

## Large deviations of the giant in supercritical kernel-based spatial random graphs

Jorritsma, Joost; Komjáthy, Júlia; Mitsche, Dieter

**DOI**

[10.1007/s00440-025-01417-1](https://doi.org/10.1007/s00440-025-01417-1)

**Publication date**

2025

**Document Version**

Final published version

**Published in**

Probability Theory and Related Fields

**Citation (APA)**

Jorritsma, J., Komjáthy, J., & Mitsche, D. (2025). Large deviations of the giant in supercritical kernel-based spatial random graphs. *Probability Theory and Related Fields*. <https://doi.org/10.1007/s00440-025-01417-1>

**Important note**

To cite this publication, please use the final published version (if applicable).  
Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights.  
We will remove access to the work immediately and investigate your claim.



# Large deviations of the giant in supercritical kernel-based spatial random graphs

Joost Jorritsma<sup>1</sup> · Júlia Komjáthy<sup>2</sup> · Dieter Mitsche<sup>3</sup>

Received: 21 August 2025 / Accepted: 21 August 2025  
© The Author(s) 2025

## Abstract

We study cluster sizes in supercritical  $d$ -dimensional inhomogeneous percolation models with long-range edges —such as long-range percolation— and/or heavy-tailed degree distributions —such as geometric inhomogeneous random graphs and the age-dependent random connection model. Our focus is on large deviations of the size of the largest cluster in the graph restricted to a finite box as its volume tends to infinity. Compared to nearest neighbor Bernoulli bond percolation on  $\mathbb{Z}^d$ , we show that long edges can increase the exponent of the polynomial speed of the lower tail from  $(d-1)/d$  to any  $\zeta_\star \in ((d-1)/d, 1)$ . We prove that this exponent  $\zeta_\star$  also governs the size of the second-largest cluster, and the distribution of the size of the cluster containing the origin  $\mathcal{C}(0)$ . For the upper tail of large deviations, we prove that its speed is logarithmic for models with power-law degree distributions. We express the rate function via the generating function of  $|\mathcal{C}(0)|$ . The upper tail in degree-homogeneous models decays much faster: the speed in long-range percolation is *linear*.

**Keywords** Cluster-size distribution · Large deviation principle · Second-largest component · Spatial random graphs · Scale-free networks

**Mathematics Subject Classification** 05C80 · 60K35

---

✉ Joost Jorritsma  
joost.jorritsma@stats.ox.ac.uk

✉ Júlia Komjáthy  
j.komjathy@tudelft.nl

✉ Dieter Mitsche  
dmitsche@gmail.com

<sup>1</sup> Department of Statistics, University of Oxford, Oxford, England

<sup>2</sup> Delft Institute of Applied Mathematics, Delft University of Technology, Delft, The Netherlands

<sup>3</sup> Instituto de Ingeniería Matemática y Computacional, Pontificia Universidad Católica de Chile, Santiago, Chile

## 1 Introduction

For supercritical nearest neighbor Bernoulli bond percolation on  $\mathbb{Z}^d$  (NNP), it is a classical result that the graph restricted to a volume- $n$  box around the origin contains a *giant*, i.e., a linear-sized component  $\mathcal{C}_n^{(1)}$ , with probability tending to one as  $n \rightarrow \infty$ . Its number of vertices satisfies a Law of Large Numbers (LLN):  $|\mathcal{C}_n^{(1)}|/n$  converges to  $\theta := \mathbb{P}_{\text{NNP}}(0 \leftrightarrow \infty)$  almost surely. Regarding large deviations of  $|\mathcal{C}_n^{(1)}|$ , the speed is known in all dimensions  $d \geq 2$ , and an interesting discrepancy occurs: the lower tail decays slower than the upper tail.

For the upper tail, the event  $\{|\mathcal{C}_n^{(1)}| > (\theta + \varepsilon)n\}$  requires linearly more edges to be present than expected, which comes at an exponential cost. So, the speed of large deviations is *linear* in the volume of the box: for each  $\rho \in (\theta, 1)$ , there exist constants  $C_\rho \geq c_\rho > 0$  such that for all  $n \geq 1$ ,

$$-C_\rho \leq \frac{1}{n} \log \mathbb{P}_{\text{NNP}}(|\mathcal{C}_n^{(1)}| > \rho n) \leq -c_\rho. \quad (1.1)$$

We refer to [31, 33, 57] for the speed of large deviations of a closely related quantity, and to Theorem 2.4 below for the precise result. For the lower tail, the  $\Omega(n^{(d-1)/d})$  many edges on the “outer surface boundary” of a linear-sized cluster have to be absent. So, the speed of the lower tail is *sublinear* [69]: for each  $\rho \in (0, \theta)$ , there exist constants  $C'_\rho \geq c'_\rho > 0$  such that

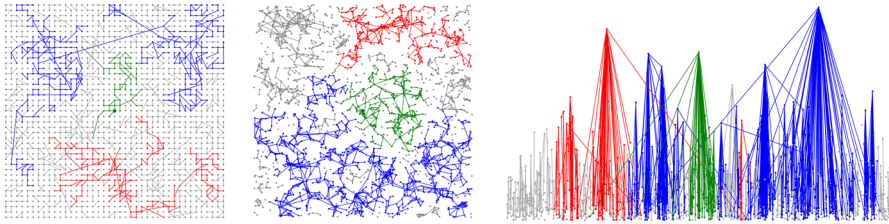
$$-C'_\rho \leq \frac{1}{n^{(d-1)/d}} \log \mathbb{P}_{\text{NNP}}(|\mathcal{C}_n^{(1)}| < \rho n) \leq -c'_\rho. \quad (1.2)$$

We can interpret that the speed of the lower tail is caused by *surface tension* (more precise results are known for closely related quantities, see [6, 21, 31, 33] and Section 2 below).

In this paper, we study large deviations of the giant in long-range percolation [71], and also in models with regularly-varying edge-length- and degree distributions. Our driving question is:

*How do high-degree vertices and long-range edges influence  
the tails of large deviations of the giant component?* (Q)

We describe three new phenomena that encompass the answer. First, high-degree vertices slow down the upper tail enormously: the speed becomes logarithmic. Second, long-range edges speed up the lower tail: the speed is still polynomial, but is no longer determined by surface tension alone. Since the lower tail decays much faster than the upper tail, the discrepancy of speeds is the opposite of the one in NNP. Third, we prove that the speed of the lower tail is closely related to the size of the second-largest connected component, and to the cluster-size distribution of the origin in the infinite model. We explain these phenomena after motivating this topic with an application and an informal model description.



**Fig. 1** Simulations of 2-dimensional long-range percolation (LRP), a 2-dimensional geometric inhomogeneous random graph (GIRG), and a 1-dimensional soft Poisson–Boolean model (sPBM). In the sPBM, the  $y$ -axis reflects the vertex marks. The presence of long edges in these models leads to delocalized components, and affects the distributions of the sizes of three components: the speed of the lower tail of large deviations of the largest component (blue), the size of the second-largest component (red), and the distribution of the size of the cluster containing the origin (green) are all governed by the exponent  $\zeta_*$ , which is defined in (1.4). The upper tail of large deviations of the giant’s size has linear speed in (mark)-homogeneous models as LRP, and logarithmic speed in inhomogeneous models as GIRG and sPBM (color figure online)

*Motivation.* Many real networks have been found to have underlying geometry, inhomogeneous degrees, clustering, and long-range connections [5, 66]. In biological contexts, the component sizes (under percolation of an initial graph) correspond to the final size of a Reed–Frost or SIR epidemic on the network [1, 10, 58], and to activity patterns in brain networks [20, 28, 53]. In the context of statistical physics, there are many models with long-range interactions, such as models of gravitational and charged systems, turbulent hydrodynamic systems, dipolar systems, and two-dimensional elasticity, see among others [19, 25, 26] for examples. In the last decade many random graph models have been invented to mimic real networks and systems with long-range interactions, see the next paragraph. For many of these models, an analysis of basic properties of their component structure—such as their approximate size-distribution—has been lacking in the literature. In this paper we fill this gap and develop new, robust proof techniques that allow for the presence of long-range edges and inhomogeneous degrees.

*Spatial inhomogeneous percolation models.* The most well-known spatial degree-inhomogeneous models are hyperbolic random graphs [54], geometric inhomogeneous random graphs [16], the (soft) Poisson–Boolean model with random radii [35, 43], and the age-dependent random connection model [36]. The degree-inhomogeneity leads to arbitrarily long edges. While the degree distribution of long-range percolation (LRP) is light-tailed, LRP also contains long edges. The presence of long edges leads to components that might be ‘delocalized in space’, see Figure 1. Our formal setup, introduced informally now, captures these models all at once.

The vertex sets are formed by a  $d$ -dimensional ergodic point process: either  $\mathbb{Z}^d$  or a homogeneous Poisson point process (PPP) on  $\mathbb{R}^d$ . Each vertex  $u \in \mathcal{V}$  at location  $x_u$  is equipped with an independent and identically distributed (iid) mark  $w_u \geq 1$ , which is Pareto-distributed for degree-inhomogeneous models, and a constant for long-range percolation. The expected degree of a vertex is proportional to its mark (sometimes also called ‘weight’). Conditionally on all vertex locations and marks  $\{(x_u, w_u)\}_{u \in \mathcal{V}}$ , each edge  $\{u, v\}$  is present independently with probability  $\varphi(\beta, w_u, w_v, \|x_u - x_v\|)$ . The (model-dependent) *edge-connectivity function*  $\varphi$  is non-increasing in the distance,

non-decreasing and symmetric in both marks. The parameter  $\beta > 0$  controls the edge-density. We often assume that the model is supercritical, i.e.,  $\beta > \beta_c$ , where  $\beta_c$  is the infimum over  $\beta \geq 0$  such that the graph contains an infinite connected component almost surely.

*Law of Large Numbers for the size of the giant.* In this general model class, even the existence of a linear-sized component in finite boxes is not guaranteed/known in the whole supercritical regime. When  $\varphi$  allows for sufficiently many long edges, we prove existence and uniqueness of the giant for all  $\beta > \beta_c$ . Its size satisfies a Law of Large Numbers:  $|\mathcal{C}_n^{(1)}|/n \rightarrow \theta := \mathbb{P}(0 \leftrightarrow \infty)$  in probability (for some cases the convergence holds almost surely). The long edges that are necessary for this LLN, come either from a slow decay of the connectivity function  $\varphi$  as  $\|x_u - x_v\| \rightarrow \infty$ , or from the presence of high vertex marks, or from a combination of both.

*Upper tail of large deviations.* We now state informally our result on the upper tail.

**Meta-theorem 1.1** (High-degree vertices cause slow upper tails). *Consider a spatial inhomogeneous percolation model with regularly-varying degree distribution, so that  $|\mathcal{C}_n^{(1)}|$ , the size of the giant, satisfies a Law of Large Numbers. The upper tail of  $|\mathcal{C}_n^{(1)}|$  has logarithmic speed. We identify the non-negative rate-function  $I(\cdot)$  such that for all  $\rho \in (\theta, 1)$ ,*

$$\begin{aligned} -\inf_{r>\rho} I(r) &\leq \liminf_{n \rightarrow \infty} \frac{1}{\log n} \log \mathbb{P}(|\mathcal{C}_n^{(1)}| > \rho n) \\ &\leq \limsup_{n \rightarrow \infty} \frac{1}{\log n} \log \mathbb{P}(|\mathcal{C}_n^{(1)}| > \rho n) = -\inf_{r \geq \rho} I(r) = I(\rho). \end{aligned} \quad (1.3)$$

*On the contrary, the speed in Bernoulli bond percolation or long-range percolation on  $\mathbb{Z}^d$  is linear.*

We formalize Meta-theorem 1.1 in Theorem 2.4 and Theorem 2.5 below. Meta-theorem 1.1 is the first result that identifies the rate function for the size of the giant in a spatial random graph model. Moreover, the rate function for the size of the giant is not even known for many classical *non-spatial* random graph models such as the rank-one inhomogeneous random graph with iid Pareto weights [22, 67], and the configuration model with iid Pareto degrees [63]. Some known results for non-spatial graphs are the following: The LDP of the giant in the configuration model with *fixed* degree sequence has linear speed when linear-degree vertices are absent [13]; in sparse inhomogeneous random graphs with bounded weights the speed is linear as well [8], similar to sparse Erdős-Rényi random graphs [9, 68]. A work in progress by the first author and Zwart shows an LDP for the giant in inhomogeneous random graphs with iid heavy-tailed vertex weights [50].

We briefly explain the reason for the logarithmic speed: with *polynomially* decaying probability, the vertex set contains a few hubs: vertices with mark  $\Theta(n)$ . These hubs connect to the giant and to linearly many small components. Small components do not merge with the giant with a probability that is exponentially small in their size and in the number of hubs. This allows us to express the rate-function  $I(\cdot)$  via the generating function of the cluster-size distribution of the infinite model.

A similar phenomenon occurs for the sum of  $n$  iid (fully asymmetric) Pareto random variables: this sum also satisfies an LDP with logarithmic speed of the upper tail. Having a single summand whose value is linear in  $n$  comes at a polynomial cost. On the contrary, the lower tail for such variables has linear speed because linearly many summands need to be small [65].

*Lower tail of large deviations.* Similar to the sum of iid Pareto random variables, the lower tail of the size of the giant decays much faster than the upper tail. However, due to the underlying geometry, the speed is not linear, but polynomial. Contrary to NNP in (1.2), in models with long-range edges the speed is not determined by surface tension alone. Instead, the speed is given by the solution of a variational problem that describes the most likely way that a box of volume  $n$ , denoted by  $\Lambda_n$ , is isolated from its complement  $\Lambda_n^c$  [48]. For most edge-connectivity functions  $\varphi$ , the probability that  $\Lambda_n$  is isolated decays as  $\exp(-\Theta(n^{\zeta_\star}))$  for some exponent  $\zeta_\star \in [0, 1)$ <sup>1</sup>. The exponent  $\zeta_\star$  can also be determined via the expected size of the *downward vertex boundary* that we define now. We say that the edge  $\{u, v\} = \{(x_u, w_u), (x_v, w_v)\}$  is a ‘downward edge’ from  $u$  if  $w_u \geq w_v$ . We define<sup>2</sup>

$$\zeta_\star := \lim_{n \rightarrow \infty} \frac{\log \mathbb{E} \left[ \left| \{u \in \Lambda_n : u \text{ has a downward edge to } \Lambda_n^c\} \right| \right]}{\log n}, \quad (1.4)$$

The restriction of counting vertices with downward edges to  $\Lambda_n^c$  is necessary for some (but not all) edge-connectivity functions  $\varphi$ . This excludes counting (potentially many) vertices on the vertex boundary that are connected by an edge to a small number of high-mark vertices spatially close to  $\Lambda_n$ . In NNP all marks are identical and  $\zeta_\star = (d-1)/d$ : the expectation in  $\zeta_\star$  is governed by surface tension. As we will see,  $\zeta_\star$  provides the right generalization of surface tension to spatial inhomogeneous percolation models with long edges. For these latter models,  $\zeta_\star = \max(\zeta_{\text{long}}, (d-1)/d)$ , where<sup>2</sup>

$$\zeta_{\text{long}} := \lim_{n \rightarrow \infty} \left( 0 \vee \frac{\log \mathbb{E} \left[ \left| \{u \in \Lambda_{n/2} : u \text{ has a downward edge to } \Lambda_n^c\} \right| \right]}{\log n} \right) \quad (1.5)$$

describes the number of vertices incident to long downwards edges of length  $\Omega(n^{1/d})$ . The value of  $\zeta_{\text{long}}$  can be explicitly computed based on the exact model specifications. Importantly,  $\zeta_{\text{long}}$  can be larger than  $(d-1)/d$  for any fixed  $d \in \mathbb{N}$ . In that case,  $\zeta_\star = \zeta_{\text{long}}$  and long edges speed up the decay of the lower tail of large deviations.

**Meta-theorem 1.2** (Long edges speed up the lower tail). *Consider a supercritical spatial inhomogeneous percolation model so that  $|\mathcal{C}_n^{(1)}|$ , the size of the giant, satisfies a Law of Large Numbers, i.e.,  $\zeta_{\text{long}} > 0$ . Assume that the expected downward vertex boundary in (1.4) has size  $\Theta(n^{\zeta_\star})$ . For each  $\rho \in (0, \theta)$ , there exist constants*

<sup>1</sup> For some connectivity functions  $\varphi$ , slowly-varying correction factors appear inside the “ $\Theta(\cdot)$ ” in the exp-function.

<sup>2</sup> The limits in the definitions of  $\zeta_\star$  and  $\zeta_{\text{long}}$  are well-defined in the models that we study. We will never use  $\zeta_{\text{long}}$  when it equals 0. The truncation ensures that the limit always exists.

$C_\rho \geq c_\rho > 0$  such that for all  $n \geq 1$ ,

$$-C_\rho \leq \frac{1}{n\zeta_\star} \log \mathbb{P}(|\mathcal{C}_n^{(1)}| < \rho n) \leq -c_\rho. \quad (1.6)$$

When the expected downward vertex boundary has slowly-varying correction factors, (potentially different) slowly-varying correction factors appear in the lower bound.

We formalize Meta-theorem 1.2 in Theorem 2.7. Meta-theorem 1.2 implies the existence of the giant with stretched exponential error probability, which is novel for several spatial inhomogeneous percolation models mentioned above on page 3. Moreover, this result includes long-range percolation and improves upon the currently known best bound there as well, as we explain now.

In long-range percolation, two vertices  $x, y \in \mathbb{Z}^d$  are connected by an edge independently with probability  $\Theta(1 \wedge \|x - y\|^{-\alpha d})$  as  $\|x - y\| \rightarrow \infty$ , for a long-range parameter  $\alpha > 1$ . An elementary computation shows that  $\zeta_{\text{long}} = 2 - \alpha$ , and  $\zeta_\star = \max(2 - \alpha, (d - 1)/d)$ . In [14], Biskup showed that there exists  $\delta > 0$  such that  $\mathbb{P}(|\mathcal{C}_n^{(1)}| < \delta n) \leq \exp(-n^{2-\alpha-o(1)})$ . Meta-theorem 1.2 improves upon [14] in two ways when  $2 - \alpha = \zeta_{\text{long}} > (d - 1)/d$ . First, the power of  $n$  in the exponent in (1.6) has no  $o(1)$ -term. Second, the decay in (1.6) holds for arbitrary  $\rho \in (0, \theta)$ , instead of a small constant  $\delta$ .

Returning to the general model setting, we give some intuition why  $\zeta_{\text{long}}$  drives the lower tail when  $\zeta_{\text{long}} > (d - 1)/d$ . The definition of  $\zeta_{\text{long}}$  in (1.5) tells us that  $\Lambda_n$  contains about  $\Theta(n^{\zeta_{\text{long}}})$  many vertices that are incident to long edges of length  $\Omega(n^{1/d})$ . We partition  $\Lambda_n$  into subboxes of size  $\Theta(n^{1-\zeta_{\text{long}}})$ , so that a subbox typically contains constantly many vertices incident to these long edges. Then, we couple the graph with an ERRG with large constant expected degree. In this coupling, each vertex corresponds to a local giant inside one of the  $\Theta(n^{\zeta_{\text{long}}})$  many subboxes, and we use the long edges to estimate the probability that two local giants are connected. Then we apply the LDP on the giant in ERRGs by O’Connell [68]. This LDP has linear speed in the number of ‘Erdős-Rényi vertices’, and yields that  $\mathbb{P}(|\mathcal{C}_n^{(1)}| < \rho n) \leq \exp(-\Theta(n^{\zeta_{\text{long}}}))$ . We apply this coupling *after* proving the LLN, which allows us to choose any  $\rho < \theta$ .

The Law of Large Numbers is a crucial ingredient to make this sketch work. The LLN ensures that local giants in subboxes exist and have the right size. We first establish the existence of the giant via two multi-scale renormalization schemes. To obtain the right size for the giant, it suffices to prove an upper bound on the cluster-size decay  $\mathbb{P}(k \leq |\mathcal{C}_n(0)|, 0 \notin \mathcal{C}_n^{(1)})$  in finite boxes [45, 48].

*Cluster-size decay and second-largest component.* In our third main result, we prove that  $\zeta_\star$  also governs the cluster-size decay of the origin and the size of the second-largest connected component  $|\mathcal{C}_n^{(2)}|$ . In [48], we already showed that both sizes are determined by  $\zeta_\star$  for some parameter settings of spatial inhomogeneous percolation models. In this paper, in Theorems 2.9 and 2.10, we prove the same relation for complementary parameter settings. Combining the present paper and [48], we obtain the following result. Here,  $\mathcal{C}(0)$  is the cluster of the origin in the model on  $\mathbb{R}^d$ .



**Meta-theorem 1.3** (Cluster-size decay and the second-largest component). *Consider a supercritical spatial inhomogeneous percolation model with  $\zeta_{\text{long}} > (d-1)/d$ . Assume that the expected downward vertex boundary in (1.4) has size  $\Theta(n^{\zeta_\star})$ . There exist constants  $C \geq c > 0$  such that for all  $k \geq 1$ ,*

$$-C \leq \frac{1}{k^{\zeta_\star}} \log \mathbb{P}(k < |\mathcal{C}(0)| < \infty) \leq -c, \quad \mathbb{P}\left(c \leq \frac{|\mathcal{C}_n^{(2)}|}{(\log n)^{1/\zeta_\star}} \leq C\right) \xrightarrow{n \rightarrow \infty} 1. \quad (1.7)$$

*When the expected downward vertex boundary has slowly-varying correction factors, (potentially different) slowly-varying correction factors appear in the lower bound.*

For long-range percolation, Crawford and Sly proved a polylogarithmic upper bound on  $\mathcal{C}_n^{(2)}$  in [24] with some unspecified exponent. Meta-theorem 1.3 finds the correct exponent of this polylog. For supercritical NNP and other homogeneous percolation models, surface-tension-driven behavior for the speed of the lower tail, for the cluster-size decay, and for the second-largest component is known, see Pisztora [69]. Moreover, when  $\zeta_{\text{long}} < (d-1)/d$  (equivalently  $\zeta_\star = (d-1)/d$ ), we proved in [49] that Meta-theorem 1.3 holds for long-range percolation with high edge-density. That is, for (high edge-density) LRP with  $\alpha > 1 + 1/d$ , the cluster-size decay and the size of the second-largest component are both driven by surface tension.

*Another exponent determining the presence of a supercritical phase.* Combined with [48], this paper completes the following picture:  $\zeta_\star = \max(\zeta_{\text{long}}, (d-1)/d)$  in (1.4) is the correct generalization of the surface-tension exponent  $(d-1)/d$  for supercritical spatial inhomogeneous percolation models. It governs the size of the second-largest component, the connected component containing a vertex at the origin, and the lower tail of large deviations of the giant component.

The value of  $\zeta_{\text{long}}$ , defined in (1.5), is related to the *effective decay exponent*  $\delta_{\text{eff}}$  defined in [47]. In words,  $2 - \delta_{\text{eff}}$  measures the size of the “long-edge” edge-boundary of a large box. To define it, say that  $u \in \Lambda_n \times [a, b]$  if its location  $x_u$  is in  $\Lambda_n$ , and its mark  $w_u$  is in the interval  $[a, b]$ , and let  $w_n^{\max}$  denote the maximal *expected* mark of a vertex in  $\Lambda_n$ .

$$2 - \delta_{\text{eff}} = \lim_{n \rightarrow \infty} \frac{\log \mathbb{E}\left[|\{\text{edges between } \Lambda_{n/2} \times [0, w_n^{\max}] \text{ and } \Lambda_n^c \times [0, w_n^{\max}]\}|\right]}{\log n}. \quad (1.8)$$

When  $2 - \delta_{\text{eff}} > 0$  and  $d = 1$ , [47] proves that the models studied here admit a supercritical phase, i.e.,  $\beta_c < \infty$ . If  $2 - \delta_{\text{eff}} < 0$  and  $d = 1$ , there is no supercritical phase [47]. For dimensions  $d \geq 2$ ,  $\beta_c < \infty$  holds in spatial inhomogeneous percolation models, since such models at high edge-density dominate supercritical random geometric graphs or Bernoulli bond percolation on  $\mathbb{Z}^d$ . Further, when  $2 - \delta_{\text{eff}} > 0$ , in this model class annulus-crossing probabilities do not decay with the size of the annulus. This is denoted by  $\hat{\lambda}_c = 0$  in the recent work in [47].

By its definition in (1.5),  $\zeta_{\text{long}}$  describes the size of the “long-edge” downward vertex-boundary. Hence,  $2 - \delta_{\text{eff}}$  has the same sign as  $\zeta_{\text{long}}$  throughout the model class. Combining this observation with the results in [47], positivity/negativity of  $\zeta_\star = \max(\zeta_{\text{long}}, (d-1)/d)$  also gives finiteness of  $\beta_c$  and  $\hat{\lambda}_c = 0$  in this model class.



The difference between  $\zeta_{\text{long}}$  and  $2 - \delta_{\text{eff}}$  becomes apparent as soon as degrees are inhomogeneous. High-degree vertices can have many edges leaving the box. Edges leaving the maximal vertex mark in  $\Lambda_n$  can form the dominating contribution to the edge-boundary. In this case, the formula in (1.8) is sensitive to truncation of marks and to the actual presence of the high-mark vertex. This is expressed in the non-concentration of the size of the edge-boundary around  $\Theta(n^{2-\delta_{\text{eff}}})$ . On the contrary, the size of the downward vertex boundary is robust to random fluctuations of the maximal degree, and its size concentrates around  $\Theta(n^{\zeta_{\text{long}}})$ . Summarizing, we observe that regarding presence/absence of supercritical phase,  $\delta_{\text{eff}}$  may be replaced by  $\zeta_*$  (or  $\zeta_{\text{long}}$ ), and that  $\zeta_*$  also drives the behavior of the three important distributions mentioned above in Meta-theorems 1.2 and 1.3. In the following section we formalize Meta-theorems 1.1–1.3 that answer the question (Q).

## 2 Formal statements of model and main results

### 2.1 Kernel-based spatial random graphs

We give a formal definition of a large class of spatial inhomogeneous percolation models. This class appeared under various names in the literature. The first general version of the model appeared in [52], called *general geometric inhomogeneous random graphs*, and independently in [37] where it was called *weight-dependent random connection model*. Afterwards, it was called *spatial inhomogeneous random graphs* in [73]. Due to the important role of a *kernel function* driving the connection probabilities, we called them *kernel-based spatial random graphs* in [48]. Since this paper ties in with [48], we keep this last name here. This model class includes classical homogeneous models such as Bernoulli bond percolation on  $\mathbb{Z}^d$  [17], long-range percolation [71], and random geometric graphs [34, 42]. It also includes (degree-) inhomogeneous models such as hyperbolic random graphs [54], the Poisson-Boolean model with random radii [43], geometric inhomogeneous random graphs [16], scale-free percolation [27], the age-dependent random connection model [36], and the scale-free Gilbert model [44]. We give now a formal definition.

**Definition 2.1** (*Kernel-based spatial random graphs (KSRG)*). Fix a dimension  $d \geq 1$ . A KSRG is then a random graph defined as follows. Let the vertex set  $V$  be formed by an ergodic point process: either  $\mathbb{Z}^d$  or a unit-intensity Poisson point process (PPP) on  $\mathbb{R}^d$ . Given  $V$ , we equip each vertex  $u \in V$  with an independent mark  $w_u \geq 1$  following distribution  $F_W$ . Let  $\kappa : \mathbb{R}_+^2 \rightarrow \mathbb{R}_+$  be a symmetric function, called the kernel function. Let  $\varphi : \mathbb{R}_+ \rightarrow [0, 1]$  be non-decreasing, called the profile function, and let  $\beta > 0$ . Conditionally on the realization of the marked vertex set  $\mathcal{V} := \{(x_u, w_u)\}_{u \in V} \subset \mathbb{R}^d \times \mathbb{R}_+$ , the edge  $\{u, v\} \in V^2$  is present in the edge-set  $\mathcal{E}$  independently of other edges with probability

$$p(u, v) := \mathbb{P}_{W, \kappa, \varphi, \beta}(u \text{ connected by an edge to } v \mid \mathcal{V}) = \varphi\left(\beta \cdot \frac{\kappa(w_u, w_v)}{\|x_u - x_v\|^d}\right). \quad (2.1)$$

We denote the obtained infinite graph by  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ , and by  $\mathcal{G}_n = (\mathcal{V}_n, \mathcal{E}_n)$  the graph induced by vertices in  $\Lambda_n := [-n^{1/d}/2, n^{1/d}/2]^d$ . We write  $\mathcal{C}_n^{(i)}$  for the  $i$ th largest component of  $\mathcal{G}_n$ , and  $\mathcal{C}_n(0)$  and  $\mathcal{C}(0)$  for the component containing a vertex at the origin in  $\mathcal{G}_n$  and in  $\mathcal{G}$ , respectively. We write  $\mathbb{P}^x$  for the Palm-measure when the vertex set of a unit-intensity Poisson point process is conditioned to contain a vertex at location  $x \in \mathbb{R}^d$  with unknown mark.

We omit the subscripts  $W, \kappa, \varphi$  of  $\mathbb{P}$  as they are fixed and always clear from the context. Sometimes we emphasize the dependency of  $\mathbb{P}$  on the edge-density parameter  $\beta$  by writing  $\mathbb{P}_\beta$ . We define

$$\theta = \theta(\beta) := \mathbb{P}_\beta^0(0 \leftrightarrow \infty) = \mathbb{P}_\beta^0(|\mathcal{C}(0)| = \infty) > 0. \quad (2.2)$$

Then, the critical edge-density  $\beta_c$  is given by

$$\beta_c := \inf \{ \beta \geq 0 : \theta(\beta) > 0 \}. \quad (2.3)$$

We call the model subcritical if  $\beta < \beta_c$ , critical if  $\beta = \beta_c$ , and supercritical if  $\beta > \beta_c$ . For a fixed choice of  $W, \varphi, \kappa$ , and  $d$ , some phases may not exist [38, 39].

Definition 2.1 allows for general kernel and profile functions. In the rest of the paper we restrict them to a few choices that are commonly used, and which cover the specific models in the introduction [16, 27, 36, 43, 44, 54, 71]. Our results remain valid for (or can be extended to) other sufficiently similar kernels, profile functions, and mark distributions. There is no additional technicality in the proofs if one includes slowly varying functions in any of these functions, and the only change in the results is that lower order correction terms (for instance, slowly varying functions multiplying the polynomial exponent of cluster-size decay) appear in the results below. These lower-order corrections can be traced throughout the proofs. For ease of presentation, we thus work with functions satisfying the following assumption. We follow the notation of our previous work [48], and write for any  $a, b, \in \mathbb{R}$ ,  $a \wedge b$  for  $\min(a, b)$ , and  $a \vee b$  for  $\max(a, b)$ .

**Assumption 2.2** We assume that the mark distribution is either  $W_v \equiv 1$  for all  $v$ , or a Pareto distribution with parameter  $\tau > 2$ :

$$1 - F_W(w) := \mathbb{P}(W_v \geq w) = w^{-(\tau-1)}, \quad w \geq 1. \quad (2.4)$$

We assume that the profile function  $\varphi$  is either threshold or polynomial: for constants  $p \in (0, 1], \alpha > 1$ , let

$$\varphi_{\text{thres}}(s) := p \mathbb{1}_{\{s \geq 1\}}, \quad \text{or} \quad \varphi_{\text{pol}}(s) := p(1 \wedge s^{-\alpha}), \quad (2.5)$$

and we assume that the kernel  $\kappa$  is one of the following:

$$\kappa_{\text{sum}}(w_1, w_2) = (w_1^{1/d} + w_2^{1/d})^d, \quad \text{or} \quad \kappa_\sigma(w_1, w_2) = (w_1 \vee w_2)(w_1 \wedge w_2)^\sigma, \quad (2.6)$$

for a parameter  $\sigma \geq 0$ . By abuse of notation, when  $W_v \equiv 1$  for all  $v \in V$  we say that  $\tau = \infty$ ; when  $\varphi = \varphi_{\text{thres}}$  we say that  $\alpha = \infty$ ; when  $\kappa = \kappa_{\text{sum}}$  we say that  $\sigma = 0$ .

Assumption 2.2 ensures that the model is parametrized so that the expected degree of a vertex is proportional to its mark when  $\sigma \leq \tau - 1$ . The restrictions  $\tau > 2$  and  $\alpha > 1$  ensure that the graph is locally finite. Increasing  $\tau$  and/or  $\alpha$  leads to less inhomogeneity (lighter-tailed degrees and fewer long edges, respectively). All our results below for  $\tau = \infty$  except Theorem 2.4 can be directly extended to all distributions  $F_W$  with lighter tail than any power-law in (2.4). The parameter  $p$  in (2.5) makes the model closed under Bernoulli percolation of the edges, while the parameter  $\sigma$  controls assortativity: increasing  $\sigma$  makes it more likely that high-degree vertices are connected by an edge. The parameter  $\sigma$  allows us to continuously interpolate between well-known models that are special cases. Therefore, we call  $\kappa_\sigma$  the *interpolation kernel*, see more in [48].

Our earlier work [48] as well as the papers [37, 73] describe how the above models can be reconstructed as special cases. In short, in long-range percolation, the vertex set is  $\mathbb{Z}^d$ , and the marks are a constant, i.e.,  $\tau = \infty$ , and  $\alpha < \infty$ . The kernel  $\kappa_1$ , corresponding to the product  $w_1 w_2$ , is used to obtain hyperbolic random graphs, geometric inhomogeneous random graphs, and scale-free percolation [16, 27, 54],  $\kappa_{\tau-2}$  to obtain the age-dependent random connection model [36], and  $\kappa_{\text{sum}}$ , with additionally  $p = 1$ , to obtain the (soft) Poisson–Boolean model [43]. The soft Poisson–Boolean model uses the polynomial profile  $\varphi_{\text{pol}}$ , while the classical Poisson–Boolean model uses the threshold profile  $\varphi_{\text{thres}}$ . The kernel  $\kappa_0$  is sometimes called the *max-kernel*, and is closely related to  $\kappa_{\text{sum}}$ , as  $\kappa_0 \leq \kappa_{\text{sum}} \leq 2\kappa_0$ . Since the latter two kernels are qualitatively similar, we simply set  $\sigma = 0$  for  $\kappa_{\text{sum}}$ ; we never prove<sup>3</sup> explicitly results for  $\kappa_{\text{sum}}$ .

## 2.2 Dominant connection types

To make Meta-theorems 1.1–1.3 formal, we give the values of  $\zeta_\star$  in (1.4) and  $\zeta_{\text{long}}$  in (1.5) in terms of the model parameters for KSRRs satisfying Assumption 2.2. We follow the notation of [48], and write  $\zeta_\star$  and  $\zeta_{\text{long}}$  as a maximum of simple expressions of the parameters  $d, \tau, \alpha, \sigma$ . As we will show below, they take the form

$$\zeta_\star = \max(\zeta_{\text{short}}, \zeta_{\text{ll}}, \zeta_{\text{hl}}, \zeta_{\text{hh}}), \quad \zeta_{\text{long}} = \max(\zeta_{\text{ll}}, \zeta_{\text{hl}}, \zeta_{\text{hh}}). \quad (2.7)$$

Each term inside the maxima describes the expected number of vertices with a certain type of downward edge leaving  $\Lambda_n$ . Each of these types contributes to the downward vertex boundary of  $\mathcal{V} \cap \Lambda_n$  in (1.4). If the maximum is unique, the size of the downward vertex boundary is dominated by exactly one of these four types. If the maximum is non-unique, we say that the parameters are on a phase-transition boundary. We provide some back-of-the-envelope calculations to compute  $\zeta_\star$  in (1.4) and  $\zeta_{\text{long}}$  in (1.5) when  $\alpha < \infty$ . When  $\alpha = \infty$ , one has to take the limit in the expressions below as  $\alpha \rightarrow \infty$ . The precise computations are given in [48].

<sup>3</sup> The corresponding results are straightforward, in lower/upper bounds we use the corresponding bound from  $\kappa_0 \leq \kappa_{\text{sum}} \leq 2\kappa_0$ , and constant factors do not matter in our calculations.

*Short edges* – of constant length: there are roughly  $\Theta(n^{(d-1)/d})$  vertices incident to such edges crossing the boundary of  $\Lambda_n$  in (1.4), giving the ‘surface-tension’ exponent

$$\zeta_{\text{short}} := (d - 1)/d. \quad (2.8)$$

Next, we start counting ‘long edges’, that is, edges of length  $\Theta(n^{1/d})$  crossing the boundary of  $\Lambda_n$ , and thus also contributing to  $\zeta_{\text{long}}$  in (1.5). These edges fall into three categories.

*Low-low edges* connect two vertices of constant (low) mark between  $\Lambda_{n/2}$  and  $\Lambda_n^c$ . The expected number of low-low edges is  $\Theta(n \cdot n \cdot n^{-\alpha})$ , and this is the same as the number of low-mark vertices having a low-low edge. Abbreviating ‘low-mark to low-mark’ by ll, we obtain

$$\zeta_{\text{ll}} := 2 - \alpha. \quad (2.9)$$

The other two types describe long edges incident to ‘high-mark’ vertices in  $\Lambda_{n/2}$ . We say a vertex has a high mark if its mark is at least  $n^{\gamma_{\text{long}}}$ , where

$$\gamma_{\text{long}} := \min \left\{ \gamma \geq 0 : \liminf_{n \rightarrow \infty} \mathbb{E}^0 \left[ |\{\text{edges between } 0 \text{ and } \Lambda_n^c\}| \mid (0, n^\gamma) \in \mathcal{V} \right] > 0 \right\}. \quad (2.10)$$

Then, a constant proportion of vertices of mark at least  $n^{\gamma_{\text{long}}}$  inside  $\Lambda_{n/2}$  contributes to the vertex boundary. By the Pareto distribution of the marks in (2.4), there are  $\Theta(n^{1-\gamma_{\text{long}}(\tau-1)})$  many high-mark vertices inside  $\Lambda_{n/2}$ . The values  $\tau$  in (2.4),  $\sigma$  in (2.6), and  $\alpha$  in (2.5) jointly determine the value of  $\gamma_{\text{long}}$ . We compute two possible values of  $\gamma_{\text{long}}$ , that then yield the two remaining types.

*High-low edges* are dominant if the other end-vertex of an edge emanating from vertex  $(0, n^\gamma)$  to  $\Lambda_n^c$  typically has constant mark. There are  $\Theta(n)$  constant-mark vertices at distance  $\Theta(n^{1/d})$ . By the connection probability in (2.1) with  $\kappa_\sigma$  or  $\kappa_{\text{sum}}$  from (2.6) and  $\varphi_{\text{pol}}$  from (2.5), for  $\gamma > 0$ , the expected number of edges between vertex  $(0, n^\gamma)$  and constant-mark vertices in  $\Lambda_n^c$  is roughly  $n(n^\gamma/n)^\alpha$ . As required in (2.10), this expression is of constant order when

$$\gamma = \gamma_{\text{hl}} := 1 - 1/\alpha, \quad (2.11)$$

which leads to

$$\zeta_{\text{hl}} := 1 - \gamma_{\text{hl}}(\tau - 1) = (\tau - 1)/\alpha - (\tau - 2). \quad (2.12)$$

*High-high edges* occur dominantly if the other end-vertex of an edge emanating from vertex  $(0, n^\gamma)$  in (2.10) typically has a high mark. There are  $\Theta(n^{1-\gamma(\tau-1)})$  many vertices of mark  $\Omega(n^\gamma)$  at distance  $\Theta(n^{1/d})$  from 0. Using the interpolation kernel in (2.6), the expected number of edges between  $(0, n^\gamma)$  and these other vertices is roughly  $n^{1-\gamma(\tau-1)}(n^{\gamma(\sigma+1)}/n)^\alpha$ . This expression is of constant order when

$$\gamma = \gamma_{\text{hh}} := \begin{cases} \frac{1-1/\alpha}{\sigma+1-(\tau-1)/\alpha}, & \text{if } \tau \leq \sigma + 2, \\ \frac{1}{\sigma+1}, & \text{if } \tau > \sigma + 2, \end{cases} \quad (2.13)$$

which leads to

$$\zeta_{hh} := 1 - \gamma_{hh}(\tau - 1) = \begin{cases} \frac{\sigma+2-\tau}{\sigma+1-(\tau-1)/\alpha}, & \text{if } \tau \leq \sigma + 2, \\ \frac{\sigma+2-\tau}{\sigma+1}, & \text{if } \tau > \sigma + 2. \end{cases} \quad (2.14)$$

We will never use the second row for  $\gamma_{hh}$  and  $\zeta_{hh}$ :  $\zeta_{hh}$  is negative there, and some other connectivity type is dominant. The choice of  $\gamma_{hh}$  in this second row ensures continuity and monotonicity in the parameters. High-low connections can be dominant when  $\sigma$  is small, for example in the age-dependent random connection model and the soft Poisson–Boolean model. The high-high type of connection can be dominant for large values of  $\sigma$ , e.g. in geometric inhomogeneous random graphs and random hyperbolic graphs. The infimum in (2.10) indicates that  $\gamma_{long}$  is the minimum of  $\gamma_{hl}$  and  $\gamma_{hh}$ , and the switch is when the denominator  $\sigma + 1 - (\tau - 1)/\alpha$  crosses 1. Indeed, the following claim makes the above back-of-the-envelope computations formal: the integrals in the definitions of  $\zeta_\star$  in (1.4) and  $\zeta_{long}$  in (1.5) are dominated by the above four cases. Vertices incident to edges of intermediate length  $\ell_n$  ( $1 \ll \ell_n \ll n^{1/d}$ ) crossing  $\Lambda_n$  do not significantly contribute to the vertex boundary.

**Claim 2.3** (*Phases of  $\zeta_\star$ , [48, Lemma 7.6]*) *Consider any KSRG with kernel and profile satisfying Assumption 2.2. Then*

$$\zeta_{long} = \max(0, \zeta_{ll}, \zeta_{hl}, \zeta_{hh}), \quad \zeta_\star = \max(\zeta_{short}, \zeta_{long}). \quad (2.15)$$

The values of  $\zeta_\star$  and  $\zeta_{long}$  do not depend on  $p$ ,  $\beta$ , and  $\zeta_{long}$  not even on  $d$ . We refer to [48] for phase diagrams that illustrate for which values of the parameters  $(\tau, \alpha, \sigma, d)$  the four different connection types are dominant. We call the areas in the parameter space corresponding to unique dominant types the *phases* of  $\zeta_\star$ . On the boundaries of these phases, multiple connection types are dominant simultaneously. These cause poly-logarithmic factors in the expectation in the numerator of  $\zeta_\star$  in (1.4). These factors vanish in the limit, but they do appear in some of our lower bounds. So, for a given parameter setting of  $\tau, \alpha, \sigma, d$ , we let  $m_Z$  count the number of dominant connection types. Formally, we define the *multiplicity* of the maximum of  $Z := \{\zeta_{ll}, \zeta_{hl}, \zeta_{hh}, \zeta_{short}\}$  as

$$m_Z := \sum_{\zeta \in Z} \mathbb{1}\{\zeta = \max(Z)\}. \quad (2.16)$$

We now formalize Meta-theorems 1.1, 1.2, and 1.3 to answer the question (Q).

### 2.3 Upper tail of large deviations

First we state the upper tail of the giant's size for LRP on  $\mathbb{Z}^d$  and Bernoulli bond percolation on  $\mathbb{Z}^d$  (NNP), making (1.1) precise.

**Theorem 2.4** (Linear speed in degree-homogeneous models). *Consider long-range percolation or Bernoulli bond percolation on  $\mathbb{Z}^d$  with parameters such that  $\theta < 1$ .*

For all  $\rho \in (\theta, 1)$ , there exists  $A > 0$  such that for all  $n \geq 1$ ,

$$-A \leq \frac{1}{n} \log \mathbb{P}(|\mathcal{C}_n^{(1)}| > \rho n) \leq -1/A.$$

Theorem 2.4 applies for *any* edge density  $\beta$ , i.e., regardless of whether the model is supercritical, subcritical or critical. Linear speed for the upper tail of large deviations of a closely related quantity is known for NNP [31, 33]: the number of vertices in the intersection of a finite box and the infinite component.

We proceed to the upper tail for models with power-law degree distributions, when the speed is logarithmic, heading towards making Meta-theorem 1.1 precise. At an increased cost of technicality our subsequent results can be extended to KSRGs on  $\mathbb{Z}^d$ ; in order to keep the proof simpler, we refrain from this and restrict to Poisson vertex sets for degree-inhomogeneous KSRGs, i.e., with  $\tau < \infty$ . To state the rate function, we introduce some notation. We write  $\mathcal{C}(0)$  for the cluster of the origin in the model on  $\mathbb{R}^d$ , and recall from Definition 2.1 that we use the superscript  $x$  when the vertex set has a vertex at location  $x$  with unknown mark. Let

$$H_{\mathcal{C}}(z) := \mathbb{E}^0[z^{|\mathcal{C}(0)|} \mathbb{1}_{\{|\mathcal{C}(0)| < \infty\}}]$$

be the probability generating function of the size of  $\mathcal{C}(0)$  restricted to be finite. The function  $H_{\mathcal{C}}(z)$  is continuous and increasing, and has range  $(0, 1 - \theta]$  for  $z \in (0, 1]$ . Hence, the inverse  $H_{\mathcal{C}}^{(-1)}(y)$  is well-defined and increasing when  $y \in (0, 1 - \theta]$ . For  $\rho \in (\theta, 1)$ , define

$$\text{hubs}(\rho) := \begin{cases} \frac{\log H_{\mathcal{C}}^{(-1)}(1-\rho)}{\log(1-\rho)}, & \text{if } \rho < 1, \\ 1, & \text{if } \rho = 1. \end{cases} \quad (2.17)$$

The distinction based on  $p$  ensures that  $\log(1 - p)$  is well-defined. We prove that  $\lceil \text{hubs}(\rho) \rceil$  many linear-mark vertices are required to increase the density of the giant from the typical density  $\theta$  to  $\rho$ , see Section 3.2 below for intuition. The function  $\text{hubs}(\rho)$  is positive, increasing in  $\rho$  and tends to infinity as  $\rho \uparrow 1$ : similar to [50], it can be shown that  $\text{hubs}(\rho) = \log(1 - \rho)/\log(1 - p) + o(1)$  as  $\rho$  approaches 1. The following theorem gives the rate function for the upper tail. Recall  $\zeta_{\text{long}}$  from (1.5).

**Theorem 2.5** (Logarithmic speed in degree-inhomogeneous models). *Consider a KSRG satisfying Assumption 2.2 with a unit-intensity Poisson point process as vertex set and power-law mark distribution with  $\tau < \infty$ . Fix  $\rho \in (\theta, 1)$ . There exists  $A > 0$  such that for all  $n \geq 1$ ,*

$$\mathbb{P}(|\mathcal{C}_n^{(1)}| > \rho n) \leq A n^{-(\tau-2)\lceil \text{hubs}(\rho) \rceil}. \quad (2.18)$$

*If either the model is supercritical with  $\zeta_{\text{long}} > 0$ , or  $\theta = 0$ , then there exists  $A > 0$  such that for all  $n \geq 1$ ,*

$$\mathbb{P}(|\mathcal{C}_n^{(1)}| > \rho n) \geq (1/A) n^{-(\tau-2) \lim_{r \downarrow \rho} \lceil \text{hubs}(r) \rceil}. \quad (2.19)$$

The formulation of the either/or above (2.19) excludes critical models from the lower bound with  $\theta(\beta_c) > 0$ , where  $\beta_c$  is the critical edge-density in (2.3). For KSRGs in general it is unknown when  $\theta(\beta_c)$  is positive: this may sensitively depend on the short edges of the model [3, 4]. The preprint [64] by Mönch shows that  $\theta(\beta_c) = 0$  for KSRGs in dimensions  $d \geq 2$  with  $\zeta_{\text{long}} > 0$ .

Combining (2.18) and (2.19) gives that when  $\text{hubs}(\rho) \notin \mathbb{N}$ , we identify the decay rate of  $\mathbb{P}(|\mathcal{C}_n^{(1)}| > \rho n)$  up to a constant factor. If  $\text{hubs}(\rho) \in \mathbb{N}$  and  $p < 1$ , the lower and upper bounds are no longer of the same order, because the necessary “wiggle room” in the proof of the lower bound is lost. Nevertheless, Theorem 2.5 gives logarithmic speed for the upper tail with rate function

$$I(\rho) := (\tau - 2)\lceil \text{hubs}(\rho) \rceil.$$

**Corollary 2.6** (Rate function for the upper tail). *Consider a KSRG satisfying Assumption 2.2 on a unit-intensity Poisson point process as vertex set and power-law mark distribution with  $\tau < \infty$ . If the model is either supercritical with  $\zeta_{\text{long}} > 0$  or  $\theta = 0$ , then for all  $\rho \in (\theta, 1)$ ,*

$$\begin{aligned} -\inf_{r>\rho} I(r) &\leq \liminf_{n \rightarrow \infty} \frac{1}{\log n} \log \mathbb{P}(|\mathcal{C}_n^{(1)}| > \rho n) \\ &\leq \limsup_{n \rightarrow \infty} \frac{1}{\log n} \log \mathbb{P}(|\mathcal{C}_n^{(1)}| > \rho n) \leq -\inf_{r \geq \rho} I(r). \end{aligned} \quad (2.20)$$

**Proof** Since  $\text{hubs}(\rho)$  in (2.17) is continuous and non-decreasing, both bounds follow directly from Theorem 2.5.  $\square$

When we compare this result to Theorem 2.4 we see that the speed drops from linear to logarithmic when degree-inhomogeneity enters the models. The assumption  $\zeta_{\text{long}} > 0$  in (2.19) reflects the aim to study the influence of degree and edge-length inhomogeneity in (Q). However,  $\zeta_{\text{long}} > 0$  is a technical condition here. It comes from our bounds for the lower tail, which serves as a tool for proving (2.19).

## 2.4 Lower tail of large deviations

Our next result determines the speed of the lower tail of large deviations of the giant component for supercritical KSRGs, i.e., when  $\beta > \beta_c$  in (2.3). This will make Meta-theorem 1.2 precise. Contrary to Bernoulli bond percolation, the lower tail decays faster than the upper tail when  $\tau < \infty$ : the speed is polynomial, with exponent  $\zeta_\star = \max(\zeta_{\text{long}}, \zeta_{\text{short}})$  where  $\zeta_{\text{short}} = (d - 1)/d$  and  $\zeta_{\text{long}}$  is defined in (1.5). Recall the definition of  $\zeta_\star$  from (1.4), its value from Claim 2.3, and the multiplicity  $m_Z$  from (2.16).

**Theorem 2.7** (Speed in the lower tail of large deviations for the giant). *Consider a supercritical KSRG satisfying Assumption 2.2 with a unit-intensity Poisson point process as vertex set, or consider supercritical long-range percolation on  $\mathbb{Z}^d$ . Assume*



that the model satisfies  $\zeta_{\text{long}} > 0$ . Fix  $\rho \in (0, \theta)$ . There exists a constant  $A > 0$  such that for all  $n \geq 1$ ,

$$\log \mathbb{P}(|C_n^{(1)}| < \rho n) \in \left( -An^{\zeta_\star}(\log n)^{m_Z-1}, -\frac{1}{A}n^{\zeta_\star} \right). \quad (2.21)$$

The lower bound in Theorem 2.7 is proved in our earlier work [48, Theorem 2.5]. When a single connection type is dominant (i.e.,  $m_Z = 1$ , see Section 2.2), the lower and upper bound on the right-hand side differ by a constant factor, matching Meta-theorem 1.2. For such parameters, Theorem 2.7 determines the exact *speed* of large deviations for the lower tail. When  $m_Z \in \{2, 3, 4\}$ , the bounds differ by a poly-log factor.

We discuss a few related results. For Bernoulli nearest neighbor percolation on  $\mathbb{Z}^d$  (NNP), the exponent of the polynomial speed of the lower tail is  $\zeta_\star = \zeta_{\text{short}} = (d-1)/d$ , see (1.2) and [69]. We expect that Theorem 2.7 remains true<sup>4</sup> also for KSRGs with  $\zeta_{\text{long}} \leq 0$  in  $d \geq 2$ . In view of (Q), this is beyond the scope of our paper. For Bernoulli NNP on  $\mathbb{Z}^2$  or  $\mathbb{Z}^3$ , the works [6] and [21] identify the *rate function* of the lower tail for the size of the intersection of the infinite component with  $\Lambda_n$ . Keep in mind that this is not necessarily the size of the largest component in  $\Lambda_n$ .

For long-range percolation, Biskup proved in [14] that  $\mathbb{P}(|C_n^{(1)}| < \delta n) \leq \exp(-\Theta(n^{2-\alpha-o(1)}))$  for  $\delta > 0$  sufficiently small, which is an almost sharp upper bound, see our comment below Meta-theorem 1.2. Hyperbolic random graphs [32, 54] are equivalent to one-dimensional threshold geometric inhomogeneous random graphs and are a special instance of the KSRGs [52] by setting  $p = 1$ ,  $\sigma = 1$ ,  $\alpha = \infty$  in Assumption 2.2. Independent of our work, the recent preprint [15] proves the upper bound  $\mathbb{P}(|C_n^{(1)}| < \delta n) \leq \exp(-\Theta(n^{(3-\tau)/2}))$  for  $\delta > 0$  sufficiently small. This is a special case of the upper bound here, and  $(3-\tau)/2$  matches the exponent  $\zeta_{\text{hh}}$  in our earlier general lower bound in [48] for  $\alpha = \infty$ . In comparison, Theorem 2.7 allows (beyond other kernels) for  $p < 1$  and long-range profiles. It also shows that  $\delta$  can be chosen arbitrarily close to  $\theta$ .

