Cluster-size decay in supercritical long-range percolation

Citation for published version (APA):

Document status and date:
Published: 01/03/2023

Please check the document version of this publication:
• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher’s website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.
CLUSTER-SIZE DECAY IN SUPERCRITICAL LONG-RANGE PERCOLATION

JOOST JORRITSMA, JÚLIA KOMJÁTHY, AND DIETER MITSCHE

ABSTRACT. We study the cluster-size distribution of supercritical long-range percolation on $\mathbb{Z}^d$, where two vertices $x, y \in \mathbb{Z}^d$ are connected by an edge with probability $p(\|x-y\|) := p \min\{1, \beta^{\alpha \|x-y\|^{-\alpha d}}\}$ for parameters $p \in (0, 1)$, $\alpha > 1$, and $\beta > 0$. We show that when $\alpha > 1 + 1/d$, and either $\beta$ or $p$ is sufficiently large, the probability that the origin is in a finite cluster of size at least $k$ decays as $\exp\left( -\Theta(k^{(d-1)/d}) \right)$. This corresponds to classical results for nearest-neighbor Bernoulli percolation on $\mathbb{Z}^d$, but is in contrast to long-range percolation with $\alpha < 1 + 1/d$, when the exponent of the stretched exponential decay changes to $2 - \alpha$. This result, together with our accompanying paper establishes the phase diagram of long-range percolation with respect to cluster-size decay. Our proofs rely on combinatorial methods that show that large delocalized components are unlikely to occur. As a side result we determine the asymptotic growth of the second-largest connected component when the graph is restricted to a finite box.

Keywords: Long-range percolation, cluster-size distribution, second-largest component, spatial random graphs.
MSC Class: 60C05, 60K35.

1. Introduction

For nearest-neighbor Bernoulli percolation on $\mathbb{Z}^d$ \cite{8} it is well known \cite{1, 3, 7, 18, 24, 26} that in the supercritical case, the distribution of the number of vertices in the cluster containing the origin follows subexponential decay. Let us write $|C(0)|$ for the number of vertices in the cluster containing the origin, and assume $p > p_c(\mathbb{Z}^d)$. Then it holds that

$$\mathbb{P}(k \leq |C(0)| < \infty) = \exp\left(-\Theta(k^{\frac{1}{d-1}})\right).$$

(1.1)

The decay rate in (1.1) – stretched exponential decay with exponent $(d-1)/d$ – can be intuitively explained as follows: a cluster $C$ with at least $k$ vertices has at least $\Theta(k^{(d-1)/d})$ edges on its (outer) boundary. All these edges need to be absent. In other words, the tail decay in (1.1) is driven by surface tension. Recently, this result was extended to supercritical Bernoulli percolation on classes of transitive graphs \cite{8, 21}. Related works are also \cite{27, 30}, that determine the size of the second-largest component in a finite box for random geometric graphs, also known as continuum percolation, obtaining the same exponent $(d-1)/d$ for the cluster-size decay.

In our accompanying papers \cite{22, 23}, we study the supercritical cluster-size decay in a large class of spatial random graph models where at least one of the degree distribution and the edge-length distribution obey heavy tails: long-range percolation \cite{24, 31}, scale-free percolation on $\mathbb{Z}^d$ and in the continuum \cite{10, 11}; geometric inhomogeneous random graphs \cite{5}, hyperbolic random graphs \cite{24}, the ultra-small scale-free geometric network \cite{32}, the scale-free Gilbert model \cite{19}, the Poisson Boolean model with random radii \cite{13}, the age- and the weight-dependent random connection models \cite{16, 17}.

In these models the tail in (1.1) stays still stretched exponential, but is at least as light as the right-hand-side of (1.1). The new exponent – say $\zeta$ – is at least $(d-1)/d$, with its formula depending on the model parameters. Generally speaking, the accompanying papers \cite{22, 23} treat cluster-sizes whenever the decay is strictly lighter than the right-hand side of (1.1). There, we leave the part of the phase diagram open where the model parameters are such that the conjectured exponent in (1.1) stays $(d-1)/d$ and the tail decay is driven by surface tension as in nearest-neighbor percolation.

In particular, the paper \cite{23} leaves open this region for long-range percolation (LRP) \cite{2, 31} (see Theorem 1.3 below for the result). This missing region for LRP is the main focus in this paper.

Definition 1.1 (Long-range percolation (LRP)). Fix constants $d \in \mathbb{N}, \alpha > 1$, $p \in (0, 1)$, and $\beta > 0$. We consider the random graph $\mathcal{G}_\infty = (V(\mathcal{G}_\infty), E(\mathcal{G}_\infty))$ with $V(\mathcal{G}_\infty) = \mathbb{Z}^d$ such that each edge $\{x, y\}$ is included in $E(\mathcal{G}_\infty)$, independently of all the other edges, with probability

$$p(\|x-y\|) := p \cdot \left( \min\{1, \frac{\beta}{\|x-y\|^\alpha}\} \right)^\alpha,$$

(1.2)
Theorem 1.2 (Second-largest component and cluster-size decay). Consider supercritical long-range percolation on $\mathbb{Z}^d$ for $d \geq 2$ and $\alpha > 1 + 1/d$. If $\beta$ in (1.2) is sufficiently large (depending on $p, \alpha, d$), or if $\beta \geq 1$ and $p$ is sufficiently close to 1, then there exist constants $A, \delta > 0$ such that for all $n$ sufficiently large,

$$P\left(\frac{1}{n}(\log(n))^{d/(d-1)} \leq |C_n^{(2)}| \leq A(\log(n))^{d/(d-1)}\right) \geq 1 - n^{-\delta}. \tag{1.3}$$

Under the same assumptions, for all $k$ sufficiently large, whenever $n(\log(n))^{2d/(d-1)} \geq k$ or $n = \infty$,

$$-k^{-(d-1)/d} \log(\mathbb{P}(|C_n(0)| \geq k, 0 \notin C_n^{(1)})) \in [1/A, A]. \tag{1.4}$$

Lastly, under the same assumptions,

$$\frac{|C_n^{(1)}|}{n} \xrightarrow{p} \mathbb{P}(|C(0)| = \infty), \quad \text{as } n \to \infty. \tag{1.5}$$

Note that (1.4) allows for $n = \infty$. Our results can be extended to hold when $\beta < 1$ and $p\beta^\alpha$ are sufficiently close to one. In Remark 1.4 we discuss a further generalization to more general connectivity functions $p(||x - y||)$. Theorem 1.2 complements the result of [23] applied to long-range percolation that we state here for completeness.

Theorem 1.3 (Complementary result for $\alpha < 1 + 1/d$ [23]). Consider supercritical long-range percolation on $\mathbb{Z}^d$ for $d \geq 1$ and $\alpha < 1 + 1/d$. There exists constants $A, \delta > 0$ such that for all $\varepsilon > 0$ and $n$ sufficiently large,

$$P\left(\frac{1}{A}(\log(n))^{1/(2-\alpha)} \leq |C_n^{(2)}| \leq A(\log(n))^{1/(2-\alpha-\varepsilon)}\right) \geq 1 - n^{-\delta}. \tag{1.6}$$

Moreover, for all $k$ sufficiently large, whenever $n \in [Ak, \infty],$

$$-k^{-(2-\alpha)/d} \log(\mathbb{P}(|C_n(0)| \geq k, 0 \notin C_n^{(1)})) \in [1/A, Ak^{\varepsilon}]. \tag{1.7}$$

Lastly, under the same assumptions,

$$\frac{|C_n^{(1)}|}{n} \xrightarrow{p} \mathbb{P}(|C(0)| = \infty), \quad \text{as } n \to \infty.$$

In Theorem 1.3 we do not require $\beta$ or $p$ sufficiently large, and also allow one-dimensional models: when $d = 1$, LRP is supercritical when $\alpha \leq 2 = 1 + 1/d$ and when $p, \beta$ are sufficiently large $[13, 31]$. When $d = 1$ and $\alpha > 2$, LRP is subcritical for any $p, \beta > 0$ such that $p(1) < 1$ $[31]$, so Theorems 1.2 and 1.3 together give a complete picture for the cluster-size decay for supercritical long-range percolation (under the additional assumption that $\beta$ or $p$ is sufficiently large when $\alpha > 1 + 1/d$). In [23], we also study the phase boundary $\alpha = (d - 1)/d$. In that case the lower bounds (1.6) and (1.7) contain lower order correction factors, that we conjecture to be sharp. We omit further details here. To the extent of our knowledge, for LRP, the only related results regarding the distribution of smaller clusters in supercritical LRP is an upper bound on the second-largest component with unidentified exponent by Crawford and Sly $[9]$ for $\alpha \in (1, 2)$ in dimension 1 and $\alpha \in (1, 1 + 2/d)$ in dimensions 2 and higher. For subcritical LRP with $\alpha \in (1, 2)$, a polynomial upper bound on $\mathbb{P}(|C(0)| \geq n)$ is established in [20].

Before proceeding to the technical contributions, we remark that our results could be generalized to a more general class of random graph models on $\mathbb{Z}^d$. Theorem 1.2 extends to random graph models on $\mathbb{Z}^d$ with independent edges for any connectivity function that has a lighter tail than $p$ in Theorem 1.2 provided that the probability of ‘short-range’ edges is still sufficiently large. In particular, it extends to spread-out percolation, in which two vertices within distance $R$ are connected independently with probability $p$, or long-range percolation models in which the connection probability
decays subpolynomially. We refrain from proving the result in this generality, since it would require many technically involved changes in our already technical companion paper \cite{22}.

**Remark 1.4.** Consider the percolation model on $\mathbb{Z}^d$ where each pair of vertices $x, y \in \mathbb{Z}^d$ is connected by an edge with probability $p(|x - y|)$ for some function $p : [0, \infty) \to [0, 1)$, independently of other vertex pairs. Let $J : [0, \infty) \to [0, 1)$ be a function that satisfies $\sup_{r > 0} J(r) < 1$, and

$$\int_{x : x \in \mathbb{R}^d} \|x\| J(\|x\|) < \infty. \tag{1.8}$$

Then we have the following two cases:

1) If the connectivity function $p$ is of the form

$$p(|x|) = J(|x|/\beta),$$

and there is an $\varepsilon > 0$ such that $J(x) > \varepsilon$ whenever $x < \varepsilon$, then \((1.3)\)–\((1.5)\) hold for all sufficiently large $\beta$ depending on $\varepsilon$.

2) If the connectivity function is of the form

$$p(|x|) = \begin{cases} p & \text{if } |x| = 1, \\ J(|x|) & \text{if } |x| > 1, \end{cases}$$

then \((1.3)\)–\((1.5)\) hold for all $p$ sufficiently close to 1.

The integral in the condition \((1.8)\) represents the order of the expected number of edges $\{x, y\}$ for which the line-segment $(x, y)$ crosses a fixed box of volume one. If this number is finite, Theorem \ref{thm:main} holds in more generality. The connectivity function $p$ from Definition \ref{def:connectivity} satisfies the integrability condition \((1.8)\) if and only if $\alpha > 1 + 1/d$. We conjecture that the upper bounds in \((1.3)\)–\((1.4)\) remain valid if \((1.8)\) is violated (but no longer match the lower bounds). However, this would require a different proof technique.

We state a proposition that contains the main technical contribution of this paper. Together with statements from our companion paper \cite{22}, where we establish the relation between the second-largest component and the cluster-size decay for spatial random graph models more generally, this proposition will readily imply Theorem \ref{thm:main}.

**Proposition 1.5** (Second-largest component, upper bound). Consider supercritical long-range percolation on $\mathbb{Z}^d$ for $\alpha > 1 + 1/d$, $d \geq 2$. If $\beta$ in \((1.2)\) is sufficiently large (depending on $p, \alpha, d$), or if $\beta \geq 1$ and $p$ is sufficiently close to 1, then there exists a constant $A > 0$ such that for all $k$ sufficiently large and for all $n$ satisfying $n(\log(n))^{-2d/(d-1)} \geq k$:

$$\mathbb{P}(|C_n^{|\infty}| \geq k) \leq n \log(n) \exp(-k^{(d-1)/d}).$$

We believe that Proposition \ref{prop:main} and hence Theorem \ref{thm:main} should hold for any values of $\beta, p$ that lead to a supercritical graph: however, this would require non-trivial adaptations of our proof techniques. This proposition also can be extended to $\beta < 1$ such that $p\beta^\alpha$ is sufficiently close to 1.

1.1. Idea of proof. The proof of Proposition \ref{prop:main} relies on a careful first-moment analysis in which we count all possible candidates of isolated components of size at least $k$. The starting point is the classic isoperimetric inequality that says that any set $\mathcal{S}$ of at least $k$ vertices has an edge-boundary of size $|\partial \mathcal{S}| = \Omega(k^{(d-1)/d})$. These edges need to be absent when $\mathcal{S}$ is a connected component (or simply component below), i.e., detached from the rest of the graph. The combinatorial difficulty arises when we account for all possible candidate components $\mathcal{S}$: the structure of $\mathcal{S}$ is more complex than for nearest-neighbor bond percolation in $\mathbb{Z}^d$, since $\mathcal{S}$ can be “delocalized” in space. The second difficulty arises in the finite box $\Lambda_n \subseteq \mathbb{Z}^d$, where we need to take boundary effects into account caused by possibly shared boundary of $\partial \mathcal{S}$ and $\partial \Lambda_n$.

To resolve these two complications, we distinguish two types of components: the first type consists of several “blocks” connected by long edges: each block is a connected subset of $\mathbb{Z}^d$ (with respect to nearest-neighbor relation in $\mathbb{Z}^d$). We consider each possible combination of blocks with fixed total outer edge-boundary size $m$, and give an upper bound on the probability that these blocks form a connected component by counting all possible spanning trees on these blocks. We show that the combinatorial factor arising from counting all potential components with boundary $m$ is at most exponential in $m$. We then use the large value of $\beta$ or $p$ in our favor to prove that the probability that such a component is formed and isolated is sufficiently small.
The second type of potential component $S$ contains a large block that has a large overlap with the boundary of $\Lambda_n$, and consequently $|\partial S|$ inside $\Lambda_n$ may be small. In this case we enumerate all such blocks and use an adapted isoperimetric inequality to still ensure that many edges need to be absent, thereby showing the right error probability.

**Organization.** In Section 2 we derive an intermediate upper bound for $\mathbb{P}(|C_n^{(2)}| \geq k)$, defining the two types of components formally. Then, we state two lemmata and show that they imply Proposition 1.5. We prove the two lemmata in separate sections. In the last section we use the result of Proposition 1.5 to prove Theorem 1.2.

**Notation.** Let $H = (V_H, E_H)$ be a graph. For two sets $A, B \subseteq V_H$, we write $A \leftrightarrow_H B$ if there exist $x \in A, y \in B$ such that $\{x, y\} \in E_H$, and $A \not\leftrightarrow_H B$ if no such pair exists. We leave out the subscript $H$ if the graph is clear from the context. For $A \subseteq V_H$, we write $H[A]$ for the induced subgraph of $H$ on vertices in $A$. Denote by $\mathbb{Z}^d_\infty$ the graph on the vertex set $\mathbb{Z}^d$ and an edge between $x, y \in \mathbb{Z}^d$ if and only if $\|x - y\|_\infty = 1$. Similarly, let $\mathbb{Z}^d_1$ be the graph on the vertex set $\mathbb{Z}^d$ and an edge between $x, y \in \mathbb{Z}^d$ if and only if $\|x - y\|_1 = 1$. As already mentioned, we write $\|\cdot\| := \|\cdot\|_2$. For two sets $A, B \subseteq \mathbb{Z}^d$, denote by $\|A - B\|_p = \min\{\|x - y\|_p \mid x \in A, y \in B\}$. We say that a path $\pi = (v_1, v_2, v_3, \ldots)$ is self-avoiding if its vertices are all distinct.

## 2. Preliminaries and setup

Throughout the rest of the paper, we assume that $d \geq 2$, $\alpha > 1 + 1/d$, and $n^{1/d} \in \mathbb{N}$. We will now formalize the concepts from the proof outline in Section 1.1 that eventually lead to two lemmas for each of the two described types of components. To ensure that the upcoming definitions naturally follow each other, we will postpone the (sometimes standard) proofs of intermediate claims to the appendix.

We start with a definition to describe sets of $\Lambda_n$ that form (subsets of) the second-largest component.

**Definition 2.1** (Connected sets and blocks). We call a set $A \subseteq \mathbb{Z}^d$ of vertices 1-connected or a block, if the graph $\mathbb{Z}^d[A]$ consists of a single connected component. We similarly define $A$ being $*$-connected if the graph $\mathbb{Z}^d_\infty[A]$ consists of a single connected component. We write

$$\mathcal{A} := \{A \subseteq \Lambda_n \mid A \text{ is 1-connected}\}, \quad \mathcal{A}_* := \{A \subseteq \Lambda_n \mid A \text{ is } *\text{-connected}\}. \quad (2.1)$$

We say that a sequence of sets $A_1, A_2, \ldots \subseteq \mathbb{Z}^d$ is 1-disconnected if $\|A_i - A_j\|_1 > 1$ for all $i \neq j$. We say that a set $A \subseteq \Lambda_n$ consists of blocks $A_1, \ldots, A_b$ if $A_i$ is a blocks for $1 \leq i \leq b$, if the sequence $(A_i)_{i \leq b}$ is 1-disconnected, and their union equals $A$.

We say that a vertex $x$ is surrounded by $A \in \mathcal{A}$ if each infinite 1-connected self-avoiding path starting from $x$ contains a vertex of $A$. We define for $A \in \mathcal{A}$ its closure $\bar{A}$ as

$$\bar{A} = A \cup \{x \in \mathbb{Z}^d : x \text{ surrounded by } A\}. \quad (2.2)$$

We call the maximal 1-connected subsets of $\bar{A} \setminus A$ the holes of $A$, and write $\mathcal{H}_A$ for the collection of holes.

We make a few comments. The closures of blocks will be used for the first type of components described in Section 1.1. Take now a block $A \subseteq \Lambda_n$. Then $x$ can only be surrounded by $A$ if $x \in \Lambda_n$ too, hence $A \subseteq \Lambda_n$ implies that $\bar{A} \subseteq \Lambda_n$.

Due to the presence of long-range edges, a component in long-range percolation may consist of multiple 1-disconnected blocks (some of them possibly consisting of a single vertex). We define the notion of a block graph.

**Definition 2.2** (Block graph). Let $A_1, \ldots, A_b \in \mathcal{A}$ be a sequence of blocks, and consider a graph $G$ on vertices $V_G \supseteq \bigcup_{i \leq b} A_i$. The block graph $\mathcal{G}(A_1)_{i \leq b}$ of $G$ on blocks $A_1, \ldots, A_b$ is defined as

$$V_{\mathcal{G}} := \{1, \ldots, b\}, \quad E_{\mathcal{G}} := \{\{i, j\} : A_i \leftrightarrow G A_j\}.$$  

In words, the vertices of each block are contracted to a single vertex in the block-graph, and two corresponding vertices for the blocks $i$ and $j$ are connected in the block graph if and only if there exists an edge in the original graph $G$ between (some vertices in) the two blocks. We continue with a simple claims, with proof in the appendix on page 23.
Claim 2.3 (Unique block-decomposition of components). Let $C$ be a finite component of a graph $G$ with vertex set either $\Lambda_{n}$ or $\mathbb{Z}^{d}$. Then $C$ can be uniquely decomposed into a 1-disconnected sequence $(A_{i})_{i \leq b}$ of blocks, with $b < \infty$, so that the block graph $\mathcal{H}_{G}((A_{i})_{i \leq b})$ is connected.

