Cycle structure of percolation on high-dimensional tori

Remco van der Hofstad [*]   Artëm Sapozhnikov [†]

October 8, 2012

Abstract

In the past years, many properties of the critical behavior of the largest connected components on the high-dimensional torus, such as their sizes and diameter, have been established. The order of magnitude of these quantities equals the one for percolation on the complete graph or Erdős-Rényi random graph, raising the question whether the scaling limits or the largest connected components, as identified by Aldous (1997), are also equal.

In this paper, we investigate the *cycle structure* of the largest critical components for high-dimensional percolation on the torus $\{-\lfloor r/2 \rfloor, \ldots, \lceil r/2 \rceil - 1\}^d$. While percolation clusters naturally have many *short* cycles, we show that the *long* cycles, i.e., cycles that pass through the boundary of the cube of width $r/4$ centered around each of their vertices, have length of order $r^{d/3}$, as on the critical Erdős-Rényi random graph. On the Erdős-Rényi random graph, cycles play an essential role in the scaling limit of the large critical clusters, as identified by Addario-Berry, Broutin and Goldschmidt (2010).

Our proofs crucially rely on various new estimates of probabilities of the existence of open paths in critical Bernoulli percolation on $\mathbb{Z}^d$ with constraints on their lengths. We believe these estimates are interesting in their own right.

Key words: Random graph, phase transition, critical behavior, percolation, cycle structure.
MSC2000 subject classification. 05C80, 60K35, 82B43.

1 Introduction and results

In the past years, the investigation of percolation on various high-dimensional tori has attracted tremendous attention. In [4, 5], the phase transition of the largest connected component was investigated for percolation on general high-dimensional tori, including the complete graph, the hypercube in high dimensions, as well as finite-range percolation in sufficiently high dimensions. The phase transition of percolation on high-dimensional tori is *mean-field*, i.e., it shares many features with that on the complete graph as identified in [8] (see, e.g., [2, 3, 16, 17, 21]).

In [4], the subcritical and critical behavior was investigated under the so-called *triangle condition*, a general assumption on the underlying graph that ensures that the model is mean-field. The critical behavior of the model was identified in terms of the blow-up of the expected cluster size, which identifies a window of critical values of the edge occupation probabilities. For any parameter value in this critical window, the largest connected component was shown to be of order $V^{2/3}$, as on the complete graph, where V denotes the number of vertices in the graph. In [5], the triangle condition was proved to hold for the above-mentioned examples.

The situation of finite-range high-dimensional tori, which in the graph sense converge to the hypercubic lattice, was brought substantially further in [13, 14], where, among others, it was shown that the

---

[*]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`

[†]ETH Zürich, Department of Mathematics, Rämistrasse 101, 8092 Zürich. E-mail: `artem.sapozhnikov@math.ethz.ch`

percolation critical value on the infinite lattice lies inside the scaling window. We now know that the largest connected components are all of order $V^{2/3}$, that the maximal connected component $|\mathcal{C}_{\max}|$ satisfies that $|\mathcal{C}_{\max}|V^{-2/3}$ and $V^{2/3}/|\mathcal{C}_{\max}|$ are tight sequences of random variables that are non-concentrated, and that the diameter of large clusters is of order $V^{1/3}$. These results (and more) are also known to hold on the Erdős-Rényi random graph, see e.g., [2, 22], as well as the monographs [3, 17]. This raises the question whether the scaling limits agree. We shall expand on this question in Section 1.4 below.

1.1 Percolation in high dimensions

We consider bond percolation on a graph G. For a given parameter $p \in [0,1]$, this is the probability measure $\mathbb{P}_p$ on subgraphs of G defined as follows. We delete edges of G with probability $(1-p)$ and otherwise keep them, independently for different edges. The edges of the resulting random subgraph of G are called *open* and the deleted edges are called *closed*. Connected components of this random subgraph are called *open clusters*. The graphs we investigate in this paper are (a) the d-dimensional torus $\mathbb{T}_r^d = \{-\lfloor r/2 \rfloor, \ldots, \lceil r/2 \rceil - 1\}^d$; and (b) the hypercubic lattice $\mathbb{Z}^d$, where the dimension d is supposed to be sufficiently large. How large we need to take d depends on the edge structure of G. We consider two different settings: (a) In the *nearest-neighbor model*, two vertices are connected by an edge if they are nearest-neighbors on G. With our choice of G, every vertex has $2d$ nearest-neighbors. In this setting we take the dimension d large enough. (b) In the *spread-out model* with a parameter L, two vertices are connected by an edge if there is a hypercube of size L in G that contains these vertices. With our choice of G, every vertex has $(2L+1)^d - 1$ neighbors. Of course, we are only interested in the case when the size of the torus is much larger than L. In the spread-out setting with large enough L, we take the dimension $d > 6$.

To justify our choice of dimension, we recall a number of well-known results about percolation on $\mathbb{Z}^d$. For bond percolation on $\mathbb{Z}^d$ with $d > 1$, there exists a critical probability $p_c \in (0,1)$ such that, for $p < p_c$, all open clusters are almost surely finite and, for $p > p_c$, there is almost surely an infinite open cluster. At $p = p_c$, it is widely believed that there is almost surely no infinite open cluster. This fact has been shown for $d = 2$ by Kesten [18] and for sufficiently large d by Hara and Slade [12]. Here, by sufficiently large d, we mean $d > 18$ for the nearest-neighbor model and $d > 6$ for the spread-out model with sufficiently large L. Showing this for all $d > 1$ remains a challenging open problem.

The main assumption that we use in the paper concerns an estimate on the probability that, at criticality, two vertices x and y are in the same open cluster of bond percolation on $\mathbb{Z}^d$, which we denote by $x \leftrightarrow y$. We assume that there exist constants D_1 and D_2 such that, for all x and y in $\mathbb{Z}^d$,

$$D_1(1 + |x - y|)^{2-d} \leq \mathbb{P}_{p_c}(x \leftrightarrow y) \leq D_2(1 + |x - y|)^{2-d}. \tag{1.1}$$

These bounds have been established using so-called *lace-expansion* techniques, for the nearest-neighbor model with large enough d by Hara [10], and for the spread-out model with $d > 6$ by Hara, van der Hofstad and Slade [11]. In fact, these papers give asymptotic formulas for such probabilities, but for our purposes, the bounds (1.1) suffice.

It is believed that the estimates (1.1) hold for the nearest-neighbor model with $d > 6$, however the proof of this fact is beyond the current methods. It has been proved by Chayes and Chayes [7] (assuming the existence of critical exponents) that the bounds (1.1) are violated for $d < 6$. The dimension $d = d_c = 6$ is usually referred to as the *upper critical dimension*.

A simple computation using the upper bound in (1.1) shows that

$$\nabla(p_c) = \sum_{x,y} \mathbb{P}_{p_c}(0 \leftrightarrow x)\mathbb{P}_{p_c}(x \leftrightarrow y)\mathbb{P}_{p_c}(y \leftrightarrow 0) < \infty. \tag{1.2}$$

The bound in (1.2) is called the *triangle condition*, and is believed to be true for $d > 6$. The triangle condition implies that the sub- and critical phases of percolation on $\mathbb{Z}^d$ behave similarly to the ones on a tree, for example, many critical exponents on $\mathbb{Z}^d$ are equal to those on the tree. Intuitively, the geometry of large critical clusters trivializes, since the space is so vast that far away clusters are close

to being independent. In recent years, a related condition has been proved to hold on the torus, which implies that the critical behavior of large connected components on the high-dimensional torus is similar to that on the complete graph. Sometimes this is called *random graph asymptotics* for percolation on the high-dimensional torus.

In this paper, we study the *cycle structure* of bond percolation on the d-dimensional torus in the above two settings. Despite the fact that, for any $p \in (0, 1)$, the vertices in open cycles occupy a positive fraction of the torus, which is not the case for the critical and subcritical Erdős-Rényi random graph (see [21]), most of such vertices belong only to *short* cycles, such as open squares of four bonds. Short cycles vanish in the scaling limit of large critical clusters, and are thus irrelevant to the scaling limit. Therefore, we focus on the existence of open long cycles, where we say that a cycle is long when it passes through the boundary of the cube of width $r/4$ centered around each of its vertices. Special cases of long cycles are *non-contractible cycles*, which are cycles that cannot, when considered as continuous curves, be contracted to a point, and thus wind around the torus at least once.

Our main results show that the mean number of vertices in open long cycles grows like $V^{1/3}$, and that such cycles (when they exist) contain order of $V^{1/3}$ vertices. Moreover, we show that the probability of the existence of at least one open long cycle in a large cluster is bounded away from 0 and 1, uniformly in the volume of the graph. As we discuss in more details below, this situation is analogous to the situation on the complete graph, as investigated in [1, 2, 21]. We also refer the reader to [21, pages 722-723] for the discussion of more refined results about the structure of connected components of the critical Erdős-Rényi random graph.

For simplicity of presentation, *we restrict ourselves hereafter to the nearest-neighbor model.* The results of this paper still hold for the spread-out model on the d-dimensional torus with $d > 6$.

The remainder of this section is organized as follows. In Section 1.2, we describe our main results, in Section 1.3, we describe some results on critical percolation on $\mathbb{Z}^d$ and the torus that are used in the proofs of our main results and are interesting in their own right, and in Section 1.4, we discuss the results and their relation to the work on the Erdős-Rényi random graph.

1.2   Main results

We start by introducing some notation. For $a \in \mathbb{R}$, we write $|a|$ for the absolute value of a, and, for a site $x = (x_1, \ldots, x_d) \in \mathbb{Z}^d$, we write $|x|$ for $\max(|x_1|, \ldots, |x_d|)$, and $|x|_1$ for $\sum_{i=1}^d |x_i|$. For $s > 0$ and $x \in \mathbb{Z}^d$, let $Q_s(x) = \{y \in \mathbb{Z}^d \colon |y - x| \leq s\}$ and $\partial Q_s(x) = \{y \in \mathbb{Z}^d \colon |y - x| = \lfloor s \rfloor\}$. We write Q_s for $Q_s(0)$ and ∂Q_s for $\partial Q_s(0)$.

For a positive integer r, we consider the torus $(\mathbb{T}_r^d, \mathbb{E}_r^d)$ with $\mathbb{T}_r^d = \{-\lfloor r/2 \rfloor, \ldots, \lceil r/2 \rceil - 1\}^d$ and the edge set $\mathbb{E}_r^d = \{\{x, y\} \in \mathbb{T}_r^d \times \mathbb{T}_r^d \colon \sum_{i=1}^d |(x_i - y_i)(\mathrm{mod}\ r)| = 1\}$. We often abuse notation and write $\mathbb{T}_r^d$ for $(\mathbb{T}_r^d, \mathbb{E}_r^d)$. The vertex $0 = (0, \ldots, 0)$ is called the origin. We denote the number of vertices in the torus or *volume* by $V = r^d$. For $p \in [0, 1]$, we consider the probability space $(\Omega_{\mathbb{T},p}, \mathcal{F}_{\mathbb{T},p}, \mathbb{P}_{\mathbb{T},p})$, where $\Omega_{\mathbb{T},p} = \{0, 1\}^{\mathbb{E}_r^d}$, $\mathcal{F}_{\mathbb{T},p}$ is the σ-field generated by the cylinders of $\Omega_{\mathbb{T},p}$, and $\mathbb{P}_{\mathbb{T},p}$ is a product measure on $(\Omega_{\mathbb{T},p}, \mathcal{F}_{\mathbb{T},p})$, $\mathbb{P}_{\mathbb{T},p} = \prod_{e \in \mathbb{E}_r^d} \mu_e$, where μ_e is given by $\mu_e(\omega_e = 1) = 1 - \mu_e(\omega_e = 0) = p$, for vectors $(\omega_e)_{e \in \mathbb{E}_r^d} \in \Omega_{\mathbb{T},p}$. We write $\mathbb{E}_{\mathbb{T},p}$ for the expectation with respect to $\mathbb{P}_{\mathbb{T},p}$.

We further consider the hypercubic lattice $(\mathbb{Z}^d, \mathbb{E}^d)$, where the edge set is given by $\mathbb{E}^d = \{\{x, y\} \in \mathbb{Z}^d \times \mathbb{Z}^d \colon |x - y|_1 = 1\}$. Again, we often abuse notation and write $\mathbb{Z}^d$ for $(\mathbb{Z}^d, \mathbb{E}^d)$. For $p \in [0, 1]$, we consider a probability space $(\Omega_{\mathbb{Z},p}, \mathcal{F}_{\mathbb{Z},p}, \mathbb{P}_{\mathbb{Z},p})$, where $\Omega_{\mathbb{Z},p} = \{0, 1\}^{\mathbb{E}^d}$, $\mathcal{F}_{\mathbb{Z},p}$ is the σ-field generated by the finite-dimensional cylinders of $\Omega_{\mathbb{Z},p}$, and $\mathbb{P}_{\mathbb{Z},p}$ is the product measure on $(\Omega_{\mathbb{Z},p}, \mathcal{F}_{\mathbb{Z},p})$, $\mathbb{P}_{\mathbb{Z},p} = \prod_{e \in \mathbb{E}^d} \mu_e$, where μ_e is given by $\mu_e(\omega_e = 1) = 1 - \mu_e(\omega_e = 0) = p$, for vectors $(\omega_e)_{e \in \mathbb{E}^d} \in \Omega_{\mathbb{Z},p}$.

In both settings, we say that an edge e is *open* or *occupied* if $\omega_e = 1$, and e is *closed* or *vacant* if $\omega_e = 0$.

The event that two sets of sites $\mathcal{K}_1, \mathcal{K}_2 \subset \mathbb{T}_r^d$ are connected by an open path is denoted by $\{\mathcal{K}_1 \leftrightarrow \mathcal{K}_2 \text{ in } \mathbb{T}_r^d\}$, and the event that $\mathcal{K}_1$ and $\mathcal{K}_2$ are connected by an open path of length (number of edges) at most k is denoted by $\{\mathcal{K}_1 \xleftrightarrow{\leq k} \mathcal{K}_2 \text{ in } \mathbb{T}_r^d\}$. We write $\mathcal{C}_{\mathbb{T}}(x)$ for the set of $y \in \mathbb{T}_r^d$ such that $x \leftrightarrow y$ in $\mathbb{T}_r^d$.

Similarly, the event that two sets of sites $\mathcal{K}_1, \mathcal{K}_2 \subset \mathbb{Z}^d$ are connected by an open path is denoted by $\{\mathcal{K}_1 \leftrightarrow \mathcal{K}_2 \text{ in } \mathbb{Z}^d\}$, and the event that $\mathcal{K}_1$ and $\mathcal{K}_2$ are connected by an open path of length at most k is denoted by $\{\mathcal{K}_1 \xleftrightarrow{\leq k} \mathcal{K}_2 \text{ in } \mathbb{Z}^d\}$. We write $\mathcal{C}_\mathbb{z}(x)$ for the set of $y \in \mathbb{Z}^d$ such that $x \leftrightarrow y$ in $\mathbb{Z}^d$.

Finally, for $x, y \in \mathbb{T}_r^d$, we write $\tau_{\mathbb{T},p}(x,y) = \mathbb{P}_{\mathbb{T},p}(x \leftrightarrow y \text{ in } \mathbb{T}_r^d)$, and $\tau_{\mathbb{T},p}(x) = \tau_{\mathbb{T},p}(0,x)$, while, for $x, y \in \mathbb{Z}^d$, $\tau_{\mathbb{Z},p}(x,y) = \mathbb{P}_{\mathbb{Z},p}(x \leftrightarrow y \text{ in } \mathbb{Z}^d)$ and $\tau_{\mathbb{Z},p}(x) = \tau_{\mathbb{Z},p}(0,x)$.

We call a nearest-neighbor path $\pi = (x(1), \ldots, x(m))$ in $\mathbb{T}_r^d$ a *cycle* if it is edge-disjoint and $x(1) = x(m)$, i.e., π is an (edge-)self-avoiding polygon. We say that a cycle π is *long* if for each $1 \leq n \leq m$, the cycle π has a vertex in $\partial Q_{r/4}(x(n))$. Finally, we denote by LC_k the event that the origin is in an open long cycle of length at most k.

For two functions g and h from a set $\mathcal{X}$ to $\mathbb{R}$, we write $g(z) \asymp h(z)$ to indicate that $g(z)/h(z)$ is bounded away from 0 and ∞, uniformly in $z \in \mathcal{X}$. All the constants (C_i) in the proofs are strictly positive and finite and depend only on the dimension, unless the dependence on other parameters is explicitly stated. Their exact values may be different from section to section.

We first give bounds on the probability that a vertex of the torus is in an open long cycle.

Theorem 1.1 (Expected number of vertices in long cycles)**.** *Assume (1.1). For $x \in \mathbb{T}_r^d$,*

$$\mathbb{P}_{\mathbb{T},p_c}(x \text{ is in an open long cycle}) \asymp V^{-2/3}. \tag{1.3}$$

Consequently,

$$\mathbb{E}_{\mathbb{T},p_c}\big[\#\{x \colon x \text{ is in an open long cycle}\}\big] \asymp V^{1/3}. \tag{1.4}$$

In the next theorem we show that, with high probability, large open clusters of the torus may only contain long cycles of length of order $V^{1/3}$.

