\varepsilon + (\varepsilon^3 V)^{-c} + \alpha_G\big)\big]. \tag{A.3}$$

Remark. The above propositions also apply to infinite transitive graphs (where $(\varepsilon^3 V)^{-c}$ is replaced by 0), assuming that (1.5) holds with $\chi(p)^3/V$ replaced by 0.

Proof of Theorem 2.2. Follows immediately from the above propositions. □

A.1. Differential inequalities. We follow [16, Section 5]. For $p, \gamma \in [0, 1]$, we define the *magnetization* by

$$M(p, \gamma) = \sum_{k=1}^{V} [1 - (1 - \gamma)^k] \mathbf{P}_p(|\mathscr{C}(0)| = k). \tag{A.4}$$

For fixed p, the function $\gamma \mapsto M(p, \gamma)$ is strictly increasing, with $M(p, 0) = 0$ and $M(p, 1) = 1$. When we color all vertices independently *green* with probability γ, and we let $\mathscr{G}$ denote the set of green vertices, then (A.4) has the appealing probabilistic interpretation of

$$M(p, \gamma) = \mathbf{P}_{p,\gamma}(0 \leftrightarrow \mathscr{G}), \tag{A.5}$$

where $\mathbf{P}_{p,\gamma}$ is the probability measure of the joint bond and site percolation model, where bonds and sites have an independent status. This representation is important for the derivation of useful differential inequalities involving the magnetization.

Lemma A.3 (Differential inequalities for the magnetization). *Let G be a finite transitive graph on V vertices and degree m. Then for any $p, \gamma \in (0, 1)$,*

$$(1 - p)\frac{\partial M}{\partial p} \le m(1 - \gamma)M \frac{\partial M}{\partial \gamma}, \tag{A.6}$$

$$M \le \gamma \frac{\partial M}{\partial \gamma} + \big[\tfrac{1}{2} mpM^2 + \gamma M\big] + \big[\tfrac{1}{2} mpM + \gamma\big] p \frac{\partial M}{\partial p}, \tag{A.7}$$

$$M \ge mp\big[\gamma + (1 - \gamma)\tfrac{1}{2} m(m - 1)p^2 \alpha(p) M^2\big] \frac{\partial M}{\partial \gamma}, \tag{A.8}$$

where

$$\alpha(p) = (1 - 2p)^2 - (1 + mp + 2(mp)^2)\nabla_p^{\max} - mpM - (mp)^2 M^2. \tag{A.9}$$

The inequality (A.6) is proved in [1], where it was used to prove the sharpness of the percolation phase transition on $\mathbb{Z}^d$, and was first stated in the context of finite graphs in [16, (5.14)]. The differential inequality in (A.7) is an adaptation of another differential inequality proved and used in [1], which is improved here in order to obtain sharp constants in our bounds. The bound in (A.8) is an adaptation of [16, (5.16)], which was used there to prove an upper bound on $M(p, \gamma)$. Again, the inequality is adapted in order to obtain the optimal constants. We will first use Lemma A.3 to obtain Propositions A.1 and A.2.

A.2. The magnetization for subcritical p. We take $p = p_c(1 - \varepsilon)$ with $\varepsilon = o(1)$ and $\varepsilon^3 V \gg 1$, and we take $\gamma = o(1)$. Then [16, Lemma 5.3] shows that $M(p, \gamma) = O(\sqrt{\gamma})$. The main aim of this section is to improve upon this bound, using the improved differential inequality in (A.8).

We have $M(p, \gamma) = O(\sqrt{\gamma})$ and $\chi(p) = O(1/\varepsilon)$ by [16, Theorem 1.5]. Furthermore, assumption (1.5) gives $\nabla_p^{\max} = O(\alpha_G + (\varepsilon^3 V)^{-1})$, and [16, (1.30)] then implies that $mp \le 1 + O(\alpha_G)$. Putting all this into (A.9) yields

$$\alpha(p) \ge 1 + O\big(\sqrt{\gamma} + (\varepsilon^3 V)^{-1} + \alpha_G\big). \tag{A.10}$$

Substituting (A.10) into (A.8) in turn gives

$$M \geq \left[1 + O\left(\sqrt{\gamma} + (\varepsilon^3 V)^{-1} + \alpha_G\right)\right]\left[\gamma + \tfrac{1}{2}M^2\right]\frac{\partial M}{\partial \gamma}. \tag{A.11}$$

We now use this to prove the following lemma:

Lemma A.4 (Upper bound on the slightly subcritical magnetization). *Let G be a finite transitive graph of degree m on V vertices such that (1.5) holds. Let $\gamma = o(1)$ and set $p = p_c(1-\varepsilon)$ with $\varepsilon = o(1)$ and $\varepsilon^3 V \gg 1$. Then*

$$M(p, \gamma) \leq \sqrt{2\gamma}\left[1 + O\left(\sqrt{\gamma} + (\varepsilon^3 V)^{-1} + \alpha_G\right)\right]. \tag{A.12}$$

A similar bound was proved in [16, Lemma 5.3], whose proof we adapt here, with $\sqrt{2\gamma}$ replaced with $\sqrt{12\gamma}$, and a less precise error bound. The precise constant $\sqrt{2}$ is important for us here as it relates to the constant 2 for the $2\varepsilon(1 + o(1))$ survival probability.

Proof of Lemma A.4. We note that (A.11) implies that

$$M \geq \frac{B}{2} M^2 \frac{\partial M}{\partial \gamma}, \tag{A.13}$$

where we abbreviate $B = 1 + O(\sqrt{\gamma} + (\varepsilon^3 V)^{-1} + \alpha_G)$. Therefore,

$$\frac{\partial [M^2]}{\partial \gamma} \leq 4/B. \tag{A.14}$$

Integrating between 0 and γ, and using $M(p, 0) = 0$, yields

$$M^2 \leq 4\gamma/B, \tag{A.15}$$

so that $M \leq \sqrt{\gamma}(2/\sqrt{B})$. Now, when we have this inequality, we can further bound

$$\gamma \geq (B/4)M^2, \tag{A.16}$$

so that by (A.11) we get

$$M \geq B[1/2 + B/4]M^2 \frac{\partial M}{\partial \gamma}. \tag{A.17}$$

Performing the same integration steps, we arrive at

$$M^2 \leq \frac{2}{B/2 + B^2/4}\gamma. \tag{A.18}$$

Therefore, the constant has become a little better (recall that B is close to 1). Iterating these steps yields, for every $k \geq 1$,

$$M^2 \leq \frac{2}{\sum_{j=1}^{k} (B/2)^j}\gamma. \tag{A.19}$$

We prove (A.19) by induction on k, the initialization for $k = 1, 2$ having been proved above. Suppose that (A.19) holds for some $k \geq 1$. Define

$$A_k = \sum_{j=1}^{k} (B/2)^j, \tag{A.20}$$

so that (A.19) is equivalent to $M^2 \leq 2\gamma/A_k$. In turn, this yields $\gamma \geq A_k M^2/2$, so that

$$M \geq B[A_k/2 + 1/2] M^2 \frac{\partial M}{\partial \gamma}, \tag{A.21}$$

which in turn yields

$$M^2 \leq \frac{2}{B[1 + A_k]/2}\gamma. \tag{A.22}$$

Note that

$$B[1 + A_k]/2 = A_{k+1}, \tag{A.23}$$

which proves the induction step. By (A.19), we obtain

$$M^2 \leq \frac{2}{\sum_{j=1}^{\infty}(B/2)^j}\gamma = 2[2 - B]\gamma/B. \tag{A.24}$$