Later, we will enumerate subsets $S \subseteq \Lambda_{n}$ of vertices that potentially form a component of LRP in $\Lambda_{n}$. To ensure that a subset is isolated from the rest of the graph, there must be no edges from $S$ to its "surrounding" inside $\Lambda_{n}$. This motivates the following definition of boundaries with respect to $\Lambda_{n}$.

Definition 2.4 (Boundaries). Let $A \subseteq \Lambda_{n}$. We define the exterior boundary of $A$ with respect to $\Lambda_{n}$ and $\mathbb{Z}^{d}$, respectively, as

$$\partial_{\text{ext}} A := \{ x \in \Lambda_{n} : \| x - A \|_{1} = 1 \}, \quad \tilde{\partial}_{\text{ext}} A := \{ x \in \mathbb{Z}^{d} : \| x - A \|_{1} = 1 \}. \quad (2.3)$$

We define the interior boundary of $A$ with respect to $\Lambda_{n}$ and $\mathbb{Z}^{d}$, respectively, as

$$\partial_{\text{int}} A := \{ x \in A : \| x - \partial_{\text{ext}} A \|_{1} = 1 \}, \quad \tilde{\partial}_{\text{int}} A := \{ x \in A : \| x - \tilde{\partial}_{\text{ext}} A \|_{1} = 1 \}. \quad (2.4)$$

If, in words, we mention the exterior, interior, or outer boundary of $A$ then – unless explicitly specified differently – we mean with respect to $\Lambda_{n}$.

We mention that $\tilde{\partial}_{\text{int}} A_{n}$ is the ‘usual’ vertex boundary of $\Lambda_{n}$. The boundary $\tilde{\partial}_{\text{ext}} A$ may contain vertices outside $\Lambda_{n}$, and will be useful in the enumeration of subsets forming isolated components below. It may happen that a block $A$ contains (many) vertices of $\tilde{\partial}_{\text{int}} A_{n}$. On such regions, $A$ may not have external boundary vertices, implying that there $\partial_{\text{int}} A$ is also empty. The next claim contains basic properties of blocks, their closures, and their boundaries, with proof in the Appendix on page 23.

Claim 2.5 (Blocks, their closures and their boundaries). The following four statements hold:

(i) For any block $B$, $\partial_{\text{int}} \tilde{B} \subseteq \tilde{\partial}_{\text{int}} B$.
(ii) Let $B_{1}, B_{2}$ be 1-disconnected blocks such that $B_{1} \cap B_{2} \neq \emptyset$. Then either $B_{1} \subseteq B_{2}$ or $B_{2} \subseteq B_{1}$.
(iii) Let $B_{1}, B_{2}$ be 1-disconnected blocks such that $B_{1} \cap B_{2} = \emptyset$. Then $B_{1}, B_{2}$ are also 1-disconnected from each other.
(iv) For any block $B$, $\partial_{\text{int}} \tilde{B}$ and $\tilde{\partial}_{\text{ext}} B$ are *-connected.
(v) For any hole $H$ of a block $B$, we have that $H = \bar{H}$, so $\partial_{\text{int}} H$ and $\tilde{\partial}_{\text{ext}} H$ are *-connected.

We point out that the fourth statement of the preceding claim is [12, Lemma 2.1]. The next claim shows that the sizes of the boundaries with respect to $\Lambda_{n}$ and $\mathbb{Z}^{d}$ are of the same order, provided that the set is smaller than $3n/4$. Moreover, it contains an isoperimetric inequality that we extensively use below. The proof is given in the appendix on page 24.

Claim 2.6 (Boundary bounds and isoperimetry). There exists $\delta > 0$ such that for all $A \subseteq \Lambda_{n}$ with $|A| \leq 3n/4$ or $A \cap \tilde{\partial}_{\text{int}} A_{n} = \emptyset$,

$$|\partial_{\text{int}} A| \geq \delta |\tilde{\partial}_{\text{int}} A| \geq \delta |A|^{(d-1)/d}, \quad |\partial_{\text{ext}} A| \geq \delta |\tilde{\partial}_{\text{ext}} A| \geq \delta |A|^{(d-1)/d}. \quad (2.5)$$

The inequalities with ($\ast$) hold for any $A \subseteq \Lambda_{n}$ without conditions on $A$.

The next lemma is due to Peierls (its proof is given in the appendix on page 26) and is crucial for the enumeration of blocks that satisfy $A = \tilde{A}$.

Lemma 2.7 (Peierls’ argument). There exists a constant $c_{\text{pel}} > 0$ such that for all $x \in \mathbb{Z}^{d}$ and $m \in \mathbb{N}$,

$$|\{ A \in \mathcal{A} : A \ni x, A = \tilde{A}, |\tilde{\partial}_{\text{int}} A| = m \}| \leq \exp(c_{\text{pel}} m). \quad (2.6)$$

We remark that the proof relies on the fact that $\tilde{\partial}_{\text{int}} A$ is *-connected [12, Lemma 2.1] (which would not hold if $A$ contained holes, or may not hold if one replaces $\tilde{\partial}_{\text{int}} A$ by $\partial_{\text{int}} A$).

In (2.6), we would like to replace $\partial_{\text{int}} A$ by $\partial_{\text{int}} \tilde{A}$ (enumeration of sets that are equal to their closure would allow us to use Peierls’ argument). However, then the isoperimetric inequality (2.5) may not hold anymore if the total boundary size of the holes in $A$ is too large compared to $\partial_{\text{int}} \tilde{A}$, which could happen if $|A| > 3n/4$. We define two types of blocks, based on the size of the closures of the blocks (that is, whether Claim 2.5 applies to $\tilde{A}$ or not), i.e.,

$$\mathcal{A}_{\text{small}} := \{ A \in \mathcal{A} : |\tilde{A}| \leq 3n/4, \text{ and } \tilde{A} = A \}, \quad \mathcal{A}_{\text{large}} := \{ A \in \mathcal{A} : |\tilde{A}| > 3n/4, \text{ and } |A| \leq n/2 \}. \quad (2.7)$$
For each of the sets we define an event, i.e.,

\[ \mathcal{E}_1(b) := \left\{ \exists \text{ 1-disconnected } (A_i)_{i \leq b} \in \mathcal{A}_{\text{small}} : \left( \cup_i \tilde{\partial}_{\text{int}} A_i \right) \not\ni g_n \left( \Lambda_n \cup \cup_i A_i \right), \left| \cup_i A_i \right| \geq k \right\}, \quad \text{(2.8)} \]

\[ \mathcal{E}_2 := \{ \exists A \in \mathcal{A}_{\text{large}} : A \not\ni g_n \partial_{\text{ext}} A \}. \quad \text{(2.9)} \]

The following deterministic claim holds for any graph on vertices in \( \Lambda_n \). It shows that the union of these events contains the event \( \{|C_n^{(2)}| \geq k\} \). In particular, the proof reveals why we could restrict to sets with \( A = A \) in the definition of \( \mathcal{A}_{\text{small}} \) in (2.7).

**Claim 2.8.** Consider the graph \( G_n \) from Definition 2.4 with \( C_n^{(2)} \) the second largest component of \( G_n \). Then

\[ \{ \{|C_n^{(2)}| \geq k\} \subseteq \mathcal{E}_2 \cup \left( \bigcup_{b=1}^{n/2} \mathcal{E}_1(b) \right). \]

**Proof of Claim 2.8.** Clearly \( \{|C_n^{(2)}| \geq k\} \subseteq \{ \exists \text{ a component } C \in G_n : |C| \in [k, |k|/2] \} \). The size restriction \( n/2 \) is needed since otherwise \( C_n^{(2)} \) would be the largest component. Take any such \( C \). We use Claim 2.8 to first uniquely decompose \( C \) into 1-disconnected (hence disjoint) blocks \( A_1, \ldots, A_b \) for some \( b \geq 1 \). Since \( k \leq |C| \leq n/2 \), \( |A_i| \leq n/2 \) also holds for all \( i \leq b \), and also \( b \leq n/2 \), and \( \sum_{i \leq b} |A_i| \geq k \).

We distinguish two cases. Either (1) there is at least one block that is in \( \mathcal{A}_{\text{large}} \) or (2) all the blocks are in \( \mathcal{A}_{\text{small}} \). In the first case the event \( \mathcal{E}_2 \) in (2.7) holds: the block of \( C \) satisfying \( \mathcal{A}_{\text{large}} \) (say \( A_1 \)) is per definition 1-disconnected from the other blocks of \( C \), and since \( C \) is a component of \( G_n \), \( A_1 \) is \( G_n \)-disconnected from its exterior boundary, hence \( \mathcal{E}_2 \) holds for \( A := A_1 \) in (2.7).

In case (2) all the blocks \( (A_i)_{i \leq b} \) are in \( \mathcal{A}_{\text{small}} \), and, since they form the component \( C \), the graph \( G_n \) spanned on \( \cup_i A_i \) is connected, while their union is disconnected in \( G_n \) from the rest of the graph. Finally their disjointness and \( |C| \in [k, n/2] \) ensures that \( \{|A_i| \leq n/2 \} \) also holds for all \( i \leq b \), and \( \sum_{i \leq b} |A_i| \geq k \).

Formally we describe this event as (changing the sets to be denoted by \( B_i \) to avoid clash of notation later):

\[ \tilde{\mathcal{E}}_1(b) := \left\{ \exists \text{ 1-disconnected } (B_i)_{i \leq b} \in \mathcal{A}_{\text{large}} : \left. \mathcal{G}_n|_{\cup_i B_i} \right|_{\cup_i B_i} \not\ni g_n \left( \Lambda_n \cup \cup_i B_i \right), \left| \cup_i B_i \right| \leq n/2 \right\}. \]

Taking a union over the number of blocks and combining the two cases, we arrive at

\[ \{ \{|C_n^{(2)}| \geq k\} \subseteq \{ \exists \text{ a component } C \in G_n : |C| \in [k, n/2] \} \subseteq \mathcal{E}_2 \cup \left( \bigcup_{b=1}^{n/2} \tilde{\mathcal{E}}_1(b) \right). \]

We will show that

\[ \left( \bigcup_{b=1}^{n/2} \mathcal{E}_1(b) \right) \subseteq \left( \bigcup_{b=1}^{n/2} \tilde{\mathcal{E}}_1(b) \right). \quad \text{(2.10)} \]

Take now any 1-connected blocks \( (B_1, \ldots, B_b) \) for which \( \tilde{\mathcal{E}}_1(b) \) holds. The conditions of Claim 2.8 are satisfied for any pair \( B_i, B_j \) with \( i \neq j \), hence for each pair, \( B_i \) and \( B_j \) are either 1-connected disjoint sets, or one contains fully the other one. Choose now those sets in \( \{B_1, \ldots, B_b\} \) that are not contained in any other set in the same list. We then obtain an integer \( b' \leq b \) and a 1-connected subset \( \{B_{i_1}, \ldots, B_{i_{b'}}\} \subseteq \{B_1, \ldots, B_b\} \) such that

\[ \left( \bigcup_{i=1}^{b} B_i \right) \subseteq \left( \bigcup_{i=1}^{b'} B_{i_j} \right) = \bigcup_{j=1}^{b'} B_{i_j}. \quad \text{(2.11)} \]

Let \( B_1, \ldots, B_b \) in \( \mathcal{A}_{\text{large}} \) be an arbitrary sequence of 1-connected blocks satisfying \( \tilde{\mathcal{E}}_1(b) \), and assume without loss of generality that the indices \( i_1, \ldots, i_{b'} \) correspond to indices \( 1, \ldots, b' \), and we may thus assume that the sets \( B_1, \ldots, B_{b'} \) satisfy (2.11). Since \( (B_1, \ldots, B_{b'}) \) is a 1-connected sequence of blocks that are equal to their own closure, (2.10) follows if

\[ \{G_n|_{\cup_i B_i} \not\ni g_n \left( \Lambda_n \cup \cup_i B_i \right) \} \subseteq \{H_{\text{ext}} \cup \cup_i B_i \} \not\ni g_n \left( \Lambda_n \cup \cup_i B_i \right), \quad \text{(2.12)} \]

\[ \{ |\cup_i B_i| \leq n/2 \} \subseteq \{ |\cup_i B_i| \geq k \}, \quad \text{(2.13)} \]

\[ \{H_{\text{ext}} \cup \cup_i B_i \not\ni g_n \left( \Lambda_n \cup \cup_i B_i \right) \} \subseteq \{H_{\text{ext}} \cup \cup_i B_i \not\ni g_n \left( \Lambda_n \cup \cup_i B_i \right) \}, \quad \text{(2.14)} \]

since then all conditions for \( \mathcal{E}_1(b') \) defined in (2.8) are satisfied by setting \( A_i = B_i \) for all \( i \leq b' \).
For the first inclusion (2.12) we observe that
\[ \{G_n \cup_{i \leq b} B_i \text{ connected} \} \subseteq \{\mathcal{H}_{G_n}((B_i)_{i \leq b}) \text{ connected} \} \subseteq \{\mathcal{H}_{G_n}((\tilde{B}_i)_{i \leq b}) \text{ connected} \}, \]

since the left-hand side was assumed in (2.1), the block graph being connected is a less demanding event than the actual spanned graph being connected, and the second containment follows since each set of edges in \( G_n \) that ensures that \( \mathcal{H}_{G_n}((B_i)_{i \leq b}) \) is connected, also ensures that the block graph on the closures of \((\tilde{B}_i)_{i \leq b}\) is connected.

The second inclusion (2.13) is trivial since \( \cup_{i \leq b} B_i \subseteq \cup_{i \leq b} \tilde{B}_i \) by (2.11). For the third inclusion (2.14) we have to argue that the set of edges that is excluded on the right-hand side is smaller than the set of excluded edges on the left-hand side. Clearly \((A_n \setminus \cup_{i \leq b} B_i) \supseteq (A_n \setminus \cup_{i \leq b} \tilde{B}_i)\), and by part (i) in Claim 2.3 it follows that \( \partial_{\text{int}} \tilde{B}_i \subseteq B_i \).

We state two lemmas that together with Claim 2.8 prove Proposition 1.5.

Lemma 2.9 (Unlike block graphs). Let \( \mathcal{G}_n \) be long-range percolation on \( \Lambda_n \) as in Definition 1.1 with \( d \geq 2 \), \( \alpha > 1 + 1/d \). There exists a constant \( \beta_* > 0 \), such that for all \( p \in (0,1) \), there exists \( \beta_\ast = \beta_\ast (p,d,\alpha) > 0 \) such that for all \( \beta \geq \beta_\ast \), and \( n \) sufficiently large,
\[ \Pr \left( \bigcup_{b = 1}^{n/2} \mathcal{E}_1(b) \right) \leq n \log(n) \exp \left( - \frac{2.10}{2.10} \log \left( \frac{1}{1-p} \right) \beta_\ast^{1/d} k^{(d-1)/2d} \right). \] (2.15)

Moreover, there exists \( p_d < 1 \) such that \( \beta_\ast \leq 1 \) for all \( p \in (p_d,1) \).

Lemma 2.10 (No large isolated component). Let \( \mathcal{G}_n \) be long-range percolation on \( \Lambda_n \) as in Definition 1.1 with \( d \geq 2 \), \( \alpha > 1 + 1/d \). There exists a constant \( \beta_* > 0 \) such that for all \( p \in (0,1) \), there exists \( \beta_\ast = \beta_\ast (p,d,\alpha) > 0 \) so that for all \( \beta \geq \beta_\ast \), and \( n \) sufficiently large
\[ \Pr (\mathcal{E}_2) \leq \exp \left( - \frac{2.10}{2.10} \log \left( \frac{1}{1-p} \right) \beta_\ast^{(d-1)/2d} \frac{C_{n/2}}{\log^2(n)} \right). \] (2.16)

Moreover, there exists \( p_d < 1 \) such that \( \beta_\ast \leq 1 \) for all \( p \in (p_d,1) \).

We prove the two lemmata in the following sections.

3. Spanning Trees on Block Graphs

We work towards proving Lemma 2.9. Recall the event \( \mathcal{E}_1(b) \) from (2.8), and \( A_{\text{small}} \) for the blocks of \( \Lambda_n \) without holes and closure of size at most \( 3n/4 \) from (2.7) on which there should be a connected block graph \( \mathcal{H}_{G_n}((A_i)_{i \leq b}) \) (see Definition 2.2). By a union bound over all possible 1-disconnected sequences of blocks \((A_1,\ldots,A_b)\subseteq A_{\text{small}}\) whose total size is at least \( k \), we obtain
\[ \Pr (\mathcal{E}_1(b)) \leq \frac{1}{b!} \sum_{A_1} \cdots \sum_{A_b} \Pr (\mathcal{H}_{G_n}((A_i)_{i \leq b}) \text{ is connected}, (\cup_{i \leq b} \partial_{\text{int}} A_i) \not\leftrightarrow \mathcal{G}_n (\Lambda_n \cup_{i \leq b} A_i)), \]

where the factor 1/\( b! \) corrects for the permutations of \((A_1,\ldots,A_b)\) yielding the same blocks, but ordered differently. Using the independence of edges in long-range percolation in Definition 1.1 we obtain
\[ \Pr (\mathcal{E}_1(b)) \leq \frac{1}{b!} \sum_{A_1} \cdots \sum_{A_b} \Pr (\mathcal{H}_{G_n}((A_i)_{i \leq b}) \text{ is connected}) \cdot \Pr ((\cup_{i \leq b} \partial_{\text{int}} A_i) \not\leftrightarrow \mathcal{G}_n (\Lambda_n \cup_{i \leq b} A_i)) \cdot \Pr (\mathcal{E}_2) . \] (3.1)

The block graph \( \mathcal{H}_{G_n}((A_i)_{i \leq b}) \) can only be connected if it contains a spanning tree on its blocks. To count these spanning trees, we introduce the rooted labeled \( f \)-tree. In the following definition, we use that each tree on \( b \) vertices has \( b-1 \) edges.

Definition 3.1 (f-tree). Let \( \mathcal{F}_b \) be the set of vectors \( f = (f_1,\ldots,f_b) \in \mathbb{N}_0^b \) satisfying \( \sum_{i \in [b]} f_i = b - 1 \), \( f_1 + \ldots + f_j \geq j \) for all \( j \in [b-1] \). We call \( f \) the vector of forward degrees. A rooted labeled tree on \( b \) vertices is an \( f \)-tree if the root has label 1, and it has an outgoing edge to each of the vertices with labels 2, \ldots, \( f_1 + 1 \), vertex 2 has an outgoing edge to each of the vertices with labels \( f_1 + 2,\ldots,f_1 + f_2 + 1 \), and so on, the vertex with label \( j \) has an outgoing edge to each of the vertices with labels \( 2 + \sum_{i=1}^{j-1} f_i,\ldots,1 + \sum_{i=1}^{j-1} f_i \). If \((i,j)\) is a directed edge in an \( f \)-tree, then we say that \( i \) is the parent of \( j \) and \( j \) is the child of \( i \). We say that the labeled block graph \( \mathcal{H}_{G_n}((A_i)_{i \leq b}) \) is \( f \)-connected, if it contains an \( f \)-tree on its vertices \((1,2,\ldots,b)\).
Given a forward-degree vector \( \mathbf{f} \) and a labeled set of vertices, the \( \mathbf{f} \)-tree is uniquely determined. The construction ensures that the block with label \( b \) must be a leaf, i.e., it has forward degree \( f_b = 0 \), and its parent corresponds to the label of the last nonzero entry of \( \mathbf{f} \). Further, given a tree \( T \) with labeling \( \mathbf{f} \in \mathcal{F}_b \), upon removing the leaf with label \( b \), we obtain a tree \( T \setminus \{ b \} \) on \( \{ 1, \ldots, b-1 \} \) with a labeling in \( \mathcal{F}_{b-1} \).