Theorem 1.2 (Long cycles have length of order $V^{1/3}$)**.** *Assume (1.1). There exists $C < \infty$ such that for any positive ε and δ, and integer $r \geq 1$,*

$$\mathbb{P}_{\mathbb{T},p_c}\left(\exists x \colon |\mathcal{C}_\mathbb{T}(x)| > \delta V^{2/3}, \ \mathcal{C}_\mathbb{T}(x) \text{ contains a long cycle of length } \leq \varepsilon V^{1/3}\right) \leq \frac{C\varepsilon}{\delta}, \tag{1.5}$$

and

$$\mathbb{P}_{\mathbb{T},p_c}\left(\exists \text{ a long cycle of length } \geq \varepsilon^{-1} V^{1/3}\right) \leq C\varepsilon. \tag{1.6}$$

We next study the number of long cycles. We start by defining what this is. For a subgraph G of the torus, we define Y_G as the smallest k for which there exist edges $e_1, \ldots, e_k$ in G such that $G \setminus \{e_1, \ldots, e_k\}$ does not contain any long cycles. For $\delta > 0$, we define

$$Y_\delta = \sum_\mathcal{C} Y_\mathcal{C} I(|\mathcal{C}| > \delta V^{2/3}),$$

where the sum is over all open clusters $\mathcal{C}$ of the torus. We prove the following theorems:

Theorem 1.3. *Assume (1.1). There exists $C < \infty$ such that for all $\delta > 0$ and integer $r \geq 1$,*

$$\mathbb{E}_{\mathbb{T},p_c}[Y_\delta] \leq C/\delta.$$

In particular, the random variables Y_δ are tight.

Theorem 1.4 (Non-trivial existence of long cycles)**.** *Assume (1.1).*
(a) There exists $c > 0$ such that for all integers $r \geq 1$,

$$\mathbb{P}_{\mathbb{T},p_c}\left(\exists \text{ a long cycle of length } > c V^{1/3}\right) > c.$$

(b) For any $\delta > 0$ there exists $c > 0$ such that for all integers $r \geq 1$,

$$\mathbb{P}_{\mathbb{T},p_c}(Y_\delta = 0) > c.$$

In other words, with positive probability uniformly in r, there are no long cycles in clusters of size $> \delta V^{2/3}$.

4

1.3   Related results on critical percolation

In this section, we state a few results about critical Bernoulli percolation on $\mathbb{Z}^d$ and $\mathbb{T}_r^d$ that are interesting in their own right, for the ease of future reference.

Theorem 1.5 (Connections inside balls). *Assume (1.1). There exists $C < \infty$ such that*
(a) for all $n \geq 1$,

$$\sum_{x \in \partial Q_n} \mathbb{P}_{\mathbb{Z},p_c}(0 \leftrightarrow x \ in \ Q_n) \leq C,$$

(b) for all $\varepsilon > 0$,

$$\limsup_{n \to \infty} \sum_{x \in \partial Q_n} \mathbb{P}_{\mathbb{Z},p_c}(0 \xleftrightarrow{\leq \varepsilon n^2} x \ in \ Q_n) \leq C\sqrt{\varepsilon}. \tag{1.7}$$

Theorem 1.6 (Short connections). *Assume (1.1). There exists $C < \infty$ such that for any $z \in \mathbb{Z}^d$ and positive integers r and k,*

$$\sum_{x \in \mathbb{Z}^d, \ x \overset{r}{\sim} z, \ |x| \geq \lfloor r/4 \rfloor} \mathbb{P}_{\mathbb{Z},p_c}\left(0 \xleftrightarrow{\leq k} x \ in \ \mathbb{Z}^d\right) \leq \begin{cases} \frac{Ck}{r^d} & if \ k \geq r^2, \\ Cr^{2-d} \cdot \mathrm{e}^{-\frac{r}{C\sqrt{k}}} & if \ k \leq r^2, \end{cases} \tag{1.8}$$

and for $k \geq r^2$,

$$\sum_{w \overset{r}{\sim} z, \ |w| \geq 3r/4} \sum_{u,v} \mathbb{P}_{\mathbb{Z},p_c}\left(0 \xleftrightarrow{\leq k} u \ in \ \mathbb{Z}^d\right) \mathbb{P}_{\mathbb{Z},p_c}\left(u \xleftrightarrow{\leq k} v \ in \ \mathbb{Z}^d\right) \mathbb{P}_{\mathbb{Z},p_c}\left(v \xleftrightarrow{\leq k} w \ in \ \mathbb{Z}^d\right) \leq \frac{Ck^3}{r^d}. \tag{1.9}$$

We note that in the special case where $r = 1$, (1.8) implies the result of [20, Theorem 1.2(i)] that

$$\sum_{x \in \mathbb{Z}^d} \mathbb{P}_{\mathbb{Z},p_c}\left(0 \xleftrightarrow{\leq k} x \ in \ \mathbb{Z}^d\right) \leq Ck. \tag{1.10}$$

Theorem 1.7 (Torus two-point function). *There exists $C < \infty$ such that for all $x \in \mathbb{T}_r^d$,*

$$\tau_{\mathbb{T},p_c}(x) \leq \tau_{\mathbb{Z},p_c}(x) + CV^{-2/3}, \tag{1.11}$$

and for all positive integers $n < r/2$,

$$\sup_{x \in \mathbb{T}_r^d} \mathbb{P}_{\mathbb{T},p_c}\left(0 \leftrightarrow x \ by \ a \ path \ which \ visits \ \partial Q_n\right) \leq Cn^{2-d} + CV^{-2/3}. \tag{1.12}$$

1.4   Discussion

In this section, we compare our results to those for the Erdős-Rényi random graph (ERRG), as proved, for example, by Aldous in [2], and formulate some open problems.

Cycle structure on the Erdős-Rényi random graph.   We refer to [2] for the extensive literature on the cycle structure of the ERRG. The ERRG is obtained by removing each edge of the complete graph K_n independently with probability p. On the critical ERRG, there is no distinction between long and short cycles, and the number of cycles of a cluster equals the tree excess of the cluster, i.e., the minimal number of edges one needs to remove from the cluster in order for it to be a tree.

On the ERRG, within the critical window $p = (1 + \lambda n^{-1/3})/n$ for some $\lambda \in \mathbb{R}$, the number of cycles of large clusters converges in distribution to a Poisson random variable with a *random* parameter. This random parameter can be described as the *area* of the cluster exploration process. Since this parameter is a bounded random variable, in particular, each large cluster has a tree excess that converges in distribution, and the probability that the ith largest cluster does not contain any cycle is strictly positive for any $i \geq 1$ fixed.

Since there is just a finite number of clusters of size at least $\delta n^{2/3}$, this immediately implies that the probability that there are no connected components of size at least $\delta n^{2/3}$ containing cycles is strictly positive. Further, all cycles have macroscopic length. Indeed, the largest connected components in the ERRG have diameter of the order $n^{1/3}$, and the length of cycles (when they exist) is also of the same order of magnitude. Cycles play a crucial role in describing the scaling limit of the largest critical clusters on the ERRG, as identified in [1]. Indeed, clusters locally look like trees, with cycles creating shortcuts between the different branches of the tree. Since cycles have a macroscopic length, these shortcuts are also macroscopic and thus the scaling limit of large critical clusters on the ERRG within the critical window is *not* a tree.

We conclude that the main features of the scaling limit of the critical ERRG are (a) the largest critical clusters being of size $O(n^{2/3})$, with a non-concentrated limit; (b) the largest critical clusters being close to trees with at most a finite number of macroscopic cycles; (c) a non-trivial probability that there exists large clusters having cycles.

Cycle structure on high-dimensional tori. We next investigate percolation on high-dimensional tori. In this case, of the above features (a-c) of the scaling limit of the critical ERRG, the feature (a) has been investigated in [4, 5, 13, 14], we focus on features (b) and (c) here. While our results clearly are not as strong as on the ERRG, they do establish the non-triviality of the probability of existence of cycles in large clusters as well as bounds on their length. Based on our results, we see that a natural split exists between cycles containing $O(V^{1/3})$ vertices that are truly macroscopic, and short cycles that basically remain within a cube without leaving its boundary. The former are essential in describing the scaling limit, the latter vanish in the scaling limit. There is no middle ground.

Our results provide yet another argument why the scaling limit of critical percolation on high-dimensional tori should be related to that for the critical ERRG. More precisely, the scaling limit of large critical clusters on the ERRG within the critical window is, or is not, a tree, each with positive probability. We see the same for large critical clusters on the high-dimensional torus, where with high probability, there are no long cycles containing $o(V^{1/3})$ vertices in large cluster. Thus, all long cycles are macroscopic, and will thus change the scaling limit of large critical clusters, in a similar way as they do on the ERRG.

Open probems. We complete this section by formulating a few open problems. The first extension deals with the values of p within the so-called *scaling window*. The results in [13, 14], in conjunction with those in [4, 5], show that when $p = p_c(1 + \varepsilon)$, where $V^{1/3}|\varepsilon|$ remains uniformly bounded, the largest clusters obey similar scaling as for $p = p_c$. An open problem is to show that Theorems 1.1–1.4 remain valid throughout the scaling window. Another extension is to more general high-dimensional tori, for example, to percolation on the hypercube as studied in [6, 15]. An open problem is to prove that Theorems 1.1–1.4 also hold on the hypercube, where a cycle is defined to be *long* when the number of edges in it is at least the random walk mixing time $n \log n$. The random walk mixing time plays a crucial role in [15] to identify the supercitical behavior of percolation on the hypercube.

Organization of this paper. This paper is organized as follows. In Section 2, we collect some preliminary results that we use in the proof. In Section 3, we prove Theorem 1.1. In Section 4, we prove Theorem 1.2. In Section 5, we prove Theorems 1.3 and 1.4. In Section 6, we prove Theorem 1.5, in Section 7 we prove Theorem 1.6, and in Section 8 we prove Theorem 1.7.

2 Preliminary results

In this section we collect some results that we use in the proofs.

2.1 Coupling clusters on the torus and the lattice

We say that two vertices x and y in $\mathbb{Z}^d$ are *r-equivalent* and write $x \overset{r}{\sim} y$, if $y = x + rz$ for some $z \in \mathbb{Z}^d$. We say that two edges $e_1 = \{x_1, y_1\}$ and $e_2 = \{x_2, y_2\}$ in $\mathbb{E}^d$ are *r-equivalent* ($e_1 \overset{r}{\sim} e_2$) if $x_1 \overset{r}{\sim} x_2$ and $y_1 \overset{r}{\sim} y_2$, or if $x_1 \overset{r}{\sim} y_2$ and $y_1 \overset{r}{\sim} x_2$. In the proof of Theorem 1.1, we need an extension of [13, Proposition 2.1]:

Proposition 2.1 (Coupling of clusters on torus and lattice)**.** *Consider bond percolation on $\mathbb{Z}^d$ and on $\mathbb{T}_r^d$ with parameter $p \in [0, 1]$. There exists a coupling $\mathbb{P}_{\mathbb{Z},\mathbb{T},p}$ of $\mathbb{P}_{\mathbb{Z},p}$ and $\mathbb{P}_{\mathbb{T},p}$ on the joint space of percolation on $\mathbb{Z}^d$ and $\mathbb{T}_r^d$ such that the following properties are satisfied $\mathbb{P}_{\mathbb{Z},\mathbb{T},p}$-almost surely for all k:*

(a) for all $x \in \mathbb{T}_r^d$,

$$\{0 \overset{\leq k}{\longleftrightarrow} x \text{ in } \mathbb{T}_r^d\} \subseteq \bigcup_{y \in \mathbb{Z}^d, y \overset{r}{\sim} x} \{0 \overset{\leq k}{\longleftrightarrow} y \text{ in } \mathbb{Z}^d\},$$

(b) for $x \overset{r}{\sim} y$, the event

$$\{0 \overset{\leq k}{\longleftrightarrow} y \text{ in } \mathbb{Z}^d\} \setminus \{0 \overset{\leq k}{\longleftrightarrow} x \text{ in } \mathbb{T}_r^d\}$$

implies that there exist distinct r-equivalent vertices v_1 and v_2 in $\mathbb{Z}^d$ and a vertex $z \in \mathbb{Z}^d$ such that the following disjoint connections take place in $\mathbb{Z}^d$:

$$\{0 \overset{\leq k}{\longleftrightarrow} z\} \circ \{z \overset{\leq k}{\longleftrightarrow} v_1\} \circ \{z \overset{\leq k}{\longleftrightarrow} v_2\} \circ \{v_1 \overset{\leq k}{\longleftrightarrow} y\}.$$

[13, Proposition 2.1] is Proposition 2.1 for $k = \infty$.

Proof. Let φ be the map from $\mathbb{E}_r^d$ to subsets of $\mathbb{E}^d$ defined by

$$\varphi(\{x, y\}) = \left\{ \{x', y'\} \in \mathbb{E}^d \ : \ x' \overset{r}{\sim} x, \ y' \overset{r}{\sim} y \right\} \qquad \text{for all } \{x, y\} \in \mathbb{E}_r^d.$$

The sets $(\varphi(e))_{e \in \mathbb{E}_r^d}$ form the equivalence classes of the equivalence relation $e \overset{r}{\sim} f$, where $e \overset{r}{\sim} f$ denotes that the endpoitns are r-equivalent. Note that for each $e \in \mathbb{E}_r^d$, $\varphi(e) \neq \varnothing$, and the sets $(\varphi(e))_{e \in \mathbb{E}_r^d}$ form a partition of $\mathbb{E}^d$. Let Φ be the map from $\{0, 1\}^{\mathbb{E}_r^d}$ to $\{0, 1\}^{\mathbb{E}^d}$ defined by

$$\Phi(\omega)_f = \omega_e \qquad \text{for all } \omega \in \{0, 1\}^{\mathbb{E}_r^d}, \ e \in \mathbb{E}_r^d, \ f \in \varphi(e).$$

Informally, we prove the proposition by defining for each percolation configuration on $\mathbb{T}_r^d$ a certain "unwrapping" of the set of edges with an end-vertex in $\mathcal{C}_{\mathbb{T}}(0)$ onto $\mathbb{E}^d$, and then revealing the status of the remaining edges of $\mathbb{E}^d$ by sampling from an independent percolation configuration on $\mathbb{Z}^d$. We achieve this by constructing an exploration of the edges of $\mathcal{C}_{\mathbb{T}}(0)$ for each percolation configuration on $\mathbb{T}_r^d$.

Let $\omega \in \{0, 1\}^{\mathbb{E}_r^d}$. For $n \geq 0$, we will define the following sets of edges in $\mathbb{E}_r^d$ recursively in n:

$$A_{\mathbb{T}}(n) \text{ (active edges)}, \qquad O_{\mathbb{T}}(n) \text{ (occupied edges)}, \qquad V_{\mathbb{T}}(n) \text{ (vacant edges)}, \qquad \text{and}$$
$$E_{\mathbb{T}}(n) = O_{\mathbb{T}}(n) \cup V_{\mathbb{T}}(n) \text{ (explored edges)};$$

and the following sets of edges in $\mathbb{E}^d$:

$$A_{\mathbb{Z}}(n) \text{ (active edges)}, \qquad O_{\mathbb{Z}}(n) \text{ (occupied edges)}, \qquad V_{\mathbb{Z}}(n) \text{ (vacant edges)},$$
$$G_{\mathbb{Z}}(n) \text{ (ghost edges)}, \qquad \text{and} \qquad E_{\mathbb{Z}}(n) = O_{\mathbb{Z}}(n) \cup V_{\mathbb{Z}}(n) \cup G_{\mathbb{Z}}(n) \text{ (explored edges)}.$$

We refer to this recursive procedure as the *exploration*.

We initiate the exploration by taking $A_{\mathbb{Z}}(0)$ to be the set of all edges that are neighbors of the origin in $\mathbb{Z}^d$, $E_{\mathbb{Z}}(0) = E_{\mathbb{T}}(0) = \varnothing$, and $A_{\mathbb{T}}(0) = \{e \in \mathbb{E}_r^d \ : \ \varphi(e) \cap A_{\mathbb{Z}}(0) \neq \varnothing\}$. Note that $A_{\mathbb{T}}(0)$ is the set of all edges in $\mathbb{E}_r^d$ that are neighbors of the origin in $\mathbb{T}_r^d$.

Let $n \geq 0$. We assume that the exploration is defined up to step n and now define it at step $(n + 1)$. If $A_{\mathbb{Z}}(n) = \varnothing$, then we stop the exploration and write $T = n - 1$. Otherwise, we take an edge $e \in A_{\mathbb{Z}}(n)$ which is *closest to the origin in terms of the graph distance in $O_{\mathbb{Z}}(n)$* (with ties broken in an arbitrary deterministic fashion). We define $G_{\mathbb{Z}}(n + 1) = G_{\mathbb{Z}}(n) \cup \{f : f \neq e, f \overset{r}{\sim} e\}$. To make the other updates, we consider two cases:

(a) if $\Phi(\omega)_e = 1$, then we define $O_{\mathbb{Z}}(n+1) = O_{\mathbb{Z}}(n) \cup \{e\}$, $V_{\mathbb{Z}}(n+1) = V_{\mathbb{Z}}(n)$, and $A_{\mathbb{Z}}(n+1) = A_{\mathbb{Z}}(n) \cup \{f : f \sim e\} \setminus E_{\mathbb{Z}}(n+1)$, where $e \sim f$ means that e and f share an end-vertex;

(b) if $\Phi(\omega)_e = 0$, then we define $O_{\mathbb{Z}}(n+1) = O_{\mathbb{Z}}(n)$, $V_{\mathbb{Z}}(n+1) = V_{\mathbb{Z}}(n) \cup \{e\}$, and $A_{\mathbb{Z}}(n+1) = A_{\mathbb{Z}}(n) \setminus E_{\mathbb{Z}}(n+1)$.

