Distances and component sizes in scale-free random graphs

Distances and component sizes in scale-free random graphs

Joost Jorritsma

Distances and component sizes in scale-free random graphs
Joost Jorritsma

Online version

Distances and component sizes in scale-free random graphs

This work was financially supported by the Netherlands Organisation for Scientific Research (NWO) through grant NWO 613.009.122.

colophon

A catalogue record is available from the
Eindhoven University of Technology Library
ISBN: 978-90-386-5704-2
DOI: 10.6100/mdm9-3191

This thesis was typeset using `classicthesis` developed by André Miede.
Cover design by Simone Golob || `www.sgiv.nl`
Printed by ProefschriftMaken || `www.proefschriftmaken.nl`

Distances and component sizes in scale-free random graphs

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Technische Universiteit Eindhoven, op gezag van de rector magnificus prof.dr.ir. F.P.T. Baaijens, voor een commissie aangewezen door het College voor Promoties, in het openbaar te verdedigen op dinsdag 18 april 2023 om 11.00 uur.

door

Joost Jorritsma

geboren te Terneuzen

Dit proefschrift is goedgekeurd door de promotoren en de samenstelling van de promotiecommissie is als volgt:

Voorzitter:	prof. dr. J.J. Lukkien
Promotor:	prof. dr. N.V. Litvak
Copromotor:	dr. J. Komjáthy
Leden:	dr. R.M. Castro
	prof. dr. P.F.A. Van Mieghem (TU Delft)
	dr. D. Mitsche (Pontificia Universidad Católica de Chile)
	prof. dr. P. Mörters (Universität zu Köln)
	prof. dr. A.P. Zwart

Het onderzoek dat in dit proefschrift wordt beschreven is uitgevoerd in overeenstemming met de TU/e Gedragscode Wetenschapsbeoefening.

ACKNOWLEDGEMENTS

When I started pursuing a PhD degree in September 2018, I was hoping that it would be a life-changing experience; as a (slightly naive) 26-year old lad I was ready to solve purely theoretical problems with fellow mathematicians, and hoping to enjoy 4 years without organisational politics. In some aspects, my expectations couldn't have been more different from reality; instead of working on *theoretical* models for disease spreading, a pandemic outbreak changed the course of my PhD drastically and made parts of the research suddenly much more *applied* than expected. Yet, I have been able to work (4.5 years) with many inspirational people, whose different views and ways of working shaped me as an independent researcher, and who made these intense years with all theirs ups and downs unforgettable.

Júlia, thank you for giving me the opportunity to return from industry and giving all you have to make me a successful researcher. You have been a devoted supervisor, and supported me in every aspect that you could. I appreciate the mode of working together that we developed, being always open and critical towards each other if we didn't agree. We spent hours and hours writing on whiteboards, discovering novel methods and results that form the basis of this thesis. You motivated me to preserve on problems, even though they seemed uninteresting or too hard at first sight, and were always available to give constructive criticism: no matter if I produced an incorrect proof, wanted to jump too quickly to a new project, or prepared some incomprehensible draft or presentation. I value the freedom that you provided to develop and pursue other activities besides my main research, and for introducing me to many of the people that contributed to this thesis. I am sure that we will stay in touch for a very long time to come.

Dieter, starting to work with you cheered me up after a difficult period. I am grateful that you always treated me as a peer and a friend, no matter what: when discussing mathematics from 7AM to 3AM in person or on Zoom/Skype, by introducing me to your colleagues and friends (Marcos and Amitai, it was great meeting and working you!), when going out after a dinner in Lausanne or Santiago, when discussing life, or when calming down others (for example, if you simply want to enjoy a terramoto). I hope that one day I can offer you and Claudia the same hospitality as you did when you welcomed me to Santiago.

Johannes, it was a wonderful experience to visit ETH Zürich for three months. Arriving to Switzerland with a carnaval-related infection was far from ideal, but afterwards you welcomed me to the world of (jokes on) clones of parents that could potentially make it hard for evolutionary algorithms to escape traps. Aiming for the paper deadline of a computer-science conference together with Dirk was a completely different experience from writing a journal article; the teamwork in the weeks leading up to the submission of [156] was incredible. I hope that we soon find the time to write and think about an extended journal version.

I believe that doing research in mathematics is an inherently social activity, and work becomes so much more fun if you have the chance to work with great co-authors: besides Júlia, Dieter, Johannes, and Dirk, I would like to thank Tim, Leslie, and John: it was a pleasure to share the experience to get out of our comfort zone and dive into the epidemiology literature that led to Part iii in this thesis. Dennis, besides being my go-to person for anything related to visualisation, it was great to work together on the first academic paper with multiple Drago authors [55], together with non-Drago's Pratik, Mykola, and Jarke.

Serte, together we have celebrated the best moment of my PhD. Nothing is better than to combine working with a friend with fantasizing about opera singers, joking about Bassie and Adriaan, moving post-its over a whiteboard, drinking a shot of tequila after each decision, all during the organisation of the RandNET workshop. Christina and Remco, Guillem and Marc, Sem, thank you for giving us the freedom to organise this workshop the way we envisioned it. It was incredible to meet so many other mathematicians and to see them having fun, solving open problems, and singing karaoke.

Thanks to all the members of my doctoral committee: Nelly Litvak, Rui Castro, Piet Van Mieghem, Dieter Mitsche, Peter Mörters, and Bert Zwart. I am sorry that I gave you the time-consuming task to read an even longer version of this thesis. Thank you for your valuable comments that particularly improved the presentation of the last part of the thesis. I am looking forward to interesting discussions during the defense. I want to specifically thank Nelly for her courage to step up to be my promotor, and Bert for talking and listening to me when I needed it.

I am also indebted to the other senior and junior members of the SPOR group. Alessandro, Edwin, Jaron, Maria, Noela, Pim, Remco, Sem, and Tim:

thank you for being always approachable, and especially for teaching me about the pros and cons that come with the life of a researcher in academia. Albert, Alexander, Angelos, Benoît, Daniel, Haodong, Ivo, Kay, Lorenzo, Marta, Martijn, Mayank, Mike, Neeladri, Richard, Rik, Rounak, Rowel, Suman, Tim (great MSc thesis!), Viktoria, Wessel, and Zsolt: thank you for playing table tennis, instructing courses together, discussing mathematics, and going out for a beer, glass of champagne, or gin and tonic. The secretaries Chantal, Ellen, Marianne, and Patty were always there for a chat, and have been a great help with administration and practicalities, also before, during, and after the RandNET workshop.

My office mates deserve some extra words. Most of the time I was lucky to share an office with Collin, Mona, and Youri. Together we shared our successes and frustrations, making the office feel like a second home. Even during the months that we all worked from home, we stayed in touch with a daily coffee break. Collin and Youri, I hope we will keep going to pub quizzes, continue our cycling trips, and enjoy more and more restaurants together with Richard. Mona, like Collin and Youri [85, 215], I don't know how you survived with all our stupid jokes, but it is good to see that they didn't scare you away from the Netherlands. Connor, Manish, and Yvonne brought life back to the office after a silent period, making it possible to have a random chat at any time of the day. Before joining MF4.086, Alberto and Clara warmly welcomed me to the 300+-page-thesis office. Lasse, you were a great office mate in "the sauna" at floor I in Zürich, thank you for introducing me to other areas of mathematics.

The one-but-last place in this section is for Louise. It is hard to express how grateful I am for your support throughout these years. You gave me freedom and encouraged me to explore the world and to develop myself in any direction, pursuing all kinds of mathematical, musical, and organisational activities, and sometimes to do just nothing. As a result, I was often physically or mentally not present, which must have been hard for you at times. Still you were there when I needed you. Every day you (try to) make me realise that there is more in life than mathematics. Thank you for always being a safe haven to me.

Lastly, I want to thank you, the reader of this thesis. I hope that this first section encourages you to continue reading the next chapters that are founded on thoughts during many sleepless nights.

Joost Jorritsma,
April 2023.

CONTENTS

1

INTRODUCTION

1.1 MOTIVATION

How is it possible that COVID-19 could spread so quickly around the world? What are the main reasons that can make a (fake) message go viral? Although these two questions are of a very different nature, they are similar from a mathematical perspective. A way to model information diffusion (or disease spreading) in a real-world network mathematically, is by representing the topological structure of the underlying network as an (un)directed graph $\mathcal{G} = (V, E)$, where the vertices V represent people, and two vertices are connected by an edge $e \in E \subseteq V \times V$ if they see each other regularly, or are friends in the social network. Each edge has a non-negative weight attached to it, standing for the time that it takes to transmit a message from one side of the edge to the other.

Since the underlying networks are often unknown (let alone the weights), *random graph* models can be used that share topological properties with the original network [201]. To *qualitatively* understand processes in real-world networks (such as information diffusion), one can model the process on a random graph, and translate the results back to the original network. Another advantage of modelling real-world networks as (weighted) random graphs, is that many networks are believed to share *universal properties*, that we highlight in the following paragraphs. As a result, by studying a simplistic random graph model that shares the same universal properties as real-world networks, one could gain qualitative insight in structures in multiple real-world networks at once. We highlight four properties that many real-world networks are believed to share.

Scale-free degree distribution. The degree sequences are scale-free, i.e., the number of connections (degree) per vertex decays as a power law: there exists $\tau > 1$ such that, writing p_k for the proportion of vertices with degree k,

$$p_k \sim k^{-\tau}. \tag{1.1.1}$$

Here, $p_k \sim k^{-\tau}$ means that $k^{-\tau-\varepsilon} < p_k < k^{-\tau+\varepsilon}$ for any $\varepsilon > 0$ when k is sufficiently large. For example in the WWW, social networks, and protein networks, the power-law exponent τ is estimated to be in $(2, 3)$, see e.g.

[7, 8, 201]. When $\tau \in (2,3)$, the asymptotic degree distribution has infinite second moment. In this case, the degree dispersion is very high, which indicates the presence of *hubs* in the network.

Supercriticality. Typically, a large proportion of nodes are part of the same connected component $\mathcal{C}^{(1)}$: the largest set of nodes in a network such that for all $u, v \in \mathcal{C}^{(1)}$ there is a path of edges in the network from u to v. Mathematically, we consider sequences of graphs $(\mathcal{G}_n)_{n\geqslant 1}$ for which the number of vertices in $\mathcal{G}_n$ tends to infinity, and say that the graph sequence $(\mathcal{G}_n)_{n\geqslant 1}$ is supercritical if a linear-sized connected component exists with probability tending to one.

Small-world property. Although real-world networks may contain billions of nodes, the graph distances, i.e., the minimal number of edges to connect two nodes, is very small compared to the size of the network [15, 191, 201, 222]. We say that $(\mathcal{G}_n)_{n\geqslant 1}$ satisfies the *small-world property* if the graph distance between two vertices chosen uniformly at random from the largest connected component in $\mathcal{G}_n$, also called *typical graph distance*, grows at most logarithmically with the number of vertices. The network is an *ultra-small world* if the typical graph distance is at most of doubly logarithmic order in the number of vertices.

Clustering. Empirically, clustering quantifies the effect commonly known as "a friend of a friend is also likely to be my friend". Mathematically, clustering refers to the presence of *triangles* in the network. A natural way to model a network that contains triangles, is by embedding the vertex set in Euclidean space, such that vertices that are close to each other are more likely to be connected. The presence of underlying geometry is easily motivated through real-world networks: people living nearby (or having similar interests, which could be represented by a so-called *feature vector*) tend to know each other with higher probability. On the other hand, there has been a large interest in embedding real-life networks (in which the spatial embedding is less obvious than just the geographical location of a person) into Euclidean space [111]. In this case, the spatial nearness of two vertices represents similarity of these vertices in the (partially) observed network: for example, the obtained spatial location could encode an unobserved feature vector of a node in the network. These spatial embeddings could then be used for machine learning tasks, such as link prediction or community detection.

1.1.1 *Aim of the thesis*

In this thesis we study (theoretical and generative) supercritical random graph models from three different perspectives.

(i) In Part i we study first-passage percolation, a model for information diffusion and epidemic spreading, and the small-world property on preferential attachment models: a class of random graph models that gives a possible explanation to the emergence of the power law in the degree sequence of real-world networks [20]. This class of model exhibits the scale-free and small-world property, is supercritical [71] and in some cases connected [124], but it does not satisfy the clustering property [99].

(ii) In Part ii we study the sizes of connected components in a framework called *kernel-based spatial random graph models*. Random graph models in this class can exhibit all four properties mentioned above. The vertices are embedded in Euclidean space and the degree distribution has a heavy tail. When the inhomogeneity in the graph is sufficiently large, the obtained scalings on the sizes of connected components are vastly different from the only known scalings for spatial random graphs, i.e., random geometric graphs and the graph obtained by nearest-neighbour percolation on $\mathbb{Z}^d$.

(iii) In Part iii we conduct two simulation-based studies of spreading processes on networks motivated by COVID-19. We demonstrate the influence of space and the presence of long-range edges in spatial networks on various spreading processes that are more realistic than first-passage percolation. We compare characteristics of these spreading processes to classical non-spatial models (e.g., effect of temporary immunity, shape of the epidemic curve, number of persons in quarantine over time) and illustrate how the spatial models can be used to give intuition for qualitative implications of the intervention strategies, and to model contact-tracing applications for infection spread.

In the remainder of this chapter we introduce the random graph models that are studied in this thesis, and explain the relations between the these models. Moreover, we give extensive references for the random graph models that have the main focus: preferential attachment models and kernel-based spatial random graphs; at the end of the chapter we summarize the main contributions of the thesis. In the following chapters that describe the detailed results, we motivate and define the properties that are investigated in the specific chapter (e.g., graph distance, first-passage percolation, cluster-size distribution, and spreading process characteristics) in detail. Eventually, in Chapter 9 we discuss the main contributions and present directions for future research.

1.2 NON-SPATIAL RANDOM GRAPH MODELS

Non-spatial random graph models are often used as null-models to compare network data to "purely random" networks. The preferential attachment model (PAM) mentioned above is an example of such a model, and this model has the main focus in Part i. To contrast PAMs against more classical null-models, and to motivate a new research line for PAMs, we first introduce two more classical models: the configuration model and the inhomogeneous random graph. Besides, the notion of inhomogeneous random graphs will serve as motivation for the kernel-based spatial random graphs (KSRGs) that have the main focus in Parts ii and iii. We introduce KSRGs below in Section 1.3.

1.2.1 *Configuration model and inhomogeneous random graph*

The configuration model and inhomogeneous random graph allow to specify the (expected) degrees of the vertices in the resulting graph. Amongst others, this makes it possible to mimic the scale-free degree distribution observed in real-world networks. The configuration model is used in Part iii as one of the underlying networks to model the spread of diseases.

Configuration model

The configuration model is a well-studied object in the mathematical literature. It dates back to Bollobás [34] and Molloy and Reed [188, 189], we refer to [124] for references.

Definition 1.2.1 (Configuration model)**.** Fix $n \geqslant 1$ the number of vertices. Prescribe to each vertex $u \in \{1, 2, \dots, n\}$ its degree $\deg(u) \geqslant 0$, so that the total degree $h_n := \sum_{u \leqslant n} \deg(u)$ is even. To form a graph, to each vertex u we assign $\deg(u)$ number of half-edges and the half-edges are then paired uniformly at random to form edges. The resulting random multi-graph is the configuration model.

The random pairing of the half-edges ensures that the configuration model is a random (multi)graph, even though the prescribed degrees might be deterministic. The configuration model naturally allows for the scale-free property, since one specifies the degree distribution. Besides that, the configuration model also satisfies the small-world property [129, 130]: when the degrees follow a power law with parameter $\tau > 2$, the typical graph distance $d_n^{(G)}(u,v)$, i.e., the number of connections on the shortest path between two vertices u, v that are sampled uniformly at random from a graph on n vertices, satisfies

$$d_n^{(G)}(u,v) = \begin{cases} \Theta(\log N), & \text{when } \tau > 3, \\ \Theta(\log\log N), & \text{when } \tau \in (2,3). \end{cases} \tag{1.2.1}$$

While the configuration model easily accommodates power-law degrees and satisfies the small-world property, it does not contain clustering [82, 217].

1.2.1.1 *Inhomogeneous random graph*

Alternatively to the configuration model, one can use the Chung-Lu [50] and the similar Norros-Reittu model [204] as null-models. These random graph models are also referred to as rank-1 inhomogeneous random graphs. Contrary to the configuration model, in a rank-1 inhomogeneous random graph only the *expected* degrees of vertices are prescribed, rather than their exact degree. The expected degree of a vertex u is encoded by a non-negative mark W_u (also called weight).

Definition 1.2.2 (Chung-Lu model)**.** Fix $n \geqslant 1$ the number of vertices. Prescribe to each vertex $u \in \{1, 2, \dots, n\}$ a non-negative mark W_u. To form a graph, we connect each pair of vertices $\{u, v\}$ by an edge with probability

$$p_{u,v} := \min\left\{\frac{W_u W_v}{n\mathbb{E}[W]}, 1\right\}, \tag{1.2.2}$$

independently of other possible edges.

For the Norros-Reittu model, one replaces the connection probability by the function $p_{u,v} = 1 - \exp\left(-W_u W_v / \sum_y W_y\right)$. Similar to the configuration model, these models satisfy the small-world property [50], but violate the clustering property [217].

We proceed to a special case of the inhomogeneous random graph introduced by Bollobás, Janson, and Riordan in [37]. As in rank-1 inhomogeneous random graphs, each vertex is equipped with a non-negative mark W_u. However, they generalise the connection probability by replacing the product of the two marks in the numerator in (1.2.2) by an arbitrary non-negative symmetric function κ, called the *kernel*.

Definition 1.2.3 (Inhomogeneous random graph)**.** Fix $n \geqslant 1$ the number of vertices, and let $\kappa(\cdot,\cdot)$ be a non-negative symmetric function. Prescribe to each vertex $u \in \{1, 2, \ldots, n\}$ a non-negative mark W_u. To form a graph, we connect each pair of vertices $\{u, v\}$ by an edge with probability

$$p_{u,v} := \min\left\{\frac{\kappa(W_u, W_v)}{n}, 1\right\},$$

independently of other possible edges.

Indeed, $\kappa(W_u, W_v) = W_u W_v$ yields the same connection probability as in (1.2.2), up to a scaling of the marks by a constant. We call this kernel also the *product kernel*. In Section 1.3 below, we generalise this model to a spatial random graph, where the vertices are embedded in Euclidean space.

Static models

The configuration model and rank-1 inhomogeneous random graph are examples of *static random graph models*. The vertex set is a fixed set, and the degree distribution of a sampled graph is determined by the specified degree or mark distribution. Moreover, these models are not projective: the graph on $n+1$ vertices cannot be obtained by simply adding a vertex and the edges emanating from this vertex to the graph on n vertices. This is in contrast to preferential attachment models, that we introduce next, which have the main focus in Part i of this thesis.

1.2.2 *Preferential attachment models*

Unlike the configuration model and the inhomogeneous random graph, the preferential attachment model (PAM) is an example of a model where the vertex and edge set grow dynamically over time. A model similar to PAMs was introduced by Yule to study biological processes [229], but the model became increasingly popular after the paper [20], that uses the model to give an explanation to the emergence of power-law degrees in various real-world networks. Bollobás, Riordan, Spencer, and Tusnády were the first ones to define and investigate the model rigorously from a mathematical perspective [33].

We introduce the models informally. The construction of a preferential attachment graph is initialized with a graph PA_1. Arrivals of vertices happen deterministically at times $t \in \{2, 3, ...\}$. We denote the graph at time t by PA_t and label all the vertices by their arrival time, also called *birth time*. In the models that we study, arriving vertices favour connecting by an edge to vertices that have a high degree. Writing $\{t \to v\}$ for the event that vertex t connects by an edge to a vertex $v < t$ that is present in PA_{t-1}, we assume for some $\tau > 2$ that

$$\mathbb{P}\big(\{t \to v\} \mid PA_{t-1}\big) = \frac{D_v(t-1)}{t(\tau-1)}(1+o(1)),$$

where $D_v(t-1)$ denotes the degree of vertex v directly after the arrival of vertex $t-1$. As a result, the asymptotic degree distribution has a power-law decay (as defined in (1.1.1)) with exponent τ [70, 124], that we therefore call the power-law exponent.

We call the number of connections that a new vertex establishes upon arrival the *outdegree*. We study three variants of this model: one with fixed outdegree, which we call fixed preferential attachment (FPA), and two with variable outdegree (i.e., random outdegree), which we call variable preferential attachment (VPA), and generalised variable preferential attachment (GVPA), respectively.

The fixed preferential attachment model appeared formally in [23, 38], using a parametrisation that only yields preferential attachment graphs with $\tau \geqslant 3$. In this thesis, we consider the model with power-law exponent $\tau \in (2, 3)$. We recursively grow the graph PA_t for $t \geqslant 2$ from a starting graph PA_1. At time t we sequentially add m outgoing edges to the arriving vertex that has label t. After the j-th edge has been formed, the defining connection probabilities are updated. We write $D^{\leftarrow}_{(t,j)}(v)$ for the indegree, the number of incoming edges, of a vertex v right after the $(tm+j)$-th edge is added to the graph. Similarly, we write $FPA_{(t,j)}$ for the constructed graph right after the moment it contains exactly $(tm+j)$ edges. Let $\{t \xrightarrow{j} v\}$ be the event that the j-th edge of vertex $t \in \mathbb{N}$ is attached to $v \in [t-1]$.

Definition 1.2.4 (FPA(m, δ))**.** Fix $m \in \mathbb{N}, \delta \in (-m, \infty)$. Let $FPA_1(m, \delta)$ be the graph with a single vertex without any edges. The model $FPA(m, \delta)$ is defined by the following sequence of conditional edge-connection probabilities

$$\mathbb{P}\big(\{t \xrightarrow{j} v\} \mid \mathrm{FPA}_{(t,j)}\big) = \frac{D^{\leftarrow}_{(t,j-1)}(v) + m\,(1+\delta/m)}{(t-2)m(2+\delta/m) + j - 1 + m + \delta}, \qquad (1.2.3)$$

where $v \in [t-1]$. The power-law exponent of the model is

$$\tau_{m,\delta} := 3 + \delta/m. \qquad (1.2.4)$$

The asymptotic degree distribution in FPA decays as a power law with exponent $\tau_{m,\delta}$ in (1.2.4), see [124, 193], so that for FPA $\tau \in (2,3)$ when $\delta \in (-m, 0)$. The constraint $\delta > -m$ ensures well-defined probabilities in (1.2.3). We remark that there are multiple versions of the fixed preferential attachment model in literature, e.g., some versions allow for self-loops, and in other versions the connection probabilities are updated only after the formation of all m edges of a single vertex. We chose to present a single version of the model here, in which no self-loops are allowed, but multi-edges are allowed. This specific version allows us to re-use directly some results from [24]. With the recent results from [98], all derived results in Part i extend to the versions of PAMs introduced in [124], to which we refer for an extensive overview of the fixed preferential attachment model.

In FPA the total number of edges in the graph is deterministic, making some explicit calculations easier. On the contrary, the events $\{t \to v_1\}$ and $\{t \to v_2\}$ are negatively correlated for vertices $v_1 \neq v_2$, yielding more involved computations.

The next models that we introduce behave to some extent as the opposite of FPA, as the edges are conditionally independent, leading to a random outdegree of vertices. These two models were introduced by Dereich and Mörters [70], and they call the graph *preferential attachment with conditionally independent edges*. It appears in [124] as *Bernoulli preferential attachment*. Let $D^{\leftarrow}_t(v)$ be the indegree of the vertex v right before time t.

Definition 1.2.5 (VPA(f), GVPA(f))**.** Let $f : \mathbb{N} \to (0, \infty)$ be a concave function satisfying $f(0) \leqslant 1$ and $f(1) - f(0) < 1$. We call f the *attachment rule*. Let $\mathrm{GVPA}_1(f)$ be the graph with a single vertex without any edges. Conditionally on $\mathrm{GVPA}_{t-1}(f)$, vertex t connects to $v \in [t-1]$ by an edge with probability

$$f\left(D^{\leftarrow}_t(v)\right)/t,$$

independently of the other existing vertices. Important parameters of the model are

$$\gamma_f := \lim_{t\to\infty} f(t)/t, \qquad \tau_f := 1 + 1/\gamma_f, \qquad (1.2.5)$$

which are well-defined by the concavity of f. We call τ_f the power-law exponent. For general f, we call the model generalised variable preferential attachment (GVPA). For affine f, i.e., $f(k) = \gamma k + \eta$ for some $\gamma, \eta > 0$, we call the model variable preferential attachment (VPA).

The asymptotic degree distribution decays as a power law with exponent τ_γ in (1.2.5), see [70]. In the more restricted model VPA, calculations become explicit and more precise results can be derived, e.g., degree distribution in [70], or the 'graph-distance evolution' in Chapter 3.

The three models FPA, VPA, and (even the more general model) GVPA, behave qualitatively similar in terms of their degree distribution and typical graph distance when $\tau \in (2, 3)$ [69]. Hence, we refer to the models by their power-law exponent τ, and to call them PA collectively. In this thesis, we distinguish them only when different proofs are required, or when referred to different results from literature.

1.2.3 *Preferential attachment: literature overview*

We recall some previous results on the three preferential attachment models FPA, VPA, and GVPA: first we reflect on the four universal properties in real-world networks described in Section 1.1. After that, we mention results in a wider range of preferential attachment models than FPA and (G)VPA, identifying that a majority of the PAM literature takes a typical approach, which we call *snapshot analysis*: this approach allows to compare the *dynamical* PAMs to *static* random graph models, such as the configuration model, or inhomogeneous random graphs. In this thesis, we take a different approach: we do not aim to compare PAMs to static models, but study the dynamical changes in the preferential attachment graph itself.

Properties of real-world networks

The models FPA and GVPA that we study in this thesis are the most commonly used pure PAMs in the literature, i.e., in these models it is solely the preferential attachment mechanism that drives the changes in the graph topology. These PAMs are mathematically defined by Bollobás and Riordan [38], and Dereich and Mörters [70]. The models satisfy the scale-free property [70, 193], small-world property [46, 69, 81] (together with results in Part i in this thesis), and supercriticality [71, 126]. In [97] it is shown that FPA does not satisfy the clustering property, and it is expected that these results extend to GVPA.

Spatial variants where vertices have a location in an underlying Euclidean space are studied in [3, 142, 143]. Here, closeness in Euclidean distance is combined with preferential attachment. The age-dependent random connection model [105, 109] is a recent spatial version, that we consider also in Part ii in this thesis as a special case. In the age-dependent random connection model, the connection probabilities are not governed by the degree of vertices, but by their relative age compared to the arriving vertex.

Snapshot analysis

Since the introduction of FPA in [38], many variants with more involved dynamics and connection functions have been introduced, of which we highlight some. In [147, 213], the vertex set is fixed and only edges are formed dynamically. The variations introduced in [10, 65] allow for edges being formed (or deleted in [57, 65]) between existing vertices. References [70, 71] consider a version where the attachment function can be sublinear in the degree. In [51, 68, 73, 94] vertices are equipped with a fitness and in [180] the arriving vertices have a power of choice: for each edge formed at time t, two present vertices i_1, i_2 are sampled with probability proportional to the degree, after which the edge $\{t, i_\star\}$ is added to the graph, where $i_\star := \arg\max_{i\in\{i_1,i_2\}} D_t(i)$. Stochastic processes on PAMs have been analysed in [23, 43] for the contact process, in [11] for bootstrap percolation, and in Part i of the thesis for first-passage percolation.

The above mentioned results and papers provide statements about *static* snapshots of the graph PA_t in the large network limit: the network is considered at a *single time* t as t tends to infinity. This snapshot analysis allows for comparison to (simpler) static random graph models, such as the configuration model [34, 188], Chung-Lu model [50], and the Norros-Reittu model [204], and aims to classify properties of random graphs as either *universal* or *model-dependent*.

An example of a universal property is the typical graph distance: in the configuration model, Chung-Lu model, Norros-Reittu model, and preferential attachment models FPA, VPA, and GVPA, with parameters such that the asymptotic degree distribution decays as a power law with exponent $\tau \in (2, 3)$, the typical graph distance grows doubly logarithmic in the network size [46, 50, 69, 81, 129].

On the other hand, the configuration model (under mild assumptions on the degree distribution) and fixed preferential attachment model lead to a connected graph as the number of vertices tends to infinity [89, 126], while the Chung-Lu and Norros-Reittu model, and the variable preferential attachment models have several connected components [49, 71], so connectivity of the graph is an example of a model-dependent property. We refer to [124, 126] and its references for other universal and model-dependent properties.

Evolution of graph properties: a novel approach

We take a different approach and deviate from the snapshot analysis: we study the temporal changes of graph properties, i.e., we fix the graph PA_t at a large time t, and study the graph sequence $(\mathrm{PA}_{t'})_{t'\geqslant t}$ at *all times* $t'\geqslant t$ from the perspective of the vertices that are present at time t.

This approach has only been considered in papers on the degree of fixed vertices [70, 193], a local graph property. In Chapter 3, we study how a *global* graph property, i.e., the graph distance between two fixed vertices (typical vertices at time t), changes as the surrounding graph evolves over time. Our obtained results reflect the temporal dynamics of the graph, and this approach commences a new type of research on PAMs.

The preferential attachment is the considered random graph model in Part i in this thesis. We proceed to the *spatial* random graph models that have the main focus in Parts ii and iii.