Further, an \( \mathbf{f} \)-tree always has vertex 1 as its root, and the forward neighbors of any vertex have consecutive labels. Hence, not all the \( b! \) labelings of a tree \( T \) are valid labelings, i.e., no vector \( \mathbf{f} \in \mathcal{F}_b \) can be associated to some labelings. However, for a fixed tree \( T \) on a connected block graph \( \mathcal{H}_n^b((A_i)_{i \leq b}) \), there is at least one permutation \( \sigma \) of \( (1, 2, \ldots, b) \) with \( \sigma(1) = 1 \) and a vector \( \mathbf{f} \in \mathcal{F}_b \) such that \( \mathcal{H}_n^b((A_{\sigma(i)})_{i \leq b}) \) is \( \mathbf{f} \)-connected. In other words, we can relabel the blocks so that the new labeling \( (1, \sigma(2), \ldots, \sigma(b)) \) is a proper labeling of \( T \), for some \( \mathbf{f} \in \mathcal{F}_b \) in Definition 3.1. We denote the set of permutations of \( (1, 2, \ldots, b) \) with 1 a fixed point by \( \mathcal{S}_b^1 \). Note that the choice of the spanning tree \( T \) may not be unique if \( \mathcal{H}_n^b((A_i)_i) \) is connected. We obtain on the first factor inside the sum in (3.1) that

\[
P(\mathcal{H}_n^b((A_i)_{i \leq b}) \text{ connected}) = P\left( \bigcup_{\mathbf{f} \in \mathcal{F}_b} \bigcup_{\sigma \in \mathcal{S}_b^1} \{ \mathcal{H}_n^b((A_{\sigma(i)})_{i \leq b}) \text{ is } \mathbf{f} \text{-connected} \} \right).
\]

If, for a given \((\mathbf{f}, \sigma)\) the block graph is \( \mathcal{H}_n^b((A_{\sigma(i)})_{i \leq b}) \) is \( \mathbf{f} \)-connected, then there are \( \prod_i f_i! \) other pairs \((\mathbf{f}', \sigma')\) such that \( \mathcal{H}_n^b((A_{\sigma'(i)})_{i \leq b}) \) is \( \mathbf{f}' \)-connected, counting the isomorphisms of rooted trees: namely, the (consecutive) labels of the forward neighbors of any vertex \( v \) may be permuted (yielding the factor \( f_v! \) for each vertex), resulting in permuting the labels in the forward-subtrees of \( v \) accordingly. For any such \((\mathbf{f}, \sigma)\) and \((\mathbf{f}', \sigma')\), we then also have that \( \prod_{i=1}^b f_i! = \prod_{i=1}^b f_i!' \). Hence, in the above union each rooted tree \( T \) (with root fixed) on \( \mathcal{H}_n^b((A_{\sigma'(i)})_{i \leq b}) \) is counted \( \prod_{i=1}^b f_i! \) times. Thus, we obtain from (3.2) that

\[
P(\mathcal{H}_n^b((A_i)_{i \leq b}) \text{ connected}) \leq \sum_{\mathbf{f} \in \mathcal{F}_b} \left( \prod_{i=1}^b \frac{1}{f_i!} \right) \sum_{\sigma \in \mathcal{S}_b^1} P(\mathcal{H}_n^b((A_{\sigma(i)})_{i \leq b}) \text{ is } \mathbf{f} \text{-connected}).
\]

We substitute this into (3.1), and use that the second factor inside the sum in (3.1) is invariant under label permutations. So we arrive at

\[
P(\mathcal{E}_1(b)) \leq \frac{1}{b} \sum_{\mathbf{f} \in \mathcal{F}_b} \left( \prod_{i=1}^b \frac{1}{f_i!} \right) \sum_{A_1} \frac{1}{(b-1)!} \sum_{A_2} \cdots \sum_{A_b} \left( P(\mathcal{H}_n^b((A_{\sigma(i)})_{i \leq b}) \text{ is } \mathbf{f} \text{-connected}) \cdot P((\cup_i \tilde{\partial}_{\text{int}} A_{\sigma(i)}) \neq \mathcal{H}_n^b (\Lambda_n \setminus \cup_i A_{\sigma(i)})) \right).
\]

We will now argue that the sum over the permutations and the factor \( 1/(b-1)! \) cancel each other. Given \( A_1 \), let \((B_2, \ldots, B_b)\) be an arbitrary 1-disconnected sequence of 1-connected blocks in \( A_{\text{small}} \) of total size at least \( k - |A_1| \) (also 1-disconnected from \( A_1 \)). Then, for any permutation \( \sigma \in \mathcal{S}_b^1 \), in the summations over the blocks \( A_2, \ldots, A_b \), there is precisely one combination of blocks such that \( A_{\sigma(i)} = B_i \) for all \( i \leq b \). Hence, when summing over all permutations \( \sigma \in \mathcal{S}_b^1 \), we counted the case that the blocks are \( A_1, B_2, \ldots, B_b \) exactly \( (b-1)! \) times. This cancels the factor \( 1/(b-1)! \), and we arrive at

\[
P(\mathcal{E}_1(b)) \leq \frac{1}{b} \sum_{\mathbf{f} \in \mathcal{F}_b} \left( \prod_{i=1}^b \frac{1}{f_i!} \right) \sum_{A_1, \ldots, A_b \in A_{\text{small}}} P(\mathcal{H}_n^b((A_i)_{i \leq b}) \text{ f-connected}) \cdot P((\cup_{i \leq b} \tilde{\partial}_{\text{int}} A_i) \neq \mathcal{H}_n^b (\Lambda_n \setminus \cup_{i \leq b} A_i)),
\]

where we omitted under the summation of the blocks \( A_1, \ldots, A_b \) are 1-disconnected from each other. Lastly, we prescribe the sizes of the boundaries of the blocks. We introduce the possible boundary-length vectors:

\[
\mathcal{M}_b(k) := \left\{ \mathbf{m} = (m_1, \ldots, m_b) \in \mathbb{N}^b : \exists A_1, \ldots, A_b \in A_{\text{small}} \text{ s.t. } |\tilde{\partial}_{\text{int}} A_i| = m_i \forall i \leq b, \quad |\cup_{i \leq b} A_i| \geq k \right\}.
\]
Then,
\[
\mathbb{P}(\mathcal{E}_1(b)) \leq \frac{1}{b} \sum_{m \in M_b(k)} \sum_{f \in F_b} \left( \prod_{i \in [b]} \frac{1}{f_i!} \right) \sum_{A_1, \ldots, A_b \in A_{\text{small}}} \mathbb{P}(\mathcal{H}_{\mathcal{G}_n}((A_i)_i) \text{ f-conn.}) \cdot \mathbb{P}(\cup_i \partial_{\text{int}} A_i \neq \mathcal{G}_n \left( A_i \setminus \cup_i A_i \right))
\]
where we omitted under the summation that the blocks are 1-disconnected from each other and in \(A_{\text{small}}\). In what follows we we omit these descriptions under the sum for readability. The next two statements will imply Lemma \ref{lem:2.9}.

**Statement 3.2 (Counting spanning trees).** Let \(\mathcal{G}_n\) be long-range percolation on \(\Lambda_n\) as in Definition \ref{def:1.1} with \(d \geq 2\), \(\alpha > 1 + 1/d\). There exists \(\epsilon_{\mathcal{G},b} > 0\) such that for all fixed \(m \in M_b(1)\)
\[
\sum_{f \in F_b} \left( \prod_{i \in [b]} \frac{1}{f_i!} \right) \sum_{A_1, \ldots, A_b \in A_{\text{small}}} \mathbb{P}(\mathcal{H}_{\mathcal{G}_n}((A_i)_i) \text{ is f-connected}) \leq n \exp \left( (\epsilon_{\mathcal{G},b} + \log(p\beta^a)) \sum_{i \in [b]} m_i \right).
\]

**Statement 3.3 (Isolation).** Let \(\mathcal{G}_n\) be long-range percolation on \(\Lambda_n\) as in Definition \ref{def:1.1} with \(d \geq 2\), \(\alpha > 1 + 1/d\). There exists \(\epsilon_{\mathcal{G},b} > 0\) such that for each \(\beta \geq 1\), any \(m \in M_b(1)\), and any 1-disconnected blocks \(A_1, \ldots, A_b \in A_{\text{small}}\) with \(\partial_{\text{int}} A_i = m_i\) for all \(i\),
\[
\mathbb{P}(\cup_i \partial_{\text{int}} A_i \neq \mathcal{G}_n \left( A_i \setminus \cup_i A_i \right)) \leq \exp \left( - \epsilon_{\mathcal{G},b} \log \left( \frac{1}{1-p} \right) \beta^{1/d} \sum_{i \in [b]} m_i \right).
\]

We prove the two statements below and show first that Lemma \ref{lem:2.9} follows from them.

**Proof of Lemma \ref{lem:2.9}, assuming Statements 3.2 and 3.3.** Substituting the bounds from Statements 3.2 and 3.3 into the right-hand side of \(3.4\) yields
\[
\mathbb{P}(\mathcal{E}_1(b)) \leq \frac{1}{b} \sum_{m \in M_b(k)} n \exp \left( - \epsilon_{\mathcal{G},b} \log \left( \frac{1}{1-p} \right) \beta^{1/d} - \epsilon_{\mathcal{G},b} - \log(p\beta^a) \right) \sum_{i \in [b]} m_i \right).
\]
In what follows we evaluate the summation over the vectors \(m \in M_b(k)\). We recall from \(3.3\) that \(m\) represents the vector of interior boundary sizes of 1-connected sets \((A_i)_i \subseteq \mathcal{A}_{\text{small}}\) with total size at least \(k\), and \(A_1 = A_1 \leq 3n/4\) for all \(i \leq b\) by the definition of \(\mathcal{A}_{\text{small}}\) in \(2.7\). In \(3.3\), the boundary is taken with respect to \(\mathbb{Z}^d\), i.e., *not* with respect to \(\Lambda_n\). By the isoperimetric inequality in Claim \ref{claim:2.6}, for all blocks \((A_i)_i \subseteq \mathcal{A}_{\text{small}}\) simultaneously holds that \(\partial_{\text{int}} A_i \geq |A_i|^{d/(d-1)}\). Since the function \(g(k) = k^{(d-1)/d}\) is concave, we obtain for all \(m \in M_b(k)\) and any \(A_1, \ldots, A_b \in A_{\text{small}}\) that
\[
m_1 + \ldots + m_b = |\partial_{\text{int}} A_1| + \ldots + |\partial_{\text{int}} A_b| \geq |A_1|^{d/(d-1)} + \ldots + |A_b|^{d/(d-1)} \geq k^{(d-1)/d}.
\]
We define the set \(M_b(k, \ell) := \{m \in M_b(k) : m_1 + \ldots + m_b = \ell\}\). By standard estimates (using that each summand is at least one), we bound \(|M_b(k, \ell)| \leq (\ell + 1)^b \leq \ell^b \leq 2^{d\ell} \leq e^{2\ell}\). Hence, separating the summation in \(3.7\) according to the possible values of \(\sum_i m_i = \ell \geq k^{(d-1)/d}\), we arrive at
\[
\mathbb{P}(\mathcal{E}_1(b)) \leq \frac{1}{b} \sum_{\ell = k^{(d-1)/d}} \sum_{\ell = k^{(d-1)/d}} n \exp \left( - \ell \left( \epsilon_{\mathcal{G},b} \log \left( \frac{1}{1-p} \right) \beta^{1/d} - \epsilon_{\mathcal{G},b} - \log(p\beta^a) - 2 \right) \right).
\]
Since \(b \leq \lfloor n/2 \rfloor\), we obtain by a union bound over the number of blocks that
\[
\mathbb{P}(\cup_{b \leq \lfloor n/2 \rfloor} \mathcal{E}_1(b)) \leq \sum_{b = 1}^{\lfloor n/2 \rfloor} \frac{1}{b} \sum_{\ell = k^{(d-1)/d}} \sum_{\ell = k^{(d-1)/d}} n \exp \left( - \ell \left( \epsilon_{\mathcal{G},b} \log \left( \frac{1}{1-p} \right) \beta^{1/d} - \epsilon_{\mathcal{G},b} - \log(p\beta^a) - 2 \right) \right).\]

We investigate the factor after \(\ell\) in the exponent. For fixed \(p \in (0, 1)\), this factor is positive whenever \(\beta\) is sufficiently large, depending on \(p, d, \alpha\), yielding \(\beta_* (p, d, \alpha)\) in Lemma \ref{lem:2.9}. Evaluating the geometric summation yields \(\frac{1}{d} 1/p d\), where the factor \(\log n\) comes from the first summation in \(3.9\). Whenever \(p\) is sufficiently close to 1 (larger than \(p_d\) for some dimension-dependent \(p_d < 1\)), \(\beta_* = 1\) can be also chosen since in this case the factor \(\log(1/p)\beta^1\) dominates the term \(- \log(p\beta^a)\) for all \(\beta \geq 1\), and the statement holds with \(\beta_* = 1\). \(\square\)
3.1. Proof of Statement 3.2 We start with a geometric claim.

Claim 3.4. There exists a constant $C_{3.4} > 0$ such that for each block $A \in \mathcal{A}_{\text{small}}$ in $\Lambda_n$ and all $r \in \mathbb{N}$
\[
\left| \left\{(x, y) \in \mathbb{Z}^d \times \mathbb{Z}^d : x \in A, y \notin A, \|x - y\| \leq (r, r + 1) \right\} \right| \leq C_{3.4} r^d \|\partial \text{int} A\|.
\] (3.10)

Proof. We start counting line-segments of the right length with endpoints in $\mathbb{Z}^d$ crossing a single unit square that will be centered later at some vertex in $\partial \text{int} A$.

Let $B_0 := [-1/2, 1/2]^d$. For two vertices $x, y \in \mathbb{Z}^d$, let $L_{x,y}$ denote the segment between $x$ and $y$ on the unique line connecting $x$ and $y$. Define
\[
\text{Cross}(r) := \{(x, y) \in \mathbb{Z}^d \times \mathbb{Z}^d : \|x - y\|_2 \leq (r, r + 1), L_{x,y} \cap B_0 \neq \emptyset\}.
\]
We will show that there exists a constant $C_{3.4} > 0$ such that
\[
|\text{Cross}(r)| \leq C_{3.4} r^d.
\] (3.11)
Indeed, for each pair $(x, y) \in \text{Cross}(r)$, at least one of the inequalities $\|x\| \geq r/2$ and $\|y\| \geq r/2$ is satisfied. Without loss of generality we may assume that $\|x\| \geq r/2$, and then also $\|y\| \leq r + 1$.
Fix then such a vertex $x$. Let $S_x$ denote the smallest spherical cone with apex at $x$ that completely contains $B_0$. This cone has then radius between $\|x\| + 1/2$ and $\|x\| + \sqrt{d}/2$. Let $S_x(r)$ denote a cone with apex $x$ that has the same boundary lines (and the same angle) as $S_x$, but radius exactly $r$. Then, every $y \in \mathbb{Z}^d$ such that $(x, y) \in \text{Cross}(r)$ must be contained in $S_x(r + 1) \setminus S_x(r)$, since all half-lines emanating from $x$ that cross $B_0$ are contained in $S_x(\infty)$, and $\|x - y\| \in (r, r + 1]$. Since the radius of $S_x(r + 1)$ is at most by a factor two larger than the radius of $S_x$ for all $r \geq 1$, by homothety of the cones, $|\partial S_x(r + 1) \setminus S_x| \leq \|\partial \text{int} A\|$ is bounded from above by a dimension-dependent constant, and so for each $x$ with $\|x\|_2 \in [r/2, r + 1]$, the number of pairs $(x, y) \in \text{Cross}(r)$ is bounded from above by a dimension-dependent constant. Summing over all the at most $O(r^d)$ many such $x$, we obtain (3.11) for some $C_{3.4} > 0$.

To arrive to (3.10), the block $A \in \mathcal{A}_{\text{small}}$ in (3.7) ensures that $A = \tilde{A}$. Its inner boundary $\partial \text{int} A$ is then $*$-connected by Claim 2.5 Part (iv) ([12] Lemma 2.1)). This implies that there exists a continuous surface fully contained in $\partial \text{int} A$ separating vertices in $A \setminus \partial \text{int} A$ from vertices in $\Lambda_n \setminus A$. Hence, for each pair $x \in A$ and $y \notin A$, there exists (at least one) vertex $z \in \partial \text{int} A$ such that $L_{x,y}$ intersects the axis-parallel box $z + B_0$. Here $x = z$ may occur. The statement of the claim now follows by (3.11) when summing the at most $C_{3.4} r^d$ such pairs for each vertex $z$ in the boundary of $A$.

We continue with a lemma treating the connectedness of the block graphs, i.e., the inner summation on the left-hand side of (3.6) in Statement 3.2. We point out that in this lemma we only bound the event that the block graph $\mathcal{G}_n$ is connected, not the event that the actual graph is connected.

Lemma 3.5. Let $\mathcal{G}_n$ be long-range percolation on $\Lambda_n$ as in Definition 1.1 with $d \geq 2$, $\alpha > 1 + 1/d$. There exists a constant $C_{3.5} > 0$ such that for all $m \in \mathcal{M}_b(1)$, $f \in \mathcal{F}_b$,
\[
\sum_{A_1, \ldots, A_b \in \mathcal{A}_{\text{small}}, |\partial \text{int} A_i| = m_i, \forall i \leq b} \mathbb{P}(\mathcal{H}_n((A_i)_{i \in [b]}) \text{ is } f\text{-connected}) \leq n(C_{3.5} \beta^\alpha)^{b-1} \prod_{i \in [b]} e^{e_{m_i} m_i f},
\] (3.12)
where we omitted from the summation that $A_1, \ldots, A_b$ are 1-disconnected from each other.

We comment that it is this lemma in the proof that crucially uses that $\alpha > 1 + 1/d$.

Proof. We will prove the statement by induction on $b$. We first define the finite constant, using that $\alpha > 1 + 1/d$ as follows:
\[
C_{3.5} = C_{3.4} \sum_{r=1}^{\infty} r^{-(\alpha-1)d},
\] (3.13)
We start with the initialization. Assume first that $b = 1$, which corresponds to a tree on a single vertex (representing the block $A_1$), so its forward degree $f_1 = 0$. A tree on a single vertex is connected by convention. We obtain
\[
\sum_{A_1 \in \mathcal{A}_{\text{small}}, |\partial \text{int} A_1| = m_1} \mathbb{P}(\mathcal{H}_n((A_1)) \text{ is } f\text{-connected}) \leq |\{A_1 \in \mathcal{A}_{\text{small}} : n(\tilde{\partial} \text{int} A_1) = m_1\}|
\leq \sum_{x \in \Lambda_n} |\{A_1 \in \mathcal{A}_{\text{small}} : A_1 \ni x, |\partial \text{int} A_1| = m_1\}|. 
\]
Since $A = \tilde{A}$ for all $A \in \mathcal{A}_{\text{small}}$ by definition in (2.7), we can apply Peierls’ argument in Lemma 2.7 that yields, since $|A_n| = n$,

$$\sum_{A_1 \in \mathcal{A}_{\text{small}}: \partial A \nmid \partial A_1 \vdash m_1} \mathbb{P}(\mathcal{H}_{\mathcal{G}}((A_i)_i) \text{ is } f\text{-connected}) \leq \sum_{x \in A_n} e^{\rho_m m_1} = ne^{\rho_m m_1}.$$  

Since $m_f^1 = m_0^1 = 1$, this finishes the induction base for (3.12). We now advance the induction. Assume (3.12) holds up to $b-1$. Let $f \in \mathcal{F}_b$ and consider the summation over $A_0 \in \mathcal{A}_{\text{small}}$ on the left-hand side in (3.11). By construction of the $f$-tree in Definition 3.1, the $b$-th block is a leaf in the $f$-tree, and $f_b = 0$. Its parent in the $f$-tree is the largest vertex-label $\ell$ in $f$ that is nonzero, and the remaining labeled graph upon removing $b$ is a tree, with a labeling in $\mathcal{F}_{b-1}$ (see the comment below Definition 3.1). Then, the forward degrees of this new tree are given by $P' : = (f_1, \ldots, f_{\ell - 1}, f_\ell - 1, f_{\ell + 1}, \ldots, f_{b - 1}) \in \mathcal{F}_{b-1}$, since the forward degree of the vertex $\ell$ decreased by one upon removing the leaf $b$. With this notation at hand,

$${\mathcal{H}}_{\mathcal{G}}((A_i)_{i \in [b]})$$

is $f$-connected \iff $\mathcal{H}_{\mathcal{G}}((A_i)_{i \in [b-1]})$ is $f'$-connected $\cap \{A_\ell \leftrightarrow \mathcal{G}_n A_b\}$.