Finally, for $S \in \{A, O, V, E\}$, we define $S_{\mathbb{T}}(n+1) = \{e \in \mathbb{E}_r^d : \varphi(e) \cap S_{\mathbb{Z}}(n+1) \neq \varnothing\}$.

Note that in the exploration the status of each edge $f \in \mathbb{E}_r^d$ (i.e., the value of ω_f) is checked at most once. Indeed, once an edge $e \in \varphi(f)$ is selected, all the remaining edges in its equivalence class $\varphi(f)$ are immediately declared ghost, and therefore, cannot become active anymore. In particular, for each $S \in \{A, O, V, E\}$, if $S_{\mathbb{Z}}(n+1) \setminus S_{\mathbb{Z}}(n) \neq \varnothing$, then $S_{\mathbb{T}}(n+1) \setminus S_{\mathbb{T}}(n) \neq \varnothing$.

Also note that the exploration eventually stops, since T is at most the number of edges in $\mathbb{E}_r^d$. Moreover, the set $O_{\mathbb{T}}(T)$ coincides with the set of open edges of $\mathcal{C}_{\mathbb{T}}(0)$ in ω, and the set $V_{\mathbb{T}}(T)$ is the set of closed edges of $\mathbb{E}_r^d$ in ω sharing an end-vertex with $\mathcal{C}_{\mathbb{T}}(0)$. In other words, the exploration stops as soon as all the edges with an end-vertex in $\mathcal{C}_{\mathbb{T}}(0)$ are explored.

To complete the construction of the coupling, we define, for each $\omega \in \{0,1\}^{\mathbb{E}_r^d}$ and $\overline{\omega} \in \{0,1\}^{\mathbb{E}^d}$, the configuration $\widetilde{\omega} \in \{0,1\}^{\mathbb{E}^d}$ such that for all $e \in \mathbb{E}^d$,

$$\widetilde{\omega}_e = \begin{cases} \Phi(\omega)_e & \text{if } e \in O_{\mathbb{Z}}(T) \cup V_{\mathbb{Z}}(T), \\ \overline{\omega}_e & \text{otherwise.} \end{cases}$$

Note that for any $p \in [0,1]$, if $(\omega, \overline{\omega})$ is sampled from $\mathbb{P}_{\mathbb{T},p} \otimes \mathbb{P}_{\mathbb{Z},p}$, then $\widetilde{\omega}$ gives a sample from $\mathbb{P}_{\mathbb{Z},p}$. Therefore, $(\omega, \widetilde{\omega})$ gives us a coupling of $\mathbb{P}_{\mathbb{T},p}$ and $\mathbb{P}_{\mathbb{Z},p}$. It remains to check that this coupling satisfies the properties defined in the statement of the proposition.

Property (a) is immediate from the construction. It remains to prove property (b). Take $x \in \mathbb{T}_r^d$, $y \in \mathbb{Z}^d$, with $x \overset{r}{\sim} y$, and an integer $k \geq 0$. Let $\omega \in \{0,1\}^{\mathbb{E}_r^d}$ and $\overline{\omega} \in \{0,1\}^{\mathbb{E}^d}$ be such that

$$\widetilde{\omega} \in \{0 \overset{\leq k}{\longleftrightarrow} y \text{ in } \mathbb{Z}^d\} \qquad \text{and} \qquad \omega \notin \{0 \overset{\leq k}{\longleftrightarrow} x \text{ in } \mathbb{T}_r^d\}.$$

We need to show that there exist distinct r-equivalent vertices v_1 and v_2 in $\mathbb{Z}^d$ and a vertex $z \in \mathbb{Z}^d$ such that

$$\widetilde{\omega} \in \{0 \overset{\leq k}{\longleftrightarrow} z\} \circ \{z \overset{\leq k}{\longleftrightarrow} v_1\} \circ \{z \overset{\leq k}{\longleftrightarrow} v_2\} \circ \{v_1 \overset{\leq k}{\longleftrightarrow} y\}.$$

We fix a shortest open path π from 0 to y in $\mathbb{Z}^d$ in $\widetilde{\omega}$. By assumption, the length k' of π is at most k. Since $\omega \notin \{0 \overset{\leq k}{\longleftrightarrow} x \text{ in } \mathbb{T}_r^d\}$, there exists an edge e on this path such that $e \in G_{\mathbb{Z}}(T)$. Let e be the first edge on π from $G_{\mathbb{Z}}(T)$ (counting from 0 to y). We denote by $(e_1, \ldots, e_{m-1})$ all the edges on the part of π from 0 to e, and by $(e_{m+1}, \ldots, e_{k'})$ all the edges on the part of π from e to y. Since π is chosen to be a shortest open path from 0 to y, there exists n such that $e_1, \ldots, e_{m-1} \in O_{\mathbb{Z}}(n)$ and $e_{m+1}, \ldots, e_{k'} \notin O_{\mathbb{Z}}(n)$. Indeed, otherwise by the definition of the exploration, there would exist an open path from 0 to y which is strictly shorter than π.

By construction, there exists an edge $f \in \mathbb{Z}^d$ such that $f \neq e$, $f \overset{r}{\sim} e$, and the origin is connected to one of the end vertices of f by an open path π' inside $O_{\mathbb{Z}}(n)$. In particular, the length of π' is at most k. We denote this vertex by v_2, and let v_1 be the end-vertex of e which is r-equivalent to v_2. Let π_1 be the part of π from 0 to v_1, and π_2 be the part of π from v_1 to y.

Note that π_1 and π_2 are edge-disjoint by definition. Moreover, since the edges of π' are all in $O_{\mathbb{Z}}(n)$, and the edges of π_2 are not in $O_{\mathbb{Z}}(n)$, we deduce that π' and π_2 are also edge-disjoint. We have thus shown that

$$\widetilde{\omega} \in \bigcup_{v_1 \neq v_2, \, v_1 \overset{r}{\sim} v_2} \{0 \overset{\leq k}{\longleftrightarrow} v_1 \text{ in } \mathbb{Z}^d, \, 0 \overset{\leq k}{\longleftrightarrow} v_2 \text{ in } \mathbb{Z}^d\} \circ \{v_1 \overset{\leq k}{\longleftrightarrow} y \text{ in } \mathbb{Z}^d\}.$$

We finish the proof by observing that if 0 is connected to v_1 and v_2 by open paths in $\mathbb{Z}^d$ of length at most k, then there exists $z \in \mathbb{Z}^d$ such that the following edge-disjoint open paths (each of length at most k) exist in $\mathbb{Z}^d$: from 0 to z, from z to v_1, and from z to v_2. $\qquad\square$

From now on, we only consider the probability measure $\mathbb{P}_{\mathbb{Z},\mathbb{T},p}$ defined in Proposition 2.1 and

in the remainder of the paper, we write $\mathbb{P}_p$ for $\mathbb{P}_{\mathbb{Z},\mathbb{T},p}$.

In particular, $\mathbb{P}_p(E) = \mathbb{P}_{\mathbb{Z},p}(E)$ for $E \in \mathcal{F}_{\mathbb{Z},p}$, and $\mathbb{P}_p(E) = \mathbb{P}_{\mathbb{T},p}(E)$ for $E \in \mathcal{F}_{\mathbb{T},p}$ (see Section 2.1 for notation), and we always assume without mentioning the coupling from Proposition 2.1 when we consider events from $\mathcal{F}_{\mathbb{Z},p}$ and $\mathcal{F}_{\mathbb{T},p}$ simultaneously.

2.2 Previous results

In the next theorem, we summarize a number of results on high-dimensional percolation on $\mathbb{Z}^d$ that we will often use in the proofs in this paper.

We start by introducing balls in the intrinsic distance. For $x \in \mathbb{Z}^d$ and $k \geq 0$, we write $B_{\mathbb{Z},k}(x) = \{y : x \xleftrightarrow{\leq k} y \text{ in } \mathbb{Z}^d\}$ to denote all vertices at graph distance at most k from x in $\mathcal{C}_{\mathbb{Z}}(x)$. Similarly, for $x \in \mathbb{T}_r^d$ and $k \geq 0$, we write $B_{\mathbb{T},k}(x) = \{y : x \xleftrightarrow{\leq k} y \text{ in } \mathbb{T}_r^d\}$ to denote all vertices at graph distance at most k from x in $\mathcal{C}_{\mathbb{T}}(x)$. Further, we let $\partial B_{\mathbb{Z},k}(x) = B_{\mathbb{Z},k}(x) \setminus B_{\mathbb{Z},k-1}(x)$ and $\partial B_{\mathbb{T},k}(x) = B_{\mathbb{T},k}(x) \setminus B_{\mathbb{T},k-1}(x)$ denote the vertices at graph distance precisely equal to k from x. Let G be a subgraph of $\mathbb{Z}^d$ or $\mathbb{T}_r^d$, respectively. We define $B_{\mathbb{Z},k}^G(x)$, $\partial B_{\mathbb{Z},k}^G(x)$, $B_{\mathbb{T},k}^G(x)$ and $\partial B_{\mathbb{T},k}^G(x)$ in the same way as $B_{\mathbb{Z},k}(x)$, $\partial B_{\mathbb{Z},k}(x)$, $B_{\mathbb{T},k}(x)$ and $\partial B_{\mathbb{T},k}(x)$, except that we are now only allowed to use edges from G.

Theorem 2.1 (Critical behavior of high-dimensional percolation). *Assume (1.1). There exist $c > 0$ and $C < \infty$ such that:*

(i) For all $x \in \mathbb{T}_r^d$, $y \in \mathbb{Z}^d$, and positive integer k,

$$\sup_G \mathbb{P}_{p_c}\left(\partial B_{\mathbb{Z},k}^G(y) \neq \varnothing\right) \leq C/k, \qquad \sup_G \mathbb{P}_{p_c}\left(\partial B_{\mathbb{T},k}^G(x) \neq \varnothing\right) \leq C/k. \tag{2.1}$$

(ii) For all positive integers n,

$$cn^{-2} \leq \mathbb{P}_{p_c}\left(0 \leftrightarrow \partial Q_n \text{ in } \mathbb{Z}^d\right) \leq Cn^{-2}, \tag{2.2}$$

(iii)

$$cn^2 \leq \sum_{x \in Q_n} \tau_{\mathbb{Z},p_c}(x) \leq Cn^2, \tag{2.3}$$

and for any given $z \in \mathbb{Z}^d$ and a positive integer r with $r \leq n$,

$$\frac{cn^2}{r^d} \leq \sum_{x \in Q_n, x \overset{r}{\sim} z, |x| \geq r/8} \tau_{\mathbb{Z},p_c}(x) \leq \frac{Cn^2}{r^d}. \tag{2.4}$$

(iv) For any $z \in \mathbb{Z}^d$,

$$\sum_{x,y \in \mathbb{Z}^d} \tau_{\mathbb{Z},p_c}(0,x)\tau_{\mathbb{Z},p_c}(x,y)\tau_{\mathbb{Z},p_c}(y,z) \leq C|z|^{6-d}. \tag{2.5}$$

Proof. The first statement is [20, Theorem 1.2(ii)], and its adaptation to the torus in [14, Verification of Theorem 4.1(b)]. Statement (ii) is [19, Theorem 1]. Statements (iii) and (iv) easily follow from (1.1). $\square$

The next theorem gives an upper bound on $\mathbb{E}_{\mathbb{T},p_c}|\mathcal{C}(0)|$, which is used often in the proofs:

Theorem 2.2 (Expected critical cluster size on torus). *There exists $C < \infty$ such that for all $r \geq 1$,*

$$\mathbb{E}_{\mathbb{T},p_c}|\mathcal{C}(0)| \leq CV^{1/3}. \tag{2.6}$$

Proof. The statement follows from [13, (1.6)] and [5, Theorem 1.6(iii)]. (Alternatively, it follows from (1.11) and (2.3).) $\square$

3 Proof of Theorem 1.1

In this section, we prove Theorem 1.1 subject to Theorems 1.5–1.7. We give the proof of the upper bound in Proposition 3.1 and the proof of the lower bound in Proposition 3.2. The results of these propositions are more general than the result of Theorem 1.1, and also give bounds on the probability that the origin is in a long cycle with small length. In particular, the upper bound on the probability of such an event will be used in the proof of Theorem 1.2 to show that large open clusters do not contain long cycles with few edges.

Recall that LC_k denotes the event that the origin is in an open long cycle of length at most k. We prove the following bounds:

Proposition 3.1 (Upper bound on long cycles). *Assume (1.1). There exists $C < \infty$ such that*

$$\mathbb{P}_{p_c}(0 \text{ is in an open long cycle}) \leq CV^{-2/3},$$

and

$$\mathbb{P}_{p_c}(\mathrm{LC}_k) \leq Ck/V.$$

Proof. We begin by proving the second statement of the proposition. Let $k \geq 1$. By the definition of a long cycle, if LC_k occurs, then there exists $x \in \partial Q_{r/4}$ such that the following connections occur disjointly:

$$\{0 \leftrightarrow x \text{ in } Q_{r/4}\} \circ \{0 \xleftrightarrow{\leq k} x \text{ in } \mathbb{T}_r^d\}.$$

By the BK inequality,

$$\mathbb{P}_{p_c}(\mathrm{LC}_k) \leq \sum_{x \in \partial Q_{r/4}} \mathbb{P}_{p_c}\left(0 \leftrightarrow x \text{ in } Q_{r/4}\right) \mathbb{P}_{p_c}\left(0 \xleftrightarrow{\leq k} x \text{ in } \mathbb{T}_r^d\right).$$

By property (a) of Proposition 2.1 and (1.8), we obtain that for any $x \in \partial Q_{r/4}$,

$$\mathbb{P}_{p_c}\left(0 \xleftrightarrow{\leq k} x \text{ in } \mathbb{T}_r^d\right) \leq \sum_{y \in \mathbb{Z}^d, y \overset{r}{\sim} x} \mathbb{P}_{p_c}\left(0 \xleftrightarrow{\leq k} y \text{ in } \mathbb{Z}^d\right) \leq C_1 k/V.$$

Combining the above inequalities and using Theorem 1.5(a), we arrive at the second statement of the proposition.

We proceed with the proof of the first statement. As in the proof of the second statement,

$$\mathbb{P}_{p_c}(0 \text{ is in an open long cycle}) \leq \sum_{x \in \partial Q_{r/4}} \mathbb{P}_{p_c}\left(0 \leftrightarrow x \text{ in } Q_{r/4}\right) \mathbb{P}_{p_c}\left(0 \leftrightarrow x \text{ in } \mathbb{T}_r^d\right).$$

By (1.11) in Theorem 1.7 and (1.1), for any $x \in \partial Q_{r/4}$,

$$\mathbb{P}_{p_c}\left(0 \leftrightarrow x \text{ in } \mathbb{T}_r^d\right) \leq C_2 r^{2-d} + C_2 V^{-2/3} \leq 2C_2 V^{-2/3},$$

where the last inequality holds for $d > 6$. Putting the bounds together and using Theorem 1.5(a), we obtain the first statement of the proposition. $\square$

Proposition 3.2 (Lower-bound on long cycles). *Assume (1.1). There exist constants $c, \varepsilon > 0$ and $K < \infty$ such that*

$$\mathbb{P}_{p_c}(0 \text{ is in an open long cycle}) \geq cV^{-2/3},$$

and for any $k \in [Kr^2, \varepsilon V^{1/3}]$,

$$\mathbb{P}_{p_c}(\mathrm{LC}_k) \geq ck/V.$$

Proof. The first statement immediately follows from the second one for $k = \varepsilon V^{1/3}$. We now prove the second statement. Let R be a large positive integer. Let K be a large positive number, and ε a small positive number. The precise choice of these numbers will be made later in the proof. Let k be an integer with $k \in [Kr^2, \varepsilon V^{1/3}]$.

First of all, note that it suffices to prove the result for $r > 16R$. Indeed, once we fix R, the result for $r \in [1, 16R]$ will follow by adjusting the constant c. Therefore, throughout the proof we assume that $r > 16R$. The proof consists of several steps.

Step 1. For any $x \in \mathbb{Z}^d$, let $u(x)$ be the vertex in $\mathbb{Z}^d$ with coordinates $(x_1, \ldots, x_{d-1}, x_d + R)$. In particular, $u(x) \in \partial Q_R(x)$. For $x \in \mathbb{Z}^d$, consider the event

$$A_x = A_x(k, R) = \{0 \overset{\leq k}{\longleftrightarrow} u(x)\}.$$

Let

$$N(A) = \left| \left\{ x \in \mathbb{Z}^d \colon x \overset{r}{\sim} 0, \; x \neq 0, \; A_x \text{ occurs} \right\} \right|.$$

We use the second moment method to show that

$$\mathbb{P}_{p_c}(N(A) \neq 0) \geq ck/V,$$

for some constant c that may depend on K, but not on r, k or R.

We first show that there exists a constant $C_3 = C_3(K)$ such that

$$\mathbb{E}_{p_c} N(A) \geq C_3 k/V. \tag{3.1}$$

We write

$$\mathbb{E}_{p_c} N(A) \geq \sum_{x \overset{r}{\sim} 0, \; r/16 \leq |x| \leq \sqrt{k/K}} \mathbb{P}_{p_c}(0 \overset{\leq k}{\longleftrightarrow} u(x)). \tag{3.2}$$

Similarly to the proof of [20, Theorem 1.3(i)], one can show that if $K = K(d)$ is chosen large enough, then for any $r/16 \leq |x| \leq \sqrt{k/K}$,

$$\mathbb{P}_{p_c}(0 \overset{\leq k}{\longleftrightarrow} u(x)) \geq \frac{1}{2} \mathbb{P}_{p_c}(0 \leftrightarrow u(x)). \tag{3.3}$$

Inequality (3.1) now follows from (3.2), (3.3), the assumptions $r > 16R$ and $k \geq Kr^2$, and the lower bound in (2.4), where $C_3 = C_3(K) = c_3/K$ for some $c_3 > 0$ independent of all other parameters. From this moment onwards, the large integer K remains unchanged.