1.3 KERNEL-BASED SPATIAL RANDOM GRAPHS

None of the models introduced in Section 1.2 satisfies the clustering property that many real-world networks are believed to share [217]. In Part ii and iii in this thesis, we study (instances of) a large class of random graph models where the vertices are embedded in Euclidean space, i.e., each vertex u has a location $x_u \in \mathbb{R}^d$. In these models, the connection probabilities are such that spatially nearby vertices are more likely to be connected by an edge. This naturally leads to clustering in those graphs [60, 91, 135, 217]. The models that we study are parameterized such that both the degree distribution and the edge-length distribution can be heavy-tailed, i.e., there is structural inhomogeneity present in the graphs. Here, the length of the edge $\{u, v\}$ refers to the Euclidean distance $\|x_u - x_v\|$ between the vertices forming the edge.

Long-range percolation and continuum long-range percolation (LRP and CLRP) [4, 216] are examples of classical models in which the edge-length distribution has a heavy tail. In these models, each potential edge $\{u, v\}$ is independently present with probability proportional to $\beta^{\alpha}\|u-v\|^{-d\alpha}$ for some $\alpha, \beta > 0$. However, in the graphs formed by these models, the degree distribution is light-tailed. We study long-range percolation in Chapter 6.

Our main focus in Chapter 5 and Part iii is on models where *both* the degree- and edge-length distribution are heavy-tailed. Recently, several such spatial random graph models have been introduced that gained significant interest in the literature; they were introduced amongst others to accommodate the need for modelling degree inhomogeneity (i.e., scale-free property) in real-life networks [7, 32, 83, 200, 198]. Examples of such models include scale-free percolation (SFP) [63]; geometric inhomogeneous random graphs (GIRG) [40], continuum scale-free percolation (CSFP) [67], hyperbolic random graphs (HRG) [169], the ultra-small scale-free geometric network [228]; scale-free Gilbert model (SGM) [122], the Poisson Boolean model with random radii [103], and the age- and the weight-dependent random connection models (ARCM) [105, 107].

We consider a framework unifying these models, which we call *Kernel-Based Spatial Random Graphs* (KSRG). Beyond the above models, our framework also contains the following classical models: the graph obtained by nearest-neighbour bond percolation on $\mathbb{Z}^d$ (NNP), and the random geometric graph (RGG) [182, 210]. The framework was essentially present in [166], and then refined in [107]. A random graph model in the KSRG class requires *four ingredients* to be specified:

- a *vertex set* $\mathcal{V}$: any stationary ergodic point process embedded in a metric space with norm $\|\cdot\|$; typically a Poisson point process (PPP) Ξ on $\mathbb{R}^d$ or a lattice $\mathbb{Z}^d$.
- a *mark distribution* W for vertices, the mark w_u of vertex u describing its ability to form connections.
- a *profile function* ρ, describing the influence of spatial distance on the probability of an edge. The choice of ρ potentially leads to a heavy-tailed edge-length distribution.
- a *kernel* κ that describes how marks *rescale* the spatial distance between vertices.

Conditionally on the vertex set and the marks, two vertices u and v are connected by an edge, independently of other possible edges, with a probability given by the *connectivity function*, which is the composition of the kernel and the profile, and a parameter $\beta > 0$ that controls the edge density (high values of β correspond to higher edge density):

$$\begin{aligned}\mathbf{P}(&\text{vertex } u \text{ is connected to vertex } v \text{ by an edge})\\ &= \rho\big(\beta\kappa(w_u, w_v)\|u-v\|^{-d}\big),\end{aligned} \tag{1.3.1}$$

where $\mathbf{P}$ is a measure on $\mathcal{V}\times\mathcal{V}$, conditionally on $\mathcal{V}$ and $\{W_v\}_{v\in\mathcal{V}}$. These components together define a random graph model $G(\mathcal{V},\mathcal{E})$, where $\mathcal{E}$ is a random set of edges.

The role of marks. Throughout the thesis, we will assume that the marks are power-law distributed with tail-exponent $\tau-1$ for some $\tau>1$, that is,

$$\mathbb{P}(W \geqslant x) = x^{-(\tau-1)}, \qquad \text{for } x \geqslant 1. \tag{1.3.2}$$

We shall see that the mark distribution regulates the tail of the degree-distribution. Marks have different names in the literature: in geometric inhomogeneous random graph [40] and scale-free percolation [63], they are called 'vertex-weight' or 'fitness', while for hyperbolic random graphs [169] the mark corresponds to the 'radius' of the location of a vertex in the native representation of the hyperbolic plane. For the age-dependent random connection model [105] the mark corresponds to the rescaled 'age' of a vertex. For nearest-neighbour percolation, random geometric graphs and long-range percolation marks are not used for the formation of the graph, then we set $W\equiv 1$, and formally $\tau=\infty$. Other nonnegative mark distributions than (1.3.2) are allowed, e.g., exponential or Poisson. However, for the problems considered in this thesis, lighter-tailed mark distributions behave qualitatively similar as setting $\tau=\infty$.

The role of the profile function. The profile function is generally either *threshold* or *long-range* type in the literature, and two typical choices are as follows:

$$\rho_{\mathrm{thr}}(t) = \mathbb{1}_{\{[1/R,\infty]\}}(t), \qquad \rho_\alpha(t) = \min\{1, t^\alpha\}. \tag{1.3.3}$$

Since the inverse distance $1/\|u-v\|^d$ is used in (1.3.1), ρ_{thr} corresponds to connecting all pairs of vertices at distance at most $R^{1/d}$ by an edge, while ρ_α – originating from long-range percolation [216] – allows for long-range edges. The parameter $\alpha>1$ in (1.3.3) calibrates this effect: the smaller α, the more long-range edges, and the condition $\alpha>1$ ensures a locally finite graph. With slight abuse of notation we say that $\alpha=\infty$ when considering a threshold profile. Both parameters α,τ *regulate exceptional connectedness*, but differently: α in (1.3.3) controls the presence of exceptionally long edges, whereas τ in (1.3.2) controls exceptionally high-degree vertices. Non-threshold profiles with subpolynomial tails can be considered. For the results in this thesis, we conjecture that those models fall in the same universality class as threshold profiles, i.e., one can set formally set $\alpha = \infty$.

The role of the kernel function. The kernel function κ in (1.3.1) is a symmetric non-negative function that rescales the spatial distance so that high-mark vertices experience 'shrinkage of spatial distance' and hence, the connection probability increases towards high-mark vertices. It has the same role as the kernel function in inhomogeneous random graphs, see Definition 1.2.1.1. Models in the literature commonly use the *trivial*, the *product*, the *max*, *sum*, *min*, and *preferential attachment* (PA) kernels, the last one mimicking the spatial preferential attachment model [3, 142]:

$$\begin{aligned} \kappa_{\mathrm{triv}}(x,y) &\equiv 1, & \kappa_{\mathrm{prod}}(x,y) &= xy, \\ \kappa_{\max}(x,y) &= \max(x,y), & \kappa_{\mathrm{sum}}(x,y) &= x+y, \\ \kappa_{\min} &= \min(x,y)^{\tau}, & \kappa_{\mathrm{pa}}(x,y) &= \max(x,y)\min(x,y)^{\tau-2}, \end{aligned} \tag{1.3.4}$$

where $\tau \in (2,\infty)$ corresponds to the tail exponent of the mark distribution in (1.3.2). The trivial kernel κ_{triv} includes classical models into the class of KSRGs, e.g., the graph obtained by bond percolation on $\mathbb{Z}^d$ (in which each edge between two vertices at Euclidean distance 1 is included independently with probability p), the random geometric graph, and (continuum) long-range percolation. The parametrisation in all nontrivial kernels in (1.3.4) give rise to a power-law degree distribution with the *same* tail-exponent $\tau - 1$ as for the marks in (1.3.2)[107], establishing the scale-free property often desired in complex network modelling [32, 83, 198, 200]. The choice of the kernel κ can influence the amount of correlation between degrees of two endpoints of an edge, called *assortativity*, see Section 1.3.1 below.

We refer to Table 1 for an overview of models that fit into the framework of KSRGs.

Properties of the giant component in KSRGs. Many graph properties and processes on graphs have been studied on special cases of KSRGs with non-trivial kernels in the past years: the *degree distribution* to establish the scale-free property in [40, 63, 105, 122, 134, 179, 228]; *graph distances* in the giant component in [39, 63, 67, 106, 122, 228], establishing the small-world property for certain parameter settings; *first-passage percolation* on the giant component in [136, 165, 166]; the *random walk* on the giant component [52, 107, 121, 161]; the *contact process* on the giant component [104, 177]; *bootstrap percolation* in [44, 164]; the *clustering coefficient* in [40, 60, 91, 135,

Model	$\mathcal{V}$	**Kernel**	**Profile**
Bond percolation on $\mathbb{Z}^d$ [113]	$\mathbb{Z}^d$	$\kappa_{triv}, \kappa_{0,0}$	ρ_{thr}
Random geometric graph [210]	PPP		ρ_{thr}
Long-range percolation [216]	$\mathbb{Z}^d$		ρ_α
Continuum long-range percolation [4]	PPP		ρ_α
Scale-free percolation [63]	$\mathbb{Z}^d$	$\kappa_{prod}, \kappa_{1,1}$	ρ_α
Continuum scale-free percolation [67]	PPP		ρ_α
Geometric inhomogeneous random graph [40]	PPP		ρ_α, ρ_{thr}
Hyperbolic random graph [169]	PPP		ρ_{thr}
Age-dependent random connection model [105]	PPP	$\kappa_{pa}, \kappa_{1,\tau-2}$	ρ_α, ρ_{thr}
Scale-free Gilbert graph [122]	PPP	$\kappa_{max}, \kappa_{1,0}$	ρ_{thr}
Ultra-small scale-free geometric network [228]	$\mathbb{Z}^d$	$\kappa_{min}, \kappa_{0,1}$	ρ_{thr}
Interpolation KSRG	**PPP**	$\kappa_{1,\sigma}$	ρ_α, ρ_{thr}

Table 1: Models belonging to the KSRG framework, their vertex sets (either the lattice $\mathbb{Z}^d$, or a Poisson point process, abbreviated by PPP), kernels, and profiles. Horizontal lines separate models with different kernels. The Interpolation KSRG is the model that has the main focus in Part ii in this thesis.

217]; and the existence of subcritical and supercritical phases in [63, 67, 108, 109], and subcritical phase in the more general model [144] that allows for dependencies between edges.

Thus, the giant component in the supercritical phase is relatively well understood. However, there are no results yet for smaller components in the graph. This is the main topic of Part ii. In Part iii we perform simulation studies of more involved stochastic processes than the contact process or first-passage percolation on the giant component in geometric inhomogeneous random graphs (a KSRG that uses the product kernel κ_{prod}). There, we use it as a model for real-life networks.

1.3.1 *The interpolation KSRG (i-KSRG)*

We will study models belonging to the class of KSRGs in a unified way, using a parameterized kernel that contains the kernels in (1.3.4) as special cases. We call this the *interpolation kernel*, and the corresponding random graph model an *interpolation KSRG*. Independently of this work it appeared before in [108, 179], and it is also used in [135]. Formally, we define the kernel as

$$\kappa_{\tilde{\sigma},\sigma}(x,y) := \max(x,y)^{\tilde{\sigma}} \min(x,y)^{\sigma}, \tag{1.3.5}$$

for some $\tilde{\sigma} \geqslant 0$ and $\sigma \in \mathbb{R}$. Non-negativity of $\tilde{\sigma}$ ensures the fairly natural requirement that a higher mark corresponds to a higher expected degree. The parameters $\tilde{\sigma}, \sigma$ affect the tail exponent of the degree distribution, as stated in the next proposition that considers the degree distribution of a typical vertex 0 with location at the o (whose mark is power-law distributed with tail-exponent $\tau - 1$, see (1.3.2)). For two functions f, g, we write $f \asymp g$ if there exist constants $c, C > 0$ such that $cg(k) \leqslant f(k) \leqslant Cg(k)$ for all $k \geqslant 1$.

Proposition 1.3.1. *Consider a supercritical i-KSRG model with kernel $\kappa_{\tilde{\sigma},\sigma}$ with $\tilde{\sigma}, \sigma \geqslant 0$ from (1.3.5), and parameters $\alpha > 1, \tau > 2$ and $d \in \mathbb{N}$. Whenever $\tau - 1 > \tilde{\sigma}$ (needed for a locally finite graph), the tail of the degree distribution is given by*

$$\mathbb{P}(\deg(0) \geqslant k) \asymp \begin{cases} k^{-(\tau-1)/\tilde{\sigma}} & \text{if } \tau - 1 > \max(\tilde{\sigma}, \sigma), \\ k^{-(\tau-1)/(\tilde{\sigma}+\sigma-(\tau-1))} & \text{if } \tilde{\sigma} < \tau - 1 \leqslant \sigma. \end{cases}$$

We refer the reader for a proof to [179], and to similar results for specific KSRGs in [63, 105, 122, 228].

In models with kernels $\kappa_{\max}, \kappa_{\mathrm{pa}}, \kappa_{\mathrm{prod}}$ in (1.3.4) we set $\tilde{\sigma} = 1$, while for $\kappa_{\min}$ we set $\tilde{\sigma} = 0$ and $\sigma = 1$. The kernel κ_{sum} cannot be directly expressed using the interpolation kernel. However, since $\max(x,y) \leqslant x + y \leqslant 2\max(x,y)$, and all known phase transitions regarding macroscopic behaviour are the same for models with κ_{sum} and with $\kappa_{\max}$, we consider this a minor restriction. By appropriately changing the mark distribution, unless $\tilde{\sigma} = 0$, any KSRG with kernel $\kappa_{\tilde{\sigma},\sigma}$ can be re-parameterized to have $\tilde{\sigma} = 1$. Even if $\tilde{\sigma} = 0$, it can still be approximated with the kernel $\kappa_{1,\sigma}$, as we show below in Section 1.3.2. Hence, from now on we fix $\tilde{\sigma} = 1$.

The parameter σ affects *assortativity* (amount of correlation between the degrees of two vertices that are connected by an edge): a larger value of σ increases $\min(x,y)^{\sigma}$ in (1.3.5), and in turn the connection probability in (1.3.1). In a natural coupling of these models using common edge-variables, edges incident to at least one low-mark vertex are barely affected by changing σ. However, edges between two high-mark vertices are created rapidly if σ increases. Thus, changing σ gradually affects the assortativity in the graph.

Using the interpolation kernel allows for unified proof techniques. We believe that i-KSRGs are of independent interest beyond the unification they allow, because of the interpretation of σ as assortativity parameter. Varying σ may affect the structure of the graph in different problems of interest, e.g. clustering coefficient or the behaviour of (the mixing time of) the simple random walk on the graph.

1.3.2 *Approximating the min-kernel*

We will argue now that the kernel $\kappa_{0,\tau}$ (corresponding to the min-kernel $\kappa_{\min}$ in (1.3.4)) can be approximated by the kernel $\kappa_{1,\sigma}$ by taking an appropriate limit in the parameter space of KSRGs. Let $\tau > 1$ be the desired tail-exponent for the degree distribution in (1.3.2) that is also reflected in the parametrised kernel $\kappa_{\min} \equiv \kappa_{0,\tau}$ in (1.3.4). We will construct a sequence $(\tau_\ell, \sigma_\ell)_{\ell \geqslant 1}$ and a coupling of KSRGs, such that the edge probabilities in a KSRG with kernel κ_{1,σ_ℓ} and mark-parameter τ_ℓ tend to the edge probabilities in a KSRG with kernel $\kappa_{0,\tau}$ and mark-parameter τ as ℓ tends to infinity. Let $(\tau_\ell)_{\ell \geqslant 1}$ be sequence tending to infinity. Let U denote a Unif$[0,1]$ random variable, and W, W_ℓ denote random variables satisfying $\mathbb{P}(W \geqslant x) = x^{-(\tau-1)}$ and $\mathbb{P}(W_\ell \geqslant x) = x^{-(\tau_\ell-1)}$ for all $x \geqslant 1$, respectively. Then,

$$U^{-1/(\tau_\ell-1)} \overset{d}{=} W_\ell; \quad U^{-1/(\tau-1)} \overset{d}{=} W \quad \Longrightarrow \quad W_\ell \overset{d}{=} W^{(\tau-1)/(\tau_\ell-1)}.$$

Thus, letting $W_\ell^{(1)}$ and $W_\ell^{(2)}$ be two independent copies of W_ℓ, we can couple them with two independent copies $W^{(1)}, W^{(2)}$ of W, such that under this coupling

$$\begin{aligned}\kappa_{1,\sigma}(W_\ell^{(1)}, W_\ell^{(2)}) &= \max\{W_\ell^{(1)}, W_\ell^{(2)}\}\min\{W_\ell^{(1)}, W_\ell^{(2)}\}^{\sigma} \\ &= \max\{W^{(1)}, W^{(2)}\}^{(\tau-1)/(\tau_\ell-1)} \\ &\qquad \cdot \min\{W^{(1)}, W^{(2)}\}^{\sigma(\tau-1)/(\tau_\ell-1)}.\end{aligned}$$

For a fixed value of $\tau > 1$, the exponent of the maximum tends to zero if τ_ℓ tends to infinity. Hence, when choosing $\sigma = \sigma_\ell = \sigma_\ell(\tau, \tau_\ell)$ as the solution of

$$\sigma_\ell(\tau-1)/(\tau_\ell-1) = \tau,$$

we obtain under a sequence of couplings that

$$\begin{aligned}\kappa_{\min}(W^{(1)}, W^{(2)}) = \kappa_{0,\tau}(W^{(1)}, W^{(2)}) &= \min\{W^{(1)}, W^{(2)}\}^{\tau} \\ &= \lim_{\tau_\ell\to\infty} \kappa_{1,\sigma_\ell}(W_\ell^{(1)}, W_\ell^{(2)}),\end{aligned}$$

concluding that under limits the min-kernel is contained in the parameter space $(\tau, \alpha, \sigma, d)$ of i-KSRGs with kernel $\kappa_{1,\sigma}$.

This finishes the introduction to i-KSRGs.

1.4 MAIN CONTRIBUTIONS AND OUTLINE

We proceed to a summary of the the main contributions of the thesis. The thesis studies various structures and processes in/on supercritical scale-free random graphs and is split into three parts: in Part i we study the non-spatial preferential attachment model introduced in Section 1.2.2, while Parts ii and iii focus on (instances of) KSRGs: instances of spatial random graph models.

In the following chapters, we define the studied properties in more and state the results formally, provide additional motivational context from the literature. Eventually, in Chapter 9 we present open questions and future research directions.

Part I: Distances in preferential attachment models.

In Chapter 2, we study three preferential attachment models (PAMs) where the parameters are such that the asymptotic degree distribution has infinite variance, i.e., with parameters such that the power-law exponent $\tau \in (2, 3)$. Each edge is equipped with a non-negative i.i.d. weight. We study the weighted distance between two vertices chosen uniformly at random from the largest component, the typical weighted distance, and the number of edges on this path, the typical hopcount. We prove that there are precisely two universality classes of edge-weight distributions, called the explosive and conservative class.

For edge-weight distributions in the *explosive* class, we show that (for any $\tau \in (2, 3)$) the typical weighted distance converges in distribution to the sum of two i.i.d. *finite* random variables, that correspond to the explosion time (originating from the study of age-dependent branching processes, see [12, 14, 117]) of the local-weak limit.

For edge-weight distributions in the *conservative* class, we prove that the typical weighted distance tends to infinity, and we give explicit expression for the main growth term. This growth term can be tuned to be any function $g = O(\log\log(t))$ by appropriately choosing the edge-weight distribution $L = L(g)$. Under a mild assumption on the edge-weight distribution, the fluctuations around the main term are tight. Our proof techniques allow to obtain tight fluctuations around the main term for the typical weighted distance in the configuration model with power-law exponent $\tau \in (2,3)$ (under the same assumption on the edge-weight distribution), partially proving a conjecture from [2].

Chapter 3 initiates a research line that studies how graph properties defined on a fixed set of vertices evolve as the surrounding preferential attachment graph grows. We consider the *graph-distance evolution*, a discrete-time stochastic process denoted by $\big(d_{t'}^{(G)}(u_t, v_t)\big)_{t' \geqslant t}$. Here, u_t and v_t are two *typical vertices*, i.e., they are sampled uniformly at random from the vertices in the largest component in PA_t. The graph distance $d_{t'}^{(G)}(u_t, v_t)$ is the number of edges on the shortest path between u_t and v_t that uses only vertices that arrived at latest at time t'. We identify a function $K_{t,t'}$ such that $\big(\sup_{t' \geqslant t} \big|d_{t'}^{(G)}(u_t, v_t) - 2K_{t,t'}\big|\big)_{t \geqslant 1}$ forms a tight sequence of random variables. We generalise the result to a setting where each edge is equipped with a non-negative i.i.d. weight.

Part II: Cluster-size decay in kernel-based spatial random graphs.

In this part, we study component sizes in i-KSRGs (introduced in Section 1.3.1). For nearest-neighbour Bernoulli percolation on $\mathbb{Z}^d$ [113] it is well known that when $p > p_c(\mathbb{Z}^d)$ – the critical percolation probability on $\mathbb{Z}^d$– the number of vertices in the cluster containing the origin $|\mathcal{C}(0)|$ satisfies that

$$\mathbb{P}(k \leqslant |\mathcal{C}(0)| < \infty) = \exp\big(-\Theta(k^{\zeta})\big), \tag{1.4.1}$$

with $\zeta = (d-1)/d$. Intuitively, the stretched exponential decay with exponent $(d-1)/d$ reflects that all $\Omega(k^{(d-1)/d})$ edges on the boundary of a cluster $\mathcal{C}$ with $|\mathcal{C}| \geqslant k$ need to be absent: the tail decay in (1.4.1) is determined by *surface tension*. In Chapter 5 and the accompanying paper [151], we study $\mathbb{P}(k \leqslant |\mathcal{C}(0)| < \infty)$ for i-KSRGs where the degree distribution and the edge-length distribution are heavy-tailed. We identify when this *structural inhomogeneity changes the surface-tension driven behaviour of finite clusters*: in those cases, the cluster-size decay in (1.4.1) is still stretched exponential, but the exponent ζ changes. The value of ζ undergoes several phase transitions with respect to the four main model parameters $(\alpha, \tau, \sigma, d)$. Besides establishing cluster-size decay (related to *smaller* components in the graph), we also determine a law of large numbers for the size of the *largest* component when the graph is restricted to the vertices in a finite box of volume n.

Chapter 5 presents the proof for the region in the phase diagram where the model is a generalisation of continuum scale-free percolation and/or hyperbolic random graphs. The component sizes depend heavily on the amount of degree-inhomogeneity *and* the amount of *long-range connections* in the graph. This is in large contrast to the degree distribution, typical distance [64, 67], first-passage percolation [166], and clustering coefficient [40, 60, 91, 135, 217], for which the main scalings are only determined by the degree inhomogeneity and not influenced by the presence of long-range connections when $\tau \in (2, 3)$.

To complement these results, we study a classical model in Chapter 6: the graph formed by supercritical long-range percolation on $\mathbb{Z}^d$ (LRP), where two vertices x, y are connected by an edge with probability proportional to $\|x - y\|^{-\alpha d}$, independently of other edges. Under the (technical) condition that the edge density is sufficiently high, we prove that the cluster-size decay is still determined by surface tension when $\alpha > 1 + 1/d$. The proofs for other regimes of the phase diagram (including LRP with $\alpha < 1 + 1/d$) are presented in an accompanying paper [151] that is not included in the thesis.

Part III: Agent-based modelling: infection spreading.

Motivated by COVID-19, we conduct two simulation-based studies of epidemics on networks: we demonstrate the influence of space and the presence of long-range edges on various characteristics of epidemics (e.g., number of infected or quarantined people over time, the influence of temporal immunity, and the time between two peaks of the epidemic) by comparing them to more traditional models.

We present how the spatial models can be used to model intervention strategies, which gives intuition for qualitative implications of the intervention strategies in Chapter 7, and present how to model contact-tracing applications for infection spread in Chapter 8.

1.5 NOTATION

We conclude this chapter by introducing some notation that is commonly used throughout the thesis.

For $\min\{m, n\}$ and $\max\{m, n\}$ we write respectively $m \wedge n$ and $m \vee n$. Furthermore, $\lceil x \rceil := \min\{y \in \mathbb{Z}, y \geqslant x\}$ and $\lfloor x \rfloor := \max\{y \in \mathbb{Z}, y \leqslant x\}$. For $n \in \mathbb{N}$, the set $\{1, 2, ..., n\}$ is denoted by $[n]$. For two functions $f(x)$

and $g(x)$, we say $f(x) = o\big(g(x)\big)$ if $\limsup_{x\to\infty} |f(x)|/g(x) = 0$, and write $f(x) = O(g(x))$ if $\limsup_{x\to\infty} |f(x)|/g(x) < \infty$.

The complement of an event $\mathcal{E}$ is denoted by $\neg\mathcal{E}$. A sequence of events $(\mathcal{E}_n)_{n\geqslant 1}$ holds with high probability (whp) if $\lim_{n\to\infty} \mathbb{P}(\mathcal{E}_n) = 1$. Let $(X_n)_{n\geqslant 0}$ and $(Y_n)_{n\geqslant 0}$ be two sequences of random variables. We say that a random variable X_0 dominates a random variable Y_0 if there exists a coupling $(\hat{X}_0, \hat{Y}_0)$ such that $\mathbb{P}(\hat{X}_0 \geqslant \hat{Y}_0) = 1$. Similarly, the sequence $(X_n)_{n\geqslant 0}$ dominates $(Y_n)_{n\geqslant 0}$ if there exists a coupling of the sequences such that $\mathbb{P}(\forall_{n\geqslant 0} : \hat{X}_n \geqslant \hat{Y}_n) = 1$. A random graph $G = (V_G, E_G)$ dominates a random graph $H = (V_H, E_H)$ if there exists a coupling such that $\mathbb{P}\big(\hat{V}_H \supseteq \hat{V}_G, \hat{E}_H \supseteq \hat{E}_G\big) = 1$. If a random object X dominates Y, we write $X \succcurlyeq Y$. We say that $(X_n)_{n\geqslant 0}$ converges in probability to a random variable X_∞, i.e., $X_n \xrightarrow{\mathbb{P}} X_\infty$, if for all $\varepsilon > 0$ it holds that $\mathbb{P}(|X_n - X_\infty| > \varepsilon) = o(1)$.

If in a graph G there is an edge incident to both u and v, we write $u \leftrightarrow_G v$, and $u \not\leftrightarrow_G v$ otherwise. We leave out the subscript G if the graph is clear from the context. Similarly, for a set of vertices S, we write $u \leftrightarrow S$ if there is a vertex v in S such that $u \leftrightarrow v$. Moreover, for a set of vertices $(\pi_i)_{i\leqslant n} \subseteq V_G$ we write $\{\pi_0 \leftrightarrow \cdots \leftrightarrow \pi_n\} := \{\pi_0 \leftrightarrow \pi_1\} \cap \cdots \cap \{\pi_{n-1} \leftrightarrow \pi_n\}$. The sequence (or path) $(\pi_i)_{i\leqslant n}$ is called self-avoiding if $\pi_i \neq \pi_j$ for all $i \neq j$.

Part I
Preferential attachment models: Distances

2 WEIGHTED DISTANCES

Based on [154]:
Weighted distances in scale-free preferential attachment models,
J. Jorritsma, J. Komjáthy,
Random Structures & Algorithms, 2020(57): 823– 859.

2.1 INTRODUCTION

First-passage percolation is a model that was introduced to describe the flow of a fluid through a porous random medium [114]. One starts with a graph $G = (V, E)$ and equips each edge $e \in E$ with an i.i.d. copy of a random weight, that represents the time for a fluid to "cross" the edge. It can also be viewed as the time to transmit a message or disease in graphs that represent real-life networks. Natural questions for FPP are amongst others:

(i) For two vertices u and v, what is the transmission time of a message from u to v? How does the passage time depend on the size of the graph?

(ii) How many edges are on the shortest weighted path from u to v? In telecommunication networks, the signal loss increases in the number of edges on a path.

In the last decades, first passage percolation on random graphs has gained increasing attention, and the process is quite well understood on *static* graphs, i.e., graphs that do not grow over time. It has been studied on the Erdős-Rényi graph in [27], on configuration models (CM) with finite variance degrees in [26, 28, 29, 74], and with infinite variance degrees in [2, 22, 21, 74]. FPP on spatial models as scale-free percolation, geometric inhomogeneous random graphs, and hyperbolic random graphs is studied for infinite variance degrees in [136, 166]. In this chapter, we focus on FPP on a dynamically growing model, the preferential attachment models introduced in Section 1.2.2 (FPA, VPA, and GVPA). To the extent of our knowledge, no formal results are known for FPP with non-trivial weights on *dynamically growing* models such as PA, except for the results in Part i in this thesis. However, if all weights are equal to one, then the time to transmit a message between two vertices u_t, v_t, sampled uniformly at random from the largest component in the graph at time t, corresponds to the *typical graph distance* $d_t^{(G)}(u_t, v_t)$. When the variance of the asymptotic degree distribution is infinite, i.e., when the power-law exponent $\tau \in (2,3)$, precise asymptotics are known as shown in the next theorem.

