EURANDOM PREPRINT SERIES
2013-026

Unlacing hypercube percolation: a survey

October 31, 2012

Remco van der Hofstad, Asaf Nachmias
ISSN 1389-2355

Noname manuscript No.
(will be inserted by the editor)

Unlacing hypercube percolation: a survey

Remco van der Hofstad · **Asaf Nachmias**

Received: October 30, 2012/ Accepted: August 1, 2013

Abstract The purpose of this note is twofold. First, we survey the study of the percolation phase transition on the Hamming hypercube $\{0,1\}^m$ obtained in the series of papers [1,2,3,4]. Secondly, we explain how this study can be performed without the use of the so-called "lace-expansion" technique. To that aim, we provide a novel simple proof that the triangle condition holds at the critical probability.

Keywords Percolation · Phase transition · Hypercube

1 Introduction

This paper is intended to be a companion to the papers [1,2,3,4] in the setting of percolation on the Hamming hypercube $\{0,1\}^m$. Our goal here is to present the most recent state of affairs of this topic, emphasizing ideas, techniques and the gaps in our understanding. We will present one novel proof of a result obtained in [2], namely, that the triangle condition on the hypercube holds at the critical probability. This proof is simpler than the one presented in [2] as it does not use the *lace-expansion* technique. See also [**?**,**?**,**?**] for different lace-expansion-free proofs of the triangle-condition on some classes of non-amenable graphs. Given this estimate, it will be indicated in this note how the study of the qualitative properties of

The work of RvdH was supported in part by the Netherlands Organization for Scientific Research (NWO). The work of AN was partially supported by NSF grant #6923910 and NSERC grants.

R. van der Hofstad
Department of Mathematics and Computer Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands.
Tel.: +31 (40) 2472910
Fax: +31 (40) 2475665
E-mail: rhofstad@win.tue.nl

A. Nachmias
Department of Mathematics, University of British Columbia, 121-1984 Mathematics Rd, Vancouver, BC, Canada V6T1Z2.

the phase transition on the hypercube can be obtained *without* the use of the lace expansion. While this has worked in the hypercube setting, “unlacing” the proofs in the setting of high-dimensional tori seems much more difficult and requires new ideas. We hope that the hypercube study will get us closer to this goal.

The paper is organized as follows. In Section **??**, we describe the phase transition in the Erdős and Rényi random graph, our main source of inspiration and in Section **??** we describe the analogous results, obtained in [1,2,4], in the setting of the hypercube. We proceed in Section **??** to introduce and discuss the role of the so-called *triangle condition* for percolation. This condition arises fairly naturally in the study of percolation and to exemplify this, we present a classical argument of Aizenman and Newman [**?**] showing how to control the expected cluster size using the triangle condition. Next, in Section **??** we present some conditions about the behavior of the random walk on the underlying graph and state a general theorem that allows us to analyze the phase transition in percolation on any graph satisfying these conditions. In Section **??** we restrict our attention back to the hypercube setting and verify that the random walk conditions of the previous section holds. In particular, we state there the required estimate on p_c (Theorem **??**) which we prove in this paper. We conclude this chapter in Section **??** by discussing open problems.

Section **??** contains an overview of the proof of Theorem **??** and Section **??** provides a proof of Theorem **??** which together with the argument in Section **??** yields a simple proof that the triangle condition holds on the hypercube.

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 j-th largest component obtained this way. An inspiring discovery of Erdős and Rényi [**?**] 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 and $|\mathscr{C}_1| = \Theta(n)$ whp when $c > 1$. Here, we say that a sequence of events E_n occurs *with high probability* (whp) when $\lim_{n\to\infty} \mathbf{P}(E_n) = 1$. We further 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 $g(m) = O(f(m))$, $f(m) = \Omega(g(m))$ if $1/f(m) = O(1/g(m))$ and $f(m) = o(g(m))$ if $f(m)/g(m)$ tends to 0 as $m \to \infty$. We say that $f(m) \gg g(m)$ when $g(m) = o(f(m))$.

The investigation of the case c close to 1, initiated by Bollobás [**?**] and further studied by Łuczak [**?**], revealed an intricate picture of the phase transition's nature. See [**?**] for results up to 1984, and [**?**,**?**,**?**,**?**] for references to subsequent work. We briefly describe these now.

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

$$\frac{|\mathscr{C}_j|}{2\varepsilon_n^{-2}\log(\varepsilon_n^3 n)} \xrightarrow{\mathbb{P}} 1\,,$$

where $\xrightarrow{\mathbb{P}}$ denotes convergence in probability.

The critical window. When $p = (1+O(n^{-1/3}))/n$, for any fixed integer $j \geq 1$,

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

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

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

$$\frac{|\mathscr{C}_1|}{2\varepsilon_n n} \xrightarrow{\mathbb{P}} 1\,,$$

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

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

1.2 The phase transition on the hypercube

Perform percolation on the hypercube $\{0,1\}^m$ with edge probability $p \in [0,1]$. Unlike percolation on infinite transitive graphs, one of the inherent difficulties with percolation on finite graphs is that is not obvious how to *define* the critical percolation probability. We would like a natural definition which would coincide with the critical value $p = (1+O(n^{-1/3}))/n$ of $G(n,p)$. The mean cluster size in the latter is of order $n^{1/3}$. This fact inspired Borgs, Chayes, the first author, Slade and Spencer [1,**?**,**?**] to suggest that the precise location $p_c = p_c(\lambda)$ 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 that is typically taken to be small. Here $2^{m/3}$ can be viewed as the cube root of the volume of the graph, i.e., its number of vertices. There are several other more intuitive definitions (see the discussion in Section 7 of [**?**]), however, in order to justify these definitions one needs to show that analogous results to the ones described in Section **??** holds with this definition. To the best of our knowledge this was only done with (**??**). Let us now describe the phase transition of percolation on the hypercube around this p_c.

From here on, we take $p_c = p_c(\lambda)$ with $\lambda \in (0,1)$ a fixed constant. The phase transition on the hypercube is described in the following three theorems, in all of which we consider edge percolation on the hypercube $\{0,1\}^m$ with varying p and write $\mathscr{C}_j$ fr the j-th largest component.

Theorem 1 **(The subcritical phase [1,2])** *Put $p = p_c(1-\varepsilon_m)$ where $\varepsilon_m = o(1)$ is a positive sequence satisfying $\varepsilon_m \gg 2^{-m/3}$. Then, for all fixed $\delta > 0$,*

$$\mathbf{P}_p\big(\varepsilon_m^{-2}/3600 \le |\mathscr{C}_1| \le (2+\delta)\varepsilon_m^{-2}\log(\varepsilon_m^3 2^m)\big) = 1 - o(1)\,, \tag{1.2}$$

and

$$\mathbb{E}_p|\mathscr{C}(0)| = \frac{1+o(1)}{|\varepsilon_m|}. \tag{1.3}$$

Theorem 2 **(The critical window [1,2,?])** *Put $p = p_c(1+\varepsilon_m)$ with $|\varepsilon_m| = O(2^{-m/3})$. Then, for every $j \ge 1$,*

$$\mathbf{P}_p(\omega^{-1}2^{2m/3} \le |\mathscr{C}_j| \le \omega 2^{2m/3}) \ge 1 - O(\omega^{-1})\,. \tag{1.4}$$

and

$$\mathbb{E}_p|\mathscr{C}(0)| = \Theta(2^{m/3})\,. \tag{1.5}$$

Theorem 3 **(The supercritical phase [4])** *Put $p = p_c(1+\varepsilon_m)$ where $\varepsilon_m = o(1)$ is a positive sequence with $\varepsilon_m \gg 2^{-m/3}$. Then*

$$\frac{|\mathscr{C}_1|}{2\varepsilon_m 2^m} \xrightarrow{\mathbb{P}} 1\,, \tag{1.6}$$

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

$$\mathbb{E}_p|\mathscr{C}(0)| = (4+o(1))\varepsilon_m^2 2^m\,. \tag{1.7}$$

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

$$\frac{|\mathscr{C}_2|}{\varepsilon_m 2^m} \xrightarrow{\mathbb{P}} 0\,. \tag{1.8}$$

Theorems **??** and **??** are proved in [1,2]. The work in [1,2] did not provide sharp estimates for the supercritical phase and the authors conjectured (see Conjecture 3.2 in [3]) the statement of Theorem **??**, proved in [4]. Thus, Theorems **??**–**??** fully identify the phase transition of the largest connected component and the critical window in the hypercube.

1.3 The role of the percolation triangle condition

Let us briefly review the general study of random subgraphs of general finite transitive graphs initiated in [1,**?**]. Let G be a finite transitive graph and write V for the number of vertices of G and m for its degree. We think of G as a member of a sequence of graphs whose volume V tends to infinity. 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 j-th 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

$$\nabla_p(x,y) = \sum_{u,v} \mathbf{P}_p(x \leftrightarrow u)\mathbf{P}_p(u \leftrightarrow v)\mathbf{P}_p(v \leftrightarrow y)\,. \tag{1.9}$$

The quantity $\nabla_p(x,y)$, known as the *triangle diagram*, was introduced by Aizenman and Newman [**?**] to study critical percolation on high-dimensional infinite lattices. In that setting, the 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 [**?**] proved that the triangle condition holds; it is expected to hold only above the upper critical dimension for percolation, that is, when the dimension of the lattice is greater than 6. It allows to deduce that numerous critical exponents attain the same values as they do on an infinite regular tree, see e.g. [**?**,**?**,**?**,**?**].

When G is a finite graph, $\nabla_p(0,0)$ is obviously always finite, but there is 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. 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 that $V \to \infty$ and that $\lambda \in (0,1)$ is a fixed and small constant. Let $p_c = p_c(\lambda)$ be defined by

$$\mathbb{E}_{p_c(\lambda)}|\mathscr{C}(0)| = \lambda V^{1/3}. \tag{1.10}$$

The finite triangle condition is the assumption that $\nabla_{p_c(\lambda)}(x,y) \leq \mathbb{1}_{\{x=y\}} + a_0$, for some $a_0 = a_0(\lambda)$ sufficiently small. The strong triangle condition, defined in [2, (1.26)], is the statement that there exists a constant C such that for all $p \leq p_c$,

$$\nabla_p(x,y) \leq \mathbb{1}_{\{x=y\}} + \frac{C\chi(p)^3}{V} + o(1), \tag{1.11}$$

where $o(1)$ tends to 0 as $m \to \infty$. In [2], (**??**) is shown to hold for various graphs: the complete graph, the hypercube and high-dimensional tori $\mathbb{Z}_n^d$. The proofs in [2] rely on the *lace expansion*, a perturbative technique to investigate the two-point function $\mathbf{P}_{p_c}(0 \longleftrightarrow x)$ that was first used for percolation on the high-dimensional infinite lattice by Hara and Slade [**?**]. The lace expansion is an extremely powerful technique, but is also quite involved. We feel that it should only be used when more elementary techniques fail. Apart from surveying the literature on hypercube percolation, our aim in this paper is to show that Theorems **??**-**??** can be proved *without* relying on the lace expansion.