Next, we bound the second moment of $N(A)$. Let $\sum'$ be the sum over all distinct $x, y \in \mathbb{Z}^d$ such that $x, y \neq 0$ and $x, y \overset{r}{\sim} 0$. We obtain

$$\mathbb{E}_{p_c} N(A)^2 \leq \mathbb{E}_{p_c} N(A) + {\sum}' \mathbb{P}_{p_c} \left(0 \overset{\leq k}{\longleftrightarrow} u(x), \; 0 \overset{\leq k}{\longleftrightarrow} u(y) \right).$$

Note that if $0 \overset{\leq k}{\longleftrightarrow} u(x)$ and $0 \overset{\leq k}{\longleftrightarrow} u(y)$, then there exists $z \in \mathbb{Z}^d$ such that the following open paths occur disjointly: $0 \overset{\leq k}{\longleftrightarrow} z$, $z \overset{\leq k}{\longleftrightarrow} u(x)$ and $z \overset{\leq k}{\longleftrightarrow} u(y)$. Therefore, the BK inequality implies

$$\mathbb{E}_{p_c} N(A)^2 \leq \mathbb{E}_{p_c} N(A) + {\sum}' \sum_{z \in \mathbb{Z}^d} \mathbb{P}_{p_c} \left(0 \overset{\leq k}{\longleftrightarrow} z \right) \mathbb{P}_{p_c} \left(z \overset{\leq k}{\longleftrightarrow} u(x) \right) \mathbb{P}_{p_c} \left(z \overset{\leq k}{\longleftrightarrow} u(y) \right).$$

Let $\sum''$ be the sum over all pairwise distinct $x, y, z \in \mathbb{Z}^d$ such that $x \overset{r}{\sim} y \overset{r}{\sim} z$. By translation invariance and the fact that $u(x) - z = u(x - z)$, we have

$$\mathbb{E}_{p_c} N(A)^2 \leq \mathbb{E}_{p_c} N(A) + {\sum}'' \mathbb{P}_{p_c} \left(0 \overset{\leq k}{\longleftrightarrow} z \right) \mathbb{P}_{p_c} \left(0 \overset{\leq k}{\longleftrightarrow} u(x) \right) \mathbb{P}_{p_c} \left(0 \overset{\leq k}{\longleftrightarrow} u(y) \right).$$

Since x, y and z are distinct and r-equivalent, at least two of them are at distance at least $r/2$ from the origin. Therefore, the above sum is at most

$$\mathbb{E}_{p_c} N(A) + \sum_{x \overset{r}{\sim} y \overset{r}{\sim} z;\ |x|,|y| \geq r/2} \left[\mathbb{P}_{p_c}\left(0 \overset{\leq k}{\longleftrightarrow} z\right) \mathbb{P}_{p_c}\left(0 \overset{\leq k}{\longleftrightarrow} u(x)\right) \mathbb{P}_{p_c}\left(0 \overset{\leq k}{\longleftrightarrow} u(y)\right) \right.$$

$$\left. + 2\, \mathbb{P}_{p_c}\left(0 \overset{\leq k}{\longleftrightarrow} x\right) \mathbb{P}_{p_c}\left(0 \overset{\leq k}{\longleftrightarrow} u(y)\right) \mathbb{P}_{p_c}\left(0 \overset{\leq k}{\longleftrightarrow} u(z)\right) \right].$$

Remember that we assume that $r > 16R$. In particular, $|u(x)| \geq r/4$ when $|x| \geq r/2$. Applying (1.8) consequently to the sums over x, y, and then (1.10) to the sum over z, and then using the assumption $k \leq \varepsilon V^{1/3}$, we obtain that there exists $\varepsilon_0 = \varepsilon_0(d) > 0$ such that for any $\varepsilon < \varepsilon_0$,

$$\mathbb{E}_{p_c} N(A)^2 \leq \mathbb{E}_{p_c} N(A) + C_4 k \cdot \left(\frac{k}{V}\right)^2 \leq \mathbb{E}_{p_c} N(A) + C_4 \varepsilon^2 V^{-1/3} \cdot \frac{k}{V} \leq 2\mathbb{E}_{p_c} N(A). \tag{3.4}$$

A second moment estimate, using (3.1) and (3.4), yields

$$\mathbb{P}_{p_c}(N(A) \neq 0) \geq \frac{(\mathbb{E}_{p_c} N(A))^2}{\mathbb{E}_{p_c} N(A)^2} \geq \frac{C_3 k}{2V}. \tag{3.5}$$

Step 2. Consider the event

$$E = E(r, k, R) = \{0 \overset{\leq k}{\longleftrightarrow} u(0) \text{ in } \mathbb{T}_r^d \text{ by an open path which visits } \partial Q_{r/2}\}.$$

We show that for small enough ε and large enough R,

$$\mathbb{P}_{p_c}(E) \geq \frac{1}{2}\mathbb{P}_{p_c}(N(A) \neq 0).$$

Since $\mathbb{P}_{p_c}(E) \geq \mathbb{P}_{p_c}(N(A) \neq 0) - \mathbb{P}_{p_c}(\{N(A) \neq 0\} \setminus E)$, we should show that

$$\mathbb{P}_{p_c}(\{N(A) \neq 0\} \setminus E) \leq \frac{1}{2}\mathbb{P}_{p_c}(N(A) \neq 0), \tag{3.6}$$

when ε is chosen small enough and R large enough.

If $N(A) \neq 0$ and E does not occur, then according to Proposition 2.1, there exist $x \in \mathbb{Z}^d$ with $x \overset{r}{\sim} 0$ and $x \neq 0$, a vertex $z \in \mathbb{Z}^d$, and distinct vertices v_1 and v_2 in $\mathbb{Z}^d$ with $v_1 \overset{r}{\sim} v_2$ such that the following disjoint connections take place in $\mathbb{Z}^d$:

$$\{0 \overset{\leq k}{\longleftrightarrow} z\} \circ \{z \overset{\leq k}{\longleftrightarrow} v_1\} \circ \{z \overset{\leq k}{\longleftrightarrow} v_2\} \circ \{v_1 \overset{\leq k}{\longleftrightarrow} u(x)\}.$$

By the BK inequality, the probability of the event $\{N(A) \neq 0\} \setminus E$ is bounded from above by

$$\sum_{x \overset{r}{\sim} 0, x \neq 0} \sum_{v_1 \overset{r}{\sim} v_2, v_1 \neq v_2} \sum_z \mathbb{P}_{p_c}\left(0 \overset{\leq k}{\longleftrightarrow} z\right) \mathbb{P}_{p_c}\left(z \overset{\leq k}{\longleftrightarrow} v_1\right) \mathbb{P}_{p_c}\left(z \overset{\leq k}{\longleftrightarrow} v_2\right) \mathbb{P}_{p_c}\left(v_1 \overset{\leq k}{\longleftrightarrow} u(x)\right). \tag{3.7}$$

Note that, since v_1 and v_2 are distinct and r-equivalent, either $|v_1 - z| \geq r/2$ or $|v_2 - z| \geq r/2$. Assume first that $|v_2 - z| \geq r/2$. It follows from (1.8) that for any v_1 and z fixed,

$$\sum_{v_2 \in \mathbb{Z}^d,\ v_2 \overset{r}{\sim} v_1,\ |v_2 - z| \geq r/2} \mathbb{P}_{p_c}\left(z \overset{\leq k}{\longleftrightarrow} v_2\right) \leq C_5 k/V,$$

where C_5 does not depend on k, v_1 or z. On the other hand, by (1.9) and using the fact that $k \leq \varepsilon V^{1/3}$,

$$\sum_{x \overset{r}{\sim} 0, x \neq 0} \sum_{v_1, z} \mathbb{P}_{p_c}\left(0 \overset{\leq k}{\longleftrightarrow} z\right) \mathbb{P}_{p_c}\left(z \overset{\leq k}{\longleftrightarrow} v_1\right) \mathbb{P}_{p_c}\left(v_1 \overset{\leq k}{\longleftrightarrow} u(x)\right) \leq \frac{C_6 k^3}{V} \leq C_6 \varepsilon^3. \tag{3.8}$$

Therefore, for v_2 and z with $|v_2 - z| \geq r/2$, the sum (3.7) is, uniformly in R, bounded from above by

$$C_5 C_6 \varepsilon^3 k / V.$$

Next, consider the sum (3.7) in the case $|v_1 - z| \geq r/2$. By translation invariance, the sum (3.7) equals

$$\sum_{x \overset{r}{\sim} 0,\ x \neq 0} \sum_{v_1 \overset{r}{\sim} v_2,\ v_1 \neq v_2} \sum_z \mathbb{P}_{p_c}\left(0 \xleftrightarrow{\leq k} z\right) \mathbb{P}_{p_c}\left(z \xleftrightarrow{\leq k} v_1\right) \mathbb{P}_{p_c}\left(z \xleftrightarrow{\leq k} v_2\right) \mathbb{P}_{p_c}\left(v_2 \xleftrightarrow{\leq k} u(x) + (v_2 - v_1)\right).$$

By the definition of $u(x)$, the translation of $u(x)$ by $(v_2 - v_1)$ equals $u(x + (v_2 - v_1))$. Since $v_1 \overset{r}{\sim} v_2$, the translation of x (r-equivalent to 0) by $(v_2 - v_1)$ is still r-equivalent to 0. However, note that it is possible that $x + (v_2 - v_1) = 0$. These observations imply that the above sum is bounded from above by

$$\sum_{x \overset{r}{\sim} 0} \sum_{v_1 \overset{r}{\sim} v_2, v_1 \neq v_2} \sum_z \mathbb{P}_{p_c}\left(0 \xleftrightarrow{\leq k} z\right) \mathbb{P}_{p_c}\left(z \xleftrightarrow{\leq k} v_1\right) \mathbb{P}_{p_c}\left(z \xleftrightarrow{\leq k} v_2\right) \mathbb{P}_{p_c}\left(v_2 \xleftrightarrow{\leq k} u(x)\right).$$

Since we only consider the above sum in the case z and v_1 satisfy $|v_1 - z| \geq r/2$, we obtain as before that for any given z and v_2,

$$\sum_{v_1 \in \mathbb{Z}^d,\ v_1 \overset{r}{\sim} v_2,\ |v_1 - z| \geq r/2} \mathbb{P}_{p_c}\left(z \xleftrightarrow{\leq k} v_1\right) \leq C_5 k / V.$$

It remains to bound the sum

$$\sum_{z, v_2} \sum_{x \overset{r}{\sim} 0} \mathbb{P}_{p_c}\left(0 \xleftrightarrow{\leq k} z\right) \mathbb{P}_{p_c}\left(z \xleftrightarrow{\leq k} v_2\right) \mathbb{P}_{p_c}\left(v_2 \xleftrightarrow{\leq k} u(x)\right).$$

There are two cases depending on whether $x \neq 0$ or $x = 0$. The case $x \neq 0$ can be considered similarly to (3.8), so the above sum is bounded from above by $C_6 \varepsilon^3$ in this case.

It remains to consider the case $x = 0$. In this case $|u(x)| = R$, and we simply bound the above sum by

$$\sum_{z, v_2} \mathbb{P}_{p_c}\left(0 \leftrightarrow z\right) \mathbb{P}_{p_c}\left(z \leftrightarrow v_2\right) \mathbb{P}_{p_c}\left(v_2 \leftrightarrow u(0)\right),$$

which is bounded from above by $C_7 R^{6-d}$ by (2.5), where C_7 is independent of R. Therefore, for v_1 and z with $|v_1 - z| \geq r/2$, the sum (3.7) is bounded from above by

$$C_5(2C_6 \varepsilon^3 + C_7 R^{6-d}) k / V.$$

Now recalling (3.5), we take R large and ε small so that (3.6) holds. We obtain from (3.5) and (3.6) that

$$\mathbb{P}_{p_c}(E) \geq C_3 k / (4V).$$

Step 3. We now show that there exists $C_8 = C_8(R) > 0$ such that

$$\mathbb{P}_{p_c}(\mathrm{LC}_{k+R^d}) \geq C_8 \mathbb{P}_{p_c}(E).$$

This follows from a local modification argument as follows. Note that if E occurs, there exist z and z' on ∂Q_R such that z is connected to z' by a path in $\mathbb{T}_r^d \setminus Q_R$ of length at most k which visits $\partial Q_{r/2}$. We can therefore modify the configuration of bonds inside Q_R to make sure that $\{0 \leftrightarrow z \text{ in } Q_R\} \circ \{0 \leftrightarrow z' \text{ in } Q_R\}$, which implies that the origin is in a long cycle of length at most $k + R^d$. Since there are only finitely many edges in Q_R, the above inequality follows.

We can now complete the proof of Proposition 3.2. We pick $K = K(d)$ so that (3.1) holds for all $r > 16R$. We then pick $\varepsilon = \varepsilon(d)$ and $R = R(d)$ to satisfy (3.4) and (3.6). It follows from Steps 2 and 3 of the proof that for all $r > 16R$ and $k \in [Kr^2 + R^d, \varepsilon V^{1/3}]$,

$$\mathbb{P}_{p_c}(\mathrm{LC}_k) \geq C_8 \frac{C_3(k - R^d)}{4V} \geq C_9 k / V.$$

Finally, we adjust the constant C_9 so that the result remained valid for $r \leq 16R$. $\qquad \square$

4 Proof of Theorem 1.2

Proof of (1.5). Let $\varepsilon > 0$ and $\delta > 0$. Let M be the number of vertices $x \in \mathbb{T}_r^d$ such that $|\mathcal{C}_{\mathbb{T}}(x)| \geq \delta V^{2/3}$ and $\mathcal{C}_{\mathbb{T}}(x)$ contains a long cycle of length at most $\varepsilon V^{1/3}$. We need to show that there exists $C < \infty$ such that

$$\mathbb{P}_{\mathbb{T},p_c} (M \neq 0) \leq C\varepsilon/\delta.$$

By the definition of M and the Markov inequality,

$$\mathbb{P}_{\mathbb{T},p_c} (M \neq 0) = \mathbb{P}_{\mathbb{T},p_c} \left(M \geq \delta V^{2/3} \right) \leq \frac{\mathbb{E}_{\mathbb{T},p_c}[M]}{\delta V^{2/3}}.$$

By translation invariance,

$$\mathbb{E}_{\mathbb{T},p_c}[M] \leq V \cdot \mathbb{P}_{\mathbb{T},p_c} \left(\mathcal{C}_{\mathbb{T}}(0) \text{ contains a long cycle of length at most } \varepsilon V^{1/3} \right).$$

Note that if $\mathcal{C}_{\mathbb{T}}(0)$ contains a long cycle of length at most $\varepsilon V^{1/3}$, then there exists $z \in \mathcal{C}_{\mathbb{T}}(0)$ such that the following events occur disjointly:

$$\{0 \leftrightarrow z \text{ in } \mathbb{T}_r^d\} \circ \mathrm{LC}_{\varepsilon V^{1/3}}.$$

Therefore, application of the BK inequality gives that

$$\mathbb{P}_{\mathbb{T},p_c} \left(\mathcal{C}_{\mathbb{T}}(0) \text{ contains a long cycle of length at most } \varepsilon V^{1/3} \right) \leq \mathbb{E}_{\mathbb{T},p_c}|\mathcal{C}_{\mathbb{T}}(0)| \cdot \mathbb{P}_{\mathbb{T},p_c}(\mathrm{LC}_{\varepsilon V^{1/3}}).$$

Using (2.6) and Proposition 3.1, and putting all the bounds together gives (1.5). $\qquad \square$

Proof of (1.6). This is a simple consequence of Theorem 1.1. Indeed, if there exists a long cycle of length at least $\varepsilon^{-1}V^{1/3}$, then the number of vertices in long cycles is at least $(2d\varepsilon)^{-1}V^{1/3}$. Denote the number of vertices in long cycles by M. Then, by the Markov inequality and Theorem 1.1,

$$\mathbb{P}_{p_c} \left(\exists \text{ a long cycle of length } \geq \varepsilon^{-1}V^{1/3} \right) \leq \mathbb{P}_{p_c}(M \geq (2d\varepsilon)^{-1}V^{1/3}) \leq 2d\varepsilon V^{-1/3}\mathbb{E}_{p_c}[M]$$

$$= 2d\varepsilon V^{2/3}\mathbb{P}_{p_c} (0 \text{ is in a long cycle}) \leq C_1\varepsilon.$$

$$\square$$

5 Proof of Theorems 1.3 and 1.4

5.1 Proof of Theorem 1.3

Let $\mathcal{I}$ be the set of $z \in \mathcal{C}_{\mathbb{T}}(0)$ such that $\{0 \leftrightarrow z\} \circ \{z \text{ is in a long cycle}\}$. Theorem 1.3 follows from Lemma 5.1:

Lemma 5.1 (Bound on the number of long cycles).

$$Y_{\mathcal{C}_{\mathbb{T}}(0)} \leq 2d \, |\mathcal{I}|.$$

Moreover, when (1.1) *holds, there exists a finite constant C such that*

$$\mathbb{E}_{p_c}|\mathcal{I}| \leq CV^{-1/3}.$$

Before we prove Lemma 5.1, we show how to use it to complete the proof of Theorem 1.3:

$$\mathbb{E}_{p_c} [Y_\delta] = \sum_x \mathbb{E}_{p_c} \left[\frac{1}{\mathcal{C}_{\mathbb{T}}(x)} Y_{\mathcal{C}_{\mathbb{T}}(x)} I(|\mathcal{C}_{\mathbb{T}}(x)| > \delta V^{2/3}) \right] \leq (\delta V^{2/3})^{-1} V \mathbb{E}_{p_c} \left[Y_{\mathcal{C}_{\mathbb{T}}(0)} \right] \leq C_1/\delta,$$

as required. In the remaining part of this section, we prove Lemma 5.1.