Theorem 2.1.1 (Graph distances for PA [46, 69, 81])**.** *Consider* PA_t *with power-law exponent* $\tau \in (2,3)$*. For all* $\varepsilon > 0$*, there exists a constant* $M > 0$ *such that for all* t *sufficiently large*

$$\mathbb{P}\big(2K_t^* - M \leqslant d_t^{(G)}(u_t, v_t) \leqslant 2(1+\varepsilon)K_t^*\big) \geqslant 1 - \varepsilon, \tag{2.1.1}$$

where

$$K_t^* := \lfloor 2 \log\log(t)/|\log(\tau - 2)| \rfloor .$$

In [190] it is shown that for VPA the upper bound in (2.1.1) can be improved to $2K_t^* + M$. Corollary 2.2.7 below shows that this upper bound can also be improved to $2K_t^* + M$ for FPA. Little is known for graph distances in FPA and GVPA when $\tau > 3$. For FPA it is shown in [81] that the diameter of the graph and the typical graph distance are of order $\Theta(\log(t))$, but the precise main order of growth remains unknown. This is in sharp contrast with the configuration model, where the graph distance grows as $\Theta(\log(t))$ and the precise order is found in [131].

We introduce *edge-weighted* preferential attachment models.

Definition 2.1.2 (Edge-weighted preferential attachment models)**.** Let L be a non-negative random variable with cumulative distribution function $F_L(x) := \mathbb{P}(L \leqslant x)$. In the edge-weighted versions of FPA and (G)VPA from Definitions 1.2.4 and 1.2.5, respectively, each edge e is equipped on its creation with an i.i.d. copy of L, denoted by L_e.

We assume throughout the chapter that indexed weight random variables with different indices are i.i.d. Let $d_t^{(L)}(u_t, v_t)$ denote *typical weighted distance*, the sum of the weights along the least-weighted path from u_t to v_t, two vertices that are sampled uniformly at random from the largest component in PA_t. We show in this chapter that there are exactly two universality classes of weight distributions for the three edge-weighted preferential attachment models defined in Section 1.2.2 when $\tau \in (2,3)$.

The universality classes are determined by a computable characteristic of the weight distribution L. We present the characteristic and an informal version of our main result here, precise results can be found in Theorems 2.2.5 and 2.2.10 below. For a random variable L, we define its cumulative distribution function as $F_L(x) := \mathbb{P}(L \leqslant x)$, and its generalised inverse by $F^{(-1)}(y) := \inf_x\{x \in \mathbb{R} : F(x) \geqslant y\}$.

Definition 2.1.3 (Explosion characteristic $I_1(L)$)**.** Let L be a non-negative random variable with distribution function F_L. We define the explosion characteristic $I_1(L)$ as

$$I_1(L) := \sum_{k=1}^{\infty} F_L^{(-1)}\left(e^{-e^k}\right). \tag{2.1.2}$$

We call $\{L : I_1(L) = \infty\}$ the conservative class, and $\{L : I_1(L) < \infty\}$ the explosive class.

The term *explosion* originates from the study of age-dependent branching processes, see e.g. [12, 14, 117]. In these branching processes we say that explosion happens if infinitely many individuals are born within finite time. The relation to explosion in trees comes from the fact that the neighbourhood of a typical vertex in PA_t converges in distribution to a random tree, the local weak limit. Local weak convergence is shown for the three models in [24, 71, 97]. It is interesting in its own right to study the edge-weighted version of the local weak limit tree. We prove that infinitely many vertices are within finite weighted distance from the root in the local weak limit if the weight distribution is in the explosive class. This fact is then used to show convergence in distribution for the typical weighted distance in PA_t if $I_1(L) < \infty$, the first part of our following main result.

Theorem 2.1.4 (Meta theorem)**.** *Consider* PA_t *with power-law exponent* $\tau \in (2, 3)$*. If* $I_1(L) < \infty$*, then the typical weighted distance converges in distribution to an almost surely finite random variable. If* $I_1(L) = \infty$*, then*

$$d_L^{(t)}(u, v) \approx 2Q_t + o_{\mathbb{P}}(Q_t),$$

where

$$\begin{aligned} K_t^* &= \lfloor 2 \log\log(t)/|\log(\tau - 2)| \rfloor, \\ Q_t &:= \sum_{k \leqslant K_t^*} F_L^{(-1)}\left(\exp\left(-(\tau-2)^{-k/2}\right)\right). \end{aligned} \tag{2.1.3}$$

Under a mild extra condition on L*, the error term* $o_{\mathbb{P}}(Q_t)$ *is tight.*

This meta theorem is formalized below in Section 2.2.1. There, we describe the limiting random variables if $I_1(L) < \infty$ and state the computable, mild condition on L that yields tight error terms. As a side result of the second part of Theorem 2.1.4, when $I_1(L) = \infty$, we show that if the weights are of the form $1+X$, $I(X) < \infty$, the typical weighted distance and typical hopcount are both tight around $4\log\log(t)/|\log(\tau-2)|$ for the models FPA and VPA. This indicates that the addition of an *excess* edge-weight X does not affect the topology of the shortest paths drastically. Constant weights are a special case of these. So, our result extends results from [46, 69, 81, 190], by showing that the fluctuations of the typical graph distance around $4\log\log(t)/|\log(\tau-2)|$ are tight.

Organisation

The next section formally introduces the necessary concepts to describe the limiting random variables for the explosive case. Afterwards, we state our main results, and discuss them by formulating some open problems and recalling relevant results from literature. In Section 2.3, we prove upper bounds for the weighted distance in the finite graphs by constructing a path, and show that the local weak limit tree is explosive if and only if the edge-weight distribution L is a member of the explosive universality class. Then, in Section 2.4, we prove the corresponding lower bounds for both the conservative as the explosive regime. In Section 2.5 we prove a theorem on the hopcount. Lastly, in Section 2.6 we extend the results for conservative distributions on finite graphs to the local weak limit.

2.2 DEFINITIONS AND MAIN RESULTS

In this chapter we look at *typical least weighted paths*, that is, we assume that every edge in PA_t is equipped with an i.i.d. weight, and we are interested in the sum of the weights on the least weighted path between two vertices, and the number of edges on this path, called *hopcount*.

Definition 2.2.1 (Distances in graphs)**.** Consider the graph $\mathcal{G} = (V, E)$ and assume every edge $e \in E$ is equipped with a weight L_e. For a path π, we define its length as $\|\pi\| := \sum_{e\in\pi} 1$ and L-length as $\|\pi\|_L := \sum_{e\in\pi} L_e$. For $u, v \in V$, let $\Omega_{u,v} := \{\pi : \pi \text{ is a path from } u \text{ to } v \text{ in } \mathcal{G}\}$. We define the distance, L-distance (also called weighted distance), and H-distance (hopcount) between u and v in the graph $\mathcal{G}$ as

$$\begin{aligned} d_{\mathcal{G}}^{(G)}(u,v) &:= \min_{\pi\in\Omega_{u,v}} \|\pi\|, \\ d_{\mathcal{G}}^{(L)}(u,v) &:= \min_{\pi\in\Omega_{u,v}} \|\pi\|_L, \\ d_{\mathcal{G}}^{(H)}(u,v) &:= \Big\| \arg\min_{\pi\in\Omega_{u,v}} \|\pi\|_L \Big\|, \end{aligned}$$

respectively. If $\Omega_{u,v} = \emptyset$, the above distance-metrics are defined as ∞. If there are several paths $\pi_1, \dots, \pi_k$ minimizing the L-distance, the hopcount is defined as $\min_{i\leqslant k} \|\pi_i\|$. For a letter $\Box \in \{G, L, H\}$, the typical $\Box$-distance of a graph $\mathcal{G}$ is defined as the $\Box$-distance between two typical vertices. For $q \in V$ and a set $A \subseteq V$, we generalise distances and define the $\Box$-diameter of A by

$$d_{\mathcal{G}}^{(\Box)}(q,A) := \min_{w\in A} d_{\mathcal{G}}^{(\Box)}(q,w), \qquad \mathrm{diam}_{\mathcal{G}}^{(\Box)}(A) := \max_{x,y\in A} d_{\mathcal{G}}^{(\Box)}(x,y).$$

For a vertex q, its $\Box$-neighbourhood with radius $r > 0$ and its boundary are defined as

$$\mathcal{B}_{\mathcal{G}}^{(\Box)}(q,r) := \{w : d_{\mathcal{G}}^{(\Box)}(q,w) \leqslant r\}, \quad \partial\mathcal{B}_{\mathcal{G}}^{(G)}(q,r) := \{w : d_{\mathcal{G}}^{(G)}(q,w) = \lfloor r \rfloor\}.$$

We write $\widetilde{\mathcal{B}}_{\mathcal{G}}^{\Box}(q,r)$ for the induced subgraph of $\mathcal{G}$ on the vertex set $\mathcal{B}_{\mathcal{G}}^{\Box}(q,r)$, with edges (u,v) from $\mathcal{G}$ if both u and v are in $\mathcal{B}_{\mathcal{G}}^{\Box}(q,r)$. If $\mathcal{G} = \mathrm{PA}_t$, we abbreviate (PA_t) by t in the subscript.

An alternative way to look at an edge-weighted graph is to view the weights as *passage times*, i.e., the time that it takes to send a message from one side of the edge to the other. The notions of time and weight are used interchangeably. We stress that the passage time of a single edge is not related to the time t in the construction of the graph PA_t.

We introduce the concepts of explosion time and local weak limit to describe the limiting random variables for the typical weighted distance in the explosive class in Theorem 2.2.10 below. We recall the definitions of the graph neighbourhood $\mathcal{B}_{\mathcal{G}}^{(G)}$ and weighted neighbourhood $\mathcal{B}_{\mathcal{G}}^{(L)}$ for a graph $\mathcal{G}$ from Definition 2.2.1.

Definition 2.2.2 (Explosive graph)**.** Let $\mathcal{G} = (V, E)$ be a weighted graph that is locally finite. For the time to reach graph distance k and the time to its n-th closest vertex in L-distance from a vertex q, we write

$$\beta_{\mathcal{G},k}(q) := d_{\mathcal{G}}^{(L)}\big(q, \partial\mathcal{B}_{\mathcal{G}}^{(G)}(q,k)\big), \quad \sigma_{\mathcal{G},n}(q) := \inf\big\{r : |\mathcal{B}_{\mathcal{G}}^{(L)}(q,r)| \geqslant n\big\}.$$

If $|V| = \infty$, we define the explosion time of q as

$$\beta_{\mathcal{G},\infty}(q) := \lim_{k\to\infty} \beta_{\mathcal{G},k}(q).$$

If there is a $q \in V$ with finite explosion time, then we call $\mathcal{G}$ explosive. For $\mathcal{G} \equiv \mathrm{PA}_t$, we write $\beta_{t,k}$ and $\sigma_{t,n}$ if the q in PA_t is a typical vertex, i.e., it is sampled uniform at random from the largest component in the graph at time t. If $\mathcal{G}$ is a tree rooted in $\circledcirc$, we abbreviate $\beta_{\mathcal{G},k} := \beta_{\mathcal{G},k}(\circledcirc)$.

The local weak limit of graphs can be used to describe the neighbourhood of a typical vertex. For an introduction we refer to [125, Chapter 2] and its references. Let $\mathbb{G}_\star$ be the space of all (possibly infinite) rooted graphs.

Definition 2.2.3 (Local weak limit in probability)**.** Let $(\mathcal{G}_t)_{t\geqslant 0}$ be a sequence of finite random rooted graphs, and let $(\mathcal{G}, q)$ be a rooted random graph following law μ. The sequence $(\mathcal{G}_t)_{t\geqslant 0}$ converges in probability in the local weak convergence sense to $(\mathcal{G}, q)$, when

$$\mathbb{E}_t[h(\mathcal{G}_t, q_t)] \xrightarrow{\mathbb{P}} \mathbb{E}[h(\mathcal{G}, q)],$$

for every bounded and continuous function $h : \mathbb{G}_\star \to \mathbb{R}$, where the expectation on the rhs is w.r.t. $(\mathcal{G}, q)$ having law μ, while the expectation on the lhs is w.r.t. the typical vertex q_t only.

Berger *et al.* [24] identify the local weak limit of FPA. They give an explicit construction of the limit that they call the *Pólya-point graph* (PPG), an infinite rooted tree derived from a multi-type branching process. While the construction of FPA in [24] is slightly different from Definition 1.2.4, it can be related to our model for $\delta \geqslant 0$. In [97, Chapter 4], it is shown that the result remains valid for a wider class of models, in particular when $\delta < 0$.

Turning to the local weak limit of GVPA, Dereich and Mörters [71] introduce a similar concept for the GVPA-model, the *idealized neighbourhood tree* (INT). While local weak convergence is only stated briefly before [71, Theorem 1.8], they construct a coupling similar to the PPG. Contrary to FPA, the graph (G)VPA model is not asymptotically almost surely connected, and the INT could consist of finitely many vertices. Throughout the remainder, we assume that the INT is conditioned to consist of infinitely many vertices. For our proofs it is not important how the local weak limits can be constructed, only that they exist. In fact, we consider the PPG and INT as a *black box* and yet obtain results. If a statement holds for both models, we refer to the INT or PPG as LWL (local weak limit). We write LWL_k for the tree restricted to vertices that have graph distance at most k from the root. We call the vertices that are at graph distance exactly k away from the root the k-th generation of the LWL. We state the combined result on local weak convergence for the reader's convenience. We call two rooted graphs $(\mathcal{G}, x)$, $(\mathcal{G}', x')$ rooted isomorphic, and write $(\mathcal{G}, x) \simeq (\mathcal{G}', x')$, if there exists an isomorphism from $\mathcal{G}$ to $\mathcal{G}'$ that maps x to x'. Recall the graph neighbourhood $\widetilde{\mathcal{B}}_G$ from Definition 2.2.1.

Proposition 2.2.4 (Local weak convergence [24, Theorem 2.2, Proposition 3.6], [71, Section 5, 6])**.** *The local weak limits of PA are the Pólya-point graph for* $\mathrm{FPA}(m, \delta)$, *and the idealized neighbourhood tree for* $\mathrm{GVPA}(f)$. *Moreover, let* q *be a typical vertex, then for all* $\delta_{2.2.4} > 0$ *there exists a function* $\kappa_{\delta_{2.2.4}}(t)$ *that tends to infinity with* t, *such that* $\mathcal{B}_t^{(G)}(q, \kappa_{\delta_{2.2.4}}(t))$ *and* $\mathrm{LWL}_{\kappa_{\delta_{2.2.4}(t)}}$ *can be coupled, such that, denoting by* $\circledcirc$ *the root of the* LWL,

$$\mathbb{P}\big(\widetilde{\mathcal{B}}_t^{(G)}(q, \kappa_{\delta_{2.2.4}}(t)) \simeq \mathrm{LWL}_{\kappa_{\delta_{2.2.4}(t)}}(\circledcirc)\big) \geqslant 1 - \delta_{2.2.4}, \qquad (2.2.1)$$

2.2.1 *Main results*

Recall the explosion characteristic $I_1(L)$ from (2.1.2). We start with the results on the typical weighted distance in PA_t, where the edge-weight distribution satisfies $I_1(L) = \infty$. In this case, we show that the typical weighted distance tends to infinity as the graph size tends to infinity. We determine the first order of growth and the number of edges used on this path. For FPA and VPA, we strengthen our results by showing that the fluctuations around the first order term are tight under a mild condition on L. Recall K_t^* and Q_t from (2.1.3).

Theorem 2.2.5 (Weighted distance, conservative case)**.** *Consider PA with power-law exponent* $\tau \in (2, 3)$, *i.i.d. weights on the edges with distribution* F_L *satisfying* $I_1(L) = \infty$. *Let* u, v *be two typical vertices. Then, for the typical weighted distance in* PA_t,

$$d_t^{(L)}(u, v)/2Q_t \xrightarrow{\mathbb{P}} 1, \qquad \textit{as } t \to \infty. \qquad (2.2.2)$$

Moreover, for the models FPA and VPA from Definition 1.2.4 and 1.2.5, if $I_1(L) = \infty$ *and* F_L *satisfies*

$$I_2(L) := \sum_{k=1}^{\infty} \frac{1}{k}\left(F_L^{(-1)}\left(e^{-e^k}\right) - \sup\{x : F_L(x) = 0\}\right) < \infty, \qquad (2.2.3)$$

then

$$\left(d_t^{(L)}(u,v) - 2Q_t\right)_{t\geqslant 1}$$

forms a tight sequence of random variables, i.e., the fluctuations are of order $O(1)$ *whp.*

We believe that (2.2.3) is only a technical condition. Only distributions L that are extremely *flat* around the origin (triple exponentially) violate it. An artificial example of such a distribution is if F_L in the neighbourhood of 0 satisfies

$$F_L(x) = \exp\left(-\exp\left(e^{x^{-\beta}}\right)\right)$$

for some $\beta \geqslant 1$. If F_L satisfies this equality for some $\beta \in (0,1)$, then condition (2.2.3) is satisfied.

We proceed to the typical hopcount for a class of conservative weight distributions.

Theorem 2.2.6 (Hopcount, weights bounded away from zero)**.** *Consider PA with power-law exponent* $\tau \in (2,3)$*, i.i.d. weights on the edges with distribution* F_L *satisfying* $a := \sup\{x : F_L(x) = 0\} > 0$[1]*. Let* u, v *be two typical vertices. Then, for the typical hopcount in* PA_t

$$d_t^{(H)}(u,v)/2K_t^* \xrightarrow{\mathbb{P}} 1, \qquad \textit{as } t \to \infty. \tag{2.2.4}$$

Moreover, for the models FPA and VPA from Definition 1.2.4 and 1.2.5, if $I(L-a) < \infty$*, then*

$$\left(d_t^{(H)}(u,v) - 2K_t^*\right)_{t\geqslant 1} \tag{2.2.5}$$

forms a tight sequence of random variables.

Setting the weights $L \equiv 1$ in Theorem 2.2.6 immediately implies the following corollary, extending results in [46, 69] on the typical graph distance in FPA up to tight error terms, and confirming the tight error terms for VPA from [190].

Corollary 2.2.7 (Tight graph distances)**.** *Consider FPA or VPA with power-law exponent* $\tau \in (2,3)$*. Let* u, v *be two typical vertices. Then, for the typical graph distance in* PA_t

$$\left(d_t^{(G)}(u,v) - 2K_t^*\right)_{t\geqslant 1}$$

forms a tight sequence of random variables.

---

1 This implies that $I_1(L) = \infty$.

Before we move on to the results on finite graphs for weight distributions satisfying $I_1(L) < \infty$, we discuss first passage percolation on the LWL. The following theorems show that the LWL is explosive if and only if $I_1(L) < \infty$. We start with conservative edge weights. Afterwards we prove explosiveness of the LWL for explosive edge weights. This is then used to state the last theorem on the weighted distances in PA_t for explosive edge weights.

Theorem 2.2.8 (FPP on the LWL, conservative case)**.** *Consider a PPG or INT rooted in* $\circledcirc$ *with power-law exponent* $\tau \in (2,3)$ *with i.i.d. weights on the edges with distribution* F_L *satisfying* $I_1(L) = \infty$*. If* F_L *satisfies* (2.2.3)*, then, for FPA and VPA,*

$$\left(\beta_{\mathrm{LWL},k}(\circledcirc) - \sum_{i=1}^{k} F_L^{(-1)}\left(\exp\left(-(\tau-2)^{-i/2}\right)\right)\right)_{k\geqslant 1} \tag{2.2.6}$$

is a tight sequence of random variables. Regardless of (2.2.3)*, for FPA, VPA, and GVPA, as* k *tends to infinity,*

$$\beta_{\mathrm{LWL},k}(\circledcirc) \Big/ \sum_{i=1}^{k} F_L^{(-1)}\left(\exp\left(-(\tau-2)^{-k/2}\right)\right) \xrightarrow{a.s.} 1. \tag{2.2.7}$$

Observe the similarities between Theorem 2.2.5 and Theorem 2.2.8. In fact, the proof of Theorem 2.2.8 heavily relies on couplings between the LWL and PA_t, which are possible by Proposition 2.2.4. These couplings allow for the intermediate lemmas and propositions to consider whichever object, i.e., PA_t vs. LWL, is more suitable and lead to the similarities between the two theorems.

We now proceed to the universality class of weight distributions satisfying $I_1(L) < \infty$. This holds for most *well-known* distributions with support starting at 0, e.g. the exponential distribution. After stating that for these weight distributions the LWL is explosive, we proceed with a theorem on the typical weighted distance in finite graphs. Recall Definition 2.2.2 of an explosive graph.

Theorem 2.2.9 (FPP on the LWL, explosive case)**.** *Consider a PPG or INT rooted in* $\circledcirc$ *with power-law exponent* $\tau \in (2,3)$ *with i.i.d. weights on the edges with distribution* F_L *satisfying* $I_1(L) < \infty$*. Then the explosion time of the LWL is an almost surely finite random variable, i.e.,*

$$\mathbb{P}\left(\beta_{\mathrm{LWL},\infty} < \infty\right) = 1.$$

Theorem 2.2.10 (Weighted distance, explosive case)**.** *Consider PA with power-law exponent $\tau \in (2,3)$, i.i.d. weights on the edges with distribution F_L satisfying $I_1(L) < \infty$. Let u, v be two typical vertices. Then, for the typical weighted distance in PA_t*

$$d_t^{(L)}(u,v) \xrightarrow{d} \beta_\infty^{(1)} + \beta_\infty^{(2)}, \qquad \text{as } t \to \infty,$$

where $\beta_{\mathrm{LWL},\infty}^{(1)}$ and $\beta_{\mathrm{LWL},\infty}^{(2)}$ are two i.i.d. copies of the explosion time of the LWL.

It is remarkable that the limiting random variable does *not* depend on t and thus the graph distance is of much larger order than the weighted distance. The underlying intuition is that in the graph neighbourhoods of u and v there is a vertex with sufficiently high degree. The weighted distances to these vertices converge in distribution to $\beta_{\mathrm{LWL},\infty}^{(u)}$ and $\beta_{\mathrm{LWL},\infty}^{(v)}$. There are many paths connecting these high degree vertices, where the number of edges on these paths is *similar* to the graph distance, allowing to bound its total weight from above and show that it tends to zero.

2.3 UPPER BOUND ON THE WEIGHTED DISTANCE

In this section we prove the upper bounds for Theorems 2.2.5 and 2.2.10, respectively. The upper bound of Theorem 2.2.8 which follows from the same proof techniques, is postponed to Section 2.6. Recall $I_1(L)$ from (2.1.2), and Q_t from (2.1.3).

Proposition 2.3.1 (Upper bound on the weighted distance, conservative case)**.** *Consider PA under the same conditions as Theorem 2.2.5. Recall $I_1(L) = \infty$. Then for every $\delta, \varepsilon > 0$, when t is sufficiently large*

$$\mathbb{P}\left(d_t^{(L)}(u,v) \leqslant (1+\varepsilon)2Q_t\right) \geqslant 1-\delta.$$

Moreover, for the models FPA and VPA from Definition 1.2.4 and 1.2.5, if F_L satisfies (2.2.3), there exists a constant $M_{2.4.1} = M_{2.4.1}(\delta)$ such that for t sufficiently large

$$\mathbb{P}\left(d_t^{(L)}(u,v) \leqslant 2Q_t + 2M_{2.4.1}\right) \geqslant 1-\delta.$$

Proposition 2.3.2 (Upper bound on the weighted distance, explosive case)**.** *Consider PA under the same conditions as Theorem 2.2.10. Recall $I_1(L) < \infty$. Then, there is a coupled probability space, such that for every $\delta, \varepsilon > 0$ there exists a constant $N \in \mathbb{N}$ such that for t sufficiently large*

$$\mathbb{P}\left(d_t^{(L)}(u,v) \leqslant \beta_{\mathrm{LWL}^{(u)},N} + \beta_{\mathrm{LWL}^{(v)},N} + \varepsilon\right) \geqslant 1-\delta, \qquad (2.3.1)$$

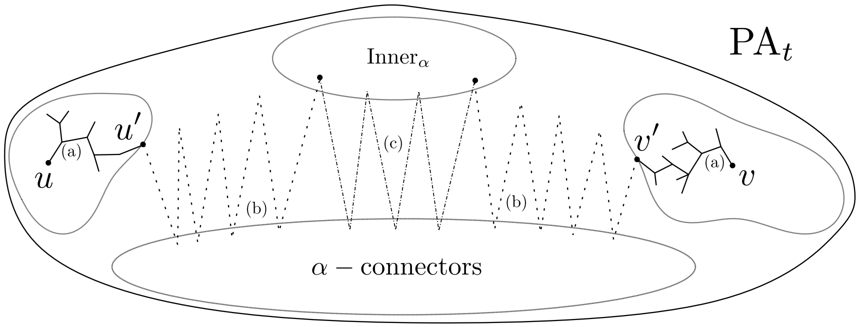


Figure 1: The constructed five-segment path from u to v, via vertices with sufficiently high degree u' and v' and Inner$_\alpha$.

where $\mathrm{LWL}^{(u)}, \mathrm{LWL}^{(v)}$ *are the LWL trees coupled to the neighbourhood of* u, v, *respectively.*

Throughout this section, we look at the graph at times t and $t' := \alpha t$, for some $\alpha \in [\frac{1}{2}, 1)$. For the upper bound on $d_t^{(L)}(u, v)$ it is enough to construct a path between u and v, and study its weight. The path that we construct, consists of five segments of three different types. First, for $q \in \{u, v\}$, we construct a path, consisting of one segment of type (a) and one of type (b), to

$$\mathrm{Inner}_\alpha := \{w \in [\alpha t] : D_{\alpha t}(w) \geqslant (\alpha t)^{\frac{1}{2(\tau-1)}} \log(\alpha t)^{-\frac{1}{2}}\}, \tag{2.3.2}$$

i.e., vertices with a very large degree, also called inner core, see Figure 1. The total weight of one path contributes almost half of the total weight of the entire path from u to v. The segment of type (a) connects $q \in \{u, v\}$ to a vertex q' that has degree at least $s_0 \in \mathbb{N}$. It only passes through vertices that arrived before time αt, *old* vertices. To do so, we shall condition on $q < \alpha t$, which happens with probability close to one if α is close to one. The segment of type (b) connects q' to the inner core and alternatingly passes through old vertices and *α-connectors*, vertices that arrived after time αt. Similarly for segment type (c), see Figure 1, we construct a path with negligible total weight that connects two vertices in the inner core, by alternatingly using α-connectors and other vertices in the inner core. By construction, all edges on segments of type (a) arrived *before* time αt, while on types (b) and (c) all edges arrived *after* αt.

We present now three segment-specific propositions, used in both the explosive and conservative case. Afterwards we introduce some necessary notation to construct the path. Lastly, we show how these propositions together prove Propositions 2.3.1 and 2.3.2. Starting with segment (a), we show that the number of edges on the path between q and q' is bounded, for $q \in \{u, v\}$. In the conservative case, its total weight is negligible compared to Q_t. In the explosive case, this part is the main contributor and later we show that its total weight tends in distribution to the (finite) explosion time of the LWL.

Proposition 2.3.3 (Bounded graph distance to a vertex with degree at least s)**.** *Consider PA with power-law exponent $\tau \in (2,3)$, i.i.d. weights on the edges with distribution F_L and fix $\delta_{2.3.3} > 0$. Let q be chosen uniformly at random from $[t]$. For any $s_{2.3.3} \in \mathbb{N}$, there is a constant $C_{2.3.3} = C_{2.3.3}(s_{2.3.3}, \delta_{2.3.3})$ such that for t sufficiently large*

$$\mathbb{P}\Big(\bigcap_{q' \in [t]: D_t(q') \geqslant s_{2.3.3}} \big\{ d_t^{(G)}(q, q') \geqslant C_{2.3.3} \big\} \Big) \leqslant \delta_{2.3.3}.$$

We write $\mathcal{E}_{2.3.3}^{(t)}(q, s_{2.3.3})$ for the complement of the above event between brackets.

The proof for FPA follows from a minor adaptation of the proof of [81, Theorem 3.6]. For GVPA it follows from an adaptation of [190, Proposition 5.10]. We refer the reader to the cited paper and thesis to fill in the details. We emphasize that we apply Proposition 2.3.3 at time αt, rather than t.

From the vertex q' with degree at least s_0 at time αt, we construct a path to the inner core, corresponding to segment (b). We show that there are many such paths, allowing to bound the weight. The next proposition is the main (technical) contribution of the chapter. Due to this statement, we obtain tight bounds for the various distances in FPA, improving upon existing results [46, 69]. Its proof can easily be adapted to obtain tight fluctuations on the typical weighted distance in the configuration model if condition (2.2.3) is satisfied, improving results in [2].

Proposition 2.3.4 (Upper bound on the weighted distance to Inner_α)**.** *Consider PA with power-law exponent $\tau \in (2,3)$, i.i.d. weights on the edges with distribution F_L. Fix $\delta_{2.3.4}, \varepsilon_{2.3.4} > 0$, $\alpha \in [1/2, 1)$. There exists $s_0 = s_0(\delta_{2.3.4}) \in \mathbb{N}$ and a constant $M_{2.3.4}$, such that for $s > s_0$, and any $q' \in [\alpha t]$ with $D_{\alpha t}(q') = s$, if t is sufficiently large, for FPA or VPA, if F_L satisfies* (2.2.3),

$$\mathbb{P}\Big(d_t^{(L)}(q', \mathrm{Inner}_\alpha) \geqslant M_{2.3.4} + \sum_{k=\lfloor h_\tau(s) \rfloor}^{K_t^* + \lfloor h_\tau(s) \rfloor + 4} F_L^{(-1)}\big(\exp\big(-(\tau-2)^{-k/2}\big)\big)\Big) \\ \leqslant \delta_{2.3.4}, \qquad (2.3.3)$$

where $h_\tau(s) = 2\log\log(s)/|\log(\tau-2)| + c_\tau$ *for some constant* c_τ. *Without* (2.2.3), *it holds that*

$$\mathbb{P}\Big(d_t^{(L)}(q', \mathrm{Inner}_\alpha) \geqslant (1+\varepsilon_{2.3.4}) \sum_{k=\lfloor h_\tau(s)\rfloor}^{K_t^*+\lfloor h_\tau(s)\rfloor+4} F_L^{(-1)}\big(\exp\big(-(\tau-2)^{-k/2}\big)\big)\Big) \leqslant \delta_{2.3.4}, \qquad (2.3.4)$$

where the constant c_τ *in the function* h_τ *might be different and can depend on* $\varepsilon_{2.3.4}$.