The main result of [1] is that the triangle condition implies strong estimates on $|\mathscr{C}_1|$ in the critical and subcritical case:

Theorem 4 ([1]) *Consider edge percolation on a finite transitive graph G having V vertices and degree m satisfying the strong triangle condition* (**??**) *where $m \to \infty$ as $V \to \infty$. Then the assertions of Theorems* **??** *and* **??** *hold when each occurrence of 2^m is replaced with V.*

The triangle condition is significant since it arises naturally in various calculation one performs. To indicate how this occurs, we provide here a classical argument of Aizenman and Newman [**?**] controlling the size of $\mathbb{E}_p|\mathscr{C}(0)|$ as p varies. We put $\chi(p) = \mathbb{E}_p[|\mathscr{C}(0)|]$ and will show that the strong triangle condition (**??**) implies that for all $p < p_c$,

$$\chi(p) = \frac{1 + O(\max_{v\neq 0} \nabla_{p_c}(0,v))}{m(p_c - p) + \chi(p_c)^{-1}}. \tag{1.12}$$

In particular, (**??**) implies that $\chi(p) = \Theta(V^{1/3})$ whenever $p \le p_c$ is in the scaling window, i.e., when $m(p_c - p) = O(V^{-1/3})$ as stated in (**??**) in Theorem **??**. Equation (**??**) also proves (**??**) in Theorem **??**.

We start by proving the lower bound on $\chi(p)$ of (**??**). We remark that this bound is valid for *all* transitive graphs (that is, we do not require here the triangle condition). We say that a (directed) edge (u,v) is *pivotal* for $0 \longleftrightarrow x$ when (a) $0 \longleftrightarrow u$ and (b) $0 \longleftrightarrow x$ in the (possibly modified) configuration where the status of (u,v) is turned to open, while 0 is not connected to x in the (possibly modified) configuration where the status of (u,v) is turned to closed. By Russo's formula,

$$\frac{d}{dp}\mathbf{P}_p(0 \longleftrightarrow x) = \sum_{(u,v)\in E(G)} \mathbf{P}_p\big((u,v) \text{ is pivotal for } 0 \longleftrightarrow x\big). \tag{1.13}$$

Summing over x yields

$$\frac{\mathrm{d}}{\mathrm{d}p}\chi(p) = \sum_{x\in V(G)} \sum_{(u,v)\in E(G)} \mathbf{P}_p\big((u,v) \text{ is pivotal for } 0 \longleftrightarrow x\big). \tag{1.14}$$

If (u,v) is pivotal for $0 \longleftrightarrow x$, then there exist two disjoint paths of open edges connecting 0 and u, and v and x, respectively. Thus, $\{0 \longleftrightarrow u\} \circ \{v \longleftrightarrow x\}$ occurs. The BK inequality [**?**,**?**] gives

$$\frac{\mathrm{d}}{\mathrm{d}p}\chi(p) \le \sum_{x\in V(G)} \sum_{(u,v)\in E(G)} \mathbf{P}_p(0 \longleftrightarrow u)\mathbf{P}_p(v \longleftrightarrow x) = m\chi(p)^2. \tag{1.15}$$

We rewrite the last inequality as $\frac{\mathrm{d}}{\mathrm{d}p}\chi(p)^{-1} \ge -m$, and integrate over $[p, p_c]$ to get

$$\chi(p_c)^{-1} - \chi(p)^{-1} \ge -m(p_c - p), \tag{1.16}$$

so that

$$\chi(p) \geq \frac{1}{m(p_c - p) + \chi(p_c)^{-1}}, \tag{1.17}$$

showing the lower bound on $\chi(p)$. For the upper bound, we let $\tilde{\mathscr{C}}^{(u,v)}(0)$ consist of those vertices which are connected to 0 without using the edge (u,v) and for a set of vertices $A \subseteq V$, the restricted two-point function $\tau^A(v,x)$ is the probability that there exists an open path from v to x that does not touch A. The event that (u,v) is pivotal for $0 \longleftrightarrow x$ equals the intersection of $0 \longleftrightarrow u$ and the event that v is connected to x such that *all* open paths from v to x do not touch $\tilde{\mathscr{C}}^{(u,v)}(0)$. When there exists an open path from v to x that touches $\tilde{\mathscr{C}}^{(u,v)}(0)$, then $v \in \tilde{\mathscr{C}}^{(u,v)}(0)$, so that in fact *all* open paths from v to x touch $\tilde{\mathscr{C}}^{(u,v)}(0)$. Therefore, the event that (u,v) is pivotal for $0 \longleftrightarrow x$ also equals the intersection of $0 \longleftrightarrow u$ and the event that v is connected to x such that there exists an open path from v to x that does not touch $\tilde{\mathscr{C}}^{(u,v)}(0)$. Hence, by [**?**, (6.5)] or [2, Lemma 3.2],

$$\mathbf{P}_p((u,v) \text{ is pivotal for } 0 \longleftrightarrow x) = \mathbb{E}_p[\mathbb{1}_{\{0 \longleftrightarrow u\}} \tau^{\tilde{\mathscr{C}}^{(u,v)}(0)}(v,x)]. \tag{1.18}$$

We note that

$$\mathbf{P}_p(v \longleftrightarrow x) - \tau^A(v,x) = \mathbf{P}_p(v \overset{A}{\longleftrightarrow} x), \tag{1.19}$$

where we write that $v \overset{A}{\longleftrightarrow} x$ when $v \longleftrightarrow x$ and all open paths from v to x have an edge touching A. Thus,

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}p}\chi(p) &= \sum_{x \in V(G)} \sum_{(u,v) \in E(G)} \mathbb{E}_p[\mathbb{1}_{\{0 \longleftrightarrow u\}}]\mathbf{P}_p(v \leftrightarrow x) \\
&\quad - \sum_{x \in V(G)} \sum_{(u,v) \in E(G)} \mathbb{E}_p[\mathbb{1}_{\{0 \longleftrightarrow u\}} \mathbf{P}_p(v \overset{\tilde{\mathscr{C}}^{(u,v)}(0)}{\longleftrightarrow} x)] \\
&= m\chi(p)^2 - \sum_{x \in V(G)} \sum_{(u,v) \in E(G)} \mathbb{E}_p[\mathbb{1}_{\{0 \longleftrightarrow u\}} \mathbf{P}_p(v \overset{\tilde{\mathscr{C}}^{(u,v)}(0)}{\longleftrightarrow} x)].
\end{aligned} \tag{1.20}$$

Now, for any $A \subseteq \mathbb{Z}^d$,

$$\mathbf{P}_p(v \overset{A}{\longleftrightarrow} x) \leq \sum_a \mathbf{P}_p\big(\{v \longleftrightarrow a\} \circ \{a \longleftrightarrow x\}\big) \mathbb{1}_{\{a \in A\}}, \tag{1.21}$$

which by BK inequality and summing over x leads to

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}p}\chi(p) &\geq m\chi(p)^2 - \chi(p) \sum_{(u,v) \in E(G)} \sum_a \mathbf{P}_p(0 \longleftrightarrow u, a \in \tilde{\mathscr{C}}^{(u,v)}(0)) \mathbf{P}_p(v \longleftrightarrow a) \\
&\geq m\chi(p)^2 - \chi(p) \sum_{(u,v) \in E(G)} \sum_a \mathbf{P}_p(0 \longleftrightarrow u, 0 \longleftrightarrow a) \mathbf{P}_p(v \longleftrightarrow a).
\end{aligned} \tag{1.22}$$

If $0 \longleftrightarrow u$ and $0 \longleftrightarrow a$, then there exists z such that $\{0 \longleftrightarrow z\} \circ \{z \longleftrightarrow u\} \circ \{z \longleftrightarrow a\}$, so by the BK inequality

$$\frac{\mathrm{d}}{\mathrm{d}p}\chi(p) \geq m\chi(p)^2 - \chi(p) \sum_{(u,v) \in E(G)} \sum_{a,z} \mathbf{P}_p(0 \longleftrightarrow z) \mathbf{P}_p(z \longleftrightarrow u) \mathbf{P}_p(z \longleftrightarrow a) \mathbf{P}_p(v \longleftrightarrow a).$$

The sum over a, z looks almost like the triangle diagram, except for the pesky $\mathbf{P}_p(0 \longleftrightarrow z)$ factor. However, by transitivity the double sum on the right hand side remains the same if we replace 0 by any other vertex. Hence we may sum this over 0 and divide by V to get

$$\frac{\mathrm{d}}{\mathrm{d}p}\chi(p) \geq m\chi(p)^2 - V^{-1}\chi(p)^2 \sum_{(u,v)\in E(G)} \nabla_p(u,v) \geq m\chi^2(p)[1 - \max_{v\neq 0}\nabla_p(0,v)],$$

giving that

$$\frac{\mathrm{d}}{\mathrm{d}p}\chi(p) \geq m\chi(p)^2[1 - \max_{v\neq 0}\nabla_{p_c}(0,v)]. \tag{1.23}$$

We now integrate as we do in (**??**) and obtain the lower bound in (**??**).

1.4 Random walk conditions for percolation

We now describe a general theorem, obtained in [4], which allows to deduce as corollaries the assertions of Theorems **??**-**??** for any underlying transitive graph sequence adhering to certain geometric conditions. These conditions are satisfied in the case of the hypercube (and so Theorems **??**-**??** follow) but in general they are more restrictive than the triangle condition, (for instance, they do not hold in the case of high-dimensional tori), however, they are easier to verify since they are expressed in terms of random walks. In particular, these conditions imply the strong triangle condition (and hence by Theorem **??** they imply Theorems **??** and **??**), but more importantly they allow us to analyze percolation in the supercritical case, where the triangle diagram ceases to be small, and obtain Theorem **??**.