Independence of edges in $\mathcal{G}_n$ by Definition 1.4 yields

$$\sum_{A_1, \ldots, A_b} \mathbb{P}(\mathcal{H}_{\mathcal{G}}((A_i)_{i \in [b]}) \text{ f-conn.}) \leq \sum_{A_1, \ldots, A_{b-1}} \mathbb{P}(\mathcal{H}_{\mathcal{G}}((A_i)_{i \in [b-1]}) \text{ f'-conn.}) \sum_{A_0} \mathbb{P}(A_\ell \leftrightarrow \mathcal{G}_n A_b) \quad (3.14)$$

where in the subscripts of the sums (and also in the remainder of the proof) we omitted the conditions that the sets have to satisfy: $A_1, \ldots, A_b \in \mathcal{A}_{\text{small}}$, and $|\partial A_i| = m_i$ for all $i \leq b$.

We focus on the summation over $A_b$. Here, $A_0 \in \mathcal{A}_{\text{small}}$ and $A_b$ is 1-disconnected from $A_\ell$. The forward degrees of this new tree are given by $P' := (f_1, \ldots, f_{\ell - 1}, f_\ell - 1, f_{\ell + 1}, \ldots, f_{b - 1}) \in \mathcal{F}_{b-1}$, since the forward degree of the vertex $\ell$ decreased by one upon removing the leaf $b$. With this notation at hand, it follows that

$$\sum_{A_b} \mathbb{P}(A_\ell \leftrightarrow \mathcal{G}_n A_b) \leq \sum_{r=1}^{\infty} \sum_{x \in A_\ell, y \in \mathbb{Z}^d \setminus A_\ell} 1\{\|x-y\| \in (r, r+1]\} \sum_{y \in A_b} \mathbb{P}(x \leftrightarrow \mathcal{G}_n, y)$$

$$\leq p_\beta^3 \sum_{r=1}^{\infty} r^{-ad} \sum_{x \in A_\ell, y \in \mathbb{Z}^d \setminus A_\ell} 1\{\|x-y\| \in (r, r+1]\} \sum_{y \in A_b} 1.$$  

In the last row the sum over $y$ runs over a larger set than the actual allowed set of vertices: we did not exclude vertices from the other blocks $\{A_1, \ldots, A_{\ell - 1}, A_{\ell + 1}, \ldots, A_{b-1}\}$. Using (2.6), we bound the last sum over $A_b$ from above by $\exp(p_\beta m_b)$ . Next, we can apply Claim 3.3 to evaluate the summation over $x \in A_\ell, y \in \mathbb{Z}^d \setminus A_\ell$, since this sum equals the set described in (3.10) with $A = A_\ell$. The conditions of the claim are satisfied since $A_\ell = \tilde{A}_\ell$ by assuming $A_\ell \in \mathcal{A}_{\text{small}}$. Hence

$$\sum_{x \in A_\ell, y \in \mathbb{Z}^d \setminus A_\ell} 1\{\|x-y\| \in (r, r+1]\} \sum_{y \in A_b} 1 = e^{\rho_m m_b} C_{\ref{eqn:rho_m_m_b}} d^{|\partial A_\ell|} = e^{\rho_m m_b} C_{\ref{eqn:rho_m_m_b}} d^{m_\ell}.$$  

Substituting this back into (3.15) yields with the constant $C_{\ref{eqn:rho_m_m_b}}$ from (3.13),

$$\sum_{A_b} \mathbb{P}(A_\ell \leftrightarrow \mathcal{G}_n A_b) \leq p_\beta^3 \sum_{r=1}^{\infty} r^{-ad} P^{(1-a)d} = C_{\ref{eqn:rho_m_m_b}} \sum_{r=1}^{\infty} r^{-ad} P^{(1-a)d} = C_{\ref{eqn:rho_m_m_b}} \sum_{r=1}^{\infty} r^{-ad} P^{(1-a)d}.$$  

We substitute this bound back into (3.14), and use the induction hypothesis:

$$\sum_{A_1, \ldots, A_b} \mathbb{P}(\mathcal{H}_{\mathcal{G}}((A_i)_{i \in [b]}) \text{ f-conn.}) \leq C_{\ref{eqn:rho_m_m_b}} \sum_{A_1, \ldots, A_{b-1}} \mathbb{P}(\mathcal{H}_{\mathcal{G}}((A_i)_{i \in [b-1]}) \text{ f'-conn.})$$

$$\leq C_{\ref{eqn:rho_m_m_b}} \sum_{i \in [b-1]} e^{\rho_m m_i f_i} \prod_{i \in [b-1]} e^{\rho_m m_i f_i}$$

$$\leq C_{\ref{eqn:rho_m_m_b}} \sum_{i \in [b-1]} e^{\rho_m m_i n(b-1)} \prod_{i \in [b-1]} e^{\rho_m m_i f_i}.$$  

To obtain the third row we used that $f_i' = f_i$ for all $i \neq \ell, i \leq b - 1$, and $f_\ell' = f_\ell - 1$ by construction, yielding the $1/m_\ell$ factor. We can rearrange the expression and obtain (3.12), using that $f_b = 0$ (the last block is a leaf). This finishes the proof. □
We are ready to prove Statement 3.2.

**Proof of Statement 3.2.** Using Lemma 3.3 in (3.5) of Statement (3.2), we arrive at

\[
\sum_{f \in F_b} \left( \prod_{i \in [b]} \frac{1}{f_i} \right) \sum_{A_1, \ldots, A_b} P(H_{G_n}((A_i)_i) \text{ is } f\text{-connected}) \leq n \sum_{f \in F_b} (C_3 p^{b\alpha} b^{-1}) \prod_{i \in [b]} e^{\frac{\log m_i}{f_i}}. \tag{3.16}
\]

We first analyze a single summand, i.e., the value for a fixed \( f \). Using again that Lemma 3.5 in (3.5) of Statement (3.2), we arrive at

\[
W \text{e first analyze a single summand, i.e., the value for a fixed } f. \text{ By Stirling's approximation, there exists a constant } c > 1 \text{ such that } f_i! \geq (f_i/e)^{f_i}/c. \text{ Thus,}
\]

\[
n(3.3) p^{b\alpha} b^{-1} \prod_{i \in [b]} e^{\frac{\log m_i}{f_i}} \leq n(3.3) p^{b\alpha} b^{-1} \exp \left( (c_{\text{pe}} + 1) \sum_{i \in [b]} m_i \right) \frac{e^{\log f_i}}{f_i}.
\]

It follows from standard differentiation techniques that for any \( a, x \geq 1 \), the function \( g_a(x) = (ae/x)^x \) is maximized at \( x = a \). Maximizing all factors \( (m_i/e)/f_i \) at \( f_i = m_i \) yields that \( (m_i/e/f_i) \leq e^{-m_i} \) for all \( i \leq b \). Since by definition of \( F_b \) in Definition 3.1, we have \( f_1 + \ldots + f_b = b - 1 \) for all \( f \in F_b \), we have

\[
n(3.3) p^{b\alpha} b^{-1} \prod_{i \in [b]} e^{\frac{\log m_i}{f_i}} \leq n(3.3) p^{b\alpha} b^{-1} \exp \left( (c_{\text{pe}} + 1) \sum_{i \in [b]} m_i \right) e^{-1} \sum_{f \in F_b} \frac{1}{f_i}.
\]

Using again that \( f_1 + \ldots + f_b = b - 1 \) for all \( f \in F_b \), and using the same combinatorial bounds as for \( m \) above (3.8), we obtain \( |F_b| \leq (2b)^b \leq 2^b \leq 2^{2b} \leq (2 \sum_{i \in [b]} m_i) \), finishing the proof with \( c'' := c'' + 2 \). □

**Proof of Statement 3.3.** We start with a geometric claim. Recall \( A_{\text{small}} \) from (2.4), and holes from Definition 2.1.

**Claim 3.6.** Let \( (A_1, \ldots, A_b) \subseteq A \) be a 1-disconnected sequence of blocks without holes (i.e., \( A_i = A \) for all \( i \leq b \), with \( A := \bigcup_{i \leq b} A_i \subseteq \Lambda_n \). Then, \( (\Lambda_n \setminus A) \cup (\bigcup_{i \leq b} \partial_{\text{int}} A_i) \supseteq \partial_{\text{int}} \Lambda_n \). Moreover, \( (\Lambda_n \setminus A) \cup (\bigcup_{i \leq b} \partial_{\text{int}} A_i) \) is *-connected.

**Proof.** We first show that \( (\Lambda_n \setminus A) \cup (\bigcup_{i \leq b} \partial_{\text{int}} A_i) \supseteq \partial_{\text{int}} \Lambda_n \). When \( x \in (\partial_{\text{int}} \Lambda_n \setminus A) \), then \( x \in \Lambda_n \setminus A \). The (only) other case is when \( x \in \partial_{\text{int}} \Lambda_n \cap A \). Then there must exist \( A_i \subseteq \Lambda_n \) such that \( x \in A_i \). Since \( x \in \partial_{\text{int}} \Lambda_n \) and \( A \subseteq \Lambda_n \), it follows from (2.3) in Definition 2.1 that there is a vertex \( y \) in \( \mathbb{Z}^d \setminus \Lambda_n \) neighboring \( x \). Since \( x \in A_i \subseteq \Lambda_n \), we evidently have \( y \not\in A_i \), hence \( y \in \partial_{\text{ext}} A_i \). As a result of (2.21), \( x \in \partial_{\text{int}} A_i \), establishing the statement.

We turn to prove *-connectedness of \((\Lambda_n \setminus A) \cup (\bigcup_{i \leq b} \partial_{\text{int}} A_i)\). Using Claim 3.3, we decompose the set \( \Lambda_n \setminus A \) into a 1-disconnected sequence of 1-connected blocks \( (B_j)_{j \leq b'} \) for some \( b' \geq 1 \):

\[
\bigcup_{j \leq b'} B_j = \Lambda_n \setminus A. \tag{3.17}
\]

To keep track of indices, for each set \( \partial_{\text{int}} A_i \) we associate a vertex \( a_i \) for \( i \leq b \), and also for each block \( B_j \) a vertex \( t_j \) for \( j \leq b' \). Define the subsets \( \mathcal{I}_a \subseteq \{a_1, \ldots, a_b\} \) and \( \mathcal{I}_t \subseteq \{t_1, \ldots, t_{b'}\} \) so that for all \( i \leq b \), \( a_i \in \mathcal{I}_a \) if and only if \( \partial_{\text{int}} A_i \cap \partial_{\text{int}} \Lambda_n \neq \emptyset \), and for all \( j \leq b' \), \( t_j \in \mathcal{I}_t \) if and only if \( B_j \cap \partial_{\text{int}} \Lambda_n \neq \emptyset \). In words, these are the vertices corresponding to the sets that intersect the boundary of \( \Lambda_n \).

By the first statement of the claim and (3.17), \( \partial_{\text{int}} \Lambda_n \) is completely contained in \((\bigcup_{b' \leq b} \partial_{\text{int}} A_i) \cup (\bigcup_{b' \leq b} B_j)\), hence \( \mathcal{I}_a \cup \mathcal{I}_t \) is non-empty. Since the indices of the sets/blocks that have an intersection with \( \partial_{\text{int}} \Lambda_n \) are all collected in \( \mathcal{I}_a \) and \( \mathcal{I}_t \), respectively, we have

\[
\partial_{\text{int}} \Lambda_n \subseteq \left( \bigcup_{i \in \mathcal{I}_a} \partial_{\text{int}} A_i \right) \cup \left( \bigcup_{j \in \mathcal{I}_t} B_j \right) =: D. \tag{3.18}
\]
Since we assumed that $A_i = \widetilde{A}_i$ for all $i \leq b$, Claim \[2.4\text{iv}\] is applicable and hence $\tilde{\partial}_{\text{int}}A_i$ is $*$-connected for each $i \leq b$. Further, $(B_j)_{j \leq b}$ are blocks, i.e., 1-connected and also $*$-connected. Thus, since $\tilde{\partial}_{\text{int}}\Lambda_n$ is 1-connected in dimensions 2 and higher, it is also $*$-connected, and hence the set $D$ on the right-hand side of \[3.18\] is also $*$-connected.

We now decompose the set $(\cup_{j \leq b} B_j) \cup (\cup_{i \leq b} \tilde{\partial}_{\text{int}}A_i)$ (using the edges of the graph $\mathbb{Z}^d_\Lambda$) into $*$-connected 

\textit{components}, in other words, sets that are $*$-connected themselves but are $*$-disconnected from each other. Our goal is then to show that there is only a single component, namely, the one containing vertices of $D$. Assume now that there exists another $*$-connected component $C$ of $(\cup_{j \leq b} B_j) \cup (\cup_{i \leq b} \tilde{\partial}_{\text{int}}A_i)$ that does not contain any vertices of $D$ in \[3.18\]. Being a component in $\mathbb{Z}^d_\Lambda$ also means $C$ is then not $*$-connected to $D$ by the $*$-connectedness of each set in $(B_j)_{j \leq b}$ and $(\tilde{\partial}_{\text{int}}A_i)_{i \leq b}$, this is only possible if $C$ is a union of some of $(B_j)_{j \leq b}$ and $(\tilde{\partial}_{\text{int}}A_i)_{i \leq b}$. Define then $\mathcal{I}_C \subseteq \{a_1, \ldots, a_b\} \cup \{t_1, \ldots, t'_b\}$ so that $a_i \in \mathcal{I}_C$ if and only if $\tilde{\partial}_{\text{int}}A_i \subseteq C$, and similarly $t_j \in \mathcal{I}_C$ if and only if $B_j \subseteq C$:

$$\mathcal{C} = (\cup_{a_i \in \mathcal{I}_C} \tilde{\partial}_{\text{int}}A_i) \cup (\cup_{t_j \in \mathcal{I}_C} B_j)$$

is a $*$-connected component, and $\mathcal{I}_C \cap (\mathcal{I}_n \cup \mathcal{I}_1) = \emptyset$. \[3.19\]

Take now the closest vertex in $\cup_{j \in \mathcal{I}_C} B_j$ to $\tilde{\partial}_{\text{int}}\Lambda_n$, i.e., let (with arbitrary tie-breaking rule)

$$x_\star := \arg \min_{x \in B_j, \ t_j \in \mathcal{I}_C} \|x - \tilde{\partial}_{\text{int}}\Lambda_n\|_1,$$

and $x_\star \in B_j$ with $t_j \in \mathcal{I}_C$. \[3.20\]

Since $t_j \in \mathcal{I}_C$, by our assumption in \[3.19\] the block $B_j$, does not intersect $\tilde{\partial}_{\text{int}}\Lambda_n$. So

$$\|x_\star - \tilde{\partial}_{\text{int}}\Lambda_n\|_1 = \|B_j - \tilde{\partial}_{\text{int}}\Lambda_n\|_1 \geq 1.$$ \[3.21\]

Hence, there exists a vertex $y_\star \in \Lambda_n \setminus B_j$ such that $\|y_\star - \tilde{\partial}_{\text{int}}\Lambda_n\|_1 = \|x_\star - \tilde{\partial}_{\text{int}}\Lambda_n\|_1 - 1$ and $\|x_\star - y_\star\|_1 = 1$. This vertex $y_\star$ cannot be part of $\cup_{j \in \mathcal{I}_C} B_j$ since $x_\star$ was minimal, and $y_\star$ cannot be part of $(\cup_{j \leq b} B_j) \setminus (\cup_{t_j \in \mathcal{I}_C} B_j)$ by 1-connectedness of the sets $(B_j)_{j \leq b}$, see above \[3.17\]. So, there exists a block $A_\ell$ such that $y_\star \in A_\ell$ for some $\ell \leq b$. Then immediately also $y_\star \in \tilde{\partial}_{\text{int}}A_\ell$, since $x_\star \in B_j$, serves as an external boundary vertex for $y_\star$ in $A_\ell$.

We also observe that via the pair $(x_\star, y_\star)$ the block $B_j$, is 1-connected (and $*$-connected) to $A_\ell$. Hence, since we assumed $\mathcal{C} \supseteq B_j$, is a $*$-connected component, $A_\ell \subseteq \mathcal{C}$, equivalently, $A_\ell \in \mathcal{I}_C$ and so by \[3.19\], $a_\ell \notin \mathcal{I}_n$. This then implies that $A_\ell \cap \tilde{\partial}_{\text{int}}\Lambda_n = \emptyset$ by definition of $\mathcal{I}_n$ above \[3.18\]. Hence, $\|A_\ell - \tilde{\partial}_{\text{int}}\Lambda_n\|_1 \geq 1$ and in turn $\tilde{\partial}_{\text{ext}}A_\ell \subseteq \Lambda_n$. By the same reasoning as for the existence of $x_\star, y_\star$ below \[3.21\], we find a vertex $y_\star \in \tilde{\partial}_{\text{int}}A_\ell$ that is closest to $\tilde{\partial}_{\text{int}}\Lambda_n$ within $\tilde{\partial}_{\text{int}}A_\ell$, and a vertex $x_\star \in \tilde{\partial}_{\text{ext}}A_\ell$ that is strictly closer to $\tilde{\partial}_{\text{int}}\Lambda_n$ than $y_\star$, with $\|x_\star - y_\star\|_1 = 1$. We arrive at the sequence of (in)equalities:

$$\|x_\star - \tilde{\partial}_{\text{int}}\Lambda_n\|_1 + 1 = \|y_\star - \tilde{\partial}_{\text{int}}\Lambda_n\|_1 = \|\tilde{\partial}_{\text{int}}A_\ell - \tilde{\partial}_{\text{int}}\Lambda_n\|_1 \leq \|y_\star - \tilde{\partial}_{\text{int}}\Lambda_n\|_1 = \|x_\star - \tilde{\partial}_{\text{int}}\Lambda_n\|_1 - 1,$$ \[3.22\]

where the inequality in the middle follows since $y_\star \in \tilde{\partial}_{\text{int}}A_\ell$ by construction, and the equalities follow by the choices of $x_\star, y_\star$, and $x_\star, y_\star$, below \[3.21\].

Since $y_\star \in \tilde{\partial}_{\text{int}}A_\ell, x_\star \in \tilde{\partial}_{\text{ext}}A_\ell$ and the sequence of blocks $(A_i)_{i \leq b}$ is 1-disconnected, $x_\star$ is not in $(\cup_{i \leq b} A_i) = A$. Hence, $x_\star \in \Lambda_n \setminus A = \cup_{j \leq b} B_j$, so there exists some $j_0 \leq b'$ so that $x_\star \in B_{j_0}$.