Finally, the fact that

$$[2 - B]/B = 1 + O\big(\sqrt{\gamma} + (\varepsilon^3 V)^{-1} + \alpha_G\big) \tag{A.25}$$

completes the proof. □

A.3. The magnetization for supercritical p**.** In this section, we use *extrapolation inequalities* to obtain a bound on the supercritical magnetization from the subcritical one derived in Lemma A.4. Our precise result is the following:

Lemma A.5 (Upper bound on the slightly supercritical magnetization). *Let G be a finite transitive graph of degree m on V vertices such that (1.5) holds, and set $p = p_c(1 + \varepsilon)$ where $\varepsilon = o(1)$ and $\varepsilon \gg V^{-1/3}$. Then, for any $c \in (0, 1/3)$,*

$$M(p, \gamma) \leq \big(\varepsilon + \sqrt{2\gamma + \varepsilon^2}\big)\big[1 + O\big(\varepsilon + \sqrt{\gamma} + (\varepsilon^3 V)^{-c} + \alpha_G\big)\big]. \tag{A.26}$$

Proof. We follow the proof in [16, Section 5.3], paying special attention to the constants and error terms. Indeed, we use (A.6) and the chain rule to deduce that, with $A = (1 - 2p_c)^{-1}$ and $\tilde{M}(p, h) = M(p, 1 - \mathrm{e}^{-h})$,

$$\frac{\partial \tilde{M}}{\partial p} \leq m A \tilde{M} \frac{\partial \tilde{M}}{\partial h}. \tag{A.27}$$

Take $P_1 = (p_c(1 + \varepsilon), h)$ and write $m_1 = \tilde{M}(P_1)$. Further, take $\eta = \varepsilon(\varepsilon^3 V)^{-c}$ for some $c \in (0, 1/3)$, so that $\eta = o(\varepsilon)$ and $\eta^3 V \to \infty$, and let $P_2 = (p_c(1 - \eta), A m_1 \varepsilon')$, where

$$\varepsilon' = \varepsilon + \eta + \frac{h}{A m_1}. \tag{A.28}$$

Then, with $m_2 = \tilde{M}(P_2)$, we have $m_2 \geq m_1$ (see e.g. [16, (5.46)]). Therefore, by Lemma A.4 and again writing $B = 1 + O(\sqrt{\gamma} + (\varepsilon^3 V)^{-1} + \alpha_G)$ with $\gamma = 1 - \mathrm{e}^{-h}$, we obtain

$$\begin{aligned} M(p, 1 - \mathrm{e}^{-h}) = m_1 \leq m_2 &\leq \sqrt{2B(1 - \mathrm{e}^{-Am_1\varepsilon'})} \\ &= (1 + O(m_1\varepsilon'))\sqrt{2ABm_1\varepsilon'} \\ &= (1 + O(m_1\varepsilon))\sqrt{2ABm_1(\varepsilon + \eta) + 2Bh} \\ &= \big(1 + O(\varepsilon + (\varepsilon^3 V)^{-c})\big)\sqrt{2ABm_1\varepsilon + 2Bh}, \end{aligned} \tag{A.29}$$

where in the last inequality we use $\eta = \varepsilon(\varepsilon^3 V)^{-c} \ll \varepsilon$ and $m_1 \leq 1$. The inequality

$$m_1 \leq \sqrt{2ABm_1\varepsilon + 2Bh}$$

has roots

$$m_\pm = AB\varepsilon \pm \sqrt{2Bh + (AB\varepsilon)^2}. \tag{A.30}$$

Since $m_1 \geq 0$ and $m_+ \geq 0$ while $m_- \leq 0$, we deduce that

$$M(p_c{+}\varepsilon/m, 1{-}\mathrm{e}^{-h}) = m_1 \leq \big(1{+}O(\varepsilon{+}(\varepsilon^3 V)^{-c})\big)\big(AB\varepsilon{+}\sqrt{2Bh + (AB\varepsilon)^2}\big). \tag{A.31}$$

We have $\gamma = 1 - \mathrm{e}^{-h} = h(1 + O(h))$ and $A = 1 + O(\alpha_G)$ (by [16, (1.30)]) and $B = 1 + O(\sqrt{\gamma} + (\varepsilon^3 V)^{-1} + \alpha_G)$. Putting all this together in the last inequality completes the proof. □

Proof of Proposition A.1. We note that, for any $l \geq k \geq 1$ and $a > 0$,

$$1 - (1 - a/k)^l \geq 1 - \mathrm{e}^{-a}. \tag{A.32}$$

Therefore, by (A.4),

$$\mathbf{P}_p(|\mathscr{C}(0)| \geq k) \leq [1 - \mathrm{e}^{-a}]^{-1} M(p, a/k). \tag{A.33}$$

Recall that $k \gg \varepsilon^{-2}$ and take $a = (\varepsilon^2 k)^{1/2}$ so that $a/k = \varepsilon^2(\varepsilon^2 k)^{-1/2} = o(\varepsilon^2)$. We note that for $\gamma = \varepsilon^2(\varepsilon^2 k)^{-1/2}$, (A.26) reduces to

$$M(p, \gamma) \leq 2\varepsilon\big[1 + O\big(\varepsilon + (\varepsilon^3 V)^{-1} + (\varepsilon^2 k)^{-1/4} + \alpha_G\big)\big]. \tag{A.34}$$

Then, by (A.34) and the fact that $1 - \mathrm{e}^{-a} = 1 + o((\varepsilon^2 k)^{-1/4})$,

$$M(p, a/k) \leq 2\varepsilon\big[1 + O\big(\varepsilon + (\varepsilon^3 V)^{-1} + (\varepsilon^2 k)^{-1/4} + \alpha_G\big)\big]. \tag{A.35}$$

This completes the proof of Proposition A.1. □

A.4. Lower bound on tail probabilities. In the remainder of this section, we shall prove Proposition A.2. Throughout this proof, we will take $p = p_c(1 + \varepsilon)$.

We shall assume that with $k_0 = \varepsilon^{-2}(\varepsilon^3 V)^\alpha \gg \varepsilon^{-2}$ and $\alpha \in (0, 1/3)$, there exists $b_{10} = b_{10}(\alpha)$ such that

$$\mathbf{P}_p\big(|\mathscr{C}(v)| \geq \varepsilon^{-2}(\varepsilon^3 V)^\alpha\big) \geq b_{10}\varepsilon. \tag{A.36}$$

The bound in (A.36) is proved for finite graphs in [16, Theorem 1.6(i)] and in [9], in conjunction with [28], on infinite lattices satisfying the triangle condition. The proof of (A.36) is similar to the argument we shall give for the improved bound, and will be omitted here. In turn, (A.36) implies that, for $\gamma = 1/k_0 = \varepsilon^2(\varepsilon^3 V)^{-\alpha} = o(\varepsilon^2)$, there exists a constant $\tilde{b}_{10}$ such that

$$M(p, \gamma) \geq [1 - [1 - \gamma]^{k_0}]\mathbf{P}_p(|\mathscr{C}(v)| \geq k_0) \geq \tilde{b}_{10}\varepsilon. \tag{A.37}$$

Inequality (A.37) will be an essential ingredient in our proof. We start by proving the following lemma:

Lemma A.6 (Lower bound on the magnetization). *Let G be a finite transitive graph of degree m on V vertices such that (1.5) holds, and set $p = p_c(1+\varepsilon)$ where $\varepsilon = o(1)$ and $\varepsilon \gg V^{-1/3}$. Then, for $\gamma = \varepsilon^2(\varepsilon^3 V)^{-\alpha}$ with $\alpha \in (0, 1/3)$ and any $c < 1$,*

$$M(p, \gamma) \geq 2\varepsilon\big[1 + O\big(\varepsilon + (\varepsilon^3 V)^{-c} + \alpha_G\big)\big]. \tag{A.38}$$

Proof. Throughout the proof, we fix $\alpha \in (0, 1/3)$. We recall the differential inequality (A.7),

$$M \leq \gamma\frac{\partial M}{\partial \gamma} + \big[\tfrac{1}{2}mpM^2 + \gamma M\big] + \big[\tfrac{1}{2}mpM + \gamma\big]p\frac{\partial M}{\partial p}. \tag{A.39}$$

By (A.37), and the fact that $\gamma \mapsto M(p, \gamma)$ is increasing, for any $\gamma = \varepsilon^2(\varepsilon^3 V)^{-\alpha}$ we have $\gamma = O(M\varepsilon)$. Further, $mp \leq 1 + O(\varepsilon + \alpha_G)$, so that, for some $A > 1$ with $A = 1 + O(\varepsilon + \alpha_G)$,

$$M \leq \gamma\frac{\partial M}{\partial \gamma} + \frac{A}{2}M^2 + \frac{A}{2}Mp\frac{\partial M}{\partial p}. \tag{A.40}$$

We rewrite (A.40) as

$$0 \leq \frac{1}{M}\frac{\partial M}{\partial \gamma} + \frac{1}{\gamma}\frac{\partial}{\partial p}\left[\frac{A}{2}pM - p\right], \tag{A.41}$$

and integrate for $\gamma \in [\gamma_0, \gamma_1]$ and $p \in [p_0, p_1]$, where $\gamma_0 = (\delta\varepsilon)^2(\delta^3\varepsilon^3 V)^{-\alpha}$. We note that (A.37) holds for $p_0 = p_c(1 + \varepsilon\delta)$ for any $\delta = o(1)$ and $\gamma = \gamma_0$. We further take

$$p_0 = p_c(1 + \delta\varepsilon), \qquad p_1 = p_c(1 + \varepsilon), \qquad \gamma_1 = \mathrm{e}^{(\log(1/\delta))^a}\gamma_0, \tag{A.42}$$

where $a > 1$ is chosen below.

Then, as in [27, (5.57) and the argument below it], since $p \mapsto M(p, \gamma)$ and $\gamma \mapsto M(p, \gamma)$ are non-decreasing,

$$0 \leq (p_1 - p_0)\log\frac{M(p_1, \gamma_1)}{M(p_0, \gamma_0)} + \log(\gamma_1/\gamma_0)\left[\frac{A}{2}p_1 M(p_1, \gamma_1) - (p_1 - p_0)\right]. \tag{A.43}$$

Now,

$$\log(\gamma_1/\gamma_0) = (\log(1/\delta))^a, \tag{A.44}$$

while, by Lemma A.5 and (A.37),

$$M(p_1, \gamma_1) \le 2\varepsilon\big[1 + O\big(\varepsilon + (\varepsilon^3 V)^{-1}\big)\big], \qquad M(p_0, \gamma_0) \ge \tilde{b}_{10}\delta\varepsilon, \tag{A.45}$$

so that, for $\delta > 0$ sufficiently small,

$$\log \frac{M(p_1, \gamma_1)}{M(p_0, \gamma_0)} \le \log(2\varepsilon/(\tilde{b}_{10}\delta\varepsilon)) \le 2\log(1/\delta). \tag{A.46}$$

Dividing (A.43) through by $(\log(1/\delta))^a$, we arrive at

$$\frac{A}{2} p_1 M(p_1, \gamma_1) \ge p_c(1-\delta)\varepsilon\big[1 - 2(\log(1/\delta))^{1-a}\big]. \tag{A.47}$$

Recalling that $p_1 = p_c(1+\varepsilon)$ and $a > 1$, as well as $A = 1 + O(\varepsilon + \alpha_G)$, this yields

$$M(p, \gamma_1) \ge 2\varepsilon\big[1 + O\big(\varepsilon + (\log(1/\delta))^{1-a} + \alpha_G\big)\big]. \tag{A.48}$$

Finally, note that

$$\begin{aligned}\gamma_1 &= \mathrm{e}^{(\log(1/\delta))^a}\gamma_0 = \mathrm{e}^{(\log(1/\delta))^a}(\delta\varepsilon)^2(\delta^3\varepsilon^3 V)^{-\alpha} \\ &= \varepsilon^2(\varepsilon^3 V)^{-\alpha}\big(\mathrm{e}^{(\log(1/\delta))^a}\delta^{2-3\alpha}\big) \ge \varepsilon^2(\varepsilon^3 V)^{-\alpha},\end{aligned} \tag{A.49}$$

when we take $\delta = \mathrm{e}^{-(\varepsilon^3 V)^{1/a}}$ for any $a > 1$. Indeed, then $\mathrm{e}^{(\log(1/\delta))^a}\delta^{2-3\alpha} \to \infty$ as $\delta \to 0$. Since $\gamma \mapsto M(p, \gamma)$ is increasing, this implies that

$$\begin{aligned}M(p, \gamma) \ge M(p, \gamma_1) &\ge 2\varepsilon\big[1 + O\big(\varepsilon + (\log(1/\delta))^{1-a} + \alpha_G\big)\big] \\ &= 2\varepsilon\big[1 + O\big(\varepsilon + (\varepsilon^3 V)^{1/a-1} + \alpha_G\big)\big],\end{aligned} \tag{A.50}$$

which proves the claim with $c = 1 - 1/a$. □

Proof of Proposition A.2. We use [16, (6.5)], which states that, for any $0 \le \gamma_0, \gamma_1 \le 1$,

$$\mathbf{P}_p(|\mathscr{C}(0)| \ge k) \ge M(p, \gamma_1) - \frac{\gamma_1}{\gamma_0}\mathrm{e}^{\gamma_0 k} M(p, \gamma_0). \tag{A.51}$$

Now we take $\gamma_1 = \varepsilon^2(\varepsilon^3 V)^{-\alpha'}$ with $\alpha' \in (0, 1/3)$ taken as in Lemma A.6, $\gamma_0 = \varepsilon^2(\varepsilon^3 V)^{-\alpha}$ with $\alpha < \alpha'$, and $k = 1/\gamma_0$. Then $\mathrm{e}^{\gamma_0 k} = \mathrm{e}$, while from Lemma A.5 and $\gamma_0 = o(\varepsilon^2)$ we obtain

$$M(p, \gamma_0) \le 2\varepsilon(1 + o(1)). \tag{A.52}$$

Therefore, by Lemma A.6 and (A.51), taking $c = 1/2$ in Lemma A.6 yields

$$\mathbf{P}_p(|\mathscr{C}(0)| \ge k) \ge 2\varepsilon\big[1 + O\big(\varepsilon + (\varepsilon^3 V)^{-1/2} + \alpha_G\big)\big] - (\varepsilon^3 V)^{\alpha'-\alpha} O(\varepsilon). \tag{A.53}$$

We deduce that

$$\mathbf{P}_p(|\mathscr{C}(0)| \ge \varepsilon^{-2}(\varepsilon^3 V)^{\alpha}) \ge 2\varepsilon\big[1 + O\big(\varepsilon + (\varepsilon^3 V)^{-1/2} + (\varepsilon^3 V)^{\alpha-\alpha'} + \alpha_G\big)\big]. \tag{A.54}$$