Let G be a finite transitive graph on V vertices and with degree m. Consider the non-backtracking random walk (NBW) on it (this is just a simple random walk not allowed to traverse back on the edge it just came from, for a formal definition see [4, Section 3.4]). For any two vertices x, y, we put $\mathbf{p}^t(x,y)$ for the probability that the walk started at x visits y at time t. We write T_{mix} for the *uniform mixing time* of the walk, that is,

$$T_{\mathrm{mix}} = \min\left\{t\colon \max_{x,y} \frac{\mathbf{p}^t(x,y) + \mathbf{p}^{t+1}(x,y)}{2} \leq (1+o(1))V^{-1}\right\}, \tag{1.24}$$

where $o(1)$ tends to 0 slowly. Then the main result in [4] is as follows:

Theorem 5 **([4])** *Let G be a transitive graph on V vertices with degree m and define p_c as in (**??**) with $\lambda = 1/10$. Assume that the following conditions hold:*

1. $m \to \infty$ *as* $V \to \infty$,
2. $[p_c(m-1)]^{T_{\mathrm{mix}}} = 1 + o(1)$,
3. *For any vertices x, y,*

$$\sum_{u,v} \sum_{\substack{t_1,t_2,t_3=0 \\ t_1+t_2+t_3\geq 3}}^{T_{\mathrm{mix}}} \mathbf{p}^{t_1}(x,u)\mathbf{p}^{t_2}(u,v)\mathbf{p}^{t_3}(v,y) = o(1/\log V). \tag{1.25}$$

Then,

*(a) the finite triangle condition (**??**) holds (and hence the assertions of Theorems **??**-**??** hold),*

(b) for any sequence $\varepsilon = \varepsilon_m$ satisfying $\varepsilon_m \gg V^{-1/3}$ and $\varepsilon_m = o(T_{\rm mix}^{-1})$,

$$\frac{|\mathscr{C}_1|}{2\varepsilon_m V} \xrightarrow{\mathbb{P}} 1, \qquad \mathbb{E}_p|\mathscr{C}(0)| = (4+o(1))\varepsilon_m^2 V, \qquad \frac{|\mathscr{C}_2|}{\varepsilon_m V} \xrightarrow{\mathbb{P}} 0. \tag{1.26}$$

In Section **??** we will prove part (a) of the theorem above, and in Section **??** we will verify the conditions of the Theorem for the case of the hypercube $\{0,1\}^m$. Hence, we will obtain a proof that the triangle condition holds on the hypercube. Note that condition (2) involves both a random walk estimate (bounding $T_{\rm mix}$) and a percolation estimate (bounding p_c). Let us now discuss how the verification of these conditions is done in a rather elementary way.

1.5 Back to the hypercube

It is a classical fact that the total-variation mixing time of the random walk on the hypercube is of order $m \log m$ [**?**]. A separate argument is needed to show that this is the correct order for $T_{\rm mix}$ since (a) we are dealing with the non-backtracking walk; and (b) we require a bound on the stronger uniform mixing time. This can be done by analyzing the transition matrix of the non-backtracking random walk using classical tools. This analysis is performed by Fitzner and the first author in [**?**, Theorem 3.5] and also allows us to verify condition (3) of Theorem **??**. We will not delve further into this part of the proof.

Thus, given that $T_{\rm mix} = O(m \log m)$, the verification of condition (2) in Theorem **??** in the case of the hypercube simply states that $p_c = \frac{1}{m-1} + o(m^{-2}\log m)$ (note that this does not follow from the results in [**?**], which prove that $p_c \le (1+\varepsilon)/m$ for every $\varepsilon > 0$ as $m \to \infty$). This estimate (and more) was already proved in the work of the first author and Slade [**?**,**?**] so no further estimates on p_c were required in [4]. However, the estimate we require is much weaker and the proofs in [**?**,**?**] are difficult and rely on the lace expansion. In this paper we provide an elementary argument giving this estimate. This is the last piece in the "unlacing" puzzle which verifies condition (2) of Theorem **??** in the case of the hypercube.

Theorem 6 (Unlacing the lace expansion in the hypercube) *Consider edge percolation on the hypercube $\{0,1\}^m$ and fix $\lambda = 1/10$. Then, there exists $C > 0$ such that*

$$p_c \le \frac{1 + 5/(2m^2) + C/m^{-3}}{m-1}. \tag{1.27}$$

*Consequently, $[(m-1)p_c]^{T_{\rm mix}} = 1 + o(1)$, so that the results in Theorems **??** and **??** apply (and hence also the assertions of Theorems **??**-**??**).*

This theorem is the only novel result of this paper, and is proved in Section **??**. Let us further briefly discuss the precise value of p_c since it is related to the early literature on hypercube percolation.

The asymptotic expansion of p_c The problem of establishing a phase transition for the appearance of a component of size of order 2^m was solved in the breakthrough work of Ajtai, Komlós and Szemerédi [**?**]. 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: for $c < 1$, the largest component has size of order m whp, while for $c > 1$, the largest component has size linear in 2^m whp. In fact, they also prove that the giant component has size $\zeta(c)2^m(1+o(1))$ whp when $p = c/m$, where $\zeta(c)$ is the survival probability of a Poisson branching process with expected offspring equal to c. Thus, $p_c \approx 1/m$, but it was unclear at that time just how close it is.

The first improvement to [**?**] was obtained by Bollobás, Kohayakawa and Łuczak [**?**]. They showed that if $p = (1+\varepsilon_m)/m$ 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. This raises the question whether $p_c = 1/(m-1)$, which is answered negatively in [**?**,**?**]. These results give the most precise estimates on p_c to date:

Theorem 7 (Asymptotic expansion of p_c [?,?]) *For edge percolation on the hypercube $\{0,1\}^m$, there exist rational coefficients $(a_i)_{i\geq 1}$ with $a_1 = a_2 = 1, a_3 = 7/2$ such that, for every $s \geq 1$, as $m \to \infty$,*

$$p_c = \sum_{i=1}^{s} a_i m^{-i} + O(m^{-(s+1)}). \tag{1.28}$$

Note that Theorem **??** implies Theorem **??** with a matching lower bound. Further, note that by Theorems **??**, **??** and **??**, whatever $s \geq 2$ is, the largest cluster jumps from $O(m^{2s-1})$ for $p = \sum_{i=1}^{s} a_i m^{-i} + \eta m^{-s}$ with $\eta < 0$ to $\Theta(2^m/m^{s-1})$ for $p = \sum_{i=1}^{s} a_i m^{-i} + \eta m^{-s}$ with $\eta > 0$. Thus, the phase transition in η is extremely sharp for every $s \geq 2$ fixed.

1.6 Open problems

We collect here a list of what we consider to be important open problems in this area. Some of these problems appear in [4, Section 8], but not all.

1. **Percolation on high-dimensional tori.** Consider edge percolation on the nearest-neighbor torus $\mathbb{Z}_n^d$ where d is a large fixed constant and $n \to \infty$ with $p = p_c(1+\varepsilon_n)$ such that $\varepsilon_n \gg n^{-d/3}$ and $\varepsilon_n = o(1)$. Show that $|\mathscr{C}_1|/(\varepsilon_n n^d)$ 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-edge percolation on the infinite lattice $\mathbb{Z}^d$.
2. **Identify the scaling limit of cluster sizes in the scaling window.** 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. Is it the same limiting distribution of critical clusters in $G(n,p)$ identified by Aldous [**?**]?
   We remark that in [**?**] it is proved that any subsequential limit of $\{|\mathscr{C}_1|2^{-2m/3}\}_{m\geq 1}$ is a proper random variable, that is, $\mathbf{P}(X = \mathbb{E}[X]) < 1$. This *non-concentration* is the hallmark of critical behavior.

3. **Prove that the discrete duality principle holds for hypercube percolation.** Show that $|\mathscr{C}_1| = (2+o(1))\varepsilon^{-2}\log(\varepsilon^3 2^m)$ when $\varepsilon_m \ll -2^{-m/3}$ and $\varepsilon_m = o(1)$ and that $|\mathscr{C}_2| = (2+o(1))\varepsilon^{-2}\log(\varepsilon_m^3 2^m)$ when $p = p_c(1+\varepsilon_m)$ with $\varepsilon_m \gg 2^{-m/3}$. This is also the content of [3, Conjectures 3.1 and 3.3] and is proved for some values of ε_m in [**?**]. In $G(n,p)$ these results are proved in [**?**] and [**?**, Theorem 5.6].
4. **Prove a central limit theorem for** $|\mathscr{C}_1|$**.** Show that $|\mathscr{C}_1|$ satisfies a central limit theorem throughout the supercritical regime. In $G(n,p)$ this and much more was established by Pittel and Wormald [**?**].
5. **Unlace the asymptotic expansion of** p_c**.** Find a proof of Theorem **??** that does not rely on the lace expansion. Possibly, the ideas in the proof of Theorem **??** in Section **??** can be used.
6. **Compute further coefficients of the asymptotic expansion of** p_c**.** Find the numerical values of a_i for $i \geq 4$ in Theorem **??**. There is a large physics literature on asymptotic expansions of critical values. See e.g., [**?**] for some of the references. We expect that $a_4 = 16$, as the first 4 coefficients of the asymptotic expansion of p_c can be expected to be the same as the ones for the asymptotic expansion of $p_c(\mathbb{Z}^d)$ in terms of inverse powers of $2d$ (see e.g. [**?**]). We also expect that a_5 is *not* equal to the 5th coefficient in the asymptotic expansion of $p_c(\mathbb{Z}^d)$ in terms of $1/(2d)$, which is predicted to be equal to 103 [**?**]. Recently, substantial progress was made for the asymptotic expansion of the connective constant for self-avoiding walk on $\mathbb{Z}^d$, for which the first 13 coefficients have been computed by Clisby, Liang and Slade (see [**?**,**?**]).

2 Overview of the proof of the supercritical phase

In this section we give an overview of the key steps in the proofs in [4]. From here on, we assume that ε_m is a sequence such that $\varepsilon_m = o(1)$ but $\varepsilon_m^3 V \to \infty$.