The previous three theorems give the *strong* Law of Large Numbers as a corollary for KSRGs with finite-variance degree distributions, which is not known for (specific) models in the KSRG class<sup>5</sup>. For KSRGs with an infinite-variance mark distribution, we obtain a weak LLN.

**Corollary 2.8** (Law of Large Numbers). *Consider either supercritical long-range percolation on  $\mathbb{Z}^d$  with  $\zeta_{\text{long}} > 0$ , or Bernoulli nearest neighbor bond percolation on  $\mathbb{Z}^d$ , or a supercritical KSRG satisfying Assumption 2.2 on a unit-intensity PPP with  $\tau > 3$  and  $\zeta_{\text{long}} > 0$ . Then*

$$\frac{|C_n^{(1)}|}{n} \xrightarrow{a.s.} \theta = \theta(\beta, p, \alpha, d), \quad \text{as } n \rightarrow \infty. \quad (2.22)$$

<sup>4</sup> In dimension  $d = 1$  there is no lower tail since the model is subcritical when  $\zeta_{\text{long}} < 0$ , see below (1.8).

<sup>5</sup> While the result seems folklore to us for Bernoulli NNP, we could not find a reference for the exact quantity we look at here. We are also not aware of strong LLN for other models in the class.

For a supercritical KSRG satisfying Assumption 2.2 on a unit-intensity PPP with  $\zeta_{\text{long}} > 0$  and  $\tau \in (2, 3]$ ,

$$\frac{|\mathcal{C}_n^{(1)}|}{n} \xrightarrow{\mathbb{P}} \theta = \theta(\beta, p, \alpha, \tau, \sigma, d), \quad \text{as } n \rightarrow \infty. \quad (2.23)$$

**Proof** For NNP, a lower bound follows from the lower tail of large deviations by Pisztor in [69], an upper bound follows from Theorem 2.4. The weak law of large numbers follows directly from Theorem 2.4, Theorem 2.5 and Theorem 2.7. The error terms are summable for LRP, NNP, and KSRGs with  $\tau > 3$  (in (2.18)): an application of the Borel–Cantelli Lemma gives the strong law, while the weak law still holds also for KSRGs with  $\tau \in (2, 3]$ .  $\square$

*Related work on LLN.* Fountoulakis and Müller in [32] proved a weak LLN for the giant in hyperbolic random graphs, which is a threshold one-dimensional KSRG with infinite variance degrees satisfying  $\zeta_{\text{hh}} > 0$ , and is contained here in (2.23). In [48, Corollary 2.3], we extended the weak LLN of [32] to general KSRGs on PPPs with  $\zeta_{\text{hh}} > 0$ . Finally, for general long-range percolation models with  $\zeta_{\text{long}} \leq 0$ , Bäumler in [11, Theorem 1.6] recently showed that  $\mathbb{P}(|\mathcal{C}_n^{(1)}| \geq \rho n) \rightarrow 1$  for any  $\rho \in (0, \theta)$ . Below, we first prove a weak LLN when  $\zeta_{\text{long}} = \max(\zeta_{\text{hh}}, \zeta_{\text{hl}}, \zeta_{\text{ll}}) > 0$ . This weak LLN serves as a tool to proving Theorems 2.5 and 2.7. When simultaneously  $\zeta_{\text{hh}} \leq 0$  and  $\max(\zeta_{\text{hl}}, \zeta_{\text{ll}}) > 0$ , the LLN requires new methods compared to the methods in [32, 48]. These theorems give the strong law in some cases.

## 2.5 Cluster-size decay and second-largest component

We proceed to formalizing Meta-theorem 1.3. We start with the cluster-size decay of the origin. Let  $\mathcal{C}(0)$  denote the cluster of the origin in the infinite model. Recall  $\zeta_{\text{long}}$  from (1.5),  $\zeta_{\text{short}} = (d-1)/d$  and  $m_{\mathcal{Z}}$  from (2.16).

**Theorem 2.9** (Cluster-size decay). *Consider a supercritical KSRG satisfying Assumption 2.2 with a unit-intensity Poisson point process as vertex set, or consider supercritical long-range percolation on  $\mathbb{Z}^d$ . Assume that the model satisfies  $\zeta_{\text{long}} > \max\{0, \zeta_{\text{hh}}\}$ . Then there exists a constant  $A > 0$  such that for all  $k \geq 1$  and  $n \in [Ak, \infty]$ ,*

$$\log \mathbb{P}^0(|\mathcal{C}_n(0)| > k, 0 \notin \mathcal{C}_n^{(1)}) \in \left(-Ak^{\max(\zeta_{\text{long}}, \zeta_{\text{short}})}(\log k)^{m_{\mathcal{Z}}-1}, -\frac{1}{A}k^{\zeta_{\text{long}}}\right). \quad (2.24)$$

*The result remains valid when  $\zeta_{\text{long}} = \zeta_{\text{hh}}$  and  $\sigma \leq \tau - 1$  [48, Theorem 2.2(i),(ii)]. When  $\zeta_{\text{long}} = \zeta_{\text{hh}}$  and  $\sigma > \tau - 1$ , the lower bound on the right-hand side remains valid while the upper bound changes to  $-(1/A)k^{1/(\sigma+1-(\tau-1)/\alpha)}$  [48, Theorem 2.2(i),(iii)].*

We proved the bounds for the case  $\zeta_{\text{long}} = \zeta_{\text{hh}}$  in our earlier work [48]: then high-high connections are dominant. Theorem 2.9 extends the result for the phases when  $\zeta_{\star}$  is either  $\zeta_{\text{ll}}$  in (2.9) and/or  $\zeta_{\text{hl}}$  in (2.12). This corresponds to parameter settings where long-range connections incident to at least one vertex of small mark dominate

the expectation in (1.4). The techniques for these latter phases are different from those in [48], see Remark 3.2 on the details.

The lower and upper bound in (2.24) are of the same order if the maximum in  $\mathcal{Z}$  is unique and  $\zeta_{\text{long}} > (d-1)/d$  (excluding the sub-optimal bound for  $\zeta_{\text{long}} = \zeta_{\text{hh}}$  and  $\sigma > \tau - 1$ ). If the maximum is non-unique, i.e.,  $m_{\mathcal{Z}} \in \{2, 3, 4\}$  and  $\zeta_{\text{long}} \geq (d-1)/d$ , then the two bounds differ by a polylog factor. The exponent of  $k$  in the two bounds only differs when  $(d-1)/d$  is *strictly larger* than  $\zeta_{\text{long}}$ . We conjecture that the lower bound is sharp, and that the correct exponent must be  $\zeta_{\text{short}} = (d-1)/d$  generally in the KSRG class in this regime. The additional restriction that  $\sigma \leq \tau - 1$  when  $\zeta_{\text{hh}} = \zeta_{\text{long}}$  comes as a technical assumption from [48].

*Related work on cluster-size decay.* There are many percolation models where the cluster-size decay is determined by surface tension. For Bernoulli nearest neighbor bond percolation on  $\mathbb{Z}^d$ , a series of papers [2, 41, 51, 55] showed that the exponent of the decay is  $\zeta_{\text{short}}$ , i.e., for some  $C_p \geq c_p > 0$ ,

$$\begin{aligned} -C_p &\leq \liminf_{k \rightarrow \infty} \frac{1}{k^{(d-1)/d}} \log \mathbb{P}(k < |\mathcal{C}(0)| < \infty) \\ &\leq \limsup_{k \rightarrow \infty} \frac{1}{k^{(d-1)/d}} \log \mathbb{P}(k < |\mathcal{C}(0)| < \infty) \leq -c_p. \end{aligned} \quad (2.25)$$

For dimension  $d \in \{2, 3\}$ , the constants  $C_p = c_p$  above are known, and determined by the Wulff construction [6, 21]. For Bernoulli percolation on more general, transitive infinite graphs with polynomial ball-growth, [23, 46] obtain similar bounds. In [49], we derived (2.25) for high edge-density long-range percolation when  $\zeta_{\star} = (d-1)/d > 2 - \alpha$ , i.e., when short-range edges are dominant. Our techniques in [49] differ from those here and in [48], see more on the difficulties when  $(d-1)/d = \zeta_{\star}$  in Remark 3.1 below.

*The second-largest component.* Our last result identifies the size  $|\mathcal{C}_n^{(2)}|$  of the second largest component in KSRGs in boxes of volume  $n$ .

**Theorem 2.10** (Second-largest component). *Consider the same setting as in Theorem 2.9. There exist constants  $A, \delta > 0$  such that for all  $n \geq 1$ , with probability at least  $1 - n^{-\delta}$ ,*

$$\left( \frac{\log n}{A(\log \log n)^{m_{\mathcal{Z}}-1}} \right)^{1/\max(\zeta_{\text{long}}, \zeta_{\text{short}})} \leq |\mathcal{C}_n^{(2)}| \leq (A \log n)^{1/\zeta_{\text{long}}}.$$

*The result remains valid when  $\zeta_{\text{long}} = \zeta_{\text{hh}}$  and  $\sigma \leq \tau - 1$  [48, Theorem 2.4(i),(ii)]. When  $\zeta_{\text{long}} = \zeta_{\text{hh}}$  and  $\sigma > \tau - 1$ , the lower bound on the right-hand side remains valid while the upper bound changes to  $(A \log n)^{\sigma+1-(\tau-1)/\alpha}$  [48, Theorem 2.4(i),(iii)].*

Theorem 2.10 accompanies the results from [48, 49] exactly as Theorem 2.9. The lower and upper bounds are matching exactly when the bounds on the cluster-size decay are matching, and differ by a “poly-loglog” factor on phase transition boundaries (excluding the sub-optimal bound for  $\zeta_{\text{long}} = \zeta_{\text{hh}}$  and  $\sigma > \tau - 1$ ). Since the size of  $\mathcal{C}_n^{(2)}$  is polylogarithmic in  $n$ , this theorem implies uniqueness of the giant component.

Moreover, it makes a result by Crawford and Sly [24] sharp: that paper proves a polylogarithmic upper bound on  $|\mathcal{C}_n^{(2)}|$  for long-range percolation with unidentified exponent.

For readers interested in classical long-range percolation models (LRP), we state all our results valid for LRP in a single corollary, gathered from [48, 49] and this paper. Here, we assume that the model is on  $\mathbb{Z}^d$  for some  $d \geq 1$  and two vertices of  $\mathbb{Z}^d$  are connected by an edge with probability that is asymptotically of order  $1/\|x - y\|^{\alpha d}$ , independently of all other edges. We assume  $\alpha \neq 1 + 1/d$  so that lower-order corrections in the lower bounds are not present (for the case  $\alpha = 1 + 1/d$  see the individual theorems above, setting  $m_{\mathcal{Z}} = 2$ ). When we talk about high edge-density below, we assume that the model follows the connectivity function

$$p(x \sim y) = p \min(1, \beta/\|x - y\|)^{d\alpha}, \quad (2.26)$$

where we assume for  $\beta \leq 1$  that  $p\beta^{d\alpha}$  is sufficiently close to 1, and we assume for  $\beta > 1$  that  $\beta$  is sufficiently large (given  $p$ ). In the proof of this corollary we give the exact place of the proof of each result.

**Corollary 2.11** (Summary of results for LRP). *Consider supercritical long-range percolation on  $\mathbb{Z}^d$  with connection probability  $p(x \sim y)$  decaying up to constant factors as  $\|x - y\|^{-\alpha d}$ , and let  $\theta = \mathbb{P}(0 \in \mathcal{C}_\infty) > 0$ . In the model spanned in a box of volume  $n$ , let  $\mathcal{C}_n(0)$ ,  $\mathcal{C}_n^{(1)}$ ,  $\mathcal{C}_n^{(2)}$ , be respectively the connected component of the origin, the largest and the second largest connected component.*

Assume  $\alpha \in (1, 2)$  and  $\alpha \neq 1 + 1/d$ .

Let  $\rho \in (\theta, 1)$ . There exists a constant  $A_0$  such that for all  $n \geq 1$ ,

$$-A_0 n \leq \log \mathbb{P}(|\mathcal{C}_n^{(1)}| > \rho n) \leq -\frac{1}{A_0} n. \quad (2.27)$$

Let  $\rho \in (0, \theta)$ . There exists a constant  $A_1$  such that for all  $n \geq 1$ ,

$$-A_1 n^{\max(2-\alpha, 1-1/d)} \leq \log \mathbb{P}(|\mathcal{C}_n^{(1)}| < \rho n) \leq -\frac{1}{A_1} n^{\max(2-\alpha, 1-1/d)}. \quad (2.28)$$

There exist constants  $A_2, \delta > 0$  such that with probability at least  $1 - n^{-\delta}$ ,

$$\frac{1}{A_2} (\log n)^{1/\max(2-\alpha, 1-1/d)} \leq |\mathcal{C}_n^{(2)}| \leq (A_2 \log n)^{1/(2-\alpha)}. \quad (2.29)$$

There exists a constant  $A_3 > 0$  such that for all  $k \geq 1$  and  $n \in [A_3 k, \infty]$ ,

$$-A_3 k^{\max(2-\alpha, 1-1/d)} \leq \log \mathbb{P}(|\mathcal{C}_n(0)| > k, 0 \notin \mathcal{C}_n^{(1)}) \leq -\frac{1}{A_3} k^{2-\alpha}. \quad (2.30)$$

All lower bounds and the upper bound in (2.27) extend to long-range percolation with  $\alpha \geq 2$ .

Assuming high edge-density and  $\alpha > 1 + 1/d$ , the upper bounds in (2.29) and (2.30) can be sharpened to match the lower bounds for all  $\alpha > 1 + 1/d$ :

$$\frac{1}{A_2} (\log n)^{1/(1-1/d)} \leq |\mathcal{C}_n^{(2)}| \leq (A_2 \log n)^{1/(1-1/d)} \quad (2.31)$$

$$-A_3 k^{1-1/d} \leq \log \mathbb{P}(|C_n(0)| > k, 0 \notin C_n^{(1)}) \leq -\frac{1}{A_3} k^{1-1/d}. \quad (2.32)$$

**Proof** *Large deviations of the giant.* The proof of the upper tail in (2.27) can be found in Section 7.1 of this paper, and the statement is present in Theorem 2.4 (note that both bounds are valid for all  $\alpha > 1$ ). The statement of the upper bound on the lower tail in (2.28) can be found in Theorem 2.7, where the condition  $\zeta_{\text{long}} > 0$  translates to  $\alpha \in (1, 2)$  for LRP. For the proof, see Section 5 (Proposition 5.1 in particular). When the exponent is  $2 - \alpha$ , the proof is in Section 5.1, while for the exponent  $1 - 1/d$ , we refer to Section 5.2. The lower bound of the lower tail in (2.28) is [48, Theorem 2.4], while its proof can be found in Section 7 of the same paper.

*Cluster-size decay.* The upper bound in (2.30) follows from Theorem 2.9 of this paper, and its proof can be found in Section 6.2. Its proof heavily relies on [48, Section 6], where we developed a general technique to extend results on the decay of small components in finite boxes to the cluster-size decay of the infinite model. Moving on to the lower bound on the cluster-size decay in (2.30), the statement itself is in Theorem 2.9, while its proof is in Section 6, after Proposition 6.8 of this paper. The proof there heavily relies on [48, Proposition 7.1, Section 7] where the actual lower bound appears as a localized event: we find, with probability given in the lower bound, a localized component in a box of volume order  $k$  centered at 0.

*Second-largest component.* The bounds on the second-largest component are a consequence of Theorem 2.10. The proof of the lower bound is in Section 6.2, and relies on the same proof as for the cluster-size-decay, namely the localized component in [48, Section 7, in particular Section 7.3]. The proof of the upper bound (2.29) on the second-largest component follows a four-step revelation scheme described in the proof of Proposition 6.1. As the lower bound on the localized component event in [48, Section 7] is valid for all  $\alpha > 1$ , all lower bounds remain valid for  $\alpha \geq 2$ .

*Upper bounds for  $\alpha > 1 + 1/d$ .* Finally, when we assume high edge density and  $\alpha > 1 + 1/d$  (which is actually below 2 for dimensions 2 and higher), the proofs of the upper bounds in (2.32) and (2.31) are the main results of [49]. The proof of (2.31) uses a decomposition of any component into nearest-neighbor connected blocks that are connected by edges longer than 2, then carries out a spanning tree counting method combined with estimates based on isoperimetry. The extension to the upper bound on the cluster-size decay in (2.32) follows the same method as here in Section 6.2, which is essentially [48, Section 6].  $\square$

We believe that many of our proof techniques can be extended to other models with long-range interaction, for instance the random cluster model (FK-percolation) with  $q \geq 1$ . To deal with the additional dependencies coming from boundary conditions in renormalization arguments, techniques from [69] and [30] could serve as a starting point. Extension to FK percolation would match the picture for large deviations of the largest component for Bernoulli bond percolation on  $\mathbb{Z}^d$  obtained by Pisztora [69].

**Notation.** Let  $G = (V, E)$  be a graph. We write  $u \leftrightarrow_G v$  if there exists a path between  $u, v \in V$ , and  $|A|$  for the number of vertices in a set  $A$  of vertices, whereas for two subsets  $V_1, V_2$  containing (marked) vertices we write  $V_1 \sim_G V_2$  if there exist vertices  $v_1 \in V_1, v_2 \in V_2$  such that  $v_1$  is connected by an edge to  $v_2$  (in case  $V_1 = \{v_1\}$  and/or  $V_2 = \{v_2\}$  we write simply  $v_1 \sim_G V_2$ , or  $v_1 \sim_G v_2$ , respectively). We also

write  $V_1 \approx_G V_2$  if there is no edge between vertex pairs in  $V_1 \times V_2$ . If the graph  $G$  is clear from the context we omit the subscript. If  $\mathcal{H} = (\mathcal{V}_{\mathcal{H}}, \mathcal{E}_{\mathcal{H}})$  is (a subgraph of) a KSRG  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ , we write  $\mathcal{H}[a, b]$  for the induced subgraph of  $\mathcal{H}$  on vertices with mark in the half-open interval  $[a, b)$ , denoted by  $\mathcal{V}_{\mathcal{H}}[a, b)$ . For a set  $B \subseteq \mathbb{R}^d$ , we write  $\mathcal{V}_B$  for the vertices with location in  $B$ . If  $B = \Lambda_n = [-n^{1/d}/2, n^{1/d}/2]^d$ , we simply write  $\mathcal{V}_n$ . For two random variables  $X$  and  $Y$ , we write  $Y \preceq X$ , if  $X$  stochastically dominates  $Y$ , that is,  $\mathbb{P}(X \geq x) \geq \mathbb{P}(Y \geq x)$  for all  $x \in \mathbb{R}$ . For two random graphs  $G_1 = (V_1, E_1)$ ,  $G_2 = (V_2, E_2)$  we say that  $G_1$  stochastically dominates  $G_2$  if there exists a coupling such that  $\mathbb{P}(E_1 \supseteq E_2, V_1 \supseteq V_2) = 1$ . A sequence of events  $(\mathcal{A}_n)_{n \geq 1}$  holds with high probability (whp) if  $\mathbb{P}(\mathcal{A}_n) \rightarrow 1$  as  $n \rightarrow \infty$ . We use standard Landau notation for asymptotics of functions: for two not necessarily positive real functions  $f(n), g(n)$  we say that  $f = O(g)$  if  $\limsup_{n \rightarrow \infty} |f(n)/g(n)| < \infty$ ,  $f = o(g)$  if  $\lim_{n \rightarrow \infty} |f(n)/g(n)| = 0$ ,  $f = \Omega(g)$  if  $g = O(f)$ ,  $f = \omega(g)$  if  $g = o(f)$ , and  $f = \Theta(g)$  if both  $f = O(g)$  and  $g = O(f)$ .

### 3 Roadmap of the paper

Theorems 2.5–2.10 are closely linked together. We prove them in (almost) reverse order, in six main steps. All except Step 4 rely on (often multi-scale) renormalization techniques. Here we sketch these steps and explain their relation, the role of the condition  $\zeta_{\text{long}} > 0$ , and the connection to the methods used in our earlier works [48, 49]. This sketch assumes a unit-intensity Poisson point process and power-law vertex marks. In the sections below we make the adaptation for the other settings. We start with the lower tail of large deviations.

#### 3.1 Route to the lower tail of large deviations

The lower bound is given in [48], so we focus on the more involved upper bound. Using  $\zeta_{\text{long}} > 0$ , for  $\rho \in (0, \theta)$  we sketch how to prove

$$\mathbb{P}(|\mathcal{C}_n^{(1)}| < \rho n) \leq \exp(-\Theta(n^{\zeta_*})). \quad (3.1)$$

*Step 1. Existence of a giant (Section 4).* For supercritical settings with  $\max(\zeta_{\text{hl}}, \zeta_{\text{ll}}) > 0$ , we prove for a small constant  $\varepsilon > 0$  that  $\{|\mathcal{C}_n^{(1)}| \geq \varepsilon n\}$  whp. For  $\zeta_{\text{hh}} > 0$  this step is already given by the weak law of large numbers for  $|\mathcal{C}_n^{(1)}|$  in our earlier work [48]. To prove  $\mathbb{P}(|\mathcal{C}_n^{(1)}| \geq \varepsilon n) \rightarrow 1$ , we first adapt a multi-scale renormalization technique by Berger [12] which relies on ergodicity and uniqueness of the infinite component. This technique yields a polynomial lower bound on  $|\mathcal{C}_n^{(1)}|$ . We use this lower bound as the initialization of another multi-scale renormalization scheme, inspired by a method that Biskup introduced for long-range percolation [14]. The two techniques combined yield that for some  $\varepsilon > 0$ ,  $A > 0$ , and  $\gamma \in (0, 1)$  satisfying  $\zeta = 1 - \gamma(\tau - 1)$ ,

$$\mathbb{P}(|\mathcal{C}_n^{(1)}[1, An^\gamma]| < \varepsilon n) \leq \exp(-n^{\max(\zeta_{\text{ll}}, \zeta_{\text{hl}}) - o(1)}), \quad (3.2)$$

where  $\mathcal{C}_n^{(1)}[1, An^\nu)$  denotes the largest connected component in the graph induced on vertices with location in  $\Lambda_n$  and mark in  $[1, An^\nu)$ .

*Step 2. Sharp bounds on the giant's existence (Section 5).* The inequality (3.2) is worse than (3.1). In the next step we improve it: we show the implication that if

$$\Lambda_n \text{ has whp a giant component of density at least } \rho > 0 \quad (\text{Req})$$

$$\text{containing all vertices with mark at least } n^{(1-o(1))/(\tau-1)},$$

then for any  $\varepsilon' > 0$  and some  $A' = A'(\varepsilon') > 0$ ,

$$\mathbb{P}(|\mathcal{C}_n^{(1)}[1, A'n^\nu)| < (\rho - \varepsilon')n) \leq \exp(-\Theta(n^{\max(\zeta_{\text{ll}}, \zeta_{\text{hl}}, \zeta_{\text{hh}}, (d-1)/d)})). \quad (3.3)$$

Compared to (3.2), the  $o(1)$  disappeared from the speed-exponent while  $(d-1)/d$  and  $\zeta_{\text{hh}}$  entered the maximum. To prove (3.3) when  $\zeta_{\text{long}} > (d-1)/d$ , we reduce the event  $\{|\mathcal{C}_n^{(1)}| < (\rho - \varepsilon')n\}$  to a large-deviation event for Erdős-Rényi random graphs, see below Meta-theorem 1.2. When  $(d-1)/d$  is the unique maximum and  $\zeta_{\text{long}} > 0$ , we prove (3.3) with a different renormalization. Using boxes of constant volume  $M$ , the renormalized model dominates supercritical site-bond percolation on  $\mathbb{Z}^d$ . Using long-edges present due to  $\zeta_{\text{long}} > 0$ , two neighboring boxes of volume  $M$  are connected by an edge with probability tending to 1 as  $M \rightarrow \infty$ .

We also prove the condition (Req) of (3.3): all vertices with mark at least  $(\log n)^{\eta'}$  for an explicit  $\eta' > 0$  are in the giant component: using an ‘extra layer’ on the renormalization leading to (3.2) we prove in the case  $\max(\zeta_{\text{hl}}, \zeta_{\text{ll}}) > 0$  that

$$\mathbb{P}(\exists \text{ component } \mathcal{C} \text{ in } \mathcal{G}_n : \mathcal{V}_n[(\log n)^{\eta'}, \infty) \supseteq \mathcal{C}, |\mathcal{C}| \geq \varepsilon n) \rightarrow 1. \quad (3.4)$$

We combine this with (3.3), and set  $\rho = \varepsilon$ ,  $\varepsilon' = \varepsilon/2$  when  $\zeta_{\text{long}} = \max(\zeta_{\text{hl}}, \zeta_{\text{ll}})$ . When  $\zeta_{\text{long}} = \zeta_{\text{hh}}$ , we use the results from [48] and set  $\rho = \theta - \varepsilon$ . In Section 5 we thus prove that for some  $\varepsilon > 0$ ,

$$\mathbb{P}(|\mathcal{C}_n^{(1)}[1, A'n^\nu)| < \varepsilon n) \leq \exp(-\Theta(n^{\zeta_{\star}})). \quad (3.5)$$

After having proven an improved value of  $\rho = \theta - \varepsilon$  in Step 5 below, requirement (Req) shall imply an improved version of (3.5) with the new density.

*Step 3. Non-giant components are small (Section 6.1).* In this step we prove the upper bound on the second-largest component in Theorem 2.10, using (3.5). For KSRGs with PPP vertex sets, for any  $k = k(n)$ , we prove

$$\mathbb{P}(|\mathcal{C}_n^{(2)}| > k) \leq 3(n/k) \exp(-\Theta(k^{\zeta_{\text{long}}})) =: \text{err}_{n,k}. \quad (3.6)$$

Given (3.6), substitute  $k = (A \log n)^{1/\zeta_{\text{long}}}$  for a large constant  $A = A(\delta)$  to obtain the upper bound in Theorem 2.10. The lower bound there follows by verifying a prerequisite from our earlier work [48]. There we also proved (3.6) when  $\zeta_{\text{long}} = \zeta_{\text{hh}}$ . The method in [48] breaks down for the cases when  $\zeta_{\text{long}} \in \{\zeta_{\text{ll}}, \zeta_{\text{hl}}\}$ , see Remark 3.2. So, we use a new method that we explain now.



We sequentially reveal the graph. After the last revealment stage a *non-largest* component of size at least  $k$  is present in  $\Lambda_n$  with probability at most  $\text{err}_{n,k}$ . We partition  $\Lambda_n$  into boxes of volume  $k$ . The probability that all boxes contain a linear-sized component on vertices with mark at most  $A'k^\gamma$  is at least  $1 - \frac{1}{3}\text{err}_{n,k}$  by (3.5). When  $\zeta_{\text{long}} = \zeta_{\text{hl}}$ , we sprinkle in vertices of mark larger than  $A'k^{\gamma_{\text{hl}}}$ , forming the PPP  $\mathcal{V}^{\text{high}}$ . We say that a vertex  $v \in \mathcal{V}^{\text{high}}$  is *box-wedging* if it connects to the local giant in its own box and also to the local giant in the neighboring box. We prove that the number of box-wedging vertices in each box is  $\Theta(k^{\zeta_{\text{long}}})$ : see the back-of-the-envelope calculations in Section 2.2 for intuition on this. So, all local giants exist and are in the same connected component in  $\mathcal{G}_n$  via box-wedging vertices with probability at least  $1 - \frac{2}{3}\text{err}_{n,k}$ . We call this component the *backbone*.

We now take care of small components: in the third stage we reveal edges between vertices that are not in local giants but in different boxes, obtaining  $\mathcal{G}_n^{\text{3rd}}$ . In the fourth revealment stage we reveal edges between the components of  $\mathcal{G}_n^{\text{3rd}}$  and the backbone. If a component of  $\mathcal{G}_n^{\text{3rd}}$  is larger than  $k$ , then we show that with probability at least  $1 - \exp(-\Theta(k^{\zeta_{\text{long}}}))$  it connects to a box-wedging vertex in the backbone. So, all components of size at least  $k$  merge with the backbone with probability at least  $1 - \text{err}_{n,k}$ .

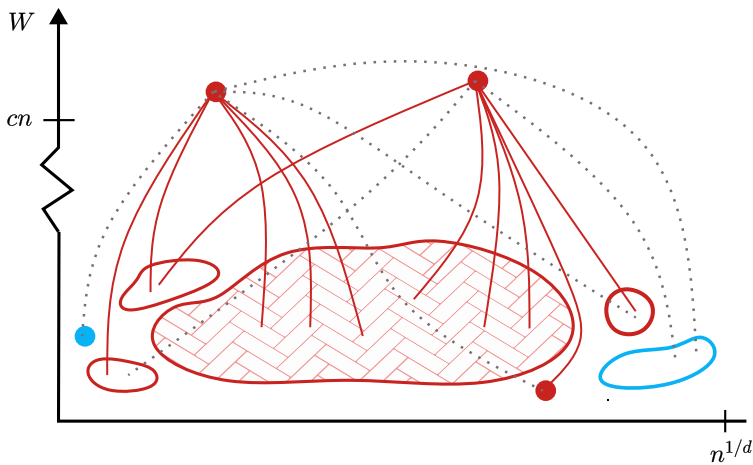
*Step 4. LLN for the giant, cluster-size decay (Section 6.2).* In [48, Proposition 6.1], we showed how the combination of the bounds (3.4), (3.6), and the whp existence of a polynomially-sized largest component (a weaker bound than (3.5)), leads to the upper bound on the cluster-size decay in Theorem 2.9 and the weak LLN

$$|\mathcal{C}_n^{(1)}|/n \xrightarrow{\mathbb{P}} \theta(\beta, p, \alpha, \tau, \sigma), \quad \text{as } n \rightarrow \infty. \quad (3.7)$$

*Step 5. Lower tail of large deviations of the giant (Section 6.3).* The LLN in (3.7) gives that  $|\mathcal{C}_n^{(1)}|/n \geq \theta - \varepsilon$  holds whp for any  $\varepsilon > 0$ . So the condition (Req) holds with  $\rho = \theta - \varepsilon$ . We can reuse (3.3), combine it with (3.4), so that the unique giant contains all highest-mark vertices whp. Then combination with (3.6) results in (3.1).

**Remark 3.1** (Difficulties when surface tension is dominant). When  $(d-1)/d > \zeta_{\text{long}}$ , Theorem 2.7 on the lower tail of LDP is sharp, while Theorems 2.9 and 2.10 are not sharp. In the proofs we connect local giants in subboxes via edges with length of the same order as the box-diameter. These edges are present under the assumption that  $\zeta_{\text{long}} > 0$ . This is enough to give Theorem 2.7: to be able to estimate the giant's size, not every subbox needs to contribute to the global giant. The condition  $\zeta_{\text{long}} > 0$  could only be dropped by developing more information about the position of local giants inside subboxes. They must come close to the boundary to be able to connect them without using long edges. This is beyond the scope of this paper.

For proving Theorems 2.9 and 2.10 with  $\zeta = (d-1)/d$  in the upper bound, we need sharper estimates. Between the giant component and any other component of size at least  $k$ , we need to find at least  $\Omega(k^{(d-1)/d})$  non-connected vertex pairs within constant distance, with probability at least  $1 - n \exp(-\Omega(k^{(d-1)/d}))$ . Our techniques do not focus on the geometry of short missing edges around components, so this case is beyond the scope of this paper. For NNP, such geometric information is known, see [40] for references. For high edge-density long-range percolation our recent work [49]



**Fig. 2** Strategy for the upper tail. The giant induced on lower-mark vertices (filled component) increases in size: hubs (vertices of mark  $\Omega(n)$ ) connect to it by an edge, and also to smaller-size components. The resulting giant component in the entire graph is colored red. If  $p < 1$ , the hubs do not connect by an edge to all vertices (missing edges are dashed in gray), leaving some small components or isolated vertices (blue). The probability for a component of size exactly  $k$  to have no edge to one of at least  $h$  many hubs is  $(1-p)^{kh}$ , so the giant's size increases by roughly  $\sum_k \mathbb{P}(C(0) = k)(1 - (1-p)^{kh})$ . The function hubs in (2.17) gives the minimal number of hubs, so that the size of the giant increases above  $\rho n$  (color figure online)

solves this issue, and obtains matching upper and lower bounds for  $|C_n^{(2)}|$  and the cluster-size decay. For long-range percolation, we refer also to the recent work by Bäumler [11] for related results.

**Remark 3.2** (High-high regime is different). In [48], we proved the bound on  $|C_n^{(2)}|$  in (3.6) when  $\zeta_{hh}$  is maximal. When high-high type connections dominate, the likeliest way to connect two boxes is via edges between two high-mark vertices. So, building a backbone with error probability  $\text{err}_{n,k}$  in (3.6) is more straightforward than the multi-scale renormalization techniques here. However, the sequential revealment method sketched in Step 3 above breaks down when we look at small components of size at least  $k$ . When  $\zeta_{hh} = \zeta_*$ , we also need to show that the giant has  $\Theta(k^{\zeta_{hh}})$  many high-mark vertices within “connection probability”  $\Theta(1)$  to every one of these components, with probability  $\text{err}_{n,k}$ . The revealment method from Step 3 does not guarantee that, since the diameter of the box  $\Theta(k^{1/d})$  is a too crude bound on the distance between previously revealed small components and high-mark vertices. So, in [48] we developed a geometric method that we call *cover-expansion*, which allows us to control this distance based on the geometry of the cluster. See [48, Section 3] for a comparison.

### 3.2 The upper tail of large deviations (Section 7.1)

First we present the intuition behind the formula of  $\text{hubs}(\rho)$  in (2.17), sketched in Figure 2. Afterwards, we outline the proof of Theorem 2.5 and Theorem 2.4.

*Hubs increase the giant linearly.* We investigate an unlikely event. Take  $R$  to be a large constant. When  $h$  vertices with mark at least  $Rn$  – so called *hubs* – are present in a box of volume  $n$ , the size of the giant increases linearly in  $n$ , since each hub connects by an edge to every other vertex in the box with probability  $p$  independently. By the Pareto distribution of the marks in (2.4), the probability of having  $h$  hubs is  $\Theta(n^{-h(\tau-2)})$ . For the exact rate function, we need to compute how the giant's size increases when there are  $h$  hubs present. We reveal the edges in two stages: first all edges between non-hub vertices, obtaining the graph  $\mathcal{G}_n[1, Rn)$ . Then we reveal all edges incident to hubs, obtaining  $\mathcal{G}_n$ . We capture  $|\mathcal{C}_n^{(1)}|$  via its complement: small components of  $\mathcal{G}_n[1, Rn)$  do not merge with the giant in the second step if none of the vertices in the component connect to the hubs. Distinguishing all component sizes, we write

$$\begin{aligned} 1 - \frac{|\mathcal{C}_n^{(1)}|}{n} &\approx \frac{|\mathcal{V}_n \setminus \mathcal{C}_n^{(1)}|}{n} \approx \sum_{v \in \mathcal{V}_n[1, Rn)} \frac{\mathbb{1}_{\{\mathcal{C}_n(v)[1, Rn) \approx \text{hubs}\}}}{n} \\ &= \sum_{\ell \geq 1} \sum_{\substack{C \in \mathcal{G}_n[1, Rn): \\ |C| = \ell}} \frac{\ell \cdot \mathbb{1}_{\{C \approx \text{hubs}\}}}{n}. \end{aligned} \quad (3.8)$$

The probability that no edge exists between a size- $\ell$  component and the  $h$  hubs is  $(1 - p)^{h\ell}$ . Moreover, by local convergence (see [72] for a detailed explanation and references), the proportion of vertices in components of size- $\ell$  converges in probability to  $\mathbb{P}^0(|\mathcal{C}(0)| = \ell)$ . Therefore, the total number of size- $\ell$  components in  $\Lambda_n$  concentrates around  $(n/\ell) \cdot \mathbb{P}^0(|\mathcal{C}(0)| = \ell)$ . Hence,

$$1 - \frac{|\mathcal{C}_n^{(1)}|}{n} \approx \sum_{\ell \geq 1} \mathbb{P}^0(|\mathcal{C}(0)| = \ell) \cdot (1 - p)^{h\ell} = \mathbb{E}^0[(1 - p)^{h|\mathcal{C}(0)|}]. \quad (3.9)$$

On the event  $\{|\mathcal{C}_n^{(1)}|/n > \rho\}$ , the right-hand side above should be at most  $1 - \rho$ . Thus, letting  $h$  to be the smallest integer satisfying

$$\mathbb{E}^0[(1 - p)^{h|\mathcal{C}(0)|}] \leq 1 - \rho, \quad (3.10)$$

the presence of  $h$  hubs implies the event  $\{|\mathcal{C}_n^{(1)}|/n > \rho\}$ . The solution to (3.10) with equality is exactly  $\text{hubs}(\rho)$  in (2.17).

*Concentration on the number of size- $\ell$  components.* To move from (3.8) to (3.9), we need that the rate of concentration of the number of size- $\ell$  components is  $o(n^{-(\tau-2)h})$ . Local convergence (LC) does not provide any convergence rate. We boost the performance of LC by applying it to disjoint subboxes of  $\Lambda_n$ , binomial concentration, and a bound on the number of components containing edges that cross box-boundaries. When vertex-marks are at most a linear function in  $n$ , this yields a polynomial convergence rate with a tunable exponent.

*Linear-scale hubs are required to increase the giant's size.* For the lower bound in (2.19), we use that the graph without hubs contains a giant component of density almost  $\theta$  by the lower tail of large deviations. The hubs connect to this giant with error probability exponential in  $n$ . When  $\text{hubs}(\rho) \notin \mathbb{N}$ , our concentration bound on

size- $\ell$ -component is strong enough to guarantee that the presence of  $\lceil \text{hubs}(\rho) \rceil$  hubs increases the giant's size to a value above  $\rho n$ . We also show that any event that has fewer than  $\lceil \text{hubs}(\rho) \rceil$  hubs means that too many size- $\ell$  components remain outside of the giant with probability at least  $\Theta(n^{-(\tau-2)h})$ . So, the complement of the giant remains larger than  $(1 - \rho)n$  on such events. This gives the upper bound.

*Homogeneous percolation models.* In Bernoulli bond percolation or long-range percolation on  $\mathbb{Z}^d$ , the degree-distribution has (super-)exponential tails, so no linear-scale hubs are present. We show that the number of size- $\ell$  components concentrate with error probability exponentially small in  $n$  for each constant  $\ell \in \mathbb{N}$ . We capture the giant's size by its complement, and Theorem 2.4 follows.

## 4 Existence of a giant

The main goal of this section is to prove the existence of a giant, i.e., a linear-sized component, for the whole supercritical regime when  $\max(\zeta_{\text{hl}}, \zeta_{\text{ll}}) > 0$ . For  $\zeta_{\text{hh}} > 0$ , we already proved existence in our recent work [48]. The first proposition corresponds to (3.2) in the roadmap in Section 3.

**Proposition 4.1** (Existence of a giant). *Consider a supercritical KSRG satisfying Assumption 2.2 with a Poisson point process as vertex set. Assume that the parameters satisfy  $\max(\zeta_{\text{hl}}, \zeta_{\text{ll}}) > 0$ . For each constant  $\delta > 0$ , there exists a constant  $A > 0$  such that for all  $n \geq 1$ ,*

$$(hl) \text{ when } \zeta_{\text{hl}} > 0, \quad \mathbb{P}(|\mathcal{C}_n^{(1)}[1, An^{\zeta_{\text{hl}}}]| \leq \tfrac{1}{A}n) \leq \exp\left(-\tfrac{1}{A}n^{\zeta_{\text{hl}}-\delta}\right), \quad (4.1)$$

$$(ll) \text{ when } \zeta_{\text{ll}} > 0, \quad \mathbb{P}(|\mathcal{C}_n^{(1)}[1, A]| \leq \tfrac{1}{A}n) \leq \exp\left(-\tfrac{1}{A}n^{\zeta_{\text{ll}}-\delta}\right). \quad (4.2)$$

The statements (4.1) and (4.2) remain valid for the Palm version  $\mathbb{P}^x$  of  $\mathbb{P}$  for any  $x \in \mathbb{R}^d$ .

Proposition 4.1 generalizes [14, Theorem 3.2] by Biskup from long-range percolation on  $\mathbb{Z}^d$  to a wider class of KSRGs. The presence of vertex marks in KSRGs, and the truncation of them in the statement of Proposition 4.1 require new ideas. We will prove Proposition 4.1 via a multi-scale renormalization scheme. To initialize this scheme, we need an initial lower bound on  $|\mathcal{C}_n^{(1)}|$ .

**Lemma 4.2** (Existence of an “almost”-giant). *Consider a supercritical KSRG under the same setting as in Proposition 4.1. For all  $\varepsilon > 0$*

$$\mathbb{P}(|\mathcal{C}_n^{(1)}| \leq n^{1-\varepsilon}) \longrightarrow 0, \quad \text{as } n \rightarrow \infty. \quad (4.3)$$

The statement remains valid for the Palm version  $\mathbb{P}^x$  of  $\mathbb{P}$  for any  $x \in \mathbb{R}^d$ .

The proof of this lemma is inspired by a multi-scale renormalization technique introduced for long-range percolation by Berger in [12]. When  $\zeta_{\text{ll}} = 2 - \alpha > 0$  and the vertex set is formed by a PPP, [37, Proposition 3.9] already proves (4.3) using

an adaptation of [12]. For the case  $\zeta_{hl} > 0$  we give a new proof in Appendix B that explicitly relies on the vertex-marks. The rest of this section is devoted to proving Proposition 4.1. We first introduce some additional preliminaries.

#### 4.1 Preliminaries

We describe a sprinkling technique that enables us to reveal the vertex set of a KSRG on a PPP in multiple stages, so that the graph is supercritical already before sprinkling.

**Definition 4.3** (Scaled-thinned KSRG on a PPP). Consider a KSRG on a PPP. Let  $q \in [0, 1]$  and  $\eta > 0$ . We write  $\eta\mathcal{G}_n^{\beta,q}$  for a scaled-thinned KSRG as follows: first, we keep each vertex of the PPP  $\mathcal{V}$  independently with probability  $q$ . Then we sample the edges according to Definition 2.1 with edge-density parameter  $\beta > 0$ . Finally, we assign a new location to each retained vertex, i.e., we map  $u = (x_u, w_u) \in \mathbb{R}^d \times [1, \infty)$  to  $(\eta x_u, w_u)$ .

It is not hard to see that the scaled-thinned KSRG with  $\eta = q$  is distributed like the original graph with a smaller parameter  $\beta$  in a smaller box. This is the content of the following observation.

**Observation 4.4** (Sprinkling trick-A). Consider a supercritical KSRG on a PPP, i.e., with  $\beta > \beta_c(p, \alpha, \tau, \sigma)$ . Let  $q \in (0, 1]$ . Then the following distributional identity holds:

$$q\mathcal{G}_n^{\beta,q} \stackrel{d}{=} \mathcal{G}_{qn}^{\beta q^d, 1} \stackrel{d}{=} \mathcal{G}_{qn}^{\beta q^d} \quad (4.4)$$

In particular, the graph  $q\mathcal{G}_n^{\beta,q}$  is supercritical whenever the new edge-intensity  $\beta q^d > \beta_c$ .

**Proof** Denote by  $\mathcal{V}^{(q,q)}$  the marked vertex set of  $q\mathcal{G}_n^{\beta,q}$ . Since the projection of  $\mathcal{V}$  onto the spatial dimensions in  $\mathbb{R}^d$  is a PPP with intensity being the Lebesgue measure, thinning  $\mathcal{V}$  gives intensity  $q\text{Leb}(\cdot)$ . Then, re-scaling with  $\eta = q$  maps the box  $\Lambda_n$  to  $\Lambda_{qn}$ . For any set  $A \subseteq \Lambda_{qn}$ , the expected number of points in it is  $q\text{Leb}(\eta^{-1}A) = \text{Leb}(A)$ . Hence  $\mathcal{V}^{(q,q)}$  has the same distribution as  $\mathcal{V} \cap \Lambda_{qn}$ , since the spatial intensity of the new point process  $\mathcal{V}^{(q,q)}$  is  $\text{Leb}(\cdot)$  again, and the distribution of the mark-coordinate is unchanged. We now inspect the connection probabilities in  $q\mathcal{G}_n^{\beta,q}$ . Let us denote the new connectivity function by  $p^{(q,q)}$ . Consider two points  $u = (x_u, w_u)$  and  $v = (x_v, w_v)$  in  $\mathcal{V}^{(q,q)}$ . Since we connect  $u$  to  $v$  by an edge in  $q\mathcal{G}_n^{\beta,q}$  iff we connect their inverse points  $(q^{-1}x_u, w_u)$  and  $(q^{-1}x_v, w_v)$  in  $\mathcal{G}_n$  by an edge, the new connection probability  $p^{(q,q)}$  is given by

$$p^{(q,q)}((x_u, w_u), (x_v, w_v)) = p((q^{-1}x_u, w_u), (q^{-1}x_v, w_v)) = \varphi\left(\beta \frac{\kappa(w_u, w_v)}{\|q^{-1}(x_u - x_v)\|^d}\right).$$

Thus, the new graph has the same distribution as a graph on  $\mathcal{V} \cap \Lambda_{qn}$  with the same connection probabilities but new edge-density parameter  $\beta' = \beta q^d$ . This graph is denoted by  $\mathcal{G}_{qn}^{\beta', 1}$ , since in this latter notation the thinning factor is  $q = 1$ , and the

box is  $\Lambda_{qn}$ . The last identity in (4.4) says that omitting  $q = 1$  from the superscript is consistent with the notation. Finally, the new graph is supercritical whenever the new edge-intensity  $\beta q^d$  is at least  $\beta_c$  as well.  $\square$

Throughout the paper we will use the following construction of KSRGs when we consider PPPs as the underlying vertex set, following the setting of Assumption 2.2.

**Definition 4.5** (Alternative KSRG construction on a PPP). Consider a KSRG on a PPP with parameters and kernel  $\kappa$  from Assumption 2.2. If  $\beta > \beta_c$ , let  $q \in (0, 1]$  be such that  $\beta q^d > \beta_c$ , i.e., the scaled-thinned graph  $q\mathcal{G}_n^{\beta, q}$  in (4.4) is supercritical. If  $\beta \leq \beta_c$ , set  $q = 1$ . Let  $\mathcal{V}^{\text{base}}$  and  $\mathcal{V}^{\text{spr}}$  be two independent PPPs with intensity  $q\text{Leb} \times F_W(dw)$  and  $(1 - q)\text{Leb} \times F_W(dw)$ , respectively, with  $F_W$  as in Assumption 2.2. Define

$$\bar{\mathcal{V}} := \mathcal{V}^{\text{base}} \cup \mathcal{V}^{\text{spr}},$$

which is equal in distribution to a Poisson point process on  $\mathbb{R}^d \times [1, \infty)$  with intensity

$$\text{Leb} \times F_W(dw) = \begin{cases} dx \times (\tau - 1)w^{-\tau}dw, & \text{if } \tau < \infty, \\ dx \times \delta_1(dw), & \text{if } \tau = \infty, \end{cases} \quad (4.5)$$

where  $\delta_1(\cdot)$  denotes a Dirac measure at one. Conditionally on the realization of  $\bar{\mathcal{V}}$ , let  $\bar{\mathcal{G}}_n^\beta$  be the graph on  $\bar{\mathcal{V}}$  where two vertices  $u, v \in \bar{\mathcal{V}}$  are connected by an edge independently with probability

$$p(u, v) = \begin{cases} p\left(\beta \cdot \frac{\kappa(w_u, w_v)}{\|x_u - x_v\|^d} \wedge 1\right)^\alpha, & \text{if } \alpha < \infty, \\ p\mathbb{1}\{\beta\kappa(w_u, w_v) \geq \|x_u - x_v\|^d\}, & \text{if } \alpha = \infty. \end{cases} \quad (4.6)$$

Let  $\mathcal{G}_n^{\beta, \text{base}} \stackrel{d}{=} \mathcal{G}_n^{\beta, q}$  be the induced subgraph on  $\mathcal{V}^{\text{base}}$ . If  $q < 1$ , we call  $\mathcal{G}_n^{\beta, \text{base}}$  the *thinned* KSRG. We omit the superscript  $\beta$  if it is clear from the context.

**Observation 4.6** (Sprinkling trick-B). Consider a supercritical KSRG on a PPP, and let  $\bar{\mathcal{G}}_n^\beta$  and  $\mathcal{G}_n^{\beta, \text{base}}$  be as in Definition 4.5. Write  $\beta' := \beta q^d$ . Let  $\mathcal{C}_{n, \text{base}}^{(1)}$  and  $\mathcal{C}_{qn, \beta'}^{(1)}$  be the largest component in  $\mathcal{G}_n^{\beta, \text{base}}$  and  $\mathcal{G}_{qn}^{\beta'}$ , respectively. Then

$$\bar{\mathcal{G}}_n^\beta \stackrel{d}{=} \mathcal{G}_n^\beta, \quad \mathcal{G}_n^{\beta, \text{base}} \stackrel{d}{=} \frac{1}{q}\mathcal{G}_{qn}^{\beta'}, \quad |\mathcal{C}_{qn, \beta'}^{(1)}| \stackrel{d}{=} |\mathcal{C}_{n, \text{base}}^{(1)}|.$$

**Proof**  $\bar{\mathcal{G}}_n^\beta$  has the same law as  $\mathcal{G}_n^\beta$  since  $\bar{\mathcal{V}}$  has the same law as  $\mathcal{V}$  and the connection probabilities are the same.  $\mathcal{G}_n^{\beta, \text{base}}$  has the same law as  $1\mathcal{G}_n^{\beta, q}$  in Definition 4.3, i.e., we just apply thinning but do not rescale. We write this now as  $q^{-1}(q\mathcal{G}_n^{\beta, q})$  and apply the identity (4.4) to the scaled-thinned graph, and we get the second distributional identity above. The last identity about the largest component holds since edges (and thus connected components) do not change under rescaling.  $\square$

Let us also state a well-known statement about the connectivity of Erdős-Rényi random graphs. We give a short (suboptimal) proof for the sake of completeness.

**Claim 4.7** (Connectivity of  $G(n, p_n)$ ). *Consider the Erdős-Rényi random graph  $G(n, p_n)$  with  $np_n = \omega(\log n)$ . For all sufficiently large  $n$ ,*

$$\mathbb{P}(G(n, p_n) \text{ is not connected}) \leq 3(en)^2(1 - p_n)^{n/2}. \quad (4.7)$$

**Proof** If  $G(n, p_n)$  is not connected, then there must exist a connected component of size  $k$ , for some  $1 \leq k \leq \lfloor n/2 \rfloor$ . For  $k = 1$ , this probability is at most  $n(1 - p_n)^{n-1} \leq (en)^2(1 - p_n)^{n/2}$  by a union bound over all vertices. For  $k \geq 2$ , by Cayley's formula, there are  $k^{k-2}$  spanning trees on  $k$  vertices, and a connected component of size  $k$  has to contain such a tree. So by a union bound over the spanning trees, the probability that a connected component of size between 2 and  $n/2$  exists without being connected to the other  $n - k$  vertices, is at most

$$\sum_{k=2}^{\lfloor n/2 \rfloor} \binom{n}{k} k^{k-2} (1 - p_n)^{k(n-k)} \leq \sum_{k=2}^{\lfloor n/2 \rfloor} \left(\frac{en}{k}\right)^k k^{k-2} (1 - p_n)^{k(n-k)} \leq \sum_{k=2}^{\lfloor n/2 \rfloor} (en(1 - p_n)^{n-k})^k.$$

Using  $n - k \geq \lceil n/2 \rceil$ , we get that the right-hand side is bounded from above by

$$\sum_{k=2}^{\lfloor n/2 \rfloor} (en(1 - p_n)^{\lceil n/2 \rceil})^k \leq \frac{(en)^2(1 - p_n)^{2\lceil n/2 \rceil}}{1 - en(1 - p_n)^{\lceil n/2 \rceil}} \leq 2(en)^2(1 - p_n)^{n/2}.$$

Combining the cases of  $k = 1$  and  $k \geq 2$  we obtain the desired statement.  $\square$

## 4.2 A linear-sized component (high-low regime)

We turn to the proof of Proposition 4.1 for the hl-regime, so we restrict to KSRGs on PPPs. We describe the adjustments for the ll-regime at the end of this section. A sketch of the argument can be found in Figure 3.

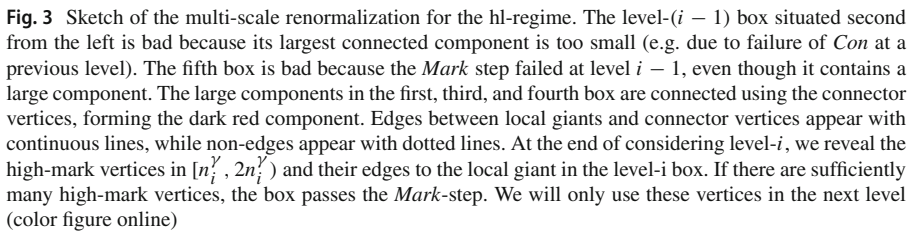
### 4.2.1 Multi-scale renormalization – sketch

We set up an iterative boxing scheme. The initial box size is a constant  $n_0$ . We check if the largest component of a thinned KSRGs inside a level-0 box has size at least  $\rho_0 n_0 := n_0^{1-\varepsilon}$ . Next, we sprinkle vertices of sufficiently high (but constant) mark that connect to this largest component. If both of these steps succeed, we call a level-0 box good. This happens with constant probability using Lemma 4.2.

For each iteration  $i \geq 1$ , we group  $m_i$  level- $(i - 1)$  boxes of volume  $n_{i-1}$  into a larger level- $i$  box  $B$  of volume  $n_i := m_i n_{i-1}$ . We always work with only the ‘good’ subboxes of the previous level  $i - 1$  inside  $B$  to determine whether the level- $i$  box  $B$  is good, and control the probabilities of the following three events:

- (Sub) Using independence across subboxes, at least  $(1 - \varepsilon_i)m_i$  of the level- $(i - 1)$  subboxes of volume  $n_{i-1}$  of  $B$  are level- $(i - 1)$ -good with sufficiently large probability. That is, the subboxes satisfy the next events (Con), (Mark) at level  $i - 1$ : they have a local component of density at least  $\rho_{i-1}$ , which contains enough level- $(i - 1)$  high-mark vertices.