This proves the claim in Proposition A.2 with $c = \alpha - \alpha' \in (0, 1/3)$. □

A.5. Derivation of (A.8). We follow the proof in [16, Appendix A.2] as closely as possible, deviating in one essential inequality. Indeed, in [16, (A.23)–(A.32)], it is proved that

$$M(p,\gamma) \geq pm\frac{\partial M}{\partial \gamma}(p,\gamma)\mathbf{P}_{p,\gamma}(0 \Leftrightarrow \mathscr{G}) - X_2 - X_3, \tag{A.55}$$

where X_2 and X_3 are defined in [16, (A.32)] and the event $0 \Leftrightarrow \mathscr{G}$ means that there are $x, y \in \mathscr{G}$ with $x \neq y$ such that $0 \leftrightarrow x$ and $0 \leftrightarrow y$ disjointly. We copy the bounds on X_2 and X_3 in [16, (A.46)] and [16, (A.53)] respectively, which prove that

$$X_2 \leq p^2 m M(p,\gamma)^2 \frac{\partial M}{\partial \gamma}(p,\gamma), \qquad X_3 \leq \nabla_p^{\max} pmM(p,\gamma)^2 \frac{\partial M}{\partial \gamma}(p,\gamma), \tag{A.56}$$

and we improve upon the lower bound on $\mathbf{P}_{p,\gamma}(0 \Leftrightarrow \mathscr{G})$ only. Our precise result is contained in the following lemma:

Lemma A.7 (Improved lower bound on the double connection). *For all $p, \gamma \in [0,1]$,*

$$\mathbf{P}_{p,\gamma}(0 \Leftrightarrow \mathscr{G}) \geq \gamma + (1-\gamma)\tfrac{1}{2}m(m-1)p^2\alpha(p)M(p,\gamma)^2, \tag{A.57}$$

where

$$\alpha(p) = (1-2p)^2 - (1 + mp + 2(mp)^2)\nabla_p^{\max} - mpM(p,\gamma) - (mp)^2 M(p,\gamma)^2.$$

Proof. Note that if $0 \in \mathscr{G}$, then $0 \Leftrightarrow \mathscr{G}$ occurs. Therefore,

$$\mathbf{P}_{p,\gamma}(0 \Leftrightarrow \mathscr{G}) = \gamma + (1-\gamma)\mathbf{P}_{p,\gamma}(0 \Leftrightarrow \mathscr{G} \mid 0 \notin \mathscr{G}). \tag{A.58}$$

Thus, it remains to obtain a lower bound on $\mathbf{P}_{p,\gamma}(0 \Leftrightarrow \mathscr{G} \mid 0 \notin \mathscr{G})$. For this, we follow the original argument in [16, Section A.2], adapting it when necessary.

For a directed bond $b = (x,y)$, we write $\underline{b} = x$ and $\overline{b} = y$ for its top and bottom. Let e, f be two distinct bonds with $\underline{e} = \underline{f} = 0$, and let $E_{e,f}$ be the event that the bonds e and f are occupied, and that in the reduced graph $G^- = (V^-, E^-)$ obtained by removing the bonds e and f, the following three events occur: $\overline{e} \leftrightarrow \mathscr{G}$, $\overline{f} \leftrightarrow \mathscr{G}$, and $\mathscr{C}(\overline{e}) \cap \mathscr{C}(\overline{f}) = \emptyset$.

Let $\mathbf{P}^-_{p,\gamma}$ denote the joint bond/vertex measure on G^-. We note that the event $0 \Leftrightarrow \mathscr{G}$ contains the event $\bigcup_{e,f} E_{e,f}$, where the (non-disjoint) union is over unordered pairs of bonds e, f incident to the origin. Then, by Bonferroni's inequality and since $E_{e,f}$ is independent of $0 \notin \mathscr{G}$, we get

$$\begin{aligned}\mathbf{P}_{p,\gamma}(0 \Leftrightarrow \mathscr{G} \mid 0 \notin \mathscr{G}) &\geq \mathbf{P}_{p,\gamma}\Big(\bigcup_{e,f} E_{e,f} \;\Big|\; 0 \notin \mathscr{G}\Big) \geq \sum_{\{e,f\}} \mathbf{P}_{p,\gamma}(E_{e,f}) - Y_1 \\ &= p^2 \sum_{e,f} \mathbf{P}^-_{p,\gamma}(\overline{e} \leftrightarrow \mathscr{G},\ \overline{f} \leftrightarrow \mathscr{G},\ \mathscr{C}(\overline{e}) \cap \mathscr{C}(\overline{f}) = \emptyset) - Y_1,\end{aligned} \tag{A.59}$$

where

$$Y_1 = \tfrac{1}{2} \sum_{\{e_1,f_1\} \neq \{e_2,f_2\}} \mathbf{P}_{p,\gamma}(E_{e_1,f_1} \cap E_{e_2,f_2} \mid 0 \notin \mathscr{G}). \tag{A.60}$$

We first bound Y_1. For this, we note that there are two contributions to Y_1, depending on the number of distinct elements in $\{e_1, f_1, e_2, f_2\}$, which can be 3 or 4, and whose contributions we denote by $Y_{1,3}$ and $Y_{1,4}$, respectively.

We start by bounding $Y_{1,3}$. The number of pairs of pairs of edges $\{e_1, f_1\} \neq \{e_2, f_2\}$ such that $|\{e_1, f_1, e_2, f_2\}| = 3$ is $m(m-1)(m-2)$. For such a pair, let x_1, x_2, x_3 denote the distinct elements of $\{\bar{e}_1, \bar{f}_1, \bar{e}_2, \bar{f}_2\}$ such that x_1 corresponds to the end of the edge that appears twice in $\{e_1, f_1, e_2, f_2\}$. If $E_{e_1,f_1} \cap E_{e_2,f_2}$ occurs, then either

$$\{(0, x_1) \text{ occ.}\} \circ \{(0, x_2) \text{ occ.}\} \circ \{(0, x_3) \text{ occ.}\} \circ \{x_1 \leftrightarrow \mathscr{G}\} \circ \{x_2 \leftrightarrow \mathscr{G}\} \circ \{x_3 \leftrightarrow \mathscr{G}\}$$

occurs, or there exists a z such that

$$\{(0, x_1) \text{ occ.}\} \circ \{(0, x_2) \text{ occ.}\} \circ \{(0, x_3) \text{ occ.}\} \circ \{x_1 \leftrightarrow \mathscr{G}\} \circ \{x_2 \leftrightarrow z\} \circ \{x_3 \leftrightarrow z\} \circ \{z \leftrightarrow \mathscr{G}\}$$

occurs. Therefore,

$$Y_{1,3} \leq (1-\gamma)\tfrac{1}{2}m(m-1)(m-2)p^3 M(p,\gamma)^2[M(p,\gamma) + \nabla_p^{\max}], \tag{A.61}$$

where we have estimated

$$\sum_z \mathbf{P}_p(x_2 \leftrightarrow z)\mathbf{P}_p(x_3 \leftrightarrow z) \leq \nabla_p^{\max},$$

which is wasteful, but sufficient for our purposes.