2.1 Notations and tools

We write $d_{G_p}(x,y)$ for the length of a shortest p-open path between x,y and put $d_{G_p}(x,y) = \infty$ if x is not connected to y in G_p. We write $x \stackrel{r}{\leftrightarrow} y$ if $d_{G_p}(x,y) \leq r$ and $x \stackrel{=r}{\longleftrightarrow} y$ if $d_{G_p}(x,y) = r$ and $x \stackrel{[a,b]}{\longleftrightarrow} y$ if $d_{G_p}(x,y) \in [a,b]$. The *intrinsic metric* ball of radius r around x and its boundary are defined by

$$B_x(r) = \{y\colon d_{G_p}(x,y) \leq r\}, \qquad \partial B_x(r) = \{y\colon d_{G_p}(x,y) = r\}. \tag{2.1}$$

Note that these are random sets of the graph and not the balls in shortest path metric of the graph G. We often drop 0 from notation and write $B(r)$ for $B_0(r)$ whenever possible.

2.2 Tails of the supercritical cluster size

We start by describing the tail of the cluster size in the supercritical regime.

Theorem 8 (Bounds on the cluster tail) *Let G be a finite transitive graph of degree m on V vertices such that the finite triangle condition (**??**) holds and put $p = p_c(1+\varepsilon_m)$ where $\varepsilon_m = o(1)$ and $\varepsilon_m \gg V^{-1/3}$. Then, for the sequence $k_0 = \varepsilon_m^{-2}(\varepsilon_m^3 V)^{1/4}$,*

$$\mathbf{P}(|\mathscr{C}(0)| \geq k_0) = 2\varepsilon_m(1+o(1)). \tag{2.2}$$

This theorem is reminiscent of the fact that a branching process with Poisson offspring distribution of mean $1+\varepsilon$ has survival probability of $2\varepsilon(1+O(\varepsilon))$. The choice of k_0 is such that $k_0 \gg 1/\varepsilon_m^2$. (In [4], we also rely on the fact that $k_0 \ll (\varepsilon_m^3 V)^\alpha/\varepsilon_m^2$ for some $\alpha < 1/3$.) This choice is inspired by the fact that $\mathbf{P}(T \geq k_0) = 2\varepsilon(1+o(1))$ precisely when $k_0 \gg 1/\varepsilon^2$, where T is the total progeny of a Poisson branching process with mean $1+\varepsilon$ offspring.

Upper and lower bounds of order ε for the cluster tail were proved already in [2] using Barsky and Aizenman's differential inequalities [**?**], and were sharpened in [4, Appendix A] to obtain the right constant 2.

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

$$Z_{\geq k} = \big|\{v\colon |\mathscr{C}(v)| \geq k\}\big|. \tag{2.3}$$

We use Theorem **??** to show that $Z_{\geq k_0}$, with k_0 as in the theorem, is concentrated.

Lemma 1 (Concentration of $Z_{\geq k_0}$) *In the setting of Theorem **??**, if $m \to \infty$, then*

$$\frac{Z_{\geq k_0}}{2\varepsilon V} \xrightarrow{\mathbb{P}} 1, \qquad \textit{and} \qquad \mathbb{E}|\mathscr{C}(0)| \leq (4+o(1))\varepsilon^2 V. \tag{2.4}$$

Lemma **??** immediately proves the upper bound on $|\mathscr{C}_1|$ in Theorem **??**:

Proof of upper bound on $|\mathscr{C}_1|$ in Theorem ??. Note that $\{|\mathscr{C}_1| \geq k\} = \{Z_{\geq k} \geq 1\}$, so that $|\mathscr{C}_1| \leq Z_{\geq k}$ on the event $\{Z_{\geq k} \geq 1\}$. Applying this to $k = k_0$ and using Lemma **??** proves the upper bound in Theorem **??**. □

2.3 Uniform connection bounds and the role of the random walk

We expand here on one of our most useful estimates on percolation connection probabilities. In its proof, a simple key connection between percolation and the mixing time of the non-backtracking walk is revealed. 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 $x \neq 0$ 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 $p_c \geq m^{-1}$, while 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 the proof of Theorem **??** in [4] 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, as shown in [4, Lemma 3.12]:

Lemma 2 (Uniform connection estimates) *Perform edge percolation on any graph G satisfying the assumptions of Theorem **??**. Then, for every $r \geq T_{\rm mix}$ and any vertex $x \in G$*

$$\mathbf{P}_{p_c}(0 \xleftrightarrow{[T_{\rm mix},r]} x) \leq (1+o(1))\frac{\mathbb{E}|B(r)|}{V}\,, \tag{2.5}$$

*where $T_{\rm mix}$ is uniform mixing time as defined above Theorem **??**. In particular,*

$$\mathbf{P}_{p_c}(0 \xleftrightarrow{[T_{\rm mix},\infty)} x) \leq (1+o(1))\frac{\mathbb{E}|\mathscr{C}(0)|}{V}\,. \tag{2.6}$$

The proof of the above lemma is short and elementary, see [4]. There it is also shown how to obtain similar estimates for $p = p_c(1+\varepsilon)$ (with an error depending on ε). The uniformity of this lemma allows us to decouple the sum in the triangle diagram and yields a simple proof of the strong triangle condition, as we now show.

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

$$\begin{aligned}\sum_{u,v}\mathbf{P}_p(x \xleftrightarrow{[T_{\rm mix},\infty)} u)\mathbf{P}_p(u \leftrightarrow v)\mathbf{P}_p(v \leftrightarrow y) &\leq \frac{(1+o(1))\mathbb{E}_p|\mathscr{C}(0)|}{V}\sum_{u,v}\mathbf{P}_p(u \leftrightarrow v)\mathbf{P}_p(v \leftrightarrow y) \\ &= \frac{(1+o(1))(\mathbb{E}_p|\mathscr{C}(0)|)^3}{V}\,,\end{aligned} \tag{2.7}$$

where we have used Lemma **??** for the first inequality. Thus, we are only left to deal with short connections:

$$\nabla_p(x,y) \leq \sum_{u,v}\mathbf{P}_p(x \xleftrightarrow{T_{\rm mix}} u)\mathbf{P}_p(u \xleftrightarrow{T_{\rm mix}} v)\mathbf{P}_p(v \xleftrightarrow{T_{\rm mix}} y) + O(\chi(p)^3/V)\,. \tag{2.8}$$

We write

$$\mathbf{P}_p(x \xleftrightarrow{T_{\rm mix}} u) = \sum_{t_1=0}^{T_{\rm mix}}\mathbf{P}_p(x \xleftrightarrow{=t_1} u)\,, \tag{2.9}$$

and do the same for all three terms so that

$$\nabla_p(x,y) \leq \sum_{u,v}\sum_{t_1,t_2,t_3}^{T_{\rm mix}}\mathbf{P}_p(x \xleftrightarrow{=t_1} u)\mathbf{P}_p(u \xleftrightarrow{=t_2} v)\mathbf{P}_p(v \xleftrightarrow{=t_3} y) + O(\chi(p)^3/V)\,. \tag{2.10}$$

We bound

$$\mathbf{P}_p(x \xleftrightarrow{=t_1} u) \leq m(m-1)^{t_1-1}\mathbf{p}^{t_1}(x,u)p^{t_1}\,, \tag{2.11}$$

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

$$\nabla_p(x,y) \leq \frac{m^3}{(m-1)^3}\sum_{u,v}\sum_{t_1,t_2,t_3}^{T_{\rm mix}}[p(m-1)]^{t_1+t_2+t_3}\mathbf{p}^{t_1}(x,u)\mathbf{p}^{t_2}(u,v)\mathbf{p}^{t_3}(v,y)+O(\chi(p)^3/V)\,. \tag{2.12}$$

Since $p \le p_c$, assumption (2) gives that $[p(m-1)]^{t_1+t_2+t_3} = 1+o(1)$, and it is a simple consequence of condition (3) that

$$\sum_{u,v} \sum_{t_1,t_2,t_3}^{T_{\text{mix}}} [p(m-1)]^{t_1+t_2+t_3} \mathbf{p}^{t_1}(x,u)\mathbf{p}^{t_2}(u,v)\mathbf{p}^{t_3}(v,y) \le \mathbf{1}_{\{x=y\}} + o(1)\,, \tag{2.13}$$

where $o(1)$ vanishes as $m \to \infty$, concluding the proof. □

2.4 Sprinkling and improved sprinkling

The sprinkling technique was invented by Ajtai, Komlós and Szemerédi [**?**] 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 $p_2 \ge \theta\varepsilon/m$ 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 an Azuma-Hoeffding concentration inequality to obtain 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 $C(c_2) > 0$. (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 which are contained in a component of size at least $2^{c_1 m}$ in G_{p_1} so that $V' \ge c_2 2^m$. We say that *sprinkling fails* when $|\mathscr{C}_1| \le c_2 2^m/3$ in the union $G_{p_1} \cup G_{p_2}$. If sprinkling fails, then we can partition $V' = A \uplus B$ such 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 which 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 that the paths guaranteed to exist by it are disjoint so that the edges in them are independent, the probability that sprinkling fails is at most

$$2^{2^m/2^{c_1 m}} \cdot \Big(1 - \big(\frac{\theta\varepsilon}{m}\big)^{C(c_2)\sqrt{m}}\Big)^{2^m/m^{100}} = \mathrm{e}^{-2^{(1+o(1))m}}\,, \tag{2.14}$$

which tends to 0.

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. In [4], this is replaced by an argument showing that percolation clusters are "close" to uniform random sets of similar size, so that two large clusters share many closed edges with the property that if we open even *one* of them, then the two clusters connect.

Let us now describe the heuristics of our improved sprinkling argument. While previously we had paths of length $\sqrt{m}$ connecting the two clusters, now we will have paths of length precisely 1. The final line of our proof, replacing (**??**), will be

$$2^{2\varepsilon V/(k_m \varepsilon^{-2})} \cdot \Big(1 - \frac{\theta\varepsilon}{m}\Big)^{m\varepsilon^2 V} \le \mathrm{e}^{-\theta\varepsilon^3 V(1+o(1))}, \tag{2.15}$$

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 (**??**), this line is rather suggestive. We will obtain that whp $2\varepsilon V$ vertices are in components of size at least $k_m \varepsilon^{-2}$, explaining the $2^{2\varepsilon V/(k_m \varepsilon^{-2})}$ term in (**??**). The most difficult part in [4] is justifying the second term showing that for any partition of these vertices into two sets of size of order εV, the number of closed edges between them is at least $\varepsilon^2 mV$. This is the content of Theorem **??** which we describe in the next section. Not surprisingly, $\varepsilon^2 mV$ is the expected number of edges that two random uniform sets of size εV have between them. The improved sprinkling argument, based on Theorem **??** is then described in more detail in Section **??**.