Proof of Lemma 5.1. Let $\mathcal{E}$ be the set of edges of the torus adjacent to at least one of the vertices from $\mathcal{I}$. We will show that $Y_{\mathcal{C}_\mathbb{T}(0)} \leq |\mathcal{E}|$.

Let $G = \mathcal{C}_\mathbb{T}(0)$, and let $\widetilde{G} = G \setminus \mathcal{E}$ denote the subgraph of G obtained by removing every edge of G that is in $\mathcal{E}$. Note that every vertex from $\mathcal{I}$ is an isolated vertex in $\widetilde{G}$. We claim that the graph $\widetilde{G}$ does not contain long cycles. Indeed, assume that there is a long cycle π in $\widetilde{G}$. Since $\widetilde{G}$ is a subgraph of G, π is a long cycle in G. In particular, there exists $z \in \pi$ such that 0 is connected to z in G by a path that does not use any edges from π. Therefore, $z \in \mathcal{I}$ and z is not an isolated vertex in $\widetilde{G}$. This is a contradiction.

We have just shown that by removing every edge adjacent to a vertex in $\mathcal{I}$, we obtain a subgraph of $\mathcal{C}_\mathbb{T}(0)$ without long cycles. This implies that $Y_{\mathcal{C}_\mathbb{T}(0)} \leq |\mathcal{E}| \leq 2d|\mathcal{I}|$.

Further, by the BK inequality, (2.6) and Theorem 1.1,

$$\mathbb{E}_{p_c}|\mathcal{I}| \leq \mathbb{E}_{p_c}|\mathcal{C}_\mathbb{T}(0)| \, \mathbb{P}_{p_c}(0 \text{ is in a long cycle}) \leq C_2 V^{-1/3}.$$

$\square$

5.2 Proof of Theorem 1.4(a): existence of long cycles

We need to show that there exists $c > 0$ such that

$$\mathbb{P}_{p_c}\left(\text{there exists a long cycle of length } > cV^{1/3}\right) > c.$$

Take $\varepsilon > 0$. The precise value of ε will be determined later. Define

$$M = |\{x \colon x \text{ is in a long cycle of length } > \varepsilon V^{1/3}\}|.$$

Then, clearly,

$$\mathbb{P}_{\mathbb{T},p_c}\left(\exists \text{ a long cycle of length } > \varepsilon V^{1/3}\right) = \mathbb{P}_{\mathbb{T},p_c}(M \neq 0). \tag{5.1}$$

By the second moment method, we can bound

$$\mathbb{P}_{\mathbb{T},p_c}(M \neq 0) \geq \frac{(\mathbb{E}_{\mathbb{T},p_c} M)^2}{\mathbb{E}_{\mathbb{T},p_c} M^2}.$$

We first show that $\mathbb{E}_{\mathbb{T},p_c} M \geq C_3 V^{1/3}$, and then that $\mathbb{E}_{\mathbb{T},p_c} M^2 \leq C_4 V^{2/3}$. By translation invariance,

$$\mathbb{E}_{\mathbb{T},p_c} M = V \mathbb{P}_{\mathbb{T},p_c}(0 \text{ in a long cycle of length } > \varepsilon V^{1/3}).$$

Recall the definition of LC_k. We write

$$\mathbb{P}_{\mathbb{T},p_c}(0 \text{ in a long cycle of length } > \varepsilon V^{1/3}) \geq \mathbb{P}_{\mathbb{T},p_c}(0 \text{ in a long cycle}) - \mathbb{P}_{\mathbb{T},p_c}(\mathrm{LC}_{\varepsilon V^{1/3}}).$$

It follows from Theorem 1.1 that

$$\mathbb{P}_{\mathbb{T},p_c}(0 \text{ in a long cycle}) > C_5 V^{-2/3},$$

and from Proposition 3.1 that

$$\mathbb{P}_{\mathbb{T},p_c}(\mathrm{LC}_{\varepsilon V^{1/3}}) < C_6 \varepsilon V^{-2/3}. \tag{5.2}$$

By taking ε small enough, we deduce that

$$\mathbb{P}_{\mathbb{T},p_c}(0 \text{ in a long cycle of length } > \varepsilon V^{1/3}) \geq \frac{1}{2}\mathbb{P}_{\mathbb{T},p_c}(0 \text{ in a long cycle}),$$

and the desired lower bound on $\mathbb{E}_{\mathbb{T},p_c} M$ follows.

It remains to prove that $\mathbb{E}_{\mathbb{T},p_c} M^2 \leq C_4 V^{2/3}$. Since

$$\mathbb{E}_{\mathbb{T},p_c} M^2 = \sum_{x,y} \mathbb{P}_{\mathbb{T},p_c}\left(x, y \text{ in long cycles of length } > \varepsilon V^{1/3}\right),$$

it suffices to show that

$$\sum_{x,y} \mathbb{P}_{\mathbb{T},p_c} (x,y \text{ in long cycles}) \leq C_4 V^{2/3}.$$

We split the above sum, depending on whether the events $\{x \text{ in long cycle}\}$ and $\{y \text{ in long cycle}\}$ occur disjointly or not. The contribution where these events do occur disjointly can be bounded, using the BK inequality and Theorem 1.1, by $C_7 V^{2/3}$, so that

$$\mathbb{E}_{\mathbb{T},p_c} M^2 \leq C_7 V^{2/3} + \sum_{x,y} \mathbb{P}_{\mathbb{T},p_c} (x,y \text{ in overlapping long cycles}), \tag{5.3}$$

where the event $\{x,y \text{ in overlapping long cycles}\}$ indicates that all pairs of long cycles, one of which contains x and the other y, share at least one edge.

Lemma 5.2 (Contribution of overlapping cycles). *Assume (1.1). There exists $C < \infty$ such that*

$$\sum_{x,y} \mathbb{P}_{\mathbb{T},p_c} (x,y \text{ in overlapping long cycles}) \leq CV^{2/3}. \tag{5.4}$$

Proof. For a pair of vertices $x, y \in \mathbb{T}_r^d$ and $s \geq 0$, we denote by $\Delta_s(x,y)$ the event that x is connected to y by a path in $\mathbb{T}_r^d$ which visits $\partial Q_s(x)$.

Note that if x and y are in overlapping long cycles, then there exist u, v such that

(a) u and v both are part of the long cycle that contains x as well as the one that contains y,

(b) the connections $x \leftrightarrow u, u \leftrightarrow v, x \leftrightarrow v, y \leftrightarrow u, y \leftrightarrow v$ all occur disjointly,

(c) at least one of the events $\Delta_{r/12}(x,u)$, $\Delta_{r/12}(u,v)$, or $\Delta_{r/12}(v,x)$ occur.

Therefore, using symmetry and the BK inequality, we can upper bound the sum in (5.4) by

$$4 \sum_{x,y,u,v} \mathbb{P}_{\mathbb{T},p_c}(\Delta_{r/12}(x,u)) \tau_{\mathbb{T},p_c}(x,v) \tau_{\mathbb{T},p_c}(u,v) \tau_{\mathbb{T},p_c}(y,u) \tau_{\mathbb{T},p_c}(y,v)$$

$$+ \sum_{x,y,u,v} \tau_{\mathbb{T},p_c}(x,u) \tau_{\mathbb{T},p_c}(x,v) \mathbb{P}_{\mathbb{T},p_c}(\Delta_{r/12}(u,v)) \tau_{\mathbb{T},p_c}(y,u) \tau_{\mathbb{T},p_c}(y,v). \tag{5.5}$$

By (1.12), we can bound $\mathbb{P}_{\mathbb{T},p_c}(\Delta_{r/12}(x,u)) \leq C_8 V^{-2/3}$. Let

$$\nabla_{\mathbb{T},p} = \sup_{x,y} \sum_{u,v} \tau_{\mathbb{T},p}(x,u) \tau_{\mathbb{T},p}(u,v) \tau_{\mathbb{T},p}(v,y). \tag{5.6}$$

It follows from [5, Theorem 1.6(iii)] and [13, (1.6)] that $\nabla_{\mathbb{T},p_c} < C_9$. Therefore, we can bound the first sum in (5.5) by

$$C_8 V^{-2/3} \sum_{x,y,u,v} \tau_{\mathbb{T},p_c}(x,v) \tau_{\mathbb{T},p_c}(u,v) \tau_{\mathbb{T},p_c}(y,u) \tau_{\mathbb{T},p_c}(y,v) \leq C_8 V^{-2/3} V \nabla_{\mathbb{T},p_c} \mathbb{E}_{\mathbb{T},p_c} |\mathcal{C}_{\mathbb{T}}(0)| \overset{(2.6)}{\leq} C_{10} V^{2/3}, \tag{5.7}$$

and the second sum in (5.5) by

$$C_8 V^{-2/3} \sum_{x,y,u,v} \tau_{\mathbb{T},p_c}(x,u) \tau_{\mathbb{T},p_c}(x,v) \tau_{\mathbb{T},p_c}(y,u) \tau_{\mathbb{T},p_c}(y,v)$$

$$\leq C_8 V^{-2/3} V \sup_v \sum_{x,y,u,v} \tau_{\mathbb{T},p_c}(0,u) \tau_{\mathbb{T},p_c}(u,y) \tau_{\mathbb{T},p_c}(y,v) \sum_v \tau_{\mathbb{T},p_c}(v)$$

$$= C_8 V^{-2/3} V \nabla_{\mathbb{T},p_c} \mathbb{E}_{\mathbb{T},p_c} |\mathcal{C}(0)| \overset{(2.6)}{\leq} C_{10} V^{2/3}.$$

These estimates complete the proof. □

16

We continue with the proof of Theorem 1.4(a). It follows from (5.3) and Lemma 5.2 that $\mathbb{E}_{\mathbb{T},p_c} M^2 \leq C_4 V^{2/3}$. By the second moment method, we obtain

$$\mathbb{P}_{\mathbb{T},p_c}\left(\exists \text{ a long cycle of length } > \varepsilon V^{1/3}\right) = \mathbb{P}_{\mathbb{T},p_c}(M \neq 0) \geq \frac{(\mathbb{E}_{\mathbb{T},p_c} M)^2}{\mathbb{E}_{\mathbb{T},p_c} M^2} \geq \frac{C_3^2}{C_4} > 0.$$

This completes the proof of the fact that a long cycle of length $> \varepsilon V^{1/3}$ exists with positive probability. $\square$

5.3 Proof of Theorem 1.4(b): non-existence of long cycles

In this section we prove that, for any positive δ, with positive probability uniformly in r, the clusters of size $> \delta V^{2/3}$ do not contain any long cycles. In other words, recalling the definition of Y_δ in Section 5.1, we will prove the following proposition:

Proposition 5.1 (Non-existence of long cycles). *For any positive δ there exists $c > 0$ such that, for all $r \geq 1$,*

$$\mathbb{P}_{\mathbb{T},p_c}(Y_\delta = 0) > c.$$

Proof. For $x \in \mathbb{T}_r^d$, run the following exploration of the edges of $\mathcal{C}_{\mathbb{T}}(x)$ started from x. Enumerate the edges of $\mathbb{T}_r^d$. (In the algorithm we describe now, if there are several edges to choose from, we always pick the edge with the smallest number.) The first stage of the algorithm is the standard depth-first exploration. At this stage, after n steps, the algorithm produces

- the set of explored vertices X_n (which will be a subset of the vertices of $\mathcal{C}_{\mathbb{T}}(x)$),

- the set of explored edges E_n (these will be the explored edges, the occupancy of which we will check),

- the set of open explored edges T_n and the open cluster induced by these edges, also denoted by T_n (which will be part of the depth-first spanning tree of $\mathcal{C}_{\mathbb{T}}(x)$), and

- the set of unexplored edges U_n (the algorithm will not check the occupancy of these edges).

Further, let $W_n = E_n \cup U_n$.

Take $x \in \mathbb{T}_r^d$. Let $X_0 = T_0 = \{x\}$, $W_0 = \varnothing$. Let $n \geq 0$. Assume that X_n, E_n, T_n and U_n are defined. If there is no edge $\{a, b\}$ with $a \in X_n$ and $\{a, b\} \notin W_n$, then we stop the algorithm and write $\mathcal{A}_x = \mathcal{A}_n$ for all $\mathcal{A} \in \{X, E, T, U, W\}$. Otherwise, pick the vertex $a \in X_n$ which is the farthest from x in T_n for which there exists $b \in \mathbb{T}_r^d$ such that $\{a, b\} \notin W_n$. Such a vertex, if it exists, is always *unique*, since we explore depth-first. (We prove this statement in Lemma 5.3(a) at the end of the section.)

Let $e = \{a, b\}$ be the smallest such edge. We distinguish two cases:

1. If $b \notin X_n$, then we define $E_{n+1} = E_n \cup \{e\}$, $U_{n+1} = U_n$, and *check* the occupancy of e.

   (a) If e is open, then we define $X_{n+1} = X_n \cup \{b\}$ and $T_{n+1} = T_n \cup \{e\}$.

   (b) If e is closed, then we define $X_{n+1} = X_n$ and $T_{n+1} = T_n$.

2. If $b \in X_n$ (in this case we call e a *surplus* edge), then we define $X_{n+1} = X_n$, $T_{n+1} = T_n$, $E_{n+1} = E_n$, and $U_{n+1} = U_n \cup \{e\}$, and *do not check* the occupancy of e.

Since the number of edges of $\mathbb{T}_r^d$ is finite, this stage of the algorithm will terminate at some step $N < \infty$. We then write $\mathcal{A}_x = \mathcal{A}_N$ for all $\mathcal{A} \in \{X, E, T, U, W\}$. In particular, X_x is the vertex set of $\mathcal{C}_{\mathbb{T}}(x)$, T_x is the "depth-first" spanning tree of $\mathcal{C}_{\mathbb{T}}(x)$ with root at x, and W_x is the set of edges with at least one end vertex in $\mathcal{C}_{\mathbb{T}}(x)$. The occupancy of edges in E_x is known. In particular, the graph induced by sets of open edges in E_x is T_x. The occupancy of edges in U_x has not been checked. The sets E_x and U_x are disjoint. Also note that, given the set of unexplored edges U_x, the edges in U_x are open independently of each other. An example of an edge $\{a, b\} \in U_x$ is given in Figure 1 below.
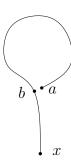
Figure 1: Example of an edge $\{a, b\} \in U_x$.

We proceed by describing the second stage of the algorithm. The aim of this second stage is to select those surplus edges $\{a, b\}$ that (i) close a long cycle; and (ii) are such that $x \longleftrightarrow b$ is completely disjoint from the long cycle that is created by the addition of the edge $\{a, b\}$; and (iii) there are no long cycles precisely when all these selected edges are closed.

After n steps, the algorithm produces

- the set of open explored edges G_n and the open cluster induced by these edges, also denoted by G_n (which will be a subgraph of $\mathcal{C}_\mathbb{T}(x)$ without long cycles),

- the set of explored edges F_n (which will be a subset of edges of U_x, the occupancy of which we will check), and

- the set of special edges Z_n (the algorithm will not check the occupancy of these edges; each $e \in Z_n$ will have the property that in the graph $G_n \cup \{e\}$, e is in a long cycle π, and there exists a path ρ connecting one of the end-vertices of e to x which is edge disjoint of π).

Note that, according to the first stage of the algorithm, each edge $e \in U_x$ can be written as $\{a, b\}$ such that the unique path from a to x in the spanning-tree T_x passes through b. (We prove this statement in Lemma 5.3(b) at the end of the section.)

Denote by B_x the set of end-vertices with this property, that is, a vertex b is in B_x if and only if there exists a vertex a such that the edge $\{a, b\}$ is in U_x and the unique path from a to x in the spanning tree T_x passes through b.

We enumerate the vertices of B_x subject to the following restriction: a vertex $b \in B_x$ receives a smaller number than $b' \in B_x$ if the unique path from b' to x in the spanning tree T_x passes through b. This ordering of the vertices in B_x can be better understood by introducing an auxiliary abstract tree $\mathbf{T}_x$ rooted at x with the vertex set $\{x\} \cup B_x$ and the following set of oriented (away from the root) edges: For $b, b' \in B_x$, there is an edge from b to b' in $\mathbf{T}_x$, if the unique path from b' to x in the depth-first spanning tree T_x passes through b, and the unique path between b and b' in T_x does not contain any other vertices from B_x. With this definition, we can alternatively say that a vertex $b \in B_x$ has a smaller number than $b' \in B_x$ if there is an oriented path from b to b' in $\mathbf{T}_x$. In other words, we enumerate the vertices of B_x according to their distance to x in the abstract tree $\mathbf{T}_x$. An example of a collection of cycles and the corresponding tree $\mathbf{T}_x$ is given in Figure 2 below.