(Mark) The sprinkling trick from Section 4.1 allows us to sprinkle level- $i$  high-mark vertices that were not exposed in previous levels. We make sure enough of them connect to the large connected component in  $B$  built in the *Con*-step. We will then use these vertices in the *Con*-step of level  $i + 1$  later.

 Springer

#### 4.2.2 Formal setup of the renormalization

We formalize the boxing scheme with all its parameters. Fix any small constant  $\delta > 0$  that appears in the statement of Proposition 4.1. We define an extra constant, that shall describe the growth rate of the box sizes  $(n_i)_{i \geq 1}$ :

$$\xi_\delta := \xi_\delta(\zeta) = \frac{\delta/2}{\zeta + (\alpha - 1) - \delta/2}. \quad (4.8)$$

The value of  $\xi_\delta$  is defined both for  $\zeta_{\text{hl}} = (\tau - 1)/\alpha - (\tau - 2)$  and for  $\zeta_{\text{ll}} = 2 - \alpha$  whenever these are positive. Since  $\alpha > 1$ ,  $\xi_\delta > 0$  for all  $\delta < \zeta$ . Unless explicitly stated otherwise, we abbreviate in the rest of the section

$$\gamma = \gamma_{\text{hl}} = 1 - 1/\alpha, \quad \zeta = \zeta_{\text{hl}} = (\tau - 1)/\alpha - (\tau - 2). \quad (4.9)$$

We recall the PPPs  $\mathcal{V}^{\text{base}}$  and  $\mathcal{V}^{\text{spr}}$  with thinning constant  $q < 1$  from Definition 4.5. We will choose any  $q$  with  $\beta q^d > \beta_c$ . We recall the parameters  $p, \beta$  from the connectivity function in (4.6). We define one extra constant depending on  $q$ :

$$\delta' := \frac{1}{4}(1 - q)(1 - 2^{-(\tau-1)}) \cdot p\beta^\alpha d^{-\alpha d/2}. \quad (4.10)$$

Slightly abusing notation, if  $\mathcal{V}$  is a set of vertices and  $\mathcal{C}$  is a component, we write  $\mathcal{V} \cap \mathcal{C}$  for the set of vertices in  $\mathcal{V}$  that belong to the component  $\mathcal{C}$ .

**Definition 4.8** (renormalization scheme). Consider a KSRG under the same setting as in Proposition 4.1. Fix  $\zeta > 0$ , and  $\delta \in (0, \zeta/2)$ , and  $\xi_\delta(\zeta)$  as in (4.8), and a thinning constant  $q < 1$  with  $\beta q^d > \beta_c$ . Let  $n_0$  be a large constant. We iteratively define for  $i \geq 1$

$$m_i := \left\lfloor n_{i-1}^{\xi_\delta/d} \right\rfloor^d, \quad n_i := m_i n_{i-1}. \quad (4.11)$$

Let  $\mathcal{Q}_i(x)$  denote a box centered at  $x \in \mathbb{R}^d$  of volume  $n_i$ . We define the level- $i$  vertex set of  $\mathcal{Q}_i(x)$  to be

$$\widehat{\mathcal{V}}_i := \mathcal{V}_{\mathcal{Q}_i(x)}^{\text{base}}[1, n_0^{2/(\tau-1)}) \cup \mathcal{V}_{\mathcal{Q}_i(x)}^{\text{spr}}[1, 2n_i^\gamma]. \quad (4.12)$$

We write  $\widehat{\mathcal{G}}_{n_i}$  for the induced subgraph on  $\widehat{\mathcal{V}}_i$ . Define the sequences

$$\varepsilon_i := \begin{cases} 1 - n_0^{-\delta/4}, & \text{if } i = 0, \\ (i+1)^{-2}, & \text{if } i \geq 1, \end{cases} \quad (4.13)$$

$$\rho_i := \prod_{j=0}^i (1 - \varepsilon_j) = n_0^{-\delta/4} \prod_{j=2}^{i+1} (1 - j^{-2}), \quad \rho := \lim_{j \rightarrow \infty} \rho_j. \quad (4.14)$$

We call  $\mathcal{Q}_i(x)$   $i$ -good (and  $i$ -bad otherwise) if it satisfies the event (with  $\delta'$  in (4.10))

$$\mathcal{A}_{i\text{-good}}(x) := \left\{ \exists \text{comp. } \mathcal{C}_i \text{ in } \widehat{\mathcal{G}}_{n_i} : |\mathcal{C}_i| \geq \rho_i n_i, \right. \\ \left. |\mathcal{V}_{\mathcal{Q}_i(x)}^{\text{spr}}[n_i^\gamma, 2n_i^\gamma] \cap \mathcal{C}_i| \geq \delta' \rho_i n_i^\zeta \right\}. \quad (4.15)$$

The definitions of  $(m_i)_{i \geq 0}$  and  $(n_i)_{i \geq 0}$  in (4.11) ensure that they increase doubly exponentially in  $i$ : indeed, ignoring integer parts,  $n_i \approx n_{i-1}^{1+\xi_\delta}$  so  $n_i \approx n_0^{(1+\xi_\delta)^i}$ . This also implies that  $2n_{i-1}^\gamma < n_i^\gamma$ , and so for any level- $(i-1)$  subbox centered at any  $y$ ,  $\widehat{\mathcal{V}}_{i-1,y}$  is disjoint of  $\mathcal{V}_{\mathcal{Q}_i(x)}^{\text{spr}}[n_i^\gamma, 2n_i^\gamma]$ . Thus, in  $\mathcal{Q}_i(x)$  we sprinkle in *new* vertices with mark in  $[n_i^\gamma, 2n_i^\gamma]$  in level- $i$  compared to previous levels. Summability of  $(\varepsilon_i)_{i \geq 0}$  in (4.13) implies that  $\rho > 0$  for any starting value  $\varepsilon_0(n_0)$  in (4.13–4.14).

The next claim provides useful bounds on the sequences that will appear in subsequent computations. Actually,  $\xi_\delta$  in (4.8) is chosen precisely such that (4.16)–(4.18) are satisfied below. The label on the inequalities indicate where we will use the bounds later.

**Claim 4.9** (Useful bounds). *Fix any  $\zeta \in (0, 1)$  and then  $\xi_\delta(\zeta)$  in (4.8). Assume  $\delta \in (0, \zeta/2)$  and consider the quantities in Definition 4.8. For each constant  $C > 0$ , there exists a constant  $n_0^* > 0$  such that for all  $n_0 \geq n_0^*$  and all  $i \geq 1$ , the following hold:*

$$\varepsilon_i m_i n_{i-1}^{\zeta-\delta} \stackrel{\text{(Sub)}}{\geq} C n_i^{\zeta-\delta}, \quad (4.16)$$

$$\rho_{i-1}^2 m_i^{1-\alpha} n_{i-1}^\zeta \stackrel{\text{(Con)}}{\geq} C n_i^{\zeta-\delta}, \quad (4.17)$$

$$\rho_i n_i^\zeta \stackrel{\text{(Mark)}}{\geq} C n_i^{\zeta-\delta}. \quad (4.18)$$

Further,

$$n_i \in \left[ n_0^{(1+\xi_\delta/2)^i}, n_0^{(1+\xi_\delta)^i} \right], \quad m_i \in \left[ \frac{1}{2} n_0^{\xi_\delta(1+\xi_\delta/2)^{i-1}}, n_0^{\xi_\delta(1+\xi_\delta)^{i-1}} \right], \quad (4.19)$$

$$\varepsilon_i m_i^{1-\zeta+\delta} \geq C, \quad (n_{i-1}/n_0)^{\zeta-\delta} \geq C \log(e/\varepsilon_i). \quad (4.20)$$

Moreover, for all  $m \leq m_{i+1}$  and  $n := n_i m$ ,

$$m n_i^{\zeta+\alpha} n^{-\alpha} \geq n^{\zeta-\delta}. \quad (4.21)$$

Finally, when  $\zeta = \zeta_{\text{hl}}$  and  $\gamma = 1 - 1/\alpha$ , then

$$n_i^\gamma > 2n_{i-1}^\gamma. \quad (4.22)$$

We postpone the proof to Appendix B.2.

### 4.2.3 Induction base of the renormalization

We now carry out the multi-scale renormalization outlined in Section 4.2.1 by induction. We first constitute the induction base.

**Claim 4.10** (Induction base). *Consider a supercritical KSRG under the setting as in Proposition 4.1 and consider Definition 4.8. There exists a constant  $n_0^* > 0$  such that for all  $n_0 \geq n_0^*$  in Definition 4.8, and all  $x \in \mathbb{R}^d$ ,*

$$\mathbb{P}(\mathcal{Q}_0(x) \text{ is 0-bad}) \leq 1/e^6.$$

**Proof** By translation invariance, we assume w.l.o.g. that  $x = 0$ , and abbreviate  $\mathcal{Q}_0 := \mathcal{Q}_0(0)$ . We use the definition of  $i$ -goodness in (4.15) for  $i = 0$  in Definition 4.8, and substitute  $\rho_0 = n_0^{-\delta/4}$  by (4.14). Then,

$$\begin{aligned} \mathbb{P}(\mathcal{Q}_0 \text{ is 0-good}) &= \mathbb{P}(\exists \text{ component } \mathcal{C} \text{ in } \widehat{\mathcal{G}}_{n_0} : |\mathcal{C}| \geq n_0^{1-\delta/4}, \\ &\quad |\mathcal{C} \cap \mathcal{V}^{\text{pr}}[n_0^\gamma, 2n_0^\gamma]| \geq \delta' n_0^{-\delta/4} n_0^\zeta). \end{aligned} \quad (4.23)$$

Following Definition 4.5, we use the alternative KSRG construction with the  $q$ -thinning of the vertex set. Let  $\mathcal{G}_{n_0}^{\text{base}}$  be the thinned KSRG restricted to  $\mathcal{Q}_0$ , which is a subgraph of  $\widehat{\mathcal{G}}_{n_0}$ ; let  $\mathcal{C}_{n_0, \text{base}}^{(1)} \left[1, n_0^{2/(\tau-1)}\right)$  be the largest component in the induced subgraph on the restricted marks  $\left[1, n_0^{2/(\tau-1)}\right)$  in  $\mathcal{G}_{n_0}^{\text{base}}$ ; and let  $\mathcal{C}_0 \supseteq \mathcal{C}_{n_0, \text{base}}^{(1)} \left[1, n_0^{2/(\tau-1)}\right)$  be the component in the whole  $\widehat{\mathcal{G}}_{n_0}$  containing this component. We define the event that this latter component is large enough:

$$\mathcal{A}_{\text{large}}^{n_0} := \{|\mathcal{C}_{n_0, \text{base}}^{(1)} \left[1, n_0^{2/(\tau-1)}\right)| \geq n_0^{1-\delta/4}\}. \quad (4.24)$$

Conditionally on  $\mathcal{A}_{\text{large}}^{n_0}$ , the component  $\mathcal{C}_0$  is large enough to play the role of  $\mathcal{C}$  in (4.23). So, by the law of total probability,

$$\mathbb{P}(\mathcal{Q}_0 \text{ is 0-good}) \geq \mathbb{P}(\mathcal{A}_{\text{large}}^{n_0}) \cdot \mathbb{P}(|\mathcal{C}_0 \cap \mathcal{V}_{\mathcal{Q}_0}^{\text{pr}}[n_0^\gamma, 2n_0^\gamma]| \geq \delta' n_0^{-\delta/4} n_0^\zeta \mid \mathcal{A}_{\text{large}}^{n_0}). \quad (4.25)$$

We show that both factors on the right-hand side tend to one as  $n_0 \rightarrow \infty$ .

(Con)-step of level-0. We first consider  $\mathbb{P}(\mathcal{A}_{\text{large}}^{n_0})$  with the event defined in (4.24).

If the thinned PPP  $\mathcal{V}_{n_0}^{\text{base}}$  contains no vertices of mark at least  $n_0^{2/(\tau-1)}$ , then  $\mathcal{C}_{n_0, \text{base}}^{(1)} \left[1, n_0^{2/(\tau-1)}\right)$  and  $\mathcal{C}_{n_0, \text{base}}^{(1)}$  coincide. Here  $\mathcal{C}_{n_0, \text{base}}^{(1)}$  is the largest component of the thinned PPP in  $\mathcal{Q}_0$ . By a union bound over the complementary events, it follows that

$$\begin{aligned} \mathbb{P}(\mathcal{A}_{\text{large}}^{n_0}) &\geq \mathbb{P}(|\mathcal{V}_{\mathcal{Q}_0}^{\text{base}}[n_0^{2/(\tau-1)}, \infty)| = 0, |\mathcal{C}_{n_0, \text{base}}^{(1)}| \geq n_0^{1-\delta/4},) \\ &\geq 1 - \mathbb{P}(|\mathcal{C}_{n_0, \text{base}}^{(1)}| \leq n_0^{1-\delta/4}) - \mathbb{P}(|\mathcal{V}_{\mathcal{Q}_0}^{\text{base}}[n_0^{2/(\tau-1)}, \infty)| > 0). \end{aligned} \quad (4.26)$$

By Observation 4.6,  $C_{n_0, \text{base}}^{(1)}$  has the same distribution as the largest component of a supercritical KSRG with edge density  $\beta q^d > \beta_c$  inside the box  $\Lambda_{qn_0}$ , with  $q \in (0, 1)$  from Definition 4.5. Hence, by Lemma 4.2, the first probability on the right-hand side in (4.26) tends to zero as  $n_0 q \rightarrow \infty$  (we absorb the factor  $q$  into the exponent  $1 - \delta/4$  by choosing  $\varepsilon < \delta/4$  in Lemma 4.2). The intensity of  $\mathcal{V}^{\text{base}}$  is  $q \text{Leb} \times F_W(dw) = q \text{Leb} \times (\tau - 1)w^{-\tau}dw$  and the box has volume  $n_0$ . So, the expected number of vertices in  $\mathcal{V}_{Q_0}^{\text{base}}[n_0^{2/(\tau-1)}, \infty)$  is  $\Theta(n_0^1 n_0^{-2})$ , which tends to zero. Hence, the second error term in (4.26) also tends to zero. This shows that  $\mathbb{P}(\mathcal{A}_{\text{large}}^{n_0}) \rightarrow 1$  in (4.25).

(Mark) *step of level-0.* We now carry out the (Mark)-step of level 0: We inspect the second factor in (4.25). The vertices  $\mathcal{V}_{Q_0}^{\text{spr}}[n_0^\gamma, 2n_0^\gamma)$  form a PPP with intensity  $(1 - q) \text{Leb} \times (\tau - 1)w^{-\tau}dw$  in a box with volume  $n_0$ . Its expected size is

$$\mathbb{E}[|\mathcal{V}_{Q_0}^{\text{spr}}[n_0^\gamma, 2n_0^\gamma)|] = (1 - q)(1 - 2^{-(\tau-1)})n_0^{1-\gamma(\tau-1)} = (1 - q)(1 - 2^{-(\tau-1)})n_0^\xi. \quad (4.27)$$

Further,  $\mathcal{V}^{\text{spr}}$  is independent of  $\mathcal{V}^{\text{base}}$  and thus also of the graph  $\mathcal{G}_{n_0}^{\text{base}}[1, n_0^{2/(\tau-1)})$ , the graph induced by vertices whose marks are in the interval  $[1, n_0^{2/(\tau-1)})$  and the event  $\mathcal{A}_{\text{large}}^{n_0}$ . On this event, there are at least  $n_0^{1-\delta/4}$  vertices in the above mentioned component, and each vertex there has mark at least one. We use that the maximal distance between two vertices in  $Q_0$  is  $\sqrt{d}n_0^{1/d}$ . Given  $\mathcal{A}_{\text{large}}^{n_0}$ , for each vertex  $v$  in  $\mathcal{V}_{Q_0}^{\text{spr}}[n_0^\gamma, 2n_0^\gamma)$ ,

$$\begin{aligned} \mathbb{P}(v \sim C_{n_0, \text{base}}^{(1)}[1, n_0^{2/(\tau-1)}) \mid \mathcal{A}_{\text{large}}^{n_0}) &\geq 1 - \left(1 - p\left(1 \wedge \frac{\beta n_0^\gamma}{d^{d/2} n_0}\right)^\alpha\right)^{n_0^{1-\delta/4}} \\ &\geq 1 - \exp\left(-p\beta^\alpha d^{-\alpha d/2} n_0^{-\delta/4} n_0^{\gamma\alpha - (\alpha-1)}\right) \\ &= 1 - \exp\left(-p\beta^\alpha d^{-\alpha d/2} n_0^{-\delta/4}\right). \end{aligned}$$

We used that  $\gamma < 1$ , so  $n_0^{\gamma-1} = o(1)$  and the minimum is attained at the second term for  $n_0$  sufficiently large to obtain the first inequality. Then we used that  $\gamma = (\alpha - 1)/\alpha$  to obtain the second row. For  $n_0$  sufficiently large, we use the bound  $1 - e^{-x} \geq x/2$  for  $x < 1$ . Therefore, the right-hand side is at least  $p\beta^\alpha d^{-\alpha d/2} n_0^{-\delta/4}/2$ . This is the connectivity estimate for a single vertex. Since each vertex in  $\mathcal{V}_{Q_0}^{\text{spr}}[n_0^\gamma, 2n_0^\gamma)$  connects independently to  $C_{n_0, \text{base}}^{(1)}[1, n_0^{2/(\tau-1)})$ , the number of connecting vertices is Poisson with mean at least the mean in (4.27) multiplied with this connectivity estimate. So, using  $\rho_0 = n_0^{-\delta/4}$  in (4.14),

$$\begin{aligned} &\mathbb{P}(|\mathcal{V}_{Q_0}^{\text{spr}}[n_0^\gamma, 2n_0^\gamma) \cap \mathcal{C}_0| < \delta' \rho_0 n_0^\xi \mid \mathcal{A}_{\text{large}}^{n_0}) \\ &\leq \mathbb{P}(\text{Poi}((1 - q)(1 - 2^{-(\tau-1)}) \cdot \tfrac{1}{2} p\beta^\alpha d^{-\alpha d/2} \cdot n_0^{\xi-\delta/4}) < \delta' n_0^{\xi-\delta/4}). \end{aligned}$$

By the choice of  $\delta'$  in (4.10) and concentration of Poisson variables in Lemma A.1, the right-hand side tends to zero as  $n_0 \rightarrow \infty$ . This proves that the second factor in (4.25) tends to 1 as well. This finishes the proof of Claim 4.10.  $\square$

#### 4.2.4 Advancing the induction: the iterative step.

Claim 4.10 shows that level-0 boxes are good with constant probability. In the next claim we advance the induction by bounding the probabilities of the steps (*Sub*), (*Con*), (*Mark*) in Section 4.2.1 for level- $i$  boxes.

**Claim 4.11** (Advancing the induction). *Consider a supercritical KS RG on a PPP with parameters such that  $\zeta_{\text{hl}} > 0$ . Assume  $\delta \in (0, \zeta/2)$ . There exists a constant  $n_0^* > 0$  such that for all  $n_0 \geq n_0^*$  in Definition 4.8, for all  $i \geq 0$ ,*

$$\mathbb{P}(\mathcal{Q}_i(x) \text{ is } i\text{-bad}) \leq \exp(-6(n_i/n_0)^{\zeta-\delta}) := \text{err}_i. \quad (4.28)$$

**Proof** When  $i = 0$  the statement follows from Claim 4.10. Let  $i \geq 1$  and assume that (4.28) holds for level- $(i-1)$  boxes. We now follow the steps (*Sub*), (*Con*), (*Mark*) in Section 4.2.1 and bound their error probabilities. Like in Claim 4.10, we assume w.l.o.g. that  $x = 0$ , and write  $\mathcal{Q}_i := \mathcal{Q}_i(0)$ .

*Enough subboxes are good.* We start with the (*Sub*) step. Since  $n_i = m_i n_{i-1}$  by (4.11), we can partition  $\mathcal{Q}_i(x)$  into  $m_i$  subboxes of volume  $n_{i-1}$ . Let  $(x_j)_{j \leq m_i}$  denote the centers of these boxes. Define the event

$$\mathcal{A}_{\text{sub}} := \left\{ \sum_{j \in [m_i]} \mathbb{1}_{\{\mathcal{Q}_{i-1}(x_j) \text{ is } (i-1)\text{-bad}\}} < \varepsilon_i m_i \right\}. \quad (4.29)$$

In words,  $\mathcal{Q}_i(x)$  contains at least  $(1 - \varepsilon_i)m_i$  many  $(i-1)$ -good subboxes. Using that disjoint boxes are i.i.d., the induction hypothesis yields that the number of bad boxes is stochastically dominated by a  $\text{Bin}(m_i, \text{err}_{i-1})$  random variable, i.e.,

$$\mathbb{P}(\neg \mathcal{A}_{\text{sub}}) \leq \mathbb{P}(\text{Bin}(m_i, \text{err}_{i-1}) \geq \varepsilon_i m_i) \leq \sum_{k=\lceil \varepsilon_i m_i \rceil}^{m_i} \binom{m_i}{k} \text{err}_{i-1}^k.$$

We use that  $\binom{m_i}{k} \leq (em_i/k)^k$ , so the binomial coefficient is at most  $(e/\varepsilon_i)^k$  for all  $k \geq \varepsilon_i m_i$ . Hence, using the definition of  $\text{err}_{i-1}$  from (4.28),

$$\mathbb{P}(\neg \mathcal{A}_{\text{sub}}) \leq \sum_{k=\lceil \varepsilon_i m_i \rceil}^{\infty} \exp(-k(6(n_{i-1}/n_0)^{\zeta-\delta} - \log(e/\varepsilon_i))).$$

We use that  $\varepsilon_i = (i+1)^{-2}$  by definition in (4.13): when  $i = 1$  we compute  $6 - \log(e/\varepsilon_1) = 6 - \log(4e) = 5 - \log(4) > 3$ ; for  $i \geq 2$ , by the second inequality in (4.20) for  $C = 4/6$ , we obtain that  $6(n_{i-1}/n_0)^{\zeta-\delta} - \log(e/\varepsilon_i) > 3(n_{i-1}/n_0)^{\zeta-\delta}$  when  $n_0$  is sufficiently large. We obtain a geometric series with base  $\exp(-3(n_{i-1}/n_0)^{\zeta-\delta})$ ,

which is at most  $1/2$  for all  $i \geq 2$  by choosing  $n_0$  sufficiently large. Evaluating the geometric sum and then using the bound (Sub) in (4.16) yields

$$\begin{aligned} \mathbb{P}(\neg \mathcal{A}_{\text{sub}}) &\leq \sum_{k=\lceil \varepsilon_i m_i \rceil}^{\infty} \exp(-3(n_{i-1}/n_0)^{\zeta-\delta})^k \leq \exp(-3\varepsilon_i m_i (n_{i-1}/n_0)^{\zeta-\delta})/2 \\ &\leq \frac{1}{3} \exp(-6(n_i/n_0)^{\zeta-\delta}). \end{aligned} \quad (4.30)$$

This finishes the (Sub)-step of level- $i$ .

*Connectivity: the local giant is big enough.* We now move on to the (Con)-step in Section 4.2.1. In level  $i-1$ , we used the vertex sets  $\mathcal{V}_{\mathcal{Q}_{i-1}(x_j)}^{\text{base}}[1, n_0^{2/(\tau-1)}] \cup \mathcal{V}_{\mathcal{Q}_{i-1}(x_j)}^{\text{spr}}[1, 2n_{i-1}^{\gamma}]$  in each subbox  $j \in [m_i]$ , see (4.12). We denote the level- $(i-1)$  graphs on these vertex sets by  $\widehat{\mathcal{G}}_{n_{i-1}, x_j}$ . We now reveal the realization of  $(\widehat{\mathcal{G}}_{n_{i-1}, x_j})_{j \leq m_i}$  for all  $j \leq m_i$  in the level- $(i-1)$  subboxes of  $\mathcal{Q}_i(x)$ , and assume that the realization satisfies the event  $\mathcal{A}_{\text{sub}}$  for  $\mathcal{Q}_i(x)$ . We write

$$\begin{aligned} \text{pre-}\widehat{\mathcal{G}}_{n_i} &:= \cup_{j \leq m_i} \widehat{\mathcal{G}}_{n_{i-1}, x_j} \quad \text{on} \\ \mathcal{V}_{\text{pre}} &:= \cup_{j \leq m_i} \mathcal{V}_{\mathcal{Q}_{i-1}(x_j)}^{\text{base}}[1, n_0^{2/(\tau-1)}] \cup \mathcal{V}_{\mathcal{Q}_{i-1}(x_j)}^{\text{spr}}[1, 2n_{i-1}^{\gamma}]. \end{aligned} \quad (4.31)$$

Since the realization satisfies  $\mathcal{A}_{\text{sub}}$ , at least  $(1-\varepsilon_i)m_i$  many subboxes are  $(i-1)$ -good. We re-label boxes so that the boxes with indices  $j \leq (1-\varepsilon_i)m_i$  are all  $(i-1)$ -good. In  $\text{pre-}\widehat{\mathcal{G}}_{n_i}$  we did not reveal yet whether vertices in *different* level- $(i-1)$  subboxes are connected to each other or not. The graph  $\widehat{\mathcal{G}}_{n_i}$  which contains  $\text{pre-}\widehat{\mathcal{G}}_{n_i}$  additionally contains these edges between subboxes.

*The Erdős-Rényi type auxiliary graph.* The definition of  $(i-1)$ -goodness in (4.15) gives us components  $\mathcal{C}_{i-1, x_j}$  in each good subbox with size at least  $\rho_{i-1}n_{i-1}$ . Consider the auxiliary graph  $\mathcal{H} = (\mathcal{V}_{\mathcal{H}}, \mathcal{E}_{\mathcal{H}})$  in which the vertex  $j \in \mathcal{V}_{\mathcal{H}} := [\lceil (1-\varepsilon_i)m_i \rceil]$  corresponds to  $\mathcal{C}_{i-1, x_j}$ , and two vertices  $j, k \in \mathcal{V}_{\mathcal{H}}$  are connected by an edge if the components  $\mathcal{C}_{i-1, x_j}$  and  $\mathcal{C}_{i-1, x_k}$  are connected by an edge in  $\widehat{\mathcal{G}}_{n_i}$ . If the auxiliary graph  $\mathcal{H}$  is connected, then  $\mathcal{Q}_i(x)$  contains a component that has size at least  $(1-\varepsilon_i)m_i \cdot \rho_{i-1}n_{i-1}$ . Since  $m_i n_{i-1} = n_i$  and  $\rho_{i-1}(1-\varepsilon_i) = \rho_i$  by (4.11) and (4.14), respectively, this size is at least  $\rho_i n_i$ . Therefore, it satisfies the first criterion for being  $i$ -good in (4.15). Hence, for all realizations  $\text{pre-}\widehat{\mathcal{G}}_{n_i}$  satisfying  $\mathcal{A}_{\text{sub}}$ ,

$$\mathbb{P}(\exists \text{ component } \mathcal{C}_i : |\mathcal{C}_i| \geq \rho_i n_i \mid \text{pre-}\widehat{\mathcal{G}}_{n_i}) \geq \mathbb{P}(\mathcal{H} \text{ connected} \mid \text{pre-}\widehat{\mathcal{G}}_{n_i}). \quad (4.32)$$

We now prove that  $\mathcal{H}$  stochastically dominates an Erdős-Rényi graph. The criteria for being  $(i-1)$ -good in (4.15) guarantee that the level- $(i-1)$  good subboxes contain a component  $\mathcal{C}_{i-1, x_j}$  with size at least  $\rho_{i-1}n_{i-1}$ , and this component contains at least  $\delta' \rho_{i-1} n_{i-1}^{\zeta}$  vertices of mark in  $[n_{i-1}^{\gamma}, 2n_{i-1}^{\gamma}]$ . We bound the probability that there is an edge between two such components in different good subboxes by checking whether any of the potential edges between the  $\delta' \rho_{i-1} n_{i-1}^{\zeta}$  many level- $(i-1)$  high mark vertices in one box and the  $\rho_{i-1}n_{i-1}$  many vertices in the component of mark at least 1 in the other box is present. Any two vertices in  $\mathcal{Q}_i(x)$  are at distance at



most  $\sqrt{d}n_i^{1/d}$ , so using the connection probability in (4.6) we compute for any *good* subboxes with index  $j, k$  that

$$\begin{aligned}\mathbb{P}(\mathcal{C}_{i-1,x_j} \sim_{\widehat{\mathcal{G}}_{n_i}} \mathcal{C}_{i-1,x_k} \mid \text{pre-}\widehat{\mathcal{G}}_{n_i}) &\leq \left(1 - p\left(1 \wedge \frac{\beta n_{i-1}^\gamma}{\sqrt{d}n_i}\right)^\alpha\right)^{\delta' \rho_{i-1}^2 n_{i-1}^{1+\zeta}} \\ &\leq \exp\left(-p\delta' \beta^\alpha d^{-\alpha d/2} \rho_{i-1}^2 n_{i-1}^{1+\zeta+\gamma\alpha} n_i^{-\alpha}\right) \\ &= \exp\left(-p\delta' \beta^\alpha d^{-\alpha d/2} \rho_{i-1}^2 n_{i-1}^{\zeta+\alpha} n_i^{-\alpha}\right) =: 1 - q_i,\end{aligned}\quad (4.33)$$

where we used that  $\gamma = 1 - 1/\alpha$  to get the last row, and also to see that the minimum is at the second term in the first row, since  $n_{i-1}^\gamma/n_i$  can be made arbitrarily small by choosing  $n_0$  sufficiently large, see (4.11). The connectivity between different local large components is conditionally independent given the realization of  $\text{pre-}\widehat{\mathcal{G}}_{n_i}$ , so the auxiliary graph  $\mathcal{H}$  dominates an Erdős-Rényi graph  $G(\lceil(1 - \varepsilon_i)m_i\rceil, q_i)$ . The reader may verify that  $m_i q_i \gg \log m_i$  for all  $i \geq 1$  if  $n_0$  is sufficiently large, so by Claim 4.7 on the connectivity of  $G$ ,

$$\begin{aligned}\mathbb{1}_{\{\mathcal{A}_{\text{sub}}\}} \mathbb{P}(\mathcal{H} \text{ not conn.} \mid \text{pre-}\widehat{\mathcal{G}}_{n_i}) &\leq 3(em_i)^2(1 - q_i)^{(1-\varepsilon_i)m_i/2} \\ &\leq 3(em_i)^2 \exp\left(-\frac{3p\delta' \beta^\alpha}{8d^{\alpha d/2}} \rho_{i-1}^2 n_{i-1}^{\zeta+\alpha} n_i^{-\alpha} m_i\right),\end{aligned}\quad (4.34)$$

where we substituted  $1 - q_i$  from (4.33), and used that  $1 - \varepsilon_i \geq 3/4$  for all  $i \geq 1$ . We substitute  $n_i = m_i n_{i-1}$  to obtain the first row below, and then use the inequality (Con) in (4.17) to see that the  $i$ -dependent factors together can be bounded from below by  $C n_i^{\zeta-\delta}$  for any constant  $C > 0$  to get the second row. Here, we can choose  $n_0$  sufficiently large so that  $C$  cancels out the constant ( $\delta', \beta, p, d$ -dependent) prefactors:

$$\begin{aligned}\mathbb{1}_{\{\mathcal{A}_{\text{sub}}\}} \mathbb{P}(\mathcal{H} \text{ not conn.} \mid \text{pre-}\widehat{\mathcal{G}}_{n_i}) &\leq 3(em_i)^2 \exp\left(-\frac{3p\delta' \beta^\alpha}{8d^{\alpha d/2}} \rho_{i-1}^2 n_{i-1}^\zeta m_i^{1-\alpha}\right) \\ &\leq 3(em_i)^2 \exp\left(-7n_i^{\zeta-\delta}\right) \leq \frac{1}{3} \exp\left(-6n_i^{\zeta-\delta}\right).\end{aligned}\quad (4.35)$$

Returning to (4.32), this shows that the first condition of being level- $i$  good in (4.15), i.e., containing a connected component with size at least  $\rho_i n_i$ , holds with probability at least  $1 - \exp(-6n_i^{\zeta-\delta})/3$ , given  $\mathcal{A}_{\text{sub}}$ . This finishes the (Con)-step. We summarize it as follows: since  $\mathbb{P}(\neg \mathcal{A}_{\text{sub}}) \leq \exp(-6(n_i/n_0)^{\zeta-\delta})/3$  by (4.30),

$$\begin{aligned}\mathbb{P}(\mathcal{Q}_i(x) \text{ is } i\text{-bad}) &\leq \mathbb{P}(\neg \mathcal{A}_{\text{sub}}) + \mathbb{E}[\mathbb{1}_{\{\mathcal{A}_{\text{sub}}\}} \mathbb{P}(\mathcal{Q}_i(x) \text{ is } i\text{-bad} \mid \text{pre-}\widehat{\mathcal{G}}_{n_i})] \\ &\leq \mathbb{P}(\neg \mathcal{A}_{\text{sub}}) + \mathbb{E}[\mathbb{1}_{\{\mathcal{A}_{\text{sub}}\}} \mathbb{P}(\mathcal{H} \text{ not conn.} \mid \text{pre-}\widehat{\mathcal{G}}_{n_i})] \\ &\quad + \mathbb{E}[\mathbb{1}_{\{\mathcal{A}_{\text{sub}}\}} \mathbb{P}(\mathcal{Q}_i(x) \text{ is } i\text{-bad} \mid \text{pre-}\widehat{\mathcal{G}}_{n_i}, \mathcal{H} \text{ connected})]\end{aligned}\quad (4.36)$$

$$\begin{aligned}&\leq \frac{2}{3} \exp\left(-6(n_i/n_0)^{\zeta-\delta}\right) \\ &\quad + \mathbb{E}[\mathbb{1}_{\{\mathcal{A}_{\text{sub}}\}} \mathbb{P}(\mathcal{Q}_i(x) \text{ is } i\text{-bad} \mid \text{pre-}\widehat{\mathcal{G}}_{n_i}, \mathcal{H} \text{ conn.})],\end{aligned}\quad (4.37)$$

where the error of the (*Con*)-step is the first term in the last row, and the second term still needs to be bounded during the (*Mark*) step.

*Enough high-mark vertices in level- $i$ .* Finally, we treat the (*Mark*)-step in Section 4.2.1. That is, we bound the conditional probability in the last row above. This corresponds to the second criterion in (4.15), i.e., the requirement that  $\mathcal{C}_i$  also contains enough level- $i$  high mark vertices. On the event  $\mathcal{A}_{\text{sub}} \cap \{\mathcal{H} \text{ is connected}\}$ , we set  $\mathcal{C}_i$  to be the connected component in  $\widehat{\mathcal{G}}_{n_i}$  that contains  $\cup_{j \in \mathcal{V}_{\mathcal{H}}} \mathcal{C}_{i-1, x_j} =: \text{pre-}\mathcal{C}_i$ . By the argument leading to (4.32), pre- $\mathcal{C}_i$  has already size at least  $\rho_i n_i$ . Hence, by the criterion in (4.15), we need to bound

$$\mathbb{1}_{\{\mathcal{A}_{\text{sub}}\}} \mathbb{P}(|\mathcal{V}_{\mathcal{Q}_i(x)}^{\text{spr}}[n_i^\gamma, 2n_i^\gamma] \cap \mathcal{C}_i| < \delta' \rho_i n_i^{1-\gamma(\tau-1)} \mid \text{pre-}\widehat{\mathcal{G}}_{n_i}, \mathcal{H} \text{ connected}). \quad (4.38)$$

We argue similarly as for  $i = 0$  around and below (4.27). Since pre- $\mathcal{C}_i \subseteq \mathcal{V}_{\text{pre}}$  in (4.31), and  $2n_{i-1}^\gamma < n_i^\gamma$  by (4.22), the already revealed vertex set  $\mathcal{V}_{\text{pre}}$  is disjoint and thus independent of  $\mathcal{V}_{\mathcal{Q}_i(x)}^{\text{spr}}[n_i^\gamma, 2n_i^\gamma]$ . Hence, the PPP  $\mathcal{V}_{\mathcal{Q}_i(x)}^{\text{spr}}[n_i^\gamma, 2n_i^\gamma]$  itself is independent of the conditioning in (4.38). Using the intensity in (4.5), its expected size is

$$\mathbb{E}[|\mathcal{V}_{\mathcal{Q}_i(x)}^{\text{spr}}[n_i^\gamma, 2n_i^\gamma]|] = (1-q)(1-2^{-(\tau-1)})n_i^{1-\gamma(\tau-1)} = (1-q)(1-2^{-(\tau-1)})n_i^\zeta, \quad (4.39)$$

with  $q \in (0, 1)$  from the Definition 4.5. There are at least  $\rho_i n_i$  vertices in pre- $\mathcal{C}_i$ , each with mark at least one, and each pair of vertices inside  $\mathcal{Q}_i$  is within distance  $\sqrt{d}n_i^{1/d}$ . So, we get using the connection probability in (4.6) that each individual vertex in  $\mathcal{V}_{\mathcal{Q}_i(x)}^{\text{spr}}[n_i^\gamma, 2n_i^\gamma]$  connects by an edge to pre- $\mathcal{C}_i \subseteq \mathcal{C}_i$  with probability at least

$$\begin{aligned} 1 - \left(1 - p\left(1 \wedge \frac{\beta n_i^\gamma}{\sqrt{d} n_i}\right)^\alpha\right)^{\rho_i n_i} &\geq 1 - \exp\left(-p\beta^\alpha d^{-\alpha d/2} \rho_i n_i^{\gamma\alpha - (\alpha-1)}\right) \\ &= 1 - \exp\left(-p\beta^\alpha d^{-\alpha d/2} \rho_i\right). \end{aligned}$$

To obtain the second expression, we used that the minimum is at the second term for  $n_0$  sufficiently large, and  $\gamma = (\alpha - 1)/\alpha$  to get the second row. Since  $\rho_i = \Theta(n_0^{-\delta/4})$  as  $n_0$  tends to infinity by (4.14), we may assume that  $n_0$  is so large so that the bound  $1 - e^{-x} \geq x/2$  applies (for all  $x \leq 1$ ), and the right-hand side is at least  $p\beta^\alpha d^{-\alpha d/2} \rho_i/2$ . Since each vertex in  $\mathcal{V}_{\mathcal{Q}_i(x)}^{\text{spr}}[n_i^\gamma, 2n_i^\gamma]$  connects independently to pre- $\mathcal{C}_i$ , the number of connecting vertices forms a PPP with mean at least the product of (4.39) multiplied with this bound. So,

$$\begin{aligned} &\mathbb{1}_{\{\mathcal{A}_{\text{sub}}\}} \mathbb{P}(|\mathcal{V}_{\mathcal{Q}_i(x)}^{\text{spr}}[n_i^\gamma, 2n_i^\gamma] \cap \mathcal{C}_i| < \delta' \rho_i n_i^\zeta \mid \text{pre-}\widehat{\mathcal{G}}_{n_i}, \mathcal{H} \text{ conn.}) \\ &\leq \mathbb{P}\left(\text{Poi}\left((1-q)(1-2^{-(\tau-1)}) \cdot \frac{1}{2} p\beta^\alpha d^{-\alpha d/2} \cdot \rho_i n_i^\zeta\right) < \delta' \rho_i n_i^\zeta\right). \end{aligned}$$

The value of  $\delta'$  in (4.10) is chosen exactly so that  $2\delta'$  is the prefactor of  $\rho_i n_i^\zeta$  in the mean of the Poisson variable. Hence, by the concentration of Poisson variables in

Lemma A.1, there exists  $c = c(\delta') > 0$  such that

$$\mathbb{1}_{\{\mathcal{A}_{\text{sub}}\}} \mathbb{P}(|\mathcal{V}_{\mathcal{Q}_i(x)}^{\text{spr}}[n_i^\gamma, 2n_i^\gamma] \cap \mathcal{C}_i| < \delta' \rho_i n_i^\zeta \mid \text{pre-}\widehat{\mathcal{G}}_{n_i}, \mathcal{H} \text{ conn.}) < \exp(-c \rho_i n_i^\zeta),$$

Finally, by the bound (Mark) in (4.18), the right-hand side is at most  $\exp(-6n_i^{\zeta-\delta})/3$  for  $n_0$  sufficiently large. Combining this bound with (4.36) finishes the proof of Claim 4.11.  $\square$

If the box  $\mathcal{Q}_i(x)$  is  $i$ -good, then it contains a giant component on the restricted marks  $[1, 2n_i^\gamma)$ . Thus, Claims 4.10 and 4.11 together prove Proposition 4.1 along the subsequence  $n_i$ , and when  $\zeta = \zeta_{\text{hl}}$ . We now prove Proposition 4.1 for arbitrary  $n$ .

*Proof of Proposition 4.1 when  $\zeta_{\text{hl}} > 0$ .* W.l.o.g. we may assume that  $\delta \in (0, \zeta/2 \wedge 1 - \zeta)$  in the statement of the lemma. Let  $n_0$  be a constant such that Claims 4.10–4.11 hold for all  $i \geq 0$ . Fix now the sequence  $n_i$  in (4.11). We may assume that  $n > n_0$ , since the constant  $A$  in the statement of Proposition 4.1 can be increased to account for values  $n \leq n_0$ . We define

$$\begin{aligned} i_0 &:= \min\{i : 2n_i^\gamma > n_0^{2/(\tau-1)}\}, & i_* &:= \max\{i : n_i < n\}, \\ m &:= \lfloor (n/n_{i_*})^{1/d} \rfloor^d, & \tilde{n} &:= n_{i_*} m. \end{aligned} \quad (4.40)$$

By increasing the constant  $A$  in the statement of Proposition 4.1, we may assume w.l.o.g. that  $i_* > i_0$ . Then, there exist  $v$  such that

$$m \geq v(n/n_{i_*}). \quad (4.41)$$

By the definition of  $m$  and  $\tilde{n}$ ,  $\Lambda_{\tilde{n}} \subseteq \Lambda_n$  can be partitioned into exactly  $m \geq 1$  disjoint boxes of volume  $n_{i_*}$ . We denote the boxes of this partition by  $\mathcal{Q}_{i_*}(x_1), \dots, \mathcal{Q}_{i_*}(x_m)$ . From here on we argue similarly as in the proof of Claim 4.11: we use (Sub) and Con-type steps. We will show that at least half of the boxes are  $i_*$ -good, and afterwards connect the large components in the first  $\lceil m/2 \rceil$  many  $i_*$ -good boxes via a domination to an Erdős-Rényi random graph.

*Enough subboxes are good.* Define

$$\mathcal{A}_{\text{sub}}(n) := \left\{ \sum_{j \in [m]} \mathbb{1}_{\{\mathcal{Q}_{i_*}(x_j) \text{ is } i_*\text{-bad}\}} < m/2 \right\}.$$

We follow the calculations from (4.29) until (4.30) and replace  $\varepsilon_i$  by  $1/2$ ,  $m_i$  by  $m$  and  $n_{i-1}$  by  $n_{i_*}$  there. The bound  $\binom{m}{k} \leq (m/ek)^k$  works for all  $m \geq 1$  and  $k \leq m$ . We also use there the second inequality in (4.20) for  $C = 4/6$ , which also holds for  $n_{i_*}$ . So all bounds except the last row in (4.30) remain valid. Using (4.41) we obtain the bound

$$\mathbb{P}(\neg \mathcal{A}_{\text{sub}}(n)) \leq \sum_{k=\lceil m/2 \rceil}^{\infty} \exp(-3(n_{i_*}/n_0)^{\zeta-\delta})^k \leq \exp(-3(m/2)(n_{i_*}/n_0)^{\zeta-\delta})/2$$

$$\leq \exp \left( - (v/2n_0^{\zeta-\delta}) \cdot nn_{i_*}^{\zeta-\delta-1} \right). \quad (4.42)$$

As  $n > n_{i_*}$  and  $\zeta < 1$ , it follows that  $n_{i_*}^{\zeta-\delta-1} > n^{\zeta-\delta-1}$ , which, when rearranged yields that  $nn_{i_*}^{\zeta-\delta-1} > n^{\zeta-\delta}$ . Hence,

$$\mathbb{P}(\neg \mathcal{A}_{\text{sub}}(n)) \leq \exp \left( - (v/2n_0^{\zeta-\delta}) \cdot n^{\zeta-\delta} \right). \quad (4.43)$$

This finishes the (Sub)-step for general  $n$ .

*Connecting local-giants of subboxes.* Now we move on to the (Con)-step in Section 4.2.1. We follow the argument around (4.31), and define the pre-vertex set  $\mathcal{V}_{\text{pre}}$  as there, with  $n_i$  replaced by  $n$ ,  $i-1$  replaced by  $i_*$ , and  $m_i$  replaced by  $m$  there:

$$\begin{aligned} \text{pre-}\widehat{\mathcal{G}}_{\tilde{n}} &:= \cup_{j \leq m} \widehat{\mathcal{G}}_{n_{i_*}, x_j} \quad \text{on} \\ \mathcal{V}_{\text{pre}}(\tilde{n}) &:= \cup_{j \leq m} \mathcal{V}_{\mathcal{Q}_{i_*}(x_j)}^{\text{base}} [1, n_0^{2/(\tau-1)}) \cup \mathcal{V}_{\mathcal{Q}_{i_*}(x_j)}^{\text{spr}} [1, 2n_{i_*}^{\gamma}). \end{aligned} \quad (4.44)$$

We reveal the realization of  $\text{pre-}\widehat{\mathcal{G}}_{\tilde{n}}$ , satisfying the  $\mathcal{A}_{\text{sub}}(n)$ . Following the argument above (4.32), we construct the auxiliary graph  $\mathcal{H}_n$  on vertices representing the first  $\lceil m/2 \rceil$  many  $i_*$ -good subboxes. Each vertex  $j$  represents a large component  $\mathcal{C}_{i_*, x_j}$  guaranteed by the first condition in  $i$ -goodness in (4.15). Two vertices  $j, k$  are connected by an edge in  $\mathcal{H}_n$  if there is an edge between  $\mathcal{C}_{i_*, x_j}$  and  $\mathcal{C}_{i_*, x_k}$ . We follow the computations between (4.32) and (4.34), and replace there  $(1 - \varepsilon_i)m_i$  with  $m/2$ , replace  $n_{i-1}, \rho_{i-1}$  with  $n_{i_*}, \rho_{i_*}$ , and replace  $n_i$  by  $\tilde{n}$ . The latter holds because  $mn_{i_*} = \tilde{n}$  by (4.40). Then, for some constant  $c > 0$ , (4.34) turns into

$$\mathbb{1}_{\{\mathcal{A}_{\text{sub}}(n)\}} \mathbb{P}(\mathcal{H}_n \text{ not conn.} \mid \text{pre-}\widehat{\mathcal{G}}_{\tilde{n}}) \leq A'm \exp \left( - c\rho_{i_*}^2 n_{i_*}^{\zeta+\alpha} \tilde{n}^{-\alpha} m \right).$$

By (4.40),  $m \leq m_{i_*+1}$ , so the bound (4.21) holds for  $\tilde{n} = mn_{i_*}$ , which gives the first row below. Then,  $\tilde{n} \geq vn$  follows by multiplying both sides of (4.41). Moreover, the prefactor  $A'm < A'n$  can be absorbed by decreasing the constant prefactor in the exponent. We obtain for some  $c' > c'' > 0$

$$\begin{aligned} \mathbb{1}_{\{\mathcal{A}_{\text{sub}}(n)\}} \mathbb{P}(\mathcal{H}_n \text{ not connected} \mid \text{pre-}\widehat{\mathcal{G}}_{\tilde{n}}) &\leq A'm \exp \left( - c'\rho_{i_*}^2 \tilde{n}^{\zeta-\delta} \right) \\ &\leq \frac{1}{2} \exp \left( - c''\rho_{i_*}^2 n^{\zeta-\delta} \right). \end{aligned}$$

The definition of  $\rho_i$  in (4.14) gives that  $\rho_{i_*} > \rho > 0$  for all  $i_*$ , which gives our desired bound for the (Con)-step. Combining the error bound in (4.43) with this, we obtain that for any  $A > 0$  sufficiently large that only depends on the fixed value of the constant  $n_0$ ,

$$\mathbb{P}(\mathcal{A}_{\text{sub}}(n) \cap \{\mathcal{H}_n \text{ connected}\}) \geq 1 - \exp \left( - \frac{1}{A} n^{\zeta-\delta} \right).$$

By (4.44) and the arguments below, connectivity of the auxiliary graph  $\mathcal{H}_n$  means that the large components  $\mathcal{C}_{i_*, x_j}$  in the first  $m/2$   $i_*$ -good boxes form a connected graph. In

$\text{pre-}\widehat{\mathcal{G}}_{\tilde{n}}$  we only revealed edges between vertices of mark at most  $2n_{i_*}' < 2n^\gamma$  by (4.44). By (4.40) and the assumption that  $i_* > i_0$  below (4.40), we also have  $2n_{i_*}' > n_0^{2/(\tau-1)}$ . Moreover, on  $\mathcal{A}_{\text{sub}}(n) \cap \{\mathcal{H}_n \text{ connected}\}$ , the  $\lceil m/2 \rceil$  components of size at least  $\rho_{i_*} n_{i_*}$  form a connected graph, and thus  $\mathcal{G}_n[1, 2n_{i_*}']$  contains a component of size at least

$$(m/2) \cdot \rho_{i_*} n_{i_*} = \rho_{i_*} \tilde{n}/2 \geq (v/2) \rho_{i_*} n \geq (v/2) \rho n.$$

This implies that the largest component must be at least this large as well. So,

$$\mathcal{A}_{\text{sub}}(n) \cap \{\mathcal{H}_n \text{ connected}\} \subseteq \{|\mathcal{C}_n^{(1)}[1, An^\gamma]| \geq \frac{1}{A}n\} \quad (4.45)$$

when  $A$  is sufficiently large. This finishes the proof of (4.1). We do not need the *(Mark)* step here since there is no further iteration. For the Palm version  $\mathbb{P}^x$  the proof is exactly the same.  $\square$

*Proof of Proposition 4.1 when  $\zeta_{\text{II}} > 0$ .* When the vertex is formed by a PPP, we obtain the second statement (4.2) by a straightforward adaptation of the above proofs. We provide a sketch. When  $\zeta_{\text{II}} = 2 - \alpha > 0$ , we re-define what we call an  $i$ -good box by leaving the requirement on the marks in (4.15) out. We now call  $\mathcal{Q}_i(x)$   $i$ -good (and  $i$ -bad otherwise) if it satisfies the event

$$\mathcal{A}_{i\text{-good}}(x) := \left\{ \exists \text{ component } \mathcal{C}_i \text{ in } \widehat{\mathcal{G}}_{n_i} : |\mathcal{C}_i| \geq \rho_i n_i \right\}. \quad (4.46)$$

Then, we replace  $\zeta_{\text{hl}}$  by  $\zeta_{\text{II}} = 2 - \alpha$  in all computations and definitions, and we skip the *(Mark)*-step on each level. We do this by setting the level- $i$  vertex set to be  $\widehat{\mathcal{V}}_i := \mathcal{V}_{\mathcal{Q}_i(x)}[1, n_0^{2/(\tau-1)})$ . That is, only the size of the box changes across levels, the mark-truncation stays put. The *(Sub)*-step leading to (4.30) carries through unchanged by the inductive assumption and the concentration of binomial random variables on the number of level- $(i-1)$  good subboxes. For the *(Con)*-step, instead of using high-mark vertices, we estimate the mark of each vertex from below by 1 in the large components  $\mathcal{C}_{i-1,x_j}$  and  $\mathcal{C}_{i-1,x_k}$ . Using this lower bound, the number of edges between the two components dominates a binomial random variable. Both components have size at least  $\rho_{i-1} n_{i-1}$  because they are contained in  $(i-1)$ -good boxes, and their distance is at most  $\sqrt{d} n_i^{1/d}$ , so (4.33) simplifies to

$$\begin{aligned} \mathbb{P}(\mathcal{C}_{i-1,x_j} \sim_{\widehat{\mathcal{G}}_{n_i}} \mathcal{C}_{i-1,x_k} \mid \text{pre-}\widehat{\mathcal{G}}_{n_i}) &\leq \left(1 - p \left(1 \wedge \frac{\beta}{\sqrt{d} n_i}\right)^\alpha\right)^{\rho_{i-1}^2 n_{i-1}^2} \\ &\leq \exp\left(- (p\beta d^{-\alpha d/2}) \cdot \rho_{i-1}^2 n_{i-1}^2 n_i^{-\alpha}\right) =: 1 - q_i. \end{aligned}$$

Observe that the exponent  $n_{i-1}^2 = n_{i-1}^{\zeta+\alpha}$  since here  $\zeta = 2 - \alpha$ . From here on the computations are identical to those between (4.34) and (4.35), and since the last term in (4.36) is not present, we conclude the proof at (4.36) for  $\zeta_{\text{II}} > 0$ . Along the lines we also use that all statements of Claim 4.9 remain valid also for  $\zeta = \zeta_{\text{II}}$ , except (4.22). But, (4.22) is only used in the *(Mark)*-step which we omit when we consider the case  $\zeta_{\text{II}} > 0$ .  $\square$

## 5 Sharp bounds on the existence of a giant

The goal of this section is to improve upon the probability of the existence of the giant component compared to Proposition 4.1, which is formalized in the next proposition. In Proposition 4.1, the exponent of  $n$  in the stretched exponential decay is  $\max(\zeta_{\text{hl}}, \zeta_{\text{ll}}) - \delta$ . Here we remove the error  $\delta$  and ‘append’  $(d-1)/d, \zeta_{\text{hh}}$  to the list in the maximum.