2.5 Most large cluster share large boundary

Since this is the most technical part of the overview, at the expense of being precise, we have chosen to reduce the clutter of notation and suppress several parameters from the notation. We ignore several dependencies between parameters and the skeptical reader is welcomed to read the more precise overview presented in [4].

However, we do emphasize the role of two important parameters. We choose r and r_0 so that $r \gg \varepsilon_m^{-1}$ but just barely, and $r_0 \gg r$ in a way that will become clear later. For vertices x, y, define the random variable

$$\begin{aligned} S_{r+r_0}(x,y) = \Big|\Big\{(u,u') \in E(G) \colon \{x \overset{r+r_0}{\longleftrightarrow} u\} \circ \{y \overset{r+r_0}{\longleftrightarrow} u'\}, \\ |B_u(r+r_0)| \cdot |B_{u'}(r+r_0)| \le \varepsilon^{-2}(\mathbb{E}|B(r_0)|)^2\Big\}\Big|. \end{aligned} \tag{2.16}$$

The important part of the definition of $S_{r+r_0}(x,y)$ is the first requirement $\{x \overset{r+r_0}{\longleftrightarrow} u\} \circ \{y \overset{r+r_0}{\longleftrightarrow} u'\}$ (the second requirement is more technical). The edges in $S_{r+r_0}(x,y)$ are such that if turned on, they enforce a connection between x and y, thus merging $\mathscr{C}(x)$ and $\mathscr{C}(y)$, so it is natural to sprinkle them. Informally, a pair of vertices (x,y) is *good* when their clusters are large and $S_{r+r_0}(x,y)$ is large, so that their clusters have many edges between them. We make this quantitative in the following definition.

Definition 1 **((r, r_0)-good pairs)** We say that x, y are (r, r_0)-good if all of the following occur:

1. $\partial B_x(r) \neq \emptyset$, $\partial B_y(r) \neq \emptyset$ and $B_x(r) \cap B_y(r) = \emptyset$,

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

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

Theorem 9 (Most large clusters share many boundary edges) *Let G be a graph on V vertices and degree m satisfying the assumptions of Theorem **??**. Assume that ε_m satisfies $\varepsilon_m \gg V^{-1/3}$ and $\varepsilon_m = o(T_{\rm mix}^{-1})$. Then,*

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

In light of Theorem **??**, we expect that the number of pairs of vertices (x,y) with $|\mathscr{C}(x)| \ge (\varepsilon_m^3 V)^{1/4}\varepsilon_m^{-2}$ and $|\mathscr{C}(y)| \ge (\varepsilon_m^3 V)^{1/4}\varepsilon_m^{-2}$ is close to $(2\varepsilon_m V)^2$. Theorem **??** shows that almost all of these pairs have clusters that share many edges between them. Theorem **??** allows us to prove Theorem **??**, as we describe in more detail in the next section.

The difficulty in Theorem **??** is the requirement (3) in Definition **??**. Indeed, conditioned on survival (that is, on $\partial B_x(r) \neq \emptyset$, $\partial B_y(r) \neq \emptyset$ and that the balls are disjoint), the random variable $S_{r+r_0}(x,y)$ is not concentrated and hence it is hard to prove that it is large. In fact, even the variable $|B(r_0)|$ is not concentrated. This is not a surprising fact: 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. Non-concentration occurs because the first generations of the process have a strong and lasting effect on the future of the population. In [4], we counteract this non-concentration by conditioning on the whole structure of $B_x(r)$ and $B_y(r)$. Since r is bigger than the correlation length ($r \gg \varepsilon_m^{-1}$), under this conditioning the variable $S_{r+r_0}(x,y)$ is concentrated (as one would expect from the branching process analogy).

In the next section, we explain how we apply Theorem **??** to improve the sprinkling argument and thus prove Theorem **??**(b). The reader that is eager to get to the proof of Theorem **??** may skip directly to Section **??**.

2.6 Improved sprinkling: Proof of Theorem **??**(b).

Recall that we have already proved the upper bound on $|\mathscr{C}_1|$ below Lemma **??**, so it remains to show that

$$\mathbf{P}_p\big(|\mathscr{C}_1| \ge (2-o(1))\varepsilon_m V\big) = 1 - o(1)\,. \tag{2.17}$$

Recall that $p = p_c(1+\varepsilon_m)$ is our percolation probability and choose p_1, p_2 satisfying

$$p_2 = \theta\varepsilon_m/m\,, \qquad p_c(1+\varepsilon_m) = p_1 + (1-p_1)p_2\,, \tag{2.18}$$

where $\theta > 0$ tends to 0 extremely slowly so that $p_1 = [1+(1-o(1))\varepsilon_m)]p_c$. Denote G_{p_1} and G_{p_2} as before. We first invoke Theorem **??** in G_{p_1} and deduce that whp

$$P_{r,r_0} = (1-o(1))4\varepsilon_m^2 V^2\,. \tag{2.19}$$

Now we wish to show that when we "sprinkle" this configuration in G_{p_1}, that is, when we add to the configuration independent p_2-open edges, most of these vertices join together to form one cluster of size roughly $2\varepsilon_m V$. We construct an auxiliary simple graph H with vertex set

$$V(H) = \left\{x \in G_{p_1} : |\mathscr{C}(x)| \geq (\varepsilon_m^3 V)^{1/4} \varepsilon_m^{-2}\right\},$$

and edge set

$$E(H) = \left\{(x,y) \in V(H)^2 : x, y \text{ are } (r, r_0)\text{-good}\right\}.$$

Lemma **??** and (**??**) now imply that whp H is almost the complete graph, that is

$$|V(H)| = (2+o(1))\varepsilon_m V, \qquad |E(H)| = (1-o(1))4\varepsilon_m^2 V^2. \tag{2.20}$$

Denote $v = |V(H)|$ so that $v = (2+o(1))\varepsilon_m V$ 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 (**??**) 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 $|M_1| \geq \Omega(\varepsilon_m V)$ and $|M_2| \geq \Omega(\varepsilon_m V)$ such that there is no open path in $G_{p_1} \cup G_{p_2}$ connecting a vertex in M_1 with 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 $(2-o(1))\varepsilon_m V$.

To show this, we first note that the number of such partitions is at most $2^{3(\varepsilon_m^3 V)^{3/4}}$ since $|\mathscr{C}(x_i)| \geq (\varepsilon_m^3 V)^{1/4}\varepsilon_m^{-2}$. Secondly, given such a partition consisting of M_1 and M_2, we claim that the number of edges $(u,u') \in E(H)$ 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 $\Omega(\varepsilon_m^2 V m)$. To see this, we consider the set of edges in H for which both sides lie in either M_1 or M_2 (more precisely, the vertices of H corresponding to M_1 and M_2). This number is clearly at most

$$M_1^2 + M_2^2 \leq (4 - \Omega(1))\varepsilon^2 V.$$

Hence, by (**??**), the number of edges in H such that one end is in M_1 and the other in M_2 is at least $\Omega(\varepsilon_m^2 V^2)$. In other words, there are at least $c\varepsilon_m^2 V^2$ pairs $(x,y) \in M_1 \times M_2$ such that $S_{r+r_0}(x,y) \geq cV^{-1} m\varepsilon_m^{-2}(\mathbb{E}|B(r_0)|)^2$. We choose r_0 so that this is a large number. In total, we counted at least order $\varepsilon^2 V^2 \cdot V^{-1} m \varepsilon_m^{-2}(\mathbb{E}|B(r_0)|)^2$ edges (u,u') and no edge is counted more than $|B_u(r+r_0)| \cdot |B_{u'}(r+r_0)|$ times, which is at most order $\varepsilon_m^{-2}(\mathbb{E}|B(r_0)|)^2$ by the definition of $S_{r+r_0}(x,y)$ and the second claim follows.

Hence if $|\mathscr{C}_1| \leq (2-\Omega(1))\varepsilon V$ after the sprinkling, 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

$$2^{3(\varepsilon_m^3 V)^{3/4}}(1-p_2)^{c\varepsilon_m^2 V m} = o(1),$$

since $p_2 = \theta\varepsilon_m/m$ and θ goes to 0 very slowly. This establishes the required estimate on $|\mathscr{C}_1|$.

We now use (**??**) to show the required bounds on $\mathbb{E}|\mathscr{C}(0)|$ and $|\mathscr{C}_2|$. The upper bound $\mathbb{E}|\mathscr{C}(0)| \leq (4+o(1))\varepsilon_m^2 V$ is stated in Lemma **??** and the lower bound follows immediately from our estimate on $|\mathscr{C}_1|$, since

$$\mathbb{E}|\mathscr{C}(0)| = V^{-1}\sum_{v\in V(G)} \mathbb{E}|\mathscr{C}(v)| = V^{-1}\sum_{j\geq 1}\mathbb{E}|\mathscr{C}_j|^2 \geq V^{-1}\mathbb{E}|\mathscr{C}_1|^2 \geq (4-o(1))\varepsilon_m^2 V,$$

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 (**??**). Furthermore, by this inequality and Lemma **??**, we deduce that

$$\sum_{j\geq 2} \mathbb{E}|\mathscr{C}_j|^2 = o(\varepsilon_m^2 V^2)\,,$$

and hence $|\mathscr{C}_2| = o(\varepsilon_m V)$ whp. This concludes the proof of Theorem **??**. □

3 Unlacing hypercube percolation: Proof of Theorem ??

The main result in this section is the following proposition:

Proposition 1 (Expectation of intrinsic balls) *Consider edge percolation on the hypercube* $\{0,1\}^m$ *with* $p = [1+5/(2m^2)+K/m^3]/(m-1)$ *for some* $K>0$ *sufficiently large. Then, for m sufficiently large there exists a* $k \geq 1$ *such that*

$$\mathbb{E}|B(k)| \geq 2^{m/2}/m^3. \tag{3.1}$$

Consequently, $p_c \leq [1+5/(2m^2)+K/m^3]/(m-1) = 1/m+1/m^2+7/(2m^3)+\Theta(1/m^4)$.

The proof uses very elementary estimates on the non-backtracking random walk transition probabilities. For completeness we provide here the crude bounds that we will use, and remark that much more precise bounds are available in [**?**].