For $Y_{1,4}$, we sum over $\{e_1, f_1\} \neq \{e_2, f_2\}$ with the constraint that all these edges are distinct. The number of such pairs of pairs is $m(m-1)(m-2)(m-3)/4$. Then, a similar computation to that for $Y_{1,3}$ yields

$$Y_{1,4} \leq (1-\gamma)\tfrac{1}{8}m(m-1)(m-2)(m-3)p^4 M(p,\gamma)^2[M(p,\gamma)^2 + 8\nabla_p^{\max}]. \tag{A.62}$$

We continue to estimate the sum over $\{e, f\}$ in (A.59) from below. Let

$$W = W_{e,f} = \{\bar{e} \leftrightarrow \mathscr{G},\ \bar{f} \leftrightarrow \mathscr{G},\ \mathscr{C}(\bar{e}) \cap \mathscr{C}(\bar{f}) = \varnothing\} \tag{A.63}$$

denote the event whose probability appears on the right side of (A.59). Conditioning on the set $\mathscr{C}(\bar{e}) = A \subset V^-$, we see that

$$\mathbf{P}^-_{p,\gamma}(W) = \sum_{A:\, \bar{f} \notin A} \mathbf{P}^-_{p,\gamma}\big(\mathscr{C}(\bar{e}) = A,\ \bar{e} \leftrightarrow \mathscr{G},\ \bar{f} \leftrightarrow \mathscr{G},\ \mathscr{C}(\bar{e}) \cap \mathscr{C}(\bar{f}) = \varnothing\big). \tag{A.64}$$

This can be rewritten as

$$\mathbf{P}^-_{p,\gamma}(W) = \sum_{A:\, \bar{f} \notin A} \mathbf{P}^-_{p,\gamma}\big(\mathscr{C}(\bar{e}) = A,\ \bar{e} \leftrightarrow \mathscr{G},\ \bar{f} \leftrightarrow \mathscr{G} \text{ in } V^- \setminus A\big), \tag{A.65}$$

where $\{\bar{f} \leftrightarrow \mathscr{G} \text{ in } V^- \setminus A\}$ is the event that there exists $x \in \mathscr{G}$ such that $\bar{f} \leftrightarrow x$ in $V^- \setminus A$. The intersection of the first two events on the right hand side of (A.65) is independent of the third event, and hence

$$\mathbf{P}^-_{p,\gamma}(W) = \sum_{A:\, \bar{e} \notin A} \mathbf{P}^-_{p,\gamma}(\mathscr{C}(\bar{e}) = A,\ \bar{e} \leftrightarrow \mathscr{G})\mathbf{P}^-_{p,\gamma}(\bar{f} \leftrightarrow \mathscr{G} \text{ in } V^- \setminus A). \tag{A.66}$$

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

$$\begin{aligned}\mathbf{P}^-_{p,\gamma}(\bar{f} \leftrightarrow \mathscr{G} \text{ in } V^- \setminus A) &= M^-(\bar{f}) - \mathbf{P}^-_{p,\gamma}(\bar{f} \leftrightarrow \mathscr{G} \text{ only on } A) \\ &\geq M^-(\bar{f}) - \sum_{y\in A} \tau_p(\bar{f}, y) M^-(y).\end{aligned} \tag{A.67}$$

By definition and the BK inequality,

$$\begin{aligned}M^-(x) &= M(p,\gamma) - \mathbf{P}_{p,\gamma}(e \text{ or } f \text{ is occ. and piv. for } x \leftrightarrow \mathscr{G}) \\ &\geq M(p,\gamma)(1-2p).\end{aligned} \tag{A.68}$$

It follows from (A.66)–(A.68) and the upper bound $M^-(x) \leq M(p,\gamma)$ that

$$\begin{aligned}\mathbf{P}^-_{p,\gamma}(W) &\geq M(p,\gamma) \sum_{A:\bar{e}\notin A} \mathbf{P}^-_{p,\gamma}(\mathscr{C}(\bar{e}) = A,\, \bar{e} \leftrightarrow \mathscr{G})\Big[(1-2p) - \sum_{y\in A} \tau_p(\bar{f}, y)\Big] \\ &= M(p,\gamma)\Big[M^-(\bar{e})(1-2p) - \sum_{y\in V^-} \tau_p(\bar{f}, y)\mathbf{P}^-_{p,\gamma}(\bar{e} \leftrightarrow y,\, \bar{e} \leftrightarrow \mathscr{G})\Big].\end{aligned} \tag{A.69}$$

It is not difficult to show, using the BK inequality, that

$$\mathbf{P}^-_{p,\gamma}(\bar{e} \leftrightarrow y,\; \bar{e} \leftrightarrow \mathscr{G}) \leq \sum_{w\in V^-} \tau_p(\bar{e}, w)\tau_p(w, y) M^-(w), \tag{A.70}$$

and hence, by (A.68)–(A.5),

$$\begin{aligned}\mathbf{P}^-_{p,\gamma}(W) &\geq M(p,\gamma)\Big[M^-(\bar{e})(1-2p) - \sum_{y,w\in V^-} \tau_p(\bar{f}, y)\tau_p(\bar{e}, w)\tau_p(w, y) M^-(w)\Big] \\ &\geq M(p,\gamma)^2[(1-2p)^2 - \nabla_p^{\max}].\end{aligned}$$

This completes the proof of (A.57). □

A.6. Derivation of (A.7). In this section, we prove (A.7), which is an adaptation of the proof of the related inequality

$$M \leq \gamma \frac{\partial M}{\partial \gamma} + M^2 + pM\frac{\partial M}{\partial p}, \tag{A.71}$$

proved in [1] (see also [27, Lemma (5.53)]). The main difference between (A.71) and (A.7) is in the precise constants. Indeed, we have $pm \approx 1$ and $M \gg \gamma$, so that (A.7) is morally equivalent to

$$M \leq \gamma \frac{\partial M}{\partial \gamma} + \frac{1}{2}M^2 + \frac{1}{2}pM\frac{\partial M}{\partial p}, \tag{A.72}$$

i.e., in the inequality of (A.71) the last two terms are multiplied by $1/2$.

We follow the proof of [27, Lemma (5.53)] as closely as possible, deviating only when necessary. Indeed,

$$\begin{aligned}M(p,\gamma) &= \mathbf{P}_{p,\gamma}(\mathscr{C}(0) \cap \mathscr{G} \neq \emptyset) \\ &= \mathbf{P}_{p,\gamma}(|\mathscr{C}(0) \cap \mathscr{G}| = 1) + \mathbf{P}_{p,\gamma}(|\mathscr{C}(0) \cap \mathscr{G}| \geq 2).\end{aligned} \tag{A.73}$$

The first term on the right hand side of (A.73) equals $\gamma\frac{\partial M}{\partial\gamma}$, as derived in [27, (5.69)]. For the second term, we define A_x to be the event that either $x \in \mathscr{G}$, or x is connected by an occupied path to a vertex $g \in \mathscr{G}$. Then

$$\begin{aligned}\mathbf{P}_{p,\gamma}(|\mathscr{C}(0)\cap\mathscr{G}|\geq 2) &= \mathbf{P}_{p,\gamma}(A_0\circ A_0)\\ &\quad+\mathbf{P}_{p,\gamma}(|\mathscr{C}(0)\cap\mathscr{G}|\geq 2,\ A_0\circ A_0 \text{ does not occur}).\end{aligned}\tag{A.74}$$