The block $B_{j_0}$ is then 1-connected via the pair $(x_\star, y_\star)$ to $\tilde{\partial}_{\text{int}}A_\ell$, which is itself $*$-connected, and via the pair $(x_\star, y_\star)$ the set $\tilde{\partial}_{\text{int}}A_\ell$ is 1-connected to $B_{j_0}$. Hence, $B_{j_0}, A_\ell, B_{j_0}$ are all in the same $*$-connected component $C$. However, $\|x_\star - \tilde{\partial}_{\text{int}}\Lambda_n\|_1 < \|x_\star - \tilde{\partial}_{\text{int}}\Lambda_n\|_1$ by the inequality \[3.22\], which contradicts that $x_\star$ was a vertex in $(\cup_{j \in \mathcal{I}_C} B_j)$ that minimized $\|x - \tilde{\partial}_{\text{int}}\Lambda_n\|_1$. So, the $*$-connected component that contains $D$ in \[3.18\] contains all blocks $(\cup_{j \leq b} B_j) = \Lambda_n \setminus A$.

Starting from \[3.20\] with reversing the role of the blocks $(B_j)_{j \leq b'}$ and the sets $(\tilde{\partial}_{\text{int}}A_i)_{i \leq b}$, the same reasoning yields that $\mathcal{I}_C \cap \{a_1, \ldots, a_b\} = \emptyset$, and hence the $*$-connected component that contains $D$ in \[3.18\] contains also all sets $\cup_{i \leq b} \tilde{\partial}_{\text{int}}A_i$. Together, we conclude thus that $\cup_{i \leq b} \tilde{\partial}_{\text{int}}A_i \cup (\Lambda_n \setminus A)$ is $*$-connected.

\textbf{Proof of Statement \[3.3\]} We assume $\beta \geq 1$. Let $A_1, \ldots, A_b \in \mathcal{A}_{\text{small}}$, and denote $A := \cup_{i \leq b} A_i$. We define the set of potential edges between the interior boundary of $A$ with respect to $\mathbb{Z}^d$ and the set of vertices outside $A$ within distance $\beta^{1/d}$ as

$$\Delta(A) := \{x, y \mid x \in \cup_{i \leq b} \tilde{\partial}_{\text{int}}A_i, y \in (\Lambda_n \setminus A) : \|y - x\| \in [1, \beta^{1/d}]\}.$$
Considering the event on the left-hand side of (3.6) in Statement 3.3, we would like \( (\cup_{i \leq b} \partial_{\text{int}} A_i) \) to be not \( G_n \)-connected to the rest of the graph. In order to achieve this, in particular, all edges in \( \Delta(A) \) must be absent. Hence,

\[
P \left( \left( \cup_{i \leq b} \partial_{\text{int}} A_i \right) \not\subset G_n \ (\Lambda_n \setminus \cup_{i \leq b} A_i) \right) \leq (1 - p)^{|\Delta(A)|} = \exp \left( -\log\left( \frac{1}{1-p} \right) |\Delta(A)| \right). \tag{3.23}
\]

Our goal is to show that for some constant \( c > 0 \),

\[
|\Delta(A)| \geq c\beta^{1/d} \sum_{i \in [b]} |\partial_{\text{int}} A_i| = c\beta^{1/d} \sum_{i \in [b]} m_i, \tag{3.24}
\]

which then immediately yields (3.6) in combination with (3.23). In what follows we estimate \( |\Delta(A)| \). In order to do so, we will make use of the boundary \( \partial_{\text{int}} A_i \), i.e., the interior boundary with respect to the box \( \Lambda_n \). Using that all blocks in \( A_{\text{small}} \) have size at most \( 3n/4 \) by definition in (2.24), the conditions of the isoperimetric inequality in Claim 2.6 are satisfied, and hence \( \delta m_i \leq |\partial_{\text{ext}} A_i| \leq m_i \) for all \( i \leq b \). Hence, (3.24) is equivalent to showing that there exist \( c' > 0 \) such that for any 1-disconnected blocks \( A_1, \ldots, A_b \in A_{\text{small}} \), with \( A = \cup_{i \leq b} A_i \)

\[
|\Delta(A)| \geq c'\beta^{1/d} \sum_{i \in [b]} |\partial_{\text{int}} A_i|, \tag{3.25}
\]

since then (3.24) holds with \( c = c'\delta \).

In order to show (3.25), our goal is to find enough pairs of vertices in \( \Delta(A) \) around a linear fraction of vertices in \( \cup_{i \leq b} \partial_{\text{int}} A_i \). For this, we claim that a set \( T := \{(x_\ell, y_\ell)\}_{\ell \geq 1} \) with the following properties exists:

(i) \( x_\ell \in \cup_i \partial_{\text{int}} A_i, \ y_\ell \in \cup_i \partial_{\text{ext}} A_i, \) and \( \|x_\ell - y_\ell\| = 1 \) for all \( \ell \geq 1 \);

(ii) each vertex \( z \in \Lambda_n \) appears at most once in a pair in \( T \);

(iii) \( |T| \geq \sum_{i \in [b]} |\partial_{\text{int}} A_i|/(2d) \).

Note that requirement (i) implies that all \((x_\ell, y_\ell)\in T\) are elements of \( \Lambda_n \times \Lambda_n \). We now show that a set \( T \) exists. Consider the following greedy algorithm: order the vertices in \( \cup_i \partial_{\text{int}} A_i \) in an arbitrary order, to obtain the list \((v_1, v_2, \ldots, v_M)\) with \( M = \sum_{i \in [b]} |\partial_{\text{int}} A_i| \). Since each \( v_j \) is in \( \cup_i \partial_{\text{int}} A_i \), for each \( v_j \) there is at least one vertex \( y_j \in \cup_i \partial_{\text{ext}} A \subseteq (\Lambda_n \setminus A) \) with \( \|v_j - y_j\| = 1 \) by Definition 2.4 (recall that the sets \( A_1, \ldots, A_b \) are 1-disconnected). Starting with \( T_1 := \{(v_1, y_1)\} \), going through the ordering of \((v_j, y_j)\) one-by-one, append the pair \((v_j, y_j)\) to the list \( T_{j-1} \), if and only if \( y_j \) has not been contained in any pair of \( T_{j-1} \) yet and so obtain \( T_j \). Then set \( T := T_M \). Since any \( y \in \cup_i \partial_{\text{ext}} A_i \) neighbors at most \( 2d \) many interior boundary vertices, adding a certain pair \((v_j, y_j)\) only affects at most \( 2d - 1 \) other indices where a pair may not be added later. Hence,

\[
|T| \geq \frac{1}{2d} \sum_{i \in [b]} |\partial_{\text{int}} A_i|. \tag{3.26}
\]

Next, assume that \( \beta^{1/d} \geq 2\sqrt{d} + 2 \) and set \( R := \left\lfloor (\beta^{1/d} - 1)/\sqrt{d} \right\rfloor \geq 1 \). Take any pair \((x_\ell, y_\ell)\in T \). Since \( \cup_{i \leq b} \partial_{\text{int}} A_i \cup (\Lambda_n \setminus A) \) is \(*\)-connected by Lemma 3.6 there exists a self-avoiding path

\[
\pi(x_\ell) = (x_\ell, z_1^{(x_\ell)}, \ldots, z_R^{(x_\ell)}) \subseteq \cup_{i \leq b} \partial_{\text{int}} A_i \cup (\Lambda_n \setminus A) \tag{3.27}
\]

(since the set on the right-hand side contains \( \partial_{\text{int}} \Lambda_n \), which has size \( \Theta(n^{(d-1)/d}) \), the set on the right-hand side has size at least \( R \) for \( n \) sufficiently large, so such a self-avoiding path of length \( R \) then exists). By the triangle inequality,

\[
\|x_\ell - z_j^{(x_\ell)}\| \leq \sqrt{d} R \leq \beta^{1/d}, \quad \text{and} \quad \|y_\ell - z_j^{(x_\ell)}\| \leq \sqrt{d} R + 1 \leq \beta^{1/d}. \tag{3.28}
\]

Define now the type \( \text{typ}(z_j^{(x_\ell)}) := x_\ell \) if \( z_j^{(x_\ell)} \in (\Lambda_n \setminus A) \) and set \( \text{typ}(z_j^{(x_\ell)}) := y_\ell \) if \( z_j^{(x_\ell)} \in \cup_{i \leq b} \partial_{\text{int}} A_i \). Then define the set of (unordered) pairs representing potential edges in \( G_n \),

\[
\Delta(x_\ell, y_\ell) := \{ \{z_1^{(x_\ell)}, \text{typ}(z_1^{(x_\ell)})\}, \ldots, \{z_R^{(x_\ell)}, \text{typ}(z_R^{(x_\ell)})\} \} \subseteq \Delta(A). \tag{3.29}
\]

The inclusion holds since for each of these pairs, exactly one element is in \( \Lambda_n \setminus A \) and the other one is in \( \cup_{i \leq b} \partial_{\text{int}} A_i \), and the distance between the two vertices of each pair is at most \( \beta^{1/d} \) by (3.28). We
claim that
\[ |\Delta(A)| \geq \left| \bigcup_{(x_t, y_t) \in \mathcal{T}} \Delta(x_t, y_t) \right| \overset{\circ}{=} (1/2) \sum_{\ell=1}^{\lvert \mathcal{T} \rvert} |\Delta(x_t, y_t)| = |\mathcal{T}| \cdot R/2. \tag{3.30} \]

To see the inequality with \( \circ \), we show that each potential edge \( \{z, z'\} \in \Lambda_n \times \Lambda_n \) appears at most twice in a set in the union in the middle. Consider \( \{z, z'\} \in \Lambda_n \times \Lambda_n \). First, assume that there exists \( \ell \) such that \((z, z') = (x_t, y_t) \in \mathcal{T} \) or \((z', z) = (x_t, y_t) \in \mathcal{T}\). Without loss of generality, we assume that the pair is ordered such that \((z, z') = (x_t, y_t) \in \mathcal{T} \). Then there is no \((x_j, y_j) \in \mathcal{T} \) different from \((x_t, y_t)\) such that \(\{z, z'\} \in \Delta(x_j, y_j)\), since each element in \(\Delta(x_j, y_j)\) contains either \(x_j\) or \(y_j\), which are different from \(x_t\) and from \(y_t\) by requirement (ii) in the construction of \(\mathcal{T}\). Moreover, the element \(\{z, z'\} = (x_t, y_t)\) is contained at most once in the set \(\Delta(x_t, y_t)\), since the first coordinates in (3.29) are different as they form a self-avoiding path, and the first coordinates do not contain \(x_t = z\) by (3.27).

Next, assume that \((z, z'), (z', z) \notin \mathcal{T}\), but \(\{z, z'\}\) is contained in some \(\Delta(x_t, y_t)\) \((x_t, y_t) \in \mathcal{T}\). Then, either \(z\) or \(z'\) must be equal to either \(x_t\) or to \(y_t\) by (3.29). Assume without loss of generality that \(z \in \{x_t, y_t\}\), and therefore \(z' \notin \{x_t, y_t\}\). Thus, \(\{z, z'\}\) is contained exactly once in \(\Delta(x_t, y_t)\). The only way that \(\{z, z'\}\) could be in a set \(\Delta(x'_t, y'_t)\) for some \((x'_t, y'_t) \neq (x_t, y_t)\), is when \(z' \in \{x'_t, y'_t\}\) and \((x'_t, y'_t) \in \mathcal{T}\) for some \(\ell' \neq \ell\). This implies that the element \(\{z, z'\}\) can be contained at most twice in a set in the union in (3.30), namely in \(\Delta(x_t, y_t)\) and \(\Delta(x'_t, y'_t)\), and the inequality \( \circ \) in (3.30) holds.

Combining (3.30) with (3.29), \(R = ([\beta^{1/d} - 1]/\sqrt{d})\), and the assumption that \(\beta^{1/d} \geq 2\sqrt{d} + 2\), (see before (3.28)), we arrive at
\[ |\Delta(A)| \geq \frac{R}{4d} \sum_{i \in [b]} |\partial_{int} A_i| = \frac{1}{4d} \left( \frac{\beta^{1/d} - 1}{\sqrt{d}} \right) \sum_{i \in [b]} |\partial_{int} A_i| \geq \frac{\beta^{1/d}}{8d\sqrt{d}} \sum_{i \in [b]} |\partial_{int} A_i|, \]

since whenever \(x \geq 2\sqrt{d} + 2\), then \(|(x - 1)/\sqrt{d}| \geq x/2\sqrt{d}\).

This proves (3.29) whenever \(\beta^{1/d} \geq 2\sqrt{d} + 2\). For the case \(1 \leq \beta^{1/d} \leq 2\sqrt{d} + 2\), we use that each vertex on the interior boundary is within distance one from a vertex on the exterior boundary, hence
\[ |\Delta(A)| \geq \sum_{i \in [b]} |\partial_{int} A_i| \geq \frac{\beta^{1/d}}{2\sqrt{d} + 2} \sum_{i \in [b]} |\partial_{int} A_i|, \]

and so (3.29) holds for both cases with \(c := 1/\max\{8d\sqrt{d}, 2\sqrt{d} + 2\}\). This finishes the proof of Statement 3.3. \(\square\)

4. Counting holes

We turn to the proof of Lemma 2.10. We set up a few preliminaries about holes. Recall that \(\mathcal{A}\) denotes the 1-connected blocks in \(\Lambda_n\) from (2.1), and that \(\mathcal{A}_{large} = \{A \in \mathcal{A} : |A| > 3n/4, |A| \leq n/2\}\) from (2.7). Recall also from Definition 2.1 that the holes \(\mathcal{H}_A\) of a 1-connected set \(A \in \mathcal{A}\) are the 1-connected subsets of \(\Lambda \setminus A\). By Definition 2.1, each hole \(H \in \mathcal{H}_A\) is surrounded by \(A\). This implies that \(H\) does not intersect the boundary of the box \(\partial_{int} \Lambda_n\). Hence, it follows by Definition 2.4 of the boundaries that for all \(H \in \mathcal{H}_A\)
\[ \partial_{int} H = \partial_{int} \mathcal{H} \subseteq \partial_{ext} A. \tag{4.1} \]

Hence, comparing this to \(\mathcal{E}_2 = \{\exists A \in \mathcal{A}_{large} : A \not\subset \partial_{ext} A\}\) from (2.9), we obtain that
\[ \mathbb{P}(\mathcal{E}_2) \leq \mathbb{P}(\exists A \in \mathcal{A}_{large} : A \not\subset \partial_{ext} A \cup \partial_{int} H). \tag{4.2} \]

The following definition (and claim) of principal holes will ensure that the total size of the boundaries of the holes of \(A\) in (1.2) is sufficiently large compared to the combinatorial factor arising from the number of possible sets \(A\) there.

Definition 4.1 (Principal holes). Let \(A \subseteq \Lambda_n\). A hole \(H \in \mathcal{H}_A\) has type \(i \in \mathbb{N}\) if \(2^{i-1} < |H| \leq 2^i\). We write \(\mathcal{H}_A(i) \subseteq \mathcal{H}_A\) for the set of holes of \(A\) of type \(i\). A hole-type \(i\) is called principal for a set \(A\) if
\[ |\mathcal{H}_A(i)| \geq \{H \in \mathcal{H}_A(i) \mid |H| \geq 2^{i-3}2^{-2n} = h_n(i)\}. \tag{4.3} \]

Since \(|\mathcal{H}_A(i)|\) is an integer for all \(i\), the inequality \(|\mathcal{H}_A(i)| \geq [h_n(i)]\) also holds whenever the inequality in (4.3) holds. Hence, we define for \(i \in \mathbb{N}\)
\[ \mathcal{A}_{large}(i) := \{A \in \mathcal{A}_{large} : |\mathcal{H}_A(i)| \geq [h_n(i)]\}, \tag{4.4} \]
and observe that \( A \) may appear in both \( \mathcal{A}_{\text{large}}(i) \) and \( \mathcal{A}_{\text{large}}(j) \) if both type \( i \) and type \( j \) are principal for \( A \). Define the following \( \beta \)-dependent constants

\[
R_2 = R_2(\beta) := \max \left\{ \frac{\beta^{1/d}}{\sqrt{d}}, 1 \right\}, \quad i_* = i_*(\beta) := 1 + \lfloor \log_2(R_2) \rfloor. \tag{4.5}
\]

**Claim 4.2** (Large blocks have a principal hole-type). For all \( A \in \mathcal{A}_{\text{large}} \), there exists \( i_A \leq \lfloor \log_2(n) \rfloor \) such that hole-type \( i \) is principal, i.e.,

\[
\mathcal{A}_{\text{large}} \subseteq \bigcup_{i \leq \lfloor \log_2(n) \rfloor} \mathcal{A}_{\text{large}}(i). \tag{4.6}
\]

There exists a constant \( c > 0 \) such that for all \( A \subseteq \Lambda_n \), with \( i_A \) a principal hole-type for \( A \),

\[
\sum_{H \in \partial_A(i_A)} |\partial_{\text{int}} H| \geq c^{|i_A| - 2} n/|A|. \tag{4.7}
\]

Moreover, for each hole \( H \) with type \( i \geq i_* \) in \( \mathcal{A}_{\text{large}} \),

\[
|\partial_{\text{ext}} H| \geq R_2^{(d-1)/d}. \tag{4.8}
\]

**Proof.** We argue by contradiction for the first part. By definition of \( \mathcal{A}_{\text{large}} \) in (2.7), \( |A| \geq 3n/4 \), and also \( |A| \leq n/2 \). Hence, the total size of the holes is at least \( n/4 \), i.e., \( \sum_{H \in \partial A} |H| \geq n/4 \) hold. Suppose \( \mathcal{A}_{\text{large}} \) holds in the opposite direction for all \( i \geq 1 \). Since the holes are 1-connected and form together the complement of \( \bar{A} \setminus A \), it follows from the size requirement in Definition 4.1 that

\[
\left| \bigcup_{H \in \partial A} H \right| = \sum_{H \in \partial A} |H| \leq \sum_{H \in \partial A} |\partial A(i_A)|^2 = \frac{n}{8} \sum_{i \geq 1} i^2 < n/4,
\]

since the sum converges increasingly to \( \pi^2/6 = 1.64... < 2 \). This contradicts the assumption that the total size is at least \( n/4 \), so there must be at least one principal hole-type, say \( i_A \). The restriction \( i \leq \lfloor \log_2(n) \rfloor \) follows since the number of vertices in \( \Lambda_n \) is \( n \), and \( H > 2^{(\lceil \log_2(n) \rceil + 1) - 1} \) thus can never be satisfied. This shows (4.6).

We turn to (4.7). As argued before (4.1), holes do not intersect the boundary of the box \( \partial A \). So, \( \partial_{\text{ext}} H = \partial_{\text{int}} H \), and \( |\partial_{\text{ext}} H| \geq |H|^{(d-1)/d} \) by Claim 2.6 for each hole. Combined with \( |H| > 2^{i_A} n \) for all \( H \in \partial A(i_A) \) and the lower bound \( |\partial A(i_A)| \geq 2^{-i_A - 3i_A - 2} n \) in (1.3), this yields that

\[
\sum_{H \in \partial A(i_A)} |\partial_{\text{int}} H| \geq \sum_{H \in \partial A(i_A)} |H|^{(d-1)/d} \geq 2^{i_A(d-1)/d} \cdot 2^{(d-1)/d} |\partial A(i_A)| \geq 2^{-i_A - 3i_A - 2} n \geq R_2^{(d-1)/d} n,
\]

for some constant \( c > 0 \). Lastly, we prove (4.8). Using again that \( \partial_{\text{ext}} H \geq |H|^{(d-1)/d} \) for each hole, we obtain for any hole \( H \) with type \( i \geq i_* \),

\[
|\partial_{\text{ext}} H| \geq |H|^{(d-1)/d} > 2^{i_* - 1}(d-1)/d \geq (2^{\log_2(R_2)})^{(d-1)/d} = R_2^{(d-1)/d}.
\]

This finishes the proof of the claim. \( \square \)

We use a union bound on (1.6) first in (1.2) (with the convention that the empty sum from 1 to \( i_* \) is 0). We arrive at

\[
P(E_2) \leq \sum_{i=1}^{i_* - 1} P(\exists A \in \mathcal{A}_{\text{large}}(i) : A \nsubseteq A \cup \partial_{\text{int}} H) + \sum_{i=i_*}^{\lfloor \log_2(n) \rfloor} P(\exists A \in \mathcal{A}_{\text{large}}(i) : A \nsubseteq A \cup \partial_{\text{int}} H). \tag{4.9}
\]

We will now bound these two sums, the first one corresponding to small principal hole types, the second one corresponding to large principal hole types.
Excluding small principal hole types. We bound the two sums on the right-hand side of (4.11) separately and start with the first one.