This part of the path is the main contributor to the upper bound in the conservative case. If $I_1(L) < \infty$, it follows that the value of the sum in (2.3.3) can be made arbitrarily small by increasing s, because $h_\tau(s)$ tends to infinity. For the conservative case, comparing the sum in Q_t in (2.1.3) to the sum in (2.3.3), one sees that they are identical up to a shift of the summation boundaries.

In the next proposition we bound the graph *and* weighted distance within the inner core, segment (c) in Figure 1.

Proposition 2.3.5 (Inner core has negligible weighted distance)**.** *Consider PA with power-law exponent* $\tau \in (2,3)$, *i.i.d. weights on the edges with distribution* F_L. *Recall* Inner_α *from* (2.3.2) *and fix* $\delta_{2.3.5} > 0, \alpha \in [1/2, 1)$. *Then there exists* $C_{2.3.5} > 0$, *such that for all sufficiently large* t, *and for any two fixed vertices* w_1, w_2 *in* Inner_α,

$$\mathbb{P}\left(d_t^{(G)}(w_1, w_2) \geqslant C_{2.3.5}\right) \leqslant \delta_{2.3.5}. \qquad (2.3.5)$$

Moreover, if F_L *satisfies* $F_L(x) > 0$ *for all* $x > 0$[2], *then for any* $\varepsilon_{2.3.5} > 0$, *if* t *is sufficiently large,*

$$\mathbb{P}\left(d_t^{(L)}(w_1, w_2) \geqslant \varepsilon_{2.3.5}\right) \leqslant \delta_{2.3.5}. \qquad (2.3.6)$$

Proof of Proposition 2.3.5. We give a proof similar to [81, Proposition 3.2]. We construct a path from w_1 to w_2 via a subset of the inner core, in which we can bound the weighted distance between two vertices whp. We show the latter first, after which we show that if w_1 or w_2 is not contained in this subset, the (weighted) distance to this subset is also small.

By Lemma 2.3.9, there are at least $n_t = \lfloor\sqrt{t}\rfloor$ vertices in the inner core. Let I be the set of the first n_t vertices that have degrees at least

2 This constraint holds whenever $I_1(L) < \infty$, but might not hold when $I_1(L) = \infty$.

$(\alpha t)^{1/(2(\tau-1))}\log(\alpha t)^{-1/2}$. We construct a graph H_t on these vertices as follows. Recall the definition of an α-connector from Definition 2.3.6. Let i, j be connected in H_t if there exists an α-connector y. The weight on the edge (i, j) is $L_{(i,y)} + L_{(j,y)}$. As explained in the proof of [81, Proposition 3.2] for the model FPA, H_t stochastically dominates a dense uniform Erdős-Rényi graph $G(n_t, p_t)$, where

$$p_t := \frac{t^{\frac{1}{\tau-1}-1}}{2\log^2 t}.$$

Here, we say that a random graph G dominates a random graph H if there exists a coupling such that every edge in H is also contained in G. Using [70, Theorem 1.1], one can verify that the same holds for GVPA. In [36, Chapter 10.2] the diameter of the dense ERRG is discussed, and it is shown that the diameter is bounded. Hence, the first assertion (2.3.5) follows. From now on, we assume that $F_L(x) > 0$ for all $x > 0$. Let $\Delta = \Delta(\tau)$ denote the diameter of the $G(n_t, p_t)$. The proof techniques in the above mentioned book chapter rely on the exploration around two vertices. In particular, it can be derived that the number of disjoint paths between two vertices of length Δ tends to infinity with the size of the graph. Hence, there is a function r_t tending to infinity with t, such that there are at least r_t disjoint paths between u and v. The weight on the i-th path is distributed as $L_\star^{(i)} := L_1^{(i)} + \ldots + L_{2\Delta}^{(i)}$. As $F_L(x) > 0$ for any $x > 0$, the same holds for $F_{L_\star}$. Thus,

$$\lim_{t\to\infty} \mathbb{P}\big(\min_{i\in r_t} L_\star^{(i)} > \varepsilon\big) = 0,$$

for any $\varepsilon > 0$. Hence, for $w_1', w_2' \in I$ and any $\varepsilon_{2.3.5}, \delta_{2.3.5} > 0$,

$$\mathbb{P}\left(d_t^{(L)}(w_1', w_2') \geqslant \varepsilon_{2.3.5}/3\right) \leqslant \delta_{2.3.5}/3. \tag{2.3.7}$$

Assume that there is an $i \in \{1, 2\}$ such that $w_i \notin I$, and observe that we are done if we prove

$$\mathbb{P}\left(d_t^{(L)}(w_i, I) \geqslant \varepsilon_{2.3.5}/3\right) \leqslant \delta_{2.3.5}/3. \tag{2.3.8}$$

Analogously to the proof of Lemma 2.3.10, one can verify that whp the number of α-connectors between w_i and I tends to infinity with t. Hence, the weighted distance becomes small, and we conclude by a union bound over (2.3.7) and (2.3.8) twice (for both w_1 and w_2) that the result (2.3.6) follows for t large. □

Before outlining the proof of Proposition 2.3.4, we state the formal definitions of notions that are of particular importance throughout the proof.

Definition 2.3.6 (Layers, α-connectors, and the greedy path)**.** Fix $\alpha \in [1/2, 1)$. Let for $k \in \mathbb{N}$

$$s_k := \min\left\{ s_{k-1}^{(1-\varepsilon_{k-1})(\tau-2)^{-1}}, (\alpha t)^{\frac{1}{2(\tau-1)}} \log(\alpha t)^{-\frac{1}{2}} \right\}, \tag{2.3.9}$$

$$K_t := \min\left\{ k : s_k \geqslant (\alpha t)^{\frac{1}{2(\tau-1)}} \log(\alpha t)^{-\frac{1}{2}} \right\}, \tag{2.3.10}$$

where $s_0 > 1$ and $(\varepsilon_k)_{k\geqslant 0}$ is a sequence that tends to 0 for FPA and VPA under condition (2.2.3), and $\varepsilon_k \equiv \varepsilon_G$ for some small constant $\varepsilon_G > 0$ otherwise. For $(s_k)_{k\geqslant 0}$, we define the k*-th layer* as

$$\mathcal{L}_k := \{x \in [\alpha t] : D_{\alpha t}(x) \geqslant s_k\}.$$

A vertex y in $[t]\backslash[\alpha t]$ is called an α*-connector* of (x, z) if it is connected both to x and z. Let $\mathcal{A}_k(x) := \{(y, z) \in [t]\backslash[\alpha t] \times \mathcal{L}_k : x \leftrightarrow y \leftrightarrow z\}$. For a vertex $\pi_0 := q' \in \mathcal{L}_0$, we construct a greedy path

$$\pi^{\mathrm{gr}} = (\pi_0, y_1, \pi_1, y_2, \pi_2, ..., y_{K_t-1}, \pi_{K_t})$$

of length $2K_t$ by sequentially choosing

$$(y_k, \pi_k) = \underset{(y,z)\in\mathcal{A}_k(\pi_{k-1})}{\arg\min} \left\{ L_{(\pi_{k-1},y)} + L_{(y,z)} \right\}$$

if it exists. If $\mathcal{A}_k(\pi_{k-1}) = \emptyset$, we say that the construction of the greedy path fails at step k.

Note that on the greedy path π^{gr}, $\pi_k \in \mathcal{L}_k$ and y_k is an α-connector of (π_k, π_{k+1}).

Outline of the proof of Proposition 2.3.4

The idea of Proposition 2.3.4 is to construct a greedy path to the inner core and use its total weight as an upper bound for the actual shortest path. We outline the proof for FPA and VPA under condition (2.2.3). The other cases follow by similar steps.

(I) The greedy path consists of K_t cherries, i.e., of $2K_t$ edges. Recall K_t^* from (2.1.3). We show for a specific choice of ε_k that

$$2K_t \leqslant K_t^* + 4.$$

(II) We show that the sizes of the sets $\mathcal{A}_k(\pi_{k-1})$ are bounded from below by a doubly exponentially growing sequence $(n_k)_{k\geqslant 0}$, i.e., for $\delta > 0$, there is an s_0 such that for large t

$$\mathbb{P}\bigg(\bigcup_{k\in[K_t]} \{|\mathcal{A}_{k+1}(\pi_k)| \leqslant n_k\}\bigg) \leqslant \delta. \tag{2.3.11}$$

As a result, the greedy path actually exists, with probability at least $1-\delta$. To prove this, the choice of the exponent of $(s_k)_{k\geqslant 0}$ in (2.3.9) is crucial, and in particular the choice of $(\varepsilon_k)_{k\geqslant 0}$. In [81, Theorem 3.1] and [69, Proposition 3.1] similar constructions of greedy paths are used. In those proofs, the exponent of $(s_k)_{k\geqslant 0}$ is equal to $1/(\tau-2)$ and every term is *corrected with a* $\log(t)$*-term* to ensure that every vertex in layer $\mathcal{L}_k$ has at least *one* t-connector. In contrary, we *correct the exponent* by $-\varepsilon_k/(\tau-2)$, implying that $(s_k)_{k\geqslant 0}$ grows slower. In return, every vertex in $\mathcal{L}_k$ has *many* t-connectors whp, that allows for small weighted distances.

(III) By the construction of π^{gr}, the weighted distance between π_0 and π_{K_t} can be bounded by

$$d_t^{(L)}(q',\pi_{K_t}) \leqslant \sum_{k\in[K_t]} \min_{(y,z)\in[\mathcal{A}_k(\pi_{k-1})]} \left\{L_{(\pi_{k-1},y)} + L_{(y,z)}\right\}. \tag{2.3.12}$$

As the minimum is non-decreasing if we consider less elements, conditionally on the complement of the event in (2.3.11), we may weaken the bound to

$$d_t^{(L)}(q',\pi_{K_t}) \leqslant \sum_{k\in[K_t]} \min_{j\in[n_k]} \left\{L_{j1}^{(k)} + L_{j2}^{(k)}\right\}. \tag{2.3.13}$$

We show that the generalised inverse $F_{L_1+L_2}^{(-1)}$ can be related to the generalised inverse $F_L^{(-1)}$. This allows us to bound the rhs of (2.3.13) and obtain the asserted bound (2.3.3) from Proposition 2.3.4 as we make the error probabilities arbitrarily small by choosing s_0 sufficiently large.

We start with Step (I) by proving an upper bound for K_t.

Lemma 2.3.7. *Let* K_t *be as in* (2.3.10), K_t^* *as in* (2.1.3), $\tau \in (2,3)$, $\alpha \in [1/2, 1)$. *For FPA and VPA, if* (2.2.3) *holds, set*

$$\varepsilon_k := (k+2)^{-2}, \qquad k \in \mathbb{N}.$$

There exists a constant $c_{2.3.7} > 0$ *such that for* s_0 *sufficiently large*

$$s_k \geqslant s_0^{c_{2.3.7}(\tau-2)^{-k}}, \qquad 2K_t \leqslant K_t^* + 4. \tag{2.3.14}$$

For FPA and VPA if (2.2.3) *does not hold and GVPA, set* $\varepsilon_G > 0$ *such that*

$$\frac{\log\big(1/(\tau-2)\big)}{\log\big((1-\varepsilon_G)/(\tau-2)\big)} = 1 + \varepsilon_{2.3.4}. \tag{2.3.15}$$

Then for s_0 *sufficiently large*

$$2K_t \leqslant (1+\varepsilon_{2.3.4})K_t^* + 4. \tag{2.3.16}$$

Proof. First we consider FPA and VPA under condition (2.2.3). Recall the definition of $(s_k)_{k\geqslant 0}$ from (2.3.9). By iterating the recursion, we obtain

$$s_k = s_0^{\prod_{j=0}^{k-1}\big((1-\varepsilon_j)/(\tau-2)\big)}.$$

By our choice of $\varepsilon_k = 1/(k+2)^2$, the product $\prod_{j=1}^{\infty}(1-\varepsilon_j) > 0$, which yields the first bound in (2.3.14) for some constant $c_{2.3.7} > 0$. Hence, by the definition of K_t in (2.3.10),

$$K_t \leqslant \min\left\{k : s_0^{c_{2.3.7}(\tau-2)^{-k}} \geqslant (\alpha t)^{\frac{1}{2(\tau-1)}} \log(\alpha t)^{-\frac{1}{2}}\right\}.$$

By taking logarithms twice in the inequality between brackets above, we obtain

$$K_t \leqslant \left\lceil \frac{\log\left(\frac{1}{2(\tau-1)}\log(\alpha t) - \frac{1}{2}\log\log(\alpha t)\right) - \log\log(s_0) - \log(c)}{|\log(\tau-2)|} \right\rceil. \tag{2.3.17}$$

If s_0 is sufficiently large, the numerator of K_t above is smaller than $\log\log(t)$ for t sufficiently large. Bounding the rounding operations yields by the definition of K_t^* in (2.1.3)

$$2K_t \leqslant 2\frac{\log\log(t)}{|\log(\tau-2)|} + 2 \leqslant K_t^* + 3.$$

For proving (2.3.16) for GVPA, and FPA and VPA without condition (2.2.3), we use a similar reasoning. By taking logarithms twice in the definition of K_t in (2.3.10), we obtain a similar formula to that in (2.3.17), and after bounding the numerator as before, as well as using the implicit definition of ε_G in (2.3.15) and that of K_t^* in (2.1.3), we arrive to

$$2K_t \leqslant (1+\varepsilon_{2.3.4})K_t^* + 3 + \varepsilon_{2.3.4} \leqslant (1+\varepsilon_{2.3.4})K_t^* + 4,$$

finishing the proof. □

We recall two preliminary lemmas from [81] that help us control the error probability in (2.3.11).

Lemma 2.3.8 (Probability on being a α-connector for an arbitrary set [81])**.** *Consider PA under the same conditions as Proposition 2.3.4. Let* $\alpha \in [1/2, 1)$*. For* $x \in [\alpha t]$*, a set* $\mathcal{V} \subset [\alpha t]$*, conditionally on* $\mathrm{PA}_{\alpha t}$*, the probability that* $y \in [t]\backslash[\alpha t]$ *is an* α*-connector of* $(x, \mathcal{V})$ *is at least*

$$\frac{\eta_{2.3.8} D_{\alpha t}(x) D_{\alpha t}(\mathcal{V})}{(\alpha t)^2} =: p_{\alpha t}(x, \mathcal{V}), \tag{2.3.18}$$

where $\eta_{2.3.8} > 0$ *is a constant, and* $D_{\alpha t}(\mathcal{V}) := \sum_{z \in \mathcal{V}} D_{\alpha t}(z)$*. Moreover, with probability at least* $p_{\alpha t}(x, \mathcal{V})$*, the event* $\{y$ *is an* α*-connector of* $(x, \mathcal{V})\}$ *happens independently of other vertices in* $[t]\backslash[\alpha t]$*.*

We use the above lemma for layers $\mathcal{V} = \mathcal{L}_k$ and a vertex $x \in \mathcal{L}_{k-1}$. The following lemma allows us to bound the total degree of vertices in $\mathcal{L}_k$, $D_{\alpha t}(\mathcal{L}_k)$, from below. Although it assumes for the model GVPA(f) that the function $f(x)$ needs to be affine for x sufficiently large, we show below that this restriction does not propagate to the requirements of Proposition 2.3.4.

Lemma 2.3.9 (Impact of high degree vertices [81, Lemma A.1], [70, Theorem 1.1(a)])**.** *Let* PA_t *satisfy the same conditions as Proposition 2.3.4, and additionally for* GVPA(f)*, assume that the function* $f(x)$ *is affine for all* x *larger than some* $x_0 \in \mathbb{R}$*. Let* $\alpha \in [1/2, 1)$*, and* $\phi \in \mathbb{R}$ *satisfy* $x_0 \leqslant \phi \leqslant (\alpha t)^{\frac{1}{2(\tau-1)}} (\log(\alpha t))^{-\frac{1}{2}}$*. There exists a constant* $c_{2.3.9} > 0$ *such that*

$$\mathbb{P}\Bigg(\sum_{z : D_{\alpha t}(z) \geqslant \phi} D_{\alpha t}(z) \geqslant c_{2.3.9} \alpha t \phi^{2-\tau} \Bigg) = 1 - o(t^{-1}).$$

Moreover, whp the number of vertices with degree at least ϕ *is at least* $\sqrt{\alpha t}$*.*

Our version of the above lemma is slightly different from [81, Lemma A.1], as we do not assume that $\phi = \phi(t)$ tends to infinity with t, while [81] does. We refer the reader to the proof of [81, Lemma A.1] to see that the proof is also valid for constant ϕ. Also [70, Theorem 1.1(a)] is slightly different from the statement here, as it states convergence of the degree distribution in total variation norm. However, especially in combination with [70, Example 3.1], it is easy to check Lemma 2.3.9 is an immediate corollary. We continue with the main lemma of Step (II). Recall K_t, $(s_k)_{k \geqslant 0}$, and $\{\mathcal{L}_k\}_{k \geqslant 0}$ from Definition 2.3.6 and ε_k from Lemma 2.3.7.

Lemma 2.3.10 (Lower bound on the number of α-connectors)**.** *Consider FPA or VPA under the same conditions as Proposition 2.3.4. Let* $\alpha \in [1/2, 1)$*. There exists constant* $c_{2.3.10} > 0$*,* $c' > 0$*, such that for an arbitrary set* $\{\pi_0, ..., \pi_{K_t-1}\}$*, where* $\pi_k \in \mathcal{L}_k$*, and* s_0, t *sufficiently large*

$$\mathbb{P}\bigg(\bigcup_{k\in[K_t]} \Big\{ |\mathcal{A}_k(\pi_{k-1})| \leqslant c_{2.3.10} s_{k-1}^{\varepsilon_{k-1}/2} \Big\} \bigg) \leqslant 2 \exp\Big(-\frac{c_{2.3.10}}{4} s_0^{c'} \Big) =: \delta_{2.3.10}^{(s_0)}. \tag{2.3.19}$$

Proof. First, we show a stochastic domination argument of GVPA to VPA. Afterwards, we prove the existence of a binomial random variable A that is dominated by $|\mathcal{A}_k(\pi_{k-1})|$. Lastly, we apply Chernoff's bound to A and show that the result follows.

In order to apply Lemma 2.3.9, $f(x)$ must be an affine function for large x. Recall γ_f from (1.2.5) and assume f is non-affine. For any $\gamma_{f_1} \in (1/2, \gamma_f)$, there are $x_0 \in \mathbb{N}$, $\eta \in \mathbb{R}$, such that

$$f(x) \geqslant f_1(x) := \begin{cases} f(x) & x \leqslant x_0 \\ \gamma_{f_1} x + \eta & x > x_0, \end{cases}$$

which is concave and affine for $x > x_0$. Hence, the model GVPA(f_1) is well-defined and is stochastically dominated by GVPA(f). Assume that $s_0 > x_0$ so that we can apply Lemma 2.3.9 on GVPA(f_1). As γ_{f_1} can be chosen arbitrarily close to γ_f, we can choose it such that the power-law exponent τ_{f_1} is close to τ_f, in particular so that the inequality

$$\frac{1 - \varepsilon_G/2}{1 - \varepsilon_G} \geqslant \frac{\tau_{f_1} - 2}{\tau - 2} \tag{2.3.20}$$

holds with ε_G from (2.3.15). Assume that $s_0 > x_0$, and write $\varepsilon_k \equiv \varepsilon_G$ for the model GVPA. Define

$$\tau' := \begin{cases} \tau & \text{for FPA, VPA,} \\ \tau_{f_1} & \text{for GVPA.} \end{cases}$$

Let $\pi_{k-1} \in \mathcal{L}_{k-1}$ for some $k \in [K_t]$. By Lemma 2.3.8, the probability that $y \in [t] \backslash [\alpha t]$ is an α-connector of $(\pi_{k-1}, \mathcal{L}_k)$ is at least $p_{\alpha t}(\pi_{k-1}, \mathcal{L}_k)$, independently of other vertices in $[t] \backslash [\alpha t]$, see (2.3.18). Since there are in total $(1 - \alpha)t$ possible α-connectors, the random variable $|\mathcal{A}_k(\pi_{k-1})|$ stochastically dominates a binomial random variable, i.e.,

$$|\mathcal{A}_k(\pi_{k-1})| \succcurlyeq \text{Bin}\left((1-\alpha)t, p_{\alpha t}(\pi_{k-1}, \mathcal{L}_k)\right) =: A_k. \tag{2.3.21}$$

Here, the notation $\succcurlyeq$ is used for stochastic domination. Conditioning on $\mathcal{D}_{\mathcal{L}_k}(\alpha t)$ yields by Lemma 2.3.9 for τ',

$$\mathbb{E}[A_k] \geqslant \mathbb{E}\big[A_k \mid D_{\mathcal{L}_k}(\alpha t) \geqslant c_{2.3.9}\alpha t s_k^{2-\tau'}\big]\mathbb{P}\Big(D_{\mathcal{L}_k}(\alpha t) \geqslant c_{2.3.9}\alpha t s_k^{2-\tau'}\Big),$$

where the latter factor equals $1 - o(t^{-1})$. Since $\pi_{k-1} \in \mathcal{L}_{k-1}$ and thus $D_{\alpha t}(\pi_{k-1}) \geqslant s_{k-1}$ by the construction of $\mathcal{L}_{k-1}$ in Definition 2.3.6, we substitute the value of $p_{\alpha t}(\pi_{k-1}, \mathcal{L}_k)$ in (2.3.18) to bound the expectation of the binomial random variable A_k further to obtain

$$\begin{aligned}
\mathbb{E}[A_k] &\geqslant (1-\alpha)t\frac{\eta_{2.3.8}c_{2.3.9}\alpha t s_k^{2-\tau'} s_{k-1}}{(\alpha t)^2}\big(1 - o(t^{-1})\big)\\
&\geqslant 2c_{2.3.10} s_k^{2-\tau'} s_{k-1}\\
&\geqslant 2c_{2.3.10} s_{k-1}^{1-(2-\tau')\left(\frac{1-\varepsilon_{k-1}}{\tau-2}\right)} \geqslant 2c_{2.3.10} s_{k-1}^{\varepsilon_{k-1}/2} \qquad (2.3.22)
\end{aligned}$$

for some constant $c_{2.3.10} \in (0, (1-\alpha)c_{2.3.9}\eta_{2.3.8}/(2\alpha))$ if t is sufficiently large, by the recursive definition of s_k in (2.3.9). The last inequality is a consequence of (2.3.20). Next we apply Chernoff's bound, see e.g. [186], in the following form: for $\psi_k > 0$,

$$\mathbb{P}(A_k \leqslant (1-\psi_k)\mathbb{E}[A_k]) \leqslant \exp\big(-\psi_k^2\mathbb{E}[A_k]/2\big).$$

Choosing $\psi_k = (1 - c_{2.3.10}s_{k-1}^{\varepsilon_{k-1}/2}/\mathbb{E}[A_k])$, yields by (2.3.22) that $\psi_k \geqslant 1/2$. Hence, we can bound $\psi_k^2\mathbb{E}[A_k] \geqslant c_{2.3.10}s_{k-1}^{\varepsilon_{k-1}/2}/2$, so that

$$\mathbb{P}\left(A_k \leqslant c_{2.3.10}s_{k-1}^{\varepsilon_{k-1}/2}\right) \leqslant \exp\left(-c_{2.3.10}s_{k-1}^{\varepsilon_{k-1}/2}/4\right).$$

Applying a union bound over $k \in [K_t]$ and switching back to the dominating random variable $|\mathcal{A}_k(\pi_{k-1})|$ as in (2.3.21) results in

$$\begin{aligned}
\mathbb{P}\Big(\bigcup_{k\in[K_t]}\big\{|\mathcal{A}_k(\pi_{k-1})| \leqslant c_{2.3.10}s_{k-1}^{\varepsilon_{k-1}/2}\big\}\Big)&\\
\leqslant \sum_{k\in[K_t]} \exp\left(-c_{2.3.10}s_{k-1}^{\varepsilon_{k-1}/2}/4\right).& \qquad (2.3.23)
\end{aligned}$$

Because ε_k is a constant for GVPA, the asserted bound in (2.3.19) follows immediately for s_0 sufficiently large. For FPA and VPA, it remains to bound the sum on the rhs. We apply the lower bound on s_k from (2.3.14) and observe that $(\tau-2)^{-k}$ grows much faster than $(k+2)^2$, so

$$s_{k-1}^{\varepsilon_{k-1}/2} \geqslant s_0^{c_{2.3.7}(\tau-2)^{-k}(k+2)^{-2}} \geqslant s_0^{c'c^k},$$

for some $c, c' > 0$, whence the asserted bound (2.3.19) follows for all sufficiently large s_0. □

The above lemma ensures that there are *many* α-connectors. However, we still need to bound the probability that the weighted distance between π_{k-1} and $\mathcal{L}_k$ is sufficiently small, given that there are *enough* α-connectors, as described in Step (III) of the outline.

Lemma 2.3.11 (Minimum of i.i.d. random variables). *Let* $L_1, \dots, L_n$ *be i.i.d. random variables having distribution* F_L*. Then for all* $\xi > 0$

$$\begin{aligned} \mathbb{P}\Big(\min_{j\in[n]} L_j \geqslant F_L^{(-1)}\left(n^{-1+\xi}\right)\Big) &\overset{(\star)}{\leqslant} e^{-n^{\xi}}, \\ \mathbb{P}\Big(\min_{j\in[n]} L_j \leqslant F_L^{(-1)}\left(n^{-1-\xi}\right)\Big) &\overset{(*)}{\leqslant} n^{-\xi}. \end{aligned} \tag{2.3.24}$$

Proof. Since the random variables are i.i.d.,

$$\mathbb{P}\Big(\min_{j\in[n]} L_j \geqslant z(n)\Big) = (1 - F_L(z(n)))^n,$$

We substitute $z(n) = F_L^{(-1)}\left(n^{-1\pm\xi}\right)$, so that applying $(1-x)^n \leqslant e^{-nx}$ yields $(\star)$ in (2.3.24), and applying $(1-x)^n \geqslant 1 - nx$ yields $(*)$. □

We are ready to prove Proposition 2.3.4.