For a regime type  $\in \{\text{ll}, \text{hl}, \text{hh}\}$ , and constants  $\rho, \delta \in (0, 1)$ , we define the event

$$\mathcal{A}_{\text{giant}}^{\text{type}}(\rho) := \begin{cases} \{\exists \text{ comp. } \mathcal{C} \text{ in } \mathcal{G}_n : |\mathcal{C}| \geq \rho n, \mathcal{C} \supseteq \mathcal{V}[n^{\frac{1-\delta}{1-\delta}}, \infty)\}, & \text{if type} \in \{\text{hl}, \text{hh}\}, \\ \{\exists \text{ comp. } \mathcal{C} \text{ in } \mathcal{G}_n : |\mathcal{C}| \geq \rho n\}, & \text{if type} = \text{ll}. \end{cases} \quad (5.1)$$

Recall the values  $\gamma_{\text{hl}}$  and  $\gamma_{\text{hh}}$  from (2.11) and (2.13), respectively. Define

$$\gamma_{\text{ll}} := (\alpha - 1)/(\tau - 1), \quad (5.2)$$

which we will only use when  $\zeta_{\text{ll}} = 2 - \alpha > 0$ . In particular, we will not use it when  $\alpha = \infty$ .

**Proposition 5.1** (Sharp bounds for the giant). *Consider a supercritical KSRG satisfying Assumption 2.2 on a PPP or consider supercritical long-range percolation on  $\mathbb{Z}^d$ . Assume that there is a type  $\in \{\text{ll}, \text{hl}, \text{hh}\}$  so that  $\zeta_{\text{type}} > 0$  and that for a constant  $\rho > 0$ ,*

$$\mathbb{P}(\mathcal{A}_{\text{giant}}^{\text{type}}(\rho)) \longrightarrow 1, \quad \text{as } n \rightarrow \infty. \quad (5.3)$$

(long) Then for all  $\varepsilon > 0$ , there exists  $A > 0$  such that for all  $n \geq 1$

$$\mathbb{P}(|\mathcal{C}_n^{(1)}[1, An^{\gamma_{\text{type}}}]| < (1 - \varepsilon)\rho n) \leq \exp\left(-\frac{1}{A}n^{\zeta_{\text{type}}}\right). \quad (5.4)$$

(short) Also, for all  $\varepsilon > 0$ , there exists  $A > 0$  such that for all  $n \geq 1$

$$\mathbb{P}(|\mathcal{C}_n^{(1)}[1, A]| < (1 - \varepsilon)\rho n) \leq \exp\left(-\frac{1}{A}n^{\zeta_{\text{short}}}\right). \quad (5.5)$$

Proposition 5.1 gives sharp decay for the value of  $\zeta_{\text{type}}$  that maximizes  $\{\zeta_{\text{ll}}, \zeta_{\text{hl}}, \zeta_{\text{hh}}, \zeta_{\text{short}}\}$ , which is equal to  $\zeta_{\star}$  by Claim 2.3. When  $\zeta_{\text{long}} = \max\{\zeta_{\text{ll}}, \zeta_{\text{hl}}, \zeta_{\text{hh}}\} > 0$  and  $(d-1)/d > 0$ , the second statement shows that the giant still exists with stretched exponential error probability even when all vertices with mark above a constant are deleted from the graph. When  $\zeta_{\star} = (d-1)/d$ , this decay is even stronger than (5.4).

The requirement (5.3) corresponds to (Req) in Section 3. It demands *some* initial density bound  $\rho$  on the size of the giant, and for the (hh) and (hl)-regimes it also demands that the highest-mark vertices are in the giant. Via edges incident to these high-mark vertices we can connect large components in subboxes in renormalization schemes below. We guarantee the presence of these long edges via the reasoning in Section 2.2. The requirement on high-mark vertices is not needed in the (ll) regime in (5.1), since we bound all marks from below by one. Thus, in the (ll)-regime, (5.3) is already satisfied for a small  $\rho$  by Proposition 4.1. At the end of this section we

verify (5.3) also for the (hl) and (hh)-regimes for a small  $\rho$  using Proposition 4.1, obtaining Corollary 5.7.

We will apply Proposition 5.1 again in Section 6.3 with density  $\rho$  close to  $\theta$ , as by then we have proved the weak LLN for  $|\mathcal{C}_n^{(1)}|$  for all parameters satisfying  $\zeta_{\text{long}} > 0$ . This will result in the upper bound on the right-hand side of (2.21) in Theorem 2.7. The exponent  $\zeta_*$  of the stretched exponential decay in Proposition 5.1 will also appear in our upper bounds on  $|\mathcal{C}_n^{(2)}|$  and the cluster-size decay for  $\zeta_{\text{type}} \in \{\zeta_{\text{hl}}, \zeta_{\text{ll}}\}$  in Theorems 2.9 and 2.10. Instead of Proposition 5.1, Proposition 4.1 could instead be used as input to obtain (weaker) upper bounds on  $|\mathcal{C}_n^{(2)}|$  and the cluster-size decay for  $\zeta_{\text{type}} \in \{\zeta_{\text{hl}}, \zeta_{\text{ll}}\}$ . That strategy would still give a weak LLN and also Theorem 2.7, but would lead to the appearance of an error  $\delta$  in exponents in Theorems 2.9 and 2.10, which would then never match the lower bounds in Theorems 2.9 and 2.10. This explains the benefit of Proposition 5.1.

**Remark 5.2** (High-level comparison of proofs in Sections 4 and 5). Section 5 sets up a single-layer renormalization with carefully chosen number and size of boxes that lead exactly to (5.4) and (5.5), respectively. It needs Section 4 as an input, because in each of these boxes a linear-sized component has to be present whp on truncated mark vertices.

It is natural to ask why one could not use ‘more carefully chosen’ box sizes already in Section 4 and obtain better bounds already there? The rough answer is the following: here in Section 5 we shall shortly perform a single-layer renormalization. Since it is a single layer, we *can* afford to lose a small linear portion of the giant. Thus, we can renormalize to a *supercritical* Erdős-Rényi random graph (ERRG), and use large deviation results for those. Then, we have  $\Theta(n^\zeta)$  many boxes, and the probability that a supercritical ERRG has a too small giant decays exponentially in the the number of boxes.

On the contrary, Section 4 uses an iterative scheme in order to improve the initial sublinear bound on the largest component to a linear bound. That is not possible in a single-layer renormalization. In an iterative scheme, one *cannot* allow to lose a small constant fraction of the largest component in each step. That would (again) result in a sublinear bound on the largest component. Therefore, in Section 4 we renormalized to a *connected* ERRG. For that we need slightly larger boxes, and thus we have  $o(n^\zeta)$  boxes when reaching total size  $n$ . Moreover, the probability of the rare event that such an ERRG is not connected, decays exponentially in the expected degree, which is smaller than the total number of boxes. These reasons explain the  $-\delta$  term in the exponent in Section 4.

We proceed to the proofs. We start with a claim that allows us to truncate the vertex marks from above; then we separately present the proof of the two statements in Proposition 5.1.

**Claim 5.3** *Consider a supercritical KSRG under the setting of Proposition 5.1. Assume that  $\zeta_{\text{type}} > 0$  and (5.3) holds for some  $\rho > 0$ ,  $\text{type} \in \{\text{hl}, \text{hh}\}$ , and  $\delta \in (0, 1)$ . Then, as  $n$  tends to infinity,*

$$\mathbb{P}\left(\exists \text{comp. } \mathcal{C} \text{ in } \mathcal{G}_n \left[1, n^{\frac{1}{\tau-1}}\right) : \mathcal{C} \supseteq \mathcal{V}\left[n^{\frac{1-\delta}{\tau-1}}, n^{\frac{1}{\tau-1}}\right), |\mathcal{C}| > \rho n\right) \longrightarrow 1. \quad (5.6)$$

**Proof** We abbreviate  $\underline{w} = n^{(1-\delta)/(\tau-1)}$ ,  $\overline{w} = n^{1/(\tau-1)}$  throughout the proof. We start from (5.3) and distinguish whether  $\mathcal{G}_n$  contains vertices of mark at least  $\overline{w}$ : if there are no vertices of mark at least  $\overline{w}$ , then  $\mathcal{G}_n[1, \overline{w}] = \mathcal{G}_n$ . Slightly abusing standard notation, we write here  $C \subset \mathcal{V}_n[1, \underline{w}] \cup \mathcal{V}_n[\underline{w}, \infty)$  for a set of vertices that is a *component* when restricted to marks below  $\underline{w}$ . Then we get

$$\begin{aligned} & \mathbb{P}(\exists \text{ comp. } C \text{ in } \mathcal{G}_n : C \supseteq \mathcal{V}[\underline{w}, \infty), |C| > \rho n) \\ &= \mathbb{P}(\{\exists \text{ comp. } C \text{ in } \mathcal{G}_n[1, \overline{w}] : C \supseteq \mathcal{V}[\underline{w}, \overline{w}), |C| > \rho n\} \cap \{\mathcal{V}[\overline{w}, \infty) = \emptyset\}) \\ &+ \mathbb{P}(\{\exists \text{ comp. } C \text{ in } \mathcal{G}_n : C \supseteq \mathcal{V}[\underline{w}, \overline{w}), |C| > \rho n\} \cap \{\mathcal{V}[\overline{w}, \infty) \neq \emptyset\}). \end{aligned}$$

Since the vertices are formed by a PPP,  $\mathcal{V}[\overline{w}, \infty)$  is independent of  $\mathcal{V}[1, \overline{w})$  and thus also of the graph  $\mathcal{G}_n[1, \overline{w})$ . So, in the first term on the right-hand side we can take the product of the probabilities of the two events. Using the intensity of  $\mathcal{V}$  in (4.5),  $\mathbb{P}(\mathcal{V}[\overline{w}, \infty) = \emptyset) = 1/e$ . Hence, conditioning on  $\{\mathcal{V}[\overline{w}, \infty) \neq \emptyset\}$  in the second term on the right-hand side yields

$$\begin{aligned} & \mathbb{P}(\exists \text{ comp. } C \text{ in } \mathcal{G}_n : C \supseteq \mathcal{V}[\underline{w}, \overline{w}), |C| > \rho n) \\ &= \frac{1}{e} \mathbb{P}(\exists \text{ comp. } C \text{ in } \mathcal{G}_n[1, \overline{w}) : C \supseteq \mathcal{V}[\underline{w}, \overline{w}), |C| > \rho n) \\ &+ \left(1 - \frac{1}{e}\right) \mathbb{P}(\exists \text{ comp. } C \text{ in } \mathcal{G}_n : C \supseteq \mathcal{V}[\underline{w}, \overline{w}), |C| > \rho n \mid \mathcal{V}[\overline{w}, \infty) \neq \emptyset). \end{aligned}$$

Since the left-hand side converges to 1 as  $n$  tends to infinity by the assumed condition (5.3), the two probabilities on the right-hand side must converge to 1 as well. Since the event in the first row on the right-hand side coincides with (5.6), the claim follows.  $\square$

## 5.1 When relying on long edges

Before the proof of Proposition 5.1(long), we state a result about the giant in Erdős-Rényi random graph  $G(n, p)$ . The result will ‘replace’ Claim 4.7 in the renormalization scheme of this section compared to the scheme in Section 4. The claim follows from a large-deviation principle for the size of the giant component obtained by O’Connell [68], see also [9]. For our result, we do not need the exact rate function of this LDP.

**Claim 5.4** (Large deviations in  $G(n, \lambda/n)$ ). *Consider the largest connected component  $C_n^{(1)}(p_n)$  of the Erdős-Rényi random graph  $G(n, p_n)$ . Fix  $\varepsilon > 0$ . There exist constants  $A > 0$ ,  $\lambda = \lambda(\varepsilon) > 1$  such that for all  $n \geq 1$*

$$\mathbb{P}(|C_n^{(1)}(\lambda/n)| < (1 - \varepsilon)n) \leq \exp\left(-\frac{1}{A}n\right).$$

**Proof** Denote by  $\varrho_\lambda$  the survival probability of a Bienaymé-Galton-Watson branching process with  $\text{Poi}(\lambda)$  offspring. By [68, Theorem 3.1], for every  $\lambda > 1$  and  $\tilde{\varepsilon} > 0$ , there



exists a constant  $A > 0$  such that

$$\mathbb{P}(|C_n^{(1)}(\lambda/n)| < (1 - \tilde{\varepsilon})\varrho_\lambda n) \leq \exp\left(-\frac{1}{A}n\right).$$

It is well known that the survival probability  $\varrho_\lambda$  tends to one as  $\lambda \rightarrow \infty$ , as  $\varrho_\lambda$  is the largest solution in  $(0, 1)$  to the equation  $\varrho_\lambda = 1 - \exp(-\lambda\varrho_\lambda)$  (see [70, Example 6.1.10]). Hence, given  $\varepsilon > 0$ , we may choose  $\lambda$  large enough, and  $\tilde{\varepsilon} > 0$  small enough, so that  $(1 - \tilde{\varepsilon})\varrho_\lambda \geq (1 - \varepsilon)$ . The statement follows.  $\square$

**Proof of Proposition 5.1(long)** We give the proof for KSRRGs on a Poisson point process. The result for long-range percolation on  $\mathbb{Z}^d$  follows analogously to the case  $\zeta_{\text{ll}} > 0$ .

Fix  $\zeta_{\text{type}} \in \{\zeta_{\text{hl}}, \zeta_{\text{hh}}, \zeta_{\text{ll}}\}$  such that  $\zeta_{\text{type}} > 0$ . Let  $\gamma_{\text{type}} \in \{\gamma_{\text{hl}}, \gamma_{\text{hh}}, \gamma_{\text{ll}}\}$  be the corresponding value that solves  $\zeta_{\text{type}} = 1 - \gamma_{\text{type}}(\tau - 1)$  by the definitions of  $\gamma_{\text{hl}}$ ,  $\gamma_{\text{hh}}$ , and  $\gamma_{\text{ll}}$  in (2.11), (2.13), and (5.2), respectively. For readability, we omit the subscript type throughout the proof.

*Definitions of constants and boxing scheme.* Below we will suitably define a constant  $M > 0$  depending on the regime (i.e., high-low, high-high, or low-low). For simplicity, we assume that  $(n^\zeta/M)^{1/d} \in \mathbb{N}$ , so that we can partition  $\Lambda_n$  into

$$m_n := n^\zeta/M \quad (5.7)$$

boxes  $\mathcal{Q}_1, \dots, \mathcal{Q}_{m_n}$ . Using that  $(1 - \zeta)/(\tau - 1) = \gamma$ , each box has volume

$$k_n := n/m_n = Mn^{1-\zeta} = (M^{1/(\tau-1)}n^\gamma)^{\tau-1}. \quad (5.8)$$

Given any  $\varepsilon > 0$  in Proposition 5.1(long), define the small  $\varepsilon_1, \varepsilon_2 > 0$  as the solutions of the following equations:

$$1 - \varepsilon = (1 - \varepsilon_1)^2, \quad \mathbb{P}(\mathcal{V}_{\varepsilon_2 k_n^{1/(\tau-1)}}, \infty) = \emptyset) = \varepsilon_1/4. \quad (5.9)$$

In a box of volume  $k_n$ , the expected number of vertices of mark at least  $\varepsilon_2 k_n^{1/(\tau-1)}$  equals  $\varepsilon_2^{-(\tau-1)}$  by (2.4), so  $\varepsilon_2$  is a constant.

*Good boxes and auxiliary graph.* We write  $\mathcal{V}_{k_n, j} := \mathcal{V} \cap (\mathcal{Q}_j \times [1, \infty))$  for the vertices in  $\mathcal{Q}_j$ ,  $\mathcal{G}_{k_n, j}$  for the subgraph of  $\mathcal{G}_n$  induced by the vertices in  $\mathcal{Q}_j$ , and similar to (4.31), we set

$$\text{pre-}\mathcal{G}_n := \cup_{j \leq m_n} \mathcal{G}_{k_n, j}. \quad (5.10)$$

We call a box  $\mathcal{Q}_j$ ,  $j \in [m_n]$  *good* if the following event holds depending on the regime type  $\in \{\text{hl}, \text{hh}, \text{ll}\}$ . When type  $\in \{\text{hl}, \text{hh}\}$ ,

$$\mathcal{A}_{\text{box}}(j) := \left\{ \begin{array}{l} \exists \text{comp. } \mathcal{C}_j \text{ in } \mathcal{G}_{k_n, j}[1, k_n^{1/(\tau-1)}): |\mathcal{C}_j| \geq \rho k_n, \\ \mathcal{C}_j \supseteq \mathcal{V}_{k_n, j}[\varepsilon_2 k_n^{1/(\tau-1)}, k_n^{1/(\tau-1)}] \neq \emptyset \end{array} \right\}. \quad (5.11)$$

When type = ll,

$$\mathcal{A}_{\text{box}}(j) := \left\{ \exists \text{comp. } \mathcal{C}_j \text{ in } \mathcal{G}_{k_n, j}[1, k_n^{1/(\tau-1)}): |\mathcal{C}_j| \geq \rho k_n \right\}.$$

We define the event that there are enough good boxes in  $\Lambda_n$  (similar to  $\mathcal{A}_{\text{sub}}$  in (4.29)):

$$\mathcal{A}_{\text{sub}} := \left\{ \sum_{j \in [m_n]} \mathbb{1}_{\{\mathcal{A}_{\text{box}}(j)\}} \geq (1 - \varepsilon_1)m_n \right\}, \quad (5.12)$$

The event  $\mathcal{A}_{\text{sub}}$  is measurable with respect to  $\text{pre-}\mathcal{G}_n$  in (5.10).

*Enough subboxes are good.* We estimate  $\mathbb{P}(\neg \mathcal{A}_{\text{sub}})$ , with  $\mathcal{A}_{\text{sub}}$  defined in (5.12). For that we analyze the individual probabilities  $\mathbb{P}(\mathcal{A}_{\text{box}}(j))$ . By the definition of  $\varepsilon_1$  in (5.9) and the hypothesis (5.3), for all  $n$  sufficiently large and any  $j \in [m_n]$

$$\mathbb{P}(\mathcal{A}_{\text{box}}(j)) \geq 1 - \varepsilon_1/2. \quad (5.13)$$

Since the graphs  $\mathcal{G}_{k_n, j}$  are iid across  $j \in [m_n]$ , by a standard Chernoff bound on binomial random variables, there exists a constant  $c > 0$  such that

$$\mathbb{P}(\neg \mathcal{A}_{\text{sub}}) \leq \exp(-cm_n). \quad (5.14)$$

*Introducing the auxiliary graph.* Similar to below (4.31) and (4.44), on any realization of  $\text{pre-}\mathcal{G}_n$  satisfying  $\mathcal{A}_{\text{sub}}$ , we re-label boxes so that the first  $\lceil (1 - \varepsilon_1)m_n \rceil$  many boxes are good. Then we define the auxiliary graph  $\mathcal{H}_n = (\mathcal{V}_{\mathcal{H}_n}, \mathcal{E}_{\mathcal{H}_n})$  on  $\lceil (1 - \varepsilon_1)m_n \rceil$  many vertices, so that vertex  $j \in \mathcal{V}_{\mathcal{H}_n}$  corresponds to the component  $\mathcal{C}_j$  inside  $\mathcal{Q}_j$  in (5.11), and two vertices  $j, \ell$  are connected in  $\mathcal{H}_n$  if the components  $\mathcal{C}_j$  and  $\mathcal{C}_\ell$  are connected by an edge in  $\mathcal{G}_n$ . We write  $\mathcal{C}_{\mathcal{H}_n}^{(1)}$  for the largest component in  $\mathcal{H}_n$ . By construction, the induced connected components  $(\mathcal{C}_j)_{j \in \mathcal{C}_{\mathcal{H}_n}^{(1)}}$  in (5.11) are all contained in the same component in  $\mathcal{G}_n$ . So if the event  $\mathcal{A}_{\text{sub}} \cap \{|\mathcal{C}_{\mathcal{H}_n}^{(1)}| \geq (1 - \varepsilon_1)|\mathcal{V}_{\mathcal{H}_n}|\}$  holds, then the graph  $\mathcal{G}_n[1, k_n^{1/(\tau-1)})$  contains a connected component  $\mathcal{C}$  with

$$|\mathcal{C}| \geq \sum_{j \in \mathcal{C}_{\mathcal{H}_n}^{(1)}} |\mathcal{C}_j| \geq (1 - \varepsilon_1)^2 m_n \cdot \rho k_n = (1 - \varepsilon_1)^2 \rho n = (1 - \varepsilon) \rho n.$$

Here we used that  $|\mathcal{C}_j| \geq \rho k_n$  by the event  $\mathcal{A}_{\text{box}}(j)$ , that  $k_n = n/m_n$  by (5.8), and  $(1 - \varepsilon_1)^2 = 1 - \varepsilon$  by (5.9). Clearly, the largest component of  $\mathcal{G}_n[1, k_n^{1/(\tau-1)})$  is at least as large as  $\mathcal{C}$ , so

$$\begin{aligned} \mathbb{P}\left(|\mathcal{C}_{\mathcal{H}_n}^{(1)}[1, k_n^{1/(\tau-1)}]| \leq (1 - \varepsilon) \rho n\right) \\ \leq \mathbb{P}(\neg \mathcal{A}_{\text{sub}}) + \mathbb{P}(\mathcal{A}_{\text{sub}} \cap \{|\mathcal{C}_{\mathcal{H}_n}^{(1)}| \leq (1 - \varepsilon)|\mathcal{V}_{\mathcal{H}_n}|\}). \end{aligned} \quad (5.15)$$

We study the two probabilities on the right-hand side separately.

*Large deviation for the giant of ERRGs.* We turn to the second term on the right-hand side in (5.15). We rewrite this second term using the law of total probability: we first reveal  $\text{pre-}\mathcal{G}_n$ , then integrate over realizations satisfying  $\mathcal{A}_{\text{sub}}$ :

$$\mathbb{P}(\mathcal{A}_{\text{sub}} \cap \{|\mathcal{C}_{\mathcal{H}_n}^{(1)}| \leq (1 - \varepsilon)^2 m_n\}) = \mathbb{E}[\mathbb{1}_{\{\mathcal{A}_{\text{sub}}\}} \mathbb{P}(|\mathcal{C}_{\mathcal{H}_n}^{(1)}| \leq (1 - \varepsilon_1)|\mathcal{V}_{\mathcal{H}_n}| \mid \text{pre-}\mathcal{G}_n)]. \quad (5.16)$$

Let  $\lambda = \lambda(\varepsilon_1)$ ,  $A > 0$  be the two constants from Claim 5.4 so that for all  $m \geq 1$ ,

$$\mathbb{P}(|\mathcal{C}_m^{(1)}(\lambda/m)| < (1 - \varepsilon_1)m) \leq A \exp\left(-\frac{1}{A}m\right). \quad (5.17)$$

We will show below that conditionally on the realization  $\widehat{\mathcal{G}}_n$  satisfying  $\mathcal{A}_{\text{sub}}$ ,

$$\mathcal{H}_n \succcurlyeq G(|\mathcal{V}_{\mathcal{H}_n}|, \lambda(\varepsilon_1)/|\mathcal{V}_{\mathcal{H}_n}|). \quad (5.18)$$

Suppose we have proven this stochastic domination. Then we can apply (5.17), and use that  $|\mathcal{V}_{\mathcal{H}_n}| = \lceil (1 - \varepsilon_1)m_n \rceil =: \widetilde{m}_n$ :

$$\begin{aligned} \mathbb{1}_{\{\mathcal{A}_{\text{sub}}\}} \mathbb{P}(|\mathcal{C}_{\mathcal{H}_n}^{(1)}| < (1 - \varepsilon_1)|\mathcal{V}_{\mathcal{H}_n}| \mid \widehat{\mathcal{G}}_n) &\leq \mathbb{P}(|\mathcal{C}_{\widetilde{m}_n}^{(1)}(\lambda/\widetilde{m}_n)| < (1 - \varepsilon_1)\widetilde{m}_n) \\ &\leq \exp\left(-\frac{1}{A}|\mathcal{V}_{\mathcal{H}_n}|\right) \leq \exp\left(-\frac{1}{A(1-\varepsilon_1)}m_n\right) \end{aligned}$$

This bound holds uniformly for all realizations of  $\text{pre-}\mathcal{G}_n$  with  $\mathcal{A}_{\text{sub}}$  in (5.16). So, combined with (5.14), this bounds the two terms on the right-hand side in (5.15). Recalling that  $m_n = n^\zeta / M$ ,

$$\mathbb{P}(|\mathcal{C}_n^{(1)}[1, k_n^{1/(\tau-1)}]| < (1 - \varepsilon)\rho n) \leq \exp\left(-\frac{c}{M}n^\zeta\right) + \exp\left(-\frac{1}{AM(1-\varepsilon_1)}n^\zeta\right).$$

The mark bound  $k_n^{1/(\tau-1)}$  in the largest component on the left-hand side is equal to  $M^{1/(\tau-1)}n^\gamma$  by (5.8). Thus, to finish the proof of (5.4), it remains to prove the stochastic domination in (5.18).

*Showing that the auxiliary graph dominates ERRG.* We now prove (5.18) case-by-case for the three regimes, for  $\zeta$  being either  $\zeta_{\text{ll}}$ ,  $\zeta_{\text{hl}}$ , or  $\zeta_{\text{hh}}$ . We reveal the subgraph  $\text{pre-}\mathcal{G}_n$  in (5.10), i.e., all induced subgraphs inside boxes  $(\mathcal{Q}_j)_{j \in [m_n]}$ , and assume that the realization satisfies  $\mathcal{A}_{\text{sub}}$ . The first  $\lceil (1 - \varepsilon_1)m_n \rceil$  boxes are good and these labels form the vertex set  $\mathcal{V}_{\mathcal{H}_n}$ . Each label corresponds to a component  $\mathcal{C}_j$  in  $\mathcal{Q}_j$ . Edges between these components  $(\mathcal{C}_j)_{j \in \mathcal{V}_{\mathcal{H}_n}}$  have not been revealed yet, and conditional on  $\text{pre-}\mathcal{G}_n$ , they are present independently of each other. Therefore, it is sufficient to show that for all  $j, \ell \in \mathcal{V}_{\mathcal{H}_n}$ ,

$$\mathbb{P}(\mathcal{C}_j \sim_{\mathcal{G}_n} \mathcal{C}_\ell \mid \text{pre-}\mathcal{G}_n) \geq \frac{\lambda}{|\mathcal{V}_{\mathcal{H}_n}|}. \quad (5.19)$$

We will prove this by suitably choosing the constant  $M$  in (5.7), case-by-case for the possible values of  $\zeta = \zeta_{\text{type}} \in \{\zeta_{\text{hl}}, \zeta_{\text{hh}}, \zeta_{\text{ll}}\}$ .

*1. High-low edges,  $\zeta_{\text{type}} = \zeta_{\text{hl}} > 0$ .* We assume that  $\alpha < \infty$ , since  $\zeta_{\text{hl}} = (\tau - 1)/\alpha - (\tau - 2) < 0$  otherwise. If a box  $\mathcal{Q}_j$  is good, see (5.11), the large component  $\mathcal{C}_j$  with size at least  $\rho k_n$  contains at least one vertex  $v_j^*$  of mark at least  $\varepsilon_2 k_n^{1/(\tau-1)}$ . If there are more such vertices, choose one arbitrarily. We estimate the probability that  $\mathcal{C}_j$  and  $\mathcal{C}_\ell$  are connected from below by the probability that either  $v_j^*$  connects to any vertex

in  $\mathcal{C}_\ell$  or  $v_\ell^*$  connects to any vertex in  $\mathcal{C}_j$ . We bound marks in the other component from below by 1. Any two vertices in  $\Lambda_n$  are within distance  $\sqrt{dn}^{1/d}$ . Since vertex weights are revealed in  $\text{pre-}\mathcal{G}_n$ , edges are present independently. So, we obtain from the connectivity probability  $p$  in (4.6):

$$\begin{aligned} \mathbb{P}(\mathcal{C}_j \sim_{\mathcal{G}_n} \mathcal{C}_\ell \mid \text{pre-}\mathcal{G}_n) &\leq \mathbb{P}(\{v_j^* \sim_{\mathcal{G}_n} \mathcal{C}_\ell\} \cap \{v_\ell^* \sim_{\mathcal{G}_n} \mathcal{C}_j\} \mid \text{pre-}\mathcal{G}_n) \\ &= \prod_{u \in \mathcal{C}_j} (1 - p(u, v_\ell^*)) \cdot \prod_{\tilde{u} \in \mathcal{C}_\ell} (1 - p(\tilde{u}, v_j^*)) \\ &\leq \left(1 - p\left(1 \wedge \frac{\beta \varepsilon_2 k_n^{1/(\tau-1)}}{d^{d/2} n}\right)^\alpha\right)^{2\rho k_n} \\ &\leq \exp\left(-2\rho k_n p\left(1 \wedge \beta^\alpha d^{-\alpha d/2} \varepsilon_2^\alpha k_n^{\alpha/(\tau-1)} n^{-\alpha}\right)\right). \end{aligned} \quad (5.20)$$

There are now two cases. If the minimum is attained at 1 in the last row, then the right-hand side tends to 0 since  $k_n = Mn^{1-\zeta} \rightarrow \infty$  since  $\zeta < 1$ . This means that (5.19) holds. If the minimum is not attained at 1, some further calculations are required. We set  $M$  as the solution of the equation

$$\rho p \beta^\alpha d^{-\alpha d/2} \varepsilon_2^\alpha = M^{-\alpha/(\tau-1)} \lambda / (1 - \varepsilon_1). \quad (5.21)$$

Then the last row in (5.20) equals

$$\exp\left(-2\rho p \beta^\alpha d^{-\alpha d/2} \varepsilon_2^\alpha k_n^{1+\alpha/(\tau-1)} n^{-\alpha}\right) = \exp\left(-\frac{2\lambda}{1-\varepsilon_1} k_n (k_n/M)^{\alpha/(\tau-1)} n^{-\alpha}\right).$$

We now use that  $k_n = n/m_n$  for the first appearance of  $k_n$  and that  $k_n/M = n^{1-\zeta_{\text{hl}}}$  in the second appearance, see (5.8). The exponent there is  $1 - \zeta_{\text{hl}} = (1 - 1/\alpha)(\tau - 1)$ , so rewriting the right-hand side yields that the exponents of  $n$  cancel each other. Using also  $|\mathcal{V}_{\mathcal{H}_n}| = \lceil (1 - \varepsilon_1)m_n \rceil$ ,

$$\exp\left(-\frac{2\lambda}{1-\varepsilon_1} (n/m_n) n^{(1-\zeta_{\text{hl}})\alpha/(\tau-1)} n^{-\alpha}\right) = \exp\left(-\frac{2\lambda}{(1-\varepsilon_1)m_n}\right) < 1 - \frac{\lambda}{|\mathcal{V}_{\mathcal{H}_n}|}$$

for all  $n$  sufficiently large. This proves the desired bound (5.19) in the high-low regime.

2. *High-high edges*,  $\zeta_{\text{type}} = \zeta_{\text{hl}} > 0$ , and  $\alpha < \infty$ . As in the high-low regime, we reveal  $\text{pre-}\mathcal{G}_n$ , and choose in each good box  $\mathcal{Q}_j$  a vertex  $v_j^*$  in  $\mathcal{C}_j$  with mark at least  $\varepsilon_2 k_n^{1/(\tau-1)}$ , see (5.11). We directly estimate the probability that  $v_j^*$  is connected to  $v_\ell^*$  for  $j, \ell \in \mathcal{V}_{\mathcal{H}_n}$ :

$$\begin{aligned} \mathbb{P}(\mathcal{C}_j \sim_{\mathcal{G}_n} \mathcal{C}_\ell \mid \text{pre-}\mathcal{G}_n) &\geq \mathbb{P}(v_j^* \sim_{\mathcal{G}_n} v_\ell^* \mid \text{pre-}\mathcal{G}_n) \\ &\geq p\left(1 \wedge \frac{\beta \varepsilon_2^{\sigma+1} k_n^{(\sigma+1)/(\tau-1)}}{d^{d/2} n}\right)^\alpha. \end{aligned} \quad (5.22)$$

If the minimum is attained at 1 on the right-hand side, then (5.19) follows, so we assume it is attained at the second term. We use that  $k_n = n/m_n$  by (5.8). The right-hand side

equals

$$p\beta^\alpha d^{-\alpha d/2} \varepsilon_2^{(\sigma+1)\alpha} n^{\alpha(\sigma+1)/(\tau-1)-\alpha} m_n^{-\alpha(\sigma+1)/(\tau-1)+1} \frac{1}{m_n},$$

where we multiplied by a factor  $m_n/m_n$ . We use now that  $m_n = n^{\zeta_{\text{hh}}}/M$  by (5.7) in the first occurrence of  $m_n$  above to compute the power of  $n$ . The above expression equals

$$p\beta^\alpha d^{-\alpha d/2} \varepsilon_2^{(\sigma+1)\alpha} M^{\alpha(\sigma+1)/(\tau-1)-1} n^{(1-\zeta_{\text{hh}})\alpha(\sigma+1)/(\tau-1)-\alpha+\zeta_{\text{hh}}} \cdot \frac{1}{m_n}.$$

It is an elementary computation to check that the power of  $n$  is 0 using that  $\zeta_{\text{hh}} = (\sigma + 1 - (\tau - 1))/(\sigma + 1 - (\tau - 1)/\alpha)$  in (2.14). Further,  $\zeta_{\text{hh}}$  is only positive when  $\sigma > \tau - 2$ . This latter inequality implies that  $(\sigma + 1)/(\tau - 1) > 1$  and so the exponent of  $M$  is strictly positive for all  $\alpha > 1$ . So, if we choose  $M$  large enough then the constant multiplying  $1/m_n$  in the above formula becomes arbitrarily large. So, returning to (5.22),

$$\mathbb{P}(\mathcal{C}_j \sim_{\mathcal{G}_n} \mathcal{C}_\ell \mid \text{pre-}\mathcal{G}_n) \geq \frac{1}{m_n} \cdot p\beta^\alpha d^{-\alpha d/2} \varepsilon_2^{(\sigma+1)\alpha} M^{\alpha(\sigma+1)/(\tau-1)-1} \geq \frac{\lambda}{|\mathcal{V}_{\mathcal{H}_n}|}$$

holds for sufficiently large  $M$ , since  $|\mathcal{V}_{\mathcal{H}_n}| = \lceil (1 - \varepsilon_1)m_n \rceil$ . This finishes the proof of (5.18) for the high-high regime with  $\alpha < \infty$ .

3. *High-high edges*,  $\zeta_{\text{type}} = \zeta_{\text{hh}} = (\sigma + 2 - \tau)/(\sigma + 1) > 0$ , and  $\alpha = \infty$ . This case is almost the same as the  $\alpha < \infty$  case. The only difference is that in (5.22) the minimum on the right-hand side *must* be attained at 1, and then the connection probability is  $p$ , otherwise it is 0. So we check whether

$$\beta \varepsilon_2^{\sigma+1} k_n^{(1+\sigma)/(\tau-1)} / (d^{d/2} n) \geq 1$$

holds. We use  $k_n = Mn^{1-\zeta_{\text{hh}}}$  in (5.8) and  $1 - \zeta_{\text{hh}} = \gamma_{\text{hh}}(\tau - 1) = (\tau - 1)/(\sigma + 1)$  when  $\alpha = \infty$ . Thus, the powers of  $n$  cancel each other again. Since  $\tau < 2 + \sigma$ , we can choose  $M$  large enough so that

$$\beta \varepsilon_2^{\sigma+1} M^{(1+\sigma)/(\tau-1)} / d^{d/2} \geq 1.$$

The bound (5.19) also follows in this regime.

4. *Low-low edges*,  $\zeta_{\text{type}} = \zeta_{\text{ll}} > 0$ . We may assume that  $\alpha < \infty$ , since otherwise  $\zeta_{\text{ll}} = 2 - \alpha < 0$ . By the definition of a good box in (5.11),  $|\mathcal{C}_j| \geq \rho k_n$  and each vertex has mark at least 1. So we compute using the connection probability  $p$  in (4.6) for  $j, \ell \in \mathcal{V}_{\mathcal{H}_n}$ :

$$\begin{aligned} \mathbb{P}(\mathcal{C}_j \sim_{\mathcal{G}_n} \mathcal{C}_\ell \mid \text{pre-}\mathcal{G}_n) &\leq (1 - p\beta^\alpha d^{-\alpha d/2} n^{-\alpha})^{|\mathcal{C}_j| \cdot |\mathcal{C}_\ell|} \\ &\leq \exp(-p\beta^\alpha d^{-\alpha d/2} \rho^2 k_n^2 n^{-\alpha}). \end{aligned} \quad (5.23)$$

Since  $k_n = Mn^{1-\zeta} = Mn^{\alpha-1}$  in (5.8),  $k_n^2 n^{-\alpha} = M^2 n^{\alpha-2}$ . With  $m_n = n^{2-\alpha}/M$ , we may write  $k_n^2 n^{-\alpha} = M/m_n$  on the right-hand side above. We choose  $M$  as the solution of the equation

$$p\beta^\alpha d^{-\alpha d/2} \rho^2 M = 2\lambda/(1 - \varepsilon_1).$$

Now we obtain for  $n$  large that, using  $|\mathcal{V}_{\mathcal{H}_n}| = \lceil (1 - \varepsilon_1)m_n \rceil$ ,

$$\mathbb{P}(C_j \approx_{\mathcal{G}_n} C_\ell \mid \text{pre-}\mathcal{G}_n) = \exp\left(-p\beta^\alpha d^{-\alpha d/2} \rho^2 M \frac{1}{m_n}\right) \leq \exp\left(-\frac{2\lambda}{|\mathcal{V}_{\mathcal{H}_n}|}\right) < 1 - \frac{\lambda}{|\mathcal{V}_{\mathcal{H}_n}|}, \quad (5.24)$$

showing (5.19). The same calculation works for long-range percolation on  $\mathbb{Z}^d$ .

*Conclusion for proving (5.5).* We conclude that the bound (5.19) holds whenever  $\zeta$  equals either  $\zeta_{ll}$ ,  $\zeta_{hl}$ ,  $\zeta_{hh}$  for a suitable choice of  $M$  in (5.7). This proves the stochastic domination stated in (5.18) and finishes the proof of (5.5) in Proposition 5.1.  $\square$

## 5.2 When relying on constant-length edges

We proceed to proving Proposition 5.1(short). Here, instead of using long edges to connect boxes and a domination towards an Erdős-Rényi random graph, we use “relatively short” edges and stochastic domination towards nearest neighbor site-bond percolation on  $\mathbb{Z}^d$ . In that model, each vertex  $x \in \mathbb{Z}^d$  is independently *active* with probability  $q$ , and each edge between active neighboring vertices is *open* with probability  $r$ , independently again. We denote the resulting graph formed by open edges between active vertices by  $\text{NNP}(q, r)$ , and write  $L_n^{(1)}(q, r)$  for the largest component inside  $\Lambda_n$ . We state a combination of two classical results [29, 59, Theorem 1.1] that will replace Claim 5.4 in this regime.

**Proposition 5.5** (High-density  $\text{NNP}(q, r)$  has a large giant). *Consider nearest neighbor site-bond percolation  $\text{NNP}(q, r)$  on  $\mathbb{Z}^d$  with probabilities  $q, r \in (0, 1]$ . For each  $\varepsilon > 0$ , there exist constants  $q_0, r_0 \in (0, 1)$ , and  $A > 0$  such that for all  $q \geq q_0$ ,  $r \geq r_0$ ,  $n \geq 1$ ,*

$$\mathbb{P}(|L_n^{(1)}(q, r)| < (1 - \varepsilon)n) \leq \exp\left(-\frac{1}{A}n^{(d-1)/d}\right).$$

**Proof** Fix  $\varepsilon > 0$ . We shall define an  $\varepsilon_1 = \varepsilon_1(\varepsilon)$  and set  $q_0 := 1 - \varepsilon_1$  and  $r_0 = 1 - \varepsilon_1$ . Let  $\omega \sim \text{NNP}(q_0, r_0)$ . We couple  $\omega$  to iid Bernoulli bond percolation  $\omega^* \sim \text{NNP}(1, 1 - 12d\varepsilon_1)$  on  $\mathbb{Z}^d$ , similarly to [7, 59]. By a union bound over the endpoints of an edge and the edge itself, each edge in is open in  $\omega$  with *marginal* probability at least  $1 - 3\varepsilon_1$ . Consider an edge with endpoints  $(x, y)$  in  $\omega$ , and let  $N_{(x,y)} := \{\text{edges incident to } x \text{ or } y \text{ in } \mathbb{Z}^d\} \setminus \{(x, y)\}$ . Clearly  $|N_{(x,y)}| = 4d - 2$ . For any set of edges  $S \subseteq N_{(x,y)}$ , by a union bound,

$$\begin{aligned} \mathbb{P}((x, y) \text{ open in } \omega \mid \text{all edges in } S \text{ open in } \omega) &\geq \mathbb{P}((x, y) \text{ open and all edges in } S \text{ open}) \\ &\geq 1 - (4d - 1) \cdot 3\varepsilon_1 \\ &\geq 1 - 12d\varepsilon_1 =: \tilde{p}. \end{aligned}$$

Edges that do not share a vertex are independently open, so one can iterate this argument: for every finite set  $E$  of edges in  $\mathbb{Z}^d$ ,  $\mathbb{P}(\text{all edges in } E \text{ are open}) \geq \tilde{p}^{|E|}$ . Hence, by Strassen's theorem [60] we can couple  $\omega$  with an independent bond percolation  $\omega^\star$  on  $\mathbb{Z}^d$  with retention-probability  $\tilde{p}$ , where every open edge in  $\omega^\star$  is open in  $\omega$ . Let us write  $C_{B_n}^{(1)}(\tilde{p})$  for the largest component of  $\omega^\star \cap \Lambda_n$ . Given  $\varepsilon > 0$ , by [29, Theorem 1.1], there exists a  $p_\star(\varepsilon) < 1$  so that whenever  $\tilde{p} > p_\star(\varepsilon)$ , for some  $A > 0$ ,

$$\mathbb{P}(|C_{B_n}^{(1)}(\tilde{p})| < (1 - \varepsilon)n) \leq \exp\left(-\frac{1}{A}n^{(d-1)/d}\right).$$

Hence, setting  $\varepsilon_1 > (1 - p_\star(\varepsilon))/12d$  yields  $\tilde{p} > p(\varepsilon)$ , and we obtain by the coupling  $\omega \succcurlyeq \omega_\star$ :

$$\mathbb{P}(|L_n^{(1)}(q, r)| < (1 - \varepsilon)n) \leq \mathbb{P}(|C_{B_n}^{(1)}(\tilde{p})| < (1 - \varepsilon)n) \leq \exp\left(-\frac{1}{A}n^{(d-1)/d}\right),$$

and the result follows for  $q_0 = r_0 = 1 - \varepsilon_1$ , and by monotonicity for all values  $q \geq q_0$  and  $r \geq r_0$ .  $\square$

**Proof of Proposition 5.1(short)** We give the proof for KS RGs on a Poisson point process. The result for long-range percolation on  $\mathbb{Z}^d$  follows analogously to the case  $\zeta_{\text{ll}} > 0$ . By the assumptions in Proposition 5.1), there is a type  $\in \{\text{hl}, \text{hh}, \text{ll}\}$  so that  $\zeta_{\text{type}} > 0$ . Positivity will ensure that “relatively long” edges are present to keep the giant together. Let  $M$  be a sufficiently large constant. We assume for simplicity that  $(n/M)^{1/d} \in \mathbb{N}$ . Tessellate  $\Lambda_n$  into  $m_n := n/M$  many boxes  $\mathcal{Q}_1, \dots, \mathcal{Q}_{m_n}$  of volume  $M$ . Similarly as before, we write  $\mathcal{V}_{M,j} := \mathcal{V} \cap (\mathcal{Q}_j \times [1, \infty))$  for the vertices in  $\mathcal{Q}_j$ ,  $\mathcal{G}_{M,j}$  for the subgraph of  $\mathcal{G}_n$  induced by these vertices in  $\mathcal{Q}_j$ , and  $\text{pre-}\mathcal{G}_n$  for the union of the subgraphs  $\mathcal{G}_{M,j}$ ,  $j \in [m_n]$ . Fix  $\delta > 0$  small as in (5.3).

*Good subboxes.* We call a box  $\mathcal{Q}_j$ ,  $j \in [m_n]$  *good* if the following event holds. When type  $\in \{\text{hl}, \text{hh}\}$ , we set

$$\mathcal{A}_{\text{box}}(j) := \left\{ \begin{array}{l} \exists \text{comp. } \mathcal{C}_j \text{ in } \mathcal{G}_{M,j}[1, M^{1/(\tau-1)}) : |\mathcal{C}_j| \geq \rho M, \\ \mathcal{C}_j \supseteq \mathcal{V}_{M,j}[M^{(1-\delta)/(\tau-1)}, M^{1/(\tau-1)}], \\ \frac{1}{2}M^\delta \leq |\mathcal{V}_{M,j}[M^{(1-\delta)/(\tau-1)}, M^{1/(\tau-1)}]| \end{array} \right\}. \quad (5.25)$$

When type = ll, we set

$$\mathcal{A}_{\text{box}}(j) := \left\{ \exists \text{comp. } \mathcal{C}_j \text{ in } \mathcal{G}_{M,j}[1, M^{1/(\tau-1)}) : |\mathcal{C}_j| \geq \rho M \right\}. \quad (5.26)$$

This event is measurable with respect to  $\text{pre-}\mathcal{G}_n$ . Our assumption (5.3) directly implies Claim 5.3, which in turn implies that the first two requirements in  $\mathcal{A}_{\text{box}}(j)$  – concerning  $\mathcal{C}_j$  – are satisfied with probability arbitrarily close to 1 by choosing  $M$  sufficiently large. For the third requirement in (5.25) we use the intensity of the PPP (4.5), and compute for fixed  $\delta > 0$

$$\mathbb{P}(|\mathcal{V}_{M,j}[M^{(1-\delta)/(\tau-1)}, M^{1/(\tau-1)}]| \geq \frac{1}{2}M^\delta) \longrightarrow 1,$$

as  $M$  tends to infinity. Since the three conditions on the right-hand side hold with probability arbitrarily close to 1, we may assume that  $M$  is so large that for a fixed

$q \in (0, 1)$  and any  $j \in [m_n]$ , independently of all other boxes,

$$\mathbb{P}(\mathcal{A}_{\text{box}}(j)) > q. \quad (5.27)$$

*The auxiliary graph and LDP for site-bond percolation.* Consider then the auxiliary graph  $\mathcal{H}_n$ , with vertex set  $\mathcal{V}_{\mathcal{H}_n} \subset [m_n]$ . By definition of the boxing, the set of vertices  $[m_n]$  can be mapped to the integer points in a box of side-length  $m_n^{1/d}$ , such that any two adjacent boxes are mapped to integer points at distance one from each other. We call  $j \in \mathcal{V}_{\mathcal{H}_n}$  active if the box  $\mathcal{Q}_j$  satisfies  $\mathcal{A}_{\text{box}}(j)$ . Two active vertices  $j, \ell \in \mathcal{V}_{\mathcal{H}_n}$  are connected in  $\mathcal{H}_n$  if they are neighboring vertices in  $B_{m_n}$ , and there exist vertices  $v \in \mathcal{C}_j, \tilde{v} \in \mathcal{C}_\ell$  such that  $v$  is connected by an edge to  $\tilde{v}$  in  $\mathcal{G}_n$ . Let us write  $\mathcal{C}_{\mathcal{H}_n}^{(1)}$  for the largest component in  $\mathcal{H}_n$ . Since each  $\mathcal{C}_j$  in a good box  $\mathcal{Q}_j$  has size at least  $\rho n$ , the event  $\{|\mathcal{C}_{\mathcal{H}_n}^{(1)}| \geq (1 - \varepsilon)m_n\}$  implies  $\{|\mathcal{C}_n^{(1)}[1, M^{1/(\tau-1)}]| \geq (1 - \varepsilon)\rho n\}$ . Let  $q, r > 0$  be such that by Proposition 5.5 for a fixed  $\varepsilon > 0, A > 0$  and all  $n \geq 1$ ,

$$\mathbb{P}(|L_{m_n}^{(1)}(q, r)| < (1 - \varepsilon)m_n) \leq \exp\left(-\frac{1}{A}m_n^{(d-1)/d}\right).$$

We will show that we can choose  $M$  sufficiently large (depending on  $\varepsilon$ ) so that

$$\mathcal{H}_n \succcurlyeq \text{NNP}(q, r). \quad (5.28)$$

This domination then implies that  $|\mathcal{C}_{\mathcal{H}_n}^{(1)}| \succcurlyeq |L_{m_n}^{(1)}(q, r)|$ , and so

$$\begin{aligned} \mathbb{P}(|\mathcal{C}_n^{(1)}[1, M^{1/(\tau-1)}]| < (1 - \varepsilon)\rho n) &\leq \mathbb{P}(|L_{m_n}^{(1)}(q, r)| < (1 - \varepsilon)m_n) \\ &\leq \exp\left(-\frac{1}{A}m_n^{(d-1)/d}\right) \\ &= \exp\left(-\frac{1}{AM^{(d-1)/d}}n^{(d-1)/d}\right), \end{aligned} \quad (5.29)$$

implying the statement of Proposition 5.1 for a larger value  $A$ .

*The auxiliary graph dominates site-bond percolation.* It remains to prove the domination (5.28). By (5.27) and the definition of  $\mathcal{H}_n$  below (5.27), each vertex  $j \in [m_n]$  is independently active with probability at least  $q$  for  $M$  sufficiently large. Thus, it remains to show that for two large components  $\mathcal{C}_j, \mathcal{C}_\ell$  in adjacent good boxes  $\mathcal{Q}_j, \mathcal{Q}_\ell$ ,

$$\mathbb{P}(\mathcal{C}_j \approx_{\mathcal{G}_n} \mathcal{C}_\ell \mid \text{pre-}\mathcal{G}_n) = \prod_{u \in \mathcal{C}_j, v \in \mathcal{C}_\ell} (1 - p(u, v)) \longrightarrow 0, \quad \text{as } M \rightarrow \infty. \quad (5.30)$$

We write  $|\mathcal{E}(\mathcal{C}_j, \mathcal{C}_\ell)|$  for the number of edges between the two components. Then

$$\begin{aligned} \mathbb{P}(\mathcal{C}_j \approx_{\mathcal{G}_n} \mathcal{C}_\ell \mid \text{pre-}\mathcal{G}_n) &\leq \exp\left(-\sum_{u \in \mathcal{C}_j, v \in \mathcal{C}_\ell} p(u, v)\right) \\ &= \exp\left(-\mathbb{E}[|\mathcal{E}(\mathcal{C}_j, \mathcal{C}_\ell)| \mid \text{pre-}\mathcal{G}_n]\right). \end{aligned}$$



We will show case-by-case for  $\zeta \in \{\zeta_{hl}, \zeta_{hh}, \zeta_{ll}\}$  that  $\mathbb{E}[|\mathcal{E}(\mathcal{C}_j, \mathcal{C}_\ell)| \mid \text{pre-}\mathcal{G}_n] \rightarrow \infty$  using either high-low, high-high, or low-low connections between two active neighboring boxes  $j, \ell$ . The maximal distance between any two vertices in neighboring boxes is  $2\sqrt{d}M^{1/d}$ .

*1. Low-low connections help:*  $\zeta_{ll} > 0$ . We may assume  $\alpha < \infty$  otherwise  $\zeta_{ll} = 2 - \alpha$  is negative. Every vertex has mark at least 1. Thus, from the connection probability  $p$  in (4.6) it follows that

$$\mathbb{E}[|\mathcal{E}(\mathcal{C}_j, \mathcal{C}_\ell)| \mid \text{pre-}\mathcal{G}_n] = \Omega(M^{-\alpha} |\mathcal{C}_j| \cdot |\mathcal{C}_\ell|) = \Omega(M^{2-\alpha}). \quad (5.31)$$

Since by assumption  $\zeta_{ll} = 2 - \alpha > 0$ , (5.30) follows for fixed  $\varepsilon$ . The same calculation works for long-range percolation on  $\mathbb{Z}^d$ .

*2. High-low connections help:*  $\zeta_{hl} > 0$ . We may again assume  $\alpha < \infty$  otherwise  $\zeta_{hl}$  is negative. On the event  $\mathcal{A}_{\text{box}}(j)$ , defined in (5.25), the component  $\mathcal{C}_j$  contains at least  $M^\delta/2$  vertices with mark at least  $M^{(1-\delta)/(\tau-1)}$ . Then, using possible edges between a high-mark vertex in  $\mathcal{C}_j$  and the at least  $\rho M$  vertices in  $\mathcal{C}_\ell$  that have mark at least 1, we obtain

$$\begin{aligned} \mathbb{E}[|\mathcal{E}(\mathcal{C}_j, \mathcal{C}_\ell)| \mid \text{pre-}\mathcal{G}_n] &= \Omega(M^{\delta+1}) \cdot \Omega\left(1 \wedge \left(\frac{M^{(1-\delta)/(\tau-1)}}{M}\right)^\alpha\right) \\ &= \Omega(M^{1+\delta}) \wedge \Omega(M^{1-\alpha(\tau-2)/(\tau-1)+\delta(1-\alpha/(\tau-1))}). \end{aligned} \quad (5.32)$$

When  $\zeta_{hl} = 1 - (1 - 1/\alpha)(\tau - 1) = (\tau - 1)/\alpha - (\tau - 2) > 0$ , the exponent of  $M$  in (5.32) is positive for sufficiently small  $\delta > 0$ . So (5.30) follows.