**Claim 4.3** (Sets with small principal hole-types are unlikely components). Let $\mathcal{G}_n$ be long-range percolation on $\Lambda_n$ as in Definition 2.4 with $d \geq 2$, $\alpha > 1$, and $i_*(\beta)$ from (1.5). Then there exists $\delta > 0$ such that for all $\beta \geq 1$

\[
\text{Err}_{\text{small}} := \sum_{i=1}^{i_*-1} \mathbb{P}(\exists A \in \mathcal{A}_{\text{large}}(i) : A \not= \mathcal{G}_n \cup H_{\mathcal{B}_A} \partial_{\text{int}} H) \leq i_* 2^n \exp \left( - n \log \left( \frac{1}{1-p} \right) \beta^{(d-1)/d} / (1 + \log_2(\beta^{1/d}))^2 \right). \tag{4.11}
\]

The claim shows that the probability that large sets with small principal hole types appear as a component of $\mathcal{G}_n$, decays exponentially in $n$ whenever $\beta$ is sufficiently large or $p$ sufficiently close to 1.

**Proof.** We may assume that $i_* > 0$ in (4.4.3), since otherwise the sum would be empty and the bound holds trivially. We start estimating a single summand on the left-hand side of (4.11). Consider some $A \in \mathcal{A}_{\text{large}}(i)$. By Definition 1.1

\[
h_A := |\delta A| \geq |\delta A(i)| \geq h_n(i) = 2^{-i-3}n^2.
\]

We now find potential edges that all must be absent in order for the event \{ $A \not= \mathcal{G}_n \cup H_{\mathcal{B}_A} \partial_{\text{int}} H$ \} in (4.11) to occur. By (4.4) in Claim 2.6, since $i$ is a principal hole-type of $A \in \mathcal{A}_{\text{large}}(i)$

\[
|H_{\mathcal{B}_A} \partial_{\text{int}} H| = \sum_{H \in \mathcal{B}_A} |\partial_{\text{int}} H| \geq (i_* - 2)^{-i-3} n =: \ell_i. \tag{4.12}
\]

We now obtain a lower bound on $|A|$ using the isoperimetric inequality of $\mathbb{R}^d$ in Claim 2.6. By definition of $\partial_{\text{int}} A$ in Definition 2.4, and since $\partial_{\text{int}} A \subseteq \partial \tilde{A}$ by Claim 2.6, it follows from Claim 2.6 applied to $\tilde{A}$ that for all $A \in \mathcal{A}_{\text{large}}$,

\[
|A| \geq |\partial_{\text{int}} A| \geq |\partial \tilde{A}| \geq |A|(d-1)/d \geq (3/4)^{(d-1)/d} n^{(d-1)/d}.
\]

Take now a vertex $x \in \partial_{\text{int}} H \subseteq \partial \tilde{A}$ and recall $R_2(\beta)$ from (1.3). Then, since $|A|$ diverges with $n$ and $A$ is 1-connected, whenever $n$ is sufficiently large compared to $\beta$, \n
\[
|\{ y \in A : \| y - x \| \leq R_2(\beta) \}| \geq R_2(\beta), \quad \forall x \in \partial_{\text{int}} H.
\]

Hence, by (4.12),

\[
|\{ x, y \} : x, y \in \partial_{\text{int}} H, y \in A : \| y - x \| \leq R_2(\beta) \} | \geq R_2 \cdot |\partial_{\text{int}} H| \geq R_2 \ell_i.
\]

These edges must be all absent in order for \{ $A \not= \mathcal{G}_n \cup H_{\mathcal{B}_A} \partial_{\text{int}} H$ \} to occur for $A \in \mathcal{A}_{\text{large}}(i)$ in (4.11). The connection probability in (1.2) ensures that two vertices within distance $R_2(\beta)$ are connected with probability $p$. Hence using a union bound and then the independence of edges, we obtain

\[
\mathbb{P}(\exists A \in \mathcal{A}_{\text{large}}(i) : A \not= \mathcal{G}_n \cup H_{\mathcal{B}_A} \partial_{\text{int}} H) \leq \sum_{A \in \mathcal{A}_{\text{large}}(i)} \mathbb{P}(A \not= \mathcal{G}_n \cup H_{\mathcal{B}_A} \partial_{\text{int}} H) \leq \sum_{A \in \mathcal{A}_{\text{large}}(i)} (1 - p)^{R_2 \ell_i} \leq 2^n (1 - p)^{R_2 \ell_i},
\]

where we used that $\mathcal{A}_{\text{large}}(i)$ counts subsets of $\Lambda_n$, and the number of subsets of $\Lambda_n$ is at most $2^n$. This bounds a single summand in (4.11). To evaluate the sum, recalling that $\ell_i = (i_* - 2)^{-i-3} n$ from (1.2), we have

\[
\text{Err}_{\text{small}} \leq \sum_{i=1}^{i_*-1} 2^n (1 - p)^{R_2 \ell_i} = 2^n \sum_{i=1}^{i_*-1} \exp \left( - n \log \left( \frac{1}{1-p} \right) R_2^{(d-1)/d} / (1 + \log_2(\beta^{1/d}))^2 \right) \leq i_* 2^n \exp \left( - n \log \left( \frac{1}{1-p} \right) R_2^{(d-1)/d} / (1 + \log_2(\beta^{1/d}))^2 \right),
\]

where for the last inequality we used that for all $i \leq i_* - 1 = \lfloor \log_2(R_2) \rfloor$, we have that $2^{-i/d} \geq 2^{-i/d} / (\log_2(R_2) + 1)^2$. Recalling that $R_2 = \max\{\lceil \beta^{1/d} \rceil, 1\}$ from (1.3), the statement in (4.11) follows by changing the constant factor in the exponent to obtain (4.3). \qed
Excluding large principal holes. We turn to the second sum in (4.10).

Claim 4.4 (Sets with large principal hole-types are unlikely components). Let $\mathcal{G}_n$ be long-range percolation on $\Lambda_n$ as in Definition (1.1) with $d \geq 2$, $\alpha > 1$, and $i_*$ from (1.4). Then there exists $c_0 > 0$ such that

$$
\text{Err}_{\text{large}} := \sum_{i = i_*}^{\lfloor \log_2(n) \rfloor} \mathbb{P} \left( \exists A \in A_{\text{large}}(i) : A \not\ni \mathcal{G}_n \cup H \in \mathfrak{H}_\Lambda, \partial \mathcal{H} \right) \leq \exp \left( -c_0 \log \left( \frac{1}{p} \right) \beta^{(d-1)/d} \log^{-2} (n^2) n^{(d-1)/d} \right)
$$

(4.13)

when $\beta^{(d-1)/d}$ is sufficiently large or $p$ is sufficiently close to one.

Proof. Similarly to the small principal hole-types, we will find enough potential edges that all must be absent in order for the events on the left-hand-side in (4.13) to occur. A fixed ordering $L$ of vertices in $\Lambda_n$ so that $x_1 < x_2 < \cdots < x_n$ with respect to this ordering (e.g., the lexicographic ordering). For a block $A \in A_{\text{large}}(i)$ (which has at least $[h_n(i)]$ holes of type $i$ by (4.3)), we order its holes $\mathfrak{H}_A$ in such a way that the holes of type $i$ are $H_A^{(i)} \ldots , H_A^{(i)}$, and that for all $r < s \leq [\mathfrak{H}_A(i)]$ the vertices smallest in the ordering within $H_A^{(r)}$ and $H_A^{(s)}$ — say $x_r \in H_A^{(r)}$ and $x_s \in H_A^{(s)}$ — satisfy $x_r < x_s$. We obtain, when excluding edges from $A$ towards only its first $[h_n(i)]$ holes of type $i$

$$
\mathbb{P} \left( \exists A \in A_{\text{large}}(i) : A \not\ni \mathcal{G}_n \cup H \in \mathfrak{H}_\Lambda, \partial \mathcal{H} \right) \leq \mathbb{P} \left( \exists A \in A_{\text{large}}(i) : A \not\ni \mathcal{G}_n \cup \partial \mathcal{H}_A^{(i)} \right)
$$

(4.14)

where to get the second row we only look at edges emanating from $A$ that are on the exterior boundaries of the holes. This is an upper bound since $\partial \mathcal{H}_A^{(i)} \subseteq A$ by (1.1) for all $j \leq [h_n(i)]$. If for two blocks $A, A' \in A_{\text{large}}(i)$, the first $[h_n(i)]$ holes coincide, also the exterior boundaries of these first $[h_n(i)]$ holes coincide, and the event in (1.13) excludes the exact same edges. So, a simple union bound over $A$ in (4.13) would overcount the non-presence of those edges too many times. Instead, we carry out a union bound over all possible lists of the first $[h_n(i)]$ holes. To this end, we consider for all $A \in A_{\text{large}}(i)$ the following:

$$
D(A) := \Lambda_n \setminus \bigcup_{j \leq [h_n(i)]} \mathfrak{H}_A^{(i)} , \quad D(i) := \{ D : \exists A \in A_{\text{large}}(i), D = D(A) \}.
$$

(4.16)

Since the set $D(A)$ shares the first $[h_n(i)]$ holes with $A$,

$$
\bigcup_{j \leq [h_n(i)]} \partial \mathfrak{H}_A^{(i)} \not\ni \mathcal{G}_n \cup \partial \mathfrak{H}_A^{(i)} = \bigcup_{j \leq [h_n(i)]} \partial \mathfrak{H}_A^{(i)} \not\ni \mathcal{G}_n \cup \partial \mathfrak{H}_A^{(i)}
$$

Hence, in (1.15) we can group the blocks in $A_{\text{large}}(i)$ that all map to the same $D \in D(i)$, and obtain that

$$
\mathbb{P} \left( \exists A \in A_{\text{large}}(i) : \partial \mathfrak{H}_A^{(i)} \not\ni \mathcal{G}_n \cup \partial \mathfrak{H}_A^{(i)} \right)
$$

$$
= \mathbb{P} \left( \exists D \in D(i) : \partial \mathfrak{H}_A^{(i)} \not\ni \mathcal{G}_n \cup \partial \mathfrak{H}_A^{(i)} \right)
$$

$$
\leq \sum_{D \in D(i)} \mathbb{P} \left( \partial \mathfrak{H}_A^{(i)} \not\ni \mathcal{G}_n \cup \partial \mathfrak{H}_A^{(i)} \right).
$$

(4.17)

We combine the following three observations to bound a single summand in the last row. First, each vertex $x \in \partial \mathfrak{H}_A^{(i)}$ is at distance one from at least one vertex $y_x \in \partial \mathfrak{H}_A^{(i)} \subseteq D \subseteq \Lambda_n$ (by definition, a hole $H_A^{(i)}$ does not intersect $\partial \mathfrak{H}_A \Lambda_n$). Second, for all $j \leq [h_n(i)]$, $|\partial \mathfrak{H}_A^{(i)}| \geq R_2^{(d-1)/d} \beta^{(d-1)/d}$ by (3.8) and the fact that $i \geq i_*$. Third, the exterior boundary of a hole $\partial \mathfrak{H}_A^{(i)}$ is $\ast$-connected by Claim (2.9 (v)).

Hence, for each vertex $x \in \partial \mathfrak{H}_A^{(i)}$, starting from $y_x \in \partial \mathfrak{H}_A^{(i)}$, one can find a $\ast$-connected set of vertices $B_x \subseteq \partial \mathfrak{H}_A^{(i)}$ that satisfies

$$
|B_x| \geq R_2^{(d-1)/d} , \quad \text{and} \quad \forall z \in B_x : \|x - z\| \leq \sqrt{d} R_2^{(d-1)/d}.
$$

Then, the edges $\{ \{x, z\} : x \in \partial \mathfrak{H}_A^{(i)}, z \in B_x \}$ all need to be absent for the event in (4.17) to occur, and the distance bound on $\|x - z\|$ ensures, using (1.2) and that $R_2 = \max \{ \beta^{1/d} \sqrt{d}, 1 \}$ in (3.9), that
all these edges are present with probability $p$ whenever $\beta \geq 1$. Combining then (4.14), (4.15) with (4.17), it follows by the independence of edges in $G_n$ that
\[
\mathbb{P}(\exists A \in A_{\text{large}}(i) : A \not\subset G_n \cup H_{\partial A} \partial H) \leq \sum_{D \in D(i)} \prod_{j \leq [h_n(i)]} (1 - p)^{|\partial H_D(j) \cap R_2^{(d-1)/d}|}.
\]
We will now encode the holes of $D \in D(i)$ similar to the encoding of the blocks in Section 3. We write $x_D := (x_1, \ldots, x_{[h_n(i)]})$ for the vertices with the smallest label in the $L$-ordering within the respective holes $H_D^{(j)}, \ldots, H_D^{([h_n(i)])}$. Let us write $\Lambda_{h_n(i)}$ for the vectors $x \in \Lambda_{h_n(i)}$ with $x_r < L x_s$ for all $r < s$. By the initial ordering of the holes above (4.13), $x_D \in \Lambda_{h_n(i)}$. Let then $m_j := |\partial H_D^{(j)}|$ for all $j \leq [h_n(i)]$, and write $m_D := (m_1, \ldots, m_{[h_n(i)]})$. Define then for all $x \in \Lambda_{h_n(i)}$, $m \in \mathbb{N}_{[h_n(i)]}$:
\[
D(i, x, m) := \{ D \in D(i) : x_D = x, m_D = m \}.
\]
The set $D \in D(i)$ has the first $[h_n(i)]$ holes of some $A \in A_{\text{large}}(i)$ where $D(A) = D$, and for that $A$, hole-type $i$ was principal in terms of Definition 4.1 and 4.4. Definition 4.4 readily implies that Claim 4.2 is applicable to already the first $[h_n(i)]$ many holes of $A$, and in turn, $D$. Hence, the total interior boundary size $m := \sum_{j=1}^{[h_n(i)]} m_j$ satisfies that $m \geq (4.14)^{-2} 2^{-i/d} n$. So for $m \geq (4.14)^{-2} 2^{-i/d} n$ we introduce the possible boundary-length vectors with total size $m$:
\[
\mathcal{M}_i(m) := \{ m \in \mathbb{N}_{[h_n(i)]} : m_1 + \ldots + m_{[h_n(i)]} = m \}, \tag{4.18}
\]
Returning to (4.17), we decompose the summation on the right hand side as follows:
\[
\mathbb{P}(\exists A \in A_{\text{large}}(i) : A \not\subset G_n \cup H_{\partial A} \partial H) \leq \sum_{m \geq (4.14)^{-2} 2^{-i/d} n} m \mathcal{M}_i(m) \sum_{x \in \Lambda_{h_n(i)}^1} \prod_{j \leq [h_n(i)]} (1 - p)^{|\partial H_D^{(j)} \cap R_2^{(d-1)/d}|} = \sum_{m \geq (4.14)^{-2} 2^{-i/d} n} m \exp(c_{\text{pe}} m j) \prod_{j \leq [h_n(i)]} 1.
\]
Now we evaluate the number of terms of the last three summations. Each block $D \in D(i, x, m)$ is uniquely characterized by its $[h_n(i)]$ holes by (4.19). Having fixed the vectors $x$ and $m$, we apply Lemma 2.7 to each hole $H_D^{(j)}$ with $H_D^{(j)} \ni x_j$ and $|\partial H_D^{(j)}| = m_j$ to count the size of $D(i, x, m)$. The lemma is applicable since for each hole $H$, we have $H = H^{\bar{1}}$ by Claim 2.5(v). Hence, there are at most $\exp(c_{\text{pe}} m_j)$ possible holes of interior boundary size $m_j$ containing $x_j$. So, for all $x \in \Lambda_{h_n(i)}$ and $m \in \mathcal{M}_i(m)$,
\[
|\mathcal{D}(i, x, m)| \leq \prod_{j \leq [h_n(i)]} \exp(c_{\text{pe}} m j) = \exp(c_{\text{pe}} m).
\]
Moreover, by (4.18), $m \geq [h_n(i)]$, and so $|\mathcal{M}_i(m)| \leq \left( \frac{m + [h_n(i)]}{m} \right) \leq 2^m \leq e^{2m}$. Next, since the vertices in $x$ are ordered, there are at most $\left( \frac{n}{m} \right)^{[h_n(i)]}$ many choices for the vector $x \in \Lambda_{[h_n(i)]}$. Using these bounds in (4.19) and evaluating the geometric sum in $m$ it follows for some $C > 0$ that
\[
\mathbb{P}(\exists A \in A_{\text{large}}(i) : A \not\subset G_n \cup H_{\partial A} \partial H) \leq \left( \frac{n}{[h_n(i)]} \right) \sum_{m \geq (4.14)^{-2} 2^{-i/d} n} (1 - p)^{m R_2^{(d-1)/d}} \exp((c_{\text{pe}} + 2) m) \leq C \left( \frac{n}{[h_n(i)]} \right) \exp((c_{\text{pe}} + 2 - \log(1/p) R_2^{(d-1)/d}) \left( (4.14)^{-2} 2^{-i/d} n \right), \tag{4.20}
\]
whenever $(1 - p)^{R_2^{(d-1)/d}} e^{c_{\text{pe}}+2} < 1$ (see $R_2(\beta)$ in (4.5)). We recall from (4.3) that $h_n(i) = 2^{-i-3} 2^{-i} n$, and using $\left( \frac{n}{n} \right) \leq (e \cdot n/h)^h$ it follows
\[
\left( \frac{n}{[h_n(i)]} \right) \leq (e^{2i+3} i^{2i+3} n^{i-3} + 1) \leq \exp((i + 2 \log(i))(i^{-2i-3} n + 1)), \tag{4.21}
\]
where we also used that $2 < e$ to obtain this bound in the right-hand-side. Using this bound in the right-hand side of (4.20), we may compare the exponents. Let $i_0$ be the smallest $i \in \mathbb{N}$ such that for all $n \geq 1$ and all $i \in [i_0, \lceil \log_2(n) \rceil]$, 

$$(i + 4 + 2 \log(i))(i^{-2}2^{-i-3}n + 1) < \frac{1}{4.2}i^{-2}2^{-i/d}n.$$ 

Then for $i \geq i_0$,

$$\mathbb{P}(\exists A \in A_{\text{large}}(i) : A \not\supseteq A_{\text{large}}(i)) \leq C \exp\left(\left(c_{\text{pe}}i + 3 - \log\left(\frac{1}{1-p}\right)R_2^{(d-1)/d}\right)\frac{i}{4.2}i^{-2}2^{-i/d}n\right). \quad (4.22)$$

Using that $i_0$ is a constant that only depends on $d$, (comparing the coefficients of (4.20) to (4.21)) we require that $R_2^{(d-1)/d} \log\left(\frac{1}{1-p}\right)$ is large enough so that for all $i \leq i_0$,

$$(c_{\text{pe}}i + 2)i^{-2}2^{-i/d} + (i + 4 + 2 \log(i))(i^{-2}2^{-i-3}n + 1) < \frac{1}{4.2}i^{-2}2^{-i/d}n.$$ 

In this case we obtain that for all $i \leq i_0$,

$$\mathbb{P}(\exists A \in A_{\text{large}}(i) : A \not\supseteq A_{\text{large}}(i)) \leq C \exp\left(- \frac{1}{2} \log\left(\frac{1}{1-p}\right)R_2^{(d-1)/d}\right)\frac{i}{4.2}i^{-2}2^{-i/d}n. \quad (4.23)$$

We recall that $R_2 = \Theta(\beta^{(d-1)/d})$ from (4.3), so by combining (4.22) and (4.23) it follows for some $c' > 0$ that, whenever $\beta^{(d-1)/d} \log\left(\frac{1}{1-p}\right)$ is sufficiently large,

$$\mathbb{P}(\exists A \in A_{\text{large}}(i) : \partial_{\text{int}}A \not\supseteq \bigcup_{i \in A} \partial_{\text{int}}A_{(i)}) \leq C \exp\left(- c' \log\left(\frac{1}{1-p}\right)\beta^{(d-1)/d}i^{-2}2^{-i/d}n\right).$$

We apply this bound to all summands in (1.13), use that the terms are increasing in $i$, that the last term is $i = \lceil \log_2(n) \rceil$ (so $2^{-i/d}i^{-2}n \geq 2^{-1/d}(\log_2(n) + 1)^{-2}n^{(d-1)/d}$ for all $i$), and obtain for some $c, d > 0$ and $n$ sufficiently large

$$\sum_{i=i_0+1}^{\lceil \log_2(n) \rceil} \mathbb{P}(\exists A \in A_{\text{large}}(i) : \partial_{\text{int}}A \not\supseteq \bigcup_{i \in A} \partial_{\text{int}}A_{(i)}) \leq \exp\left(- c \log\left(\frac{1}{1-p}\right)\beta^{(d-1)/d}(\log(n))^{-2}n^{(d-1)/d}\right).$$

This finishes the proof of Lemma 2.10. \hfill \Box

We are ready to give the final proofs of the section.