Proof of Proposition 2.3.4. Consider the greedy path π^{gr} starting from some $q' = q'(s_0)$ as defined in Definition 2.3.6. The definition of K_t in (2.3.10) ensures that π^{gr} ends in Inner_α. By construction of π^{gr}, its total weight is bounded by the formula in (2.3.12). Lemma 2.3.10 ensures that with probability at least $1 - \delta_{2.3.10}^{(s_0)}$, the number of α-connectors $(n_k)_{k\geqslant 1}$ in (2.3.11) is at least

$$n_k := c_{2.3.10} s_0^{c_{2.3.7}(\tau-2)^{-(k-1)}\varepsilon_{k-1}/2}.$$

By Lemma 2.3.7, $n_k \gg 1$ for all $k \geqslant 0$ when s_0 is sufficiently large. We bound the total weight on the greedy path using (2.3.13). Applying $(\star)$ in

(2.3.24) from Lemma 2.3.11 and a union bound over $k \in [K_t]$, we obtain for some $\xi > 0$ and all s_0 sufficiently large

$$\begin{aligned}\mathbb{P}\Bigg(\bigcup_{k\in[K_t]}\Big\{\min_{j\in[n_k]}\Big(L_{1j}^{(k)}+L_{2j}^{(k)}\Big)\geqslant F_{L_1+L_2}^{(-1)}\Big(n_k^{-(1-\xi)}\Big)\Big\}\Bigg)&\\ \leqslant \sum_{k\in[K_t]}\exp\Big(-c_{2.3.10}^{\xi}s_{k-1}^{\xi\varepsilon_{k-1}/2}\Big)&\\ \leqslant \sum_{k\in\mathbb{N}}\exp\Big(-c_{2.3.10}^{\xi}s_{k-1}^{\xi\varepsilon_{k-1}/2}\Big) =: \delta_{(2.3.25)}^{(s_0)},& \qquad (2.3.25)\end{aligned}$$

which can be made arbitrarily small by choosing s_0 large enough, similar to the reasoning below (2.3.23). We can choose s_0 so large that the error probabilities fulfil $\delta_{(2.3.25)}^{(s_0)}+\delta_{2.3.10}^{(s_0)} \leqslant \delta_{2.3.4}$. As we recall the formula for $(s_k)_{k\geqslant 0}$ from (2.3.9) and the *weakened* upper bound on the total weight (2.3.13), we obtain for any q' such that $D_{q'}(\alpha t) \geqslant s_0$, with error probability at most $\delta_{2.3.4}$,

$$d_t^{(L)}(q',\mathcal{L}_{K_t}) \leqslant \begin{cases}\sum_{k\in[K_t]} F_{L_1+L_2}^{(-1)}\Big(c_{2.3.10}^{-(1-\xi)}s_0^{-(1-\xi)c_{2.3.7}(\tau-2)^{-(k-1)}(k+1)^{-2}/2}\Big), & \text{(FV)},\\ \sum_{k\in[K_t]} F_{L_1+L_2}^{(-1)}\Big(c_{2.3.10}^{-(1-\xi)}s_0^{-(1-\xi)c_{2.3.7}((1-\varepsilon_G)/(\tau-2))^{-(k-1)}\varepsilon_G/2}\Big), & \text{(G)},\end{cases}$$

where we annotated the lines with (FV) if it holds for FPA and VPA under condition (2.2.3), and with (G) otherwise. We continue to do so below. Using (2.3.14), there exists a constant c, such that

$$d_t^{(L)}(q',\mathcal{L}_{K_t}) \leqslant \begin{cases}\sum_{k\in[K_t]} F_{L_1+L_2}^{(-1)}\left(\exp\left(-c\log(s_0)(\tau-2)^{-k}(k+1)^{-2}\right)\right), & \text{(FV)}\\ \sum_{k\in[K_t]} F_{L_1+L_2}^{(-1)}\left(\exp\left(-c\log(s_0)(1-\varepsilon_G)^{k}(\tau-2)^{-k}\varepsilon_G\right)\right), & \text{(G)}\end{cases} \qquad (2.3.26)$$

with probability at least $1-\delta_{2.3.4}$ if s_0 is sufficiently large. We bound the above sums, so that the terms match the terms of Q_t in (2.1.3): first we remove the convolution, then we switch to an integral, apply a variable transformation, and eventually switch back to a sum. The sum in (2.3.26) is taken over $F_{L_1+L_2}^{(-1)}$, while the summand in Q_t is taken over $F_L^{(-1)}$. We relate the two inverses for $x < 0$ by

$$F_{L_1+L_2}(x) = \mathbb{P}(L_1+L_2 \leqslant x) \geqslant \mathbb{P}\left(\max\{L_1,L_2\} \leqslant x/2\right) = \left(F_L\left(x/2\right)\right)^2.$$

Hence, for any $z > 0$, it holds that $F^{(-1)}_{L_1+L_2}(z) \leqslant 2F^{(-1)}_{L}(\sqrt{z})$. Applying this to the rhs of (2.3.26), we obtain for a different constant c

$$d_t^{(L)}(q', \mathcal{L}_{K_t}) \tag{2.3.27}$$
$$\leqslant \begin{cases} 2 \sum_{k \in [K_t]} F_L^{(-1)}\Big(\exp\left(-c \log(s_0)(\tau-2)^{-k}(k+1)^{-2}\right) \Big), & \text{(FV)}, \\ 2 \sum_{k \in [K_t]} F_L^{(-1)}\Big(\exp\left(-c \log(s_0)(1-\varepsilon_G)^k(\tau-2)^{-k}\varepsilon_G\right) \Big), & \text{(G)}. \end{cases}$$

For technical convenience, we define $a := \inf\{x : F_L^{(-1)}(x) > 0\}$ and $L' := L - a$, so that for (FV)

$$d_t^{(L)}(q', \mathcal{L}_{K_t}) \tag{2.3.28}$$
$$\leqslant 2aK_t + 2 \sum_{k \in [K_t]} F_{L'}^{(-1)}\Big(\exp\left(-c \log(s_0)(\tau-2)^{-k}(k+1)^{-2}\right)\Big).$$

Observe that for monotone non-increasing functions g, $g(1) < \infty$

$$\sum_{k=\lfloor a \rfloor}^{\lfloor b \rfloor} g(k) \overset{(*)}{\geqslant} \int_a^b g(x)dx, \qquad \int_a^b g(x)dx \overset{(\star)}{\geqslant} \sum_{k=\lceil a \rceil+1}^{\lfloor b \rfloor} g(k). \tag{2.3.29}$$

We apply $(\star)$ to switch in (2.3.28) from a sum to an integral. We discuss FPA and VPA under condition (2.2.3) first. We apply the variable transformation

$$\frac{y}{2} = x + \frac{\log\log s_0 + \log(c) - 2\log(x+1)}{|\log(\tau-2)|}, \tag{2.3.30}$$

that has a solution for all $x \geqslant 1$ if s_0 is sufficiently large. Now the integrand matches the summands of Q_t in (2.1.3). Differentiating both sides and rearranging terms gives an implicit formula for dx, that by (2.3.30) can be bounded by a function that only depends on y for some $C > 0$, i.e.,

$$dx = \frac{1}{2}\left(1 + \frac{2/|\log(\tau-2)|}{x+1-2/|\log(\tau-2)|}\right) dy \leqslant \frac{1}{2}\left(1 + \frac{C}{y}\right). \tag{2.3.31}$$

Thus, (2.3.28) and (2.3.31) together yield

$$d_t^{(L)}(q', \mathcal{L}_{K_t})$$
$$\leqslant 2aK_t + \int_{h_\tau(s_0)}^{h_\tau(s_0)+2K_t-\frac{4\log(K_t+1)}{|\log(\tau-2)|}} \left(1 + \frac{C}{y}\right) F_{L'}^{(-1)}\Big(\exp\Big(-(\tau-2)^{-y/2}\Big)\Big)\, dy.$$

where $h_\tau(s_0) = 2(\log\log s_0 + \log(c))/|\log(\tau-2)|$ and s_0 is chosen so large that (2.3.31) holds for all x, y in the integration domain. When condition (2.2.3) on L holds, there exists $M > 0$, such that

$$\int_{h_\tau(s_0)}^{h_\tau(s_0)+2K_t} \frac{C}{y} F_{L'}^{(-1)}\left(\exp\left(-(\tau-2)^{-y/2}\right)\right) dy < M.$$

We apply $(*)$ in (2.3.29) to switch back to a sum. As this sum contains at most $2K_t + 1$ terms, there exists a larger M, recalling that $L = L' + a$, so that

$$\begin{aligned} d_t^{(L)}(q', \mathcal{L}_{K_t}) &\leqslant 2aK_t + \sum_{k=\lfloor h_\tau(s_0)\rfloor}^{\lfloor h_\tau(s_0)+2K_t\rfloor} F_{L'}^{(-1)}\left(\exp\left(-(\tau-2)^{-y/2}\right)\right) + M \\ &\leqslant \sum_{k=\lfloor h_\tau(s_0)\rfloor}^{\lfloor h_\tau(s_0)+2K_t\rfloor} F_{L}^{(-1)}\left(\exp\left(-(\tau-2)^{-k/2}\right)\right) + M. \end{aligned}$$

Application of the bound (2.3.14) on $2K_t$ yields the assertion (2.3.3) for FPA and VPA under (2.2.3).

For FPA and VPA if (2.2.3) does not hold, and for the model GVPA, we use a similar variable transformation to (2.3.31) for the integral in (2.3.28), i.e.,

$$\frac{y}{2}\frac{|\log(\tau-2)|}{\log((1-\varepsilon_G)/(\tau-2))} = x + \frac{\log\log s_0 + \log(c\varepsilon_G)}{\log((1-\varepsilon_G)/(\tau-2))},$$

which yields combined with the second line in (2.3.27)

$$d_t^{(L)}(q', \mathcal{L}_{K_t}) \leqslant (1+\varepsilon_{2.3.4}) \int_{\tilde{h}_\tau(s_0)}^{\tilde{h}_\tau(s_0)+\frac{2}{1+\varepsilon_{2.3.4}}K_t} F_L^{(-1)}\left(\exp\left(-(\tau-2)^{-x}\right)\right) dx,$$

where $\tilde{h}_\tau(s_0) = 2\left(\log\log s_0 + \log(c\varepsilon_G)\right)/|\log(\tau-2)|$. After applying the bound (2.3.16) on K_t and switching back to a sum using $(*)$ from (2.3.29), we obtain the desired bound (2.3.4). □

This establishes all prerequisites. We turn to the upper bound for the conservative case.

2.3.1 *Conservative case*

Proof of Proposition 2.3.1. We first show the result for FPA and VPA under (2.2.3). At the end of this proof, we argue how to adapt the proof for GVPA and VPA or FPA if (2.2.3) does not hold. Fix $\delta > 0$, and let $\delta_{2.3.3} = \delta_{2.3.4} = \delta/8$, $\alpha = 1 - \delta/16$. We can choose t sufficiently large, such that

(i) by Proposition 2.3.3 for $s_{2.3.3} = s_0$ that we choose below in (ii), there exists a constant $C_{2.3.3} = C_{2.3.3}(\delta, s_0)$, such that for $q \in \{u, v\}$ there is a vertex $q' \in [\alpha t]$ with degree at least s_0 within graph distance $C_{2.3.3}$ with probability at least $1 - \delta/8$.

(ii) by Proposition 2.3.4 there is an $s_0 = s_0(\delta, \varepsilon_{2.3.4}) > 0$, such that for $q' \in [\alpha t]$ with $D_{\alpha t}(q') \geqslant s_0$, the weighted distance to the inner core is not too large. As the terms in the sum in (2.3.3) are decreasing, we shift the summation bounds to match the bounds from Q_t in (2.1.3), so that

$$\mathbb{P}\Big(d_t^{(L)}(q', \mathrm{Inner}_\alpha) \geqslant M_{2.3.4} + Q_t\Big) \leqslant \frac{\delta}{8}. \tag{2.3.32}$$

(iii) by Proposition 2.3.5 for $\delta_{2.3.5} = \delta/8$, the graph distance between $w_1, w_2 \in \mathrm{Inner}_\alpha$ is smaller than $C_{2.3.5}$ with probability at least $1 - \delta/8$.

(iv) for $(L_j)_{j \geqslant 0}$ i.i.d. copies of L, the sum of constantly many weights is negligible compared to the diverging sequence Q_t defined in (2.1.3), i.e., there exists $M' > 0$ such that

$$\mathbb{P}\Big(\sum_{j \in [2C_{2.3.3} + C_{2.3.5}]} L_j \geqslant M'\Big) \leqslant \frac{\delta}{8}. \tag{2.3.33}$$

Conditionally on the intersection of the complements of the events in (2.3.32) and (i) for $q \in \{u, v\}$, we can construct greedy paths from u and v to the inner core. Hence,

$$\mathbb{P}\Big(\bigcup_{q \in \{u,v\}} \{d_t^{(L)}(q, \mathrm{Inner}_\alpha) \geqslant Q_t + M_{2.3.4} + \sum_{j \in [C_{2.3.3}]} L_j^{(q)}\} \Big| u, v < \alpha t\Big) \leqslant \frac{3\delta}{4}, \tag{2.3.34}$$

where the sum between brackets in (2.3.34) represents the weighted distance from q to some vertex with degree at least s_0. By (iii), we can bound the distance between two vertices in the inner core, and thus we have constructed a path from u to v. Hence, by a union bound over the events in (iii) and (2.3.34), we can bound the total weight on this path, i.e.,

$$\mathbb{P}\Big(d_t^{(L)}(u, v) \geqslant 2Q_t + \sum_{j \in [2C_{2.3.3} + C_{2.3.5}]} L_j \;\Big|\; u, v < \alpha t\Big) \leqslant \frac{7\delta}{8}. \tag{2.3.35}$$

Combining (2.3.33) with (2.3.35) yields by our choice of $\alpha = 1 - \delta/16$ that

$$\mathbb{P}\left(d_t^{(L)}(u, v) \geqslant 2Q_t + 2M_{2.3.4} + M'\right) \leqslant \delta, \tag{2.3.36}$$

finishing the proof for FPA and VPA under (2.2.3). For the more general case, only (2.3.32) does not hold, which can be replaced by

$$\mathbb{P}\Big(d_t^{(L)}(q', \mathrm{Inner}_\alpha) \geqslant (1+\varepsilon_{2.3.4})Q_t\Big) \leqslant \frac{\delta}{8}.$$

Propagating the rhs between brackets through (2.3.34) and (2.3.36) yields the result for GVPA. □

2.3.2 *Explosive case*

Similarly to the conservative case we create a path from u to v and use the total weight on the path as an upper bound for the weighted distance. Again, by applying Propositions 2.3.3, 2.3.4, and 2.3.5, this path goes from $q \in \{u, v\}$ to a vertex q' with degree at least s_0, after which we connect this vertex to the inner core. The total weight on the segments to the inner core can be made arbitrarily small for large t and s_0. However, it is not straightforward to see that the weight on the first parts of the paths converge to the explosion times of two LWLs, as we increase the degree s_0. We start by showing that the explosion time of the LWL is finite, for which we use Propositions 2.3.3 and 2.3.4. Recall the formal statement of Theorem 2.2.9.

Proof of Theorem 2.2.9. Fix $\delta > 0$. Recall K_t from (2.3.10) and $C_{2.3.3}$ for $\delta_{2.3.3} = \delta/4$ from Proposition 2.3.3. Let $a(t) := \min\{\kappa(t), 2K_t + C_{2.3.3}\}$, where $\kappa(t)$ denotes the maximum number of generations in the LWL to maintain a coupling with PA_t with probability at least $1-\delta/4$, from (2.2.1). As $\kappa(t)$ and K_t tend to infinity with t, the same holds for $a(t)$. Denote the first k generations of the LWL rooted in $\circledcirc$ by $\mathrm{LWL}_k(\circledcirc)$. Define

$$\begin{aligned} X_t &:= d_t^{(L)}\left(q^{(t)}, \partial\mathcal{B}_t^{(G)}(q^{(t)}, 2K_t + C_{2.3.3})\right), \\ Y_t &:= d_t^{(L)}\left(q^{(t)}, \partial\mathcal{B}_t^{(G)}(q^{(t)}, a(t))\right)\mathbb{1}_{\{\text{coupling succesful}\}} \\ &= \beta_{\mathrm{LWL}_{a(t)}(q^{(t)}), a(t)}\mathbb{1}_{\{\text{coupling succesful}\}}. \end{aligned}$$

Observe that X_t can be viewed as an upper bound of the weighted distance to Inner_α. By the choice of $a(t)$, X_t stochastically dominates Y_t, i.e., $\mathbb{P}(Y_t \leqslant X_t) = 1$. Consider the subsequence of times, defined recursively as

$$t_i := \begin{cases} 1, & i = 0, \\ \min\{t : a(t) > a(t_{i-1}) \text{ and } K_t^* > K_{t_{i-1}}^*\}, & i > 0. \end{cases}$$

By construction of Y_t and $(t_i)_{i\geqslant 0}$,

$$\mathbb{P}(\beta_{\mathrm{LWL},\infty} \leqslant y) = \lim_{i\to\infty} \mathbb{P}(Y_{t_i} \leqslant y).$$

Since X_{t_i} dominates Y_{t_i},

$$\begin{aligned} \mathbb{P}(\beta_{\mathrm{LWL},\infty} = \infty) &= \lim_{M\to\infty} \mathbb{P}(\beta_{\mathrm{LWL},\infty} \geqslant M) \\ &= \lim_{M\to\infty} \lim_{i\to\infty} \mathbb{P}(Y_{t_i} \geqslant M) \\ &\leqslant \lim_{M\to\infty} \lim_{i\to\infty} \mathbb{P}(X_{t_i} \geqslant M). \end{aligned} \tag{2.3.37}$$

Below we find a bound on $\mathbb{P}(X_{t_i} \geqslant M)$ that does not depend on t or i to obtain the result. Recall Q_t from (2.1.3) and define $Q_\infty := \lim_{t\to\infty} Q_t$, which is finite since $I_1(L) < \infty$. By a union bound on the events in Proposition 2.3.3 and Proposition 2.3.4, for any $\delta, \varepsilon > 0$, if i is sufficiently large,

$$\mathbb{P}\Bigg(X_{t_i} \geqslant (1+\varepsilon)Q_{t_i} + \sum_{j\in[C_{2.3.3}]} L_j\Bigg) \leqslant \frac{\delta}{2}.$$

Note that $(1+\varepsilon)Q_{t_i} \leqslant (1+\varepsilon)Q_\infty$. Choose $M' = M'(\delta)$ so that $M'/2 \geqslant (1+\varepsilon)Q_\infty$ and $\sum_{j\in[C_{2.3.3}]} L_j \geqslant M'/2$ with probability at most $\delta/2$, similarly to (2.3.33). Hence,

$$\lim_{i\to\infty} \mathbb{P}(X_{t_i} \geqslant M') \leqslant \delta.$$

Recall (2.3.37) to obtain $\mathbb{P}(\beta_{\mathrm{LWL},\infty} = \infty) \leqslant \delta$. Since $\delta > 0$ was arbitrary, $\beta_{\mathrm{LWL},\infty} < \infty$ almost surely. □

Recall the definitions of the graph neighbourhoods in Definition 2.2.1 and the explosion time in Definition 2.2.2. Using the finiteness of the explosion time, we bound the weight on a path to a vertex with degree at least s_0. To do so, we need the following general lemma.

Lemma 2.3.12 (Reaching a high-degree vertex in an explosive tree)**.** *Consider a (possibly random) locally finite tree T, rooted in $\circledcirc$, where every edge is equipped with an i.i.d. edge-weight from distribution L, where $F_L(0) = 0$, such that T is explosive. Write $D(x)$ for the degree of vertex x. Fix $\delta_{2.3.12} > 0$. For any $s \in \mathbb{N}$, there exists $N = N(\delta_{2.3.12}, s) < \infty$ such that*

$$\mathbb{P}\left(\mathcal{B}_T^{(L)}(\circledcirc, \sigma_{T,N}) \cap \{x \in T : D(x) \geqslant s\} \neq \emptyset\right) \geqslant 1 - \delta_{2.3.12}. \tag{2.3.38}$$

Proof. The event in (2.3.38) means that among the N closest vertices to $\circledcirc$, there is a vertex with degree at least s. We argue by contradiction. If the tree T is explosive and the vertices contained in $\bigcup_{n\in\mathbb{N}} \mathcal{B}^{(L)}_T(\circledcirc, \sigma_{T,n})$ would all have degree at most s, then the forest restricted to vertices in T with degree at most s is also explosive. This forest consists of trees with at most exponentially growing generation sizes. Hence, [166, Lemma 4.3] applies and explosion is impossible, i.e.,

$$\mathbb{P}\Big(\bigcap_{n\geqslant 1}\big\{\mathcal{B}^{(L)}_T(\circledcirc, \sigma_{T,n}) \cap \{x \in T : D(x) \geqslant s\} = \emptyset\big\}\Big) = 0.$$

Since $\bigcup_{n\in[N]} \mathcal{B}^{(L)}_T(\circledcirc, \sigma^T_n) = \mathcal{B}^{(L)}_T(\circledcirc, \sigma^T_N)$ by definition, we obtain the result, i.e.,

$$\begin{aligned}\lim_{N\to\infty}&\mathbb{P}\Big(\big\{\mathcal{B}^{(L)}_T(\circledcirc, \sigma_{T,N}) \cap \{x \in T : D(x) \geqslant s\} \neq \emptyset\big\}\Big)\\ &= \lim_{N\to\infty}\mathbb{P}\Big(\bigcup_{n\in[N]}\big\{\mathcal{B}^{(L)}_T(\circledcirc, \sigma_{T,n}) \cap \{x \in T : D(x) \geqslant s\} \neq \emptyset\big\}\Big) = 1,\end{aligned}$$

hence, the lhs will have probability at least $1-\delta_{2.3.12}$ for a sufficiently large N. □

In the next lemma, we exploit the coupling to LWL to find a vertex with sufficiently high degree. This lemma can be viewed as an extension of Proposition 2.3.3.

Lemma 2.3.13 (Weighted distance to a vertex with degree at least s)**.** *Consider PA under the same conditions as Theorem 2.2.10 at time* t*. Let* q *be a typical vertex. For any* $\delta_{2.3.13}, s_0 > 0$*, there exists* $N_{2.3.13} \in \mathbb{N}$*, such that when* t *is sufficiently large*

$$\mathbb{P}\Big(\mathcal{B}^{(L)}_t(q, \sigma_{t,N_{2.3.13}}) \cap \{x \in [t] : D_t(x) \geqslant s_0\} = \emptyset\Big) \leqslant \delta_{2.3.13}.$$

Proof. From Lemma 2.3.12 for $\delta_{2.3.12} = \delta/2$, we obtain that in the limiting object LWL($\circledcirc$) for a root $\circledcirc$, there is an N such that a vertex of degree at least s_0 is reached before time σ_N with probability at least $1-\delta/2$. By Theorem 2.2.9 the LWL is explosive. By applying Proposition 2.2.4, we can assume t is so large that we can maintain a coupling between with probability at least $1-\delta/2$ up to graph distance N, so $\widetilde{\mathcal{B}}^{(G)}_t(q, N) \simeq \mathrm{LWL}_N(\circledcirc)$. As the N-th closest vertex in L-distance is within the G-neighbourhood of size N, we obtain the result. □

We are now ready to prove the upper bound for the explosive case.

Proof of Proposition 2.3.2. Let s_0 be the constant we obtain from Proposition 2.3.4 when setting $\delta_{2.3.4} = \delta/8$, $\alpha_{2.3.4} = 1 - \delta/16$ and $\varepsilon_{2.3.4} = \varepsilon/3$, such that the sum in (2.3.3) is smaller than $\varepsilon/6$. Abbreviate $t' := \lfloor \alpha t \rfloor$. Observe that

$$\mathbb{P}\left(\{u \notin [t']\} \cup \{v \notin [t']\}\right) \leqslant \delta/8. \tag{2.3.39}$$

We call the above event between brackets $\mathcal{E}_1$. Recall σ_n from (2.2.2). Let for $q \in \{u, v\}$

$$N^*(q) = \min\{n : \exists x \in \mathcal{B}_{t'}^{(L)}(q, \sigma_{t',n}) : D_{t'}(x) \geqslant s_0\}.$$

Let u' be the vertex in $\mathcal{B}_{t'}^{(L)}\big(u, \sigma_{t',N^*}(u)\big)$ that has degree at least s_0, and v' analogously. The key observation is that, conditionally on $\mathcal{E}_1^c$, we can use a sprinkling argument and look at the graph at two moments in time, i.e.,

$$d_t^{(L)}(u, v) \leqslant d_{t'}^{(L)}(u, u') + d_{t'}^{(L)}(v, v') + d_t^{(L)}(u', v').$$

The above is true, since for *fixed* vertices the weighted distance between them is non-increasing in time. Conditionally on the event $\mathcal{E}_1^c$, u and v are uniform vertices in $[t']$, so that we can still apply Lemma 2.3.13 for s_0 and $\delta_{2.3.13} = \delta/8$, i.e., for t sufficiently large and some $N_{2.3.13} = N_{2.3.13}(\delta, s_0)$,

$$\mathbb{P}\Bigg(\bigcup_{q \in \{u,v\}} \{N^*(q) \geqslant N_{2.3.13}\}\Bigg) \leqslant \frac{\delta}{4}. \tag{2.3.40}$$

We call the above event between the $\mathbb{P}$-sign $\mathcal{E}_2$. By the choice of u' and v', we obtain that

$$\mathbb{P}\left(d_t^{(L)}(u, v) \leqslant \sigma_{N_{2.3.13}}^{(t')}(u) + \sigma_{N_{2.3.13}}^{(t')}(v) + d_t^{(L)}(u', v') \mid \mathcal{E}_1^c \cap \mathcal{E}_2^c\right) = 1.$$

Assume t' is so large that we can maintain a coupling with LWL with probability at least $1 - \delta/8$ for $q \in \{u, v\}$ up to generation $N_{2.3.13}$, possible by Propositions 2.2.4. If the coupling is successful, then for $q \in \{u, v\}$ we have $\sigma_{t',N_{2.3.13}}(q) \leqslant \beta_{\mathrm{LWL}^{(q)},N_{2.3.13}}$, where the two random variables are independent copies of $\beta_{\mathrm{LWL},N_{2.3.13}}$. Thus, combining the coupling error and the error probabilities on $\mathcal{E}_1$ and $\mathcal{E}_2$ in (2.3.39) and (2.3.40),

$$\mathbb{P}\left(d_t^{(L)}(u, v) \leqslant \beta_{\mathrm{LWL}^{(u)},N_{2.3.13}} + \beta_{\mathrm{LWL}^{(v)},N_{2.3.13}} + d_t^{(L)}(u', v')\right) \geqslant 1 - 5\delta/8, \tag{2.3.41}$$

where $\mathbb{P}$ denotes the probability measure on the coupled probability space. Recall (2.3.1), we observe that we are left with proving

$$\mathbb{P}\left(d_t^{(L)}(u', v') \geqslant \varepsilon\right) \leqslant 3\delta/8. \tag{2.3.42}$$

By our choice of $\delta_{2.3.4} = \delta/8$ and $\varepsilon_{2.3.4} = \varepsilon/3$ above, we obtain by Proposition 2.3.4

$$\mathbb{P}\Bigg(\bigcup_{q' \in \{u', v'\}} \big\{ d_t^{(L)}(q', \text{Inner}_\alpha) \geqslant \varepsilon/3 \big\} \Bigg) \leqslant \delta/4. \tag{2.3.43}$$

It is important to note that although u' and v' are *special* vertices at time αt, (2.3.43) holds independently of $N_{2.3.13}$, as the path to inner core, constructed in the proof of Proposition 2.3.4, uses only edges that arrived after αt. To apply Proposition 2.3.5, we let $\delta_{2.3.5} = \delta/8$ and $\varepsilon_{2.3.5} = \varepsilon/3$. By a union bound over the events in Propositions 2.3.4 and 2.3.5 there is a path from u' to v' of weight at most $\varepsilon/3$, yielding (2.3.42). Combining (2.3.41) and (2.3.42) yields the desired bound (2.3.1). □

2.4 LOWER BOUND ON THE WEIGHTED DISTANCE

The next propositions state the lower bounds for Theorem 2.2.5 and 2.2.10, which are the counterparts of Propositions 2.3.1 and 2.3.2. The lower bound of Theorem 2.2.8 which follows from the same proof techniques, is postponed to Section 2.6.

Proposition 2.4.1 (Lower bound on the weighted distance, conservative case)**.** *Consider PA under the same conditions as Theorem 2.2.5. Recall* $I_1(L) = \infty$*. Let* u, v *be two typical vertices. Then for every* $\delta, \varepsilon > 0$*, if* t *is sufficiently large,*

$$\mathbb{P}\left(d_t^{(L)}(u, v) \geqslant (1-\varepsilon) 2Q_t \right) \geqslant 1 - \delta. \tag{2.4.1}$$

Moreover, for the models FPA and VPA from Definition 1.2.4 and 1.2.5, there exists a constant $M_{2.4.1} = M_{2.4.1}(\delta)$ *such that for* t *sufficiently large*

$$\mathbb{P}\left(d_t^{(L)}(u, v) \geqslant 2Q_t - 2M_{2.4.1} \right) \geqslant 1 - \delta. \tag{2.4.2}$$

Proposition 2.4.1 gives a tight lower bound on the weighted distance for any distribution, i.e., it does not depend on the tightness condition (2.2.3). Thus, only the upper bound in (2.3.1) needs be improved to prove (2.2.2) regardless of (2.2.3).

We proceed to the main proposition for the explosive edge-weight distributions.

Proposition 2.4.2 (Lower bound on the weighted distance, explosive case)**.** *Consider PA under the same conditions as Theorem 2.2.10. Recall* $I_1(L) < \infty$*. Let* u, v *be two typical vertices. Then, there is a coupled probability space, such* *that for any* $\delta > 0$, *there exists some function* $a(t)$ *that tends to infinity with* t, *such that*

$$\mathbb{P}\left(d_t^{(L)}(u,v) \geqslant \beta_{\mathrm{LWL}_{a(t)}^{(u)},a(t)} + \beta_{\mathrm{LWL}_{a(t)}^{(v)},a(t)} \right) \geqslant 1-\delta,$$

where $\beta_{\mathrm{LWL}_{a(t)}^{(u)},a(t)}$ *and* $\beta_{\mathrm{LWL}_{a(t)}^{(u)},a(t)}$ *are the times to reach graph distance* $a(t)$ *in the LWLs coupled to* u, v, *respectively.*

As the proofs of the propositions are partly based on the same principles, we outline the proofs simultaneously, after which we prove the required lemmas, and then give a separate proof for the above propositions. For notational convenience, we only outline the models FPA and VPA.

Outline of the proofs

Let u, v be two typical vertices. Recall K_t^* from (2.1.3). From [69, Theorem 2] it follows that any path connecting u and v has whp at least $2K_t^* - M_t'$ edges with probability close to 1, for some function $M_t' = o(K_t^*)$. The idea of the proof of Proposition 2.4.1 is to show that any path starting from u or v that has $K_t^* - M_t'$ edges, has total weight at least $Q_t - M_{2.4.1}$, or total weight at least $\beta_{\mathrm{LWL},a(t)}$, for some function $a(t)$ that tends to infinity, for the conservative and explosive case respectively. The main steps that make this succeed, which we formalize further below, are as follows, where (I-III) apply for the conservative case, while the explosive case follows from (I).

(I) We use results from [69, 81] to conclude that the neighbourhoods $\mathcal{B}_t^{(G)}(u,k)$, $\mathcal{B}_t^{(G)}(v,k)$ are disjoint whp for $k < K_t^* - M_t'$. These neighbourhoods up to graph distance $a(t)$ are coupled to two independent LWLs, from which the lower bound for the explosive case will follow.