*3. High-high connections help:*  $\zeta_{hh} > 0$ . We first assume that  $\alpha < \infty$ . Using that we only consider good boxes, and that there are at least  $M^\delta/2$  many high-mark vertices with mark in the interval  $[M^{(1-\delta)/(\tau-1)}, M^{1/(\tau-1)})$  in both  $\mathcal{C}_j$  and  $\mathcal{C}_\ell$ , we compute

$$\mathbb{E}[|\mathcal{E}(\mathcal{C}_j, \mathcal{C}_\ell)| \mid \text{pre-}\mathcal{G}_n] = \Omega(M^{2\delta}) \cdot \Omega\left(1 \wedge \left(\frac{M^{(\sigma+1)(1-\delta)/(\tau-1)}}{M}\right)^\alpha\right).$$

From (2.14),  $\zeta_{hh} > 0$  only when  $\sigma > \tau - 2$ . So, for  $\delta$  is sufficiently small, the second factor is  $\Theta(1)$ . Thus, the right-hand side tends to infinity with  $M$  and (5.30) follows. When  $\alpha = \infty$ , it follows analogously that  $\mathbb{E}[|\mathcal{E}(\mathcal{C}_j, \mathcal{C}_\ell)| \mid \widehat{\mathcal{G}}_n] = \Omega(M^{2\delta})$  if  $\delta$  is sufficiently small since the minimum is attained at 1 on the right-hand side above.

*Conclusion.* We conclude that (5.30) follows in all three cases, proving the stochastic domination (5.28) and finishing the proof of (5.5).  $\square$

### 5.3 Verification of the requirement

We now verify the requirement (5.3). To do so, we prove the following lemma that we only need for models with non-trivial vertex marks. We define one more exponent for each regime hl, ll, hh:

$$\eta_{hl} := 1 - \gamma_{hl}(\tau - 1)/\alpha, \quad \eta_{ll} := 1/\alpha, \quad \eta_{hh} := 1 - \sigma\gamma_{hh}. \quad (5.33)$$

Using the respective values of  $\gamma_{ll}$ ,  $\gamma_{hl}$ ,  $\gamma_{hh}$  in (5.2), (2.11), and (2.13), it is elementary to verify that if  $\zeta_{\text{type}} = \zeta_\star$  for some type  $\in \{ll, hl, hh\}$ , then  $\eta_{\text{type}} > \gamma_{\text{type}}$  holds (see the proof of Lemma 5.6). Then there are values of  $\alpha$  such that  $\gamma_{ll} > \eta_{ll}$ . We write  $\mathcal{C}_n(0)[1, w)$  for the component of the origin spanned on vertices in  $\mathcal{V}_n[1, w)$ .

**Lemma 5.6** (Poly-logarithmic mark-thresholds). *Consider a supercritical KSRG satisfying Assumption 2.2 on a PPP. Assume for a type  $\in \{ll, hl\}$  that  $\zeta_{\text{type}} = \max(\zeta_{hl}, \zeta_{ll}) > 0$ . Whenever there exist constants  $\delta \in [0, \zeta)$ ,  $A', \rho > 0$ , such that for all  $n \geq 1$ ,*

$$\mathbb{P}(|\mathcal{C}_n^{(1)}[1, A'n^{\gamma_{\text{type}}}]| < \rho n) \leq \exp\left(-\frac{1}{A'}n^{\zeta_{\text{type}}-\delta}\right), \quad (5.34)$$

*then, for all  $C > 0$ , there exists a constant  $A > 0$  such that for all  $n$  sufficiently large,*

$$\mathbb{P}(|\mathcal{C}_n^{(1)}[1, (A \log n)^{\gamma_{\text{type}}/(\zeta_{\text{type}}-\delta)}]| \geq \rho n) \geq 1 - n^{-C}, \quad (5.35)$$

*and*

$$\mathbb{P}(\exists \text{ comp. } \mathcal{C} \text{ in } \mathcal{G}_n : \mathcal{C} \supseteq \mathcal{V}[(A \log n)^{\eta_{\text{type}}/(\zeta_{\text{type}}-\delta)}, \infty), |\mathcal{C}| \geq \rho n) \geq 1 - n^{-C}. \quad (5.36)$$

*The statements remain valid for the Palm version  $\mathbb{P}^x$  of  $\mathbb{P}$  for any  $x \in \mathbb{R}^d$ .*

The proof is an adaptation of [48, Proposition 5.12] and we present it in Appendix B.3. In a nutshell, we apply a boxing scheme—with box sizes  $\Theta((\log n)^{1/(\zeta-\delta)})$ —and apply (5.34) to each box. To form  $\mathcal{C}$ , we connect boxes to each other using high- and/or low-mark vertices depending on the regime. Then, we connect vertices of mark at least  $(A \log n)^{\eta/(\zeta-\delta)}$  to the constant-mark vertices in  $\mathcal{C}$ . Computations are similar to those in Section 6.1.

The corollaries of Proposition 5.1 and Lemma 5.6 motivate the somewhat counterintuitive formulation of Lemma 5.6, i.e., the requirement (5.34) with both  $\delta > 0$  and  $\delta = 0$  allowed.

**Corollary 5.7** (A giant with sharper probability). *Consider a supercritical KSRG satisfying Assumption 2.2 on a PPP or consider supercritical long-range percolation on  $\mathbb{Z}^d$ . Assume that  $\zeta_{\text{long}} > 0$ , and  $\zeta_\star = \zeta_{\text{type}}$  for some type  $\in \{ll, hl, hh, \text{short}\}$ . Set  $\gamma_{\text{short}} := 0$ . There exist constants  $A_1 > 0$  such that for all  $n \geq 1$ ,*

$$\mathbb{P}(|\mathcal{C}_n^{(1)}[1, A_1 n^{\gamma_{\text{type}}}]| < \frac{1}{A_1} n) \leq \exp\left(-\frac{1}{A_1} n^{\max(\zeta_{\text{long}}, (d-1)/d)}\right). \quad (5.37)$$

**Proof** We first consider KSRGs on a PPP. By Claim 2.3,  $\zeta_{\text{long}} = \max(\zeta_{ll}, \zeta_{hl}, \zeta_{hh})$ . When  $\zeta_{\text{long}} \in \{\zeta_{ll}, \zeta_{hl}\}$ , Proposition 4.1 proves the requirement (5.34) of Lemma 5.6 for some  $\delta > 0$  and a small value of  $\rho > 0$ . Hence, (5.35)–(5.36) holds for these values of  $\delta$  and  $\rho$ . This implies the requirement (5.3) of Proposition 5.1. When  $\zeta_\star = \zeta_{hh} > 0$ , (5.3) follows from the combination of the law of large numbers [48, Corollary 2.3], and that  $|\mathcal{C}_n^{(1)}|$  contains all vertices of at least some polylogarithmic weight in  $n$  by [48, Proposition 5.12], which then gives (5.34) and thus Lemma 5.6. This proves (5.37) when  $\zeta_\star \in \{\zeta_{ll}, \zeta_{hl}, \zeta_{hh}\}$ , and gives the requirement (5.3) of Proposition 5.1 whenever

$\zeta_{\text{long}} > 0$ . Further, the case for  $\zeta_{\star} = \zeta_{\text{short}}$  follows from (5.5). For long-range percolation on  $\mathbb{Z}^d$ ,  $\zeta_{\text{hl}}$  and  $\zeta_{\text{hh}}$  are negative, so  $\zeta_{\text{long}} = \zeta_{\text{ll}} = 2 - \alpha$ . The requirement (5.34) of Lemma 5.6 is verified by [14, Theorem 3.2].  $\square$

The actual mark-threshold so that *all* vertices above that mark are in the giant is as follows.

**Corollary 5.8** (Poly-logarithmic mark-thresholds). *Consider a supercritical KSRG satisfying Assumption 2.2 on a PPP or consider long-range percolation on  $\mathbb{Z}^d$ . Assume  $\zeta_{\text{type}} = \max(\zeta_{\text{ll}}, \zeta_{\text{hl}}) > 0$  for some type  $\in \{\text{ll}, \text{hl}\}$ . For all  $C > 0$ , there exist constants  $A, \varepsilon > 0$  such that for all  $n$  sufficiently large,*

$$\mathbb{P}\left(|\mathcal{C}_n^{(1)}[1, (A \log n)^{\zeta_{\text{type}}/\zeta_{\text{type}}}]| \geq \varepsilon n\right) \geq 1 - n^{-C}, \quad (5.38)$$

and

$$\mathbb{P}(\exists \text{ comp. } \mathcal{C} \text{ in } \mathcal{G}_n : \mathcal{C} \supseteq \mathcal{V}_n[(A \log n)^{\eta_{\text{type}}/\zeta_{\text{type}}}, \infty), |\mathcal{C}| \geq \varepsilon n) \geq 1 - n^{-C}. \quad (5.39)$$

**Proof** In Corollary 5.7 we have just shown that the requirement (5.34) in Lemma 5.6 is met with  $\delta = 0$  for the case  $\zeta_{\text{type}} = \max(\zeta_{\text{ll}}, \zeta_{\text{hl}})$ . This yields the first two bounds (5.38)–(5.39) for KSRGs on a PPP. In long-range percolation, all marks are equal to one, so that  $\mathcal{V}_n[(A \log n)^{\eta_{\text{type}}/\zeta_{\text{type}}}, \infty) = \emptyset$ , and the two bounds are implied by Corollary 5.7.  $\square$

In Corollary 5.7 we obtain the sharp decay on the probability of having a giant, i.e., we obtain the target decay of the lower tail in Theorem 2.7. However, the proportion of vertices in the giant is not yet close to  $\theta$ .

## 6 Lower tail of large deviations

In this section we prove Theorem 2.7. We first show uniqueness of the giant via an upper bound on the size of the second-largest component  $\mathcal{C}_n^{(2)}$ . In Section 6.2, we extend this upper bound on  $|\mathcal{C}_n^{(2)}|$  to the upper bound on the cluster-size decay in Theorem 2.9 using our previous work [48]. A Law of Large Numbers for  $|\mathcal{C}_n^{(1)}|$  will follow. Eventually we prove the lower tail of large deviations.

### 6.1 Non-giant components are small

We prove an upper bound on  $|\mathcal{C}_n^{(2)}|$  when  $\max(\zeta_{\text{hl}}, \zeta_{\text{ll}}) > 0$ . When  $\zeta_{\text{hh}} > 0$ , we proved a similar upper bound in [48] using different techniques, see Remark 3.2. See also Remark 3.1 on the difficulties one needs to overcome to extend this result to the case when surface tension dominates connections, i.e., when  $\zeta_{\star} = \zeta_{\text{short}}$ .

**Proposition 6.1** (Upper bound second-largest component). *Consider a supercritical KSRG on a PPP or supercritical long-range percolation on  $\mathbb{Z}^d$ . Assume  $\zeta_{\text{type}} =$*

$\max(\zeta_{\text{ll}}, \zeta_{\text{hl}}) > 0$  for some type  $\in \{\text{ll}, \text{hl}\}$ . There exists a constant  $A > 0$  such that for all  $n > k$ ,

$$\mathbb{P}(|\mathcal{C}_n^{(2)}| > k) \leq (n/k) \exp\left(-\frac{1}{A} k^{\zeta_{\text{type}}}\right). \quad (6.1)$$

The statement remains valid for the Palm version  $\mathbb{P}^x$  of  $\mathbb{P}$  for any  $x \in \mathbb{R}^d$ .

The proof of Proposition 6.1 is inspired by Sly and Crawford [24] who obtain a polylogarithmic upper bound with unidentified exponent on the second-largest component for long-range percolation. Contrary to long-range percolation, we have to use the vertex-marks as well when  $\tau < \infty$ . We present the proof for KSRGs on Poisson point processes satisfying  $\zeta_{\text{hl}} > 0$ , and explain the needed changes for  $\zeta_{\text{ll}} > 0$  and for KSRGs on  $\mathbb{Z}^d$  at the end. We abbreviate  $\gamma = \gamma_{\text{hl}} = 1 - 1/\alpha$  and  $\zeta = \zeta_{\text{hl}} = 1 - \gamma_{\text{hl}}(\tau - 1) = (\tau - 1)/\alpha - (\tau - 2)$ . Clearly  $\zeta_{\text{hl}} > 0$  implies  $\alpha < \infty$ .

We let  $A_{5.7}$  be the constant from (5.37) in Corollary 5.7 so that for all  $n \geq 1$ ,

$$\mathbb{P}(|\mathcal{C}_n^{(1)}[1, A_{5.7} n^\gamma]| \leq \frac{1}{A_{5.7}} n) \leq \exp\left(-\frac{1}{A_{5.7}} n^\zeta\right). \quad (6.2)$$

For simplicity we assume that  $(n/k)^{1/d} \in \mathbb{N}$ , so that  $\Lambda_n$  can be partitioned into volume- $2^{-d}k$  boxes.

**Definition 6.2** (Boxing scheme). Fix  $k \geq 1$  sufficiently large. Consider the boxes  $\mathcal{Q}_1, \dots, \mathcal{Q}_{2^d n/k}$  that partition  $\Lambda_n$ , and that are labeled so that  $\mathcal{Q}_j$  is adjacent to  $\mathcal{Q}_{j+1}$  for all  $j \in [2^d n/k - 1]$ . We write  $\mathcal{V}_{k,j}$  for the PPP restricted to the box  $\mathcal{Q}_j$ ,  $\mathcal{G}_{k,j}$  for the graph  $\mathcal{G}$  restricted to  $\mathcal{Q}_j$ . We call vertices with mark respectively below and at least  $A_{5.7} k^\gamma$  low-mark and high-mark vertices. We call  $\mathcal{C}_{k,j}^{(1)}[1, A_{5.7} k^\gamma]$  the local low-largest component of  $\mathcal{Q}_j$ . We also define

$$\mathcal{U}_j := \{v \in \mathcal{V}_{k,j}[A_{5.7} k^\gamma, \infty) : v \sim \mathcal{C}_{k,j}^{(1)}[1, A_{5.7} k^\gamma]\}. \quad (6.3)$$

We then define the local giant component of box  $\mathcal{Q}_j$  as

$$\text{LG}_j := \mathcal{U}_j \cup \mathcal{C}_{k,j}^{(1)}[1, A_{5.7} k^\gamma]. \quad (6.4)$$

In words,  $\mathcal{U}_j$  is the set of high-mark vertices in  $\mathcal{Q}_j$  that connect by at least one edge to the local low-largest component in the same box. So, together they form the local giant component  $\text{LG}_j$ . We mention that  $\mathcal{U}_j$  typically does not contain all high-mark vertices.

**Definition 6.3** (Revelment stages). Consider a KSRG under the same setting as Proposition 6.1 such that  $\zeta_{\text{hl}} > 0$ , with the boxing scheme from Definition 6.2. Initially, we reveal the realization of vertex set  $\mathcal{V}$ , and none of the edges. We define the following four edge-revelment stages:

- 1 Reveal edges between pairs of vertices both in the same box  $\mathcal{Q}_j$ .
- 2 Reveal edges between pairs of vertices both in  $\cup_{j \leq 2^d(n/k)} \text{LG}_j$ .
- 3 Reveal remaining edges between pairs of vertices both in  $\mathcal{V}_n \setminus (\cup_{j \leq 2^d(n/k)} \text{LG}_j)$
- 4 Reveal all remaining edges.

We write  $\text{stage-}j\text{-}\mathcal{G}_n$  for the graph obtained after the  $j$ th revelation stage for  $j \in \{1, 2, 3, 4\}$ .

Thus, in stage (1) we reveal all ‘local’ edges between vertices inside the same box. In stage (2), edges between local giants of different boxes. In stage (3), we reveal edges with both endpoints outside local giants in different boxes. Finally, in stage (4) we reveal all remaining edges: edges in different boxes where one endpoint is in a local giant, and the other endpoint is outside a local giant. The set  $\text{LG}_j$ , (the local-low largest component and high-mark vertices connecting to it in the same box) is determined during stage (1). So stages (2–4) are well-defined. These latter stages reveal edges between *different* boxes.

**Definition 6.4** (Stage (1) and (2) good events). Let  $\delta \in (0, 1/(2^d A_{5.7}))$  be a small constant. We call a box  $\mathcal{Q}_j$  1-good if it satisfies the event

$$\mathcal{A}_{1\text{-good}}(\mathcal{Q}_j) := \{\mathcal{Q}_j \text{ 1-good}\} := \{|\mathcal{C}_{k,j}^{(1)}[1, A_{5.7}k^\gamma]| > \delta k\} \cap \{|\mathcal{U}_j| > \delta k^\zeta\}. \quad (6.5)$$

For  $j < 2^d(n/k)$ , we call  $u \in \mathcal{U}_j$  box-wedging if  $u \sim \text{LG}_{j+1}$ . We say that stage-2- $\mathcal{G}_n$  is 2-good if

$$\{\text{stage-2-}\mathcal{G}_n \text{ 2-good}\} := \bigcap_{j \leq 2^d(n/k)} \mathcal{A}_{1\text{-good}}(\mathcal{Q}_j) \cap \bigcap_{j < 2^d(n/k)} \{\exists u \in \mathcal{U}_j : u \sim \text{LG}_{j+1}\} \quad (6.6)$$

Whether  $\mathcal{Q}_j$  is 1-good is entirely determined by the revealed edges in stage (1), and whether stage-2- $\mathcal{G}_n$  is 2-good, is entirely determined by edges revealed by the end of stage (2). The second intersection in (6.6) means that there is at least one box-wedging vertex in each local giant, implying that all local giants are part of the same component in  $\mathcal{G}_n$  that we will call the backbone.

**Claim 6.5** (Stage (1) error bound). *Consider a KS RG under the same setting as Proposition 6.1 such that  $\zeta_{\text{hl}} > 0$ . There exists a constant  $\varepsilon > 0$  such that for all  $k$  sufficiently large, for all  $j \leq 2^d(n/k)$ ,*

$$\mathbb{P}(\mathcal{Q}_j \text{ is 1-good}) \geq 1 - \exp(-\varepsilon k^\zeta).$$

**Proof** We estimate the probabilities of the two events in (6.5). By Corollary 5.7,

$$\mathbb{P}(\mathcal{C}_{k,j}[1, A_{5.7}k^\gamma] \leq \delta k) \leq \exp(-2^{-d\zeta_{\text{hh}}} k^{\zeta_{\text{hh}}} / A_{5.7}). \quad (6.7)$$

Next, consider  $v \in \mathcal{V}_n[A_{5.7}k^\gamma, \infty)$  in box  $\mathcal{Q}_j$ . We use that each vertex in  $\mathcal{C}_{k,j}[1, A_{5.7}k^\gamma]$  has mark at least 1, and that the distance between any two vertices in  $\mathcal{Q}_j$  is at most  $\sqrt{d}k^{1/d}$ . Thus, by the connection probability  $p$  in (4.6) (with  $\alpha < \infty$ ), conditional on the vertex set  $\mathcal{V}_n$  and the subgraph  $\mathcal{C}_{k,j}[1, A_{5.7}k^\gamma]$  having size at least  $\delta k$ ,

$$\begin{aligned} \mathbb{P}(v \approx \mathcal{C}_{k,j}[1, A_{5.7}k^\gamma] \mid |\mathcal{C}_{k,j}[1, A_{5.7}k^\gamma]| \geq \delta k, \mathcal{V}_n) \\ \leq \left(1 - p\left(1 \wedge \beta A_{5.7} \frac{k^\gamma}{d^{d/2}k}\right)^\alpha\right)^{\delta k}. \end{aligned}$$

If the minimum is at 1, then this probability is at most  $e^{-p\delta k}$ . If the minimum is at the second expression, we substitute  $\gamma = (\alpha - 1)/\alpha$ . When we bound  $1 - x \leq e^{-x}$ , the powers of  $k$  cancel, and the right-hand side is at most

$$\mathbb{P}(v \notin \mathcal{U}_j \mid |\mathcal{C}_{k,j}[1, A_{5.7}k^\gamma]| \geq \delta k, \mathcal{V}_n) \leq \exp(-p\beta A_{5.7}^\alpha/d^{d/2}). \quad (6.8)$$

In both cases, there is  $\delta_1 > 0$  such that any such  $v$  connects by an edge to  $\mathcal{C}_{k,1}^{(1)}[1, A_{5.7}k^\gamma]$  with probability at least  $\delta_1$ , independently of all other vertices. Now,  $\mathcal{V}_{k,j} \cap [A_{5.7}k^\gamma, \infty)$  is by (4.5) a PPP with intensity  $\text{Leb} \times F_W(dw)$ . Therefore, the vertices in  $\mathcal{V}_{k,1} \cap [A_{5.7}k^\gamma, \infty)$  that connect by an edge to  $\mathcal{C}_{k,1}^{(1)}[1, A_{5.7}k^\gamma]$  dominate a Poisson random variable with mean

$$\delta_1 \cdot \text{Vol}(\mathcal{Q}_j)(A_{5.7}k^\gamma)^{-(\tau-1)} = \delta_1 2^{-d} A_{5.7}^{-(\tau-1)} k^{1-\gamma(\tau-1)} = \Theta(k^\zeta).$$

Thus, by Lemma A.1, there exists  $\delta = \delta(A_{5.7}) > 0$  small enough so that

$$\mathbb{P}(|\mathcal{U}_j| \leq \delta k \mid |\mathcal{C}_{k,j}[1, A_{5.7}k^\gamma]| \geq \delta k) \leq \exp(-\varepsilon k^\zeta).$$

Combined with (6.7) and the definition of  $\{\mathcal{Q}_1 \text{ is 1-good}\}$  in Definition 6.4 this finishes the proof for  $k$  sufficiently large and  $\delta, \varepsilon > 0$  sufficiently small.  $\square$

**Claim 6.6** (Stage (2) error bound). *Consider a KSRG under the same setting as Proposition 6.1 such that  $\zeta_{\text{hl}} > 0$ . There exists a constant  $\varepsilon > 0$  such that for all  $k$  sufficiently large,*

$$\mathbb{P}(\text{stage-2-}\mathcal{G}_n \text{ is 2-good}) \geq 1 - 2^{d+1}(n/k) \exp(-\varepsilon k^\zeta).$$

**Proof** By Claim 6.5, each box  $\mathcal{Q}_i$  is 1-good with probability at least  $1 - \exp(-\varepsilon k^\zeta)$ . Denote the first event in (6.6) that all boxes  $\mathcal{Q}_j$  are 1-good by  $\mathcal{A}_{\text{sub}}$ . By a union bound,

$$\mathbb{P}(\mathcal{A}_{\text{sub}}) \geq 1 - 2^d(n/k) \exp(-\varepsilon k^\zeta). \quad (6.9)$$

We move on to estimating the probability of the second event in (6.6) conditioned on  $\mathcal{A}_{\text{sub}}$ . Fix some  $j < 2^d(n/k)$ . Before Stage (2), none of the edges of  $\mathcal{U}_j$  towards vertices in  $\mathcal{Q}_{j+1}$  have been revealed yet. The maximal Euclidean distance between two neighboring boxes is  $2\sqrt{d}k^{1/d}/2$ , so we can bound the probability that a single high-mark vertex is box-wedging using (4.6). When  $\mathcal{Q}_{j+1}$  is 1-good, by (6.5) there are at least  $\delta n$  vertices of mark at least 1 in  $\text{LG}_{j+1}$ , so

$$\mathbb{P}(u \in \mathcal{U}_j \text{ not box-wedging} \mid \mathcal{Q}_{j+1} \text{ 1-good}, \mathcal{V}_n) \leq \left(1 - p\left(1 \wedge \frac{\beta A_{5.7}k^\gamma}{d^{d/2}k}\right)^\alpha\right)^{\delta k}. \quad (6.10)$$

Then, we can use that there are at least  $\delta k^\zeta$  candidate vertices to be box-wedging in a 1-good  $\mathcal{Q}_j$ , and that edges are present independently given the vertex set  $\mathcal{V}$ . Thus,

$$\begin{aligned} & \mathbb{P}(\nexists u \in \mathcal{U}_j : u \text{ box-wedging} \mid \mathcal{Q}_j, \mathcal{Q}_{j+1} \text{ 1-good}, \mathcal{V}_n) \\ & \leq \left(1 - p\left(1 \wedge \frac{\beta A_{5.7} k^\gamma}{d^{d/2} k}\right)^\alpha\right)^{(\delta k) \cdot (\delta k^\zeta)}. \end{aligned} \quad (6.11)$$

If the minimum on the right-hand side is at 1, this probability is at most  $\exp(-\varepsilon k^{1+\zeta})$ . If the minimum is at the second expression, using that  $1 - x \leq \exp(-x)$  and  $\gamma = 1 - 1/\alpha$ , the powers of  $k$  simplify since  $(k^\gamma/k)^\alpha \cdot k = 1$ , so this probability is at most  $\exp(-\varepsilon k^\zeta)$  for  $\varepsilon > 0$  small enough. So, integrating over the realizations  $\mathcal{V}_n$  and a union bound over all  $2^d(n/k)$  boxes, the probability that there is at least one box  $\mathcal{Q}_j$  with no box-wedging vertices is at most  $2^d(n/k) \exp(-\varepsilon k^\zeta)$ . Combined with (6.6) and (6.9) the claim follows.  $\square$

**Claim 6.7** (Stage (3–4) error bound: second-largest component). *Consider a KSRG under the same setting as Proposition 6.1 such that  $\zeta_{\text{hl}} > 0$ . There exists a constant  $\varepsilon > 0$  such that for all  $k$  sufficiently large and  $n > k$*

$$\mathbb{P}(|\mathcal{C}_n^{(2)}| > k \mid \text{stage-2-}\mathcal{G}_n \text{ is 2-good}) \leq (n/k) \exp(-\varepsilon k^\zeta). \quad (6.12)$$

**Proof** We condition on the full realization of the stage-2- $\mathcal{G}_n$  that is 2-good. By (6.6), any such realization satisfies that all local giants  $(\text{LG}_j)_{j \leq 2^d(n/k)}$  have size at least  $\delta k$  and they are all connected by box-wedging vertices. This gives a component in  $\mathcal{G}_n$  of size at least  $\delta n$  that contains these local giants. We call this component the backbone  $\text{BB} = \text{BB}_{n,k}$ . In Stage (3) we expose all edges with both endpoints outside the local giants. If no component of size larger than  $k$  except for the backbone remains after Stage (3), then we are done: In stage (4) all remaining edges to be revealed have one endpoint in local giants, hence in BB, and one outside.

We may interpret now (6.12) as follows: the second-largest component being larger than  $k$  means that there is at least one component above size  $k$  that is not BB. If after Stage (3) there is a component  $\mathcal{C} \neq \text{BB}$  above size  $k$ , then we aim to show that this component merges with BB in Stage (4) with probability at least  $1 - \exp(-\varepsilon k^\zeta)$ . Since there are at most  $(n/k)$  components above size  $k$  after Stage (3), a union bound over all such components will yield (6.12).

So, suppose that there exists a component  $\mathcal{C} \neq \text{BB}$  of size larger than  $k$  after Stage (3) and fix a realization of stage-2- $\mathcal{G}_n$ . After Stage (3) we revealed already all within-box edges, but we did not reveal edges between  $\mathcal{C}$  and local giants in neighboring boxes ‘around’ vertices of  $\mathcal{C}$ . Since the number of boxes is at least  $2^d(n/k)$  by Definition 6.2, a subbox always has a neighboring box. Let  $j(v)$  denote index of the box that vertex  $v \in \mathcal{C}$  belongs to, which is a.s. unique. Since all boxes are 1-good, by (6.5), for each  $v \in \mathcal{C}$ , there are at least  $\delta k^\zeta$  high-mark vertices at distance at most  $\sqrt{d}k^{1/d}$  from  $v$  in the vertex set  $\mathcal{U}_{j(v)+1} \subseteq \text{LG}_{j(v)+1}$  in the neighboring boxes  $\mathcal{Q}_{j(v)-1}$  and  $\mathcal{Q}_{j(v)+1}$ . Each vertex of  $v$  has mark at least 1. Hence, for a given component  $\mathcal{C}$ , by the same

calculations as in the proof of Claim 6.6 around (6.10)–(6.11),

$$\begin{aligned} \mathbb{P}(\mathcal{C} \approx \text{BB} \mid 2\text{-good-}\mathcal{G}_n, \mathcal{V}_n, \text{stage-3-}\mathcal{G}_n) &\leq \left(1 - p\left(1 \wedge \beta A_{5.7} \frac{k^\gamma}{d^{d/2}k}\right)^\alpha\right)^{k \cdot (\delta k^\zeta)} \\ &\leq \exp(-\varepsilon k^\zeta). \end{aligned}$$

for  $\varepsilon > 0$  sufficiently small. Here we used that  $\gamma = 1 - 1/\alpha$  and thus some powers of  $k$  cancel. By a union bound over all at most  $|\mathcal{V}_n|/k$  components of size at least  $k$ ,

$$\begin{aligned} \mathbb{P}(|\mathcal{C}_n^{(2)}| > k) \cap \{\text{stage-2-}\mathcal{G}_n \text{ is 2-good}\} &\leq \mathbb{E}[\mathbb{P}(|\mathcal{C}_n^{(2)}| \geq k \mid 2\text{-good-}\mathcal{G}_n, \mathcal{V}_n, \text{stage-3-}\mathcal{G}_n)] \\ &\leq \mathbb{E}[|\mathcal{V}_n|/k] \exp(-\varepsilon k^\zeta) \leq (n/k) \exp(-\varepsilon k^\zeta). \end{aligned}$$

This finishes the proof of (6.12).  $\square$

We are ready to prove Proposition 6.1 that bounds  $\mathbb{P}(|\mathcal{C}_n^{(2)}| \geq k)$ .

*Proof of Proposition 6.1 when  $\zeta_{\text{hl}} > 0$*  By the law of total probability,

$$\mathbb{P}(|\mathcal{C}_n^{(2)}| \geq k) \leq \mathbb{P}(|\mathcal{C}_n^{(2)}| \geq k) \cap \{\text{stage-2-}\mathcal{G}_n \text{ is 2-good}\} + \mathbb{P}(\text{stage-2-}\mathcal{G}_n \text{ is not 2-good}).$$

Proposition 6.1 follows from Claims 6.6–6.7.  $\square$

We now give a sketch on how to modify the above proof when  $\zeta_{\text{ll}} > 0$ .

*Proof of Proposition 6.1 when  $\zeta_{\text{ll}} > 0$*  The proofs are identical for KSRGs on PPPs and on  $\mathbb{Z}^d$ . We follow the  $\zeta_{\text{hl}}$ -computations. By Corollary 5.7,

$$\mathbb{P}(|\mathcal{C}_n^{(1)}[1, A_{5.7}]| \leq \frac{1}{A_{5.7}}n) \leq \exp\left(-\frac{1}{A_{5.7}}n^{\zeta_{\text{ll}}}\right). \quad (6.13)$$

We set  $\text{LG}_j := \mathcal{C}_n^{(1)}[1, A_{5.7}]$  to be the local giant of box  $\mathcal{Q}_j$ . We ignore high-mark vertices here. The revealment scheme in Definition 6.3 is then well-defined with this  $\text{LG}_j$ . We modify the box being 1-good in (6.5) by only requesting the first event  $|\text{LG}_j| = |\mathcal{C}_n^{(1)}[1, A_{5.7}]| \geq \delta k$  there. We define  $u \in \text{LG}_j$  to be box-wedging if  $u \sim \text{LG}_{j+1}$ . In the definition of the stage 2- $\mathcal{G}_n$  being 2-good in (6.6), we then require at least one box-wedging vertex in each box:

$$\bigcap_{j < 2^d(n/k)} \{\exists u \in \text{LG}_j : u \sim \text{LG}_{j+1}\}.$$

Claim 6.5 follows then from Corollary 5.7 as before. We describe how to prove the equivalent of Claim 6.6 for this regime. Estimating that all boxes are 1-good goes as before in (6.9). For the second part, estimating that each box has a box-wedging vertex, we use only the fact that vertices have weight at least 1. We get

$$\mathbb{P}(\nexists u \in \text{LG}_j : u \sim \text{LG}_{j+1} \mid \mathcal{Q}_j, \mathcal{Q}_{j+1} \text{ are 1-good}) \leq \left(1 - p\left(1 \wedge \beta \frac{1}{d^{d/2}k}\right)^\alpha\right)^{(\delta k)^2}.$$



If the first term is the minimum, this probability is at most  $\exp(-\varepsilon k^2) \leq \exp(-\varepsilon k^{2-\alpha}) = \exp(-\varepsilon k^{\zeta_{\text{ll}}})$  for  $\varepsilon > 0$  sufficiently small. If the second term is the minimum, this probability is also at most  $\exp(-\varepsilon k^{2-\alpha}) = \exp(-\varepsilon k^{\zeta_{\text{ll}}})$ . The rest of the claim is as for the hl-regime. For Claim 6.7, we use again that each local giant contains at least  $\delta k$  vertices of mark at least 1. Thus,

$$\mathbb{P}(C \approx \text{BB} \mid 2\text{-good-}\mathcal{G}_n, \mathcal{V}_n) \leq \left(1 - p\left(1 \wedge \beta \frac{1}{d^{d/2}k}\right)^\alpha\right)^{(\delta k) \cdot k}.$$

As before, the right-hand side is at most  $\exp(-\varepsilon k^{2-\alpha}) = \exp(-\varepsilon k^{\zeta_{\text{ll}}})$ . The proof of the claim is finished as before.  $\square$

We move towards the proofs of Theorem 2.10. We recall a proposition from our recent paper to prove the lower bounds. To make notation lighter, we state it only for the case that  $\zeta_{\text{long}} > 0$ . We recall from (2.16) that  $\max(\mathcal{Z})$  is the maximum and  $m_{\mathcal{Z}}$  is the multiplicity of the maximum in the set  $\mathcal{Z} = \{\zeta_{\text{short}}, \zeta_{\text{ll}}, \zeta_{\text{hl}}, \zeta_{\text{hh}}\}$ . We write here  $C_n(0)[1, A)$  for the component of a vertex with unknown weight at  $0 \in \mathbb{R}^d$  inside the vertex set on  $\Lambda_n \times [1, A)$ , with the idea that if  $w_0 > A$  then  $C_n(0)[1, A) := \emptyset$ .

**Proposition 6.8** (Lower bound prerequisite [48, Prop. 7.1]). *Consider a supercritical KSRG on a PPP or on  $\mathbb{Z}^d$  satisfying Assumption 2.2 with  $\zeta_{\text{long}} > 0$ . Assume that there exist constants  $\eta', \rho > 0$  such that for all  $n$  sufficiently large,*

$$\mathbb{P}^0(|C_n(0)[1, (\log n)^{\eta'}| \geq \rho n) \geq \rho. \quad (6.14)$$

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

$$\mathbb{P}^0(|C_n(0)| > k, 0 \notin C_n^{(1)}) \geq \exp(-Ak^{\max(\mathcal{Z})}(\log k)^{m_{\mathcal{Z}}-1}). \quad (6.15)$$

*Moreover, there exist  $\delta, \varepsilon > 0$ , such that for all  $n$  sufficiently large,*

$$\mathbb{P}(|C_n^{(2)}| > ((\varepsilon \log n)/(\log \log n)^{m_{\mathcal{Z}}-1})^{1/\max(\mathcal{Z})}) \geq 1 - n^{-\delta}. \quad (6.16)$$

The proof of this proposition can be found in [48, Section 7]. The key idea is to find a localized component containing 0 in the ball of radius  $\Theta(k)$  around 0, simultaneously ensuring that no edges leave this ball. The probabilistic cost of this event gives the lower bound on the cluster size-decay. The lower bound on the second-largest component can be derived from there, using a tessellation of the original space with boxes of volume  $\Theta(n^\varepsilon)$ . Each tile corresponds to a repeated trial of the local event, and we prove that the small balls inside the tiles do not connect to other tiles with sufficiently high probability. See [48, Section 7.3] for more details. The next claim verifies the requirement (6.14) for applying Proposition 6.8.

**Claim 6.9** (The prerequisite holds). *Consider a KSRG under the same setting as Theorem 2.10 and Theorem 2.9. There exist constants  $\eta', \rho > 0$  such that*

$$\mathbb{P}^0(|C_n(0)[1, (\log n)^{\eta'}| \geq \rho n) \geq \rho. \quad (6.17)$$

**Proof** We prove the bound (6.17) for KSRGs on a Poisson point process and leave it to the reader to adapt it to long-range percolation. We rely on Corollary 5.8. Fix  $(\zeta, \gamma) \in \{(\zeta_{\text{hl}}, \gamma_{\text{hl}}), (\zeta_{\text{ll}}, \gamma_{\text{ll}})\}$  such that  $\zeta > 0$ . Let  $A, \varepsilon > 0$  be so that (5.38) holds with  $C = 2$ . For any  $u \in \mathcal{V}_n$ , let  $\tilde{\mathcal{C}}_n(u)$  be the component of  $u$  in the induced graph in  $\Lambda_n \times [1, (A \log n)^{\gamma/\zeta}]$ ; and let  $\tilde{\mathcal{C}}_{2^{d_n}, u}(u)$  be the connected component of  $u$  in  $\Lambda_{2^{d_n}}(u) \times [1, (A \log n)^{\gamma/\zeta}]$ . Since  $\Lambda_n \subseteq \Lambda_{2^{d_n}}(u)$ , clearly  $\tilde{\mathcal{C}}_n(u) \subseteq \tilde{\mathcal{C}}_{2^{d_n}, u}(u)$ . If the giant of  $\Lambda_n \times [1, (A \log n)^{\gamma/\zeta}]$  has size at least  $\varepsilon n$ , then there are at least  $\varepsilon n$  vertices in a component of size at least  $\varepsilon n$  in  $\Lambda_n \times [1, (A \log n)^{\gamma/\zeta}]$ . This happens with probability at least  $1 - n^{-2}$  by (5.38). So,

$$(1 - n^{-2})\varepsilon n \leq \mathbb{E} \left[ \sum_{u \in \mathcal{V}_n} \mathbb{1}_{\{|\tilde{\mathcal{C}}_n(u)| \geq \varepsilon n\}} \right] \leq \mathbb{E} \left[ \sum_{u \in \mathcal{V}_n} \mathbb{1}_{\{|\tilde{\mathcal{C}}_{2^{d_n}, u}(u)| \geq \varepsilon n\}} \right]. \quad (6.18)$$

We now show that

$$\mathbb{E} \left[ \sum_{u \in \mathcal{V}_n} \mathbb{1}_{\{|\tilde{\mathcal{C}}_{2^{d_n}, u}(u)| \geq \varepsilon n\}} \right] = n \mathbb{P}^0(\mathcal{C}_{2^{d_n}}(0)[1, (A \log n)^{\gamma/\zeta}] \geq \varepsilon n). \quad (6.19)$$

For this, we use an alternative KSRG construction. Consider the collection  $\mathbf{U} = (U_{ij})_{j>i \geq 1}$  of independent  $\text{Unif}[0, 1]$  random variables, independent of the PPP  $\mathcal{V}_n$ . We can use  $\mathbf{U}$  to encode the presence of edges: given the PPP  $\mathcal{V}$ , order the points with respect to their distance to the origin, and include an edge between the  $i$ th and  $j$ th vertex  $u_i$  and  $u_j$  if  $U_{ij} \leq p(u_i, u_j)$ . The function  $\mathbb{1}_{\{|\tilde{\mathcal{C}}_{2^{d_n}, u}(u)| \geq \varepsilon n\}}$  is determined by the realization of the PPP  $\mathcal{V}_n$  and  $\mathbf{U}$  (i.e., it is measurable w.r.t. the sigma-algebras generated by these processes/collections). By Mecke's formula and Fubini's theorem,

$$\begin{aligned} \mathbb{E} \left[ \sum_{u \in \mathcal{V}_n} \mathbb{1}_{\{|\tilde{\mathcal{C}}_{2^{d_n}, u}(u)| \geq \varepsilon n\}} \right] &= \mathbb{E} \left[ \mathbb{E} \left[ \sum_{u \in \mathcal{V}_n} \mathbb{1}_{\{|\tilde{\mathcal{C}}_{2^{d_n}, u}(u)| \geq \varepsilon n\}} \mid \mathbf{U} \right] \right] \\ &= \mathbb{E} \left[ \int_{x \in \Lambda_n} \mathbb{P}^x(\tilde{\mathcal{C}}_{2^{d_n}, u}(u) \geq \varepsilon n \mid \mathbf{U}, x_u = x) dx \right] \\ &= \int_{x \in \Lambda_n} \mathbb{P}^x(\tilde{\mathcal{C}}_{2^{d_n}, u}(u) \geq \varepsilon n \mid x_u = x) dx. \end{aligned}$$

The box  $\Lambda_{2^{d_n}}(u)$  is centered at  $u$ . So, by translation invariance, the integrand is constant and equal to  $\mathbb{P}^0(\tilde{\mathcal{C}}_{2^{d_n}}(0) \geq \varepsilon n)$ . Since  $\tilde{\mathcal{C}}_{2^{d_n}}(0) = \mathcal{C}_{2^{d_n}}(0)[1, (A \log n)^{\gamma/\zeta}]$  by definition above (6.18), this proves (6.19). Combined with (6.18), this proves (6.17) for any  $\eta' > \gamma/\zeta$ ,  $\rho = 2^{-d}\varepsilon$ , and any  $n$  sufficiently large. The statement is trivial for smaller  $n$  by decreasing  $\rho$ , finishing the proof.  $\square$

We are ready to prove Theorem 2.10 on the second-largest component.

**Proof of Theorem 2.10** We give the proof for KSRGs on PPPs only: the proof for long-range percolation  $\mathbb{Z}^d$  is identical to the proof of the  $\zeta_{\text{ll}}$ -regime for KSRGs. Assume  $\zeta_{\text{hl}}$  or  $\zeta_{\text{ll}}$  is strictly positive and maximal in  $\mathcal{Z} = \{\zeta_{\text{short}}, \zeta_{\text{ll}}, \zeta_{\text{hl}}, \zeta_{\text{hh}}\}$ . The lower bound

holds by Proposition 6.8, since its prerequisite (6.14) is satisfied by Claim 6.9. The upper bound follows from Proposition 6.1, substituting  $k = k_n = (A \log n)^{1/\max(\zeta_{\text{hl}}, \zeta_{\text{ll}})}$  for a large constant  $A = A(\delta) > 0$ .  $\square$

## 6.2 Cluster-size decay and law of large numbers

We proceed to the proof of Theorem 2.9 on the cluster-size decay. The upper bound (6.1) is the main prerequisite (6.20) of the following general proposition. We proved this proposition for KSRGs in general in [48]. We recall that  $\mathbb{P}^x$  is the Palm measure of the PPP with a vertex at  $x$ .

**Proposition 6.10** (Prerequisites for cluster-size decay [48, Proposition 6.1]). *Consider a supercritical KSRG on a PPP or on  $\mathbb{Z}^d$  satisfying Assumption 2.2. Assume that there exist  $\zeta, \eta', c, c', M_c > 0$ , and a function  $n_0(k) = O(k^{1+c'})$  such that for all  $k$  sufficiently large and whenever  $n \in [n_0(k), \infty)$ , with  $\bar{w}(n, c) := M_c(\log n)^{\eta'}$  it holds for all  $x \in \Lambda_n$ ,*

$$\mathbb{P}^x(|\mathcal{C}_n^{(2)}| > k) \leq n^{c'} \exp(-ck^\zeta), \quad (6.20)$$

$$\mathbb{P}^x(|\mathcal{C}_n^{(1)}| < n^c) \leq n^{-1-c}, \quad (6.21)$$

$$\mathbb{P}^x(\exists v \in \mathcal{V}_n[\bar{w}(n, c), \infty) : v \notin \mathcal{C}_n^{(1)}) \leq n^{-c}. \quad (6.22)$$

Then there exists a constant  $A > 0$  such that for all  $k$  sufficiently large and  $n \in [n_0(k), \infty]$ ,

$$\mathbb{P}^0(|\mathcal{C}_n(0)| > k, 0 \notin \mathcal{C}_n^{(1)}) \leq \exp(-\frac{1}{A}k^\zeta), \quad (6.23)$$

and

$$\frac{|\mathcal{C}_n^{(1)}|}{n} \xrightarrow{\mathbb{P}} \theta(\beta, p, \alpha, \tau, \sigma), \quad \text{as } n \rightarrow \infty. \quad (6.24)$$

In our current setting, (6.1) in Proposition 6.1 proves the first prerequisite and Corollary 5.8 proves the second, and implies the third prerequisite. We spell out the details. This is similar to [48, Proposition 5.12].

**Claim 6.11** *Consider a supercritical KSRG on a PPP satisfying Assumption 2.2 or consider supercritical long-range percolation on  $\mathbb{Z}^d$ . Assume  $\zeta_{\text{type}} = \max(\zeta_{\text{ll}}, \zeta_{\text{hl}}) > 0$  for some  $\text{type} \in \{\text{ll}, \text{hl}\}$ . For each constant  $C > 0$ , there exists a constant  $A > 0$  such that all  $n$  sufficiently large,*

$$\mathbb{P}(\exists v \in \mathcal{V}_n[(A \log n)^{\eta_{\text{type}}/\zeta_{\text{type}}}, \infty) : v \notin \mathcal{C}_n^{(1)}) \leq n^{-C}.$$

The statement remains true for the Palm-version  $\mathbb{P}^x$  for any  $x \in \mathbb{R}^d$ .

**Proof** We omit the subscript type in the following proof for  $\text{type} \in \{\text{ll}, \text{hl}\}$ . Let  $A = A(C) > 0$  be sufficiently large. Using the upper bound on  $|\mathcal{C}_n^{(2)}|$  in Proposition 6.1, and the lower bound on  $|\mathcal{C}_n^{(1)}|$  from Corollary 5.8, we obtain

$$\mathbb{P}^x(\{|\mathcal{C}_n^{(2)}| \leq (A \log n)^{1/\zeta}\} \cap \{|\mathcal{C}_n^{(1)}| \geq (1/A)n\}) \geq 1 - o(n^{-C}). \quad (6.25)$$

Hence,

$$\begin{aligned} & \mathbb{P}^x(\exists v \in \mathcal{V}_n[(A \log n)^{\eta/\zeta}, \infty) \setminus \mathcal{C}_n^{(1)}) \\ & \leq \mathbb{P}^x \left( \begin{aligned} & \{\exists v \in \mathcal{V}_n[(A \log n)^{\eta/\zeta}, \infty) \setminus \mathcal{C}_n^{(1)}\} \\ & \cap \{|\mathcal{C}_n^{(2)}| \leq (A \log n)^{1/\zeta}\} \\ & \cap \{|\mathcal{C}_n^{(1)}| \geq \frac{1}{A}n\} \end{aligned} \right) + o(n^{-C}). \end{aligned}$$

On the intersection of the three events on the right-hand side, there is no linear-sized component in  $\mathcal{G}_n$  that contains all vertices with mark above  $(A \log n)^{\eta/\zeta}$ , which is exactly the negation of the event in (5.39). This has probability at most  $o(n^{-C})$  by (5.39) in Corollary 5.8, finishing the proof.  $\square$

Propositions 6.8 and 6.10 give the respective lower and upper bounds for Theorem 2.9.

**Proof of Theorem 2.9** We prove Theorem 2.9 when  $\max(\zeta_{hl}, \zeta_{ll}) > \max(\zeta_{hh}, 0)$ . The lower bound follows from (6.15) in Proposition 6.8, since the prerequisite (6.14) is satisfied by Claim 6.9. For the upper bound we verify the three prerequisites needed in Proposition 6.10. Prerequisite (6.20) about the second-largest component is proven in Proposition 6.1; prerequisite (6.21) requiring that the largest component is at least polynomial in size follows from Corollary 5.8; prerequisite (6.22) is proven in Claim 6.11. The upper bound in Theorem 2.9 then follows from (6.23).  $\square$

### 6.3 Lower tail of large deviations

We are ready to combine everything and give the proof of the lower tail of large deviations for the giant in (2.21). We cite the lower bound from [48]. We write  $\mathcal{Z} = \{\zeta_{hh}, \zeta_{hl}, \zeta_{ll}, \zeta_{short}\}$ , with maximum value  $\zeta_*$ , and  $m_{\mathcal{Z}}$  the multiplicity of the maximum, see (2.16).

**Theorem 6.12** (Lower bound for the lower tail [48, Theorem 2.7]). *Consider a supercritical KSRG on a PPP satisfying Assumption 2.2 or consider supercritical long-range percolation on  $\mathbb{Z}^d$ . There exists a constant  $A > 0$  such that for all  $\rho > 0$ , and  $n$  sufficiently large,*

$$\mathbb{P}(|\mathcal{C}_n^{(1)}| < \rho n) \geq \exp\left(-\frac{1}{A\rho}n^{\zeta_*}(\log n)^{m_{\mathcal{Z}}-1}\right).$$

**Proof of Theorem 2.7** The lower bound follows from Theorem 6.12, since the above theorem is valid for all  $\rho > 0$  (but is far from being sharp when  $\rho \geq \theta$ ). When  $\max(\zeta_{hl}, \zeta_{ll}) > 0$  we argue as follows for the upper bound: since we verified the three prerequisites needed in Proposition 6.10 already in the proof of Theorem 2.9 above, the second conclusion of Proposition 6.10 gives that

$$|\mathcal{C}_n^{(1)}|/n \xrightarrow{\mathbb{P}} \theta.$$

Hence, we may choose any  $\tilde{\rho} \in (\rho, \theta)$ , and it follows that  $\{|\mathcal{C}_n^{(1)}| \geq \tilde{\rho}n\}$  whp. All highest-mark vertices are in the giant whp by Claim 6.11. This proves the requirement (5.3) in Proposition 5.1, i.e.,  $\mathbb{P}(A_{\text{giant}}^{\text{type}}(\tilde{\rho})) \rightarrow 1$  holds. Applying Proposition 5.1 for  $\varepsilon > 0$  such that  $\rho = (1 - \varepsilon)\tilde{\rho}$  proves the upper bound when  $\max(\zeta_{\text{hl}}, \zeta_{\text{ll}}) > 0$ .

When  $\zeta_{\text{hh}} > 0$ ,  $|\mathcal{C}_n^{(1)}|$  contains all vertices of mark above some polylog of  $n$  by [48, Proposition 5.12], and an LLN for  $|\mathcal{C}_n^{(1)}|$  also holds, see [48, Corollary 2.3]. So, the same reasoning as above can be applied also when  $\zeta_{\text{hh}} > 0$ . Since  $\zeta_{\text{long}} = \max(\zeta_{\text{hl}}, \zeta_{\text{hh}}, \zeta_{\text{ll}})$  by Claim 2.3, this concludes the proof of Theorem 2.7.  $\square$

## 7 Upper tail of large deviations

In this section we deal with the upper tail, that is, with the event of the giant being too large, and prove Theorems 2.4 and 2.5 following the roadmap outlined in Section 3.1. We first show a Law of Large Numbers (LLN) for components of size  $\ell$  that applies to both homogeneous models, i.e.,  $\tau = \infty$ , and degree-inhomogeneous models, i.e.,  $\tau < \infty$ . Afterwards, we obtain bounds on the convergence rate of this LLN. We prove that the concentration is exponentially fast when  $\tau = \infty$ , and show that Theorem 2.4 follows. Then we analyze the effect of the presence of high-mark vertices (when  $\tau < \infty$ ) to prove Theorem 2.5.

### 7.1 Concentration of size- $\ell$ components

We introduce notation for the probability of the component of the origin being of size  $\ell$ , and for the total number of size- $\ell$  components in the graph induced by vertices with mark at most  $\bar{w}_n$ :

$$\begin{aligned}\theta_\ell &:= \mathbb{P}^0(|\mathcal{C}(0)| = \ell), \\ S_{n,\ell}(\bar{w}_n) &:= |\{\text{comp. } \mathcal{C} \text{ in } \mathcal{G}_n[1, \bar{w}_n) : |\mathcal{C}| = \ell\}|, \\ S_{n,\ell} &:= S_{n,\ell}(\infty).\end{aligned}\tag{7.1}$$

Clearly  $\sum_{1 \leq \ell < \infty} \theta_\ell = \mathbb{P}^0(|\mathcal{C}(0)| < \infty) = 1 - \theta$ . In models with all marks equal 1 (e.g. long-range percolation) it holds  $S_{\ell,n}(\bar{w}_n) = S_{\ell,n}$  for any  $\bar{w}_n > 1$ .

**Lemma 7.1** (LLN for size- $\ell$  components). *Consider a KSRG on a PPP or on  $\mathbb{Z}^d$  satisfying Assumption 2.2. Let  $\underline{w}_n = \omega(n^{1/(\tau-1)})$  be a sequence. Then, for every fixed  $\ell \in \mathbb{N}$ ,*

$$\frac{\ell S_{n,\ell}(\underline{w}_n)}{n} \xrightarrow{\mathbb{P}} \theta_\ell, \quad \text{as } n \rightarrow \infty.\tag{7.2}$$

**Proof** We give the proof for KSRGs on a PPP with  $\tau < \infty$ , i.e., non-constant marks. The proof is analogous when  $\tau = \infty$  or when vertex locations are given by  $\mathbb{Z}^d$ . By the intensity of  $\mathcal{V}_n$  in (4.5),  $\Lambda_n$  contains no vertices of mark at least  $\underline{w}_n$  whp. If such vertices are absent, then  $S_{n,\ell}(\underline{w}_n) = S_{n,\ell}(\infty)$ . Therefore, for any  $\psi > 0$ ,

$$\begin{aligned}
\mathbb{P}(|\ell S_{n,\ell}(\underline{w}_n)/n - \theta_\ell| > \psi) &\leq \mathbb{P}(|\ell S_{n,\ell}(\infty)/n - \theta_\ell| > \psi) \cap \mathcal{V}_n[\underline{w}_n, \infty) = \emptyset) \\
&\quad + \mathbb{P}(\mathcal{V}_n[\underline{w}_n, \infty) \neq \emptyset) \\
&\leq \mathbb{P}(|\ell S_{n,\ell}(\infty)/n - \theta_\ell| > \psi) + o(1).
\end{aligned} \tag{7.3}$$

We use local convergence to show that the probability on the right-hand side vanishes. We refer to [72] for an elaborate discussion on local convergence and give a short definition.