Proof of Lemma 2.7a We recall the bound (1.10) on the probability of the event $E_2 = \{\exists A \in A_{\text{large}} : A \not\supseteq \bigcup_{i \in A} \partial_{\text{int}}A\}$, splitting it into two sums: one sum (1.10) for small principal holes, and one sum (4.10) for large principal holes. As an immediate corollary of Claims 4.3 and 4.4 we obtain for the constant $i^*$(4.5) that

$$\mathbb{P}(E_2) \leq i^*2^{n} \exp\left(- n \left(\frac{1}{1-p}\right)\beta^{(d-1)/d}/(1 + \log_2(\beta^{(d-1)/d}))^2\right) + \exp\left(- \frac{n^{(d-1)/d}}{\log\left(\frac{1}{1-p}\right)}\beta^{(d-1)/d}\right).$$

If $\log\left(\frac{1}{1-p}\right)\beta^{(d-1)/d}$ is sufficiently large, the first term decays exponentially in $n$. Under this assumption, the second term on the right-hand side dominates the expression. The existence of $\beta^*(p, d, \alpha)$ for example is immediate. Similarly, whenever $p$ is sufficiently close to 1, then $\log\left(\frac{1}{1-p}\right)\beta^{(d-1)/d}$ is already sufficiently large for $\beta = 1$, and hence $\beta^*(p, d, \alpha) = 1$ can be set in this interval, finishing the proof of the final statement. \hfill \Box

Proof of Proposition 1.5 The proof is immediate from Claim 2.8 and Lemmata 2.4 and 2.10. The condition $n(\log(n))^{-2d/(d-1)} \geq k$ implies that $k^{(d-1)/d} \leq n^{(d-1)/d}/(\log(n))^2$, so that the error bound from (2.15) dominates the error bound from (2.16). The coefficient of $-k^{(d-1)/d}$ in the exponent can be made at least 1 by choosing either $\beta$ sufficiently large or $p$ sufficiently close to 1 and $\beta \geq 1$ so that the total coefficient of $\log(\frac{1}{1-p})\beta^{1/d} + \log(\frac{1}{1-p})\beta^{(d-1)/d}$ is already sufficiently close to 1. \hfill \Box

5. Proof of Theorem 1.2

We will verify the statements in Theorem 1.2 based on Proposition 1.5.
Upper bounds. The upper bound on the second-largest component in \([1.3]\) follows immediately from Proposition \([1.6]\) by substituting \(k = A(\log(n))^{d/(d-1)}\) for some large constant \(A = A(\delta)\). For the upper bound on the cluster-size decay and the lower bounds, we cite two statements from our paper \([22]\) which considers a more general class of percolation models. The paper \([22]\) considers models where vertices have associated vertex marks, and the connection probability \([1.2]\) contains an additional factor to \(\beta\) in the numerator on the right-hand side of \([1.2]\), that depends on these vertex-marks. The model long-range percolation in Definition \([1.1]\) hence forms a subclass of the model in \([22]\) in which the vertex set is \(\mathbb{Z}^d\) and all the vertex marks are identical to 1. Due to this, conditions that regard vertices with high vertex-marks in \([22]\) are automatically satisfied for the long-range percolation model in Definition \([1.1]\). As a result, we rephrase the results of \([22]\) to the setting of long-range percolation of Definition \([1.1]\) by setting all vertex marks identical to 1 in \([22]\).

**Proposition 5.1** (Prerequisites for the upper bound \([22\) Proposition 6.1]). Consider supercritical long-range percolation with parameters \(\alpha > 1\), and \(d \in \mathbb{N}\). Assume that there exist \(\zeta, c, c' > 0\) and a function \(g(k) = O(k^{1+c'})\) such that for all \(n, k\) sufficiently large, whenever \(n \geq g(k)\), it holds that

\[
\mathbb{P}(|C_n^{(2)}| \geq k) \leq n^c \exp\left(-ck^\zeta\right),
\]

(5.1)

\[
\mathbb{P}(|C_n^{(1)}| \leq n^c) \leq n^{-1-c}.
\]

(5.2)

Then there exists a constant \(A > 0\) such that for all \(n, k\), sufficiently large such that \(g(k) \leq n \leq \infty\)

\[
\mathbb{P}(\{|C_n(0)| \geq k, 0 \not\in C_n^{(1)}\} \leq \exp\left(-\frac{1}{A}k^\zeta\right),
\]

and

\[
\frac{|C_n^{(1)}|}{n} \xrightarrow{p} \mathbb{P}(|C(0)| = \infty), \quad \text{as } n \to \infty.
\]

We have just proved the prerequisite \([5.1]\) for our case in Proposition \([1.6]\) with \(\zeta = (d-1)/d\), \(c = 1\) and \(c' = 2\). The other prerequisite \([5.2]\) is a consequence of the following lemma ((5.3) below in particular). The second statement of the lemma, \([5.4]\) will be needed for the lower bound of Theorem \([1.2]\) shortly.

**Lemma 5.2.** Consider long-range percolation in Definition \([1.1]\) with \(\alpha > 1\) and \(d \geq 2\). For all \(p \in (0, 1)\), there exists \(\beta_* = \beta_*(p, d, \alpha) > 0\) such that for all \(\beta \geq \beta_*\) there exists \(\rho > 0\) such that for \(n\) sufficiently large

\[
\mathbb{P}(\{|C_n^{(1)}| \geq \rho n\}) \geq 1 - \exp\left(-\rho n^{(d-1)/d}\right),
\]

(5.3)

\[
\mathbb{P}(\{|C_n(0)| \geq \rho n\}) \geq \rho.
\]

(5.4)

For each \(d \geq 2\), there exists \(p_d < 1\) such that \(\beta_* \leq 1\) for all \(p \in (p_d, 1)\).

**Proof.** We start with showing \([5.3]\). If \(p\) is sufficiently close to one, then we can do the following. Taking \(G_n\) as a realization of LRF in \(\Lambda_n\), and retaining only the edges of \(G_n\) that are between nearest neighbor vertices in \(\mathbb{Z}^d_n\), we obtain the classical iid nearest-neighbor Bernoulli percolation in \(\Lambda_n\). Denote this graph by \(G_n^{(nn)}\) and its largest component by \(C_n^{(1)}(G_n^{(nn)})\). Then since \(\mathcal{E}(G_n^{(nn)}) \subseteq \mathcal{E}(G_n)\), for the sizes of the largest components it holds that \(|C_n^{(1)}| \geq |C_n^{(1)}(G_n^{(nn)})|\). Since \(p\) is sufficiently close to 1, the surface order large deviation result of \([12\) Theorem 1.1\] applies to \(C_n^{(1)}(G_n^{(nn)})\), and \([5.3]\) immediately follows.

Assume now that \(p\) is not sufficiently close to 1 for the nearest-neighbor subgraph to ensure the required result. Let

\[
m(n) := \left\lceil \frac{n^{1/d}}{\beta^{1/d}(2\sqrt{d})^d} \right\rceil^d.
\]

(5.5)

Partition \(\Lambda_n\) into \(m(n)\) identical boxes \(Q_1, Q_2, Q_{m(n)}\), each of sidelength \(r(n) = n^{1/d}/m(n)^{1/d} \leq \beta^{1/d}/(2\sqrt{d})\). Denote \(n_i := |\mathbb{Z}^d \cap Q_i|\) the number of vertices in box \(Q_i\). Then, for \(n\) large enough, for all \(i \leq m(n)\) it holds that \(n_i \in [C(\beta), 4C(\beta)]\), where for all \(n\) sufficiently large

\[
C(\beta) := \text{Vol}(Q_1)/2 = r(n)^d/2 \in \left[\frac{\beta}{4(2\sqrt{d})^d}, \frac{\beta}{2(2\sqrt{d})^d}\right].
\]

(5.6)
Let us say that \( Q_i, Q_j \) are adjacent boxes if they share a \((d-1)\)-dimensional face. Since the diameter of each box is at most \( \beta^{1/d}/2 \), by (5.2), if \( Q_i, Q_j \) are adjacent boxes,
\[
\mathbb{P}(\{x, y\} \in \mathcal{E}(G_n)) = p \quad \text{for all } x \in Q_i \cap \mathbb{Z}^d, y \in Q_j \cap \mathbb{Z}^d,
\]
(5.7) independently of other edges. The same is true when \( x, y \in Q_i \cap \mathbb{Z}^d \) both.

Let us denote the subgraph of \( G_n \) induced by the vertices in the box \( Q_i \) by \( G_n(Q_i) \). \( G_n(Q_i) \) then stochastically dominates an Erdős-Rényi random graph with \( n_i \in [C(\beta), 4C(\beta)] \) vertices and edge probability \( p \). Since \( p \) is constant, for any fixed \( \varepsilon > 0 \), by choosing \( \beta \) (and hence also \( C(\beta) \)) large enough depending on \( \varepsilon \), the probability that \( G_n(Q_i) \) is connected is at least \( 1 - \varepsilon \) by [1]. Further, the graphs \( (G_n(Q_i))_{i \leq m(n)} \) are independent since they are induced subgraphs of long-range percolation on vertices in disjoint boxes, and edges are present independently in \( G_n \) by Definition [1.1].

We define an auxiliary graph \( G \). First we define the vertex set. Every box \( Q_i \) corresponds to a vertex \( v_i \), for each \( i \leq m(n) \), and two vertices \( v_i, v_j \) in \( G \) are adjacent if the corresponding boxes \( Q_i, Q_j \) are adjacent, i.e., they share a \((d-1)\)-dimensional face. Similarly, one can define the 1-distance between any two vertices \( v_i, v_j \) via adjacent vertices. Hence, the vertices of \( G \) then form a box \( \Lambda_m(n) \) of volume \( m(n) \) of \( \mathbb{Z}^d_1 \). This we call the re-normalised lattice.

Now we define the edge-set of \( G \). We declare a vertex \( v_i \) of \( G \) active when \( G_n(Q_i) \) is connected. Edges of \( G \) will be only present between active (and adjacent) vertices. Assuming that two vertices \( v_i, v_j \) in \( G \) are adjacent and both active, we declare the edge between \( v_i \) and \( v_j \) open, equivalently, present in \( \mathcal{E}(G) \), if there exist vertices \( x \in Q_i \cap \mathbb{Z}^d, y \in Q_j \cap \mathbb{Z}^d \) with the edge \( \{u, v\} \in \mathcal{E}(G_n) \). Conditional on \( v_i, v_j \) being active, by (5.4), the nearest-neighbor edge between \( v_i \) and \( v_j \) is open with probability \( 1 - (1 - p)^{n_i n_j} \geq 1 - (1 - p)^{C(\beta)^2} \geq 1 - \varepsilon \), where the last inequality holds for arbitrarily small \( \varepsilon > 0 \) by making \( C = C(\beta) \) in (5.6) large enough. Different edges of \( \mathcal{E}(G) \) are present conditionally independently given that the end-vertices are active.

Let us denote by \( H \) the induced graph obtained from \( G \) on active vertices and open edges \( \mathcal{E}(G) \). By the observation above, the vertices \( (v_i)_{i \leq m(n)} \) form a box \( \Lambda_m(n) \) of volume \( m(n) \) of \( \mathbb{Z}^d_1 \). Then \( H \) stochastically dominates a site-bond percolation of \( \mathbb{Z}^d_1 \) in \( \Lambda_m(n) \).

More precisely, since vertices of \( G \) are active independently with probability at least \( 1 - \varepsilon \), and edges of \( G \) between adjacent vertices are present conditionally independently again with probability at least \( 1 - \varepsilon \), each edge in the renormalised lattice \( \Lambda_m(n) \) is open with probability at least \((1 - \varepsilon)^3 \). The model is 1-dependent, since the state of any edge \( \{v_i, v_j\} \) of \( H \) depends only on edges sharing at least one vertex with \( \{v_i, v_j\} \).

Since \( \varepsilon \) can be chosen arbitrarily small, by [28] Theorem 0.0], the graph \( H \) therefore stochastically dominates iid nearest-neighbor bond percolation \( G^* \) on \( \Lambda_m(n) \) with parameter \( p^* \) that can also be made arbitrarily close to 1. Hence for the sizes of the largest connected components \( |C_n^{(1)}(H)| \geq |C_n^{(1)}(G^*)| \) holds. Thus, [24] Theorem 1.1] applies to \( |C_n^{(1)}(G^*)| \), and so for some \( c(\beta) > 0 \) we obtain that using (5.5)
\[
\mathbb{P}(\{|C_n^{(1)}(H)| \geq \rho m(n)\}) \geq e^{-c m(n)(d-1)/d} \geq 1 - e^{-c(\beta)n(d-1)/d} \quad (5.8)
\]
Since in each box \( Q_i \) that an active \( v_i \) in \( G \) corresponds to contains at least \( C(\beta) \) vertices and the graphs \( G(Q_i) \) are connected, it holds deterministically that \( |C_n^{(1)}(G_n)| \geq C(\beta)|C_n^{(1)}(H)| \). This, combined with (5.8) implies (5.3).

We turn now to prove (5.4). Consider a smaller box \( \Lambda_2^{-d}n \). Define then
\[
Z_{\ell} := \sum_{x \in \Lambda_2^{-d}n} 1_{\{|C_2^{-d}n(x)| \geq \ell\}}.
\]
We argue that \( \{C_2^{(1)} \geq \ell\} \subseteq \{Z_{\ell} \geq \ell\} \). Indeed, if the largest component is at least of size \( \ell \) then in \( Z_{\ell} \) at least \( \ell \) many indicators are 1. Then applying a Markov’s inequality with \( \ell = \rho 2^{-d}n \) followed by a union bound yields that
\[
\mathbb{P}(C_2^{(1)} \geq \rho 2^{-d}n) \leq \mathbb{P}(Z_{\rho 2^{-d}n} \geq \rho 2^{-d}n) \leq \frac{E[Z_{\rho 2^{-d}n}]}{\rho 2^{-d}n} \leq \frac{1}{\rho 2^{-d}n} \sum_{x \in \Lambda_2^{-d}n} \mathbb{P}(\{|C_2^{-d}n(x)| \geq \rho 2^{-d}n\}).
\]
If for all \( x \in \Lambda_2^{-d}n \) it would hold that \( \mathbb{P}(\{|C_2^{-d}n(x)| \geq \rho 2^{-d}n\}) \leq \rho/2 \), then the right hand side would be at most 1/2. This would then contradict (5.3) for \( 2^{-d}n \) in place of \( n \).
Hence, there must exist $x \in \Lambda_{2-\epsilon_{n}}$ such that $\mathbb{P}(|C_{2-\epsilon_{n}}(x)| \geq \rho 2^{-d_{n}}) \geq \rho / 2$. Let $x \in \Lambda_{2-\epsilon_{n}}$ be such a vertex. Then, by the translation invariance of the infinite model $G_{\infty}$, looking at the component of the origin $C_{2-d_{n}}^{(-x)}(0)$ inside the box $\Lambda_{2-\epsilon_{n}}(-x)$, it holds that

$$\mathbb{P}(|C_{2-d_{n}}^{(-x)}(0)| \geq \rho 2^{-d_{n}}) = \mathbb{P}(|C_{2-\epsilon_{n}}(x)| \geq \rho 2^{-d_{n}}).$$

However, the shifted box $\Lambda_{2-\epsilon_{n}}(-x) \subseteq \Lambda_{n}$ for any $x \in \Lambda_{2-\epsilon_{n}}$, and hence $C_{2-d_{n}}^{(-x)}(0) \subseteq C_{n}(0)$. Hence, we obtain

$$\mathbb{P}(|C_{n}(0)| \geq \rho 2^{-d_{n}}) \geq \mathbb{P}(|C_{2-\epsilon_{n}}(x)| \geq \rho 2^{-d_{n}}) \geq \rho / 2.$$ 

Hence, (5.4) follows by adapting the constant $\rho$.

Since both prerequisites of Proposition 5.1 are satisfied, this finishes the proof of the upper bounds of Theorem 1.2.

**Lower bounds.** For the lower bound we adapt the lower bound from [22], rephrased to the model of long-range percolation of Definition 1.1 by setting the vertex set to $\mathbb{Z}^{d}$ and all vertex marks to 1 in [22]. The lower bound of cluster-size decay and second-largest component that we are about the cite — Proposition 7.1 — requires that in a box of volume $\ell$ a linear sized (at least $\rho \ell$) giant component on vertices with marks in the interval $[1, \text{polylog} (\ell)]$ exists, with probability at least $\rho > 0$. Since in long-range percolation, all vertex marks are identical to 1, this requirement of [22] Proposition 7.1] turns into the requirement (5.9) below for LRP.

**Proposition 5.3 (Lower bound [22 Proposition 7.1]).** Consider supercritical long-range percolation with parameters $\alpha > 1 + 1/d$, $d \geq 2$, and assume that $\min\{p, p^{\beta_{1}}\} \in (0, 1)$. Assume that there exists a constant $\rho > 0$ such that for all $n$ sufficiently large,

$$\mathbb{P}(|C_{n}(0)| \geq \rho n) \geq \rho.$$  

Then there exists $A > 0$ such that for all $n \in [Ak, \infty)$,

$$\mathbb{P}(|C_{n}(0)| \geq k, 0 \notin C_{n}^{(1)}) \geq \exp \left(1 - A k^{(d-1)/d}\right).$$  

Moreover, there exists $\delta, \varepsilon > 0$, such that for all (finite) $n$ sufficiently large

$$\mathbb{P}(|C_{n}^{(2)}| \leq \varepsilon (\log (n))^{d/(d-1)}) \leq n^{-\delta}.$$  

Since Lemma 5.2 has just proved in (5.4) the requirement (5.9), and in Theorem 1.2 we have also assumed that $\min\{p, p^{\beta_{1}}\} < 1$, the lower bounds in Theorem 1.2 follow from Proposition 5.3 in particular. (5.11) implies the lower bound in (1.3), and after taking logarithm of both sides, (5.10) implies the lower bound in (1.4).