(II) We show that $\partial\mathcal{B}_t^{(G)}(u,k)$ has whp at most $\exp\left(B(\tau-2)^{-k/2}\right)$ vertices, for B sufficiently large and $k < K_t^* - M_t'$, by counting the number of paths of length k starting from $q \in \{u,v\}$. The number of such paths is an upper bound for the number of vertices in $\partial\mathcal{B}_t^{(G)}(q,k)$.

(III) From (I) it follows that $d_t^{(L)}(u,v)$ is at least the weighted distance from u to $\partial\mathcal{B}_t^{(G)}(u, K_t^* - M_t')$ plus the weighted distance from v to

$\partial\mathcal{B}_t^{(G)}(v, K_t^* - M_t')$. Along the same reasoning as shown by Adriaans and Komjáthy [2, Lower bound (2.11)], we obtain the lower bound

$$\begin{aligned} d_t^{(L)}(u,v) &\geqslant d_t^{(L)}\left(u, \partial\mathcal{B}_G^{(t)}(u, K_t^* - M_t')\right) \\ &\quad + d_t^{(L)}\left(v, \partial\mathcal{B}_G^{(t)}(v, K_t^* - M_t')\right) \\ &\geqslant \sum_{q\in\{u,v\}} \sum_{i=0}^{\lfloor K_t^* - M_t' \rfloor - 1} \min_{x\in\partial\mathcal{B}_t^{(G)}(q,i), y\in\partial\mathcal{B}_G^{(t)}(q,i+1)} L_{(x,y)}. \end{aligned} \tag{2.4.3}$$

The weight $L_{(x,y)}$ is the weight attached to the edge (x, y) in the graph if it is present. Otherwise, it is a new i.i.d. copy of L. Note that this is a valid lower bound, as the minimum is non-increasing when adding more edges. Using the bounds established in (II), we show that this sum of minima is at least $2Q_t - 2M_{2.4.1}$.

Before proving Proposition 2.4.1 and 2.4.2, we formally introduce the lemmas and proposition that correspond to Step (I-III). We start with a proposition from [69] that implies the first part of (I).

Proposition 2.4.3 (Typical graph distance [69, Theorem 2])**.** *Consider PA with power-law exponent $\tau \in (2, 3)$. Let u, v be two typical vertices. For any $\delta_{2.4.3} > 0$, there exists a function $M_{2.4.3}(t)$ such that if t is sufficiently large, then*

$$\mathbb{P}\left(\left\{d_t^{(G)}(u,v) > 2K_t^* - 2M_{2.4.3}(t)\right\}\right) \geqslant 1 - \delta_{2.4.3}, \tag{2.4.4}$$

where the function $M_{2.4.3}(t)$ is of order $O(1)$ for the models FPA and VPA, and $o(K_t^*)$ for the model GVPA. We denote the above event between brackets by $\mathcal{E}_{2.4.3}^{(t)}$.

Indeed, as a result of Proposition 2.4.3, on the event $\mathcal{E}_{2.4.3}^{(t)}$,

$$\mathcal{B}_t^{(G)}(u, k) \cap \mathcal{B}_t^{(G)}(v, k) = \emptyset \qquad \text{for } k \leqslant K_t^* - M_{2.4.3}(t),$$

establishing (I). We continue with the lemmas implying (II) and (III) from the outline and prove the conservative case first. We make use of some results from Dereich, Mönch, and Mörters [69, Theorem 2] where they prove lower bounds on graph distances for random graphs. They work under the following general condition and prove that it holds for FPA and GVPA.

Proposition 2.4.4 (PA(γ) [69, Proposition 3.1, 3.2]). *We say that PA satisfies the condition PA($\gamma_{2.4.4}$) for some $\gamma_{2.4.4} \in (0,1)$, if there exists a constant $\nu_{2.4.4} > 0$, such that for all t and pairwise distinct vertices $v_0, ..., v_l \in [t]$*

$$\mathbb{P}(v_0 \leftrightarrow v_1 \leftrightarrow v_2 \leftrightarrow \cdots \leftrightarrow v_\ell) \leqslant \prod_{k=1}^{\ell} p(v_{k-1}, v_k) =: p(v_0, ..., v_\ell) \qquad (2.4.5)$$

where $p(m,n) := \nu_{2.4.4}(m \wedge n)^{-\gamma_{2.4.4}}(m \vee n)^{\gamma_{2.4.4}-1}$. The above condition is satisfied for FPA or VPA, with $\gamma_{2.4.4} = 1/(\tau-1)$; and for GVPA for any $\gamma_{2.4.4} > \gamma_f$.

Step (II) in the outline follows from the next lemma.

Lemma 2.4.5 (Small probability of too large neighbourhoods). *Consider PA under the same conditions as Proposition 2.4.1, that satisfies Proposition 2.4.4 for some $\gamma_{2.4.4} \in (1/2, 1)$. Let q be a typical vertex. There exist a constant $B_{2.4.5} = B_{2.4.5}(\gamma_{2.4.4}, \nu_{2.4.4}) > 0$, such that for any $\delta_{2.4.5} > 0$, there exists a sequence of events $\mathcal{E}^{(t)}_{2.4.5}(q)$, such that for all t sufficiently large*

$$\mathbb{P}\left(\mathcal{E}^{(t)}_{2.4.5}(q)\right) \geqslant 1 - \delta_{2.4.5}, \qquad (2.4.6)$$

and for any $B > B_{2.4.5}$, $\gamma \in [\gamma_{2.4.4}, 1)$ and $M_{2.4.3}(t)$ from Lemma 2.4.3 with parameter $\delta_{2.4.3} = \delta_{2.4.5}$,

$$\mathbb{P}\left(\bigcap_{k \in [K^*_t - M_{2.4.3}(t)]} \left\{ |\partial \mathcal{B}^{(G)}_t(q,k)| \leqslant \exp\left(2B\left(\gamma/(1-\gamma)\right)^{k/2}\right)\right\} \;\middle|\; \mathcal{E}^{(t)}_{2.4.5}(q)\right) \geqslant 1 - 2\exp(-B). \qquad (2.4.7)$$

Proof. For notational convenience we abbreviate $\gamma = \gamma_{2.4.4}$ and $A_\gamma = \gamma/(1-\gamma)$. We use a path counting technique similar to [69, Theorem 2], see also [46, Proposition 4.9]. Consider paths $\pi = (\pi_0, \pi_1, \ldots, \pi_K)$ of length $K < K^*_t - M_{2.4.3}(t)$, such that $\pi_0 = q, \pi_K = w$, and where $K \leqslant K^*_t - M_{2.4.3}(t)$. For a non-increasing sequence $(\ell_k)_{\geqslant 0}$, we call a path π *good* if $\pi_k \geqslant \ell_k$ for all $k \leqslant K$, and *bad* otherwise. Let $\delta' := \delta_{2.4.5}/3$, $\ell_0 := \lceil \delta' t \rceil$. We define for a vertex $q \in [t]$ the event

$$\mathcal{E}_{\mathrm{bad}}(q) := \{q < \ell_0\} \cup \left\{ \begin{matrix} \exists \text{ bad path of length } K \text{ from } q, \\ \text{for some } K \leqslant K^*_t - M_{2.4.3}(t) \end{matrix} \right\}. \qquad (2.4.8)$$

The event $\{q < \ell_0\}$ occurs with probability at most $\delta' + 1/t$. The second event in the union of (2.4.8) occurs with probability at most δ', which follows from [69, first inequality after (18)] for a given choice of ℓ_k that we also use here, see (2.4.14) below. Hence, for a vertex q chosen uniformly at random from [t]

$$\mathbb{P}(\mathcal{E}_{\mathrm{bad}}(q)) \leqslant 2\delta' + 1/t \leqslant \delta_{2.4.5},$$

when $t \geqslant 3/\delta_{2.4.5}$. Let $\mathcal{E}^{(t)}_{2.4.5}(q) := (\mathcal{E}_{\mathrm{bad}}(q))^c$. The key element is to prove that for B large enough

$$\mathbb{E}[|\partial\mathcal{B}^{(G)}_t(q,k)| \mid \mathcal{E}^{(t)}_{2.4.5}(q)] \leqslant \exp\big(BA_\gamma^{k/2}\big). \tag{2.4.9}$$

Indeed, once we have shown that (2.4.9) holds, applying Markov's inequality yields

$$\begin{aligned}\mathbb{P}\Big(|\partial\mathcal{B}^{(G)}_t(q,k)| \geqslant \exp\big(2BA_\gamma^{k/2}\big) \mid \mathcal{E}^{(t)}_{2.4.5}(q)\Big)&\\ \leqslant \mathbb{E}[|\partial\mathcal{B}^{(G)}_t(q,k)| \mid \mathcal{E}^{(t)}_{2.4.5}(q)]\, \exp\big(-2BA_\gamma^{k/2}\big)&\\ \leqslant \exp\big(-BA_\gamma^{k/2}\big).& \qquad (2.4.10)\end{aligned}$$

Given (2.4.9), applying a union bound on (2.4.10) leads to the result in (2.4.7):

$$\begin{aligned}\mathbb{P}\Big(\bigcup_{k=1}^{K}\big\{|\partial\mathcal{B}^{(G)}_t(q,k)| \geqslant \exp\big(2BA_\gamma^{k/2}\big) \mid \mathcal{E}^{(t)}_{2.4.5}(q)\big\}\Big)&\\ \leqslant \sum_{k\in[K]} \exp\big(-BA_\gamma^{k/2}\big) \leqslant 2\exp(-B).&\end{aligned}$$

We are left with proving (2.4.9). We use the number of paths to $\partial\mathcal{B}^{(G)}_t(q,k)$ as an upper bound for the number of vertices in $\partial\mathcal{B}^{(G)}_t(q,k)$. For $k \leqslant K^*_t - M_{2.4.3}(t)$, conditionally on $\mathcal{E}^{(t)}_{2.4.5}$, there are only good paths of length k emanating from q. Let $\Pi^{(g)}_{k,t}(q)$ be the set of such paths. Recall Proposition 2.4.4, and observe that we can bound

$$\begin{aligned}\mathbb{E}[|\partial\mathcal{B}^{(G)}_t(q,k)| \mid \mathcal{E}^{(t)}_{2.4.5}(q)] &\leqslant \mathbb{E}\Big[\Big|\Pi^{(g)}_{k,t}(q)\Big| \mid \mathcal{E}^{(t)}_{2.4.5}(q)\Big]\\ &\leqslant \frac{1}{1-\delta_{2.4.5}}\mathbb{E}\Big[\Big|\Pi^{(g)(q)}_{k,t}\Big|\Big].\end{aligned}$$

We can bound $\mathbb{E}[|\Pi^{(g)(q)}_{k,t}|]$ using (2.4.5), with p defined in (2.4.5), as

$$\begin{aligned}\mathbb{E}\Big[\Big|\Pi^{(g)(q)}_{k,t}\Big|\Big] &\leqslant \sum_{w=\lceil \ell_k\rceil}^{t}\sum_{\pi_1=\ell_1}^{t}\cdots\sum_{\pi_{k-1}=\ell_{k-1}}^{t} p(q,\pi_1,\ldots,\pi_{k-1},w)\\ &=: \sum_{w=\lceil \ell_k\rceil}^{t} f_{k,t}(q,w).\end{aligned}$$

Intuitively, when $w \geqslant \ell_k$, $f_{k,t}(q,w)$ is an upper bound for the expected number of good paths from q to w of length k. From [69, Section 4.1] it follows that for $q \geqslant \ell_0$ there is a majorant of the form

$$f_{k,t}(q,w) \leqslant \alpha_k w^{-\gamma} + \mathbb{1}_{\{w>\ell_{k-1}\}}\beta_k w^{\gamma-1}, \tag{2.4.11}$$

where the sequences $\alpha_k, \beta_k, \ell_k$ are defined recursively as

$$\alpha_k := \begin{cases} \nu_{2.4.4}(\delta' t)^{\gamma-1} & \text{if } k=1, \\ c\left(\alpha_{k-1}\log(t/\ell_{k-1}) + \beta_{k-1}t^{2\gamma-1}\right) & \text{if } k>1, \end{cases} \tag{2.4.12}$$

$$\beta_k := \begin{cases} \nu_{2.4.4}(\delta' t)^{-\gamma} & \text{if } k=1, \\ c(\alpha_{k-1}\ell_{k-1}^{1-2\gamma} + \beta_{k-1}\log(t/\ell_{k-1})) & \text{if } k>1, \end{cases} \tag{2.4.13}$$

$$\ell_k := \begin{cases} \lceil \delta' t \rceil & \text{if } k=0, \\ \underset{x\in\mathbb{N}\setminus\{0,1\}}{\arg\max}\left\{\frac{1}{1-\gamma}\alpha_k x^{1-\gamma} \leqslant \frac{6\delta'}{\pi^2 k^2}\right\} & \text{if } k>0, \end{cases} \tag{2.4.14}$$

with a constant $c = c(\gamma, \nu_{2.4.4}) > 1$ chosen in [69, Lemma 1]. Recall $\gamma \in (1/2, 1)$. Using the above definitions and majorant in (2.4.11), we return to the bound (2.4.5), and see

$$\begin{aligned} \mathbb{E}\left[\left|\Pi_{k,t}^{(g)(q)}\right|\right] &\leqslant \sum_{w=\ell_k}^{t}\left(\alpha_k w^{-\gamma} + \mathbb{1}_{\{w>\ell_{k-1}\}}\beta_k w^{\gamma-1}\right) \\ &= \alpha_k\left(\ell_k^{-\gamma} + \sum_{w=\ell_k+1}^{t} w^{-\gamma}\right) + \beta_k \sum_{w=\ell_{k-1}+1}^{t} w^{\gamma-1}. \end{aligned} \tag{2.4.15}$$

Observe that for $a, b > 0$, $\mu \in (0,1)$

$$\sum_{w=a+1}^{b} w^{-\mu} \leqslant \int_a^b w^{-\mu}dw \leqslant \frac{b^{1-\mu}}{1-\mu},$$

so that we can bound (2.4.15) from above by

$$\begin{aligned} \mathbb{E}\left[\left|\Pi_{k,t}^{(g)(q)}\right|\right] &\leqslant \alpha_k\ell_k^{-\gamma} + \frac{\alpha_k\ell_k^{1-\gamma}}{1-\gamma}(t/\ell_k)^{1-\gamma} + \frac{\beta_k}{\gamma}t^{\gamma} \\ &=: T_1 + T_2 + T_3. \end{aligned} \tag{2.4.16}$$

As a result of (2.4.14), we obtain for the first term, since $\ell_k \geqslant 1$,

$$T_1 = \alpha_k\ell_k^{-\gamma} \leqslant \frac{6\delta'}{\pi^2 k^2}(1-\gamma)\ell_k^{-1} \leqslant \frac{3\delta'}{\pi^2 k^2}(1-\gamma) \leqslant \delta'. \tag{2.4.17}$$

To bound T_2 and T_3, we use a claim from [69, Theorem 2, (19)] that the sequence t/ℓ_k does not increase *too fast*, i.e., there exists $B^*(\gamma, \nu_{2.4.4})$ such that for t sufficiently large

$$t/\ell_k \leqslant \exp\left(BA_\gamma^{k/2}\right) \tag{2.4.18}$$

for any $B \geqslant B^*(\gamma, \nu_{2.4.4})$. Consider T_2, and substitute the bounds from (2.4.14) and (2.4.18)

$$\begin{aligned} \frac{\alpha_k \ell_k^{1-\gamma}}{1-\gamma}(t/\ell_k)^{1-\gamma} &\leqslant \frac{6\delta'}{\pi^2 k^2} \exp\left(B(1-\gamma)A_\gamma^{k/2}\right) \\ &\leqslant \delta' \exp\left(BA_\gamma^{k/2}\right). \end{aligned} \tag{2.4.19}$$

For T_3, we substitute (2.4.13) for β_k to obtain

$$\begin{aligned} T_3 = \frac{1}{\gamma}\beta_k t^\gamma &= \frac{c}{\gamma}\alpha_{k-1}(t/\ell_{k-1})^\gamma \ell_{k-1}^{1-\gamma} + \frac{c}{\gamma}\beta_{k-1} t^\gamma \log(t/\ell_{k-1}) \\ &=: T_{31} + T_{32}. \end{aligned} \tag{2.4.20}$$

We use (2.4.14) and (2.4.18) to bound

$$\begin{aligned} T_{31} = \frac{c}{\gamma}(t/\ell_{k-1})^\gamma \alpha_{k-1}\ell_{k-1}^{1-\gamma} &\leqslant \frac{6c\delta'(1-\gamma)}{\gamma\pi^2(k-1)^2} \exp\left(B\gamma A_\gamma^{(k-1)/2}\right) \\ &\leqslant \frac{c\delta'}{\gamma} \exp\left(BA_\gamma^{k/2}\right). \end{aligned} \tag{2.4.21}$$

Observe that by (2.4.12), $\beta_{k-1} \leqslant t^{1-2\gamma}\alpha_k/c$. Hence, the second term in (2.4.20) is bounded by

$$T_{32} = \frac{c}{\gamma}\beta_{k-1} t^\gamma \log(t/\ell_{k-1}) \leqslant \frac{1}{\gamma} t^{1-\gamma}\alpha_k \log(t/\ell_{k-1}).$$

By (2.4.14), if $\delta' < \pi^2/6$, then $\alpha_k \leqslant (1/\ell_k)^{1-\gamma}$. After substituting (2.4.18) for B sufficiently large

$$T_{32} \leqslant \frac{1}{\gamma}(t/\ell_k)^{1-\gamma} \log(t/\ell_{k-1}) \leqslant \frac{B}{\gamma} A_\gamma^{(k-1)/2} \exp\left(B(1-\gamma)A_\gamma^{k/2}\right).$$

As the first factor grows exponentially in k, by increasing B, we obtain $T_{32} \leqslant \exp\left(BA_\gamma^{k/2}\right)$. Combining this with (2.4.16), (2.4.17), (2.4.19), (2.4.20), and (2.4.21) gives us that there exists a constant $B_{2.4.5} > B^*(\gamma, \nu_{2.4.4}) > 0$ such that for $B \geqslant B_{2.4.5}$ we obtain the desired bound (2.4.9). □

We are ready to prove the lower bound of the conservative case.

Proof of Proposition 2.4.1. We recall $\mathcal{E}_{2.4.3}^{(t)}$ and $\mathcal{E}_{2.4.5}^{(t)}(q)$ and their error probabilities $\delta_{2.4.3}, \delta_{2.4.5}$ in (2.4.4), (2.4.6). First we introduce some notation to work with the models FPA, VPA, and GVPA simultaneously. Eventually we dinstinguish GVPA vs. FPA and VPA again. Recall ε from the statement of Proposition 2.4.1. Observe that if t is sufficiently large, by Proposition 2.4.3,

$$\mathbb{P}\big(d_t^{(G)}(u,v) \geqslant 2(1-\varepsilon/2)K_t^*\big) \geqslant 1-\delta_{2.4.3}.$$

In order to apply Lemma 2.4.5, relying on Proposition 2.4.4, for GVPA we set $\gamma_{2.4.4}$ as the solution of

$$\frac{\log\left(\gamma_{2.4.4}/(1-\gamma_{2.4.4})\right)}{|\log(\tau-2)|} = \frac{1}{1-\varepsilon/2}, \tag{2.4.22}$$

so that indeed $\gamma_{2.4.4} > 1/(\tau-1)$ as required for Proposition 2.4.4. For FPA and VPA, we set $\gamma_{2.4.4} = 1/(\tau-1)$, where τ is the power-law exponent of the considered model, as defined in Definition 1.2.4 and Definition 1.2.5. To avoid double notation, we define

$$K_t' := \begin{cases} \lfloor (1-\varepsilon/2)K_t^* \rfloor, & \text{for GVPA}(f), \\ K_t^* - M_{2.4.3}(t), & \text{for FPA and VPA.} \end{cases} \tag{2.4.23}$$

Apply Proposition 2.4.3 and Lemma 2.4.5 for $\delta_{2.4.3} = \delta_{2.4.5} = \delta/6$ so that with probability at least $1-\delta/6$ the neighbourhoods are disjoint and not *too large* for $k < K_t'$. Assume that t is so large that

$$\mathbb{P}\big(\mathcal{E}_{2.4.3}^{(t)} \cap \mathcal{E}_{2.4.5}^{(t)}(u) \cap \mathcal{E}_{2.4.5}^{(t)}(v)\big) \geqslant 1-\delta/2, \tag{2.4.24}$$

by a union bound on the complements of the events between brackets. Throughout the remainder of this proof, we write $A = \gamma_{2.4.4}/(1-\gamma_{2.4.4})$ and condition on the above event between brackets. On this event, the bound in (2.4.3) holds. We now bound the minimal weights in (2.4.3). Assume $B > B_{2.4.5}$. By (2.4.7) in Lemma 2.4.5, for $q \in \{u, v\}$, between $\partial\mathcal{B}_t^{(G)}(q,k)$ and $\partial\mathcal{B}_t^{(G)}(q,k+1)$ there are at most

$$|\partial\mathcal{B}_t^{(G)}(q,k)| \cdot |\partial\mathcal{B}_t^{(G)}(q,k+1)| \leqslant \exp\left(2B\big(1+\sqrt{A}\big)A^{k/2}\right) =: n_k$$

edges with cumulative error probability (over k) at most $2\exp(-B) =: \delta_B^{(1)}$. By $(*)$ in (2.3.24),

$$\mathbb{P}\Big(\bigcup_{k\in[K_t']}\Big\{\min_{j\in[n_k]} L_{j,k} \leqslant F_L^{(-1)}\big(n_k^{-1-\xi}\big)\Big\}\Big) \leqslant \sum_{k\in[K_t']} \exp\big(-2B\xi(1+\sqrt{A})A^{k/2}\big) \leqslant 2\exp\big(-2B\xi(1+\sqrt{A})\big) =: \delta_B^{(2)}, \qquad (2.4.25)$$

following from a union bound over $k \in [K_t']$. Combining (2.4.3) and (2.4.25) gives for $q \in \{u, v\}$

$$\mathbb{P}\Big(d_t^{(L)}(q, \partial\mathcal{B}_t^{(G)}(q, K_t')) \leqslant \sum_{k\in[K_t']} F_L^{(-1)}\Big(\exp\big(-2B(1+\xi)\big(1+\sqrt{A}\big)A^{k/2}\big)\Big)\Big) \leqslant \delta_B^{(1)} + \delta_B^{(2)}, \qquad (2.4.26)$$

when $B > B_{2.4.5}$. In particular, we choose B so that $\delta_B^{(1)} + \delta_B^{(2)} \leqslant \delta/4$. Recall (2.4.24) and the reasoning before (2.4.3), so that by the law of conditional probability and a union bound for $q \in \{u, v\}$ on the event in (2.4.26), we obtain with probability at least $1 - \delta$ that

$$d_L^{(t)}(u, v) \geqslant 2\sum_{k\in[K_t']} F_L^{(-1)}\Big(\exp\big(-2B(1+\xi)\big(1+\sqrt{A}\big)A^{k/2}\big)\Big) =: 2S_{K_t'}. \qquad (2.4.27)$$

It remains to show that $2S_{K_t'}$ is larger than the rhs between brackets in (2.4.1) and (2.4.2) for the corresponding models. Similarly to the upper bound, we rewrite the sum in $S_{K_t'}$ to match the summands in Q_t in (2.1.3), then we bound the sum from below by switching to integrals, apply a variable transformation, and go back to sums. Applying $(*)$ in (2.3.29) to $S_{K_t'}$ yields

$$S_{K_t'} \geqslant \int_1^{K_t'} F_L^{(-1)}\Big(\exp\big(-2B(1+\xi)\big(1+\sqrt{A}\big)A^{x/2}\big)\Big)\,dx. \qquad (2.4.28)$$

For the models FPA and VPA, $A = 1/(\tau-2)$, while for GVPA, $A = (1+\varepsilon')/(\tau-2)$, for some $\varepsilon' = \varepsilon'(\varepsilon)$. After the variable transformation $2B(1+\xi)(1+\sqrt{A})A^{x/2} = (\tau-2)^{-y/2}$ on the rhs of (2.4.28), the function over which we integrate matches the function in the sum in Q_t in (2.1.3), i.e.,

$$S_{K_t'} \geqslant \frac{1}{r}\int_{r+s(B)}^{rK_t'+s(B)} F_L^{(-1)}\Big(\exp\big(-(\tau-2)^{-y/2}\big)\Big)\,dy, \qquad (2.4.29)$$

where $s(B) := 2\log\big(2B(1+\xi)(1+\sqrt{A})\big)/|\log(\tau-2)|$, and by our choice of γ in (2.4.22)

$$r := \begin{cases} \dfrac{\log(A)}{|\log(\tau-2)|} = \dfrac{1}{1-\varepsilon/2}, & \text{for the model GVPA,} \\ 1, & \text{for the models FPA, VPA.} \end{cases} \tag{2.4.30}$$

We abbreviate $a_k := F_L^{(-1)}\big(\exp(-(\tau-2)^{-k/2}\big)$. Applying $(\star)$ from (2.3.29) to the rhs of (2.4.29) gives

$$S_{K'_t} \geqslant \frac{1}{r} \sum_{k=\lceil r+s(B)\rceil+1}^{\lfloor rK'_t+s(B)\rfloor} a_k.$$

For FPA and VPA, using $r=1$ and K'_t in (2.4.23), we bound

$$S_{K'_t} \geqslant \sum_{k=\lceil s(B)\rceil+2}^{K^*_t-M_{2.4.3}(t)+\lfloor s(B)\rfloor} a_k =: \widetilde{Q}_t.$$

Up to a shift in the boundaries, the above summands match the summands in Q_t in (2.1.3). To obtain (2.4.2), we should choose $M_{2.4.1}$ such that $Q_t - M_{2.4.1} \leqslant \widetilde{Q}_t$, which is equivalent to

$$\begin{aligned} M_{2.4.1} \geqslant Q_t - \widetilde{Q}_t &= \sum_{k=1}^{K^*_t} a_k - \sum_{k=\lceil s(B)\rceil+2}^{K^*_t-M_{2.4.3}(t)+\lfloor s(B)\rfloor} a_k \\ &= \sum_{k=1}^{\lceil s(B)\rceil+1} a_k + \sum_{k=K^*_t+1}^{K^*_t-M_{2.4.3}(t)+\lfloor s(B)\rfloor} a_k, \end{aligned}$$

where we define the second sum as 0 if $\lfloor s(B)\rfloor - M_{2.4.3} < 1$. The first sum on the rhs can be bounded by a constant. As the sequence a_k is decreasing, we shift the summation boundaries of the second sum on the rhs and choose

$$M_{2.4.1} := \sum_{k=1}^{\lceil s(B)\rceil+1} a_k + \sum_{k=1}^{-M_{2.4.3}(t)+\lfloor s(B)\rfloor} a_k.$$

Observe that we can bound the second sum here by a constant, since $M_{2.4.3}(t) = O(1)$ for FPA and VPA by Proposition 2.4.3. As a result, we bound the lhs (2.4.27) from above and obtain the result

$$\begin{aligned} \mathbb{P}\big(d_t^{(L)}(u,v) \geqslant 2(Q_t - M_{2.4.1})\big) &\geqslant \mathbb{P}\big(d_t^{(L)}(u,v) \geqslant 2\widetilde{Q}_t\big) \\ &\geqslant \mathbb{P}\big(d_t^{(L)}(u,v) \geqslant 2S_{K'_t}\big) \geqslant 1-\delta. \end{aligned}$$

We turn now to GVPA. Recalling K'_t from (2.4.23) and r from (2.4.30), we bound rK'_t in the upper summation boundary in (2.4.27) from below by

$$\frac{1}{1-\varepsilon/2}\left\lfloor\left(1-\frac{\varepsilon}{2}\right)K^*_t\right\rfloor \geqslant \frac{1}{1-\varepsilon/2}\left(\left(1-\frac{\varepsilon}{2}\right)K^*_t-1\right) \geqslant K^*_t - \frac{1}{1-\varepsilon/2}.$$

Thus, we further bound $S_{K'_t}$ by

$$S_{K'_t} \geqslant (1-\varepsilon/2)\sum_{k=\lceil 1/(1-\varepsilon/2)+s(B)\rceil+1}^{\lfloor K^*_t - 1/(1-\varepsilon/2)+s(B)\rfloor} a_k.$$

We rewrite the boundaries of the sum and use that $\varepsilon < 1$ and hence $1/(1-\varepsilon/2) \leqslant 2$, so

$$\begin{aligned} S_{K'_t} &\geqslant (1-\varepsilon/2)\Bigg(\sum_{k\in[K^*_t]} a_k + \sum_{k=K^*_t+1}^{\lfloor K^*_t-1/(1-\varepsilon/2)+s(B)\rfloor} a_k \\ &\qquad - \sum_{k\in[\lceil 1/(1-\varepsilon/2)+s(B)\rceil]} a_k\Bigg) \\ &\geqslant (1-\varepsilon/2)\Bigg(\sum_{k\in[K^*_t]} a_k - \sum_{k\in[2+\lceil s(B)\rceil]} a_k\Bigg), \end{aligned} \tag{2.4.31}$$

As the second sum in (2.4.31) is a constant, it can be bounded by $\varepsilon/2$ times the first sum in (2.4.31) when t is sufficiently large. Thus, we have that

$$d_L^{(t)}(u,v) \geqslant 2(1-\varepsilon/2)^2 \sum_{k\in[K^*_t]} a_k \geqslant 2(1-\varepsilon)\sum_{k\in[K^*_t]} a_k = 2(1-\varepsilon)Q_t$$

with probability $1-\delta$, for any fixed B and t sufficiently large. This is the asserted bound in (2.4.1). □

To finish the section, we prove the lower bound for the explosive class.