A rooted graph is a couple  $(G, o)$  of a graph  $G$  and some, possibly random, distinguished vertex  $o$  of  $G$ , which we call the root of  $G$ . A finite rooted graph  $(G, o)$  is said to be uniformly rooted, if  $o$  is chosen uniformly at random among the vertices of  $G$ . Let  $\mathcal{G}_\star$  be the space of all rooted locally finite graphs. Let  $B_G(v, r)$  denote the induced subgraph of  $G$  on all vertices that are at graph distance at most  $r$  from a vertex  $v$ . For  $(G_1, o_1), (G_2, o_2) \in \mathcal{G}_\star$ , we define  $R(G_1, G_2)$  to be the largest value  $r \in \mathbb{N}$  such that  $B_{G_1}(o_1, r)$  is isomorphic to  $B_{G_2}(o_2, r)$ : there exists a bijection  $\phi : V(B_{G_1}(o_1, r)) \mapsto V(B_{G_2}(o_2, r))$  such that  $\phi(o_1) = o_2$  and  $\{u, v\}$  is an edge in  $B_{G_1}(o_1, r)$  if and only if  $\{\phi(u), \phi(v)\}$  is an edge in  $B_{G_2}(o_1, r)$ . We set  $d((G_1, o_1), (G_2, o_2)) := 1/(1 + R(G_1, o_1), (G_2, o_2))$  for the distance between two rooted graphs. A sequence of uniformly rooted graphs  $(G_n, o_n)_{n \geq 1}$  converges locally in probability towards  $(G_\infty, o)$  having law  $\mu$ , if for every bounded and continuous function  $h : \mathcal{G}_\star \mapsto \mathbb{R}$ ,

$$\mathbb{E}[h(G_n, o_n) \mid G_n] \xrightarrow{\mathbb{P}} \mathbb{E}_\mu[h(G_\infty, o)], \quad \text{as } n \rightarrow \infty, \tag{7.4}$$

where the expectation on the left-hand side is only with respect to the uniform root. (The limit object  $(G_\infty, o)$  may be random and infinite, we here omit this discussion).

Considering KSRGs in Definition 2.1, the sequence  $(\mathcal{G}_n)_{n \geq 1}$  of finite KSRGs converges locally in probability to the infinite graph  $\mathcal{G}_\infty$  conditioned to contain a vertex at the origin, the root, see [73]. We let  $h(G_n, o_n) = \mathbb{1}_{\{|\mathcal{C}(o_n)|=\ell\}}$  be the indicator that the component of  $o_n$  has size  $\ell$ . Then,  $h$  is bounded, and continuous on the Polish space  $(\mathcal{G}_\star, d)$ : it only depends on the induced subgraph up to distance  $\ell + 1$  from the root, so  $h(G_1, o_1) = h(G_2, o_2)$  for any two rooted graphs  $(G_1, o_1), (G_2, o_2)$  within distance at most  $1/(\ell + 2)$ . So (7.4) applies and the right-hand side there equals  $\mathbb{P}^0(|\mathcal{C}(0)| = \ell) = \theta_\ell$ . Instead of explicitly writing out the left-hand side, we first notice that  $\sum_{o_n \in \mathcal{V}_n} h(G_n, o_n) = \ell S_{n,\ell}(\infty)$ . Therefore,

$$\frac{\ell S_{n,\ell}(\infty)}{n} = \frac{|\mathcal{V}_n|}{n} \frac{1}{|\mathcal{V}_n|} \sum_{o_n \in \mathcal{V}_n} h(G_n, o_n) = \frac{|\mathcal{V}_n|}{n} \cdot \mathbb{E}[h(G_n, o_n) \mid G_n] \xrightarrow{\mathbb{P}} 1 \cdot \theta_\ell, \tag{7.5}$$

where we used that the ratio  $|\mathcal{V}_n|/n$  tends to one in probability. Hence, the probability on the right-hand side in (7.3) tends to zero as  $n$  tends to infinity. This finishes the proof for PPP vertex sets. When the locations are given by  $\mathbb{Z}^d$ , or  $\tau = \infty$ , the result follows analogously.  $\square$

## 7.2 Homogeneous models

We prove a bound on the convergence rate in Lemma 7.1 using a boxing argument.

**Lemma 7.2** (Concentration for size- $\ell$  components). *Consider long-range percolation with  $\alpha > 1$  or Bernoulli bond percolation on  $\mathbb{Z}^d$ . For all fixed  $\ell \in \mathbb{N}$  and  $\psi > 0$ , there exists a constant  $\varepsilon_{\ell, \psi} > 0$  such that for all  $n \geq 1$ ,*

$$\mathbb{P}\left(|S_{n, \ell}(\infty)/(n\theta_{\ell}/\ell) - 1| \leq \psi\right) \geq 1 - \exp(-\varepsilon_{\ell, \psi}n). \quad (7.6)$$

**Proof** Let  $k$  be a large constant, and w.l.o.g. assume that  $\Lambda_n$  can be partitioned into  $(n/k)$  volume- $k$  subboxes. Let  $S_{k, \ell}^{(j)}$  denote the number of size- $\ell$  components in the induced subgraph inside the  $j$ th box. Then, by Lemma 7.1, for a small constant  $\psi_1 > 0$ , independently across boxes it holds that

$$\mathbb{P}(\mathcal{A}_{\ell}(j, k)) := \mathbb{P}\left(|S_{k, \ell}^{(j)}/(k\theta_{\ell}/\ell) - 1| \leq \psi_1\right) \geq 1 - \psi_1/2.$$

The number of boxes for which this event holds dominates a binomial random variable with parameters  $n/k$  and  $1 - \psi_1/2$ . So if we set

$$\mathcal{A}_{\text{sub}}(k) := \left\{ \left| \sum_{j \in [n/k]} \mathbb{1}_{\{\mathcal{A}_{\ell}(j, k)\}} - n/k \right| \leq \psi_1 n/k \right\},$$

then by a Chernoff bound, there exists a constant  $c = c(\psi_1) > 0$  such that

$$\mathbb{P}(\neg \mathcal{A}_{\text{sub}}(k)) \leq \exp(-c(n/k)). \quad (7.7)$$

On  $\mathcal{A}_{\text{sub}}(k)$ , since  $(n/k)k\theta_{\ell}/\ell = n\theta_{\ell}/\ell$ , the number of size- $\ell$  components of in the union of the induced subgraphs satisfies

$$(1 - \psi_1)^2 n\theta_{\ell}/\ell \leq \sum_{j \in [n/k]} S_{k, \ell}^{(j)} \leq (1 + \psi_1)^2 n\theta_{\ell}/\ell \quad \text{on } \mathcal{A}_{\text{sub}}(k). \quad (7.8)$$

We now look at edges across subboxes. Let  $\mathcal{V}_n^{\text{cross}}(k)$  denote the set of vertices having an edge to a vertex in a different subbox. Each vertex in  $\mathcal{V}_n^{\text{cross}}(k)$  changes the number of size- $\ell$  components by at most 1: indeed, this number increases by one when its crossing edge(s) form a new component of size- $\ell$ , and it decreases by one if without crossing edges its size was  $\ell$ . Therefore,

$$\left| S_{n, \ell}(\infty) - \sum_{j \in [n/k]} S_{k, \ell}^{(j)} \right| \leq |\mathcal{V}_n^{\text{cross}}(k)|. \quad (7.9)$$

Hence, for sufficiently small  $\psi_1$  depending on  $\psi$  and  $\ell$ , on the event on  $\{|\mathcal{V}_n^{\text{cross}}(k)| < \psi_1 n\} \cap \mathcal{A}_{\text{sub}}(k)$ ,

$$(1 - \psi)n\theta_{\ell}/\ell \leq (1 - \psi_1)^2 n\theta_{\ell}/\ell - \psi_1 n \leq S_{n, \ell}(\infty)$$

$$\leq (1 + \psi_1)^2 n \theta_\ell / \ell + \psi_1 n \leq (1 + \psi) n \theta_\ell / \ell.$$

Combining this with (7.7), the exponential decay in (7.6) holds if

$$\forall \psi_1 > 0 : \exists k, n_0 \in \mathbb{N}, \varepsilon > 0 \text{ s.t. } \forall n \geq n_0 : \mathbb{P}(|\mathcal{V}_n^{\text{cross}}(k)| \geq \psi_1 n) \leq \exp(-\varepsilon n). \quad (7.10)$$

For Bernoulli bond percolation on  $\mathbb{Z}^d$ , the number of vertices with a crossing edge in each subbox is deterministically at most  $O(k^{(d-1)/d})$  as  $k \rightarrow \infty$ , so the total number of vertices with a crossing edge across the  $(n/k)$  many subboxes is  $n O(k^{-1/d})$ . So, given  $\psi_1$ , we choose  $k$  large enough so that  $O(k^{-1/d}) < \psi_1$ , and then (7.10) holds (with error probability 0).

Consider long-range percolation on  $\mathbb{Z}^d$  with  $\alpha > 1$ . Let  $\mathcal{V}_n^{\text{surf}}(t)$  be the set of vertices within distance  $t$  from the union of the boundaries of all subboxes  $j \leq n/k$ ; and let  $\mathcal{E}_n^{\text{long}}(t)$  denote the edges of length at least  $t$  inside  $\Lambda_n$ . Clearly,  $|\mathcal{V}_n^{\text{surf}}(t)| \leq (n/k) 2^d k^{(d-1)/d} (2t + 1)$ . For any  $\psi_1 > 0$ , fix  $\psi_2 > 0$  such that  $|\mathcal{V}_n^{\text{surf}}(\psi_2 k^{1/d})| \leq \psi_1 n/2$  holds deterministically for all  $k$  sufficiently large (simultaneously). Then, since each edge has two endpoints,

$$|\mathcal{V}_n^{\text{cross}}(k)| \leq |\mathcal{V}_n^{\text{surf}}(\psi_2 k^{1/d})| + 2|\mathcal{E}_n^{\text{long}}(\psi_2 k^{1/d})| \leq \psi_1 n/2 + 2|\mathcal{E}_n^{\text{long}}(\psi_2 k^{1/d})|.$$

Therefore,

$$\mathbb{P}(|\mathcal{V}_n^{\text{cross}}(k)| \geq \psi_1 n) \leq \mathbb{P}\left(2|\mathcal{E}_n^{\text{long}}(\psi_2 k^{1/d})| \geq \psi_1 n/2\right). \quad (7.11)$$

By a standard isoperimetric inequality, the total number of possible edges inside  $\Lambda_n$  with length in the interval  $[r_{i,k}, r_{i+1,k}) := [2^i \psi_2 k^{1/d}, 2^{i+1} \psi_2 k^{1/d})$  is at most  $n \cdot O(r_{i,k}^d)$  as  $i \rightarrow \infty$ . Each such edge is present independently with probability at most  $\beta^\alpha r_{i,k}^{-\alpha d}$  by the connection probability (4.6). Hence, there exists a constant  $C > 0$  such that for  $k$  sufficiently large

$$|\mathcal{E}_n^{\text{long}}(\psi_2 k^{1/d})| \leq \sum_{i=0}^{\infty} \text{Bin}(C \psi_2^d n 2^{id} k, \beta^\alpha \psi_2^{-\alpha d} 2^{-i\alpha d} k^{-\alpha}),$$

where all binomial random variables in the sum are independent of each other. Thus,

$$\mathbb{P}(|\mathcal{V}_n^{\text{cross}}(k)| \geq \psi_1 n) \leq \mathbb{P}\left(2 \sum_{i=0}^{\infty} \text{Bin}(C \psi_2^d n 2^{id} k, \beta^\alpha \psi_2^{-\alpha d} (2^{id} k)^{-\alpha}) \geq \psi_1 n/2\right).$$

First we apply  $x \mapsto \exp(x/2)$  to both sides of the inequality inside the probability on the right-hand side, then we use Markov's inequality and independence of the random variables:

$$\mathbb{P}(|\mathcal{V}_n^{\text{cross}}(k)| \geq \psi_1 n) \leq e^{-\psi_1 n/4} \prod_{i=0}^{\infty} \mathbb{E}\left[\exp\left(\text{Bin}(C \psi_2^d n 2^{id} k, \beta^\alpha \psi_2^{-\alpha d} (2^{id} k)^{-\alpha})\right)\right].$$



Each factor corresponds to the moment-generating function (MGF) of a binomial random variable evaluated at 1. The MGF of  $\text{Bin}(n, p)$  at 1 equals  $(1 + p(e - 1))^n \leq \exp((e - 1)np)$ . Applying this bound to each of the factors yields

$$\mathbb{P}(|\mathcal{V}_n^{\text{cross}}(k)| \geq \psi_1 n) \leq \exp\left(-\psi_1 n/4 + (e - 1)C\beta^\alpha \psi_2^{d(1-\alpha)} n k^{1-\alpha} \sum_{i=0}^{\infty} 2^{id(1-\alpha)}\right).$$

Since  $\alpha > 1$  by assumption, the sum converges. When  $k$  is sufficiently large, the right-hand side is at most  $\exp(-\psi_1 n/5)$ . This proves (7.10) for long-range percolation and hence (7.6).  $\square$

We are ready to prove Theorem 2.4.

**Proof of Theorem 2.4** For the lower bound, we observe that  $|\mathcal{C}_n^{(1)}| = n$  when all nearest neighbor edges are present. Its probability decays as  $\exp(-\Theta(n))$ . For the upper bound, we argue as follows. We fix  $\rho \in (\theta, 1)$ , and a small  $\psi < 1 - (1 - \rho)/(1 - \theta)$ . Since  $\sum_{\ell=1}^{\infty} \theta_\ell = 1 - \theta$ , this gives us a constant  $\ell^* = \ell^*(\rho, \psi)$  that satisfies  $(1 - \psi) \sum_{\ell=1}^{\ell^*} \theta_\ell > 1 - \rho$ . We shall shortly show that

$$\{|\mathcal{C}_n^{(1)}| > \rho n\} \subseteq \{\exists \ell \leq \ell^* + 1 : S_{n,\ell}(\infty) \leq (1 - \psi)n\theta_\ell/\ell\} =: \mathcal{A}_{\text{comp}}. \quad (7.12)$$

If (7.12) holds, then Lemma 7.2 and a union bound over the finite set  $\ell \leq \ell^* + 1$  gives that  $\mathcal{A}_{\text{comp}}$  occurs with probability exponentially small in  $n$ , finishing the proof of the upper bound. To prove (7.12), we show that  $\neg \mathcal{A}_{\text{comp}} \subseteq \{|\mathcal{C}_n^{(1)}| \leq \rho n\}$ . Since we excluded the case in LRP that nearest neighbor edges are present with probability one,  $\theta_\ell = \mathbb{P}(|\mathcal{C}(0)| = \ell) > 0$  for all  $\ell \geq 1$ . So, for all  $n$  sufficiently large, on  $\neg \mathcal{A}_{\text{comp}}$ , there is at least one component of size  $\ell^* + 1$ ; i.e.,  $|\mathcal{C}_n^{(1)}| \geq \ell^* + 1$ . Hence, vertices in components of size  $\ell \leq \ell^*$  are not in the giant. Thus,

$$n - |\mathcal{C}_n^{(1)}| = |\mathcal{V}_n \setminus \mathcal{V}(\mathcal{C}_n^{(1)})| \geq \sum_{\ell=1}^{\ell^*} \ell S_{n,\ell}(\infty) \geq \sum_{\ell=1}^{\ell^*} (1 - \psi)n\theta_\ell > n(1 - \rho).$$

Rearranging yields that  $|\mathcal{C}_n^{(1)}| < \rho n$  on  $\neg \mathcal{A}_{\text{comp}}$ , finishing the proof.  $\square$

### 7.3 Inhomogeneous models

Here we start again with the convergence rate of the number of size- $\ell$  components. Since the degree distribution is heavy-tailed, this is more involved than Lemma 7.2. We write  $\deg_v$  for the degree of vertex  $v$  in  $\mathcal{G}_n$ , and  $\deg_v[a, b]$  for the number of neighbors of  $v$  in  $\mathcal{G}_n$  with mark in  $[a, b]$ .

**Claim 7.3** (Total degree from high-mark vertices). *Consider a KSRG on a PPP satisfying Assumption 2.2 with  $\tau < \infty$ . For any constants  $\psi, C > 0$ , there exists  $\phi_0 > 0$*

such that for all  $n$  sufficiently large and  $\phi \in (0, \phi_0)$ , writing  $\underline{w}_n = n^{(1+\psi)/(\tau-1)}$ ,

$$\mathbb{P}\left(\sum_{v \in \mathcal{V}_n[\underline{w}_n, \phi n]} \deg_v[1, \phi n] \geq \psi n\right) \leq n^{-C}. \quad (7.13)$$

**Claim 7.4** (High-degree vertices). *Consider a KSRG on a PPP satisfying Assumption 2.2 with  $\tau < \infty$ . For any constants  $\ell, C > 0$ , there exist constants  $n_0, C_1$  such that for all  $n \geq n_0$ , and any  $\underline{w}_n > C_1 \log n$ ,*

$$\mathbb{P}(\exists v \in \mathcal{V}_n[\underline{w}_n, \infty) : \deg_v[1, \underline{w}_n] < \ell) \leq n^{-C}.$$

**Claim 7.5** (Crossing edges). *Consider a KSRG on a PPP satisfying Assumption 2.2 with  $\tau < \infty$ . Let  $\psi > 0$  be a constant such that  $(1 + \psi)/(\tau - 1) < 1$ . Then, for any constant  $C > 0$ , there exists  $n_0$  such that for all  $n \geq n_0$  and any  $\underline{w}_n \leq n^{(1+\psi)/(\tau-1)}$ ,*

$$\mathbb{P}(|\{v \in \mathcal{V}_n[1, \underline{w}_n) : v \sim \Lambda_n^c \times [1, \underline{w}_n)\}| \geq \psi n) \leq n^{-C}. \quad (7.14)$$

The proofs of the three claims are based on concentration inequalities for Poisson/binomial random variables and postponed to Appendix C.2. Their combination gives a *polynomial* convergence rate for the size- $\ell$  components in Lemma 7.1, which is in stark contrast to the exponential convergence rate for degree-homogeneous models in Lemma 7.2.

**Lemma 7.6** (Concentration for size- $\ell$  components). *Consider a KSRG on a PPP satisfying Assumption 2.2 with  $\tau < \infty$ . For all constants  $\ell, \psi, C > 0$ , there exist constants  $\phi_0 > 0$  and  $n_0$  such that for all  $n \geq n_0$  and all  $\phi \in (0, \phi_0)$ ,*

$$\mathbb{P}(|S_{n,\ell}(\phi n)/n - \theta_\ell/\ell| > \psi) \leq n^{-C}. \quad (7.15)$$

**Proof** We first reduce the upper mark threshold  $\bar{w}_n := \phi n$  in the event in (7.15) to a smaller value  $\underline{w}_n = n^{(1+\psi)/(\tau-1)}$  (we assume w.l.o.g. that  $(1 + \psi)/(\tau - 1) < 1$ ). We use that each vertex with degree at least  $\ell$  is in a component of size at least  $\ell + 1$ , and distinguish the case that each vertex of mark  $\underline{w}_n$  is in a component of size larger than  $\ell$ , i.e.,

$$\mathbb{P}(|S_{n,\ell}(\bar{w}_n)/n - \theta_\ell/\ell| > \psi) \leq \mathbb{P}\left(\begin{aligned} & \{|S_{n,\ell}(\bar{w}_n)/n - \theta_\ell/\ell| > \psi\} \\ & \cap \{\forall v \in \mathcal{V}_n[\underline{w}_n, \bar{w}_n) : |\mathcal{C}_n(v)[1, \bar{w}_n]| > \ell\} \end{aligned}\right) \quad (7.16)$$

$$+ \mathbb{P}(\exists v \in \mathcal{V}_n[\underline{w}_n, \bar{w}_n) : \deg_v[1, \bar{w}_n] \leq \ell). \quad (7.17)$$

Now we ‘switch’  $\bar{w}_n$  to  $\underline{w}_n$  in  $S_{n,\ell}(\cdot)$  on the right-hand side. Let  $T_{n,\ell}$  denote the number of size- $\ell$  components in  $\mathcal{G}_n[1, \underline{w}_n)$  that have an edge towards a vertex with mark in  $[\underline{w}_n, \bar{w}_n)$ . The events on the right-hand side of (7.16) imply that the size- $\ell$  components

of  $\mathcal{G}_n[1, \bar{w}_n)$  do not contain any vertices with mark at least  $\bar{w}_n$ . So, on those events,  $S_{n,\ell}(\bar{w}_n) = S_{n,\ell}(\underline{w}_n) - T_{n,\ell}$ . Distinguishing also whether  $T_{n,\ell}/n \leq \psi/2$  or not yields

$$\begin{aligned} \mathbb{P}(|S_{n,\ell}(\bar{w}_n)/n - \theta_\ell/\ell| > \psi) &\leq \mathbb{P}\left(\begin{aligned} &\{|S_{n,\ell}(\underline{w}_n)/n - T_{n,\ell}/n - \theta_\ell/\ell| > \psi\} \\ &\cap \{T_{n,\ell}/n \leq \psi/2\} \\ &\cap \{\forall v \in \mathcal{V}_n[\underline{w}_n, \bar{w}_n) : |C_n(v)[1, \bar{w}_n]| > \ell\} \end{aligned}\right) \\ &\quad + \mathbb{P}(T_{n,\ell}/n \geq \psi/2) + \mathbb{P}(\exists v \in \mathcal{V}_n[\underline{w}_n, \bar{w}_n) : \deg_v[1, \bar{w}_n] \leq \ell). \end{aligned} \quad (7.18)$$

The first two events in (7.18) imply that  $\{|S_{n,\ell}(\underline{w}_n)/n - \theta_\ell/\ell| > \psi/2\}$ . Moreover,  $T_{n,\ell}$  is at most the total number of edges between  $\mathcal{V}[\underline{w}_n, \bar{w}_n)$  and  $\mathcal{V}_n[1, \underline{w}_n)$ , which is at most the total degree of vertices in  $\mathcal{V}[\underline{w}_n, \bar{w}_n)$ . Hence,

$$\begin{aligned} \mathbb{P}(|S_{n,\ell}(\bar{w}_n)/n - \theta_\ell/\ell| > \psi) &\leq \mathbb{P}(|S_{n,\ell}(\underline{w}_n)/n - \theta_\ell/\ell| > \psi/2) \\ &\quad + \mathbb{P}\left(\sum_{v \in \mathcal{V}_n[\underline{w}_n, \bar{w}_n)} \deg_v[1, \bar{w}_n] \geq n\psi/2\right) \\ &\quad + \mathbb{P}(\exists v \in \mathcal{V}_n[\underline{w}_n, \bar{w}_n) : \deg_v[1, \bar{w}_n] \leq \ell). \end{aligned}$$

By Claim 7.3, the second term can be made at most  $n^{-2C}$  if  $\phi$  in the definition of  $\bar{w}_n = \phi n$  is sufficiently small. Since  $\underline{w}_n = n^{(1+\psi)/(\tau-1)}$ , the third term is also at most  $n^{-2C}$  by Claim 7.4. So,

$$\mathbb{P}(|S_{n,\ell}(\bar{w}_n)/n - \theta_\ell/\ell| > \psi) \leq \mathbb{P}(|S_{n,\ell}(\underline{w}_n)/n - \theta_\ell/\ell| > \psi/2) + o(n^{-C}). \quad (7.19)$$

It remains to bound the first term on the right-hand side with  $\underline{w}_n = n^{(1+\psi)/(\tau-1)}$ . Let  $A$  be a sufficiently large constant with  $(A \log n)^{1/d} \in \mathbb{N}$  and partition  $\Lambda_n$  into boxes  $\mathcal{Q}_1, \dots, \mathcal{Q}_{A \log n}$  of volume  $k_n := n/(A \log n)$ . Let  $S_{k_n, \ell}^{(j)}(\underline{w}_n)$  denote the number of components of size exactly  $\ell$  in the  $j$ th box, and  $\mathcal{V}_{k_n}^{(j)}$  the set of vertices in the  $j$ th box.

$$\mathcal{A}^{(j)} := \{|S_{k_n, \ell}^{(j)}(\underline{w}_n)/k_n - \theta_\ell/\ell| \leq \psi/12\}, \quad (7.20)$$

and otherwise we call it bad. The probability of being good tends to 1 with  $k_n$  by Lemma 7.1, and so it is at least  $1 - \psi/32$  for all  $n$  sufficiently large. Let  $X := \sum_{j \in [k_n]} \mathbb{1}_{\{\mathcal{A}^{(j)}\}}$  be the number of good boxes. Independence across boxes gives that  $X$  dominates a binomial random variable with parameters  $A \log n$  and  $1 - \psi/32$ . So, by a Chernoff bound there exists a constant  $c = c(\psi) > 0$  such that

$$\mathbb{P}(X \geq (1 - \psi/16)A \log n) \geq 1 - \exp(-cA \log n) = 1 - n^{-cA}. \quad (7.21)$$

Since  $k_n = n/A \log n$ , using concentration for Poisson random variables, for all  $n$  sufficiently large,

$$\mathbb{P}(\forall j : |\mathcal{V}_{k_n}^{(j)}[1, \underline{w}_n]|/k_n < 4/3) \geq 1 - (A \log n) \exp(-\Omega(n/\log n)) \geq 1 - n^{-cA}. \quad (7.22)$$

Thus, in the following, we will find bounds on  $S_{n,\ell}(\underline{w}_n)$  conditional on the event

$$\mathcal{A}_{\text{good}} := \{\forall i : |\mathcal{V}_{k_n}^{(j)}[1, \underline{w}_n]|/k_n < 4/3\} \cap \{X \geq (1 - \psi/16)A \log n\}. \quad (7.23)$$

We write  $S_{n,\ell}^{\text{good}}$  for the sum of  $S_{k_n,\ell}^{(j)}$  over  $j$  for which the  $j$ th box is good, and similarly  $S_{k_n,\ell}^{\text{bad}}$  for the same sum over bad boxes. Finally, let  $\mathcal{V}_n^{\text{cross}}[1, \underline{w}_n]$  denote the set of vertices incident to an edge that crosses a box-boundary. Since each edge changes the number of size- $\ell$  components by at most 2, similarly to (7.9), it holds that

$$|S_{n,\ell}(\underline{w}_n) - (S_{n,\ell}^{\text{good}} + S_{n,\ell}^{\text{bad}})| \leq |\mathcal{V}_n^{\text{cross}}[1, \underline{w}_n]|. \quad (7.24)$$

On the event  $\mathcal{A}_{\text{good}}$  in (7.23), the at most  $A \log n$  good boxes contain together at most  $(\theta_\ell/\ell + \psi/12)n$  components of size  $\ell$ . There are at most  $n(\psi/16)4/3 = n\psi/12$  vertices in bad boxes. Assuming the worst case that all of them are in size- $\ell$  components inside their boxes, we obtain the upper bound  $S_{n,\ell}^{\text{bad}} \leq n\psi/(12\ell)$ . So, by a triangle inequality on the left-hand side above,

$$S_{n,\ell}(\underline{w}_n) - n\theta_\ell/\ell \leq n\psi/12 + n\psi/(12\ell) + |\mathcal{V}_n^{\text{cross}}[1, \underline{w}_n]| \leq n\psi/6 + |\mathcal{V}_n^{\text{cross}}[1, \underline{w}_n]|.$$

Similarly, for a lower bound we may assume that there are no size- $\ell$  components in bad boxes. On the event (7.23), this still gives

$$\begin{aligned} S_{n,\ell}(\underline{w}_n) - n\theta_\ell/\ell &\geq n(1 - \psi/16)(\theta_\ell/\ell - \psi/12) - |\mathcal{V}_n^{\text{cross}}[1, \underline{w}_n]| \\ &\geq -n\psi/48 - |\mathcal{V}_n^{\text{cross}}[1, \underline{w}_n]| \end{aligned}$$

Recalling the bounds (7.19), (7.21), and (7.22), we obtain

$$\mathbb{P}(|S_{n,\ell}(\underline{w}_n)/n - \theta_\ell/\ell| > \psi) \leq 2n^{-cA} + o(n^{-C}) + \mathbb{P}(|\mathcal{V}_n^{\text{cross}}|/n \geq \psi/3).$$

We bound the probability concerning the crossing edges. By the pigeon-hole principle,  $|\mathcal{V}_n^{\text{cross}}| \geq n\psi/3$  implies that there exists a box that contains at least  $(\psi/3)k_n$  vertices with an outgoing edge towards the complement of the box. Hence, by translation invariance of the model on  $\mathbb{R}^d$ ,

$$\mathbb{P}(|\mathcal{V}_n^{\text{cross}}|/n \geq \psi/3) \leq (A \log n) \mathbb{P}(|\{v \in \mathcal{V}_{k_n}[1, \underline{w}_n] : v \sim \Lambda_{k_n}^c \times [1, \underline{w}_n]\}| \geq (\psi/3)k_n).$$

Claim 7.5 is applicable here for  $n_{7.5} = k_n = n/(A \log n)$ , since  $\underline{w}_n = n^{(1+\psi)/(\tau-1)} \leq k_n^{(1+\psi')/(\tau-1)}$  for any  $\psi' > \psi$ . So the right-hand side is smaller than  $n^{-cA}$  for  $n$  sufficiently large. We conclude that (7.15) follows for  $\phi$  small, setting  $A \geq C/c$  such that  $(A \log n)^{1/d} \in \mathbb{N}$ .  $\square$

We need a last preliminary claim. Recall  $\text{hubs}(\rho)$  from (2.17). We introduce notation for the functions appearing in the upper and lower bound of Theorem 2.5:

$$\begin{aligned} h_{\text{up}} &:= \lceil \text{hubs}(\rho) \rceil, \\ h_{\text{lo}} &:= \lim_{r \downarrow \rho} \lceil \text{hubs}(r) \rceil = \begin{cases} \lceil \text{hubs}(\rho) \rceil, & \text{if } \text{hubs}(\rho) \notin \mathbb{N} \text{ or } p = 1, \\ \text{hubs}(\rho) + 1, & \text{if } \text{hubs}(\rho) \in \mathbb{N} \text{ and } p < 1. \end{cases} \end{aligned} \quad (7.25)$$

Clearly  $h_{\text{lo}} \geq h_{\text{up}} \geq \text{hubs}(\rho)$ , and for all  $\rho \in (\theta, 1)$  and  $p < 1$  also  $h_{\text{lo}} > \text{hubs}(\rho)$  strictly.

**Claim 7.7** (Truncation of the generating function). *Consider a KSRG on a PPP satisfying Assumption 2.2 with  $\tau < \infty$ . Let  $\rho \in (\theta, 1)$ . There exist constants  $\psi > 0$  and  $\ell_0 \in \mathbb{N}$  such that for all  $\ell^* > \ell_0$ ,*

$$(1 - \psi)^2 \left( \theta + \sum_{\ell=1}^{\ell^*} (1 - (1 - p)^{h_{\text{lo}} \ell}) \theta_{\ell} \right) > \rho, \quad (7.26)$$

and

$$(1 - \psi)^2 \sum_{\ell=1}^{\ell^*} \theta_{\ell} (1 - p)^{(h_{\text{up}} - 1) \ell} > 1 - \rho + \psi. \quad (7.27)$$

We postpone the proof to Appendix C.3 and proceed to the lower bound of Theorem 2.5.

**Proposition 7.8** (Lower bound of the upper tail). *Consider a KSRG on a PPP satisfying Assumption 2.2 with  $\tau < \infty$ . Fix  $\rho \in (\theta, 1)$ . Assume either supercriticality and  $\zeta_{\text{long}} > 0$ , or assume  $\theta = 0$ . Then there exists  $A = A(\rho) > 0$  such that for all  $n \geq 1$ ,*

$$\mathbb{P}(|\mathcal{C}_n^{(1)}| > \rho n) \geq (1/A) n^{-(\tau-2)h_{\text{lo}}}.$$

**Proof** The proof is similar in spirit to the proof of Theorem 2.4 on page 63, but we also need to deal with high-mark vertices. See Section 3.2 for an outline that we make formal now. Let  $R, \ell^*$  be two sufficiently large constants, and let  $\phi, \psi > 0$  be two small constants so that given  $\rho \in (\theta, 1)$ , (7.26) holds. Write  $T_{n,\ell}$  for the number of components of size  $\ell$  in  $\mathcal{G}_n[1, \phi n]$  that connect by an edge to the vertices with mark in  $\mathcal{V}_n[Rn, \infty)$ . Define the events

$$\mathcal{A}_{\text{comp}} := \{|\mathcal{C}_n^{(1)}[1, \phi n]|/n \geq \theta(1 - \psi)\} \cap \{\forall \ell \leq \ell^* : S_{n,\ell}(\phi n) \geq (1 - \psi)n\theta_{\ell}/\ell\}, \quad (7.28)$$

$$\mathcal{A}_{\text{hub}} := \{|\mathcal{V}_n[Rn, \infty)| = h_{\text{lo}}\} \cap \{\forall v \in \mathcal{V}_n[Rn, \infty) : v \sim \mathcal{C}_n^{(1)}[1, \phi n]\}, \quad (7.29)$$

$$\mathcal{A}_{\text{conn}} := \{\forall \ell \leq \ell^* : T_{n,\ell} \geq (1 - \psi)(1 - (1 - p)^{h_{\text{lo}} \ell})((1 - \psi)n\theta_{\ell}/\ell)\}. \quad (7.30)$$

We shall show that

$$\mathcal{A}_{\text{comp}} \cap \mathcal{A}_{\text{hub}} \cap \mathcal{A}_{\text{conn}} \subseteq \{|\mathcal{C}_n^{(1)}| > \rho n\}. \quad (7.31)$$

By its definition,  $T_{n,\ell}$  many clusters of size  $\ell$  connect to the hubs listed in  $\mathcal{A}_{\text{hub}}$ , which then all are part of the giant component in the whole graph. Hence, on  $\mathcal{A}_{\text{hub}}$ , the giant  $|\mathcal{C}_n^{(1)}[1, \phi n]|$  increases by at least  $\ell T_{n,\ell}$  when fully revealing the graph. By the bounds on  $T_{n,\ell}$  and the initial size  $(1 - \psi)\theta n$  of the giant inside the events, and (7.26) in Claim 7.7, we obtain that

$$\begin{aligned} |\mathcal{C}_n^{(1)}|/n &\geq \theta(1 - \psi) + \frac{1}{n} \sum_{\ell=1}^{\ell^*} \ell T_{n,\ell} \geq \theta(1 - \psi) + (1 - \psi)^2 \sum_{\ell=1}^{\ell^*} (1 - (1 - p)^{h_{10}\ell}) \theta_\ell \\ &\geq (1 - \psi)^2 \left( \theta + \sum_{\ell=1}^{\ell^*} (1 - (1 - p)^{h_{10}\ell}) \theta_\ell \right) > \rho, \end{aligned}$$

so (7.31) holds. We bound from below the probability of the intersection using the law of total probability:

$$\mathbb{P}(\mathcal{A}_{\text{comp}} \cap \mathcal{A}_{\text{hubs}} \cap \mathcal{A}_{\text{conn}}) = \mathbb{P}(\mathcal{A}_{\text{comp}} \cap \mathcal{A}_{\text{hubs}}) \mathbb{P}(\mathcal{A}_{\text{conn}} \mid \mathcal{A}_{\text{comp}} \cap \mathcal{A}_{\text{hubs}}). \quad (7.32)$$

To study the first factor, we use the law of total probability again and then also  $\mathbb{P}(A \cap B) \geq \mathbb{P}(A) - \mathbb{P}(\neg B)$  (see (7.28)-(7.29)):

$$\begin{aligned} \mathbb{P}(\mathcal{A}_{\text{comp}} \cap \mathcal{A}_{\text{hubs}}) &\geq \mathbb{P}\left(\forall v \in \mathcal{V}_n[Rn, \infty) : v \sim \mathcal{C}_n^{(1)}[1, \phi n] \mid \mathcal{A}_{\text{comp}} \cap \{|\mathcal{V}_n[Rn, \infty)| = h_{10}\}\right) \\ &\quad \cdot \left(\mathbb{P}(|\mathcal{V}_n[Rn, \infty)| = h_{10}) - \mathbb{P}(\neg \mathcal{A}_{\text{comp}})\right). \end{aligned} \quad (7.33)$$

For the first term in the row (7.33) we use that  $|\mathcal{V}_n[Rn, \infty)| \sim \text{Poi}(n(Rn)^{-(\tau-1)})$  by the intensity in (4.5). Note here that the parameter tends to 0 since  $\tau > 2$ . Therefore, since  $h_{10}$  is a  $\rho$ -dependent constant that does not depend on  $n$ ,

$$\begin{aligned} \mathbb{P}(|\mathcal{V}_n[Rn, \infty)| = h_{10}) &= \exp\left(-R^{-(\tau-1)}n^{-(\tau-2)}\right) \frac{\left(R^{-(\tau-1)}n^{-(\tau-2)}\right)^{h_{10}}}{h_{10}!} \\ &\geq c_{R,\rho} n^{-(\tau-2)h_{10}} \end{aligned} \quad (7.34)$$

for some constant  $c_{R,\rho} > 0$ . We now bound  $\mathbb{P}(\neg \mathcal{A}_{\text{comp}})$  in (7.33).

$$\begin{aligned} \mathbb{P}(\neg \mathcal{A}_{\text{comp}}) &\leq \mathbb{P}\left(\exists \ell \leq \ell^* : \ell S_{n,\ell}(\phi n)/n \leq (1 - \psi)\theta_\ell\right) \\ &\quad + \mathbb{P}(|\mathcal{C}_n^{(1)}[1, \phi n]|/n < \theta(1 - \psi)). \end{aligned} \quad (7.35)$$

Lemma 7.6 and a union bound over the  $\ell^*$  many  $\ell$  shows that the first term is at most  $n^{-C}$  for arbitrary large  $C$  when  $\phi < \phi_0(\psi)$ . For the second term, there are two cases. Either  $\theta = 0$  and then this probability is 0; or,  $\theta$  is non-zero. In the latter case, when additionally  $\zeta_{\text{long}} > 0$ , our result for the *lower* tail of large deviations in Theorem 2.7 guarantees that this event has stretched exponentially decaying probability in  $n$ . This

is where we use that the model must be supercritical with  $\zeta_{\text{long}} > 0$  when<sup>6</sup>  $\theta > 0$ . So, under these conditions, we know that  $\mathbb{P}(\neg \mathcal{A}_{\text{comp}})$  is of smaller order than the lower bound in (7.34) on the first term in (7.33).

We now analyze the conditional probability on the line above (7.33). Take  $R > \sqrt{d}/\beta$ . Then a vertex with mark at least  $Rn$  connects by an edge to any other vertex in  $\Lambda_n$  with probability  $p$  by (4.6). The giant  $\mathcal{C}_n^{(1)}[1, \phi n]$  has size at least  $(1 - \psi)\theta n$  on the event  $\mathcal{A}_{\text{comp}}$  in (7.28) in the conditioning. So, by a union bound, the (conditional) probability that all the  $h_{10}$  many vertices of mark at least  $Rn$  connect by an edge to  $\mathcal{C}_n^{(1)}[1, \phi n]$  is at least  $1 - h_{10}(1 - p)^{\Theta(n)}$ , which tends to 1 as  $n \rightarrow \infty$ . Combining these with (7.33)–(7.34) gives in (7.32)

$$\mathbb{P}(\mathcal{A}_{\text{comp}} \cap \mathcal{A}_{\text{hubs}}) \geq c_{R,\rho} n^{-(\tau-2)h_{10}}. \quad (7.36)$$

We show that the other term in (7.32),  $\mathbb{P}(\mathcal{A}_{\text{conn}} \mid \mathcal{A}_{\text{comp}} \cap \mathcal{A}_{\text{hubs}})$  tends to 1 as  $n \rightarrow \infty$ . Recall these events from (7.28)–(7.30) and that under the conditioning,  $T_{n,\ell}$  denotes the number of size- $\ell$  components of  $\mathcal{G}_n[1, \phi n]$  that connect to one of the  $h_{10}$  many vertices of mark at least  $Rn$  (the hubs). As before, each vertex connects independently to these hubs with probability  $p$ . Further, the vertex set counted in  $T_{n,\ell}$  is disjoint of the vertices in  $\mathcal{C}_n^{(1)}[1, \phi n]$  so their connection to the hubs is independent of whether the vertices in  $\mathcal{C}_n^{(1)}[1, \phi n]$  connect to the hubs, which is an event inside the conditioning  $\mathcal{A}_{\text{comp}} \cap \mathcal{A}_{\text{hubs}}$ . So, under the conditioning, each component of size  $\ell$  in  $\mathcal{G}_n[1, \phi n]$  connects to one of the  $h_{10}$  hubs with probability  $1 - (1 - p)^{h_{10}\ell}$  independently across size- $\ell$  components and across  $\ell < \ell^*$ . Hence, given the number of size- $\ell$  components in  $\mathcal{G}_n[1, \psi n]$ ,  $(S_{n,\ell}(\phi n))_{\ell \leq \ell^*}$ , the variables  $T_{n,\ell}$  are binomially distributed with parameters  $S_{n,\ell}(\phi n)$  and  $(1 - (1 - p)^{h_{10}\ell})$ . Further, under the conditioning in  $\mathcal{A}_{\text{comp}}$ , each  $S_{n,\ell}(\phi n)$  is at least  $(1 - \psi)n\theta_\ell/\ell$ , so  $T_{n,\ell}$  dominates a binomial random variable with parameters  $(1 - \psi)n\theta_\ell/\ell$  and  $(1 - (1 - p)^{h_{10}\ell})$  for each  $\ell \leq \ell^*$ . A union bound over  $\ell \leq \ell^*$  and then a Chernoff bound for each  $\ell$  give that  $\mathbb{P}(\mathcal{A}_{\text{conn}} \mid \mathcal{A}_{\text{comp}} \cap \mathcal{A}_{\text{hubs}}) \rightarrow 1$  as  $n \rightarrow \infty$ . Using this in (7.32) and (7.36) gives in (7.31) that for some constant  $c_{R,\rho} > 0$ ,

$$\mathbb{P}(\mathcal{C}_n^{(1)} > \rho n) \geq \mathbb{P}(\mathcal{A}_{\text{comp}} \cap \mathcal{A}_{\text{hubs}} \cap \mathcal{A}_{\text{conn}}) \geq c_{R,\rho} n^{-(\tau-2)h_{10}}.$$

□

We proceed to the upper bound of Theorem 2.5.

**Proposition 7.9** (Upper bound of the upper tail). *Consider a KSRG satisfying Assumption 2.2 on a PPP with  $\tau < \infty$ . Fix  $\rho \in (\theta, 1)$ . There exists  $A = A(\rho) > 0$  such that for all  $n \geq 1$ ,*

$$\mathbb{P}(|\mathcal{C}_n^{(1)}| > \rho n) \leq A n^{-(\tau-2)h_{\text{up}}}.$$

**Proof** Given  $\rho$ , fix  $\phi, \psi > 0$  sufficiently small and  $\ell^*$  sufficiently large so that (7.27) holds for  $\psi$  and  $\ell^*$ . We write  $\bar{T}_{n,\ell}$  for the number of components of size  $\ell$  in  $\mathcal{G}_n[1, \phi n]$

<sup>6</sup> Inspection of the proof of Theorem 2.7 shows that the sprinkling technique used in Section 4 needs supercriticality. The assumption  $\zeta_{\text{long}} > 0$  is the technical condition of Theorem 2.7 that guarantees enough long edges.

that do *not* connect by an edge to the vertices with mark at least  $\phi n$ , and define

$$\begin{aligned}\mathcal{A}_{\text{comp}} &:= \{|\mathcal{V}_n|/n \leq 1 + \psi\} \cap \{\forall \ell \leq \ell^* + 1 : S_{n,\ell}(\phi n) \geq (1 - \psi)n\theta_\ell/\ell\}, \\ \mathcal{A}_{\text{hubs}} &:= \{|\mathcal{V}_n[\phi n, \infty)| \leq h_{\text{up}} - 1\}, \\ \mathcal{A}_{\text{conn}} &:= \{\forall \ell \in [\ell^*] : \bar{T}_{n,\ell} > (1 - \psi)(1 - p)^{(h_{\text{up}}-1)\ell}((1 - \psi)n\theta_\ell/\ell)\}.\end{aligned}\quad (7.37)$$

We show in the next paragraph that

$$\mathcal{A}_{\text{comp}} \cap \mathcal{A}_{\text{hubs}} \cap \mathcal{A}_{\text{conn}} \subseteq \{|\mathcal{C}_n^{(1)}| \leq \rho n\}.\quad (7.38)$$

To show this, we bound the size of the complement of the giant on the intersection of the three events. Since  $S_{n,\ell^*+1}(\phi n) \geq 1$  for all sufficiently large  $n$ , the largest component has size at least  $\ell^*$ . Hence, vertices in components of size  $\ell \leq \ell^*$  are not in the giant of  $\mathcal{G}_n[1, \psi n)$ . Combining the bounds in the events with Claim 7.7, we obtain for sufficiently small  $\psi$  that

$$\begin{aligned}1 - |\mathcal{C}_n^{(1)}|/n &\geq |\mathcal{V}_n \setminus \mathcal{C}_n^{(1)}|/n - \psi \geq \frac{1}{n} \sum_{\ell=1}^{\ell^*} \ell \bar{T}_{n,\ell} - \psi \\ &\geq (1 - \psi)^2 \sum_{\ell=1}^{\ell^*} \theta_\ell (1 - p)^{(h_{\text{up}}-1)\ell} - \psi > 1 - \rho.\end{aligned}$$

Rearranging shows (7.38), which implies that  $\{|\mathcal{C}_n^{(1)}| > \rho n\} \subseteq (\neg \mathcal{A}_{\text{comp}}) \cup (\neg \mathcal{A}_{\text{hubs}}) \cup (\neg \mathcal{A}_{\text{conn}})$ . Thus, it follows that

$$\begin{aligned}\mathbb{P}(|\mathcal{C}_n^{(1)}| > \rho n) &\leq \mathbb{P}(\neg \mathcal{A}_{\text{comp}}) + \mathbb{P}(\neg \mathcal{A}_{\text{hubs}}) + \mathbb{P}(\neg \mathcal{A}_{\text{conn}} \cap \mathcal{A}_{\text{comp}} \cap \mathcal{A}_{\text{hubs}}) \\ &\leq \mathbb{P}(\neg \mathcal{A}_{\text{comp}}) + \mathbb{P}(\neg \mathcal{A}_{\text{hubs}}) + \mathbb{P}(\neg \mathcal{A}_{\text{conn}} \mid \mathcal{A}_{\text{comp}} \cap \mathcal{A}_{\text{hubs}}).\end{aligned}\quad (7.39)$$

By Lemma 7.6 and the concentration of Poisson random variables in Lemma A.1,  $\mathbb{P}(\neg \mathcal{A}_{\text{comp}}) = o(n^{-C})$  for any  $C > 0$  if  $\phi = \phi(\psi)$  is sufficiently small. For  $\mathbb{P}(\neg \mathcal{A}_{\text{hubs}})$ , we use similarly to (7.34) that the number of hubs is Poisson. So, there exists a constant  $C_{\phi,\rho} > 0$  such that for all  $n \geq 1$ ,

$$\begin{aligned}\mathbb{P}(\neg \mathcal{A}_{\text{hubs}}) &= \mathbb{P}(\text{Poi}(n(\phi n)^{-(\tau-1)}) \geq h_{\text{up}}) \\ &= \exp(-\phi^{-(\tau-1)}n^{-(\tau-2)}) \sum_{j=h_{\text{up}}}^{\infty} \frac{\phi^{-j(\tau-1)}n^{-j(\tau-2)}}{j!} \\ &\leq C_{\phi,\rho}n^{-(\tau-2)h_{\text{up}}}.\end{aligned}$$

For the conditional probability in (7.39), with the corresponding events defined in (7.37), we argue similarly as in the proof of the lower bound. The probability that a single vertex in a size- $\ell$  component of  $\mathcal{G}_n[1, \phi n)$  connects by an edge to each of the at most  $h_{\text{up}} - 1$  hubs is at most  $p$ , independently of each other under the conditioning.



Hence, the probability that a component of size  $\ell$  does not connect by an edge to one of the hubs is at least  $(1-p)^{\ell(h_{\text{up}}-1)}$ , and connected components connect independently by an edge to the hubs. Thus, conditionally on  $S_{n,\ell}(\phi n)$  and  $\mathcal{A}_{\text{comp}} \cap \mathcal{A}_{\text{hubs}}$ ,  $\bar{T}_{n,\ell}$  dominates a binomial random variable with parameters  $S_{n,\ell}(\phi n)$  and  $(1-p)^{\ell(h_{\text{up}}-1)}$ . On the event  $\mathcal{A}_{\text{comp}}$ ,  $S_{n,\ell}(\phi n) \geq (1-\psi)n\theta_\ell/\ell$ , so  $\bar{T}_{n,\ell}$  dominates a binomial random variable with parameters  $(1-\psi)n\theta_\ell/\ell$  and  $(1-p)^{\ell(h_{\text{up}}-1)}$ . A union bound over  $\ell \leq \ell^* + 1$  and then a Chernoff bound shows that  $\mathbb{P}(\neg \mathcal{A}_{\text{conn}} \mid \mathcal{A}_{\text{comp}} \cap \mathcal{A}_{\text{hubs}}) = \exp(-\Theta(n))$ . Substituting these three bounds into (7.39) finishes the proof.  $\square$

**Proof of Theorem 2.5** Recall  $h_{\text{lo}}$  and  $h_{\text{up}}$  from (7.25). The lower bound is given by Proposition 7.8, the upper bound by Proposition 7.9.  $\square$

## Appendix A Preliminary lemmas

**Lemma A.1** (Concentration of Poisson random variables [62]). For  $x > 1$ ,

$$\mathbb{P}(\text{Poi}(\lambda) \geq x\lambda) \leq \exp(-\lambda(1 + x \log(x) - x)), \quad (\text{A1})$$

and for  $x < 1$ ,

$$\mathbb{P}(\text{Poi}(\lambda) \leq x\lambda) \leq \exp(-\lambda(1 - x - x \log(1/x))). \quad (\text{A2})$$

## Appendix B Existence of a giant: postponed proofs

### B.1 Polynomially-sized components

The proof of Lemma 4.2 uses a similar strategy as the proof of Proposition 4.1: we first prove the statement along a subsequence  $(n_i)_{i \geq 1}$  using a multi-scale renormalization method, and then extend the result to all  $n$ . For the multi-scale renormalization, we follow the steps (Sub), (Con), (Mark) outlined in Section 4.2.1. The initialization of the renormalization is different. Here, we shall use the uniqueness of the infinite cluster to construct large connected components in level-0 boxes. We use slightly overlapping boxes, so the events that two overlapping boxes contain a large connected component in their induced subgraph are no longer independent. Therefore, we cannot use Chernoff bounds as we did in the proof of Proposition 4.1. Markov's inequality can still be applied, but it leads to weaker bounds. We start by formally setting up the renormalization.

We recall the parameter  $\beta$  from the connectivity function in (4.6), and the critical value  $\beta_c$  from (2.3). From 4.5, we recall that we can construct the vertex set  $\mathcal{V}$  as the union of two independent PPPs  $\mathcal{V}^{\text{base}}$  and  $\mathcal{V}^{\text{spr}}$  with respective spatial intensities  $q$  and  $1-q$ , with  $\beta q^d > \beta_c$  gives two independent PPPs. The choice of  $q$  ensures that the graph  $\mathcal{G}^{\text{base}} = (\mathcal{V}^{\text{base}}, \mathcal{E}^{\text{base}})$  induced by vertices in  $\mathcal{V}^{\text{base}}$  is supercritical itself. Let  $\theta^{\text{base}}$  be the percolation probability of  $\mathcal{G}^{\text{base}}$ . Lastly, w.l.o.g. we assume that  $\alpha < \infty$ , as otherwise  $\zeta_{\text{hl}} = (\tau - 1)/\alpha - (\tau - 2) < 0$ .