**Appendix A. Proofs of preliminary claims**

**Proof of Claim [22].** Identify a path on $\mathbb{Z}^{d}$ with its vertex set. We define an equivalence class $\sim_{C, 1}$ on the vertices of $C$, where $x \sim_{C, 1} y$ if and only if there is a 1-connected path $\pi$ consisting of vertices of $C$ that connects $x$ and $y$ (i.e., the edges of this path are not necessarily part of the edges of $G$). We then define the blocks $A_{1}, A_{2}, \ldots, A_{b}$ as the equivalence classes of $\sim_{C, 1}$. In other words, start from any vertex $x \in C$ and define its block as all vertices that $x$ is 1-connected to using paths only vertices of $C$ (but the edges of $\mathbb{Z}^{d}$), and then we iterate this over all $x \in C$, yielding the (different) blocks $A_{1}, A_{2}, \ldots, A_{b}$.

Each $A_{i}$ is 1-connected since every pair of vertices in $A_{i}$ is connected by a 1-connected path by the definition of $\sim_{C, 1}$, i.e., $A_{i}$ is a block. Further, if $i \neq j$ then $|A_{i} - A_{j}|_{1} > 1$ must hold, since otherwise there would be a 1-connected path from some $x \in A_{i}$ to some $y \in A_{j}$, and that would contradict $x \notin_{C, 1} y$. Uniqueness of this decomposition follows because $\sim_{C, 1}$ is an equivalence relation.

We show that the block graph $H_{G}(\{A_{i}\}_{i \leq b})$ is connected. Suppose otherwise. This means that there is a subset of blocks whose union is not connected to the union of all the other blocks. Suppose, we may assume that for the first $k$ blocks for some $k \in [1, b - 1]$ this happens, i.e., $(\cup_{i \leq k} A_{i}) \not\leftrightarrow_{G} (\cup_{i > k} A_{i})$. However, this contradicts that $C$ is a component of $G$.

**Proof of Claim [22].** We show that $\tilde{\partial}_{\text{int}} \bar{B} \subseteq \tilde{\partial}_{\text{int}} B$. We argue by contradiction. Assume that there exists $x \in \tilde{\partial}_{\text{int}} \bar{B} \setminus \tilde{\partial}_{\text{int}} B$. Since $\partial_{\text{int}} \bar{B} \subseteq \bar{B} = B \cup (\cup_{H \in \Delta_{B}} H)$ and $B$ is disjoint from $\bar{B} \setminus B$, there are two cases. Either $x \in B \setminus \tilde{\partial}_{\text{int}} B$ or $x \in (\cup_{H \in \Delta_{B}} H) \setminus \tilde{\partial}_{\text{int}} B$. 
For the first case assume that $x \in B \setminus \partial_{\text{int}} B$. Then all its $Z_i^d$-neighboring vertices were also in $B$ by Definition 2.4 of the interior boundary. Thus $x$ is surrounded by $B$, and hence $x \in \tilde{B}$. Similarly, the neighboring vertices are also in $\tilde{B}$, contradicting that $x \in \partial_{\text{int}} \tilde{B}$.

For the second case assume that $x \in (\cup_{H \in \mathcal{D}_n} H)$. Then $x$ was surrounded by $B$, but then all its $Z_i^d$-neighboring vertices were either a member of $B$ or surrounded by $B$, contradicting again that $x \in \partial_{\text{int}} \tilde{B}$.

We move on to part (ii). Assume that $\hat{B}_1 \cap \hat{B}_2 \neq \emptyset$, then there exists $x \in \hat{B}_1 \cap \hat{B}_2$. Since $B_1$ and $B_2$ are $Z_i^d$-disconnected, it is excluded that $x \in \hat{B}_1 \cap \hat{B}_2$. Assume $x \in (\hat{B}_2 \setminus \hat{B}_1) \cap \hat{B}_1$. Since $x$ is surrounded by $B_2$, any vertex $y \in B_1$ must be surrounded by $B_2$, since $|B_1 - B_2| \geq 2$ and $B_1$ itself is $Z_i^d$-connected. Hence, $\hat{B}_1 \subseteq \hat{B}_2$. The argument for $x \in (\hat{B}_1 \setminus \hat{B}_1) \cap \hat{B}_2$ follows analogously. Lastly, assume that there exists $x \in (\hat{B}_1 \setminus \hat{B}_1) \cap (\hat{B}_2 \setminus \hat{B}_2)$. Then there exists, for some $j \geq 1$, a path $\pi = (x, x_1, \ldots, x_j)$ on $Z_i^d$ such that $x_j \in \partial_{\text{int}} B_1 \cup \partial_{\text{int}} B_2$ and $x_{\ell} \in (\hat{B}_1 \setminus \hat{B}_1) \cap (\hat{B}_2 \setminus \hat{B}_2)$ for all $\ell < j$. Assume w.l.o.g. that $x_j \in \partial_{\text{int}} \hat{B}_1$. Since there were no vertices from $B_2$ on the path and $x$ is surrounded by $B_2$, it must follow that also $x_j$ is surrounded by $B_2$. Similar to the previous case, it follows that $\hat{B}_1 \subseteq \hat{B}_2$.

Part (iii) claims that when $\hat{B}_1 \cap \hat{B}_2 = \emptyset$, and initially $B_1, B_2$ are 1-connected, then $|\hat{B}_1 - \hat{B}_2| \geq 2$. By definition of the distance $\| \cdot \|$ between sets, we have

$$\|\hat{B}_1 - \hat{B}_2\| = \min_{x_i \in \hat{B}_1 \setminus \hat{B}_2, x_j \in \hat{B}_2 \setminus \hat{B}_1} \{ \|x_i - x_j\| : \|B_1 - B_2\|, \|B_1 - x_i\|, \|B_2 - x_j\| \}.$$  \hfill (A.1)

Each 1-connected path from $x_i \in \hat{B}_1 \setminus B_1$ to any $y \not\in \hat{B}_1$ must cross a vertex in $\partial_{\text{int}} \hat{B}_1$. Consequently,

$$\|x_i - x_j\| \geq \|x_i - \partial_{\text{int}} \hat{B}_1\| + \|\partial_{\text{int}} \hat{B}_1 - x_j\| \geq \|x_i - \partial_{\text{int}} \hat{B}_1\| + \|\partial_{\text{int}} \hat{B}_1 - x_j\| \geq 2,$$

where the second inequality follows since $\partial_{\text{int}} \hat{B}_1 \subseteq \partial_{\text{int}} B_1$ by part (i), and the third inequality since by 1-disconnectedness of $B_1$ and $B_2$ the second term on the right-hand side is at least two. The third and fourth term in the minimum in (A.1) can be bounded similarly. It follows that $B_1$ and $B_2$ are 1-connected.

The fourth statement is immediate from [12, Lemma 2.1] which states that the interior and exterior boundaries of a hole are $\bar{Z}_i$-connected. Hence, for all $A \subseteq \Lambda_n \setminus B$ surrounded by $B$, see below (2.2) in Definition 2.1. So if there were a hole $J$ inside $H$, then $J$ must be part of $B$, which would then contradict the 1-connectedness of $B$, since $\|J - \partial_{\text{ext}} H\| \geq 2$ as $H$ fully surrounds $J$. The fact that the interior and exterior boundaries of a hole are $\ast$-connected follows now from Part (iv).

\textit{Proof of Claim 2.7}. The proof is inspired by an argument by Deuschel and Pisztora [12, Proof of (A.3)]. The inequalities with $(\ast)$ in (2.5) follow by standard isoperimetric inequalities, but we will also derive them below.

We will first show the bounds for $\partial_{\text{int}}$ and $\partial_{\text{int}}$. At the end of the proof we adjust it to $\partial_{\text{ext}}$ and $\partial_{\text{ext}}$. We start by showing that there exists $\delta > 0$ such that $\partial_{\text{int}} A \geq \delta \partial_{\text{int}} A$ for all $A \subseteq \Lambda_n$ of size at most $3n/4$. Fix such a set $A \subseteq \Lambda_n$. We recall an inequality related to the isoperimetric inequality by Loomis and Whitney [20, Theorem 2]. For a set $A \subseteq \Lambda_n$, let $S_i$ denote the projection of $A$ onto the $i$-th coordinate hyperplane. That is, for a vertex with coordinates $x = (x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_d)$ we define $\pi_i x := (x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_d)$. Then

$$|A|^{d-1} \leq \prod_{i \in [d]} |S_i|.$$  \hfill (A.2)

Let $i_*$ be the dimension that contains the largest projected set $S_{i_*}$ (ties broken arbitrarily), so that as a result of (A.2),

$$|S_{i_*}| \geq |A|^{(d-1)/d}.$$  \hfill (A.3)

We abbreviate $S = S_{i_*}$, and write $\pi_{**}$ for the the $i_*$-th projection. For $s \in S$ we define the preimage of $s$ as

$$\pi_{**}^{-1} := \{ y \in A : \pi_{**}(y) = s \}$$
We now call a vertex \( s \in S \) a fiber if there is no vertex in \( \partial_{\text{int}} A \) that projects to \( s \) via \( \pi_* \). Formally we define the fibers of \( S \) as
\[
F := \{ s \in S : \exists y \in \partial_{\text{int}} A \text{ with } \pi_*(y) = s \}.
\]

The pre-image of any fiber vertex does not contain any vertex of \( \partial_{\text{int}} A \) within \( \Lambda_n \), hence, it contains a full length-\( n^{1/d} \) line \( \mathcal{L}_s \) connecting the two opposite faces of \( \Lambda_n \), with \( \pi_*(\mathcal{L}_s) = s \). This is because all vertices that share all coordinates with \( s \) except the \( i \)-th coordinate, project to \( s \) via \( \pi_* \), so \( A \) must contain all of them (the possibility of \( A \) containing none of \( \mathcal{L}_s \) is excluded by assuming \( s \in S \)), otherwise there would be a boundary vertex of \( A \) among them. Then the pre-image of any such fiber vertex intersects the box-boundary \( \tilde{\partial}_{\text{int}} \Lambda_n \) in exactly 2 vertices:
\[
|\pi_*^1(s) \cap \tilde{\partial}_{\text{int}} \Lambda_n| = |\pi_*^2(s) \cap \tilde{\partial}_{\text{int}} A| = 2, \quad \forall s \in F.
\]

The pre-image of vertices in \( F \) does not contribute to \( \partial_{\text{int}} A \) by their definition in (A.4). By the same definition, the pre-image of each vertex \( z \in S \setminus F \) contains at least one vertex in \( \partial_{\text{int}} A \). A similar argument as the one for fibers shows that the pre-image of each vertex \( z \in S \setminus F \) contains at least two vertices in \( \tilde{\partial}_{\text{int}} A \). Formally
\[
|\pi_*^1(z) \cap \partial_{\text{int}} A| \geq 1 \quad \text{and} \quad |\pi_*^2(z) \cap \tilde{\partial}_{\text{int}} A| \geq 2 \quad \forall z \in S \setminus F.
\]

The difference between the above two intersections is the number of vertices that \( A \cap \mathcal{L}_s \cap \tilde{\partial}_{\text{int}} \Lambda_n \) contains. We obtain
\[
|\pi_*^1(z) \cap \tilde{\partial}_{\text{int}} A| - |\pi_*^2(z) \cap \partial_{\text{int}} A| \leq 2 \quad \forall z \in S \setminus F.
\]

We characterize \( z \in S \setminus F \) according to this difference. For \( i \in \{0, 1, 2\} \) we define
\[
(S \setminus F)_i := \{ z \in S \setminus F : |\pi_*^1(z) \cap \tilde{\partial}_{\text{int}} A| - |\pi_*^2(z) \cap \partial_{\text{int}} A| = i \}.
\]

Then
\[
|\tilde{\partial}_{\text{int}} A| = \sum_{s \in F} |\pi_*^1(s) \cap \tilde{\partial}_{\text{int}} A| + \sum_{i \in \{0, 1, 2\}} \sum_{z \in (S \setminus F)_i} |\pi_*^1(z) \cap \tilde{\partial}_{\text{int}} A| = 2|F| + \sum_{i \in \{0, 1, 2\}} \sum_{z \in (S \setminus F)_i} |\pi_*^1(z) \cap \tilde{\partial}_{\text{int}} A| + i.
\]

Now we consider the ratio of the boundaries, i.e.,
\[
\frac{|\partial_{\text{int}} A|}{|\tilde{\partial}_{\text{int}} A|} = \frac{\sum_{i \in \{0, 1, 2\}} \sum_{z \in (S \setminus F)_i} |\pi_*^1(z) \cap \tilde{\partial}_{\text{int}} A|}{2|F| + \sum_{i \in \{0, 1, 2\}} \sum_{z \in (S \setminus F)_i} |\pi_*^1(z) \cap \tilde{\partial}_{\text{int}} A| + i}.
\]

Taking the set sizes \( |(S \setminus F)_i|, |F| \) fixed, it is elementary to see that the ratio is increasing in the summands of the double sum, i.e., its minimal value is attained when all summands are minimal. Now we use that each of the summands is at least 1, and obtain
\[
\frac{|\partial_{\text{int}} A|}{|\tilde{\partial}_{\text{int}} A|} \geq \frac{|S \setminus F|}{2|F| + |(S \setminus F)_1| + 2|(S \setminus F)_2| + |S \setminus F|}.
\]

We bound the denominator from above, i.e.,
\[
|F| + |S \setminus F| + |F| + |(S \setminus F)_1| + |(S \setminus F)_2| + |S \setminus F| \leq 3|S|,
\]

so
\[
\frac{|\partial_{\text{int}} A|}{|\tilde{\partial}_{\text{int}} A|} \geq \frac{|S \setminus F|}{3|S|}.
\]

We focus on the ratio on the right-hand side. For each vertex \( s \in F \), there are \( n^{1/d} \)-many vertices in \( A \) that are projected onto it (namely \( \mathcal{L}_s \)). Using also \( |S| \geq |A|^{(d-1)/d} \) by (A.3) and \( n \geq (4/3)|A| \), we get
\[
|A| \geq |F|n^{1/d} = \frac{|F|}{|S|} |S|n^{1/d} \geq \frac{|F|}{|S|} |A|^{(d-1)/d}n^{1/d} \geq \frac{4}{3}(4/3)^{1/d} |F| \geq |\tilde{\partial}_{\text{int}} A|.
\]

After rearranging we obtain \( |F| \leq (3/4)^{1/d} |S| \), and \( |\tilde{\partial}_{\text{int}} A| \geq (1 - (3/4)^{1/d})|\tilde{\partial}_{\text{int}} A| \) follows immediately from (A.3), since each projected vertex corresponds to at least two boundary vertices with respect to \( \mathbb{Z}^d \), i.e., \( |\partial_{\text{int}} A| \geq 2|S| \geq 2|A|^{(d-1)/d} \).

We turn to the inequality concerning \( \partial_{\text{ext}} A \) and \( \tilde{\partial}_{\text{ext}} A \) in (2.5). The inequality with \((*)\) in (2.6) holds for the same reason as for \( \partial_{\text{int}} A \). To obtain a lower bound on the ratio \( |\partial_{\text{ext}} A|/|\tilde{\partial}_{\text{ext}} A| \), we use that
each exterior boundary vertex is within distance one from an interior boundary vertex, which holds for both for \( \partial \) and \( \bar{\partial} \). Since each vertex has at most 2d vertices within distance one, it follows that
\[
|\partial_{\text{ext}}A| \geq \frac{1}{2d} |\partial_{\text{int}}A| \geq \frac{1}{2d} \left( 1 - \frac{3}{4d} \right)^{1/d} \cdot |\bar{\partial}_{\text{int}}A| \geq \frac{1}{2d} \left( 1 - \frac{3}{4d} \right)^{1/d} \cdot \frac{1}{2d} |\bar{\partial}_{\text{ext}}A|,
\]
and the proof is finished for \( \delta = (2d)^{-2} \left( 1 - \frac{3}{4d} \right)^{1/d} / 3 \).

\[\square\]

Proof of Lemma 2.7. We first show that there exists \( c_{\text{pei}} > 0 \) such that for all \( x \in \mathbb{Z}^d \) and \( m \in \mathbb{N} \)
\[
|\{ A \subseteq \Lambda_n : A \ni x, |A| = m, A \text{ is } \ast\text{-connected} \}| \leq \exp(c_{\text{pei}}m).
\]

Let \( A \) be in the set on the left-hand side. Since \( A \) is \( \ast\)-connected, the induced subgraph \( \mathbb{Z}_n^d[A] \) contains a spanning tree containing \( x \), which can be associated to a walk on the spanning tree (for example, walking through the tree in depth-first order), visiting each vertex in \( A \) at most twice. Since the degree of any vertex is \( 3^d - 1 \) in \( \mathbb{Z}_n^d \), the walk has at most \( 3^d - 1 \) options at each step for its next vertex, and has length at most \( 2m \). This shows (A.5).

For (2.6), we observe that each set \( A \) without holes (\( A = \bar{A} \)) can be uniquely reconstructed from its interior boundary \( \partial_{\text{int}}A \) (a vertex is in \( \bar{A} \setminus \partial_{\text{int}}A \) iff it is surrounded by \( \partial_{\text{int}}A \)), which is \( \ast\)-connected [12, Lemma 2.1]. Since we assume \( |\partial_{\text{int}}A| = m \), the isoperimetric inequality (2.5) ensures that \( |A| \leq C_1 m^{d/(d-1)} \) for some \( C_1 > 0 \) for all \( m \in \mathbb{N} \). This interior boundary must either contain \( x \) or surround \( x \) as defined in Definition 2.4.

We claim that there is a constant \( C > 0 \) such that \( \| x - \partial_{\text{int}}A \|_2 \leq C m^{1/(d-1)} \) for all \( x, A \) with \( A \ni x \) and \( A = \bar{A} \). Indeed, suppose otherwise. Then, on \( \mathbb{Z}^d \), vertices in the Euclidean ball of radius \( C m^{1/(d-1)} \) around \( x \) would be contained fully in \( A \) (without containing a vertex of \( \partial_{\text{int}}A \)). This would mean, for some dimension-dependent constant \( c_d \), that \( |A| \geq c_d (C m^{1/(d-1)})^d \), which contradicts that \( |A| \leq m^{d/(d-1)} \) by Claim 2.6 when \( C \) is chosen sufficiently large.

Hence, we may find a vertex \( y \in \partial_{\text{int}}A \cap \text{Ball}(C m^{1/(d-1)}, x) \) where the latter set denotes the Euclidean ball of radius \( C m^{1/(d-1)} \) around \( x \). Then, since \( \partial_{\text{int}}A \) is a \( \ast\)-connected set of size \( m \), (A.5) ensures that the number of possible sets \( S \) that may form \( \partial_{\text{int}}A \) is \( \exp(c_{\text{pei}}m) \). Summing over the possible choices of \( y \in \text{Ball}(C m^{1/(d-1)}, x) \), we arrive at
\[
|\{ A \in \mathcal{A} : A \ni x, A = \bar{A}, |\partial_{\text{int}}A| = m \}| \leq c_d C^d m^{d/(d-1)} \exp(c_{\text{pei}}m).
\]
The result follows by absorbing the factor \( c_d C^d m^{d/(d-1)} \) into the constant \( c_{\text{pei}} \).

\[\square\]

REFERENCES