Proof of Proposition 2.4.2. Recall K^*_t from (2.1.3) and let $M_{2.4.3}(t)$ be the function we obtain by applying Proposition 2.4.3 for $\delta_{2.4.3} = \delta/2$. As a result, we obtain for any $K \leqslant K^*_t - M_{2.4.3}(t)$,

$$\mathbb{P}\Bigg(d_t^{(L)}(u,v) < \sum_{q\in\{u,v\}} d_t^{(L)}(q, \mathcal{B}_t^{(G)}(q,K))\Bigg) \leqslant \frac{\delta}{2}.$$

Let $a(t) := \min\{\kappa_{\delta/6}(t), K^*_t - M_{2.4.3}(t)\}$, where $\kappa_{\delta/4}(t)$ denotes the maximum number of generations in LWL to maintain a coupling with PA_t with probability at least $1-\delta/4$, as in Proposition 2.2.4 and 2.2.4. Hence, we obtain $d_t^{(L)}\big(q, \partial\mathcal{B}_t^{(G)}(q, a(t))\big) = \beta_{LWL^{(q)}_{a(t)}, a(t)}$ if the coupling is successful for $q \in \{u, v\}$ and thus

$$\mathbb{P}\left(d_t^{(L)}(u,v) < \sum_{i\in\{1,2\}} \beta_{LWL^{(i)}_{a(t)}, a(t)} \right) \leqslant \delta,$$

which finishes the proof. □

2.5 HOPCOUNT

In this section we prove Theorem 2.2.6. It follows from an adaptation of the proof of Theorem 2.2.5.

Proof of Theorem 2.2.6. Since the hopcount is at least the graph distance, Proposition 2.4.3 implies that for any $\varepsilon > 0$

$$\mathbb{P}\big(d_t^{(H)}(u,v) \geqslant 2(1-\varepsilon)K_t^*\big) \longrightarrow 1, \qquad \text{as } t \to \infty, \tag{2.5.1}$$

for the model GVPA(f), and shows lower tightness for FPA and VPA. It suffices to prove the matching upper bounds. By rescaling the weights to L/a, all weights are at least one. For any two vertices u and v, the shortest path with the unscaled weights uses the same edges as the shortest path with the scaled weights. Hence, $d_t^{(H)}(u,v) \leqslant d_t^{(L/a)}(u,v)$. Observe that for any $x > 0$, $F^{(-1)}_{L/a}(x) = 1 + F^{(-1)}_{L/a-1}(x)$. Fix a small $\delta > 0$ and recall steps (i), (iii), and (iv) from the proof of Proposition 2.3.1. Analogously to the reasoning leading to (2.3.36), for any s_0 there exists constants $C_{2.3.3}, C_{2.3.5}, M$ such that for t sufficiently large

$$\mathbb{P}\left(d_t^{(L/a)}(u,v) \leqslant 2K_t^* + 2\sum_{k=h_\tau(s_0)}^{h_\tau(s_0)+K_t^*} F^{(-1)}_{L/a-1}\left(\exp\left(\tau-2\right)^{-k/2}\right) + M \right) \geqslant 1-\delta/2,$$

where we used the upper bound in (2.3.3) from Proposition 2.3.4 on the weight of the segment reaching the inner core. Thus, we did *not* shift the summation bounds as in step (ii) in the proof of Proposition 2.3.1. If $I(L/a-1) < \infty$, then the above sum is finite and thus yields the result for tightness in (2.2.5). If the sum is infinite, we choose s_0 so large that all the terms are bounded by a fixed $\varepsilon/2 > 0$, hence the sum is bounded by $\varepsilon K_t^*/2$. As K_t^* is increasing in t, for t sufficiently large, $\varepsilon K_t^* > M$. Thus,

$$\mathbb{P}\left(d_t^{(H)}(u,v) \leqslant 2(1+\varepsilon)K_t^*\right) \geqslant \mathbb{P}\left(d_t^{(L/a)}(u,v) \leqslant 2(1+\varepsilon)K_t^*\right) \geqslant 1-\delta.$$

Combining this upper bound with the lower bound (2.5.1) yields the desired asymptotics in (2.2.4). □

2.6 CONSERVATIVE WEIGHTS ON THE LWL

We are left with proving Theorem 2.2.8.

Proof of Theorem 2.2.8. We prove (2.2.6) in Theorem 2.2.8 for the LWL of FPA and VPA such that the weight distributions satisfies (2.2.3). At the end of the proof we describe briefly how to prove (2.2.7). The proof is split in an upper bound and a lower bound. Both bounds rely on a coupling of the LWL to the finite graph for some large t and follow from adaptations of the results in Section 2.3 and Section 2.4. Fix k and $\delta > 0$.

For the upper bound we show that there exists $M = M(\delta)$, not depending on k, such that

$$\mathbb{P}\big(\beta_{\mathrm{LWL},k}(\circledcirc) \leqslant Q^{(k)} + M\big) \geqslant 1 - \delta/2, \tag{2.6.1}$$

where

$$Q^{(k)} := \sum_{i=1}^{k} F_L^{(-1)}\left(\exp\left(-(\tau-2)^{-i/2}\right)\right).$$

Recall Q_t in (2.1.3) and observe that $Q_t = Q^{(K_t^*)}$. Set $\delta_{2.2.4} = \delta_{2.3.3} = \delta/12$, $\delta_{2.3.4} = \delta/6$, and $\alpha = 1 - \delta/12$. Let t be sufficiently large such that all the following hold.

(i) By Proposition 2.2.4 the LWL and the graph neighbourhood of a typical vertex q can be coupled up to generation k with probability at least $1 - \delta/12$. Recall $\beta_{\mathrm{LWL},k}(\circledcirc)$ from Definition 2.2.2. Conditionally on the event that the coupling is successful,

$$\beta_{\mathrm{LWL},k}(\circledcirc) = d_t^{(L)}(q, \partial\mathcal{B}_t^{(G)}(q,k)).$$

Thus, the upper bound in (2.6.1) follows if we show that

$$\mathbb{P}\big(d_t^{(L)}(q, \partial\mathcal{B}_t^{(G)}(q,k)) \leqslant Q^{(k)} + M \mid \text{coupling successful}\big) \geqslant 1 - \tfrac{11}{12}\delta.$$

Since q is a typical vertex, the event $\{q < \alpha t\}$ holds with probability at least $1 - 11\delta/12$. By a union bound, the event

$$\mathcal{E}^{(i)} := \{q < \alpha t\} \cap \{\text{coupling successful}\}$$

holds with probability at least $1 - 5\delta/6$, leaving to show that

$$\mathbb{P}\big(d_t^{(L)}(q, \partial\mathcal{B}_t^{(G)}(q,k)) \leqslant Q^{(k)} + M \mid \mathcal{E}^{(i)}\big) \geqslant 1 - \delta/3. \tag{2.6.2}$$

(ii) By Proposition 2.3.3 for $s_{2.3.3} = s_0$ that we choose below in (iv), there exists $C_{2.3.3} = C_{2.3.3}(\delta, s_0)$, such that, for

$$\mathcal{E}_1^{(\mathrm{ii})} := \{\exists q' \in \mathcal{B}^{(G)}_{(1-\alpha)t}(q, C_{2.3.3}) : D_{(1-\alpha)t}(q') \geqslant s_0\},$$

$\mathbb{P}(\mathcal{E}_1^{(\mathrm{ii})} \mid \mathcal{E}^{(\mathrm{i})}) \geqslant 1 - \delta/12$. All edges on a possible path of length $C_{2.3.3}$ from q to q' are equipped with i.i.d. copies of L. Hence, there exists $M^{(\mathrm{ii})} = M^{(\mathrm{ii})}(\delta, C_{2.3.3}) > 0$ such that

$$\mathbb{P}\Big(\sum_{i=1}^{C_{2.3.3}} L_i \leqslant M^{(\mathrm{ii})}\Big) \geqslant 1 - \delta/12.$$

Let

$$\mathcal{E}^{(\mathrm{ii})} := \Big\{\exists q' \in \mathcal{B}^{(G)}_{(1-\alpha)t}(q, C_{2.3.3}) : d^{(L)}_{(1-\alpha)t}(q, q') \leqslant M^{(\mathrm{ii})}\Big\} \cap \mathcal{E}_1^{(\mathrm{ii})}.$$

By a union bound we obtain that

$$\mathbb{P}(\mathcal{E}^{(\mathrm{ii})} \mid \mathcal{E}^{(\mathrm{i})}) \geqslant 1 - \delta/6.$$

(iii) With K_t in (2.3.10), the inequality $C_{2.3.3} + 2K_t \geqslant k$ holds for all sufficiently large t. Then, we construct a greedy path to the inner core as described in the proof of Proposition 2.3.4.

(iv) By Proposition 2.3.4 there is an $s_0 = s_0(\delta, \varepsilon_{2.3.4}) > 0$, such that for $q' \in [\alpha t]$ with $D_{\alpha t}(q') \geqslant s_0$, the weighted distance to the inner core in PA_t is not too large. The graph PA_t is coupled to LWL for at least k generations. Hence, if the coupling succeeds, then any path of $k - C_{2.3.3}$ many edges emanating from q' is also present in the LWL, as q' is at graph distance at most $C_{2.3.3}$ from q by (ii). Hence, we can use the first $k - C_{2.3.3}$ edges of the greedy path described in Definition 2.3.6 and follow the proof of Proposition 2.3.4 to estimate its total weight. By (iii), $k \leqslant 2K_t + C_{2.3.3}$, so ultimately, we obtain that for some that $M^{(\mathrm{iv})} > 0$, that is not dependent on k, that

$$\mathbb{P}\Big(d_t^{(L)}\big(q', \partial\mathcal{B}_t^{(G)}(q', 0 \vee (k - C_{2.3.3}))\big) \leqslant Q^{(k)} + M^{(\mathrm{iv})}\Big) \geqslant 1 - \delta/6.$$

Denote the above event between brackets by $\mathcal{E}^{(\mathrm{iv})}$.

On the event $\mathcal{E}^{(\mathrm{i})} \cap \mathcal{E}^{(\mathrm{ii})} \cap \mathcal{E}^{(\mathrm{iv})}$, there is a path from q to a vertex in $\mathcal{B}_t^{(G)}(q', k)$, whose total weight is bounded from above by $Q^{(k)} + M^{(\mathrm{ii})} +$

$M^{(iv)}$. If the coupling succeeds, this path is also present in the LWL by Proposition 2.2.4.

To conclude the proof of (2.6.1), we recall that showing (2.6.1) was reduced to showing (2.6.2), which holds for $M = M^{(ii)} + M^{(iv)}$, since $\mathbb{P}(\mathcal{E}^{(ii)} \cap \mathcal{E}^{(iv)} \mid \mathcal{E}^{(i)}) \geqslant 1 - \delta/3$ by a union bound.

We proceed to the lower bound of Theorem 2.2.8, i.e., we show that for some $M = M(\delta)$

$$\mathbb{P}\big(\beta_{\mathrm{LWL},k}(\circledcirc) \leqslant Q^{(k)} - M\big) \geqslant 1 - \delta/2. \tag{2.6.3}$$

Let $\delta_{2.4.5} = \delta/8$, and let B in Lemma 2.4.5 be sufficiently large. Let t be so large that

1. $K_t^* - M_{2.4.5}(t) \geqslant k$.
2. by Proposition 2.2.4 the local weak limit can be coupled to the neighbourhood of radius k of a typical vertex q with probability at least $1 - \delta/8$. Conditionally on the event that this coupling is successful, $\beta_{\mathrm{LWL},k}(\circledcirc) = d_t^{(L)}(q, \partial\mathcal{B}_t^{(G)}(q,k))$.

Hence, by (2) it is for (2.6.3) sufficient to show that there exists M such that

$$\mathbb{P}\big(d_t^{(L)}(q, \partial B_t^{(G)}(q,k)) \leqslant Q^{(k)} - M \mid \{\text{coupling successful}\}\big) \geqslant 1 - 3\delta/8.$$

This follows from easy modifications of the proof of Proposition 2.4.1. We leave it to the reader to fill in the details.

Now that (2.6.1) and (2.6.3) have been established, (2.2.6) follows for weight distributions satisfying (2.2.3). For other weight distributions, (2.2.7) can be proven for FPA and VPA using the same steps. For the model GVPA, one can prove (2.2.7) using the same couplings from GVPA to VPA as in the proofs of Proposition 2.3.1 and Proposition 2.4.1. We leave it to the reader to fill in the details. □

3 DISTANCE EVOLUTIONS

Based on [153]:
Distance evolutions in preferential attachment models,
J. Jorritsma, J. Komjáthy,
The Annals of Applied Probability, 32(6), 4356-4397, 2022.

3.1 INTRODUCTION

In 1999, Faloutsos, Faloutsos, and Faloutsos studied the topology of the early Internet network, discovering power-laws in the degree distribution and short average hopcounts between routers [88]. Undoubtedly, the Internet has grown immensely in the last two decades. It would be interesting to investigate what has happened to the graph structure surrounding the early routers (or their direct replacements) that were already there in 1999, ever since. Natural questions about the evolving graph surrounding these early routers are:

- How did the number of connections of the routers gradually change? Did the early routers become important hubs in the network?
- Can we quantify the number of hops needed to connect two early routers? Particularly, did the hopcount decrease or increase while their importance in the network changed, and more and more connections arrived? If so, how did the distance gradually evolve?

These kinds of questions drive the mathematics in the present chapter. We initiate a research line that studies how certain graph properties defined on a fixed set of vertices evolve as the surrounding graph grows. We consider the weighted-distance evolution in two classical *preferential attachment models* (PAMs), the fixed preferential attachment model (FPA) and the variable preferential attachment model, both defined in Section 1.2.2. Studying the evolution of a property on fixed vertices may sound as a natural mathematical question. Yet, only the evolution of the degree of fixed vertices has been addressed so far in the PAM literature [70, 193].

The *graph-distance evolution* is a discrete-time stochastic process that we denote by $\big(d_{t'}^{(G)}(u_t, v_t)\big)_{t'\geqslant t}$ and define formally in Definition 3.3.2 below.

Here, u_t and v_t are two *typical vertices*, i.e., they are sampled uniformly at random from the vertices in largest component in PA_t. The graph distance $d^{(G)}_{t'}(u_t, v_t)$ is the number of edges on the shortest path between u_t and v_t that uses only vertices that arrived at latest at time t'. The distance evolution $\big(d^{(G)}_{t'}(u_t, v_t)\big)_{t' \geqslant t}$ is nonincreasing in t', since new edges arrive in the graph that may form a shorter path between u_t and v_t. We will now state our main result for the graph-distance evolution. To describe the graph distance we define for $t' \geqslant t$, writing $a \vee b := \max\{a, b\}$,

$$K_{t,t'} = 2 \left\lfloor \frac{\log\log(t) - \log\big(\log(t'/t) \vee 1\big)}{|\log(\tau-2)|} \right\rfloor \vee 1. \tag{3.1.1}$$

A sequence of random variables $(X_n)_n$ is tight if $\lim_{M\to\infty} \sup_n \mathbb{P}(|X_n| > M) = 0$.

Theorem 3.1.1 (Graph-distance evolution)**.** *Consider the preferential attachment model with power-law exponent $\tau \in (2,3)$. Let u_t, v_t be two typical vertices in PA_t. Then*

$$\left(\sup_{t' \geqslant t} \left| d^{(G)}_{t'}(u_t, v_t) - 2K_{t,t'} \right| \right)_{t \geqslant 1} \tag{3.1.2}$$

is a tight sequence of random variables.

Theorem 3.1.1 tracks the evolution of $d^{(G)}_{t'}(u_t, v_t)$ as time passes and the graph around u_t and v_t grows, since in (3.1.2) the supremum is taken over t'. Below, in Theorem 3.3.3, we extend Theorem 3.1.1 to a general setting and consider the so-called *weighted-distance evolution* $\big(d^{(L)}_{t'}(u_t, v_t)\big)_{t' \geqslant t}$. There, we equip every edge in the graph with a weight, an i.i.d. copy of a random variable L. We consider the evolution of the weighted distance, the sum of the weights along the least-weighted path from u_t to v_t that is present at time t'. We obtain results for *any* non-negative random variable L that serves as edge-weight distribution.

As a consequence of Theorem 3.1.1, we obtain a hydrodynamic limit, i.e., a scaled version of the distance evolution converges under proper time scaling uniformly in probability to a non-trivial deterministic function.

Corollary 3.1.2 (Hydrodynamic limit for the graph-distance evolution)**.** *Consider the preferential attachment model with power-law exponent $\tau \in (2,3)$. Define $T_t(a) := t\exp(\varepsilon \log^a(t))$ for $a \geqslant 0$ and arbitrary $\varepsilon > 0$. Let u_t, v_t be two typical vertices in PA_t. Then, as t tends to infinity,*

$$\sup_{a \geqslant 0} \left| \frac{d^{(G)}_{T_t(a)}(u_t, v_t)}{\log\log(t)} - (1 - \min\{a, 1\}) \frac{4}{|\log(\tau-2)|} \right| \xrightarrow{\mathbb{P}} 0. \tag{3.1.3}$$

This can be verified by computing the value of $K_{t,T_t(a)}$ using (3.1.1), substituting this value into (3.1.2), and then dividing all terms by $\log\log(t)$.

Observe that in Corollary 3.1.2 all $\log\log(t)$-terms have vanished when $a = 1$. The following consequence of Corollary 3.1.1 illustrates the rate at which smaller order terms appear and vanish. In particular, the graph distance is of constant order as soon as t'/t is of polynomial order in t.

Corollary 3.1.3 (Lower-order terms). *Consider the preferential attachment model with power-law exponent* $\tau \in (2,3)$. *Let* u_t, v_t *be two typical vertices in* PA_t. *Let* $g(t)$ *be any function that is bounded from above by* $2K_{t,t}$, *and set* $T_g(t) := t^{1+(\tau-2)^{-g(t)/4}}$. *Then, for two typical vertices* u_t *and* v_t *in* PA_t,

$$\left(d^{(G)}_{T_g(t)}(u_t, v_t) - g(t)\right)_{t\geqslant 1}$$

is a tight sequence of random variables.

Indeed, setting any $g(t)$ that tends to infinity with t results in a time scale $T_g(t) \sim t^{1+\varepsilon_{g(t)}}$ for some $\varepsilon_{g(t)} \to 0$ as t tends to infinity.

Remark 3.1.4. *Using a similar martingale argument as in Lemma 3.5.3 below for the degree of the vertices* u_t *and* v_t, *one can show that when* $t' = \Theta(t^{2/(3-\tau)})$, *there will be a vertex that connects to both* u_t *and* v_t. *Hence, the distance evolution settles on two.*

3.2 METHODOLOGY

The proof of Theorem 3.1.1 and Theorem 3.3.3 below consist of a lower bound and an upper bound. For the upper bound we prove that at all times $t' \geqslant t$ there is a path from u_t to v_t that has length at most $2K_{t,t'} + M_G$ (from (3.1.1)) for some constant M_G and contains only vertices born, i.e., arrived, before time t'. We first heuristically argue that the scaling for the graph distance in (3.1.3) is a natural scaling. After that, we turn to the difficulties that arise in handling the dynamics. The degree $D_{q_t}(t')$ of a vertex $q_t \in \{u_t, v_t\}$ at time t' is of order $(t'/q_t)^{1/(\tau-1)}$. Writing $t' = T_t(a) := t\exp(\log^a(t))$ and approximating the birth time of the uniform vertex q_t by t, we have that

$$D_{q_t}\big(T_t(a)\big) \approx \exp\left(\log^a(t)/(\tau-1)\right) =: s^{(a)}.$$

Generally, a vertex of degree s is at graph distance two from many vertices that have degree approximately $s^{1/(\tau-2)}$. This allows for an iterative *two-connector procedure* that starts from an initial vertex with degree at least

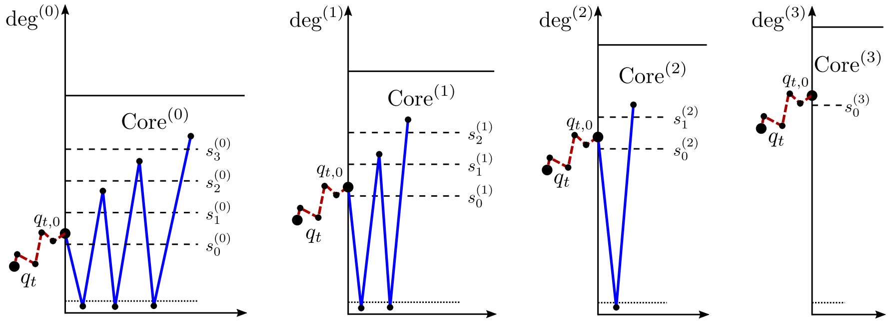


Figure 2: Construction of a path from q_t to the inner core at the times $(t_i)_{i\leqslant 3}$, where $t_0 = t$. The y-axis represents the degree of vertices at time t_i and the connected dots the vertices on the path segments from q_t via $q_{t,0}$ to the inner core. The x-axis represents graph distance from $q_{t,0}$. The black dashed horizontal lines represent the degree-threshold sequence, while the continuous black lines represent the maximal degree in the graph at time t_i. The degree of $q_{t,0}$, the maximal degree in the graph and the degree threshold for the inner core all increase over time. The degree of vertex $q_{t,0}$ satisfies the inner-core threshold at time t_3. The red dashed segment from q_t to $q_{t,0}$ is the same for all i, while the blue segment from $q_{t,0}$ is constructed at the times $(t_i)_{i\leqslant 3}$.

$s^{(a)}$ and reaches in the k-th iteration a vertex with degree approximately $(s^{(a)})^{(\tau-2)^{-k}}$. We call $k \mapsto (s^{(a)})^{(\tau-2)^{-k}}$ the degree-threshold sequence. At each iteration, we greedily extend the path by two edges, arriving to such a higher-degree vertex. In the edge-weighted version, these two edges are chosen to minimize the total edge-weight among all such two edges. This two-connector procedure to vertices with increasing degree is iterated until the well-connected inner core is reached. The inner core is the set of vertices with degree roughly $T_t(a)^{1/(2(\tau-1))}$ at time $T_t(a)$. Hence, for $a < 1$, the total number of iterations to reach the inner core is approximately

$$\min\left\{k : (s^{(a)})^{(\tau-2)^{-k}} \geqslant T_t(a)^{1/(2(\tau-1))}\right\} \approx (1-a)\frac{\log\log(t)}{|\log(\tau-2)|}. \quad (3.2.1)$$

By construction, the graph distance from u_t and v_t to the inner core is two times the right-hand side (rhs) in (3.2.1). The graph and weighted distance between vertices in the inner core are negligible, yielding the scaling in (3.1.3), as well as the upper bounds in Theorems 3.1.1 and 3.3.3.

There are three main difficulties in the outlined procedure. Firstly, it is not good enough to start the two-connector procedure from u_t (or

v_t) because the error terms coming from controlling the growth of the degree of u_t (or v_t) at t' close to t are too large. To resolve this, we start the procedure from a vertex – say $q_{t,0}$ – that has degree at least $s_0^{(0)}$ at time t for some large but universally bounded constant $s_0^{(0)}$. The segment between q_t and $q_{t,0}$ is fixed for all $t' \geqslant t$, so that we only have to account for a possible error once. Secondly, we need to bound the degree of the vertex $q_{t,0}$ from below over the entire time interval $[t, \infty)$, not just at a specific time t'. For this we employ martingale arguments. Lastly, to make the error probabilities summable in t', we argue that the two-connector procedure does not have be executed for every time $t' \geqslant t$, but only along a specific subsequence of times $(t_i)_{i \geqslant 0}$, where

$$t_i \in \left[t \exp\left((\tau-2)^{-i+1}\right), t \exp\left((\tau-2)^{-i-1}\right)\right].$$

This sequence is chosen such that at time t_{i+1} one iteration less than at time t_i is needed to reach the inner core from the initial vertices, and these are exactly the times when $K_{t,t'}$ crosses an integer and hence a previously present path is no longer short enough. See Figure 2 for a sketch. On the time scale $T_t(a)$, the number of iterations scales linearly in $a \in [0, 1]$.

For the lower bound we first bound the probability that the graph distance $d_{t'}^{(G)}(u_t, v_t)$ is ever too short and then extend it to weighted distances. To estimate the probability of a too short path being ever present, we develop a refined truncated path-counting method inspired by [69]. Let, for a fixed t, $(\ell_{k,t'}^{(t)})_{k \geqslant 0, t' \geqslant t}$ be an array of birth times, i.e., arrival times of vertices. The path-counting method first excludes possible paths from u_t to v_t that are unlikely to be present in $PA_{t'}$, called *bad paths*. A bad path of length k reaches a vertex born before time $\ell_{k,t'}^{(t)}$ using only vertices born before t'. The longer a path is, the more likely it is that an old vertex can be reached. Moreover, as the graph grows, it becomes more likely that there is a short path to an old vertex. The array of birth times $(\ell_{k,t'}^{(t)})_{k \geqslant 0, t' \geqslant t}$ is therefore nonincreasing in both parameters. Among the other possible paths that are too short, the *good paths*, the method counts the expected number of paths from u_t to v_t that are present in $PA_{t'}$. More precisely, the expected number of these paths of length at most $2K_{t,t'} - M_G$ is shown to be much smaller than one for some $M_G > 0$. The decomposition of good and bad paths is done for every $t' > t$, in an interlinked way. The crucial observation is that if there is no too short path present at time t', but there is a too short path present at time $t' + 1$, then the vertex labelled $t' + 1$ must be on this connecting path and thus it must be either on a bad path or on a too short good path. This trick allows us to develop a first moment

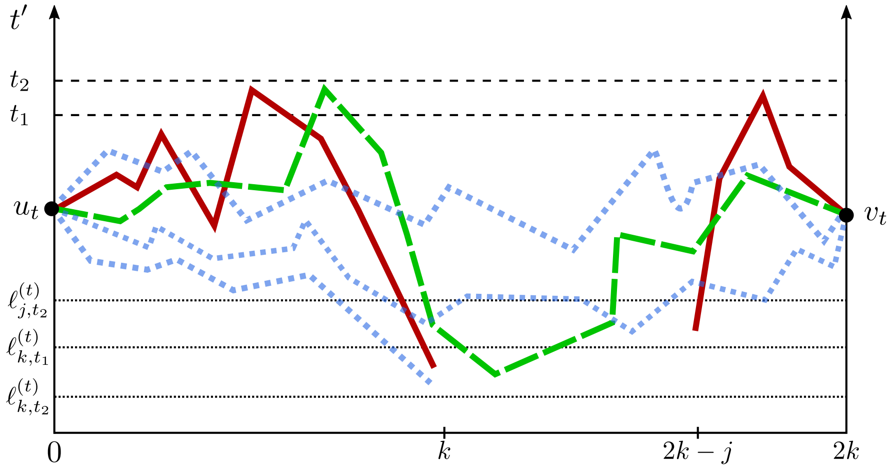


Figure 3: Good and bad path decomposition for the lower bound. The y-axis represents arrival time of vertices and the x-axis the graph distance from u_t. Bad paths are displayed in red, good paths are green and dashed. The blue dotted paths represent possible paths that are absent at time t_1. Let $t_2 > t_1$ and $k > j$. The black tiny-dotted horizontal lines represent the birth-threshold array $(\ell^{(t)}_{k,t'})_{k\geqslant 0,t'\geqslant t}$ which is decreasing in t' at the times t_1, t_2 and also decreasing in k. At time t_1 there is neither a bad path of length at most k present, nor a good path of length $2k$ that connects u and v. Then, if u and v are at time t_2 at graph distance $2k$, there must be either a bad path of length at most k emanating from u or v that traverses a vertex in $(t_1, t_2]$, or there must be a short good path traversing such a vertex. Observe that the good path is allowed to traverse a vertex in $[\ell^{(t)}_{k,t_2}, \ell^{(t)}_{k,t_1})$. In particular, this holds for $t_2 = t_1 + 1$.

method much sharper than a union bound simply over t', since we only need to bound the expected number of bad or too short good paths that are restricted to pass through the newly arrived vertex t'. These bounds are a factor $1/t'$ smaller than similar bounds without the restriction. As a result, the error bound is summable in t' and tends to zero as t tends to infinity. See Figure 3 for a sketch of the argument.

To extend the result from graph distances to weighted distances for Theorem 3.3.3 below, we observe that if the graph distance between u_t and v_t is at least $2K_{t,t'} - M_G/2$, then the graph neighbourhoods of radius $K_{t,t'} - M_G/2$ must be disjoint. A path that connects u_t to v_t must cross the boundaries of these graph neighbourhoods. We bound the number of vertices at distance precisely k from $q_t \in \{u_t, v_t\}$ from above, for

$k \leqslant K_{t,t'} - M_G/2$. This allows to bound the weight of the least-weight edge between vertices at distance k and $k+1$ from q_t from below. The sum of these minimal weight bounds is then a lower bound to reach the boundary. However, the error probabilities are not summable in t'. To resolve this, we show that it is sufficient to consider only a subsequence of times, similarly to the upper bound.

Organisation

In the next section we rigorously state the results. The lower bound is proven in Section 3.4. In Section 3.5 we present the proof of the upper bound.