**Definition B.1** (*Renormalization scheme with overlapping boxes*). Let  $\zeta = \zeta_{\text{hl}} > 0$ , and let  $\varepsilon' > 0$  be a sufficiently small constant. Let

$$\eta := \lceil 2(1/\varepsilon' - 1) \rceil. \quad (\text{B1})$$

Given  $\alpha \in (1, \infty)$ , let  $s = s(\alpha, \varepsilon')$  be the smallest integer satisfying  $s^{\varepsilon'+\alpha}(s+1)^{-\alpha} \geq 1$  (so that  $\tilde{s}^{(\varepsilon'+\alpha)}(\tilde{s}+1)^{-\alpha} \geq 1$  for all  $\tilde{s} \geq s$ ). Let  $n_0 \geq \underline{n}_0 \geq 1$  be two large constants. Then, we iteratively define for  $i \geq 1$ ,

$$m_i := (i+s)^{(2+\eta)d}, \quad \underline{n}_i := \underline{n}_{i-1}m_i = \underline{n}_0 \prod_{\ell=1}^i (\ell+s)^{(2+\eta)d}, \quad n_i := (\underline{n}_i^{1/d} + n_0^{1/d} - \underline{n}_0^{1/d})^d. \quad (\text{B2})$$

Let  $Q_i^{\text{sub}}(x)$  and  $Q_i(x)$  denote boxes of volume  $\underline{n}_i$  and  $n_i$  centered at  $x$ . We define a subgraph  $\widehat{\mathcal{G}}_{i,x} := (\widehat{\mathcal{V}}_{i,x}, \widehat{\mathcal{E}}_{i,x})$  of the graph  $\mathcal{G}$  on  $Q_i^{\text{sub}}(x)$ . We call  $\widehat{\mathcal{V}}_{i,x}$  the level- $i$  vertex set inside  $Q_i(x)$  and define it as

$$\widehat{\mathcal{V}}_{i,x} := \mathcal{V}_{Q_i(x)}^{\text{base}} \cup \mathcal{V}_{Q_i^{\text{sub}}(x)}^{\text{spr}}[1, 2\underline{n}_i^\gamma]. \quad (\text{B3})$$

Let  $\mathcal{E}_{i,x}^{\text{base}}$  be the edges of  $\mathcal{G}^{\text{base}}$  inside  $Q_i(x)$ . Denote then by  $\mathcal{E}_{i,x}^{\text{spr-base}}$  the set of edges of  $\mathcal{G}$  inside  $Q_i^{\text{sub}}(x)$  with one endpoint in  $\mathcal{V}^{\text{spr}}[1, 2\underline{n}_i^\gamma]$  and one in  $\mathcal{V}^{\text{base}}$ , formally  $\{\{u, v\} \in \mathcal{E}_{Q_i^{\text{sub}}(x)} : u \in \mathcal{V}_{Q_i^{\text{sub}}(x)}^{\text{spr}}[1, 2\underline{n}_i^\gamma], v \in \mathcal{V}_{Q_i^{\text{sub}}(x)}^{\text{base}}\}$ . Then we say that the level- $i$  edge set inside  $Q_i(x)$  is

$$\widehat{\mathcal{E}}_{i,x} := \mathcal{E}_{i,x}^{\text{base}} \cup \mathcal{E}_{i,x}^{\text{spr-base}}. \quad (\text{B4})$$

We define the graphs  $\widehat{\mathcal{G}}_{i,x} := (\widehat{\mathcal{V}}_{i,x}, \widehat{\mathcal{E}}_{i,x})$ . Let  $\rho_0 := \theta^{\text{base}}/2$ , and define for  $i \geq 1$

$$\rho_i := 1/(i+s)^{2d}. \quad (\text{B5})$$

We call a box  $Q_i(x)$   $(i, \varepsilon')$ -good if the graph  $\widehat{\mathcal{G}}_{i,x}$  satisfies the event

$$\begin{aligned} \mathcal{A}_{i\text{-good}, \varepsilon'}(x) := \Big\{ \exists \text{comp. } \mathcal{C}_{i,x} \text{ in } \widehat{\mathcal{G}}_{i,x} : |\mathcal{C}_{i,x} \cap Q_i^{\text{sub}}(x)| \geq \underline{n}_i \prod_{\ell=0}^i \rho_\ell, \\ |\mathcal{V}_{Q_i^{\text{sub}}(x)}^{\text{spr}}[\underline{n}_i^\gamma, 2\underline{n}_i^\gamma] \cap \mathcal{C}_{i,x}| \geq \underline{n}_i^{\zeta-2\varepsilon'} \Big\}. \end{aligned} \quad (\text{B6})$$

The picture about the boxes  $Q_i^{\text{sub}}(x)$ ,  $Q_i(x)$  one should have in mind is as follows:  $Q_i(x)$ , the box of volume  $n_i$ , is the “fattening” of  $Q_i^{\text{sub}}(x)$  by at most an  $n_0$ -dependent constant. By construction,  $Q_i^{\text{sub}}(x)$  can be partitioned into exactly  $\underline{n}_i/\underline{n}_0$  many non-overlapping small boxes  $(Q_0^{\text{sub}}(x_j))_{j \geq 1}$  of volume  $\underline{n}_0$ , centered at some  $(x_j)_{j \geq 1}$ . If we take the union of the overlapping fatter boxes  $(Q_0(x_j))_{j \geq 1}$  of volume  $n_0 > \underline{n}_0$ , we obtain  $Q_i(x)$ . Since  $\underline{n}_0, n_0$  are both constants, we obtain that

$$n_i = (1 + o(1))\underline{n}_i, \quad \text{as } i \rightarrow \infty. \quad (\text{B7})$$

The edge set at level  $i$  consists of all edges in  $\mathcal{E}^{\text{base}}$  whose vertices are in the larger box  $\mathcal{Q}_i(x)$ , and additionally, the edges from certain “high-mark vertices” in  $\mathcal{V}^{\text{spr}}$  to  $\mathcal{V}^{\text{base}}$  with both endpoints inside the smaller box  $\mathcal{Q}_i^{\text{sub}}(x)$ .

For the connected component that realizes the event  $\mathcal{A}_{i-\text{good}, \varepsilon'}(x)$ , we have a lower bound on the proportion of vertices inside the smaller box  $\mathcal{Q}_i^{\text{sub}}(x)$ , and also on the amount of vertices in the range  $[\underline{n}_i^\gamma, 2\underline{n}_i^\gamma)$  in the sprinkled vertex set in  $\mathcal{Q}_i^{\text{sub}}(x)$ . These criteria in (B6) do not require any information about vertices in the ‘box-annulus’  $\mathcal{Q}_i(x) \setminus \mathcal{Q}_i^{\text{sub}}(x)$ .

We can compute the density of the component  $\mathcal{C}_{i,x}$  inside the smaller box  $\mathcal{Q}_i^{\text{sub}}(x)$ . By definition of  $\underline{n}_i$  in (B2) and  $\rho_i$  in (B5), for any  $\underline{n}_0$  sufficiently large, using that  $\eta/(\eta+2) \leq \varepsilon'$  due to (B1), this density is

$$\prod_{\ell=0}^i \rho_\ell = \rho_0 \prod_{\ell=1}^i (\ell + s)^{-2d} = \rho_0 (\underline{n}_i / \underline{n}_0)^{-2/(2+\eta)} \geq (\theta^{\text{base}}/2) \cdot n_i^{-\varepsilon'}. \quad (\text{B8})$$

Hence, combined with (B7), for all  $i$  sufficiently large,

$$\mathbb{P}(|\mathcal{C}_{n_i}^{(1)}| \geq n_i^{1-\varepsilon'} \theta^{\text{base}}/4) \geq \mathbb{P}(\mathcal{A}_{i-\text{good}, \varepsilon'}(0)). \quad (\text{B9})$$

This is closely related to what we aim for in Lemma 4.2. In the next three claims, we will show that the right-hand side can be made arbitrarily close to 1 for all  $i$  when  $\underline{n}_0$  is sufficiently large. Afterwards we prove Lemma 4.2 to bound the left-hand side for arbitrary  $n$ .

**Claim B.2** (Ergodicity and uniqueness). *Consider a supercritical KSRG as in Lemma 4.2 and Definition B.1. Then the measure  $\mathbb{P}$  is ergodic, and there is a unique infinite component.*

**Proof** We start with ergodicity. We construct the graph as a marked Poisson point process  $\tilde{\mathcal{V}}$  such that the graph remains invariant under translation. These marks are ‘additional’ to the vertex marks  $w_v$ . Namely, we equip each vertex  $v = (x_v, w_v) \in \mathcal{V}$  with a sequence of iid  $\text{Unif}[0, 1]$  random variables  $(U_{vi})_{i \geq 1}$ . Given this marked PPP, we write  $\mathcal{U}(v) = (u_i)_{i \geq 1}$  for all vertices whose first spatial coordinate is strictly larger than the first spatial coordinate of  $v$ , increasingly ordered with respect to their Euclidean distance to  $v$ . We include an edge between  $v$  and  $u_i$  iff  $U_{vi} \leq p(v, u_i)$ . This yields a graph that has the same distribution as in Definition 2.1. By the adaptation of [56, Proposition 8.13] to marked Poisson point processes (see also [56, Exercise 10.1]), such a process is ergodic: any event in the invariant  $\sigma$ -algebra has either probability 0 or 1.

The uniqueness of the infinite cluster follows by the classical Burton-Keane argument [18], which makes implicit use of ergodicity. We do *not* sketch the Burton-Keane argument here, just highlight one point: the argument requires that the expected size of the vertex boundary of  $\Lambda_n$  (i.e., the number of vertices in  $\Lambda_n$  with an edge leaving  $\Lambda_n$ ) is of smaller order than the total number of vertices inside  $\Lambda_n$ . Formally, we have

to show that

$$\text{VB}_n := \mathbb{E} \left[ |\{v \in \Lambda_n : v \text{ connected by an edge to } \Lambda_n^c\}| \right] \quad (\text{B10})$$

is  $o(n)$ . Fix  $\varepsilon > 0$  arbitrarily small, and let  $\delta = \delta(\varepsilon)$  be a small constant depending on  $\varepsilon$ , such that the expected number of vertices in  $\Lambda_n$  within distance  $\delta n^{1/d}$  from the boundary  $\partial\Lambda_n$  is at most  $\varepsilon n$ . Since  $\varepsilon > 0$  is arbitrary, we may bound from above the vertex boundary in (B10) by counting all these vertices in  $\text{VB}_n$ . The remaining vertices are in  $\Lambda_{n(1-\delta)^d}$ , and they only contribute to the vertex boundary in (B10) if they have an edge of length  $\delta n^{1/d}$ . Thus,

$$\text{VB}_n \leq \varepsilon n + \mathbb{E}[|\{v \in \Lambda_n : v \text{ incident to edge of length at least } \delta n^{1/d}\}|]. \quad (\text{B11})$$

We use translation invariance of the graph and Mecke's formula, and arrive at

$$\text{VB}_n \leq \varepsilon n + n(1 - \delta)^d \mathbb{P}^0(0 \text{ incident to edge of length at least } \delta n^{1/d}).$$

When  $\alpha > 1$  and  $\tau > 2$  in Assumption 2.2, the graph is locally finite, see also [27, 44, 61], which implies that the probability on the left-hand side tends to 0 as  $n \rightarrow \infty$ . Therefore, the expected size of the vertex boundary is  $o(n)$ . By the Burton-Keane argument, the uniqueness of the infinite cluster follows.  $\square$

**Claim B.3** (Induction base for overlapping boxes). *Consider a KSRG in the setting of Lemma 4.2 and Definition B.1. Let  $\varepsilon' > 0$  is a sufficiently small constant. For each  $\delta > 0$ , there exists a constant  $\underline{n}_0^* > 0$  such that for each  $\underline{n}_0 \geq \underline{n}_0^*$  there exists  $n_0 > 0$  such that for all  $x \in \mathbb{R}^d$ ,*

$$\mathbb{P}(\mathcal{Q}_0(x) \text{ is } (0, \varepsilon')\text{-good}) \geq 1 - \delta.$$

**Proof** We estimate the probability of  $\mathcal{A}_{0\text{-good}, \varepsilon'}(x)$  defined in (B6). We set  $x = 0$ , so that  $\mathcal{Q}_0^{\text{sub}}(0) = \Lambda_{\underline{n}_0}$  while  $\mathcal{Q}_0(0) = \Lambda_{n_0}$ . We write  $\mathcal{C}_\infty^{\text{base}}$  for the unique infinite component of  $\mathcal{G}^{\text{base}}$ . By definition,  $\theta^{\text{base}} = \mathbb{P}(0 \in \mathcal{C}_\infty^{\text{base}})$ , so that  $\mathcal{C}_\infty^{\text{base}}$  typically contains a  $\theta^{\text{base}}$ -proportion of the vertices in  $\Lambda_{\underline{n}_0}$  by translation invariance. By ergodicity (Claim B.2), the proportion of vertices in  $\Lambda_{\underline{n}_0}$  in  $\mathcal{C}_\infty^{\text{base}}$  converges in probability to  $\theta^{\text{base}}$  as  $\underline{n}_0 \rightarrow \infty$ . In Definition B.1 we defined  $\rho_0 = \theta^{\text{base}}/2$ . Hence, for  $\underline{n}_0$  sufficiently large,

$$\mathbb{P}(|\Lambda_{\underline{n}_0} \cap \mathcal{C}_\infty^{\text{base}}| \geq \rho_0 \underline{n}_0) \geq 1 - \delta/4.$$

Since the infinite component is unique (Claim B.2) and the graph is locally finite, given  $\delta$ , there exists  $n_0 = n_0(\underline{n}_0)$  such that with probability at least  $1 - \delta/2$ , the separate components that  $\mathcal{C}_\infty^{\text{base}}$  induces inside  $\Lambda_{\underline{n}_0}$  are all in a single component inside the larger box  $\Lambda_{n_0}$ . Formally, let  $n_0$  be so large that

$$\mathbb{P}(|\Lambda_{\underline{n}_0} \cap \mathcal{C}_\infty^{\text{base}}| \geq \rho_0 \underline{n}_0, \Lambda_{\underline{n}_0} \cap \mathcal{C}_\infty^{\text{base}} \text{ in same component in } \mathcal{G}_{n_0}^{\text{base}}) \geq 1 - \delta/2 \quad (\text{B12})$$

holds. This also implies that the event  $\mathcal{A}_{\text{base}} := \{\exists \text{ comp. } \mathcal{C} \text{ in } \mathcal{G}_{n_0} : |\mathcal{C} \cap \Lambda_{\underline{n}_0}| \geq \rho_0 \underline{n}_0\}$  occurs with probability at least  $1 - \delta/2$ . We now condition on the realization of the graph  $\mathcal{G}^{\text{base}}$  satisfying  $\mathcal{A}_{\text{base}}$ . Let  $\mathcal{C}$  be an arbitrary component in  $\mathcal{G}_{n_0}^{\text{base}}$  that the event  $\mathcal{A}_{\text{base}}$  describes. This component only uses vertices and edges of  $\mathcal{G}_{n_0}^{\text{base}}$ , which is also a subgraph of the graph  $\widehat{\mathcal{G}}_{0,x}$ . The component  $\widehat{\mathcal{C}} \supseteq \mathcal{C}$  in  $\widehat{\mathcal{G}}_{i,0}$  thus satisfies the first condition of  $\mathcal{A}_{0-\text{good},\varepsilon'}$  in (B6). For the second criterion of  $\mathcal{A}_{0-\text{good},\varepsilon'}$  we will compute how many vertices of  $\mathcal{V}_{\underline{n}_0}^{\text{spr}}[\underline{n}_0^\gamma, 2\underline{n}_0^\gamma]$  connect to  $\mathcal{C}$  (and thus form  $\widehat{\mathcal{C}}$ ). Since we revealed a realization of  $\mathcal{G}^{\text{base}}$  satisfying  $\mathcal{A}_{\text{base}}$ ,  $\mathcal{C}$  is known, and we obtain

$$\begin{aligned} & \mathbb{1}_{\{\mathcal{A}_{\text{base}}\}} \mathbb{P}(\neg \mathcal{A}_{0-\text{good},\varepsilon'}(x) \mid \mathcal{G}^{\text{base}}) \\ & \leq \mathbb{1}_{\{\mathcal{A}_{\text{base}}\}} \mathbb{P}(|\{u \in \mathcal{V}_{\underline{n}_0}^{\text{spr}}[\underline{n}_0^\gamma, 2\underline{n}_0^\gamma] : u \sim \mathcal{C}\}| < \underline{n}_0^{\zeta-2\varepsilon'} \mid \mathcal{G}^{\text{base}}). \end{aligned}$$

The vertices in  $\mathcal{V}_{\underline{n}_0}^{\text{spr}}[\underline{n}_0^\gamma, 2\underline{n}_0^\gamma]$  form a PPP with intensity  $(1-q)\text{Leb}(\cdot) \times (\tau-1)w^{-\tau}dw$ . We compute the mean and estimate it from below using that  $1 - \gamma(\tau-1) = \zeta$ :

$$\mathbb{E}[|\mathcal{V}_{\underline{n}_0}^{\text{spr}}[\underline{n}_0^\gamma, 2\underline{n}_0^\gamma]|] \geq (1-q)\underline{n}_0^{1-\gamma(\tau-1)}(1-2^{-(\tau-1)}) \geq \underline{n}_0^{\zeta-\varepsilon'}.$$

We estimate the mark of each vertex in  $\mathcal{C} \cap \Lambda_{\underline{n}_0}$  from below by one, the spatial distance by at most the diameter of  $\Lambda_{\underline{n}_0}$ , and the size of  $\mathcal{C} \cap \Lambda_{\underline{n}_0}$  by  $\rho_0 \underline{n}_0$ . So, a single vertex in  $\mathcal{V}_{\underline{n}_0}^{\text{spr}}[\underline{n}_0^\gamma, 2\underline{n}_0^\gamma]$  connects to  $\mathcal{C}$  with probability at least

$$\begin{aligned} 1 - \left(1 - p\left(1 \wedge \beta \frac{\underline{n}_0^\gamma}{d^{d/2}\underline{n}_0}\right)^\alpha\right)^{\rho_0 \underline{n}_0} & \geq 1 - \exp(-p\beta^\alpha \rho_0 d^{-\alpha d/2} \underline{n}_0^{\gamma\alpha-\alpha+1}) \\ & = 1 - \exp(-p\beta^\alpha \rho_0 d^{-\alpha d/2}) \geq 2\underline{n}_0^{-\varepsilon'}. \end{aligned}$$

Here we used that  $\gamma = 1 - 1/\alpha$  to obtain the second row so that powers of  $\underline{n}_0$  vanish, and that the last inequality holds for all  $\underline{n}_0$  sufficiently large. Given the vertex set, vertices in  $\mathcal{V}^{\text{spr}}$  connect independently by an edge to  $\mathcal{C}$ . So, the total number of vertices in  $\mathcal{V}_{\underline{n}_0}^{\text{spr}}[\underline{n}_0^\gamma, 2\underline{n}_0^\gamma]$  dominates a Poisson random variable with mean  $2\underline{n}_0^{\zeta-2\varepsilon'}$ . By concentration for Poisson random variables (Lemma A.1), there exists  $c > 0$  such that

$$\mathbb{1}_{\{\mathcal{A}_{\text{base}}\}} \mathbb{P}(\mathcal{Q}_0(x) \text{ is } (0, \varepsilon')\text{-bad}) \leq \exp(-c\underline{n}_0^{\zeta-2\varepsilon'}).$$

The right-hand side is at most  $\delta/2$  if  $\underline{n}_0$  is large. Together with (B12), this finishes the proof.  $\square$

The key idea of Claim B.3 above is that the separate components of the infinite component in  $\mathcal{G}^{\text{base}}$  induced in a box  $\Lambda_{\underline{n}_0}$  connect to each other in a larger box  $\Lambda_{n_0}$ . However, we had no control over the relation between  $n_0$  and  $\underline{n}_0$ , and thus lose the density estimate of the connected component in the larger box. The multi-scale renormalization in the following claim makes this control possible, using that the added annulus in the ‘fattened’ box has constant width across levels, see (B7).

**Claim B.4** (Iterative renormalization with overlapping boxes). *Consider a KSRG in the setting of Lemma 4.2, with the boxing scheme from Definition B.1. Assume  $\varepsilon' > 0$  is a sufficiently small constant. For each  $\delta > 0$ , there exists a constant  $\underline{n}_0^* = \underline{n}_0^*(\varepsilon', \delta) > 0$  such that for each  $\underline{n}_0 \geq \underline{n}_0^*$  there exists  $n_0 > 0$  so that for all  $x \in \mathbb{R}^d$  and all  $i \geq 0$ ,*

$$\mathbb{P}(\mathcal{Q}_i(x) \text{ is } (i, \varepsilon')\text{-good}) \geq 1 - \delta. \quad (\text{B13})$$

**Proof** We follow the same steps (*Sub*), (*Con*), and (*Mark*), similar to the proof of Claim 4.11. By Claim B.3 the statement holds for  $i = 0$ , so we may assume inductively that it holds until  $i - 1$  and that  $i \geq 1$ .

**Good subboxes.** Consider  $\mathcal{Q}_i(x)$ , and take a partition of the smaller box  $\mathcal{Q}_i^{\text{sub}}(x)$  into  $m_i$  many subboxes of volume  $\underline{n}_{i-1}$  with centers  $(x_j)_{j \leq m_i}$ . We then obtain  $((\mathcal{Q}_{i-1}^{\text{sub}}(x_j), \mathcal{Q}_{i-1}(x_j)))_{j \leq m_i}$ , where  $(\mathcal{Q}_{i-1}^{\text{sub}}(x_j))_{j \leq m_i}$  are the subboxes forming the partitioning, and  $(\mathcal{Q}_{i-1}(x_j))_{j \leq m_i}$  their fattened versions. Since two neighboring fattened subboxes overlap, the events that two fattened boxes are  $(i - 1, \varepsilon')$ -good are not independent. Let  $\mathcal{A}_{\text{sub}}$  be the event that at least  $\rho_i m_i$  fattened boxes are  $(i - 1, \varepsilon')$ -good. By Markov's inequality and translation invariance,

$$\begin{aligned} \mathbb{P}(\neg \mathcal{A}_{\text{sub}}) &= \mathbb{P}(|\{j : \mathcal{Q}_{i-1}(x_j) \text{ is } (i, \varepsilon')\text{-bad}\}| \geq (1 - \rho_i)m_i) \\ &\leq \frac{1}{1 - \rho_i} \mathbb{P}(\mathcal{Q}_{i-1}(0) \text{ is } (i, \varepsilon')\text{-bad}) \\ &\leq \left(1 + \frac{2}{(i + s)^{2d}}\right) \mathbb{P}(\mathcal{Q}_{i-1}(0) \text{ is } (i, \varepsilon')\text{-bad}), \end{aligned} \quad (\text{B14})$$

where for the second bound we used  $\rho_i = (i + s)^{-2d}$ , so that  $1/(1 - \rho_i) \leq 1 + 2(i + s)^{-2d}$  for all  $i \geq 1$  and  $d \geq 1$ .

**Connection between subboxes.** We write  $\text{pre-}\widehat{\mathcal{G}}_{i,x}$  for the union of the graphs  $\cup_{j \leq m_i} \widehat{\mathcal{G}}_{i-1,x_j}$  inside the subboxes at level  $i - 1$ , see (B3) and (B4), and assume that  $\mathcal{A}_{\text{sub}}$  holds. Let  $\text{inter-}\widehat{\mathcal{G}}_{i,x}$  denote the graph on the same vertex set as  $\text{pre-}\widehat{\mathcal{G}}_{i,x}$ , but including also all edges between the vertices in  $\cup_{j \leq m_i} \widehat{\mathcal{V}}_{i-1,j}$ .

**The auxiliary graph.** Similar to below (4.31), we re-label the subboxes so that the subboxes with indices  $j \leq \rho_i m_i$  are all  $(i - 1, \varepsilon')$ -good. We consider the auxiliary graph  $\mathcal{H}_i = (\mathcal{V}_{\mathcal{H}_i}, \mathcal{E}_{\mathcal{H}_i})$ : each vertex  $j \in [\rho_i m_i] =: \mathcal{V}_{\mathcal{H}_i}$  corresponds to a component  $\mathcal{C}_{i-1,x_j}$  that realizes the event  $\mathcal{A}_{i\text{-good},\varepsilon'}(x_j)$  in the  $j$ -th subbox, see (B6). To determine the edges of the auxiliary graph, we use the vertices at level  $i - 1$  that belong to  $\text{inter-}\widehat{\mathcal{G}}_{i,x}$  as follows: we connect vertex  $j$  and  $k$  in  $\mathcal{H}_i$  if there is an edge between a sprinkled vertex in  $\mathcal{C}_{i-1,j} \cap \mathcal{Q}_{i-1}^{\text{sub}}(x_j)$  to some base-vertex in  $\mathcal{C}_{i-1,k} \cap \mathcal{Q}_{i-1}^{\text{sub}}(x_k)$  or vice versa:

$$\begin{aligned} \{j \sim_{\mathcal{H}_i} k\} &\iff \left\{ (\mathcal{C}_{i-1,j} \cap \mathcal{V}_{\mathcal{Q}_{i-1}^{\text{sub}}(x_j)}^{\text{spr}} [\underline{n}_{i-1}^\gamma, 2\underline{n}_{i-1}^\gamma]) \sim_{\widehat{\mathcal{G}}_{i,x}} (\mathcal{C}_{i-1,k} \cap \mathcal{V}_{\mathcal{Q}_{i-1}^{\text{sub}}(x_k)}^{\text{base}}) \right\} \\ &\cup \left\{ (\mathcal{C}_{i-1,k} \cap \mathcal{V}_{\mathcal{Q}_{i-1}^{\text{sub}}(x_k)}^{\text{spr}} [\underline{n}_{i-1}^\gamma, 2\underline{n}_{i-1}^\gamma]) \sim_{\widehat{\mathcal{G}}_{i,x}} (\mathcal{C}_{i-1,j} \cap \mathcal{V}_{\mathcal{Q}_{i-1}^{\text{sub}}(x_j)}^{\text{base}}) \right\}. \end{aligned}$$

Each good subbox  $\mathcal{Q}_{i-1}(x_j)$  is contained in  $\mathcal{Q}_i^{\text{sub}}(x)$ , and the connected component  $\mathcal{C}_{i-1,j}$  satisfies  $|\mathcal{C}_{i-1,j} \cap \mathcal{Q}_{i-1}^{\text{sub}}(x_j)| \geq \underline{n}_{i-1} \prod_{\ell=0}^{i-1} \rho_\ell$ . On  $\mathcal{A}_{\text{sub}}$  there are at least  $\rho_i m_i$  good subboxes. Thus, if  $\mathcal{H}_i$  is connected, then there is a component  $\mathcal{C}_{i,x}$  satisfying:

$$|\mathcal{C}_{i,x} \cap \mathcal{Q}_i^{\text{sub}}(x)| \geq \rho_i m_i \cdot \underline{n}_{i-1} \prod_{\ell=0}^{i-1} \rho_\ell = \underline{n}_i \prod_{\ell=0}^i \rho_\ell.$$

We bound from above the probability that  $\mathcal{H}_i$  is not connected by the probability that  $\mathcal{H}_i$  is not the complete graph. We already lost sharp bounds in (B14) as opposed to (4.32) and (4.35) where we used the connectivity of the Erdős-Rényi random graph. So even though the edges in  $\mathcal{H}_i$  are conditionally independent given  $\text{pre-}\widehat{\mathcal{G}}_{i,x}$ , we will use Markov's inequality to bound edge-presence events. We compute conditionally on the realization of  $\text{pre-}\widehat{\mathcal{G}}_{i,x}$ , the union of subgraphs inside level- $(i-1)$ -subboxes that

$$\begin{aligned} \mathbb{1}_{\{\mathcal{A}_{\text{sub}}\}} \mathbb{P} \left( \# \text{component } \mathcal{C}_{i,x} \text{ in inter-}\widehat{\mathcal{G}}_{i,x} : |\mathcal{C}_{i,x} \cap \mathcal{Q}_i^{\text{sub}}(x)| \geq \underline{n}_i \prod_{\ell=0}^i \rho_\ell \mid \text{pre-}\widehat{\mathcal{G}}_{i,x} \right) \\ \leq \mathbb{1}_{\{\mathcal{A}_{\text{sub}}\}} \mathbb{P}(\mathcal{H}_i \text{ is not connected} \mid \text{pre-}\widehat{\mathcal{G}}_{i,x}) \\ \leq m_i^2 \max_{j,k} \mathbb{1}_{\{\mathcal{A}_{\text{sub}}\}} \mathbb{P}(j \sim_{\mathcal{H}_i} k \mid \text{pre-}\widehat{\mathcal{G}}_{i,x}). \end{aligned}$$

To bound the probability on the right-hand side, we use the two properties of  $\mathcal{C}_{i-1,j}$  of a good  $(i-1, \varepsilon')$  subbox in (B6) about its density in  $\mathcal{Q}_{i-1}^{\text{sub}}(x_j)$  and its level- $(i-1)$  high-mark vertices. Since the union of the smaller level- $(i-1)$ -boxes forms  $\mathcal{Q}_i^{\text{sub}}(x)$  (see reasoning below Definition B.1), the maximal distance between two vertices in  $\mathcal{Q}_{i-1}(x_j)$  and  $\mathcal{Q}_{i-1}(x_k)$  is  $\sqrt{d} \underline{n}_i^{1/d}$ . Moreover, by the definition of the level- $(i-1)$  edge set in (B4), none of the edges between  $\mathcal{V}_{\mathcal{Q}_{i-1}^{\text{sub}}(x_j)}^{\text{pr}}$  and  $\mathcal{V}_{\mathcal{Q}_{i-1}^{\text{sub}}(x_k)}^{\text{base}}$  have been revealed in  $\text{pre-}\widehat{\mathcal{G}}_{i,x}$  and they are thus conditionally independently present. We bound this probability from above by only requiring that there is no edge between  $\mathcal{C}_{i-1,j} \cap \mathcal{V}_{\mathcal{Q}_{i-1}^{\text{sub}}(x_j)}^{\text{pr}}[\underline{n}_{i-1}^\gamma, 2\underline{n}_{i-1}^\gamma]$  and  $\mathcal{C}_{i-1,k} \cap \mathcal{V}_{\mathcal{Q}_{i-1}^{\text{sub}}(x_k)}^{\text{base}}$ , and we use that we have lower bounds on the sizes of these sets in (B6):

$$\begin{aligned} \mathbb{1}_{\{\mathcal{A}_{\text{sub}}\}} \mathbb{P}(\mathcal{H}_i \text{ is not connected} \mid \text{pre-}\widehat{\mathcal{G}}_{i,x}) \\ \leq m_i^2 \prod_{u \in \mathcal{C}_{i-1,j} \cap \mathcal{V}_{\mathcal{Q}_{i-1}^{\text{sub}}(x_{j_1})}^{\text{pr}}[\underline{n}_{i-1}^\gamma, 2\underline{n}_{i-1}^\gamma]} \prod_{v \in \mathcal{C}_{i-1,k} \cap \mathcal{V}_{\mathcal{Q}_{i-1}^{\text{sub}}(x_{j_2})}^{\text{base}}} (1 - p(u, v)) \\ \leq m_i^2 \left( 1 - p \left( 1 \wedge \beta \frac{\underline{n}_{i-1}^\gamma}{d^{d/2} \underline{n}_i} \right)^\alpha \right)^{\underline{n}_{i-1}^{1+\zeta-2\varepsilon'} \prod_{\ell=0}^{i-1} \rho_\ell}. \end{aligned}$$

We use that  $\underline{n}_i = m_i \underline{n}_{i-1}$  and  $\gamma = 1 - 1/\alpha$ , so that the minimum is attained at the second term for any  $n_0$  sufficiently large. Hence,

$$\begin{aligned} \mathbb{1}_{\{\mathcal{A}_{\text{sub}}\}} \mathbb{P}(\mathcal{H}_i \text{ is not connected} \mid \text{pre-}\widehat{\mathcal{G}}_{i,x}) &\leq m_i^2 \exp \left( -p \beta^\alpha m_i^{-\alpha} \underline{n}_{i-1}^{(\gamma-1)\alpha} \underline{n}_{i-1}^{\zeta+1-2\varepsilon'} \prod_{\ell=0}^{i-1} \rho_\ell \right) \\ &= m_i^2 \exp \left( -m_i^{-\alpha} \underline{n}_{i-1}^{\zeta-2\varepsilon'} p \beta^\alpha \prod_{\ell=0}^{i-1} \rho_\ell \right). \quad (\text{B15}) \end{aligned}$$

The expression  $\prod_{\ell=0}^{i-1} \rho_\ell$  inside the exponent is at least  $\underline{n}_i^{-\varepsilon'}$  by (B8). We would also like to replace the factor  $m_i^{-\alpha}$  by  $\underline{n}_{i-1}^{-\varepsilon'}$ . For this it is enough to show that  $\underline{n}_{i-1}^{\varepsilon'} \geq m_i^\alpha$  with  $m_i$  and  $\underline{n}_i$  from (B2). We show this by induction. For  $i = 1$ ,  $m_1^\alpha = (1+s)^{(2+\eta)d\alpha} \leq \underline{n}_0^{\varepsilon'}$  holds for any  $\underline{n}_0$  sufficiently large. We advance the induction, assuming  $\underline{n}_{i-2}^{\varepsilon'} \geq m_{i-1}^\alpha$ . Using the recursion defining  $\underline{n}_i$  and the induction hypothesis,

$$\begin{aligned} \underline{n}_{i-1}^{\varepsilon'} &= m_{i-1}^{\varepsilon'} \underline{n}_{i-2}^{\varepsilon'} \geq m_{i-1}^{\varepsilon'+\alpha} \\ &= m_i^\alpha \left( \frac{m_{i-1}}{m_i} \right)^\alpha m_{i-1}^{\varepsilon'} = m_i^\alpha \left( \frac{i-1+s}{i+s} \right)^{(2+\eta)d\alpha} (i-1+s)^{(2+\eta)d\varepsilon'} \geq m_i^\alpha, \end{aligned}$$

by the choice of  $s$  just above (B2). Substituting this bound into (B15), and discounting the prefactor  $m_{i-1}^2$  and constant factors inside the exponential by another factor  $\underline{n}_{i-1}^{\varepsilon'}$ , we obtain

$$\begin{aligned} \mathbb{1}_{\{A_{\text{sub}}\}} \mathbb{P} \left( \nexists \text{comp. } C_i \text{ in inter-}\widehat{\mathcal{G}}_{i,x} : |C_i \cap \mathcal{Q}_i^{\text{sub}}(x)| \geq \underline{n}_i \prod_{\ell=0}^i \rho_\ell \mid \Big|_{\ell=0} \text{pre-}\widehat{\mathcal{G}}_{i,x} \right) \\ \leq \exp \left( - \underline{n}_{i-1}^{\varepsilon'-5\varepsilon'} \right). \end{aligned} \quad (\text{B16})$$

*Sprinkling high-mark vertices.* Let us denote the opposite event of (B16) by

$$\mathcal{A}_{\text{conn}} := \left\{ \exists \text{comp. } C_i \text{ in inter-}\widehat{\mathcal{G}}_{i,x} : |C_i \cap \mathcal{Q}_i^{\text{sub}}(x)| \geq \underline{n}_i \prod_{\ell=0}^i \rho_\ell \right\}.$$

On  $\mathcal{A}_{\text{conn}}$ ,  $\widehat{\mathcal{G}}_{i,x}$  satisfies the first criterion of being  $(i, \varepsilon')$ -good in (B6). We now look at the second criterion of  $\mathcal{A}_{i\text{-good}, \varepsilon'}(0)$  in (B6), conditioned on the graph  $\text{inter-}\widehat{\mathcal{G}}_{i,x}$  satisfying the event  $\mathcal{A}_{\text{conn}} \cap A_{\text{sub}}$ . We reveal the sprinkled vertices  $\mathcal{V}^{\text{spr}}[\underline{n}_i^\gamma, 2\underline{n}_i^\gamma)$  with location in  $\mathcal{Q}_i^{\text{sub}}(x)$ . These vertices are independent of  $\text{inter-}\widehat{\mathcal{G}}_{i,x}$ , which only contains vertices with mark below  $2\underline{n}_{i-1}^\gamma$ . This mark threshold is below  $\underline{n}_i^\gamma$  for sufficiently large  $\underline{n}_0$  and small  $\varepsilon'$  see (B1–B2). The expected number of revealed vertices is then  $(1-q)\underline{n}_i^{1-\gamma(\tau-1)}(1-2^{-(\tau-1)})$  using Definition 4.5.

We bound from below the probability that such a vertex connects by an edge to one of the at least  $\underline{n}_i \prod_{\ell=0}^i \rho_\ell$  many vertices in  $C_i \cap \mathcal{Q}_i^{\text{sub}}(x)$  of  $\text{inter-}\widehat{\mathcal{G}}_{i,x}$ . The probability that a single vertex of mark at least  $\underline{n}_i^\gamma$  does not connect by an edge to  $C_i$  is at most

$$\begin{aligned} \left( 1 - p \left( 1 \wedge \beta \frac{\underline{n}_i^\gamma}{d^{d/2} \underline{n}_i} \right)^\alpha \right)^{\underline{n}_i \prod_{\ell=0}^i \rho_\ell} &\leq \exp \left( - p \beta^\alpha d^{-\alpha d/2} \underline{n}_i^{(\gamma-1)\alpha+1} \prod_{\ell=0}^i \rho_\ell \right) \\ &= \exp \left( - p \beta^\alpha d^{-\alpha d/2} \prod_{\ell=0}^i \rho_\ell \right), \end{aligned}$$



using again  $\gamma = 1 - 1/\alpha$  and cancellation of the powers of  $\underline{n}_i$ . Each vertex in  $\mathcal{V}^{\text{sp}}$  connects by an edge independently to the vertices in  $\mathcal{C}_i$ . Thus, the total number of vertices from  $\mathcal{V}^{\text{sp}}[\underline{n}_i', 2\underline{n}_i'] \cap \mathcal{Q}_i^{\text{sub}}(x)$  connecting to  $\mathcal{C}_i$  dominates a Poisson random variable with mean

$$\begin{aligned} & (1-q)(1-2^{-(\tau-1)})\underline{n}_i^\zeta \cdot \left(1 - \exp\left(-p\beta^\alpha d^{-\alpha d/2} \prod_{\ell=0}^i \rho_\ell\right)\right) \\ & \geq c\underline{n}_i^\zeta \prod_{\ell=0}^i \rho_\ell \geq c\rho_0 \underline{n}_i^{\zeta-\varepsilon'} \geq 2\underline{n}_i^{\zeta-2\varepsilon'}, \end{aligned}$$

where the second bound follows for some constant  $c > 0$ , the third bound by (B8), and the last bound if  $\underline{n}_0$  is sufficiently large. We apply concentration for Poisson random variables (Lemma A.1). We recall the definition of  $\mathcal{A}_{i-\text{good}, \varepsilon'}$  from (B6), and combine it with (B16) to obtain for  $\underline{n}_0$  sufficiently large that

$$\mathbb{1}_{\{\mathcal{A}_{\text{sub}} \cap \mathcal{A}_{\text{conn}}\}} \mathbb{P}(\neg \mathcal{A}_{i-\text{good}, \varepsilon'}(0)) \leq 2 \exp\left(-\underline{n}_i^{\zeta-4\varepsilon'} \mid \text{inter-}\widehat{\mathcal{G}}_{i,x}\right).$$

We combine this bound with  $\mathbb{P}(\neg \mathcal{A}_{\text{sub}}) \leq (1 + 2(i+s)^{-2d})\mathbb{P}(\neg \mathcal{A}_{(i-1)-\text{good}, \varepsilon'}(0))$  from (B14). Then

$$\mathbb{P}(\neg \mathcal{A}_{i-\text{good}, \varepsilon'}(0)) \leq (1 + 2(i+s)^{-2d})\mathbb{P}(\neg \mathcal{A}_{(i-1)-\text{good}, \varepsilon'}(0)) + 2 \exp\left(-\underline{n}_i^{\zeta-5\varepsilon'}\right). \quad (\text{B17})$$

*Combining everything: a recursive bound.* We finish the proof of (B13) by some analysis. We show that the right-hand side of (B17) is at most  $\delta$  by working out the recursion in  $i$ . Fix  $\delta' > 0$  sufficiently small. By Claim B.3,  $\mathbb{P}(\neg \mathcal{A}_{0-\text{good}, \varepsilon'}(0)) \leq \delta'$  for all  $\underline{n}_0$  sufficiently large. We may assume inductively that for all  $1 \leq j \leq i-1$ ,

$$\mathbb{P}(\neg \mathcal{A}_{j-\text{good}, \varepsilon'}(0)) \leq \delta' \prod_{\ell=1}^j \left(1 + \frac{3}{(\ell+s)^{2d}}\right).$$

Then, using this bound in (B17) we obtain

$$\mathbb{P}(\neg \mathcal{A}_{i-\text{good}, \varepsilon'}(0)) \leq \delta' \left(1 + \frac{2}{(i+s)^{2d}}\right) \prod_{\ell=1}^{i-1} \left(1 + \frac{3}{(\ell+s)^{2d}}\right) + 2 \exp\left(-\underline{n}_i^{\zeta-5\varepsilon'}\right).$$

Showing that the right-hand side is at most  $\delta' \prod_{\ell=1}^i \left(1 + \frac{3}{(\ell+s)^{2d}}\right)$  is equivalent to showing that

$$1 + \frac{2}{(i+s)^{2d}} + \frac{2 \exp\left(-\underline{n}_i^{\zeta-5\varepsilon'}\right)}{\delta' \prod_{\ell=1}^{i-1} (1 + 3/(\ell+s)^{2d})} \leq 1 + \frac{3}{(i+s)^{2d}}. \quad (\text{B18})$$

The product in the denominator on the left-hand side is bounded away from 0. Since  $\underline{n}_i$  grows superpolynomially by its definition in (B2), the bound (B18) is satisfied for all  $i \geq 1$  if  $\underline{n}_0$  is sufficiently large. This advances the induction, and we conclude that for all  $i \geq 0$  and a proper choice of the constant  $\delta' > 0$ ,

$$\mathbb{P}(\neg \mathcal{A}_{i-\text{good}, \varepsilon'}(0)) \leq \delta' \prod_{\ell=1}^{\infty} \left(1 + \frac{3}{(\ell + s)^{2d}}\right) = \delta.$$

□

We now prove Lemma 4.2. Similarly to the proof of Proposition 4.1 on page 37, this step generalizes the previous claim to arbitrary box size  $n$  (instead of  $n$  restricted along the subsequence  $(n_i)_{i \geq 1}$ ).

**Proof of Lemma 4.2** Fix  $\varepsilon' \in (0, \varepsilon/2)$ . Let  $c > 0$  be a small constant. It suffices to show that for all  $\delta$ , there exists  $n_*$  such that for all  $n \geq n_*$ ,

$$\mathbb{P}(|C_n^{(1)}| < n^{1-2\varepsilon'}) \leq \delta. \quad (\text{B19})$$

Consider the sequence  $(n_i)_{i \geq 0}$  from Definition B.1, with  $n_0$  given by Claim B.4 such that

$$\mathbb{P}(\neg \mathcal{A}_{i-\text{good}, \varepsilon'}(x)) \leq \delta/4. \quad (\text{B20})$$

for all  $i \geq 0$ . Assume  $n \gg n_0$ , and define  $m$ ,  $i_*$ , and  $\tilde{n}$  as in (4.40), i.e.,

$$i_* = \max\{i : n_i < n\}, \quad m := \lfloor (n/n_{i_*})^{1/d} \rfloor^d, \quad \tilde{n} := n_{i_*} m. \quad (\text{B21})$$

We will assume that  $i_*$  is a sufficiently large constant. We consider a box  $\Lambda_{\tilde{n}}$  of volume  $\tilde{n}$  inside  $\Lambda_n$  and partition this smaller box into *disjoint* subboxes of volume  $n_{i_*}$ . By construction,  $\Lambda_{\tilde{n}}$  covers a uniformly positive proportion of the larger box  $\Lambda_n$ , i.e., exactly as in (4.41), there exists an absolute constant  $\nu > 0$  such that

$$m \geq \nu(n/n_{i_*}). \quad (\text{B22})$$

Moreover, we may assume without loss of generality that  $m$  is at least a large constant  $m_0$ , as otherwise the statement is immediately implied by (B9) and Claim B.4. We denote the boxes of the partition of  $\Lambda_{\tilde{n}}$  by  $\mathcal{Q}_{i_*}(x_1), \dots, \mathcal{Q}_{i_*}(x_m)$ . By (B20) and Markov's inequality,

$$\mathbb{P}(\mathcal{A}_{\text{sub}}) := \mathbb{P}\left(\sum_{j \in [m]} \mathbb{1}_{\{\mathcal{Q}_{i_*}(x_j) \text{ is } (i_*, \varepsilon')\text{-good}\}} \geq m/2\right) \geq 1 - \delta/2. \quad (\text{B23})$$

We reveal the subgraphs  $\widehat{\mathcal{G}}_{i_*, x_j}$  from Definition B.1 in the subboxes, and denote their union by  $\text{pre-}\widehat{\mathcal{G}}_{\tilde{n}}$ . We condition the graph on satisfying  $\mathcal{A}_{\text{sub}}$ : each good subbox contains

a large component  $\mathcal{C}_j$  of size at least  $n_{i_*}^{1-\varepsilon'}$ . If the components  $\mathcal{C}_j$  are all connected by an edge to each other, then  $\mathcal{G}_{\tilde{n}}$  contains a connected component of size at least

$$(1/2)m\underline{n}_{i_*} \prod_{\ell=1}^{i_*} \rho_{\ell} \geq (1/2)m\underline{n}_{i_*}^{1-\varepsilon'} \geq (1/4)mn_{i_*}^{1-\varepsilon'} \geq (v/4)nn_{i_*}^{-\varepsilon'} \geq n^{1-2\varepsilon'}$$

using (B8), (B22), and the fact that  $n_i = \underline{n}_i(1+o(1))$  as  $i \rightarrow \infty$  for all  $n$  is sufficiently large.

We thus need to study the connectivity between the large connected components  $\mathcal{C}_j$  in the good subboxes. By the definition of  $(i, \varepsilon')$  goodness in (B6),  $\mathcal{C}_j$  has size at least  $n_{i_*}^{1-\varepsilon'}$ , and contains at least  $\underline{n}_{i_*}^{\zeta-2\varepsilon'}$  vertices of mark at least  $\underline{n}_{i_*}^{\gamma}$ . The edges between subboxes have not yet been revealed. The largest possible distance between two vertices in  $\Lambda_{\tilde{n}}$  is  $\sqrt{d}\tilde{n}^{1/d} = \sqrt{d}m\tilde{n}_{i_*}^{1/d}$ . Thus, the probability that the large components  $\mathcal{C}_j$  and  $\mathcal{C}_k$  in two good subboxes  $j, k$  are not connected by an edge in  $\mathcal{G}_n$  is at most

$$\begin{aligned} \prod_{u \in \mathcal{C}_j, v \in \mathcal{C}_k} (1 - p(u, v)) &\leq \left(1 - p\left(\beta \frac{\underline{n}_{i_*}^{\gamma}}{d^{d/2}n_{i_*}m}\right)\right)^{\alpha} 2n_{i_*}^{1+\zeta-3\varepsilon'} \\ &\leq \exp\left(-p\beta^{\alpha}d^{-\alpha d/2}\underline{n}_{i_*}^{\gamma\alpha+1-\alpha+\zeta-3\varepsilon'}m^{-\alpha}\right) \leq \exp\left(-\underline{n}_{i_*}^{\zeta-4\varepsilon'}m^{-\alpha}\right), \end{aligned}$$

having used  $\gamma = 1 - 1/\alpha$  in the last step, combined with a bound of  $\underline{n}_i^{-\varepsilon'}$  to account for the constant prefactors. Considering the at most  $m^2$  potential edges, and combining it with (B23), we obtain

$$\mathbb{P}(|\mathcal{C}_n^{(1)}| < n^{1-2\varepsilon'}) \leq \delta/2 + m^2 \exp\left(-\underline{n}_{i_*}^{\zeta-4\varepsilon'}m^{-\alpha}\right).$$

By (B21),  $m$  is at most  $m_{i_*}$  defined in (B2). By the same reasoning as below (B15), it follows that  $m^{-\alpha} \geq \underline{n}_{i_*}^{-\varepsilon}$ . Thus, if  $\underline{n}_0$  is a sufficiently large constant, the second term on the right-hand side is smaller than  $\delta/2$ . This finishes the proof.  $\square$

## B.2 Useful bounds for the iterative renormalization

**Proof of Claim 4.9** We first discuss (4.19). Its proof is standard and based on simple estimates relating to integer parts and induction, so we leave it to the reader.

For the first inequality (Sub) in (4.16) we substitute  $n_i = m_i n_{i-1}$  on its right-hand side. Simplification gives that (Sub)-inequality is equivalent to  $\varepsilon_i m_i^{1-\zeta+\delta} \geq C$  for all  $i \geq 1$ . This follows for any sufficiently large  $n_0$  since  $m_i$  is doubly exponentially increasing,  $\varepsilon_i = (i+1)^{-2}$ , and  $1-\zeta = \gamma(\tau-1) > 0$ . For the second inequality (Con) in (4.17) we define the constant

$$\tilde{C} := C / \prod_{j=1}^{\infty} (1 - \varepsilon_j) = C / \prod_{j=2}^{\infty} (1 - 1/j^2),$$

which does not depend on  $n_0$ . It is elementary to check that  $C/\rho_{i-1}^2 \leq \tilde{C}n_0^{\delta/2}$  by definition of  $\rho_i$  in (4.14). Hence, dividing both sides with  $\rho_{i-1}^2$  in (4.17), and using this inequality, it is sufficient to show that for all  $n_0$  sufficiently large

$$m_i^{1-\alpha} n_{i-1}^\zeta \geq \tilde{C} n_0^{\delta/2} n_i^{\zeta-\delta}.$$

Using that  $n_i = m_i n_{i-1}$ , after rearranging this is equivalent to  $n_{i-1}^\delta \geq \tilde{C} n_0^{\delta/2} m_i^{\zeta-\delta+\alpha-1}$ . By (4.11),  $m_i \leq n_{i-1}^{\xi_\delta}$  with  $\xi_\delta = (\delta/2)/(\zeta + \alpha - 1 - \delta/2)$ , and raising then both sides to the power  $2/\delta$  gives that (4.17) holds if  $n_{i-1}^2 \geq \tilde{C}^{2/\delta} n_0 n_{i-1}^{1-\xi_\delta}$ . This inequality holds for all  $i \geq 1$  when  $n_0$  is sufficiently large (the choice of  $n_0$  depends on  $\tilde{C}$ ).

For the bound (Mark) in (4.18) we have to show that  $\rho_i n_i^\delta \geq C$ . The infinite product of  $(1 - 1/j^2)$  is strictly positive in (4.14), so  $\rho_i = \Theta(\rho_0) = \Theta(n_0^{-\delta/4})$  as  $n_0 \rightarrow \infty$ . Since  $(n_i)_{i \geq 0}$  is increasing, (Mark) follows again for sufficiently large  $n_0$ .

The bounds in (4.20) follow since  $\varepsilon_i = (i+1)^{-2}$  for  $i \geq 1$  and  $m_i$  and  $n_i$  increase doubly exponentially in  $i$ , while the powers appearing in (4.20) are all positive (e.g.  $1 - \zeta - \delta > 0$ ). The same is true for (4.22). We turn to the last bound (4.21). We substitute  $n = n_i m$  and rearrange to obtain that (4.21) is equivalent to  $n_i^\delta \geq m^{\zeta-\delta+\alpha-1}$ . Since  $m \leq m_{i+1}$  by assumption, this inequality holds for all such  $m$  if  $n_i^\delta \geq m_{i+1}^{\zeta-\delta+\alpha-1}$ . Since  $m_{i+1} \leq n_i^{\xi_\delta}$  by (4.11), it follows by definition of  $\xi_\delta$  in (4.8) that this last inequality holds if  $n_i^\delta \geq n_i^{\delta/2-\xi_\delta\delta/2}$ , which clearly holds since  $\xi_\delta > 0$ .  $\square$

### B.3 Poly-logarithmic mark-thresholds to belong to the giant

**Proof of Lemma 5.6 High-low regime.** Assuming (5.34), we will adapt the proof of the upper bound on  $|\mathcal{C}_n^{(2)}|$  from Section 6 to show (5.35) and (5.36). Let  $\zeta = \zeta_{\text{hl}}$ ,  $\gamma = \gamma_{\text{hl}}$ ,  $\eta = \eta_{\text{hl}}$ . Since we assume  $\zeta_{\text{hl}} > 0$ , we may assume that  $\alpha < \infty$ .

Take  $k = k_n = (A_1 \log n)^{1/(\zeta-\delta)}$  for a sufficiently large  $A_1 > 0$  and partition  $\Lambda_n$  into boxes of volume  $k$ , denoted by  $\mathcal{Q}_1, \dots, \mathcal{Q}_{n/k}$  similarly to Section 6, and labeled so that boxes with consecutive labels share a  $(d-1)$ -dimensional side. We write  $\mathcal{C}_{k,j}^{(1)}[1, A'k^\gamma]$  for the largest connected component induced by the vertices in  $\mathcal{Q}_j \times [1, A'k^\gamma]$ , where  $A'$  is the constant from the hypothesis (5.34). Now, we construct a backbone with the following revealment scheme (again similar to Section 6):

1. For each  $j = 1, \dots, n/k$ , reveal all edges inside box  $j$  belonging to  $\mathcal{V}_{k,j}[1, A'k^\gamma]$ .
2. Let  $\mathcal{U}_j = \{v \in \mathcal{V}_{k,j}[A'k^\gamma, 2A'k^\gamma] \mid v \sim \mathcal{C}_{k,j}^{(1)}[1, A'k^\gamma]\}$ , call the local giant  $\text{LG}_j := \mathcal{U}_j \cup \mathcal{C}_{k,j}^{(1)}[1, A'k^\gamma]$  and reveal all edges between any pair of vertices in  $\bigcup_{j \leq n/k} \text{LG}_j$ .
3. Reveal all vertices of mark at least  $2A'k^\eta = 2A'(A_1 \log n)^{\eta/(\zeta-\delta)}$  and their edges towards vertices in  $\bigcup_{j \leq n/k} \text{LG}_j$ .

Define a box  $\mathcal{Q}_j$  then to be 1-good if, for some  $\delta_1 > 0$ , it satisfies the event

$$\{|\mathcal{C}_{k,j}^{(1)}[1, A'k^\gamma]| > \rho k\} \cap \{|\mathcal{U}_j| \geq \delta_1 k^\gamma\}. \quad (\text{B24})$$

We call (stage-2-) $\mathcal{G}_n$  2-good if

- (2a) all boxes  $\mathcal{Q}_1, \dots, \mathcal{Q}_{n/k}$  are 1-good,  
 (2b)  $|\{u \in \mathcal{U}_j : u \sim \text{LG}_{j+1}\}| \geq 1$  for all  $j < n/k$ , i.e., each box contains a box-wedging vertex.

We call (stage-3-) $\mathcal{G}_n$  3-good if

- (3a) the stage-2- $\mathcal{G}_n$  is 2-good,  
 (3b) all vertices of mark at least  $2A'k^\eta$  are connected by an edge to a vertex in  $\cup_{j \leq n/k} \text{LG}_j$ .