Lemma 3 (NBW computations) *Let* $e_1 = (1,0,\ldots,0) \in \{0,1\}^m$ *and* $e_{1,1} = (1,1,0,\ldots,0) \in \{0,1\}^m$ *be hypercube vertices. Then*

$$\mathbf{p}^2(0,e_{1,1}) = \frac{2}{m(m-1)}\,,$$

and for any fixed $t_0 \geq 2$ *there exists* $C = C(t_0) < \infty$ *such that for all* $t \geq t_0$,

$$\mathbf{p}^{2t}(0,e_{1,1}) \leq Cm^{-t_0-1}\,.$$

Furthermore,

$$\mathbf{p}^3(0,e_1) = \frac{1}{m(m-1)}\,,$$

and for any fixed $t_0 \geq 2$ *there exists* $C = C(t_0) < \infty$ *such that for all* $t \geq t_0$,

$$\mathbf{p}^{2t+1}(0,e_1) \leq Cm^{-t_0-1}\,.$$

Proof The equality involving $\mathbf{p}^2(0,e_{1,1})$ is immediate since the probability that the non-backtracking walk takes any one of the two paths of length two from 0 to $e_{1,1}$ is $[m(m-1)]^{-1}$. For the second inequality, denote by $X_t \in \{0,1\}^m$ the location of non-backtracking random walk after t steps and by N_t the number of 1's in X_t. First note that by symmetry,

$$\mathbf{P}(X_{2t} = e_{1,1} \mid X_0 = 0, N_{2t} = 2) = \frac{2}{m(m-1)}\,.$$

So let us estimate the probability that $N_{2t} = 2$. If $N_t = k$ and in the last step a 1 was turned in 0, then the probability of $N_{t+1} = k+1$ is $(m-k-1)/(m-1)$, and if in the last step a 0 was turned into 1, then this probability is $(m-k)/(m-1)$. This reasoning shows that the process $\{N_t\}_{t\geq 0}$ satisfies

$$\mathbf{P}(N_{t+1} = N_t - 1 \mid N_t = k) \leq \frac{k+1}{m-1}\,. \tag{3.2}$$

The event $\{N_{2t} = 2\}$ implies that $N_{2t-2t_0+k} \leq 2t_0 + 2$ for all $0 \leq k \leq 2t_0$. For any $j \leq 2t_0 + 2$ and any sequence $(\varepsilon_1, \ldots, \varepsilon_{2t_0}) \in \{-1,1\}^{2t_0}$ such that $j + \varepsilon_1 + \ldots + \varepsilon_{2t_0} = 2$ we must have that at least $t_0 - 1$ of the ε_i's are -1 and $j + \varepsilon_1 + \ldots + \varepsilon_i \leq 4t_0 + 2$ for all $i = 1, \ldots, 2t_0$. Hence, by iterating (**??**) we get that

$$\mathbf{P}\Big(N_{2t-2t_0} = j, N_{2t-2t_0+1} = j+\varepsilon_1, N_{2t-2t_0+2} = j+\varepsilon_1+\varepsilon_2, \ldots, N_{2t} = 2\Big) \leq \Big(\frac{4t_0+3}{m-1}\Big)^{t_0-1}.$$

The number of j's and sequences $(\varepsilon_1, \ldots, \varepsilon_{2t_0})$ is at most $(2t_0+2)4^{t_0}$ so we get that

$$\mathbf{P}(N_{2t} = 2 \mid X_0 = 0) \leq (2t_0+2)4^{t_0}\Big(\frac{4t_0+3}{m-1}\Big)^{t_0-1} \leq Cm^{-t_0+1}\,,$$

where $C = C(t_0) < \infty$. This and the previous estimate concludes the proof of the bound on $\mathbf{p}^{2t}(0, e_{1,1})$. The equality $\mathbf{p}^3(0, e_1) = [m(m-1)]^{-1}$ stems from the fact that there are precisely $(m-1)$ non-backtracking paths of length 3 from 0 to e_1, and the probability of taking each is $[m(m-1)^2]^{-1}$. The bound on $\mathbf{p}^{2t+1}(0, e_1)$ is performed almost identically to the bound on $\mathbf{p}^{2t}(0, e_{1,1})$, we omit the details.

We prove recursive bounds on $\mathbb{E}|\partial B(k)|$ that form the key ingredient in the proof of Proposition **??**. Before doing so, we recall some notation. For a subset of vertices A, we say that an event $\mathcal{M}$ occurs *off* A, intuitively, if it occurs in $G_p \setminus A$. Formally, for a percolation configuration ω and a set of vertices A, we write ω_A for the configuration obtained from ω by turning all the edges touching A to closed. The event "$\mathcal{M}$ occurs off A" is defined to be $\{\omega\colon \omega_A \in \mathcal{M}\}$. 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 being $A = B_0(r)$ for some $r \geq 1$. In this case, the event $\{\mathcal{M}$ occurs off $A(\omega)\}$ is defined to be

$$\{\mathcal{M} \text{ occurs off } A(\omega)\} = \{\omega\colon \omega_{A(\omega)} \in \mathcal{M}\}\,. \tag{3.3}$$

For this example, we shall rely on the fact that, for an arbitrary event $\mathcal{M}$ and $A = B_x(s)$ (see [4, (3.1)]),

$$\mathbf{P}(\mathcal{M} \text{ off } B_x(s)) = \sum_A \mathbf{P}(B_x(s) = A)\mathbf{P}(\mathcal{M} \text{ off } A)\,. \tag{3.4}$$

Lastly, let us recall the van den Berg-Kesten-Reimer (BKR) inequality, see [**?**,**?**,**?**]. For a set of edges B we say that an event E *occurs on* B if and only if it occurs independently of the status of the edges not in B, i.e., it is the event

$$\big\{\omega \in E\colon \forall \omega', \omega' = \omega \text{ on } B \Rightarrow \omega' \in E\big\}\,,$$

and denote this event by $E_{|B}$. For two events E, F, we let $E \circ F$ denote the event

$$E \circ F = \{\omega : \exists B_1, B_2 \subset E(G), B_1 \cap B_2 = \emptyset, \omega \in E_{|B_1} \cap F_{|B_2}\}.$$

We will refer to the random sets of edges B_1, B_2 as *witnesses* for the events E and F, respectively. The BKR-inequality states that

$$\mathbf{P}_p(E \circ F) \le \mathbf{P}_p(E)\mathbf{P}_p(F). \tag{3.5}$$

Lemma 4 (Recursive bounds on $\mathbb{E}|\partial B(k)|$) *For any $c > 0$ there exists $K > 0$ such that if*

$$p = \frac{1 + 5/(2m^2) + K/m^3}{m-1},$$

then, for $m = m(K)$ sufficiently large and for any $k \ge 1$ satisfying $\mathbb{E}|B(k)| \le 2^{m/2}/m^3$,

$$\mathbb{E}|\partial B(k)| \ge [1 + c/m^3]\mathbb{E}|\partial B(k-1)|. \tag{3.6}$$

Proof We prove the claim by induction on k. Given $c > 0$ we will choose K to be large at the end of the proof — this choice will not depend on k or m. Given K, we choose m so large that

$$p = (1 + 5/(2m^2) + K/m^3)/(m-1) \le (1 + 10/m^2)/(m-1),$$

so our upper bound on p is independent of K.

We start by initializing the induction. We have that $\mathbb{E}|\partial B(1)| = mp$, while $\mathbb{E}|\partial B(0)| = 1$, so that indeed (**??**) holds for $k = 1$ (for any $K > 0$ and $c > 0$). Let $k \ge 1$ such that $\mathbb{E}|B(k)| \le 2^{m/2}/m^3$ and assume the induction hypothesis holds for any $\ell \le k-1$.

We will now estimate the conditional expectation of $|\partial B(k)|$ given $B(k-1)$. To be precise, when we condition on $B(k-1)$ we condition on all the open and closed edges touching a vertex of $B(k-2)$ (observe that since the graph is bipartite there cannot be two vertices of $\partial B(k-1)$ that are connected by an edge). This allows us to calculate $B(k-1)$ and note that edges from $\partial B(k-1)$ to $\partial B(k)$ are not revealed. Given this information, for each vertex $x \in \partial B(k-1)$ the number of edges that we have not revealed any information on is precisely $m - \sum_{y:\, y\sim x} \mathbb{1}_{\{y \in B(k-2)\}}$. Hence,

$$\begin{aligned}\mathbb{E}\big[|\partial B(k)|\,\big|B(k-1)\big] \ge mp|\partial B(k-1)| \;-\; p \sum_{y:\, y\sim x} \mathbb{1}_{\{x\in\partial B(k-1), y\in B(k-2)\}} \\ -\; \frac{p^2}{2} \sum_{x,y:\, d(x,y)=2} \mathbb{1}_{\{x,y\in\partial B(k-1)\}},\end{aligned} \tag{3.7}$$

where the last term comes from subtracting the vertices of $|\partial B(k)|$ we counted more than once, which happens if they have more than one "ancestor" in $\partial B(k-1)$.

We take expectations in both sides and bound the two subtracted sums. We split the first sum according to whether the edge (x, y) is open or not. If (x, y) is open, then we must have that $y \in \partial B(k-2)$ and that (x, y) is open off $B(k-2)$ (in the sense of (**??**)). Otherwise, that is, if $x \in \partial B(k-1), y \in B(k-2)$ and the edge (x, y) is closed, then there exists $\ell \le k-3$ and a vertex w such that $\{0 \overset{=\ell}{\longleftrightarrow} w\} \circ P_{k-l-1}(w, x) \circ \{w \overset{k-\ell-2}{\longleftrightarrow} y\}$ occurs, where $P_n(w, x)$ denotes the event that there exists a path of open edges of length *precisely* n connecting w and x (this is a monotone event since the path is not necessarily the shortest path). Indeed, let γ_x and γ_y be two shortest paths connecting 0 to x and y, respectively and take w to be the last intersection of these paths and ℓ to be index of w in the paths (this has to be the same number for both paths since they are shortest paths). Note that since the edge (x,y) is closed, w has to be at distance at most $k-3$ from 0. Then the witness for the first event is $B(\ell)$ (that is, all the open and closed edges touching $B(\ell-1)$ and the two other witnesses are the parts of γ_x and γ_y starting at w and ending at x and y, respectively. Similarly for the second sum in (**??**), if $x,y \in \partial B(k-1)$, then there must exists $\ell \le k-2$ and a vertex w such that $\{0 \xleftrightarrow{=\ell} w\} \circ P_{k-l-1}(w,x) \circ P_{k-l-1}(w,y)$ occurs.