The second stage of the algorithm goes as follows. Let $G_0 = T_x$, $F_0 = \varnothing$, $Z_0 = \varnothing$ and $B_0 = B_x$. Assume that, for $n \geq 0$, the sets G_n, F_n, Z_n and B_n are defined. If $B_n = \varnothing$, then we stop the algorithm and define $G_x = G_n$, $F_x = F_n$ and $Z_x = Z_n$.

Otherwise, pick a vertex $b \in B_n$ with the biggest number. We distinguish two cases:

1. If there are at least two vertices a and a' such that the edges $\{a, b\}$ and $\{a', b\}$ are in $U_x \setminus (F_n \cup Z_n)$, then define $B_{n+1} = B_n$ and we select the admissible edge with the smallest number. (This is the same numbering of the edges of the torus as in the first stage of the algorithm.)

2. If the vertex a such that the edge $\{a, b\}$ is in $U_x \setminus (F_n \cup Z_n)$ is unique, then we define $B_{n+1} = B_n \setminus \{b\}$ and select this edge.
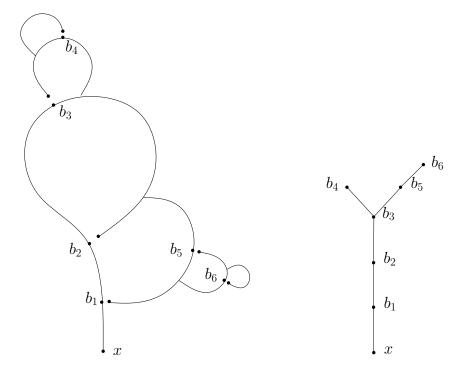
Assume that the edge $e = \{a, b\}$ is selected.

18

Figure 2: Example of a collection of cycles and the corresponding tree $\mathbf{T}_x$.

1. If the graph $G_n \cup \{e\}$ *does not contain* a long cycle, then we define $F_{n+1} = F_n \cup \{e\}$ and $Z_{n+1} = Z_n$ and check the occupancy of e.

   (a) If e is open, then define $G_{n+1} = G_n \cup \{e\}$.
   (b) If e is closed, then define $G_{n+1} = G_n$.

2. If the graph $G_n \cup \{e\}$ *does contain* a long cycle, then we define $G_{n+1} = G_n$, $F_{n+1} = F_n$ and $Z_{n+1} = Z_n \cup \{e\}$ and *do not check* the occupancy of e. Note that, by the special ordering of the vertices in B_x, every long cycle of the graph $G_n \cup \{e\}$ passes through e and it is *edge disjoint with the unique path from b to x in the tree T_x.*

Since the number of edges of $\mathbb{T}_r^d$ is finite, this stage of the algorithm will terminate at some step $N' < \infty$. We then write $\mathcal{A}_x = \mathcal{A}_{N'}$ for all $\mathcal{A} \in \{G, F, Z\}$. The sets F_x and Z_x are disjoint, and their union is U_x. The occupancy of edges in F_x is known. In particular, the graph induced by set of open edges in $E_x \cup F_x$ is G_x. The occupancy of edges in Z_x has not been checked. In particular, given the set Z_x, the edges in Z_x are open independently of each other. By the definition of Z_x, any edge $e \in Z_x$ is in a long cycle in the graph $G_x \cup \{e\}$, and every long cycle of the graph $G_x \cup \{e\}$ passes through e. Moreover, by the special ordering of the vertices in B_x, any edge $e \in Z_x$ can be written as $\{a, b\}$ so that the unique path from b to x in the spanning tree T_x is edge disjoint from some long cycle of the graph $G_x \cup \{e\}$ (but not necessarily from all long cycles of the graph $G_x \cup \{e\}$).

We run the above defined exploration algorithm for all the open clusters of the torus. We pick a vertex x_1 uniformly on the torus and determine the depth-first spanning tree of the cluster of x_1 with root at x_1, T_{x_1}, the set of explored edges $E_{x_1} \cup F_{x_1}$, and the set of special edges Z_{x_1}. We then pick a vertex x_2 uniformly from the remaining vertices and determine the depth-first spanning tree of $\mathcal{C}_{\mathbb{T}}(x_2)$, T_{x_2}, the set of explored edged $E_{x_2} \cup F_{x_2}$, and the set of special edges Z_{x_2}. We then proceed similarly by selecting $x_3, \ldots, x_M$ and determining T_{x_i}, $E_{x_i} \cup F_{x_i}$ and Z_{x_i}. Here $M = M(\omega)$ is the number of open clusters in the realization ω.

Given the sets of explored edges $E_{x_i} \cup F_{x_i}$, the number of long cycles is defined by the status of the special edges Z_{x_i}. In particular, if all the edges in Z_{x_i} are closed, then $\mathcal{C}_{\mathbb{T}}(x_i)$ does *not* contain long cycles. Note that given the set of explored edges $E_{x_i} \cup F_{x_i}$, the event that all the edges in Z_{x_i} are closed has probability

$$(1 - p_c)^{|Z_{x_i}|}.$$

19

Also, the size of a cluster is determined by the number of vertices in a spanning tree. Therefore, (remember that M is the number of open clusters in the torus)

$$\mathbb{P}_{p_c}\left(Y_\delta = 0\right) = \mathbb{E}_{p_c}\left[(1-p_c)^{\sum_{i=1}^{M}|Z_{x_i}|I(|\mathcal{C}(x_i)|>\delta V^{2/3})}\right] \geq (1-p_c)^{\mathbb{E}_{p_c}\left[\sum_{i=1}^{M}|Z_{x_i}|I(|\mathcal{C}(x_i)|>\delta V^{2/3})\right]}. \tag{5.8}$$

The last step follows from Jensen's inequality.

Let $\mathcal{C}_{(1)},\ldots,\mathcal{C}_{(M)}$ be all the clusters of the torus sorted from the largest (in the number of vertices) to the smallest with ties broken in an arbitrary way. We will show that

$$\mathbb{E}_{p_c}\left[\sum_{i=1}^{M}|Z_{x_i}|I(|\mathcal{C}_{\mathbb{T}}(x_i)| > \delta V^{2/3})\right] = \mathbb{E}_{p_c}\left[\sum_{i=1}^{M}\sum_{x\in\mathcal{C}_{(i)}}\frac{1}{|\mathcal{C}_{(i)}|}|Z_x|I(|\mathcal{C}_{(i)}| > \delta V^{2/3})\right]. \tag{5.9}$$

Fix a percolation realization on the torus. Remember the way we select the vertices $x_1,\ldots,x_M$: select x_1 uniformly on the torus, select x_2 uniformly on $\mathbb{T}_r^d \setminus \mathcal{C}_{\mathbb{T}}(x_1)$, select x_3 uniformly on $\mathbb{T}_r^d \setminus (\mathcal{C}_{\mathbb{T}}(x_1) \cup \mathcal{C}_{\mathbb{T}}(x_2))$, and so on. Given a percolation realization on the torus, we can select the vertices $x_1,\ldots,x_M$ in two steps: first select a permutation σ of $\{1,\ldots,M\}$ (the distribution of σ is irrelevant to us), and then select x_i uniformly from $\mathcal{C}_{(\sigma(i))}$. Note that the sum $\sum_{i=1}^{M}|Z_{x_i}|I(|\mathcal{C}_{\mathbb{T}}(x_i)| > \delta V^{2/3})$ does not depend on σ, i.e., on the order in which we select clusters, and only depends on which points in clusters we select as x_i's. This implies (5.9).

Note that

$$\mathbb{E}_{p_c}\left[\sum_{i=1}^{M}\sum_{x\in\mathcal{C}_{(i)}}\frac{1}{|\mathcal{C}_{(i)}|}|Z_x|I(|\mathcal{C}_{(i)}| > \delta V^{2/3})\right] \leq \frac{1}{\delta V^{2/3}}\mathbb{E}_{p_c}\left[\sum_{i=1}^{M}\sum_{x\in\mathcal{C}_{(i)}}|Z_x|\right] = \frac{1}{\delta V^{2/3}}\sum_{x\in\mathbb{T}_r^d}\mathbb{E}_{p_c}[|Z_x|].$$

Therefore, it follows from (5.8) and (5.9) that

$$\mathbb{P}_{p_c}\left(Y_\delta = 0\right) \geq (1-p_c)^{\delta^{-1}V^{1/3}\mathbb{E}_{p_c}[|Z_0|]}.$$

Proposition 5.1 follows once we show that

$$\mathbb{E}_{p_c}[|Z_0|] \leq C_{11}V^{-1/3}. \tag{5.10}$$

Recall that $\mathcal{I}$ is the set of $z \in \mathcal{C}_{\mathbb{T}}(0)$ such that $\{0 \leftrightarrow z\} \circ \{z \text{ is in a long cycle}\}$, and let $\mathcal{E}$ be the set of edges with at least one end-vertex in $\mathcal{I}$. By the properties of Z_0, if $e \in Z_0$ is open, then $e \in \mathcal{E}$. Therefore,

$$\mathbb{E}_{p_c}[|Z_0|] = \frac{1}{p_c}\sum_e \mathbb{P}_{p_c}(e \in Z_0, e \text{ is open}) \leq \frac{1}{p_c}\mathbb{E}_{p_c}[|\mathcal{E}|] \leq \frac{1}{p_c}2d\mathbb{E}_{p_c}[|\mathcal{I}|].$$

The claim (5.10) now follows from Lemma 5.1. This completes the proof of Proposition 5.1. $\qquad\square$

In the remainder of this section, we prove some properties of the exploration algorithm defined in the proof of Proposition 5.1. Remember the notation used in the description of the algorithm.

Lemma 5.3 (Structure depth-first tree). *(a) For $n \geq 0$, let $\widetilde{X}_n$ be the set of vertices $a' \in X_n$ for which there exists $b' \in \mathbb{T}_r^d$ such that the edge $\{a',b'\} \notin W_n$. For each $n \geq 0$, there exists a unique vertex $a \in \widetilde{X}_n$ which is the farthest from x in the tree T_n, and all the other vertices from $\widetilde{X}_n$ belong to the unique path from a to the root x in T_n.*
(b) For all $e \in U_x$, there exist $a, b \in X_x$ such that $e = \{a,b\}$ and the unique path from a to x in the tree T_x passes through b.

Proof. The proof of part (a) is by induction on n. The result is obvious for $n = 0$, since $\widetilde{X}_0 = \{x\}$. Assume that the result holds for all $n' \leq n$. Pick the unique vertex $a \in \widetilde{X}_n$ which is the farthest from x in T_n. Let $\{a,b\} \notin W_n$. (If there are several choices, then we pick the smallest edge according to the numbering of the edges of the torus.) If $\{a,b\}$ satisfies 1(a) of the first stage of the algorithm, then

$b \in \widetilde{X}_{n+1}$, and the unique path from b to x in T_{n+1} contains $\widetilde{X}_n$, by the induction assumption. If $\{a, b\}$ satisfies 1(b) or 2 of the first stage of the algorithm, then either $\widetilde{X}_{n+1} = \widetilde{X}_n$ (if there is more than one edge $\{a, b'\} \notin W_n$) or $\widetilde{X}_{n+1} = \widetilde{X}_n \setminus \{a\}$ (if there is the unique edge $\{a, b'\} \notin W_n$). In both cases $\widetilde{X}_{n+1}$ satisfies the statement in part (a) of the lemma, by the induction hypothesis. This completes the proof of (a).

To prove part (b), let $e \in U_x$. There exists $n \geq 0$ such that $U_{n+1} \setminus U_n = \{e\}$. By the definition of the algorithm, there exist a and b such that $e = \{a, b\}$ and a is the farthest vertex in $\widetilde{X}_n$ from x in T_n. Note that the edge e satisfies condition 2 of the first stage of the algorithm. In particular, $b \in X_n$ and $e \notin W_n$. Therefore, $b \in \widetilde{X}_n$. The result in part (b) now follows from part (a) of the lemma. $\qquad \square$

6 Proof of Theorem 1.5

In this section, we restrict to percolation on $\mathbb{Z}^d$. In particular, all the paths are assumed by default to be in $\mathbb{Z}^d$, and we write here $\{x \leftrightarrow y\}$ for $\{x \leftrightarrow y \text{ in } \mathbb{Z}^d\}$ and $\mathcal{C}(x)$ for $\mathcal{C}_{\mathbb{Z}}(x)$. This section is organized as follows. In Section 6.1, we start with some preparatory lemmas based on the techniques in [19, 20]. We prove Theorem 1.5(a) in Section 6.2, and Theorem 1.5(b) in Section 6.3.

6.1 Preparatory lemmas

The following lemma produces the factor $\sqrt{\varepsilon}$ that is present in Theorem 1.5(b):

Lemma 6.1. *There exists $C < \infty$ such that for any $\varepsilon > 0$, positive integer n, and $x \in \mathbb{Z}^d$ with $|x| \geq n^2$,*

$$\mathbb{P}_{p_c}\left(0 \overset{\leq \varepsilon n^2}{\longleftrightarrow} \partial Q_n, 0 \leftrightarrow x\right) \leq C\sqrt{\varepsilon}\,\mathbb{P}_{p_c}(0 \leftrightarrow x).$$

Proof of Lemma 6.1. This proof is a slight modification of the proof of [20, Lemma 2.5]. The event $E = \{0 \overset{\leq \varepsilon n^2}{\longleftrightarrow} \partial Q_n\}$ is measurable with respect to $\mathcal{B}_I(\varepsilon n^2) = \{x \in \mathbb{Z}^d \colon 0 \overset{\leq \varepsilon n^2}{\longleftrightarrow} x\}$. Therefore, [20, Lemma 2.5] implies that for any $x \in \mathbb{Z}^d$ with $|x|$ sufficiently large,

$$\mathbb{P}_{p_c}\left(0 \overset{\leq \varepsilon n^2}{\longleftrightarrow} \partial Q_n, 0 \leftrightarrow x\right) \leq C_1 \sqrt{\varepsilon n^2 \mathbb{P}_{p_c}(E)}\,\mathbb{P}_{p_c}(0 \leftrightarrow x).$$

In fact, it follows from the proof of [20, Lemma 2.5] that the above inequality holds for all $x \in \mathbb{Z}^d$ with $|x| \geq n^2$. Finally, remember that $\mathbb{P}_{p_c}(E) \leq C_2 n^{-2}$ by (2.2). This completes the proof. $\qquad \square$

It follows from Lemma 6.1 and (2.3) that

$$\sum_{x \in Q_{2n^2} \setminus Q_{n^2}} \mathbb{P}_{p_c}\left(0 \overset{\leq \varepsilon n^2}{\longleftrightarrow} \partial Q_n, 0 \leftrightarrow x\right) \leq C_3 \sqrt{\varepsilon}\, n^4, \tag{6.1}$$

which shall be used crucially later on.

We next recall some notation from [19]. Recall the definition of a K-regular vertex from [19, Definition 4.1]: For $A \subseteq \mathbb{Z}^d$, let $\mathcal{C}(y; A)$ be the set of vertices z such that $y \leftrightarrow z$ in A. For $y \in \partial Q_n$ and positive integers s and K, we say that y is *s-bad* if $\mathcal{C}(y; Q_n)$ satisfies

$$\mathbb{P}_{p_c}\left(|\mathcal{C}(y) \cap Q_s(y)| < s^4 \log^7 s \mid \mathcal{C}(y; Q_n)\right) \leq 1 - \exp(-\log^2 s).$$

We further say that $y \in \partial Q_n$ is *K-irregular* if there exists $s \geq K$ such that y is s-bad. Otherwise we say that y is *K-regular*.

We say that a pair of vertices (x, y) is (n, K, ε)-admissible if the following conditions hold: (a) $y \in \partial Q_n$ and $x \in Q_{2n^2} \setminus Q_{n^2}$; (b) $0 \overset{\leq \varepsilon n^2}{\longleftrightarrow} y$ in Q_n and $y \leftrightarrow x$; (c) y is K-regular; and (d) the edge $\{y, \widetilde{y}\}$ is pivotal for the event $0 \leftrightarrow x$, where $\widetilde{y}$ is the neighbor of y not in Q_n (if more than one exist, then we choose the first in lexicographical order).

We define $Y(n, K, \varepsilon)$ as the number of (n, K, ε)-admissible pairs. The random variable $Y(n, K, \varepsilon)$ is very similar to $Y(j, K, L)$ defined in the proof of [19, Lemma 5.1].

Remark 1. Note that [19, Lemma 5.1] holds for all $M \geq 1$ and not just for $M \geq L^2/2$ as it is stated. Indeed, [19, Lemma 5.1] follows directly from [19, Lemmas 5.3 and 5.5], both of which hold (and are stated) for all $M \geq 1$.