By the hypothesis (5.34), for any fixed  $1 \leq i \leq n/k$ ,

$$\mathbb{P}(|\mathcal{C}_{k,j}^{(1)}[1, A'k^\gamma]| \leq \rho k) \leq A_2 \exp\left(-\frac{1}{A'}k^{\zeta-\delta}\right),$$

and by the same calculations as in Claim 6.5, there exists  $\varepsilon > 0$  such that

$$\mathbb{P}(|\mathcal{U}_j| \leq \delta_1 k^\zeta \mid |\mathcal{C}_{k,j}^{(1)}[1, A'k^\gamma]| > \rho k) \leq \exp(-\varepsilon k^\zeta).$$

Let us denote by 2a-good the event that condition (2a) is satisfied. Then, by a union bound over all  $n/k$  boxes, for some  $\varepsilon > 0$ ,

$$\mathbb{P}(\mathcal{G}_n \text{ is 2a-good}) \geq 1 - 2(n/k) \exp(-\varepsilon k^{\zeta-\delta}).$$

We estimate the probability of having a box-wedging vertex using the same calculations as in Claim 6.6. Again for some  $\varepsilon > 0$ ,

$$\mathbb{P}(|\{u \in \mathcal{U}_j : u \sim \text{LG}_{j+1}\}| = 0 \mid \mathcal{Q}_j, \mathcal{Q}_{j+1} \text{ are 1-good}) \leq \exp(-\varepsilon k^\zeta).$$

Hence, by a union bound over all  $n/k$  boxes, there exists  $\varepsilon > 0$  such that

$$\mathbb{P}(\mathcal{G}_n \text{ is 2-good} \mid \mathcal{G}_n \text{ is 2a-good}) \geq 1 - 2(n/k) \exp(-\varepsilon k^\zeta).$$

Thus, there exists a constant  $\varepsilon' > 0$  such that for  $n$  sufficiently large, using  $k = (A_1 \log n)^{1/(\zeta-\delta)}$ ,

$$\mathbb{P}(\mathcal{G}_n \text{ is 2-good}) \geq 1 - (n/\varepsilon'k) \exp(-\varepsilon'k^{\zeta-\delta}) \geq 1 - \exp(-\varepsilon' A_1 \log n + \log(n)).$$

Recall now  $C > 0$  from the statement of Lemma 5.6. Then, for all  $A_1 = A_1(\varepsilon)$  sufficiently large,

$$\mathbb{P}(\mathcal{G}_n \text{ is 2-good}) \geq 1 - o(n^{-C}). \quad (\text{B25})$$

In the first two stages (1.)–(2.) of the revealment scheme, we only revealed vertices with mark below  $2A'k^\gamma = 2A'(A_1 \log n)^{\gamma/(\zeta-\delta)}$  to obtain stage-2- $\mathcal{G}_n$ . Moreover, conditional on  $\mathcal{G}_n$  being 2-good, the local giants in all the subboxes are connected to each other by box-wedging vertices, and their size is at least  $\rho k$  by (B24). So,

$$|\mathcal{C}_n^{(1)}[1, 2A'k^\gamma]| = |\mathcal{C}_n^{(1)}[1, 2A'(A_1 \log n)^{\gamma/(\zeta-\delta)}]| \geq \sum_{j \leq n/k} |\text{LG}_j| \geq \rho n.$$

Combining this with (B25) gives that

$$\mathbb{P}(|\mathcal{C}_n^{(1)}[1, 2A'(A_1 \log n)^{\gamma/(\zeta-\delta)}]| \geq \rho n) \geq 1 - o(n^{-C}).$$

This proves (5.35) for  $A := 2A'A_1^{\gamma/(\zeta-\delta)}$ .

We proceed to the proof of (5.36). We use the same revelation scheme as above in (1.)–(3.). Choosing  $k := (A_1 \log n)^{1/(\zeta-\delta)}$ , for (5.36) it suffices if all vertices of mark at least  $2A'k^\eta$  connect by an edge to  $\mathcal{C}_n^{(1)}[1, 2A'k^\gamma]$  with probability at least  $1 - o(n^{-C})$ .

We first show that vertices with marks at least  $2A'k^\eta$  have not been revealed in the stage-2- $\mathcal{G}_n$  in (1.)–(2.) where marks up to  $2A'k^\gamma$  are revealed. To this end, we recall that  $\eta = 1 - \gamma(\tau - 1)/\alpha$  from (5.33). We then need to check that  $\eta \geq \gamma$ , equivalently,  $1 - \gamma(\tau - 1)/\alpha \geq \gamma$ . Elementary rearrangements yield that this inequality holds whenever  $1 - \gamma(\tau - 1) = \zeta_{\text{hl}} \geq 0$ . This holds by our assumption in the statement of Lemma 5.6. In conclusion, any vertex of mark  $\geq 2A'k^\eta$  is not part of stage-2- $\mathcal{G}_n$ .

Excluding first the event that stage-2- $\mathcal{G}_n$  is not 2-good, we can then assume that stage-2- $\mathcal{G}_n$  is 2-good and condition on its realization, and also on the realization of vertices with mark at least  $2A'k^\eta$ . We obtain for the complement of the event in (5.36) that

$$\begin{aligned} & \mathbb{P}(\nexists \mathcal{C} : \mathcal{C} \not\subseteq \mathcal{V}_n[2A'k^\eta, \infty), |\mathcal{C}| \geq \rho n) \\ & \leq \mathbb{E} \left[ \mathbb{P}(\exists v \in \mathcal{V}_n[2A'k^\eta, \infty) : v \sim \mathcal{C}_n^{(1)}[1, 2A'k^\gamma] | \mathcal{G}_n \text{ 2-good}, \mathcal{V}_n[2A'k^\eta, \infty)) \right] \\ & \quad + \mathbb{P}(\mathcal{G}_n \text{ not 2-good}). \end{aligned} \quad (\text{B26})$$

The second term on the right-hand side is at most  $o(n^{-C})$  by the previous calculations leading to (B25) for any  $A_1$  sufficiently large. For the first term, we consider all possible connections of a high-mark vertex  $u$  to the lower-mark vertices in the ‘local giant’ in the subbox containing  $u$ . By a union bound over all vertices of mark at least  $2A'k^\eta$ , and recalling that the maximum distance between two vertices in the same box is at most  $\sqrt{d}k^{1/d}$ , we have

$$\begin{aligned} & \mathbb{P}(\nexists \mathcal{C} : \mathcal{C} \not\subseteq \mathcal{V}_n[2A'k^\eta, \infty), |\mathcal{C}| \geq \rho n) \\ & \leq \mathbb{E} \left[ |\mathcal{V}_n[2A'k^\eta, \infty)| \left( 1 - p \left( 1 \wedge \beta \frac{2A'k^\eta}{d^{d/2}k} \right)^\alpha \right)^{\rho k} \right] + o(n^{-C}) \\ & \leq n \exp \left( -p\beta^\alpha d^{-\alpha d/2} \rho (2A')^\alpha k^{1-\alpha(1-\eta)} \right) + o(n^{-C}), \end{aligned} \quad (\text{B27})$$

where we used that the minimum is attained at the second term for  $k$  sufficiently large (this follows because  $\eta = 1 - \gamma(\tau - 1)/\alpha < 1$ ). We compute that  $1 - \alpha(1 - \eta) = 1 - \gamma(\tau - 1) = \zeta$  and substituting  $k = (A_1 \log n)^{1/(\zeta-\delta)}$ , we see that the right-hand side is at most  $o(n^{-C})$  if  $A_1$  is sufficiently large. Thus, (5.36) follows for the hl-regime for some constant  $A > 0$ .

*Low-low regime.* We only sketch the adaptation of the method to the low-low regime, where the calculations are less involved. Set now  $\zeta = \zeta_{\text{ll}} = 2 - \alpha$ ,  $\gamma = \gamma_{\text{ll}} = (\alpha -$

$1)/(\tau - 1)$ , set  $k = k_n = (A_1 \log n)^{1/(\zeta - \delta)}$  and partition  $\Lambda_n$  into boxes  $\mathcal{Q}_1, \dots, \mathcal{Q}_{n/k}$  of volume  $k$ . Set also  $\eta = \eta_{\text{II}} = 1/\alpha$ . We modify slightly the revelation scheme in (1.)–(3.) as follows: there are no sets  $\mathcal{V}_j$ ; a box  $\mathcal{Q}_j$  is called 1-good if  $|\mathcal{C}_{k,j}^{(1)}[1, A'k^\gamma]| \geq \rho k$ ; the stage-2- $\mathcal{G}_n$  is called 2-good if all subboxes  $\mathcal{Q}_1, \dots, \mathcal{Q}_{n/k}$  are 1-good and additionally there are box wedging vertices between each consecutive box:

$$|\{u \in \mathcal{C}_{k,j}^{(1)}[1, A'k^\gamma] : u \sim \mathcal{C}_{k,j+1}^{(1)}[1, A'k^\gamma]\}| \geq 1 \text{ for all } i < n/k.$$

By calculations similar to those in the proof of Proposition 6.1 when  $\zeta_{\text{II}} > 0$  on page 55, one can show that there exists  $\varepsilon_2 > 0$  such that for all  $n$  sufficiently large

$$\mathbb{P}(\mathcal{G}_n \text{ is 2-good}) \geq 1 - n \exp(-\varepsilon_2 k^{\zeta - \delta}) \geq 1 - o(n^{-C}),$$

substituting  $k = (A_1 \log n)^{1/(\zeta - \delta)}$  for any sufficiently large constant  $A_1$ . Now (5.35) follows as in (B25).

We turn to (5.36), for which we first prove that vertices of mark above  $2A'k^{\eta_{\text{II}}}$  have not been revealed yet in stage-2- $\mathcal{G}_n$ . For this we need that  $\eta_{\text{II}} \geq \gamma_{\text{II}}$ . When  $\zeta_{\text{II}} > \max(\zeta_{\text{hl}}, 0)$ , it follows that  $\alpha < \tau - 1$  and  $\alpha \in (1, 2)$  by definition of  $\zeta_{\text{II}}$  and  $\zeta_{\text{hl}}$  in (2.9) and (2.12), respectively. For such values of  $\alpha$ ,  $\alpha(\alpha - 1) < \tau - 1$  holds, which implies that  $\eta_{\text{II}} = 1/\alpha > (\alpha - 1)/(\tau - 1) = \gamma_{\text{II}}$ . Thus, no vertices of mark at least  $2A'k^{\eta_{\text{II}}}$  have been exposed in the stage-2- $\mathcal{G}_n$ . Similar to (B26) and (B27), we obtain by a union bound over all vertices of mark at least  $2A'k^{\eta_{\text{II}}}$ ,

$$\begin{aligned} & \mathbb{P}(\nexists \mathcal{C} : \mathcal{C} \not\subseteq \mathcal{V}_n[2A'k^{\eta_{\text{II}}}, \infty), |\mathcal{C}| \leq \rho n) \\ & \geq \mathbb{E} \left[ |\mathcal{V}_n[2A'k^{\eta_{\text{II}}}, \infty)| \left( 1 - p \left( 1 \wedge \beta \frac{2A'k^\eta}{d^{d/2k}} \right)^\alpha \right)^{\rho k} \right] + o(n^{-C}), \\ & \geq n \exp \left( -p\beta^\alpha d^{-\alpha d/2} \rho (2A')^\alpha k^\zeta \right) + o(n^{-C}), \end{aligned}$$

Here we used that  $\eta = 1/\alpha$  implies that the minimum is again attained at the second term, and we then used that  $\zeta = 2 - \alpha$ . Substituting  $k = (A_1 \log n)^{1/(\zeta - \delta)}$  yields that when  $A_1$  is large

$$\mathbb{P}(\mathcal{G}_n \text{ is 2-good}) \geq 1 - n \exp(-\varepsilon_2 k^{\zeta - \delta}) \geq 1 - o(n^{-C}),$$

showing (5.36) for some  $A > 0$ . The corresponding statements for the Palm version  $\mathbb{P}^x$  of  $\mathbb{P}$  for any  $x \in \mathbb{R}^d$ , are proven in exactly the same way.  $\square$

## Appendix C Postponed proofs: upper tail

### C.2 Concentration of degrees and crossing edges

**Proof of Claim 7.3** We write  $\underline{w} = \underline{w}_n$ . To bound (7.13), we first distinguish whether there are more or fewer than  $\lceil C/\psi \rceil$  such vertices, and in the latter case we condition on their number:

$$\begin{aligned} & \mathbb{P}\left(\sum_{v \in \mathcal{V}_n[\underline{w}, \phi n]} \deg_v[1, \phi n] \geq \psi n\right) \\ & \leq \mathbb{P}(|\mathcal{V}_n[\underline{w}, \phi n]| > \lceil C/\psi \rceil) \\ & \quad + \mathbb{E}\left[\mathbb{1}_{\{|\mathcal{V}_n[\underline{w}, \phi n]| \leq \lceil C/\psi \rceil\}} \mathbb{P}\left(\sum_{v \in \mathcal{V}_n[\underline{w}, \phi n]} \deg_v[1, \phi n] \geq \psi n \mid |\mathcal{V}_n[\underline{w}, \phi n]| \right)\right]. \end{aligned}$$

For the first probability, the intensity of  $\mathcal{V}$  in (4.5) yields that  $|\mathcal{V}_n[\underline{w}, \phi n]| \preceq \text{Poi}(n\underline{w}^{-(\tau-1)}) = \text{Poi}(n^{-\psi})$ . Hence,  $\mathbb{P}(|\mathcal{V}_n[\underline{w}, \phi n]| > \lceil C/\psi \rceil) = \Theta(n^{-\psi(\lceil C/\psi \rceil + 1)}) = o(n^{-C})$ . We turn to the term inside the expectation. For the sum  $\sum_{v \in \mathcal{V}_n[\underline{w}, \phi n]} \deg_v[1, \phi n]$  to be at least  $\psi n$ , by the pigeon-hole principle, there must be a vertex among the at most  $\lceil C/\psi \rceil$  vertices in  $\mathcal{V}_n[\underline{w}, \phi n]$  that has total degree at least  $\psi n / \lceil C/\psi \rceil \geq 2n\psi^2/(3C)$  (we may assume  $C$  large so that this inequality holds). Since at most  $\lceil C/\psi \rceil$  edges go to other vertices in  $\mathcal{V}_n[\underline{w}, \phi n]$ , this vertex must have at least  $2n\psi^2/(3C) - \lceil C/\psi \rceil \geq n\psi^2/(2C)$  edges towards vertices with mark below  $\underline{w}$  (the bound holds for  $n$  large). We take a union bound over these high-mark vertices, and consider the worst-case location  $y$  of such a vertex to estimate the second term above:

$$\begin{aligned} & \mathbb{P}\left(\sum_{v \in \mathcal{V}_n[\underline{w}, \phi n]} \deg_v[1, \phi n] \geq \psi n\right) \leq o(n^{-C}) \\ & \quad + \lceil \frac{C}{\psi} \rceil \sup_{(y, w) \in \Lambda_n \times [\underline{w}, \phi n]} \mathbb{P}^y\left(\deg_{(y, w)}[1, \underline{w}] \geq \frac{\psi^2}{2C} n \mid (y, w) \in \mathcal{V}_n[\underline{w}, \phi n]\right). \end{aligned} \tag{C1}$$

Since  $\mathcal{V}_n[1, \underline{w}]$  is independent of  $\mathcal{V}_n[\underline{w}, \infty)$ , the PPP restricted to vertices of mark at most  $\underline{w}$  is independent of the location  $y$  of a high-mark vertex  $v = (y, w)$ . Since edges are present conditionally independently, conditionally on  $v = (y, w)$ , the neighbors of  $v$  in the graph form an inhomogeneous PPP on  $\Lambda_n \times [1, \underline{w}] \subseteq \mathbb{R}^{d+1}$  with intensity  $p((y, w), (x, z))(\mathrm{d}x \times F_W(\mathrm{d}z))$ . By integrating this intensity over  $\Lambda_n \times [1, \underline{w}]$  (see [61, Proposition 2.1] for an example) the reader may verify that there exists a constant  $c = c(d, \alpha, \tau, \sigma, \beta, p) > 0$  such that for any vertex  $v = (y, w)$  with location  $y \in \Lambda_n$  and  $w \geq \underline{w}$

$$\deg_{(y, w)}[1, \underline{w}] \preceq \begin{cases} \text{Poi}(cw), & \text{if } \sigma < \tau - 1, \\ \text{Poi}(cw \log(n/w)), & \text{if } \sigma = \tau - 1, \\ \text{Poi}(cw^{(\tau-1)/\sigma} n^{1-(\tau-1)/\sigma}), & \text{if } \sigma > \tau - 1. \end{cases}$$

The intensity is increasing in  $w$  in all three cases, so the supremum in (C1) is at  $w = \phi n$ . Substituting this value, the parameter of the Poisson distribution is order  $O(n)$  in all three cases with constant prefactor at most  $c_\phi := c \max(\phi, \phi^{(\tau-1)/\sigma}, \phi \log(1/\phi))$ . Clearly  $c_\phi$  tends to zero as  $\phi \downarrow 0$ . We conclude that

$$\mathbb{P}\left(\sum_{v \in \mathcal{V}_n[\underline{w}, \phi n]} \deg_v[1, \phi n] \geq \psi n\right) \leq o(n^{-C}) + \lceil \frac{C}{\psi} \rceil \mathbb{P}\left(\text{Poi}(nc_\phi) \geq n \frac{\psi^2}{2C}\right).$$



By concentration for Poisson random variables (Lemma A.1), the second term decays exponentially in  $n$  for any  $\phi$  sufficiently small, finishing the proof.  $\square$

**Proof of Claim 7.4** We first condition on the set of vertices with mark at least  $\underline{w}$  and apply a union bound over these vertices:

$$\mathbb{P}(\exists v \in \mathcal{V}_n[\underline{w}, \infty) : \deg_v[1, \underline{w}] < \ell) \leq \mathbb{E} \left[ \sum_{v \in \mathcal{V}_n[\underline{w}, \infty)} \mathbb{P}(\deg_v[1, \underline{w}] < \ell \mid \mathcal{V}_n[\underline{w}, \infty)) \right].$$

We analyze a single summand: since the PPP  $\mathcal{V}_n[1, \underline{w}]$  is independent of  $\mathcal{V}_n[\underline{w}, \infty)$ , and edges are present conditionally independently, it follows that

$$\mathbb{P}(\deg_v[1, \underline{w}] < \ell \mid \mathcal{V}_n[\underline{w}, \infty)) \leq \sup_{(y, w) \in \Lambda_n \times [\underline{w}, \infty)} \mathbb{P}(\deg_v[1, \underline{w}] < \ell \mid v = (y, w) \in \mathcal{V}). \quad (\text{C2})$$

The neighbors of the vertex  $v$  form an independent inhomogeneous PPP on  $\Lambda_n \times [1, \underline{w}]$  with intensity  $p((y, w), (x, z))(dx \times F_W(dz))$ . By integrating the intensity over  $\Lambda_n \times [1, \underline{w}]$  (see [61, Proposition 2.1] for an example) there exists a constant  $c = c(d, \alpha, \tau, \sigma, \beta, p) > 0$  such that  $\deg_v[1, \underline{w}] \succcurlyeq \text{Poi}(c(w \wedge n))$  for any  $y \in \Lambda_n$ . This is also true when  $y$  is close to the boundary of  $\Lambda_n$ , since at least a  $d$ -dependent constant fraction of the ball of volume  $w \wedge n$  centered at any  $y$  falls inside  $\Lambda_n$ . Since  $w \geq \underline{w}$ , the supremum in (C2) is attained at  $w = \underline{w}$ , so we obtain that

$$\begin{aligned} \mathbb{P}(\exists v \in \mathcal{V}_n[\underline{w}, \infty) : \deg_v[1, \underline{w}] < \ell) &\leq \mathbb{E}[|\mathcal{V}_n[\underline{w}, \infty)|] \mathbb{P}(\text{Poi}(c\underline{w}) < \ell) \\ &\leq n \mathbb{P}(\text{Poi}(c\underline{w}) < \ell). \end{aligned}$$

By concentration inequalities for Poisson random variables (see Lemma A.1), the probability on the right-hand side is of order  $o(n^{-C-1})$  when  $\underline{w} \geq C_1 \log n$ , where  $C_1 = C_1(C, c)$  is a sufficiently large constant.  $\square$

**Proof of Claim 7.5** We distinguish three types of vertices inside  $\Lambda_n$  that may be connected by an edge to a vertex outside  $\Lambda_n$ : (1) those close to the boundary of  $\Lambda_n$ , (2) those with mark at least a very large constant not close to the boundary, and (3) the rest. By the pigeon-hole principle, in order for the event inside of the probability of the statement (7.14) to hold, at least one of these three disjoint sets must have size at least  $(\psi/3)n$ . We write  $\tilde{n} = (1 - \psi/4)n$  and  $A_0 := \Lambda_n \setminus \Lambda_{\tilde{n}}$  for an annulus inside  $\Lambda_n$ . Formally, the three cases become

$$\begin{aligned} \mathbb{P}(|\{v \in \mathcal{V}_n[1, \underline{w}) : v \sim \Lambda_n^c \times [1, \underline{w})\}| \geq \psi n) \\ \leq \mathbb{P}(|\mathcal{V}_{A_0}| \geq (\psi/3)n) + \mathbb{P}(|\mathcal{V}_{\tilde{n}}[(4/\psi)^{1/(\tau-1)}, \infty)| \geq (\psi/3)n) \quad (\text{C3}) \\ + \mathbb{P}(|\{v \in \mathcal{V}_{\tilde{n}}[1, (4/\psi)^{1/(\tau-1)}) : v \sim \Lambda_n^c \times [1, \underline{w})\}| \geq (\psi/3)n). \end{aligned}$$

The size of the annulus in the first term and the mark-truncation value  $(4/\psi)^{1/(\tau-1)}$  in the second term are both chosen so that  $\mathbb{E}[|\mathcal{V}_{A_0}|] = \mathbb{E}[|\mathcal{V}_{\tilde{n}}[(4/\psi)^{1/(\tau-1)}, \infty)|] = (\psi/4)n$ . By concentration inequalities for Poisson variables (Lemma A.1), the first

two terms on the right-hand side decay exponentially in  $n$  and are thus  $o(n^{-C})$ . In the remainder, we focus on the third term.

When  $\alpha = \infty$ , the third term equals zero for  $n$  sufficiently large: by the connection probability  $p$  in (4.6), a vertex with mark at most  $(4/\psi)^{1/(\tau-1)}$  cannot be connected by an edge to a vertex of mark at most  $\underline{w} = o(n)$  at distance  $\Omega(n^{1/d})$ .

We now study the third term in (C3) when  $\alpha < \infty$ . We use a discretization argument: we decompose  $\Lambda_n^c \times [1, \underline{w})$  into space-mark annuli and count edges according to where their endpoint falls. Set for  $i \geq 1$  and  $j \leq \lceil \log_2(\underline{w}) \rceil$ ,

$$A_{i,j} := (\Lambda_{2^i n} \setminus \Lambda_{2^{i-1} n}) \times [2^j, 2^{j+1}),$$

$$\mathbb{E}[\mathcal{V} \cap A_{i,j}] = 2^{i-1} n \cdot (2^{-j(\tau-1)} - 2^{-(j+1)(\tau-1)}) \leq 2^{i-j(\tau-1)-1} n. \quad (\text{C4})$$

To avoid too many crossing edges, a good event is when none of the sets  $A_{i,j}$  contains more than  $2(\log n)^2$  times the expectation many vertices. Truncating also at 1, we define this event as

$$\mathcal{A}_{\text{out}} := \bigcap_{i \geq 1} \bigcap_{j \leq \lceil \log_2(\underline{w}) \rceil} A_{i,j}, \quad \text{with} \quad \mathcal{A}_{i,j} := \left\{ |\mathcal{V} \cap A_{i,j}| \leq (\log n)^2 \max(2^{i-j(\tau-1)} n, 1) \right\}. \quad (\text{C5})$$

Let us introduce the event that the number of vertices inside  $\Lambda_{\bar{n}} = \Lambda_{(1-\psi/4)n}$  does not exceed  $n$ :

$$\mathcal{A}_{\text{in}} := \{ |\mathcal{V}_{\bar{n}}[1, (4/\psi)^{1/(\tau-1)}]| \leq n \} \quad \text{and} \quad \mathcal{A} = \mathcal{A}_{\text{in}} \cap \mathcal{A}_{\text{out}}. \quad (\text{C6})$$

We first use concentration inequalities for Poisson random variables to show that  $\mathbb{P}(\neg \mathcal{A})$  decays superpolynomially. To make another case distinction below, we let

$$j_*(i) := \max \{ j \in [\lceil \log_2 \underline{w} \rceil] : 2^{i-j(\tau-1)-1} n \geq (\log n)^2 \},$$

$$C_1 := \inf_{C' \geq 2\psi} \left\{ \forall i \geq C_1 \log_2 n : j_*(i) = \lceil \log_2 \underline{w} \rceil \right\}. \quad (\text{C7})$$

The number of vertices in  $A_{i,j}$  is Poisson distributed with expectation at most  $2^{i-j(\tau-1)-1} n$ , see (C4). Therefore, by a union bound over all space-mark annuli,

$$\begin{aligned} \mathbb{P}(\neg \mathcal{A}_{\text{out}}) &\leq \sum_{i \geq 1} \sum_{j=1}^{j_*(i)} \mathbb{P} \left( \text{Poi}(2^{i-j(\tau-1)-1} n) \geq (\log n)^2 2^{i-j(\tau-1)} n \right) \\ &\quad + \sum_{i=1}^{C_1 \log_2 n} \sum_{j=j_*(i)+1}^{\lceil \log_2 \underline{w} \rceil} \mathbb{P} \left( \text{Poi}(2^{i-j(\tau-1)-1} n) \geq (\log n)^2 \max(2^{i-j(\tau-1)} n, 1) \right). \end{aligned}$$

We use that the function  $f(\lambda) := \mathbb{P}(\text{Poi}(\lambda) > 2(\log n)^2(\lambda \vee 1))$  is increasing on the interval  $[0, 1]$ , maximal at  $\lambda = 1$ , and then decreasing on  $[1, \infty)$ . So, the summands on the second line can be all bounded from above by  $\mathbb{P}(\text{Poi}(1) \geq (\log n)^2) = O(1/(\log n)^2!)$ . Since  $\underline{w} = n^{(1+\psi)/(\tau-1)}$ , there are  $O((\log n)^2)$  summands, so the second row decays faster than any polynomial. Turning to the first

row, since the means here are all above 1, for each fixed  $i$ , the largest summand is  $j = j_*(i) \leq \lceil \log_2 \underline{w} \rceil \leq \lceil \log_2 n \rceil$ . We then employ concentration for Poisson random variables (Lemma A.1) and obtain that there exists  $c > 0$  such that

$$\mathbb{P}(\neg \mathcal{A}_{\text{out}}) \leq o(n^{-C}) + \sum_{i \geq 1} \left( \lceil \log_2 n \rceil \exp(-c 2^{i-j_*(i)(\tau-1)} n) \right).$$

If  $i \geq \lceil C_1 \log_2 n \rceil$ , then  $j_*(i) = \lceil \log_2 w \rceil$  by definition in (C7). For smaller  $i$ , again by definition of  $j_*(i)$ , the argument of the exponential is always at least  $c(\log n)^2$ . Therefore,

$$\begin{aligned} \mathbb{P}(\neg \mathcal{A}_{\text{out}}) &\leq o(n^{-C}) + C_1 \lceil \log_2 n \rceil^2 \exp(-c(\log n)^2) \\ &\quad + \lceil \log_2 n \rceil \sum_{i \geq \lceil C_1 \log_2 n \rceil} \exp(-c 2^{i-\lceil \log_2 \underline{w} \rceil(\tau-1)} n). \end{aligned}$$

The second term decays superpolynomially and is absorbed by the  $o(n^{-C})$  term. Since  $\underline{w} \leq n^{(1+\psi)/(\tau-1)}$  with  $(1+\psi)/(\tau-1) < 1$ , we obtain for a constant  $c' > 0$  that

$$\mathbb{P}(\neg \mathcal{A}_{\text{out}}) \leq o(n^{-C}) + \lceil \log_2 n \rceil \sum_{i \geq \lceil C_1 \log_2 n \rceil} \exp(-c' 2^i n^{-\psi}).$$

Since  $C_1 > 2\psi$  by definition in (C7), it follows that  $2^i \geq n^{2\psi}$  for all  $i$ . Therefore, the sum decays faster than any polynomial. We recall that  $\mathcal{A} = \mathcal{A}_{\text{in}} \cap \mathcal{A}_{\text{out}}$  by (C6). By Lemma A.1, also  $\mathbb{P}(\neg \mathcal{A}_{\text{in}}) = \exp(-\Omega(n))$ . This proves that

$$\mathbb{P}(\neg \mathcal{A}) = o(n^{-C}). \quad (\text{C8})$$

We return to the third term in (C3) and show that it decays exponentially in  $n$  when we condition on any realization of the vertex set  $\mathcal{V}$  that satisfies  $\mathcal{A}$ . Our first goal is to show that the probability that any fixed vertex  $v \in \mathcal{V}_{\bar{n}}[1, (4/\psi)^{1/(\tau-1)}]$  is connected by an edge to a vertex outside  $\Lambda_n$  tends to zero, conditionally on  $\mathcal{V}$  that satisfies  $\mathcal{A}$ . Using the upper bounds on  $|\mathcal{V} \cap A_{i,j}|$  in (C5) and the connection probability  $p$  from (4.6), Markov's inequality yields for any  $v \in \Lambda_{\bar{n}} \times [1, (4/\psi)^{1/(\tau-1)}]$

$$\begin{aligned} &\mathbb{P}(\exists u \in \Lambda_n^c \times [1, \underline{w}) : v \sim u \mid \mathcal{V}, \mathcal{A}) \\ &\leq p \beta^\alpha (\log n)^2 \sum_{i \geq 1} \sum_{j \in [\lceil \log_2 \underline{w} \rceil]} (2^{i-j(\tau-1)} n \vee 1) \left( \frac{\kappa_\sigma((C/\psi)^{1/(\tau-1)}, 2^{j+1})}{\|\Lambda_{\bar{n}} - A_{i,j}\|^d} \right)^\alpha. \end{aligned}$$

By the definition of  $\kappa_\sigma$  in (2.6) the numerator is at most  $c_1 2^{j\alpha}$  for some constant  $c_1 = c_1(C, \psi, \sigma, \alpha)$ . There exists a constant  $c_2 > 0$  such that  $\|\Lambda_{\bar{n}} - A_{i,j}\| \geq (c_2 2^i n)^{1/d}$  for all  $i \geq 1$ , and hence there exists another constant  $c > 0$  such that

$$\begin{aligned} \mathbb{P}(\exists u \in \Lambda_n^c \times [1, \underline{w}) : v \sim u \mid \mathcal{V}, \mathcal{A}) \\ \leq c(\log n)^2 n^{-\alpha} \sum_{i \geq 1} \sum_{j \in [\lceil \log_2 \underline{w} \rceil]} (2^{i-j(\tau-1)} n \vee 1) 2^{(j-i)\alpha}. \end{aligned}$$

Studying the maximum inside the summation,  $2^{i-j(\tau-1)} n < 1$  if  $j \in [(i + \log_2 n)/(\tau - 1), \lceil \log_2 \underline{w} \rceil]$ . Since  $\log_2 \underline{w} \leq \log_2 n$ , this interval is non-empty only when  $i < C_2 \log_2 n$  for some constant  $C_2 > 0$ . Using now that  $x \vee 1 \leq x + \mathbb{1}_{\{x \leq 1\}}$  with these criteria, we obtain the upper bound

$$\begin{aligned} \mathbb{P}(\exists u \in \Lambda_n^c \times [1, \underline{w}) : v \sim u \mid \mathcal{V}, \mathcal{A}) \\ \leq c(\log n)^2 n^{1-\alpha} \sum_{i \geq 1} \sum_{j \in [\lceil \log_2 \underline{w} \rceil]} 2^{-i(\alpha-1)+j(\alpha-(\tau-1))} \\ + c(\log n)^2 n^{-\alpha} \sum_{i=1}^{\lceil C_2 \log_2 n \rceil} \sum_{j=\lceil (i+\log_2 n)/(\tau-1) \rceil}^{\lceil \log_2 \underline{w} \rceil} 2^{(j-i)\alpha}. \quad (\text{C9}) \end{aligned}$$

We continue to bound the second line. Recall that  $\underline{w} \leq n^{(1+\psi)/(\tau-1)}$  where  $(1 + \psi)/(\tau - 1) < 1$ . Therefore, there exists  $\varepsilon > 0$  such that  $2^j < n^{1-\varepsilon}$  for all  $j \leq \lceil \log_2 \underline{w} \rceil$  when  $n$  is sufficiently large. Furthermore, we bound  $2^{-i\alpha} \leq 1$ . Consequently, there exists  $C_3 > 0$  such that

$$\begin{aligned} \mathbb{P}(\exists u \in \Lambda_n^c \times [1, \underline{w}) : v \sim u \mid \mathcal{V}, \mathcal{A}) \\ \leq c(\log n)^2 n^{1-\alpha} \sum_{i \geq 1} 2^{-i(\alpha-1)} \sum_{j \in [\lceil \log_2 \underline{w} \rceil]} 2^{j(\alpha-(\tau-1))} + C_3(\log n)^4 n^{-\varepsilon\alpha}. \end{aligned}$$

We now bound the first line in (C9) on the right-hand side. Since  $\alpha > 1$ , the sum over  $i$  is finite. The sum over  $j$  is finite when  $\alpha < \tau - 1$ , and at most logarithmic when  $\alpha = \tau - 1$ . Since  $\alpha > 1$ , in both of these cases the right-hand side on the first line is at most  $O(n^{-\tilde{\varepsilon}})$  for some  $\tilde{\varepsilon} > 0$ . When  $\alpha > \tau - 1$ , we use again that  $2^j < n^{1-\varepsilon}$ , so the maximal term is at most  $n^{(1-\varepsilon)(\alpha-(\tau-1))}$ . When multiplied with  $n^{1-\alpha}$ , this is at most  $n^{-(\tau-2)-\varepsilon(\alpha-(\tau-1))}$ . We conclude that in all cases there exist  $\tilde{\varepsilon}, C_4 > 0$  such that

$$\mathbb{P}(\exists u \in \Lambda_n^c [1, \underline{w}) : v \sim u \mid \mathcal{V}, \mathcal{A}) \leq C_4(\log n)^3 n^{-\tilde{\varepsilon}} + C_2(\log n)^4 n^{-\varepsilon\alpha}.$$

Assuming  $n$  is sufficiently large, the right-hand side is at most  $\psi/6$ . By definition of  $\mathcal{A}_{\text{in}}$  in (C6) there are at most  $n$  vertices inside  $\mathcal{V}_{\tilde{n}}[1, (4/\psi)^{1/(\tau-1)})$ . Recall that edges are present conditionally independently on  $\mathcal{V}$ . Therefore, conditionally on any realization of  $\mathcal{V}$  that satisfies  $\mathcal{A}$ ,

$$|\{v \in \mathcal{V}_{\tilde{n}}[1, (4/\psi)^{1/(\tau-1)}) : v \sim \Lambda_n^c \times [1, \underline{w})\}| \succcurlyeq \text{Bin}(n, \psi/6).$$

Returning to the third term in (C3), we obtain that

$$\begin{aligned} \mathbb{P}(|\{v \in \mathcal{V}_{\tilde{n}}[1, (4/\psi)^{1/(\tau-1)}) : v \sim \Lambda_n^c \times [1, \underline{w})\}| \geq (\psi/3)n) \\ \leq \mathbb{P}(\neg \mathcal{A}) + \mathbb{E}[\mathbb{1}_{\{\mathcal{A}\}} \mathbb{P}(\text{Bin}(n, \psi/6) \geq (\psi/3)n \mid \mathcal{V}, \mathcal{A})]. \end{aligned}$$

The first term is of order  $o(n^{-C})$  by (C8). The second term decays exponentially in  $n$  by a Chernoff bound. Thus all terms in (C3) are of order  $o(n^{-C})$ , finishing the proof.  $\square$

### C.3 Truncation of the generating function

**Proof of Claim 7.7** We start with the first bound. Using elementary rearrangements of (7.26), we leave it to the reader to verify that it is sufficient that there exists  $\psi' > 0$  and  $\ell_0$  such that

$$\rho - \theta + \psi' < \sum_{\ell=1}^{\ell_0} (1 - (1-p)^{h_{10}\ell})\theta_\ell = \sum_{\ell=1}^{\ell_0} \theta_\ell - \sum_{\ell=1}^{\ell_0} (1-p)^{h_{10}\ell}\theta_\ell.$$

Since  $\sum_{\ell=1}^{\infty} \theta_\ell = 1 - \theta$ , we may choose  $\ell_0$  is sufficiently large (depending on  $\psi'$ ) so that  $\sum_{\ell=1}^{\ell_0} \theta_\ell = \mathbb{P}^0(|\mathcal{C}(0)| \leq \ell_0) > 1 - \theta - \psi'$ . Rearranging both sides, it is sufficient to show that we can choose  $\psi'$  satisfying

$$2\psi' < 1 - \rho - \sum_{\ell=1}^{\ell_0} (1-p)^{h_{10}\ell}\theta_\ell. \quad (\text{C10})$$

If  $p = 1$ , then the sum is 0 and the bound is satisfied for any  $\psi' \in (0, (1 - \rho)/2)$ . Now we assume that  $p < 1$ , so that  $\text{hubs}(\rho)$ , defined in (2.17), satisfies the equation

$$\mathbb{E}^0[(1-p)^{\text{hubs}(\rho)|\mathcal{C}(0)|}] = \sum_{\ell=1}^{\infty} (1-p)^{\text{hubs}(\rho)\ell}\theta_\ell = 1 - \rho. \quad (\text{C11})$$

Now the definition of  $h_{10}(\rho)$  matters: If  $\text{hubs}(\rho)$  is an integer, then  $h_{10}(\rho) = \lceil \text{hubs}(\rho) \rceil + 1 = \text{hubs}(\rho) + 1$  by definition of  $h_{10}$  in (7.25). If  $\text{hubs}(\rho)$  is not an integer, then  $h_{10}(\rho) = \lceil \text{hubs}(\rho) \rceil$ , and in both cases we have  $h_{10}(\rho) > \text{hubs}(\rho)$ . So, using the infinite sum on the right-hand side of (C10) in place of  $1 - \rho$  gives that (C10) holds if

$$2\psi' < \sum_{\ell=1}^{\ell_0} ((1-p)^{\text{hubs}(\rho)\ell} - (1-p)^{h_{10}\ell}).$$

Since  $\text{hubs}(\rho) < h_{10}$ , each summand is nonnegative, so the above bound is satisfied if  $\psi'$  is chosen as, for instance,  $\psi' := ((1-p)^{\text{hubs}(\rho)} - (1-p)^{h_{10}})/2 > 0$ .

We turn to the proof of (7.27) for which we first assume  $p < 1$ . Using elementary rearrangements, it suffices to find for an  $\psi' > 0$  sufficiently small, and an  $\ell_0$  large, so that for all  $\ell^* \geq \ell_0$ ,

$$\sum_{\ell=1}^{\ell^*} \theta_\ell (1-p)^{(h_{\text{up}}-1)\ell} > 1 - \rho + \psi'. \quad (\text{C12})$$

Since we assumed  $p < 1$ , the equality (C11) is satisfied. Because  $h_{\text{up}} = \lceil \text{hubs}(\rho) \rceil$ , it holds that  $h_{\text{up}} - 1 < \text{hubs}(\rho)$ . Since the left-hand side in (C11) is decreasing in  $\text{hubs}(\rho)$ , it follows that

$$\sum_{\ell=1}^{\infty} (1-p)^{(h_{\text{up}}-1)\ell} \theta_{\ell} > 1 - \rho. \quad (\text{C13})$$

Thus, there exists  $\delta > 0$ , and  $\ell_0 > 0$  such that for all  $\ell^* \geq \ell_0$

$$\sum_{\ell=1}^{\ell^*} (1-p)^{(h_{\text{up}}-1)\ell} \theta_{\ell} > 1 - \rho + \delta.$$

We combine this bound with (C12), and choose  $\psi' = \delta$ . This finishes the proof of (7.27) when  $p < 1$ . When  $p = 1$ ,  $h_{\text{up}} = 1$  by its own definition in (7.25) and the definition of  $\text{hubs}(\rho) = 1$  in (2.17). Hence,  $(1-p)^{(h_{\text{up}}-1)\ell} = 1$  for all  $\ell \geq 1$ . The proof follows as before since  $\sum_{\ell=1}^{\infty} \theta_{\ell} = 1 - \theta$ , choose  $\ell_0$  is sufficiently large (depending on  $\psi'$ ) so that  $\sum_{\ell=1}^{\ell_0} \theta_{\ell} = \mathbb{P}^0(|\mathcal{C}(0)| \leq \ell_0) > 1 - \theta - \psi' > 1 - \rho$ .  $\square$

**Acknowledgements** The work of JJ and JK has been partly supported through grant NWO 613.009.122. The work of DM has been partially supported by grant Fondecyt grant 1220174 and by grant GrHyDy ANR-20-CE40-0002.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Abbey, H.: An examination of the Reed-Frost theory of epidemics. *Hum. Biol.* **24**(3), 201 (1952)
2. Aizenman, M., Delyon, F., Souillard, B.: Lower bounds on the cluster size distribution. *J. Stat. Phys.* **23**(3), 267–280 (1980)
3. Aizenman, M., Kesten, H., Newman, C.M.: Uniqueness of the infinite cluster and continuity of connectivity functions for short and long range percolation. *Commun. Math. Phys.* **111**(4), 505–531 (1987)
4. Aizenman, M., Newman, C.M.: Discontinuity of the percolation density in one dimensional  $1/|x-y|^2$  percolation models. *Commun. Math. Phys.* **107**(4), 611–647 (1986)
5. Albert, R., Barabási, A.-L.: Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74**(1), 47 (2002)
6. Alexander, K., Chayes, J., Chayes, L.: The Wulff construction and asymptotics of the finite cluster distribution for two-dimensional Bernoulli percolation. *Commun. Math. Phys.* **131**(1), 1–50 (1990)
7. Andjel, E.D.: Characteristic exponents for two-dimensional bootstrap percolation. *Ann. Probab.* **21**(2), 926–935 (1993)
8. Andreis, L., König, W., Langhammer, H., Patterson, R.I.: A large-deviations principle for all the components in a sparse inhomogeneous random graph. *Prob. Theory Related Fields* 1–100 (2023)
9. Andreis, L., König, W., Patterson, R.I.: A large-deviations principle for all the cluster sizes of a sparse Erdős-Rényi graph. *Random Struct. Algorithms* **59**(4), 522–553 (2021)

10. Barbour, A.D., Utev, S.: Approximating the Reed-Frost epidemic process. *Stoch. Proc. Appl.* **113**(2), 173–197 (2004)
11. Bäumlér, J.: Continuity of the critical value and a shape theorem for long-range percolation. (2023). Preprint [arXiv:2312.04099](https://arxiv.org/abs/2312.04099)
12. Berger, N.: Transience, recurrence and critical behavior for long-range percolation. *Commun. Math. Phys.* **226**(3), 531–558 (2002)
13. Bhamidi, S., Budhiraja, A., Dupuis, P., Wu, R.: Rare event asymptotics for exploration processes for random graphs. *Ann. Appl. Prob.* **32**(2), 1112–1178 (2022)
14. Biskup, M.: On the scaling of the chemical distance in long-range percolation models. *Ann. Probab.* **32**(4), 2938–2977 (2004)
15. Bläsius, T., Friedrich, T., Katzmann, M., Ruff, J., Zeif, Z.: On the giant component of geometric inhomogeneous random graphs. *Eur. Symp. Algorithms (ESA)*, 20:1–20:13 (2023)
16. Bringmann, K., Keusch, R., Lengler, J.: Geometric inhomogeneous random graphs. *Theoret. Comput. Sci.* **760**, 35–54 (2019)
17. Broadbent, S.R., Hammersley, J.M.: Percolation processes: I. crystals and mazes. *Math. Proc. Cambridge Philos. Soc.* **53**(3), 629–641 (1957)
18. Burton, R.M., Keane, M.: Density and uniqueness in percolation. *Commun. Math. Phys.* **121**, 501–505 (1989)
19. Campa, A., Dauxois, T., Ruffo, S.: Statistical mechanics and dynamics of solvable models with long-range interactions. *Phys. Rep.* **480**(3–6), 57–159 (2009)
20. Carvalho, T.T., Fontenele, A.J., Girardi-Schappo, M., Feliciano, T., Aguiar, L.A., Silva, T.P., De Vasconcelos, N.A., Carelli, P.V., Copelli, M.: Subsampled directed-percolation models explain scaling relations experimentally observed in the brain. *Frontiers Neural Circuits* **14**, 576727 (2021)
21. Cerf, R.: Large deviations for three dimensional supercritical percolation. *Société mathématique de France*, (2000)
22. Chung, F., Lu, L.: Connected components in random graphs with given expected degree sequences. *Ann. Comb.* **6**(2), 125–145 (2002)
23. Contreras, D., Martineau, S., Tassion, V.: Supercritical percolation on graphs of polynomial growth. *Duke Math. J.* **173**(4), 745–806 (2024)
24. Crawford, N., Sly, A.: Simple random walk on long range percolation clusters I: heat kernel bounds. *Probab. Theory Relat. Fields* **154**(3–4), 753–786 (2012)
25. Dauxois, T., Ruffo, S., Arimondo, E., Wilkens, M.: Dynamics and thermodynamics of systems with long range interactions. *Lecture Notes in Physics Vol. 602*, Springer, (2002)
26. Defenu, N., Donner, T., Macrì, T., Pagano, G., Ruffo, S., Trombettoni, A.: Long-range interacting quantum systems. *Rev. Mod. Phys.* **95**, 035002 (2023)
27. Deijfen, M., van der Hofstad, R., Hooghiemstra, G.: Scale-free percolation. *Ann. Inst. Henri Poincaré Probab. Stat.* **49**(3), 817–838 (2013)
28. Del Ferraro, G., Moreno, A., Min, B., Morone, F., Pérez-Ramírez, Ú., Pérez-Cervera, L., Parra, L.C., Holodny, A., Canals, S., Makse, H.A.: Finding influential nodes for integration in brain networks using optimal percolation theory. *Nat. Commun.* **9**(1), 2274 (2018)
29. Deuschel, J.-D., Pisztora, A.: Surface order large deviations for high-density percolation. *Probab. Theory Relat. Fields* **104**(4), 467–482 (1996)
30. Duminil-Copin, H., Garban, C., Tassion, V.: Long-range models in 1D revisited. *Ann. Inst. Henri Poincaré Probab. Stat.* **60**(1), 232–241 (2024)
31. Durrett, R., Schonmann, R.H.: Large deviations for the contact process and two dimensional percolation. *Probab. Theory Relat. Fields* **77**(4), 583–603 (1988)
32. Fountoulakis, N., Müller, T.: Law of large numbers for the largest component in a hyperbolic model of complex networks. *Ann. Appl. Prob.* **28**(1), 607–650 (2018)
33. Gandolfi, A.: Clustering and uniqueness in mathematical models of percolation phenomena. PhD thesis, Technische Universiteit Delft, (1989)
34. Gilbert, E.N.: Random plane networks. *J. Soc. Ind. Appl. Math.* **9**, 533–553 (1961)
35. Gouéré, J.-B.: Subcritical regimes in the Poisson Boolean model of continuum percolation. *Ann. Probab.* **36**(4), 1209–1220 (2008)
36. Gracar, P., Grauer, A., Lüchtrath, L., Mörters, P.: The age-dependent random connection model. *Queueing Syst.* **93**(3), 309–331 (2019)
37. Gracar, P., Heydenreich, M., Mönch, C., Mörters, P.: Recurrence versus transience for weight-dependent random connection models. *Electron. J. Probab.* **27**, 1–31 (2022)

38. Gracar, P., Lüchtrath, L., Mönch, C.: Finiteness of the percolation threshold for inhomogeneous long-range models in one dimension. (2022). Preprint [arXiv:2203.11966](https://arxiv.org/abs/2203.11966)
39. Gracar, P., Lüchtrath, L., Mörters, P.: Percolation phase transition in weight-dependent random connection models. *Adv. Appl. Probab.* **53**(4), 1090–1114 (2021)
40. Grimmett, G.: *Percolation*. Grundlehren der Mathematischen Wissenschaften, Springer (1999)
41. Grimmett, G.R., Marstrand, J.M.: The supercritical phase of percolation is well behaved. *Proc. R. Soc. Lond. A* **430**(1879), 439–457 (1990)
42. Hafner, R.: The asymptotic distribution of random clumps. *Computing* **10**, 335–351 (1972)
43. Hall, P.: On continuum percolation. *Ann. Prob.* 1250–1266 (1985)
44. Hirsch, C.: From heavy-tailed Boolean models to scale-free Gilbert graphs. *Braz. J. Prob. Stat.* **31**(1), 111–143 (2017)
45. Hofstad, R.v.d.: The giant in random graphs is almost local. (2021). [arXiv:2103.11733](https://arxiv.org/abs/2103.11733)
46. Hutchcroft, T.: Transience and anchored isoperimetric dimension of supercritical percolation clusters. *Electron. J. Probab.* **28**, 1–15 (2023)
47. Jacob, E.: Subcritical annulus crossing in spatial random graphs. (2025). arXiv preprint [arXiv:2411.10333](https://arxiv.org/abs/2411.10333)
48. Jorritsma, J., Komjáthy, J., Mitsche, D.: Cluster-size decay in supercritical kernel-based spatial random graphs. *Ann. Probab.* **53**(4), 1537–1597 (2025)
49. Jorritsma, J., Komjáthy, J., Mitsche, D.: Cluster-size decay in supercritical long-range percolation. *Electron. J. Probab.* **29**, 1–36 (2024)
50. Jorritsma, J., Zwart, B.: Large deviations of the giant component in scale-free inhomogeneous random graphs. (2024). eprint [arXiv:2407.01224](https://arxiv.org/abs/2407.01224)
51. Kesten, H., Zhang, Y.: The probability of a large finite cluster in supercritical Bernoulli percolation. *Ann. Prob.* 537–555 (1990)
52. Komjáthy, J., Lodewijks, B.: Explosion in weighted hyperbolic random graphs and geometric inhomogeneous random graphs. *Stoch. Proc. Appl.* **130**(3), 1309–1367 (2020)
53. Kozma, R., Puljic, M.: Random graph theory and neuropercolation for modeling brain oscillations at criticality. *Curr. Opin. Neurobiol.* **31**, 181–188 (2015)
54. Krioukov, D., Papadopoulos, F., Kitsak, M., Vahdat, A., Boguñá, M.: Hyperbolic geometry of complex networks. *Phys. Rev. E* **82**(3), 036106 (2010)
55. Kunz, H., Souillard, B.: Essential singularity in percolation problems and asymptotic behavior of cluster size distribution. *J. Stat. Phys.* **19**(1), 77–106 (1978)
56. Last, G., Penrose, M.: *Lectures on the Poisson process*, vol. 7. Cambridge University Press, Cambridge (2017)
57. Lebowitz, J., Schonmann, R.H.: Pseudo-free energies and large deviations for non-Gibbsian FKG measures. *Probab. Theory Relat. Fields* **77**(1), 49–64 (1988)
58. Lefèvre, C.: Stochastic epidemic models for SIR infectious diseases: a brief survey of the recent general theory. In *Stochastic Processes in Epidemic Theory: Proceedings of a Conference held in Luminy, France, October 23–29, 1988*, pages 1–12. Springer, (1990)
59. Liggett, T.M., Schonmann, R.H., Stacey, A.M.: Domination by product measures. *Ann. Probab.* **25**(1), 71–95 (1997)
60. Lindvall, T.: On Strassen’s theorem on stochastic domination. *Electron. Commun. Probab.* **4**, 51–59 (1999)
61. Lüchtrath, L.: *Percolation in weight-dependent random connection models*. PhD thesis, Universität zu Köln, (2022)
62. Mitzenmacher, M., Upfal, E.: *Probability and Computing: Randomization and Probabilistic Techniques in Algorithms and Data Analysis*. Cambridge University Press, Cambridge (2017)
63. Molloy, M., Reed, B.: The size of the giant component of a random graph with a given degree sequence. *Comb. Probab. Comput.* **7**(3), 295–305 (1998)
64. Mönch, C.: Inhomogeneous long-range percolation in the weak decay regime. *Probab. Theory Relat. Fields* **189**, 1129–1160 (2024)
65. Nagaev, S.V.: Large deviations of sums of independent random variables. *Ann. Probab.* **7**, 745–789 (1979)
66. Newman, M.E.: Communities, modules and large-scale structure in networks. *Nat. Phys.* **8**(1), 25–31 (2012)
67. Norros, I., Reittu, H.: On a conditionally Poissonian graph process. *Adv. Appl. Probab.* **38**(1), 59–75 (2006)



68. O'Connell, N.: Some large deviation results for sparse random graphs. *Probab. Theory Relat. Fields* **110**, 277–285 (1998)
69. Pisztora, A.: Surface order large deviations for Ising, Potts and percolation models. *Probab. Theory Relat. Fields* **104**, 427–466 (1996)
70. Roch, S.: *Modern Discrete Probability: An Essential Toolkit*. Cambridge Series in Statistical and Probabilistic Mathematics, (2024)
71. Schulman, L.S.: Long range percolation in one dimension. *J. Phys. A: Math. Gen.* **16**(17), L639–L641 (1983)
72. van der Hofstad, R.: *Random Graphs and Complex Networks: Volume 2*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, (2024)
73. van der Hofstad, R., van der Hoorn, P., Maitra, N.: Local limits of spatial inhomogeneous random graphs. *Adv. Appl. Prob.* 1–48 (2023)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.