In the derivation of (A.71), we simply apply the BK inequality to obtain

$$\mathbf{P}_{p,\gamma}(A_0\circ A_0)\leq \mathbf{P}_{p,\gamma}(A_0)^2 = M(p,\gamma)^2,\tag{A.75}$$

leading to the second term in (A.71). Instead, we split depending on whether $0 \in \mathscr{G}$ or not. If $0\in\mathscr{G}$, then $0\in\mathscr{G}$ occurs disjointly from A_0, so that the BK inequality yields

$$\mathbf{P}_{p,\gamma}(A_0\circ A_0, 0\in\mathscr{G})\leq \mathbf{P}_{p,\gamma}(A_0\circ\{0\in\mathscr{G}\})\leq\gamma\mathbf{P}(A_0)=\gamma M(p,\gamma).\tag{A.76}$$

When, instead, $0\notin\mathscr{G}$, there must be *at least two* neighbors e of the origin for which the event

$$A_e\circ A_0\circ\{(0,e)\text{ occ.}\}\tag{A.77}$$

occurs. Therefore, we can bound, with N denoting the number of neighbors e for which the event in (A.77) occurs, so that $N\geq 2$ a.s. and Markov's inequality yields

$$\begin{aligned}\mathbf{P}_{p,\gamma}(A_0\circ A_0, 0\notin\mathscr{G}) &\leq \sum_{e\sim 0}\mathbb{E}_{p,\gamma}\Big[\frac{1}{N}\mathbf{1}_{\{A_e\circ A_0\circ\{(0,e)\text{ occ.}\}}\Big]\\ &\leq \tfrac{1}{2}\sum_{e\sim 0}\mathbf{P}_{p,\gamma}(A_e\circ A_0\circ\{(0,e)\text{ occ.}\}).\end{aligned}\tag{A.78}$$

Therefore, again by the BK inequality,

$$\mathbf{P}_{p,\gamma}(A_0\circ A_0, 0\notin\mathscr{G})\leq\tfrac{1}{2}\sum_{e}\mathbf{P}_{p,\gamma}(A_e)\mathbf{P}_{p,\gamma}(A_0)p=\tfrac{1}{2}pmM(p,\gamma)^2,$$

so that

$$\mathbf{P}_{p,\gamma}(A_0\circ A_0)\leq\tfrac{1}{2}pmM(p,\gamma)^2+\gamma M(p,\gamma),\tag{A.79}$$

which yields the second term in (A.7).

We move on to the bound on the probability of the event that $|\mathscr{C}(0)\cap\mathscr{G}|\geq 2$, but $A_0\circ A_0$ does not occur. This event is equivalent to the existence of an edge $b=(x,y)$ for which the following occurs:

(i) the edge b is occupied; and
(ii) in the subgraph of G obtained by deleting b, the following events occur:

  (a) no vertex of $\mathscr{G}$ is joined to the origin by an open path;
  (b) x is joined to 0 by an occupied path;
  (c) the event $A_y\circ A_y$ occurs.

The events in (ii) are independent of the occupation status of the edge $b=(x,y)$ so that

$$\begin{aligned}
&\mathbf{P}_{p,\gamma}\big(|\mathscr{C}(0)\cap\mathscr{G}|\geq 2,\ A_0\circ A_0 \text{ does not occur}\big)\\
&\qquad=\frac{p}{1-p}\sum_{x\sim y}\mathbf{P}_{p,\gamma}\big((x,y)\text{ closed}, x\in\mathscr{C}(0),\mathscr{C}(0)\cap\mathscr{G}=\emptyset, A_y\circ A_y\big)\\
&\qquad\leq\frac{p}{1-p}\sum_{x\sim y}\mathbf{P}_{p,\gamma}\big(x\in\mathscr{C}(0),\mathscr{C}(0)\cap\mathscr{G}=\emptyset, A_y\circ A_y\big),
\end{aligned}\tag{A.80}$$

where we write $x\sim y$ to denote that (x,y) is a bond. We condition on $\mathscr{C}(0)$ to obtain

$$\begin{aligned}
&\mathbf{P}_{p,\gamma}(\mathscr{C}(0)\cap\mathscr{G}=\emptyset, A_y\circ A_y)\\
&\ =\frac{p}{1-p}\sum_{x\sim y}\sum_{A}\mathbf{P}_p(\mathscr{C}(0)=A)\mathbf{P}_{p,\gamma}\big(\mathscr{C}(0)\cap\mathscr{G}=\emptyset,\ A_y\circ A_y\ \big|\ \mathscr{C}(0)=A\big),
\end{aligned}\tag{A.81}$$

where the sum over A is over all sets of vertices which contain 0 and x but not y. Conditionally on $\mathscr{C}(0)=A$, the events $\mathscr{C}(0)\cap\mathscr{G}=\emptyset$ and $A_y\circ A_y$ are *independent*, since $\mathscr{C}(0)\cap\mathscr{G}=\emptyset$ is defined on the vertices in A, while $A_y\circ A_y$ depends on the vertices in A^c and the edges between them. Thus,

$$\begin{aligned}
&\mathbf{P}_{p,\gamma}\big(\mathscr{C}(0)\cap\mathscr{G}=\emptyset, A_y\circ A_y\ \big|\ \mathscr{C}(0)=A\big)\\
&\qquad\qquad=\mathbf{P}_{p,\gamma}(\mathscr{C}(0)\cap\mathscr{G}=\emptyset\mid\mathscr{C}(0)=A)\mathbf{P}_{p,\gamma}(A_y\circ A_y\text{ off }A),
\end{aligned}\tag{A.82}$$

where we write $\{A_y\circ A_y \text{ off } A\}$ for the event that $A_y\circ A_y$ occurs in the graph where all edges with at least one endpoint in the set A are removed. So far, the derivation follows the proof of [27, Lemma (5.53)]. Now we shall deviate from it. We split, depending on whether $y\in\mathscr{G}$ or not, to obtain

$$\mathbf{P}_{p,\gamma}(A_y\circ A_y\text{ off }A)=\mathbf{P}_{p,\gamma}(A_y\circ A_y\text{ off }A, y\in\mathscr{G})+\mathbf{P}_{p,\gamma}(A_y\circ A_y\text{ off }A, y\notin\mathscr{G}).$$

When $y\in\mathscr{G}$,

$$\{A_y\circ A_y\text{ off }A, y\in\mathscr{G}\}=\big\{(\{y\in\mathscr{G}\}\circ A_y)\text{ off }A\big\},\tag{A.83}$$

so that, by the BK inequality,

$$\mathbf{P}_{p,\gamma}(A_y\circ A_y\text{ off }A,\ y\in\mathscr{G})\leq\gamma\mathbf{P}_{p,\gamma}(A_y\text{ off }A).\tag{A.84}$$