Lemma 6.2. *There exists a positive constant $C_4 = C_4(K)$ such that for K sufficiently large, any positive integer n and $\varepsilon > 0$,*

$$\mathbb{E}_{p_c} Y(n, K, \varepsilon) \geq C_4(K) n^4 \mathbb{E}_{p_c} |\{y \in \partial Q_n : 0 \overset{\leq \varepsilon n^2}{\longleftrightarrow} y \text{ in } Q_n, y \text{ is } K\text{-regular}\}|.$$

Proof. A word for word repetition of the proof of [19, Lemma 5.1] (taking into account Remark 1) gives Lemma 6.2. Indeed, with the notation of [19], the only difference in the proofs arises in the proof (and the statement) of [19, Lemma 5.3], where instead of the event

$$\mathcal{E}_1 = \{0 \leftrightarrow y \text{ in } Q_j, y \text{ is } K - \text{regular and } X_j^{K-reg} = M\},$$

we use the event

$$\widetilde{\mathcal{E}}_1 = \{0 \overset{\leq \varepsilon j^2}{\longleftrightarrow} y \text{ in } Q_j, y \text{ is } K - \text{regular and } X_j^{K-reg} = M\}.$$

However, the event $\widetilde{\mathcal{E}}_1$ still can be determined by observing only the edges of $\mathcal{C}(0; Q_j)$. Therefore, the proof of Lemma 5.3 in [19] remains unchanged if we replace the event $\mathcal{E}_1$ with the event $\widetilde{\mathcal{E}}_1$. The proof of Lemma 5.5 in [19] also requires only that the event $\mathcal{E}_1$ must be determined by observing only the edges of $\mathcal{C}(0; Q_j)$, and therefore also holds with $\mathcal{E}_1$ replaced with $\widetilde{\mathcal{E}}_1$. $\qquad\square$

Note that for every $x \in Q_{2n^2} \setminus Q_{n^2}$ there exists at most one $y \in \partial Q_n$ such that the pair of vertices (x, y) is (n, K, ε)-admissible. Therefore,

$$\mathbb{E}_{p_c} Y(n, K, \varepsilon) = \sum_{x \in Q_{2n^2} \setminus Q_{n^2}} \mathbb{P}_{p_c} \Big(\exists y \in \partial Q_n : (x, y) \text{ is } (n, K, \varepsilon)\text{-admissible} \Big).$$

If (x, y) is (n, K, ε)-admissible, then $\{0 \overset{\leq \varepsilon n^2}{\longleftrightarrow} y \text{ in } Q_n\}$ and $\{0 \leftrightarrow x\}$ both occur. We use this observation to bound the expected number of (n, K, ε)-admissible pairs from above by

$$\sum_{x \in Q_{2n^2} \setminus Q_{n^2}} \mathbb{P}_{p_c} \left(\exists y \in \partial Q_n : 0 \overset{\leq \varepsilon n^2}{\longleftrightarrow} y \text{ in } Q_n, 0 \leftrightarrow x \right) = \sum_{x \in Q_{2n^2} \setminus Q_{n^2}} \mathbb{P}_{p_c} \left(0 \overset{\leq \varepsilon n^2}{\longleftrightarrow} \partial Q_n, 0 \leftrightarrow x \right).$$

We can now combine these bounds with the results of (6.1) and Lemma 6.2 to get

$$\mathbb{E}_{p_c} \left| \{y \in \partial Q_n : 0 \overset{\leq \varepsilon n^2}{\longleftrightarrow} y \text{ in } Q_n, y \text{ is } K\text{-regular}\} \right| \leq \frac{C_3}{C_4(K)} \sqrt{\varepsilon} \tag{6.2}$$

for all large enough K, positive integers n and for all $\varepsilon > 0$ with the constants C_3 from (6.1) (not depending on K, n and ε) and $C_4(K)$ from Lemma 6.2 (not depending on n and ε). We next investigate the contribution from K-irregular y's:

Lemma 6.3. *For all large enough K and for all $\varepsilon > 0$,*

$$\limsup_{n \to \infty} \sum_{y \in \partial Q_n} \mathbb{P}_{p_c} \left(0 \overset{\leq \varepsilon n^2}{\longleftrightarrow} y \text{ in } Q_n, y \text{ is } K\text{-irregular} \right) \leq \frac{1}{2} \limsup_{n \to \infty} \sum_{y \in \partial Q_n} \mathbb{P}_{p_c} \left(0 \overset{\leq 2\varepsilon n^2}{\longleftrightarrow} y \text{ in } Q_n \right).$$

Proof of Lemma 6.3. Recall the definition of an s-locally bad vertex from [19, Definition 4.3]: Let $\mathcal{T}_s^{loc}(y)$ be the event that (a) for all $z \in Q_s(y)$, $|\mathcal{C}(z; Q_{s^{2d}}(y)) \cap Q_s(y)| < s^4 \log^4 s$; and (b) there exist at most $\log^3 s$ disjoint open paths starting in $Q_s(y)$ and ending at $\partial Q_{s^{2d}}(y)$. For $y \in \partial Q_n$ and positive integers s and K, we say that a cluster $\mathcal{C}$ in $Q_{s^{4d^2}}(y) \cap Q_n$ is a "spanning cluster" if (a) $\mathcal{C} \cap Q_n$ intersects both

$\partial Q_{s^{4d^2}}(y)$ and $Q_{s^{2d}}(y)$, or (b) $\mathcal{C} = \mathcal{C}(y)$. We say that y is s-*locally bad* if there exist spanning clusters $\mathcal{C}_1, \ldots, \mathcal{C}_m$ in $Q_{s^{4d^2}}(y) \cap Q_n$ such that

$$\mathbb{P}_{p_c}\left(\mathcal{T}_s^{loc}(y) \mid \mathcal{C}_1, \ldots, \mathcal{C}_m\right) \le 1 - \exp(-\log^2 s).$$

Note that the event that y is s-locally bad is determined by the status of the edges in the box $Q_{s^{4d^2}}(y) \cap Q_n$. Moreover, it follows from [19, Claim 4.2] that if y is not K-regular, then there exists $s \ge K$ such that y is s-locally bad. Therefore, we need to bound from above the probabilities

$$\mathbb{P}_{p_c}\left(0 \overset{\le \varepsilon n^2}{\longleftrightarrow} y \text{ in } Q_n, y \text{ is } s\text{-locally bad}\right)$$

for $y \in \partial Q_n$, $s \ge K$ and large enough n.

Since we are only interested in large n, we may assume that $n\varepsilon > 1$. We consider two different cases: $2d(s^{4d^2})^d < n$ and $2d(s^{4d^2})^d \ge n$. We start with the case $2d(s^{4d^2})^d < n$. Note that in this case the ball $Q_{s^{4d^2}}$ contains at most εn^2 edges. We bound the sum

$$\sum_{y \in \partial Q_n} \mathbb{P}_{p_c}\left(0 \overset{\le \varepsilon n^2}{\longleftrightarrow} y \text{ in } Q_n, y \text{ is } s\text{-locally bad}\right)$$

from above by

$$\sum_{y \in \partial Q_n} \mathbb{P}_{p_c}\left(0 \overset{\le \varepsilon n^2}{\longleftrightarrow} Q_{s^{4d^2}}(y) \text{ in } Q_n, y \text{ is } s\text{-locally bad}\right).$$

Since the events $\{0 \overset{\le \varepsilon n^2}{\longleftrightarrow} Q_{s^{4d^2}}(y) \text{ in } Q_n\}$ and $\{y \text{ is } s\text{-locally bad}\}$ depend on the states of edges in disjoint subsets of $\mathbb{Z}^d$, they are independent. In particular, the above sum equals

$$\sum_{y \in \partial Q_n} \mathbb{P}_{p_c}\left(0 \overset{\le \varepsilon n^2}{\longleftrightarrow} Q_{s^{4d^2}}(y) \text{ in } Q_n\right) \mathbb{P}_{p_c}\left(y \text{ is } s\text{-locally bad}\right).$$

By [19, Lemma 1.1] and the FKG inequality (see, e.g., [9, Theorem 2.4]), there exist a positive constant C_5 and a finite constant C_6 such that for all m and $z, z' \in \partial Q_m$,

$$\mathbb{P}_{p_c}(z \leftrightarrow z' \text{ in } Q_m) \ge C_5 \exp\left(-C_6 \log^2 m\right).$$

We apply this result to "extend" the path $0 \overset{\le \varepsilon n^2}{\longleftrightarrow} Q_{s^{4d^2}}(y)$ in Q_n to a path $0 \overset{\le 2\varepsilon n^2}{\longleftrightarrow} y$ in Q_n:

$$\mathbb{P}_{p_c}\left(0 \overset{\le \varepsilon n^2}{\longleftrightarrow} Q_{s^{4d^2}}(y) \text{ in } Q_n\right) \le C_7 \exp\left(C_8 \log^2 s\right) \mathbb{P}_{p_c}\left(0 \overset{\le 2\varepsilon n^2}{\longleftrightarrow} y \text{ in } Q_n\right).$$

Here we also use the fact that the number of edges in $Q_{s^{4d^2}}(y)$ is at most εn^2, which implies that if two vertices z and z' in $Q_{s^{4d^2}}(y)$ are connected by an open path in $Q_{s^{4d^2}}(y)$ then the length of this path is at most εn^2.

It follows from [19, Lemma 4.3] that

$$\mathbb{P}_{p_c}\left(y \text{ is } s\text{-locally bad}\right) \le C_9 \exp\left(-C_{10} \log^4 s\right).$$

We now put these bounds together. Let $\sum'$ be the sum over all s such that $s \ge K$ and $2d(s^{4d^2})^d < n$. We obtain that

$$\sum' \sum_{y \in \partial Q_n} \mathbb{P}_{p_c}\left(0 \overset{\le \varepsilon n^2}{\longleftrightarrow} y \text{ in } Q_n, y \text{ is } s\text{-locally bad}\right)$$

is bounded from above by

$$C_{11} \exp\left(-C_{12} \log^4 K\right) \sum_{y \in \partial Q_n} \mathbb{P}_{p_c}(0 \overset{\le 2\varepsilon n^2}{\longleftrightarrow} y \text{ in } Q_n),$$

where the constants C_{11} and C_{12} do not depend on n, ε or K.

We now consider the case $2d(s^{4d^2})^d \geq n$. Let $s_0 = (n/2d)^{1/(4d^3)}$. For $s \geq s_0$, we simply bound

$$\sum_{s=s_0}^{\infty} \sum_{y \in \partial Q_n} \mathbb{P}_{p_c} \left(0 \overset{\leq \varepsilon n^2}{\longleftrightarrow} y \text{ in } Q_n, y \text{ is } s\text{-locally bad} \right) \leq |\partial Q_n| \sum_{s=s_0}^{\infty} \sup_{y \in \partial Q_n} \mathbb{P}_{p_c} \left(y \text{ is } s\text{-locally bad} \right).$$

We again use [19, Lemma 4.3] to bound the above expression by

$$|\partial Q_n| \sum_{s=s_0}^{\infty} C_9 \exp \left(-C_{10} \log^4 s \right) \leq C_{13} \exp \left(-C_{14} \log^4 n \right),$$

since $s_0 = (n/2d)^{1/(4d^3)}$, and where the constants C_{13} and C_{14} do not depend on n, K, or ε.

We take K so large that $C_{11} \exp \left(-C_{12} \log^4 K \right) < 1/2$. Remember [19, Claim 4.2] which states that if y is K-irregular, then there exists $s \geq K$ such that y is s-locally bad. Therefore, for such choice of K, the sum

$$\sum_{y \in \partial Q_n} \mathbb{P}_{p_c} \left(0 \overset{\leq \varepsilon n^2}{\longleftrightarrow} y \text{ in } Q_n, y \text{ is } K\text{-irregular} \right) \leq \frac{1}{2} \sum_{y \in \partial Q_n} \mathbb{P}_{p_c} \left(0 \overset{\leq 2\varepsilon n^2}{\longleftrightarrow} y \text{ in } Q_n \right) + C_{13} \exp \left(-C_{14} \log^4 n \right).$$

$$(6.3)$$

The result of Lemma 6.3 follows. □

6.2 Proof of Theorem 1.5(a)

The proof of Theorem 1.5(a) is similar to the proof of Lemma 6.3, but easier. We refer the reader to Section 6.1 for definitions and notation. Remember the definition of a K-regular vertex from Section 6.1. Let X_n^{K-reg} be the number of K-regular vertices on the boundary of Q_n connected to the origin by an open path in Q_n.

Let $Y(n, K, L)$ be the random variable defined in the proof of [19, Theorem 2]: We say that a pair of vertices (x, y) are (n, K, L)-admissible if the following conditions hold: (a) $y \in \partial Q_n$ and $x \in Q_L(y)$; (b) $0 \leftrightarrow y$ in Q_n and $y \leftrightarrow x$; (c) y is K-regular; and (d) the edge $\{y, \widetilde{y}\}$ is pivotal for the event $0 \leftrightarrow x$, where $\widetilde{y}$ is the neighbor of y not in Q_n (if more than one exist, we choose the first in lexicographical order). We define $Y(n, K, L)$ as the number of (n, K, L)-admissible pairs.

It follows from [19, Lemma 5.1] and Remark 1 that there exists $C_{15} = C_{15}(K)$ such that for all large enough K and for all n and L,

$$\mathbb{E}_{p_c} Y(n, K, L) \geq C_{15}(K) L^2 \mathbb{E}_{p_c} X_n^{K-reg}. \tag{6.4}$$

Lemma 6.4. *For all large enough K and for all n,*

$$\mathbb{E}_{p_c} X_n^{K-reg} \geq \frac{1}{3} \sum_{y \in \partial Q_n} \mathbb{P}_{p_c}(0 \leftrightarrow y \text{ in } Q_n).$$

Proof. The proof of this lemma is very similar to the proof of Lemma 6.3, but simpler. In the same way as in the proof of Lemma 6.3 and with the same choice of K, we bound

$$\sum_{y \in \partial Q_n} \mathbb{P}_{p_c} \left(0 \leftrightarrow y \text{ in } Q_n, y \text{ is not } K\text{-regular} \right) \leq \frac{1}{2} \sum_{y \in \partial Q_n} \mathbb{P}_{p_c} \left(0 \leftrightarrow y \text{ in } Q_n \right) + C_{13} \exp \left(-C_{14} \log^4 K \right).$$

We then use the result of [19, Lemma 3.1]: For all positive integers n,

$$\sum_{y \in \partial Q_n} \mathbb{P}_{p_c}(0 \leftrightarrow y \text{ in } Q_n) \geq 1.$$

We increase K if necessary to fulfill the bound $C_{13} \exp \left(-C_{14} \log^4 K \right) < 1/6$. The result follows. □

Theorem 1.5(a) follows in a straightforward way from (6.4), Lemma 6.4 and the fact that

$$\mathbb{E}_{p_c} Y(n, K, n) \leq \sum_{x \in Q_{2n}} \mathbb{P}_{p_c}(0 \leftrightarrow x) \leq C_{16} n^2.$$

The last inequality follows from (2.3). □

6.3 Proof of Theorem 1.5(b)

Let

$$F(\varepsilon) = \limsup_{n \to \infty} \sum_{y \in \partial Q_n} \mathbb{P}_{p_c}\left(0 \overset{\leq \varepsilon n^2}{\longleftrightarrow} y \text{ in } Q_n\right).$$

Theorem 1.5(a) implies that there exists a finite constant C_{17} such that $F(\varepsilon) \leq C_{17}$ for all $\varepsilon > 0$. It follows from (6.2) and Lemma 6.3 that for all $\varepsilon > 0$,

$$F(\varepsilon) \leq \frac{C_3}{C_4}\sqrt{\varepsilon} + \frac{1}{2}F(2\varepsilon).$$

We apply the above inequality k times to get

$$F(\varepsilon) \leq C_{18}\sqrt{\varepsilon} + \frac{1}{2^k}F(2^k\varepsilon) \leq C_{18}\sqrt{\varepsilon} + \frac{C_{17}}{2^k},$$

with $C_{18} = C_3\sqrt{2}/C_4(\sqrt{2} - 1)$ and where we use that $F(2^k\varepsilon) \leq C_{17}$ by Theorem 1.5(a). This inequality holds for any fixed k. We complete the proof of Theorem 1.5(b) by taking k such that $2^k\sqrt{\varepsilon} > 1$. □

7 Proof of Theorem 1.6

Proof of (1.8). Let k and r be positive integers and $z \in \mathbb{Z}^d$. For brevity, we write $r/4$ instead of $\lfloor r/4 \rfloor$. It suffices to prove the result for all large enough k. By Theorem 1.5(b), there exist A and K such that for any $k \geq K$,

$$\sum_{y \in \partial Q_{A\sqrt{k}}} \mathbb{P}_{p_c}\left(0 \overset{\leq k}{\longleftrightarrow} y \text{ in } Q_{A\sqrt{k}}\right) \leq \frac{1}{2}. \tag{7.1}$$

Indeed, fix $\varepsilon > 0$ such that the right hand side of (1.7) is strictly smaller than $1/2$. Then by (1.7), inequality (7.1) holds for all large k with $A = 1/\sqrt{\varepsilon}$.

From now on we assume that $k \geq K$. We first consider the case $A\sqrt{k} \leq r/8$. We write

$$\sum_{x \in \mathbb{Z}^d,\ x \overset{r}{\sim} z,\ |x| \geq r/4} \mathbb{P}_{\mathbb{Z}, p_c}\left(0 \overset{\leq k}{\longleftrightarrow} x \text{ in } \mathbb{Z}^d\right) \leq \sum_{n=0}^{\infty} \sum_{x \overset{r}{\sim} z,\ r/4 + rn \leq |x| < r/4 + r(n+1)} \mathbb{P}_{\mathbb{Z}, p_c}\left(0 \overset{\leq k}{\longleftrightarrow} x \text{ in } \mathbb{Z}^d\right)$$

$$\leq \sum_{n=0}^{\infty} C_1(n+1)^{d-1} \sup_{r/4 + rn \leq |x| < r/4 + r(n+1)} \mathbb{P}_{\mathbb{Z}, p_c}\left(0 \overset{\leq k}{\longleftrightarrow} x \text{ in } \mathbb{Z}^d\right).$$

Here we use the fact that, uniformly in r, the number of $x \overset{r}{\sim} z$ such that $r/4 + rn \leq |x| < r/4 + r(n+1)$ is of order $(n+1)^{d-1}$.

Take $n \geq 0$ and $x \in \mathbb{Z}^d$ with $r/4 + rn \leq |x| < r/4 + r(n+1)$. Let $M_n = \lfloor \frac{r/8 + rn}{A\sqrt{k}} \rfloor \geq 1$, since $A\sqrt{k} \leq r/8$.