We sum over ℓ and w and use the BKR inequality (**??**) to obtain that

$$\mathbb{E}|\partial B(k)| \ge mp\mathbb{E}|\partial B(k-1)| - (\mathrm{I}) - (\mathrm{II}) - (\mathrm{III})\,, \tag{3.8}$$

where

$$(\mathrm{I}) = p \sum_{w,x: x\sim w} \mathbf{P}(0 \xleftrightarrow{=k-2} w, (x,w) \text{ is open off } B(k-2))\,,$$

$$(\mathrm{II}) = p \sum_{\ell=0}^{k-3} \sum_{w,x,y:\ x\sim y} \mathbf{P}(0 \xleftrightarrow{=\ell} w)\mathbf{P}(P_{k-\ell-1}(w,x) \circ \{w \xleftrightarrow{k-\ell-2} y\})\,,$$

$$(\mathrm{III}) = \frac{p^2}{2} \sum_{\ell=0}^{k-2} \sum_{w,x,y:\ d(x,y)=2} \mathbf{P}(0 \xleftrightarrow{=\ell} w)\mathbf{P}(P_{k-\ell-1}(w,x) \circ P_{k-\ell-1}(w,y))\,.$$

We start by bounding the sum (I) from above. By conditioning on $B(k-2)$, we may rewrite (I) as

$$(\mathrm{I}) = p\sum_{w} \sum_{A:0 \xleftrightarrow{=k-2} w} \mathbf{P}(B(k-2) = A) \sum_{x:\ x\sim w} \mathbf{P}((x,w) \text{ is open off } A \mid B(k-2) = A)\,.$$

Note that the probability that (x,w) is open off A equals p for any x such that $x \sim w$ and the edge (x,w) is not in A. Since A is such that $0 \xleftrightarrow{=\ell} w$ and $\mathbf{P}(B(k-2) = A) > 0$, we learn that there are at most $m-1$ such possible x's (instead of m, the total number of neighbors of w). Hence

$$(\mathrm{I}) \le (m-1)p^2\mathbb{E}|\partial B(k-2)| \le (m-1)p^2\mathbb{E}|\partial B(k-1)|\,, \tag{3.9}$$

where in the last line we used the induction hypothesis. We proceed by bounding the two sums (II) and (III) from above. We handle the sums separately according to whether $k-\ell-1 \le T_{\mathrm{mix}}$ or not. To that aim, we define

$$(\mathrm{II})_1 = p \sum_{\ell=0}^{k-1-T_{\mathrm{mix}}} \sum_{w,x,y:\ x\sim y} \mathbf{P}(0 \xleftrightarrow{=\ell} w)\mathbf{P}(P_{k-\ell-1}(w,x) \circ \{w \xleftrightarrow{k-\ell-2} y\})\,,$$

and

$$(\mathrm{II})_2 = p \sum_{\ell=k-1-T_{\mathrm{mix}}}^{k-3} \sum_{w,x,y:\ x\sim y} \mathbf{P}(0 \xleftrightarrow{=\ell} w)\mathbf{P}(P_{k-\ell-1}(w,x) \circ \{w \xleftrightarrow{k-\ell-2} y\})\,,$$

and similarly we define $(\mathrm{III})_1$ and $(\mathrm{III})_2$. Our convention is that if $k-1 \le T_{\mathrm{mix}}$, then $(\mathrm{I})_2 = (\mathrm{II})_2 = 0$.

It turns out that $(\mathrm{I})_1$ and $(\mathrm{II})_1$ contribute a negligible amount to (**??**). Indeed, when $k-\ell-1 \ge T_{\mathrm{mix}}$ we use Lemma **??** to bound

$$\mathbf{P}(P_{k-\ell-1}(w,x)) \le \frac{C\mathbb{E}|B(k-\ell-1)|}{2^m}.$$

We use this estimate and the BK inequality to bound

$$(\mathrm{II})_1 \le Cmp2^{-m} \sum_{\ell \le k-1-T_{\mathrm{mix}}} \mathbb{E}|\partial B(\ell)| \big(\mathbb{E}|B(k-\ell-1)|\big)^2.$$

We bound $\mathbb{E}|B(k-\ell-1)| \le \mathbb{E}|B(k)| \le m^{-3}2^{m/2}$ by our assumption on k to get that

$$(\mathrm{II})_1 \le Cm^{-5}p \sum_{\ell \le k-1-T_{\mathrm{mix}}} \mathbb{E}|\partial B(\ell)| \le Cm^{-5}p \sum_{\ell \le k-1-T_{\mathrm{mix}}} \frac{\mathbb{E}|\partial B(k-1)|}{[1+cm^{-3}]^{k-\ell-1}},$$

where the last inequality is due to our induction hypothesis. This yields the bound

$$(\mathrm{II})_1 \le Cm^{-2}p\mathbb{E}|\partial B(k-1)| \le Cm^{-3}\mathbb{E}|\partial B(k-1)|, \tag{3.10}$$

since $p = O(m^{-1})$ and $C > 0$ may depend on c. An almost identical calculation gives that

$$(\mathrm{III})_1 \le Cm^{-3}\mathbb{E}|\partial B(k-1)|. \tag{3.11}$$

Bounding $(\mathrm{II})_2$ and $(\mathrm{III})_2$ is more delicate and the local structure of the hypercube comes into play. Let us start by bounding $(\mathrm{III})_2$ since it is slightly simpler. We start by bounding

$$\mathbf{P}(P_{k-\ell-1}(w,x) \circ P_{k-\ell-1}(w,y)) \le m(m-1)^{2k-2\ell-2}p^{2k-2\ell-2}\mathbf{p}^{k-\ell-1,k-\ell-1}(x,w,y),$$

where $\mathbf{p}^{t_1,t_2}(x,w,y)$ is the probability that a non-backtracking random walk starting from x visits w at time t_1 and visits y at time t_1+t_2. The reason for this bound is that if the event on the left hand side occurs, then there exists a simple open path of length precisely $2k-2\ell-2$ from x to y going through w at time $k-\ell-1$. The number of such paths is bounded above by $m(m-1)^{2k-2\ell-2}\mathbf{p}^{k-\ell-1,k-\ell-1}(x,w,y)$ and the estimate follows by the union bound. By transitivity,

$$\begin{aligned}(\mathrm{III})_2 \le \frac{(1+O(m^{-1}))p^2}{2} &\sum_{\ell=k-1-T_{\mathrm{mix}}}^{k-2} \mathbb{E}|\partial B(\ell)| \\ &\times \sum_{x,y:\, d(x,y)=2} [(m-1)p]^{2k-2\ell-2}\mathbf{p}^{k-\ell-1,k-\ell-1}(x,0,y).\end{aligned}$$

Note that the sum over x, y in the right hand side does not depend on the 0, so we may rewrite this sum as

$$\begin{aligned}
&\sum_{x,y:\ d(x,y)=2} [(m-1)p]^{2k-2\ell-2}\mathbf{p}^{k-\ell-1,k-\ell-1}(x,0,y) \\
&\qquad = 2^{-m} \sum_{x,y:\ d(x,y)=2} \sum_{v} [(m-1)p]^{2k-2\ell-2}\mathbf{p}^{k-\ell-1,k-\ell-1}(x,v,y) \\
&\qquad = 2^{-m}[(m-1)p]^{2k-2\ell-2} \sum_{x,y:\ d(x,y)=2} \mathbf{p}^{2k-2\ell-2}(x,y) \\
&\qquad = \frac{m(m-1)}{2}[(m-1)p]^{2k-2\ell-2}\mathbf{p}^{2k-2\ell-2}(0,e_{1,1}),
\end{aligned}$$

where in the last equality we used the fact that on the hypercube, $\mathbf{p}^t(x,y)$ is the same for any pair x, y such that $d(x,y) = 2$, and $e_{1,1}$ is the hypercube vector $(1,1,0,\ldots,0)$. We now use the induction hypothesis which implies that $\mathbb{E}|\partial B(\ell)| \le \mathbb{E}|\partial B(k-1)|$ to get the bound of

$$(\mathrm{III})_2 \le \frac{(1+O(m^{-1}))p^2 m(m-1)\mathbb{E}|\partial B(k-1)|}{4} \sum_{t=1}^{T_{\mathrm{mix}}} [(m-1)p]^{2t}\mathbf{p}^{2t}(0,e_{1,1}).$$

We now appeal to Lemma **??** and use the fact that $T_{\mathrm{mix}} = O(m\log m)$ and that $1 \le (m-1)p \le 1+10/m^2$. A straightforward calculation with these gives that

$$\sum_{t=1}^{T_{\mathrm{mix}}} [(m-1)p]^{2t}\mathbf{p}^{2t}(0,e_{1,1}) = \frac{2+O(m^{-1})}{m(m-1)}. \tag{3.12}$$

Thus,

$$(\mathrm{III})_2 \le \frac{p^2}{2}(1+O(m^{-1}))\mathbb{E}|\partial B(k-1)|. \tag{3.13}$$

We proceed by bounding $(\mathrm{II})_2$. We begin by estimating

$$\mathbf{P}(P_{k-\ell-1}(w,x) \circ \{w \xleftrightarrow{k-\ell-2} y\}) \le \sum_{s=0}^{k-\ell-2} \mathbf{P}(P_{k-\ell-1}(w,x) \circ P_s(w,y)),$$

and further bound, for each s,

$$\mathbf{P}(P_{k-\ell-1}(w,x) \circ P_s(w,y)\}) \le m(m-1)^{k-\ell-1+s-1}p^{k-\ell-1+s}\mathbf{p}^{k-\ell-1,s}(x,w,y),$$

where $\mathbf{p}^{t_1,t_2}(x,w,y)$ was defined earlier, and the reasoning for this bound is as before. We get

$$\begin{aligned}
(\mathrm{II})_2 \le (1+O(m^{-1}))p &\sum_{\ell=k-1-T_{\mathrm{mix}}}^{k-3} \mathbb{E}|\partial B(l)| \\
&\times \sum_{s=0}^{k-\ell-2} [(m-1)p]^{k-\ell-1+s} \sum_{x,y:\ x\sim y} \mathbf{p}^{k-\ell-1,s}(x,0,y).
\end{aligned} \tag{3.14}$$