As a result,

$$\begin{aligned}
&\frac{p}{1-p}\sum_{x\sim y}\mathbf{P}_{p,\gamma}\big(x\in\mathscr{C}(0),\ \mathscr{C}(0)\cap\mathscr{G}=\emptyset,\ A_y\circ A_y, y\in\mathscr{G}\big)\\
&\quad\leq\frac{\gamma p}{1-p}\sum_{x\sim y}\sum_{A}\mathbf{P}_p(\mathscr{C}(0)=A)\mathbf{P}_{p,\gamma}\big(\mathscr{C}(0)\cap\mathscr{G}=\emptyset\ \big|\ \mathscr{C}(0)=A\big)\mathbf{P}_{p,\gamma}(A_y\text{ off }A)\\
&\quad=\frac{\gamma p}{1-p}\sum_{x\sim y}\sum_{A}\mathbf{P}_p(\mathscr{C}(0)=A)\mathbf{P}_{p,\gamma}\big(\mathscr{C}(0)\cap\mathscr{G}=\emptyset,\ A_y\ \big|\ \mathscr{C}(0)=A\big)\\
&\quad=\frac{\gamma p}{1-p}\sum_{x\sim y}\mathbf{P}_{p,\gamma}\big(x\in\mathscr{C}(0),\mathscr{C}(0)\cap\mathscr{G}=\emptyset,\ \mathscr{C}(y)\cap\mathscr{G}\neq\emptyset\big)=\gamma p\frac{\partial M}{\partial p},
\end{aligned}\tag{A.85}$$

where the first equality follows again by conditional independence, and the last equality from the fact that (see [27, (5.67)])

$$(1-p)\frac{\partial M}{\partial p} = \sum_{x\sim y} \mathbf{P}_{p,\gamma}\big(x \in \mathscr{C}(0),\ \mathscr{C}(0)\cap\mathscr{G}=\emptyset,\ \mathscr{C}(y)\cap\mathscr{G}\neq\emptyset\big). \tag{A.86}$$

It remains to bound the contribution from $y \notin \mathscr{G}$. For this, we note that if $A_y \circ A_y$ occurs off A and $y \notin \mathscr{G}$, then there must be *at least two* neighbors z of y for which the event

$$\big\{(A_z \circ A_y \circ \{(y,z) \text{ occ.}\}) \text{ off } A\big\} \tag{A.87}$$

occurs. Therefore, by a similar argument to (A.78),

$$\mathbf{P}_{p,\gamma}(A_y \circ A_y \text{ off } A,\ y \notin \mathscr{G}) \le \tfrac{1}{2}\sum_{z\sim y} \mathbf{P}_{p,\gamma}\big((A_z \circ A_y \circ \{(y,z) \text{ occ.}\}) \text{ off } A\big). \tag{A.88}$$

By the BK inequality,

$$\begin{aligned}\mathbf{P}_{p,\gamma}\big((A_z \circ A_y \circ \{(y,z) \text{ occ.}\}) \text{ off } A\big) &\le p\mathbf{P}_{p,\gamma}(A_z \text{ off } A)\mathbf{P}_{p,\gamma}(A_y \text{ off } A)\\ &\le pM(p,\gamma)\mathbf{P}_{p,\gamma}(A_y \text{ off } A).\end{aligned}$$

Repeating the steps in (A.85), we thus arrive at

$$\frac{p}{1-p}\sum_{x\sim y}\mathbf{P}_{p,\gamma}\big(x \in \mathscr{C}(0), \mathscr{C}(0)\cap\mathscr{G}=\emptyset,\ A_y\circ A_y,\ y\notin\mathscr{G}\big) \le \tfrac{1}{2}mp^2M(p,\gamma)\frac{\partial M}{\partial p}. \tag{A.89}$$

Therefore, summing the two bounds in (A.85) and (A.89), we arrive at

$$\mathbf{P}_{p,\gamma}(|\mathscr{C}(0)\cap\mathscr{G}| \ge 2,\ A_0\circ A_0 \text{ does not occur}) \le \big[\tfrac{1}{2}mpM(p,\gamma)+\gamma\big]p\frac{\partial M}{\partial p},$$

which is the third term in (A.7). This completes the proof of (A.7). □

Acknowledgments. The work of RvdH was supported in part by the Netherlands Organisation for Scientific Research (NWO) through VICI grant 639.033.806 and the Gravitation NETWORKS grant 024.002.003. The work of AN was partially supported by NSF grant #6923910. RvdH warmly thanks the Theory Group at Microsoft Research Redmond for the hospitality during his visit in January 2010, where this work was initiated, the Department of Mathematics at M.I.T. for their hospitality during his visit in April 2011, and the Department of Mathematics at U.B.C. for their hospitality during his visit in November 2011. We thank Gady Kozma for useful discussions and Gordon Slade for valuable comments on a preliminary version of this paper. AN thanks Christian Borgs for introducing him to [18, Conjecture 3.2] when he was an intern at Microsoft Research in the summer of 2005.

References

[1] Aizenman, M., Barsky D.: Sharpness of the phase transition in percolation models. Comm. Math. Phys. **108**, 489–526 (1987) Zbl 0618.60098 MR 0874906

[2] Aizenman, M., Newman, C. M.: Tree graph inequalities and critical behavior in percolation models. J. Statist. Phys. **436**, 107–143 (1984) Zbl 0586.60096 MR 0762034

[3] Ajtai, M., Komlós, J., Szemerédi, E.: Largest random component of a k-cube. Combinatorica **2**, 1–7 (1982) Zbl 0489.05053 MR 0671140

[4] Aldous, D.: Brownian excursions, critical random graphs and the multiplicative coalescent. Ann. Probab. **25**, 812–854 (1997) Zbl 0877.60010 MR 1434128

[5] Alon, N., Benjamini, I., Lubetzky, E., Sodin, S.: Non-backtracking random walks mix faster. Comm. Contemp. Math. **9**, 585–603 (2007) Zbl 1140.60301 MR 2348845

[6] Alon, N., Benjamini, I., Stacey A.: Percolation on finite graphs and isoperimetric inequalities. Ann. Probab. **32**, 1727–1745 (2004) Zbl 1046.05071 MR 2073175

[7] Balogh, J., Bollobás, B., Morris, R.: Majority bootstrap percolation on the hypercube. Combin. Probab. Comput. **18**, 17–51 (2009) Zbl 1198.60041 MR 2497373

[8] Balogh, J., Bollobás, B., Morris, R.: Bootstrap percolation in high dimensions. Combin. Probab. Comput. **19**, 643–692 (2010) Zbl 1263.60082 MR 2726074

[9] Barsky, D., Aizenman, M.: Percolation critical exponents under the triangle condition. Ann. Probab. **19**, 1520–1536 (1991) Zbl 0747.60093 MR 1127713

[10] Benjamini, I., Kozma, G., Wormald, N.: The mixing time of the giant component of a random graph. Random Structures Algorithms **45**, 383–407 (2014) Zbl 06370216 MR 3252922

[11] van den Berg, J., Fiebig, U.: On a combinatorial conjecture concerning disjoint occurrences of events. Ann. Probab. **15**, 354–374 (1987) Zbl 0617.60011 MR 0877608

[12] van den Berg, J., Kesten, H.: Inequalities with applications to percolation and reliability. J. Appl. Probab. **22**, 556–569 (1985) Zbl 0571.60019 MR 0799280

[13] Bollobás, B.: The evolution of random graphs. Trans. Amer. Math. Soc. **286**, 257–274 (1984) Zbl 0579.05046 MR 0756039

[14] Bollobás, B.: Random Graphs. Academic Press, London (1985) Zbl 0567.05042 MR 0809996

[15] Bollobás, B., Kohayakawa, Y., Łuczak, T.: The evolution of random subgraphs of the cube. Random Structures Algorithms **3**, 55–90 (1992) Zbl 0779.05045 MR 1139488

[16] Borgs, C., Chayes, J. T., van der Hofstad, R., Slade, G., Spencer, J.: Random subgraphs of finite graphs: I. The scaling window under the triangle condition. Random Structures Algorithms **27**, 137–184 (2005) Zbl 1076.05071 MR 2155704

[17] Borgs, C., Chayes, J. T., van der Hofstad, R., Slade, G., Spencer, J.: Random subgraphs of finite graphs: II. The lace expansion and the triangle condition. Ann. Probab. **33**, 1886–1944 (2005) Zbl 1079.05087 Zbl 1079.05087