Note that the event $\{0 \overset{\leq k}{\longleftrightarrow} x \text{ in } \mathbb{Z}^d\}$ implies the existence of $x_1, \ldots, x_{M_n}$ such that for all $i \in \{1, \ldots, M_n\}$, $x_i \in \partial Q_{A\sqrt{k}}(x_{i-1})$ (where we assume $x_0 = 0$) and the following connections all occur disjointly:

$$\{x_{i-1} \overset{\leq k}{\longleftrightarrow} x_i \text{ in } Q_{A\sqrt{k}}(x_{i-1})\}, \quad i \in \{1, \ldots, M_n\}, \quad \text{and} \quad \{x_{M_n} \leftrightarrow x\}.$$

By the BK inequality, translation invariance, (7.1), the fact that $|x_{M_n} - x| \geq r/8$, and (1.1), for any $n \geq 0$,

$$\sup_{r/4 + rn \leq |x| < r/4 + r(n+1)} \mathbb{P}_{\mathbb{Z}, p_c} \left(0 \overset{\leq k}{\longleftrightarrow} x \text{ in } \mathbb{Z}^d \right) \leq C_2 r^{2-d} \left(\frac{1}{2} \right)^{M_n}.$$

Putting all the bounds together and using that $M_n \geq n + M_0$ (since $A\sqrt{k} \leq r/8$), we have in the case $A\sqrt{k} \leq r/8$ that

$$\sum_{x \in \mathbb{Z}^d, \ x \overset{r}{\sim} z, \ |x| \geq r/4} \mathbb{P}_{\mathbb{Z}, p_c} \left(0 \overset{\leq k}{\longleftrightarrow} x \text{ in } \mathbb{Z}^d \right) \leq C_1 C_2 r^{2-d} \sum_{n=0}^{\infty} (n+1)^{d-1} \left(\frac{1}{2} \right)^{M_n} \leq C_3 r^{2-d} 2^{-\frac{r}{8A\sqrt{k}}}.$$

This finishes the proof of (1.8) in the case $A\sqrt{k} \leq r/8$.

It remains to consider the case $A\sqrt{k} > r/8$. We write

$$\sum_{x \in \mathbb{Z}^d, \ x \overset{r}{\sim} z, \ |x| \geq r/4} \mathbb{P}_{\mathbb{Z}, p_c} \left(0 \overset{\leq k}{\longleftrightarrow} x \text{ in } \mathbb{Z}^d \right) \ \leq \sum_{x \in Q_{8A\sqrt{k}}, \ x \overset{r}{\sim} z, \ |x| \geq r/8} \mathbb{P}_{\mathbb{Z}, p_c} \left(0 \leftrightarrow x \text{ in } \mathbb{Z}^d \right)$$

$$+ \sum_{n=8}^{\infty} \sum_{x \overset{r}{\sim} z, \ nA\sqrt{k} < |x| \leq (n+1)A\sqrt{k}} \mathbb{P}_{\mathbb{Z}, p_c} \left(0 \overset{\leq k}{\longleftrightarrow} x \text{ in } \mathbb{Z}^d \right).$$

Using (2.4), the first sum can be bounded from above by $C_4 k / r^d$.

Take $n \geq 8$ and $x \in \mathbb{Z}^d$ with $nA\sqrt{k} < |x| \leq (n+1)A\sqrt{k}$. Note that the event $\{0 \overset{\leq k}{\longleftrightarrow} x \text{ in } \mathbb{Z}^d\}$ implies the existence of $x_1, \ldots, x_{n-1}$ such that for all $i \in \{1, \ldots, n-1\}$, $x_i \in \partial Q_{A\sqrt{k}}(x_{i-1})$ (where we assume $x_0 = 0$) and the following connections all occur disjointly:

$$\{ x_{i-1} \overset{\leq k}{\longleftrightarrow} x_i \text{ in } Q_{A\sqrt{k}}(x_{i-1}) \}, \qquad i \in \{1, \ldots, n-1\}, \qquad \text{and} \qquad \{ x_{n-1} \leftrightarrow x \}.$$

By the BK inequality,

$$\sum_{x \overset{r}{\sim} z, \ nA\sqrt{k} < |x| \leq (n+1)A\sqrt{k}} \mathbb{P}_{\mathbb{Z}, p_c} \left(0 \overset{\leq k}{\longleftrightarrow} x \text{ in } \mathbb{Z}^d \right)$$

$$\leq \sum_{x_1, \ldots, x_{n-1}} \prod_{i=1}^{n-1} \mathbb{P}_{\mathbb{Z}, p_c} \left(x_{i-1} \overset{\leq k}{\longleftrightarrow} x_i \text{ in } Q_{A\sqrt{k}}(x_{i-1}) \right) \sum_{x \overset{r}{\sim} z, \ nA\sqrt{k} < |x| \leq (n+1)A\sqrt{k}} \mathbb{P}_{\mathbb{Z}, p_c} \left(x_{n-1} \leftrightarrow x \text{ in } \mathbb{Z}^d \right).$$

Note that for any choice of $x_1, \ldots, x_{n-1}$ as above, we have $r/8 < A\sqrt{k} \leq |x_{n-1} - x| \leq 2nA\sqrt{k}$. Therefore, by translation invariance and (2.4), we obtain

$$\sum_{x \overset{r}{\sim} z, \ nA\sqrt{k} < |x| \leq (n+1)A\sqrt{k}} \mathbb{P}_{\mathbb{Z}, p_c} \left(x_{n-1} \leftrightarrow x \text{ in } \mathbb{Z}^d \right) \ \leq \sum_{x \in Q_{2nA\sqrt{k}}, \ x \overset{r}{\sim} z - x_{n-1}, \ |x| \geq r/8} \mathbb{P}_{\mathbb{Z}, p_c} \left(0 \leftrightarrow x \text{ in } \mathbb{Z}^d \right)$$

$$\leq \frac{C_5 k}{r^d}.$$

Putting all the bounds together and using translation invariance and (7.1), we get

$$\sum_{x \in \mathbb{Z}^d, \ x \overset{r}{\sim} z, \ |x| \geq r/4} \mathbb{P}_{\mathbb{Z}, p_c} \left(0 \overset{\leq k}{\longleftrightarrow} x \text{ in } \mathbb{Z}^d \right) \leq \frac{C_4 k}{r^d} + \sum_{n=8}^{\infty} \frac{C_5 k}{r^d} \left(\frac{1}{2} \right)^{n-1} \leq \frac{C_6 k}{r^d}.$$

This finishes the proof of (1.8) in the case $A\sqrt{k} > r/8$. □

Proof of (1.9). We bound the sum over u, v, and w by distinguishing three cases: $|u| \geq r/4$, $|v-w| \geq r/4$, and the remaining term.

If $|v - w| \geq r/4$, then (1.9) follows by applying (1.8). Indeed, by (1.8) we get one factor of k/r^d from summing over w, and by (1.10) two factors of k from the remaining two sums.

If $|u| \geq r/4$, we let $u' = u + w$ and $v' = v + w$, and replace the sums over u and v by sums over u' and v' and use translation invariance to obtain that (1.9) over this range equals

$$\sum_{w \overset{r}{\sim} z,\ |w| \geq 3r/4} \sum_{u',v':\ |u'-w| \geq r/4} \mathbb{P}_{\mathbb{Z},p_c}\left(w \overset{\leq k}{\longleftrightarrow} u' \text{ in } \mathbb{Z}^d\right) \mathbb{P}_{\mathbb{Z},p_c}\left(u' \overset{\leq k}{\longleftrightarrow} v' \text{ in } \mathbb{Z}^d\right) \mathbb{P}_{\mathbb{Z},p_c}\left(0 \overset{\leq k}{\longleftrightarrow} v' \text{ in } \mathbb{Z}^d\right) \leq \frac{Ck^3}{r^d},$$
(7.2)

where we use that $|u' - w| = |u| \geq r/4$, together with (1.8) and (1.10).

It remains to bound the sum over all u, v, and w with $w \overset{r}{\sim} z$, $|w| \geq 3r/4$, $|u| < r/4$, and $|v - w| < r/4$, which we denote by $\sum'$. By the triangle inequality, we have $|u - v| \geq r/4$. We write

$$\sum{}' \mathbb{P}_{\mathbb{Z},p_c}\left(0 \overset{\leq k}{\longleftrightarrow} u \text{ in } \mathbb{Z}^d\right) \mathbb{P}_{\mathbb{Z},p_c}\left(u \overset{\leq k}{\longleftrightarrow} v \text{ in } \mathbb{Z}^d\right) \mathbb{P}_{\mathbb{Z},p_c}\left(v \overset{\leq k}{\longleftrightarrow} w \text{ in } \mathbb{Z}^d\right)$$
$$= \sum{}' \mathbb{P}_{\mathbb{Z},p_c}\left(0 \overset{\leq k}{\longleftrightarrow} u \text{ in } \mathbb{Z}^d\right) \mathbb{P}_{\mathbb{Z},p_c}\left(u \overset{\leq k}{\longleftrightarrow} v \text{ in } \mathbb{Z}^d\right) \mathbb{P}_{\mathbb{Z},p_c}\left(0 \overset{\leq k}{\longleftrightarrow} v - w \text{ in } \mathbb{Z}^d\right).$$

Note that $v \overset{r}{\sim} v - w + z$ and $|u - v| \geq r/4$. With the change of variables $(v, x) = (v, v - w + z)$, we observe that the above sum is bounded from above by

$$\sum_{u,x} \sum_{v \overset{r}{\sim} x,\ |v-u| \geq r/4} \mathbb{P}_{\mathbb{Z},p_c}\left(0 \overset{\leq k}{\longleftrightarrow} u \text{ in } \mathbb{Z}^d\right) \mathbb{P}_{\mathbb{Z},p_c}\left(u \overset{\leq k}{\longleftrightarrow} v \text{ in } \mathbb{Z}^d\right) \mathbb{P}_{\mathbb{Z},p_c}\left(0 \overset{\leq k}{\longleftrightarrow} x \text{ in } \mathbb{Z}^d\right).$$

Applying (1.8) to the sum over v, and then (1.10) to the sums over u and x, we obtain that the above sum is bounded from above by $(C_7 k/r^d) \cdot C_7 k \cdot C_7 k$. Putting all the cases together, we arrive at (1.9). $\square$

8 Proof of Theorem 1.7

Proof of (1.11). Take $x \in \mathbb{T}_r^d$. Let $k = \lfloor V^{1/3} \rfloor$. We write

$$\mathbb{P}_{p_c}\left(0 \leftrightarrow x \text{ in } \mathbb{T}_r^d\right) = \mathbb{P}_{p_c}\left(0 \overset{\leq 3k}{\longleftrightarrow} x \text{ in } \mathbb{T}_r^d\right) + \mathbb{P}_{p_c}\left(\{0 \leftrightarrow x \text{ in } \mathbb{T}_r^d\} \setminus \{0 \overset{\leq 3k}{\longleftrightarrow} x \text{ in } \mathbb{T}_r^d\}\right).$$

It follows from Proposition 2.1, (1.8), and the choice of k that

$$\mathbb{P}_{p_c}\left(0 \overset{\leq 3k}{\longleftrightarrow} x \text{ in } \mathbb{T}_r^d\right) \leq \tau_{\mathbb{Z},p_c}(0,x) + C_1 V^{-2/3}.$$

Note that the event $\{0 \leftrightarrow x \text{ in } \mathbb{T}_r^d\} \setminus \{0 \overset{\leq 3k}{\longleftrightarrow} x \text{ in } \mathbb{T}_r^d\}$ implies that (a) $\partial B_{\mathbb{T},k}(0) \neq \varnothing$, (b) $\partial B_{\mathbb{T},k}(x) \neq \varnothing$, and (c) $B_{\mathbb{T},k}(0)$ and $B_{\mathbb{T},k}(x)$ do not intersect. Let G' be the subgraph of $\mathbb{T}_r^d$ obtained by removing all edges needed to calculate $B_{\mathbb{T},k}(x)$. Note that the events (a)-(c) imply that $\partial B_{\mathbb{T},k}(x) \neq \varnothing$ and $\partial B_{\mathbb{T},k}^{G'}(0) \neq \varnothing$. By (2.1),

$$\sup_G \mathbb{P}_{p_c}\left(\partial B_{\mathbb{T},k}^G(0) \neq \varnothing\right) \cdot \mathbb{P}_{p_c}\left(\partial B_{\mathbb{T},k}(x) \neq \varnothing\right) \leq (C_2/k)^2 \leq C_3 V^{-2/3}.$$

This completes the proof of (1.11). $\square$

Proof of (1.12). Let $x \in \mathbb{T}_r^d$. Let n be a positive integer smaller than $r/2$. We distinguish two cases: $|x| < 2n/3$ and $|x| \geq 2n/3$. In the first case, we observe that the event $\{0 \leftrightarrow x \text{ by a path which visits } \partial Q_n\}$ implies that there exists a vertex $y \in \partial Q_n$ such that

$$\{0 \leftrightarrow y \text{ in } Q_n\} \circ \{y \leftrightarrow x\}.$$

27

By the BK inequality, Theorem 1.5(a) and (1.11),

$$\mathbb{P}_{p_c}(0 \leftrightarrow x \text{ in } \mathbb{T}_r^d \text{ by a path which visits } \partial Q_n) \tag{8.1}$$
$$\leq \sum_{y \in \partial Q_n} \mathbb{P}_{p_c}(0 \leftrightarrow y \text{ in } Q_n) \cdot \mathbb{P}_{p_c}(y \leftrightarrow x \text{ in } \mathbb{T}_r^d) \leq C_4 \big(\sup_{y \in \partial Q_n} \tau_{\mathbb{Z}, p_c}(y, x) + C_5 V^{-2/3} \big).$$

Since $|x| < 2n/3$, the distance between x and any $y \in \partial Q_n$ is at least $n/3$. Therefore, by (1.1), we have $\tau_{\mathbb{Z}, p_c}(y, x) \leq C_6 n^{2-d}$, and (1.12) follows.

In the case $|x| \geq 2n/3$, we simply use the bound

$$\mathbb{P}_{p_c}(0 \leftrightarrow x \text{ in } \mathbb{T}_r^d \text{ by a path which visits } \partial Q_n) \leq \mathbb{P}_{p_c}(0 \leftrightarrow x \text{ in } \mathbb{T}_r^d) \overset{(1.11)}{\leq} \tau_{\mathbb{Z}, p_c}(0, x) + C_5 V^{-2/3}.$$

Since $|x| \geq 2n/3$, we obtain by (1.1) that $\tau_{\mathbb{Z}, p_c}(0, x) \leq C_7 n^{2-d}$, and (1.12) follows. $\qquad \square$

References

[1] L. Addario-Berry, N. Broutin, and C. Goldschmidt. (2010) Critical random graphs: limiting constructions and distributional properties. *Electron. J. Probab.*, **15(25)** 741-775.

[2] D. Aldous (1997) Brownian excursions, critical random graphs and the multiplicative coalescent. *Ann. Probab.*, **25** 812-854.

[3] B. Bollobás (2001) *Random graphs*, volume **73** of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition.

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

[5] C. Borgs, J.T. Chayes, R. van der Hofstad, G. Slade and J. Spencer (2005) Random subgraphs of finite graphs: II. The lace expansion and the triangle condition. *Annals of Probability* **33** 1886-1944.

[6] C. Borgs, J. Chayes, R. van der Hofstad, G. Slade, and J. Spencer (2006) Random subgraphs of finite graphs. III. The phase transition for the n-cube. *Combinatorica*, **26(4)** 395–410.

[7] J.T. Chayes and L. Chayes (1987) On the upper critical dimension of Bernoulli percolation. *Commun. Math. Phys.* **113(1)** 27-48.

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

[9] G. Grimmett (1999) *Percolation*. Springer-Verlag, Berlin, second edition.

[10] T. Hara (2008) Decay of correlations in nearest-neighbor self-avoiding walk, percolation, lattice trees and animals. *Ann. Probab.* **36(2)** 530-593.

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

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

[13] M. Heydenreich and R. van der Hofstad (2007) Random graph asymptotics on high-dimensional tori. *Comm. Math. Phys.* **270(2)** 335-358.

[14] M. Heydenreich and R. van der Hofstad (2011) Random graph asymptotics on high-dimensional tori II: Volume, diameter and mixing time. *Probab. Theory Related Fields* **149(3-4)** 397-415.

[15] R. van der Hofstad and A. Nachmias (2012) Hypercube percolation. Preprint. Available at `arXiv:1201.3953`.

[16] S. Janson, D.E. Knuth, T. Łuczak, and B. Pittel (1993) The birth of the giant component. *Random Structures Algorithms*, **4(3)** 231-358.

[17] S. Janson, T. Łuczak, and A. Rucinski (2000) *Random graphs.* Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York.

[18] H. Kesten (1980) The critical probability of bond percolation on the square lattice equals 1/2. *Commun. Math. Phys.* **74(1)** 41-59.

[19] G. Kozma and A. Nachmias (2011) Arm exponents in high-dimensional percolation. *Journal of the American Mathematical Society* **24(2)** 375-409.

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

[21] T. Łuczak, B. Pittel, and J. Wierman (1994) The structure of a random graph at the point of the phase transition. *Trans. Amer. Math. Soc.* **341(2)** 721-748.

[22] A. Nachmias and Y. Peres (2008) Critical random graphs: diameter and mixing time. *Ann. Probab.* **36** 1267-1286.