As before, the sum over x, y does not depend on 0, that is

$$\begin{aligned}\sum_{x,y\colon x\sim y}\mathbf{p}^{k-\ell-1,s}(x,0,y) &= 2^{-m}\sum_{v}\sum_{x,y\colon x\sim y}\mathbf{p}^{k-\ell-1,s}(x,v,y)\\ &= 2^{-m}\sum_{x,y\colon x\sim y}\mathbf{p}^{k-\ell-1+s}(x,y) = m\mathbf{p}^{k-\ell-1+s}(0,e_1)\,,\end{aligned}$$

where e_1 is just the vector $(1,0,\ldots,0)$. As before we use the induction hypothesis to derive that $\mathbb{E}|\partial B(\ell)| \le \mathbb{E}|\partial B(k-1)|$ to get that

$$\begin{aligned}(\mathrm{II})_2 &\le (1+O(m^{-1}))mp\mathbb{E}|\partial B(k-1)|\\ &\times\sum_{\ell=k-1-T_{\mathrm{mix}}}^{k-3}\sum_{s=0}^{k-\ell-2}[(m-1)p]^{k-\ell-1+s}\mathbf{p}^{k-\ell-1+s}(0,e_1)\,.\end{aligned}\tag{3.15}$$

Denote $t = k-\ell-1+s$ and note that $t \le T_{\mathrm{mix}}$ and $t \ge 2$. By parity $\mathbf{p}^2(0,e_1) = 0$, so in fact $t \ge 3$. Further, for fixed t and k, there are $t-1$ pairs of (ℓ, s) with $3 \le k-\ell+s-1 \le t+1$ such that $k-\ell+s-1 = t$. Therefore,

$$(\mathrm{II})_2 \le (1+O(m^{-1}))mp\mathbb{E}|\partial B(k-1)|\sum_{t=3}^{2T_{\mathrm{mix}}}(t-1)[(m-1)p]^t\mathbf{p}^t(0,e_1)\,.$$

The dominant term here is $t = 3$. We appeal to Lemma **??**, the fact that $T_{\mathrm{mix}} = O(m\log m)$ and that $1 \le (m-1)p \le 1+10/m^2$ to obtain that

$$\sum_{t=3}^{2T_{\mathrm{mix}}}(t-1)[(m-1)p]^t\mathbf{p}^t(0,e_1) = \frac{2+O(m^{-1})}{m(m-1)}\,.\tag{3.16}$$

We get the bound

$$(\mathrm{II})_2 \le (1+O(m^{-1}))mp\mathbb{E}|\partial B(k-1)|[2m^{-2}+O(m^{-3})]\,.\tag{3.17}$$

Finally, we put this together with (**??**), (**??**), (**??**) and (**??**) into (**??**) to obtain

$$\mathbb{E}|\partial B(k)| \ge \big[mp-(m-1)p^2-2m^{-1}p-p^2/2-Cm^{-3}\big]\mathbb{E}|\partial B(k-1)|\,.\tag{3.18}$$

Here we stress that since $p \le (1+10/m^2)/(m-1)$, the constants hidden in the $O(\cdot)$ in (**??**)-(**??**) are *independent* of the constant K from the definition of p, which implies that also $C > 0$ may depend on $c > 0$ but not on K. We plug in the value of p and a straightforward calculation shows that we can choose $K > 0$ large enough such that

$$mp-(m-1)p^2-2m^{-1}p-p^2/2-Cm^{-3} \ge 1+cm^{-3}\,,$$

concluding our proof. □

Proof of Proposition **??**. By Lemma **??**, $\mathbb{E}|\partial B(k)| \ge (1+c/m^3)^k$ as long as $\mathbb{E}|B(k)| \le 2^{m/2}/m^3$. So, suppose by contradiction that there is no k such that $\mathbb{E}|B(k)| > 2^{m/2}/m^3$. Then, $\mathbb{E}|\partial B(k)| \ge (1+c/m^3)^k$ for every k, leading to a contradiction for k large enough. Hence, for k large enough, $\mathbb{E}|B(k)| \ge 2^{m/2}/m^3$, proving the claim. □

Acknowledgements This work was presented by RvdH on the occasion of the Stochastik Tage 2012, held in Mainz March 6-9, 2012.

References

1. Aizenman M. and Newman C.M. (1984), Tree graph inequalities and critical behavior in percolation models, J. Stat. Phys., 44: 393-454.
2. Ajtai M., Komlós J. and Szemerédi E. (1982), Largest random component of a k-cube. Combinatorica 2(1): 1–7.
3. Aldous, D. (1997). Brownian excursions, critical random graphs and the multiplicative coalescent. Ann. Probab. 25: 812–854.
4. Barsky D. and Aizenman M. (1991), Percolation critical exponents under the triangle condition. Ann. Probab. 19(4): 1520–1536.
5. van den Berg, J. and Kesten, H. (1985), Inequalities with applications to percolation and reliability. J. Appl. Prob. 22: 556–569.
6. Bollobás B. (1984), The evolution of random graphs. Trans. Amer. Math. Soc. 286: 257–274.
7. Bollobás, B. (1985), *Random Graphs.* Academic Press, London.
8. Bollobás B., Kohayakawa Y. and Łuczak T. (1992), The evolution of random subgraphs of the cube. Random Structures Algorithms 3(1): 55–90.
9. Borgs C., Chayes J.T., van der Hofstad R., Slade G. and Spencer J. (2005), Random subgraphs of finite graphs: I. The scaling window under the triangle condition. Random Structures Algorithms 27: 137–184.
10. Borgs C., Chayes J.T., van der Hofstad R., Slade G. and Spencer J. (2005), Random subgraphs of finite graphs: II. The lace expansion and the triangle condition. Ann. Probab. 33: 1886-1944.
11. Borgs C., Chayes J.T., van der Hofstad R., Slade G. and Spencer J. (2006), Random subgraphs of finite graphs: III. The phase transition for the n-cube. Combinatorica 26: 395-410.
12. Borgs C., Chayes, J. T. and Randall D. (1999), The van den Berg-Kesten-Reimer inequality: a review, *Perplexing problems in probability*, 159–173, Progr. Probab., 44, Birkhhäuser Boston, Boston, MA.
13. Clisby N., Liang R., and Slade G. (2007), Self-avoiding walk enumeration via the lace expansion. *J. Phys. A*, 40(36):10973–11017.
14. Clisby N. and Slade G. (2009), Polygons and the lace expansion. In *Polygons, polyominoes and polycubes*, volume 775 of *Lecture Notes in Phys.*, pages 117–142. Springer, Dordrecht.
15. Erdős P. and Rényi A. (1960), On the evolution of random graphs, Magyar Tud. Akad. Mat. Kutató Int. Kőzl. 5: 17–61.
16. Erdős P. and Spencer J. (1979), Evolution of the n-cube, Comput. Math. Appl. 5, 33-39.
17. Fitzner R. and van der Hofstad R. (2013), Non-backtracking random walk, J. Stat. Phys. 150(2): 264–284.
18. Gaunt D.S. and Ruskin H. (1978), Bond percolation processes in d dimensions. J. Phys. A: Math. Gen., 11:1369âĂŞ-1380.
19. Grimmett G. (1999), *Percolation.* Second edition. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 321. Springer-Verlag, Berlin.
20. Hara T. and Slade G. (1990), Mean-field critical behaviour for percolation in high dimensions. Comm. Math. Phys. 128: 333-391.
21. Heydenreich M. and van der Hofstad R. (2007), Random graph asymptotics on high-dimensional tori. Comm. Math. Phys. 270(2): 335-358.
22. Heydenreich M. and van der Hofstad R. (2011), Random graph asymptotics on high-dimensional tori II. Probability Theory and Related Fields, 149(3-4): 397-415.
23. van der Hofstad R. and Nachmias A. (2012). Hypercube percolation. Preprint available at `http://arxiv.org/abs/1201.3953`.
24. van der Hofstad, R. and Slade G. (2005), Asymptotic expansions in n^{-1} for percolation critical values on the n-cube and $\mathbb{Z}^n$. Random Structures Algorithms 27(3): 331–357.
25. van der Hofstad R. and Slade G. (2006), Expansion in n^{-1} for percolation critical values on the n-cube and $\mathbb{Z}^n$: the first three terms. Combin. Probab. Comput. 15(5): 695–713.
26. Janson S., Knuth D.E.,Łuczak T. and Pittel, B. (1993), The birth of the giant component. Random Structures and Algorithms 4: 71–84.
27. Janson S., Łuczak T. and Rucinski, A. (2000), *Random Graphs.* Wiley, New York.
28. Kozma, G. (2011), Percolation on a product of two trees, Annals of Probability 39(5):1864–1895.
29. Kozma G. and Nachmias A. (2009), The Alexander-Orbach conjecture holds in high dimensions, Invent. Math. 178(3): 635–654.
30. Kozma G. and Nachmias A. (2011), Arm exponents in high-dimensional percolation, Journal of the American Mathematical Society 24: 375-409.

31. Levin D., Peres, Y. and Wilmer, E. (2009), *Markov Chains and Mixing Times.* With a chapter by James G. Propp and David B. Wilson. American Mathematical Society, Providence, RI.
32. Łuczak T. (1990), Component behavior near the critical point of the random graph process. Random Structures Algorithms 1: 287–310.
33. Łuczak T., Pittel B. and Wierman, J.C. (1994), The structure of a random graph at the point of the phase transition. Trans. Amer. Math. Soc. 341:721–748.
34. Nachmias A. and Peres Y. (2008), Critical random graphs: diameter and mixing time, Ann. Probab. 36(4): 1267-1286.
35. Nachmias A. and Peres Y. (2012), Non-amenable Cayley graphs of high girth have $p_c < p_u$ and mean-field exponents, Preprint available at `http://arxiv.org/abs/1207.1480`.
36. Pittel, B. and Wormald, N. (2005), Counting connected graphs inside-out. J. Combin. Theory Ser. B, 93(2):127–172.
37. Reimer, D. (2000), Proof of the van den Berg-Kesten conjecture. Combin. Probab. Comput. 9(1):27–32.
38. Schonmann, R. (2001), Multiplicity of phase transitions and mean-field criticality on highly non-amenable graphs. Comm. Math. Phys. 219 (2): 271–322.