[18] Borgs, C., Chayes J. T., van der Hofstad, R., Slade, G., Spencer, J.: Random subgraphs of finite graphs: III. The phase transition for the n-cube. Combinatorica **26**, 395-410 (2006) Zbl 1121.05108 MR 2260845

[19] Borgs, C., Chayes, J. T., Kim, J. H., Wilson, D. B.: The scaling window of the 2-SAT transition. Random Structures Algorithms **18**, 201–256 (2001) Zbl 0979.68053 MR 1824274

[20] Borgs, C., Chayes, J. T., Randall, D.: The van den Berg–Kesten–Reimer inequality: a review. In: Perplexing Problems in Probability, Progr. Probab. 44, Birkhäuser Boston, Boston, MA, 159–173 (1999) Zbl 1115.82320 MR 1703130

[21] Chung, F. R. K., Graham, R. L., Wilson, R. M.: Quasi-random graphs. Combinatorica **9**, 345–362 (1989) Zbl 0715.05057 MR 1054011

[22] Ding, J., Lubetzky, E., Peres, Y.: Mixing time of near-critical random graphs. Ann. Probab. **40**, 979–1008 (2012) Zbl 1243.05217 MR 2962084

[23] Erdős, P., Rényi, A.: On the evolution of random graphs. Magyar Tud. Akad. Mat. Kutató Int. Közl. **5**, 17–61 (1960) Zbl 0103.16301 MR 0125031

[24] Erdős, P., Spencer, J.: Evolution of the n-cube. Comput. Math. Appl. **5**, 33–39 (1979) Zbl 0399.05041 MR 0534014

[25] Fitzner, R., van der Hofstad, R.: Non-backtracking random walk. J. Statist. Phys. **150**, 264–284 (2013) Zbl 1259.82043 MR 3022459

[26] Fountoulakis, N., Reed, B.: The evolution of the mixing rate of a simple random walk on the giant component of a random graph. Random Structures Algorithms **33**, 68–86 (2008) Zbl 1147.60316 MR 2428978

[27] Grimmett, G.: Percolation. 2nd ed., Grundlehren Math. Wiss. 321, Springer, Berlin (1999) MR 1707339

[28] Hara, T., Slade, G.: Mean-field critical behaviour for percolation in high dimensions. Comm. Math. Phys. **128**, 333–391 (1990) Zbl 0698.60100 MR 1043524

[29] Heydenreich, M., van der Hofstad, R.: Random graph asymptotics on high-dimensional tori. Comm. Math. Phys. **270**, 335–358 (2007) Zbl 1128.82010 MR 2276449

[30] Heydenreich, M., van der Hofstad, R.: Random graph asymptotics on high-dimensional tori II. Probab. Theory Related Fields **149**, 397–415 (2011) Zbl 1229.60108 MR 2776620

[31] Heydenreich, M., van der Hofstad, R.: Progress in High-Dimensional Percolation and Random Graphs. CRM Short Courses Ser., Springer (2016), to appear

[32] van der Hofstad, R., Luczak, M.: Random subgraphs of the 2D Hamming graph: the supercritical phase. Probab. Theory Related Fields **47**, 1–41 (2010) Zbl 1188.05141 MR 2594346

[33] van der Hofstad, R., Luczak, M., Spencer, J.: The second largest component in the supercritical 2D Hamming graph. Random Structures Algorithms **36**, 80–89 (2010) Zbl 1208.05139 MR 2591048

[34] van der Hofstad, R., Nachmias, A.: Unlacing hypercube percolation: a survey. Metrika **77**, 23–50 (2014) Zbl 06272735 MR 3152019

[35] van der Hofstad, R., Slade, G.: Asymptotic expansions in n^{-1} for percolation critical values on the n-cube and $\mathbb{Z}^n$. Random Structures Algorithms **27**, 331–357 (2005) Zbl 1077.60077 MR 2162602

[36] van der Hofstad, R., Slade, G.: Expansion in n^{-1} for percolation critical values on the n-cube and $\mathbb{Z}^n$: the first three terms. Combin. Probab. Comput. **15**, 695–713 (2006) Zbl 1102.60090 MR 2248322

[37] Hoory, S., Linial, N., Wigderson, A.: Expander graphs and their applications. Bull. Amer. Math. Soc. **43**, 439–561 (2006) Zbl 1147.68608 MR 2247919

[38] Janson, S., Knuth, D. E., Łuczak, T., Pittel, B.: The birth of the giant component. Random Structures Algorithms **4**, 71–84 (1993) Zbl 0795.05127 MR 1220220

[39] Janson, S., Łuczak, T., Ruciński, A.: Random Graphs. Wiley, New York (2000) Zbl 0968.05003 MR 1782847

[40] Kozma, G., Nachmias, A.: The Alexander–Orbach conjecture holds in high dimensions. Invent. Math. **178**, 635–654 (2009) Zbl 1180.82094 MR 2551766

[41] Kozma, G., Nachmias, A.: A note about critical percolation on finite graphs. J. Theoret. Probab. **24**, 1087–1096 (2011) Zbl 1235.60138

[42] Kozma, G., Nachmias, A.: Arm exponents in high-dimensional percolation. J. Amer. Math. Soc. **24**, 375–409 (2011) Zbl 1219.60085 MR 2748397

[43] Levin, D., Peres, Y., Wilmer, E.: Markov Chains and Mixing Times. Amer. Math. Soc., Providence, RI (2009) Zbl 1160.60001 MR 2466937

[44] Łuczak, T.: Component behavior near the critical point of the random graph process. Random Structures Algorithms **1**, 287–310 (1990) Zbl 0745.05048 MR 1099794

[45] Łuczak, T., Pittel, B., Wierman, J. C. (1994), The structure of a random graph at the point of the phase transition. Trans. Amer. Math. Soc. **341**, 721–748 (1990) Zbl 0807.05065 MR 1138950

[46] Molloy, M., Reed, B.: A critical point for random graphs with a given degree sequence. Random Structures Algorithms **6**, 161–179 (1995) Zbl 0823.05050 MR 1370952

[47] Nachmias, A.: Mean-field conditions for percolation on finite graphs. Geom. Funct. Anal. **19**, 1171–1194 (2009) Zbl 1182.60025 MR 2570320

[48] Nachmias, A., Peres, Y.: Critical percolation on random regular graphs. Random Structures Algorithms **36**, 111–148 (2010) Zbl 1209.05228 MR 2583058

[49] Nachmias, A., Peres, Y.: Critical random graphs: diameter and mixing time. Ann. Probab. **36**, 1267–1286 (2008) Zbl 1160.05053 MR 2435849

[50] Penrose, M. D.: Random minimal spanning tree and percolation on the N-cube. Random Structures Algorithms **12**, 63–82 (1998) Zbl 0899.60083 MR 1637395

[51] Peres, Y.: Probability on trees: an introductory climb. In: École d'été de Probabilités de Saint-Flour XXVII, Lecture Notes in Math. 1717, Springer, 193–280 (1999) Zbl 0957.60001 MR 1746302

[52] Pittel, B., Wormald, N.: Counting connected graphs inside-out. J. Combin. Theory Ser. B **93**, 127–172 (2005) Zbl 1057.05044 MR 2117934

[53] Reimer, D.: Proof of the van den Berg–Kesten conjecture. Combin. Probab. Comput. **9**, 27–32 (2000) Zbl 0947.60093 MR 1751301