3.3 GENERAL RESULTS

We now formalize the notion of paths for $(\mathrm{PA}_t)_{t \geqslant 1}$. This definition is used below to define distance evolutions.

Definition 3.3.1 (Paths)**.** We call a vertex tuple $(\pi_0, \dots, \pi_n) =: \boldsymbol{\pi}$ a q-path if $\pi_0 = q$, and we call it a (u,v)-path if $\pi_0 = u$, $\pi_n = v$, and $u \neq v$. Let $t' \in \mathbb{N}$. The path $\boldsymbol{\pi}$ is called t'-possible if $\max_{i \leqslant n} \pi_i \leqslant t'$ and t'-present if it is t'-possible and all edges $\{(\pi_0, \pi_1), \dots, (\pi_{n-1}, \pi_n)\}$ are present in $\mathrm{PA}_{t'}$.

For $u_t, v_t \in V_t$, let $\Omega^{(t')}(u_t, v_t) := \{\boldsymbol{\pi} : \boldsymbol{\pi} \text{ is a } t'\text{-present } (u_t, v_t)\text{-path}\}$ denote the set of t'-present paths. Since the edge set and vertex set are increasing in t', new paths between u_t and v_t emerge. Hence, we have that $\Omega^{(t')}(u_t, v_t) \subseteq \Omega^{(\tilde{t})}(u_t, v_t)$ for $\tilde{t} \geqslant t'$. Similar to Chapter 2, we equip every edge with a weight, an i.i.d. copy of a non-negative random variable L. Recalling the definitions of graph distance $d_{t'}^{(G)}$ and weighted distance $d_{t'}^{(L)}$ from Definition 2.2.1, we define distance evolutions.

Definition 3.3.2 (Distance evolutions)**.** Consider the edge-weighted graph $\mathrm{PA}_t = (V_t, E_t)$ where every edge e is equipped with a weight L_e. Let u_t and v_t be two typical vertices, i.e., they are sampled uniformly at random from the largest component in PA_t. We call the discrete-time stochastic processes $\big(d_{t'}^{(G)}(u_t, v_t)\big)_{t' \geqslant t}$ and $\big(d_{t'}^{(L)}(u_t, v_t)\big)_{t' \geqslant t}$ the graph-distance evolution and weighted-distance evolution, respectively.

The following function will be used to describe the weighted distance evolution. Define for $a, b \in \mathbb{N}$

$$\mathcal{Q}(a, b] := \sum_{k=a+1}^{b} F_L^{(-1)}\big(\exp\big(-(\tau-2)^{-k/2}\big)\big), \tag{3.3.1}$$

so that $\mathcal{Q}(a, b]$ is a sum consisting of $b-a$ terms. Recall $K_{t,t'}$ from (3.1.1) that describes the graph distance. We define for $t' \geqslant t$ its weighted-distance counterpart

$$Q_{t,t'} := \mathcal{Q}(K_{t,t} - K_{t,t'}, K_{t,t}]. \tag{3.3.2}$$

Lastly, we recall for the edge-weight distribution F_L the explosion characteristic $I_1(L)$ from (2.1.3), and the tightness-criterion $I_2(L)$ from (2.2.3). We are mostly interested in distributions satisfying $I_1(L) = \infty$ to observe how the distance decreases from $\Theta(\log\log(t))$ to constant order: by Theorem 2.2.10 the typical weighted distance is already of constant order if $I_1(L) < \infty$, making it less suitable to study evolutions. Observe also that in this case $Q_{t,t'}$ in (3.3.2) is bounded from above by some constant.

Theorem 3.3.3 (Main result). *Consider the preferential attachment model with power-law exponent $\tau \in (2, 3)$. Equip every edge upon creation with an i.i.d. copy of the non-negative random variable L. Let u_t, v_t be two typical vertices at time t. If $I_2(L) < \infty$, then*

$$\Big(\sup_{t' \geqslant t} \big|d_{t'}^{(L)}(u_t, v_t) - 2Q_{t,t'}\big|\Big)_{t \geqslant 1} \tag{3.3.3}$$

is a tight sequence of random variables. Regardless of the value of $I_2(L)$, for any $\delta, \varepsilon > 0$, there exists $M_L > 0$ such that

$$\mathbb{P}\Big(\forall t' \geqslant t : 2Q_{t,t'} - M_L \leqslant d_{t'}^{(L)}(u_t, v_t) \leqslant 2(1+\varepsilon)Q_{t,t'} + M_L\Big) \leqslant \delta. \tag{3.3.4}$$

Theorem 3.3.3 tracks the evolution of $d_{t'}^{(L)}(u, v)$ as time passes and the graph around u and v grows, since in (3.3.3) the supremum is taken over t' and t' is inside the $\mathbb{P}$-sign in (3.3.4). It is the $(1+\varepsilon)$-factor in the upper bound in (3.3.4) that makes (3.3.4) different from (3.3.3). Thus, the lower bound is tight for any non-negative weight distribution. A special case of Theorem 3.3.3 is when the edge-weight distribution $L \equiv 1$. Then the weighted distance and graph distance coincide, yielding Theorem 3.1.1, since $I_2(1) = 0$.

Observe that $2Q_{t,t'} = 2\mathcal{Q}(K_{t,t} - K_{t,t'}, K_{t,t}]$ in (3.3.2) could be seen as two sums, each consisting of $K_{t,t'}$ terms: the number of terms in $Q_{t,t'}$ is equal to the number of edges on the shortest graph-distance path. The additive constant M_L ensures that there will be many almost-shortest paths, from which we are able to choose one with low edge-weight. As time passes, the degrees of u_t and v_t increase, so that it becomes more likely that there are edges close to u_t and v_t that have small edge-weights. Since the terms in $Q_{t,t'}$ are decreasing in k, this intuitively explains that $Q_{t,t'}$ consists of the *smallest* $K_{t,t'}$ terms of $Q_{t,t}$, rather than the *largest* $K_{t,t'}$ terms of the sum defining $Q_{t,t}$. However, if $L \equiv 1 + X$ for some random variable X that satisfies $I_1(X) < \infty$ (e.g., X exponential, gamma, or a power of uniform on $[0, 1]$), then $|K_{t,t'} - Q_{t,t'}| \leqslant M$ for some constant M. Consequently, the graph distance and weighted distance are of the same order (up to additive constants). This phenomenon has also been observed for the Configuration Model [21]. As a result, for weight distributions with $I_1(X) < \infty$ the location of the summation interval in (3.3.2) does not influence the main result: there exists a constant M_1 such that $\mathcal{Q}(0, K_{t,t'}] - \mathcal{Q}(K_{t,t} - K_{t,t'}, K_{t,t}] \leqslant M_1$ for all $t' \geqslant t \geqslant 0$. For the other case, if $L = 1 + X$ such that $I_1(X) = \infty$, such a constant does not exist. For such distributions, the fact that the lower summation boundary in (3.3.2) is shifted to $K_{t,t} - K_{t,t'}$ from 0 matters and influences the growth rate. As an example, we set L such that the terms in the sum in (3.3.1) are equal to $1 + 1/k$, yielding $\mathcal{Q}(0, K_{t,t'}] \approx K_{t,t'} + \log(K_{t,t'})$, while for t' large enough that $K_{t,t} - K_{t,t'} \gg 1$:

$$\mathcal{Q}(K_{t,t} - K_{t,t'}, K_{t,t}] \approx K_{t,t'} + \log(K_{t,t}) - \log(K_{t,t} - K_{t,t'}) \ll \mathcal{Q}(0, K_{t,t'}].$$

We now recall the hydrodynamic limit for the graph-distance evolution in Corollary 3.1.2. A similar limit can be derived for the weighted-distance evolution if the weight distribution satisfies $I_1(L) = \infty$. The proper scaling and the constant prefactor, similar to (3.1.3), can be determined through studying the main growth term of $Q_{t,t'}$ in (3.3.2) if $F_L^{(-1)}$ is explicitly known.

Like Remark 3.1.4, one can show that at the time scale $\Theta(t^{2/(3-\tau)})$ the weighted distance between u_t and v_t tends to $2b$ where $b := \inf\{x \in \mathbb{R} : F_L(x) > 0\}$. At this time scale, many vertices connect to both u_t and v_t, allowing to bound the weighted distance from above by $2(b + \varepsilon)$ for arbitrarily small $\varepsilon > 0$.

3.4 PROOF OF THE LOWER BOUND

Now we prove the lower bound of Theorem 3.3.3, i.e., we show that with probability close to one there is no *too* short path between u_t and v_t for any $t' \geqslant t$. The main contribution of this section versus Chapter 2 and existing literature, e.g. [46, 69], is the following proposition concerning the graph distance. In its proof we develop a path-decomposition technique that uses the dynamical construction of PA_t in a refined way to get strong error bounds that are summable over $t' \geqslant t$. After the notational and conceptual set-up of the argument, we state and prove some technical lemmas. In the end of the section, we extend Proposition 3.4.1 to the edge-weighted setting, using refinements of the error bounds in Chapter 2. We abbreviate $u := u_t$ and $v := v_t$, respectively.

Proposition 3.4.1 (Lower bound graph distance)**.** *Consider the preferential attachment model with power-law exponent $\tau \in (2, 3)$. Let u, v be two typical vertices in PA_t. Then for any $\delta > 0$, there exists $M_G > 0$ such that*

$$\mathbb{P}\big(\exists t' \geqslant t : d^{(G)}_{t'}(u, v) \leqslant 2K_{t,t'} - 2M_G\big) \leqslant \delta. \tag{3.4.1}$$

Observe that t' is inside the $\mathbb{P}$-sign. Hence, (3.4.1) tracks the evolution of $d^{(G)}_{t'}(u, v)$ as time passes, and the graph around u and v grows. To estimate the probability of a too short path, we use a truncated path-counting method similar to [69]. This method first excludes possible paths that are unlikely to be present, called *bad paths*. Then, among the rest, the *good paths*, it counts the expected number of paths that are too short and present in $\mathrm{PA}_{t'}$. More precisely, the expected number of paths between u and v of length at most

$$2\underline{K}_{t,t'} := 2K_{t,t'} - 2M_G \tag{3.4.2}$$

is shown to be much smaller than one. We do this decomposition in an interlinked way that ensures that paths are only counted once.

3.4.1 *Set-up for the graph-distance evolution*

Recall that the arrival time of a vertex is also called birth time. The decomposition of good and bad paths is based on an array of birth times $(\ell^{(t)}_{k,t'})$ for which we make the following assumption throughout this section.

Assumption 3.4.2. *The array of birth times $(\ell^{(t)}_{k,t'})_{k \geqslant 0, t' \geqslant t}$ is a positive integer-valued array that is nonincreasing in both parameters and satisfies $\ell^{(t)}_{0,t'} \leqslant t$. We call it the birth-threshold array.*

Recall the definition of paths in Definition 3.3.1.

Definition 3.4.3. Let $(\ell^{(t)}_{k,t'})_{k\geqslant 0,t'\geqslant t}$ be an array satisfying Assumption 3.4.2. A t'-possible q-path $(\pi_0,\dots,\pi_k)$ is called t'-good if $t' \geqslant \pi_j \geqslant \ell^{(t)}_{j,t'}$ for all $j \leqslant k$, otherwise it is called t'-bad. A t'-possible (u,v)-path $(\pi_0,\dots,\pi_n)$ is called (n,t')-good if $\pi_j \wedge \pi_{n-j} \geqslant \ell^{(t)}_{j,t'}$ for all $j \leqslant \lfloor n/2 \rfloor$, otherwise it is called t'-bad.

This definition calls any path bad if it has a too *old* vertex, where the threshold $\ell^{(t)}_{k,t'}$ depends on the distance from π_0. Thus, all vertices on a good path are sufficiently *young*. We decompose t'-bad paths according to their first vertex violating the threshold.

Definition 3.4.4. Let $(\ell^{(t)}_{k,t'})_{k\geqslant 0,t'\geqslant t}$ be an array satisfying Assumption 3.4.2. We say that a π_0-path $(\pi_0,\dots,\pi_n)$ of length n is (k,t')-bad if the path is t'-possible and $\pi_j \geqslant \ell^{(t)}_{j,t'}$ for all $j<k$, but $\pi_k < \ell^{(t)}_{k,t'}$.

Observation 3.4.5. *Let $(\ell^{(t)}_{k,t'})_{k\geqslant 0,t'\geqslant t}$ be an array satisfying Assumption 3.4.2. Then*

1. *if a path is t'-good, then it is $\tilde{t}$-good for all $\tilde{t} \geqslant t'$.*
2. *if a path is t'-bad, it is possible that it turns $\tilde{t}$-good for some $\tilde{t} > t'$.*
3. *if a path $(\pi_0,\dots,\pi_n)$ is t'-bad, then it is $\tilde{t}$-bad for any $\tilde{t} \in [\max\{\pi_i\}, t']$.*
4. *if for all $i \leqslant k$ no $(i, t'-1)$-bad path is present in $\mathrm{PA}_{t'-1}$, then a (k,t')-bad path can only be present in $\mathrm{PA}_{t'}$ if it passes through vertex t'.*

All four observations follow directly from the definitions of good and bad paths, and the fact that $(\ell^{(t)}_{k,t'})_{k\geqslant 0,t'\geqslant t}$ is decreasing in both parameters, see Figure 4(A-B). The fourth observation turns out to be crucial in our decomposition argument. We define the events whose union implies the event between brackets in (3.4.1). We start with the event of having a bad path emanating from $q \in \{u,v\}$ for $k \geqslant 1$, i.e.,

$$\mathcal{E}^{(q)}_{\text{bad}}(k,t) := \{\exists (k,t)\text{-bad } q\text{-path}\}, \tag{3.4.3}$$

and for $t' \geqslant t$

$$\mathcal{E}^{(q)}_{\text{bad}}(k,t') := \{\exists (k,t')\text{-bad } q\text{-path}, \forall_{i\leqslant k} : \nexists (i,t'-1)\text{-bad } q\text{-path}\} \tag{3.4.4}$$

Here the sign $\exists$ indicates that a path is present. For completeness, we define for $t' \geqslant t$, $k=0$,

$$\mathcal{E}^{(q)}_{\text{bad}}(0,t') := \{q < \ell^{(t)}_{0,t'}\} \subseteq \mathcal{E}^{(q)}_{\text{bad}}(0,t), \tag{3.4.5}$$

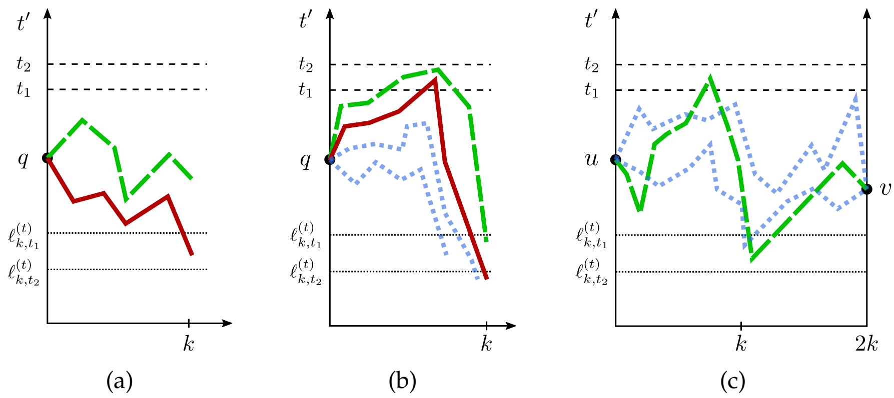


Figure 4: Good and bad path decomposition for the lower bound. Bad paths are displayed in red, the green dashed lines are the good paths, and the blue dotted lines represent possible paths that are absent. The y-axis represents the birth time of the vertices and the x-axis the graph distance from q and u, respectively. In Figure (A) we see that if a path is t_1-good, then it is also t_2-good since $t' \mapsto \ell^{(t)}_{k,t'}$ is decreasing. However, the red (k, t_1)-bad path turns (k, t_2)-good. Figure (B) shows that if there is no (k, t_1)-bad path, a red (k, t_2)-bad path must pass through a vertex in $(t_1, t_2]$. Note that the green path in Figure (B) is t_2-good. Although it violates a birth threshold valid at time time t_1, the path is not t_1-bad because it is not t_1-possible. Figure (C) shows that if there is neither a good, nor a bad t_1-present (u, v)-path of length $2k$, then a t_2-present (good) (u, v)-path must pass through a vertex in $(t_1, t_2]$. We apply the observations in these figures for $t_2 = t_1 + 1$.

where the inclusion follows since $t' \mapsto \ell^{(t)}_{0,t'}$ is nonincreasing. By the additional restriction on bad paths in (3.4.4), the events $\mathcal{E}^{(q)}_{\mathrm{bad}}(k, t')$ are disjoint in both parameters. For $\mathcal{E}^{(q)}_{\mathrm{bad}}(k, t')$ and $t' > t$, as a result of Observation 3.4.5(4) and the restriction in the definition (3.4.4) of not having a bad path at time $t' - 1$, we only have to consider paths that pass through the vertex t'. This motivates to decompose the (k, t')-bad paths passing through vertex t' according to the number of edges between the initial vertex $q \in \{u, v\}$ and t'. Indeed, consider a (k, t')-bad q-path where t' is the i-th vertex, i.e., it is of the form $(q, \pi_1, \dots, \pi_{i-1}, t', \pi_{i+1}, \dots, \pi_k)$. Then, by Definition 3.4.4, the constraints that this path satisfies is that for $j < k$, $\pi_j \geqslant \ell^{(t)}_{j,t'}$. This means that on the segment $(\pi_{i+1}, \dots, \pi_k) =: (\sigma_1, \dots, \sigma_{k-i})$ the indices of the constraints have to be shifted by i, giving rise to $\sigma_j \geqslant \ell^{(t)}_{i+j,t'}$ for $j \leqslant k - i$. Hence, we introduce *good paths on a* *segment*. Recall that $\{\pi_0 \leftrightarrow \cdots \leftrightarrow \pi_n\}$ means that $(\pi_0, \dots, \pi_n)$ is t'-present for $t' = \max_{i \leqslant n} \pi_i$.

Definition 3.4.6. Given an array $(\ell^{(t)}_{k,t'})_{k \geqslant 0, t' \geqslant t}$ satisfying Assumption 3.4.2, let

$$\{x \overset{[i,n)}{\rightsquigarrow} y\}_{t'} := \left\{ \text{disjoint}(\pi_i, \dots, \pi_{n-1}) \;\middle|\; \begin{array}{l} x = \pi_i \leftrightarrow \cdots \leftrightarrow \pi_{n-1} \leftrightarrow y \\ \forall_{j<n} : \pi_j \in [\ell^{(t)}_{j,t'}, t'] \end{array} \right\}.$$

If $\{x \overset{[i,n)}{\rightsquigarrow} y\}_{t'} \neq \varnothing$, we say that there is a t'-good x-path on segment $[i, n)$. We write

$$\{x \overset{[i_1,n_1)}{\rightsquigarrow} y\}_{t'} \circ \{y \overset{[i_2,n_2)}{\rightsquigarrow} z\}_{t'}$$

for the set of self-avoiding (x, z)-paths that are t'-good on the segment $[i_1, n_1)$ from x to y and t'-good on the segment $[i_2, n_2)$ from y to z.

Note that there is no birth restriction on the last vertex on the segment, explaining the half-open interval superscript $[i, n)$. Thus, if $\pi_k < \ell^{(t)}_{k,t'}$, then

$$\left|\{q \overset{[0,i)}{\rightsquigarrow} t'\}_{t'} \circ \{t' \overset{[i,k)}{\rightsquigarrow} \pi_k\}_{t'}\right| \geqslant 1$$

precisely means that there is a (k, t')-bad q-path from q to π_k that has t' as its i-th vertex. For notational convenience we omit the subscript t'.

Having set up the definitions for the bad paths, we define the events that allow to count the expected number of too short good (u, v)-paths. Let, for $n \geqslant 1$, and $t' = t$

$$\mathcal{E}^{(u,v)}_{\text{short}}(n, t') := \{\exists (n, t)\text{-good } (u, v)\text{-path}\}, \tag{3.4.6}$$

and for $n \geqslant 1$, $t' > t$

$$\mathcal{E}^{(u,v)}_{\text{short}}(n, t') := \left\{ \begin{array}{l} \exists (n, t')\text{-good } (u, v)\text{-path}, \\ \forall_{\tilde{t} < t'} : \nexists (u, v)\text{-path of length } n \end{array} \right\}, \tag{3.4.7}$$

setting for completeness

$$\mathcal{E}^{(u,v)}_{\text{short}}(0, t') := \{u = v\}.$$

Observe that in (3.4.7) we require that at previous times there was neither a good, nor a bad path of length n between u and v. This is a stronger requirement than the one in (3.4.4), where we do not put any restrictions on good paths at a previous time, but there only one endpoint of the path

(u or v) is fixed. By definition, for a fixed n, the events $\mathcal{E}^{(u,v)}_{\text{short}}(n, t')$ are disjoint. Moreover, we observe that if $\mathcal{E}^{(u,v)}_{\text{short}}(n, t')$ holds, then there is a t'-present (u, v)-path of length n connecting u and v that traverses the vertex t', which is a similar observation to Observation 3.4.5(4), see Figure 4(C). Using the definitions of the events $\mathcal{E}_{\text{bad}}$ and $\mathcal{E}_{\text{short}}$, we can bound the event between brackets in (3.4.1), and hence its probability of occurring, as stated in the following lemma.

Lemma 3.4.7. *Let* $(\ell^{(t)}_{k,t'})_{k\geqslant 0, t'\geqslant t}$ *be an array satisfying Assumption 3.4.2. Then*

$$\begin{aligned}\mathbb{P}\big(\exists t' \geqslant t : d^{(G)}_{t'}(u, v) \leqslant 2\underline{K}_{t,t'}\big)\\ \leqslant \sum_{q\in\{u,v\}} \sum_{t'=t}^{\infty} \sum_{k=0}^{\underline{K}_{t,t'}} \mathbb{1}_{\{k\geqslant 2 \text{ or } t'=t\}} \mathbb{P}\big(\mathcal{E}^{(q)}_{\text{bad}}(k, t')\big) \qquad (3.4.8)\\ + \sum_{t'=t}^{\infty} \sum_{n=0}^{2\underline{K}_{t,t'}} \mathbb{1}_{\{n\geqslant 2 \text{ or } t'=t\}} \mathbb{P}\big(\mathcal{E}^{(u,v)}_{\text{short}}(n, t')\big). \qquad (3.4.9)\end{aligned}$$

Proof. To prove the assertions in the statement, we will first bound the event between brackets on the left-hand side (lhs) in (3.4.8). Eventually, the bound then follows by a union bound.

Bounding the events. We write $\sigma(u, v) := \{(u, v), (v, u)\}$. We aim to show that if $(\ell^{(t)}_{k,t'})_{k\geqslant 0, t'\geqslant t}$ is an array satisfying Assumption 3.4.2, then

$$\big\{\exists t' \geqslant t : d^{(L)}_{t'}(u, v) \leqslant 2\underline{K}_{t,t'}\big\} \subseteq \bigcup_{t'=t}^{\infty} \left(\left(\bigcup_{k=0}^{\underline{K}_{t,t'}} \bigcup_{q\in\{u,v\}} \mathcal{E}^{(q)}_{\text{bad}}(k, t')\right) \cup \bigcup_{n=0}^{2\underline{K}_{t,t'}} \mathcal{E}^{(u,v)}_{\text{short}}(n, t')\right). \qquad (3.4.10)$$

Moreover, for $k, n \geqslant 2$ and $t' > t$

$$\mathcal{E}^{(u)}_{\text{bad}}(k, t') \subseteq \bigcup_{x=1}^{\ell^{(t)}_{k,t'}-1} \bigcup_{i=1}^{k-1} \Big\{\big|\{u \overset{[0,i)}{\rightsquigarrow} t'\} \circ \{t' \overset{[i,k)}{\rightsquigarrow} x\}\big| \geqslant 1\Big\}, \qquad (3.4.11)$$

$$\begin{aligned}\mathcal{E}^{(u,v)}_{\text{short}}(n, t') \subseteq \bigcup_{\substack{(q_1,q_2)\\ \in\sigma(u,v)}} \bigcup_{x=\ell^{(t)}_{\lfloor n/2\rfloor,t'}}^{t'-1} \bigcup_{i=1}^{\lfloor n/2\rfloor-1} \qquad (3.4.12)\\ \Big\{\big|\{q_1 \overset{[0,i)}{\rightsquigarrow} t'\} \circ \{t' \overset{[i,\lfloor n/2\rfloor)}{\rightsquigarrow} x\} \circ \{q_2 \overset{[0,\lceil n/2\rceil)}{\rightsquigarrow} x\}\big| \geqslant 1\Big\}\\ \cup \Big\{\big|\{q_1 \overset{[0,\lfloor n/2\rfloor)}{\rightsquigarrow} t'\} \circ \{q_2 \overset{[0,\lceil n/2\rceil)}{\rightsquigarrow} t'\}\big| \geqslant 1\Big\}.\end{aligned}$$

We first prove (3.4.12). Let $\boldsymbol{\pi}$ be any path of length $n \geqslant 2$ whose presence implies $\mathcal{E}_{\text{short}}^{(u,v)}(n, t')$ for some $t' > t$, so that $\boldsymbol{\pi}$ is a t'-good (u, v)-path by the definition of $\mathcal{E}_{\text{short}}^{(u,v)}(n, t')$ in (3.4.7). From (3.4.7) it also follows that t' is on $\boldsymbol{\pi}$, as there was neither a good, nor a bad (u, v)-path of length n before time t'. Thus, the t'-good (u, v)-path $\boldsymbol{\pi}$ can be decomposed in a t'-good u-path of length $\lfloor n/2 \rfloor$ and a t'-good v-path of length $\lceil n/2 \rceil$. Considering all possible positions of t' on the path, the presence of $\boldsymbol{\pi}$ implies the event on the rhs in (3.4.12). There, we denoted by $x \neq t'$ the vertex at distance $\lfloor n/2 \rfloor$ from q_1 that satisfies the constraint $x \geqslant \ell_{\lfloor n/2 \rfloor, t'}^{(t)}$. Thus, x is at distance $\lceil n/2 \rceil$ from q_1, and since $j \mapsto \ell_{j,t'}^{(t)}$ is nonincreasing also $x \geqslant \ell_{\lceil n/2 \rceil, t'}^{(t)}$. So the inclusion in (3.4.12) holds, since $\boldsymbol{\pi}$ was an arbitrary path.

Similarly, let $\boldsymbol{\pi} = (q, \dots, \pi_k)$ be any path of length $k \geqslant 2$ whose presence implies $\mathcal{E}_{\text{bad}}^{(q)}(k, t')$ for some $t' > t, q \in \{u, v\}$, so that $\pi_k < \ell_{k,t'}^{(t)}$. By Observation 3.4.5(4), vertex t' must be on $\boldsymbol{\pi}$ and by a similar reasoning as before we obtain (3.4.11).

Lastly, we prove (3.4.10) for which we rewrite the lhs as a union over time and paths, i.e.,

$$\left\{\exists t' \geqslant t : d_{t'}^{(L)}(u, v) \leqslant 2\underline{K}_{t,t'}\right\} = \bigcup_{t'=t}^{\infty} \bigcup_{n=0}^{2\underline{K}_{t,t'}} \bigcup_{\substack{(\pi_1, \dots, \pi_{n-1}) \\ \in [t']^{n-1}, \\ \text{disjoint}}} \{u \leftrightarrow \pi_1 \leftrightarrow \cdots \leftrightarrow \pi_{n-1} \leftrightarrow v\}.$$

Let $\boldsymbol{\pi} := (\pi_0, \dots, \pi_n)$ be any self-avoiding path from $\pi_0 := u$ to $\pi_n := v$ in this set. The smallest time t' at which $\boldsymbol{\pi}$ can be present in the union on the rhs is at $t' := t \vee \max_{i \leqslant n} \pi_i$. Then, $n \leqslant 2\underline{K}_{t,t'}$ must hold due to the fact that $t' \mapsto \underline{K}_{t,t'}$ is nonincreasing. We will show now that the event that $\boldsymbol{\pi}$ is t'-present is captured in either $\mathcal{E}_{\text{short}}^{(u,v)}(n, t')$ or $\mathcal{E}_{\text{bad}}^{(q)}(k, \tilde{t})$ for some $\tilde{t} \leqslant t', k \leqslant n/2$, $q \in \{u, v\}$. For any length $n \geqslant 0$, if $u \wedge v < \ell_{0,t'}^{(t)}$, then

$$\{\boldsymbol{\pi} \text{ present}\} \subseteq \mathcal{E}_{\text{bad}}^{(u)}(0, t) \cup \mathcal{E}_{\text{bad}}^{(v)}(0, t),$$

since $t' \mapsto \ell_{0,t'}^{(t)}$ is nonincreasing. From now on we assume that $u \wedge v \geqslant \ell_{0,t'}^{(t)}$. If $n \leqslant 1$, that is when $\{u = v\}$ or $\{u \leftrightarrow v\}$, then $\boldsymbol{\pi}$ must already be present at time t, i.e.,

$$\{\boldsymbol{\pi} \text{ present}\} \subseteq \cup_{i \in \{0,1\}} \mathcal{E}_{\text{short}}^{(u,v)}(i, t).$$

From now on we assume that the length $n \geqslant 2$. Moreover, if $\boldsymbol{\pi}$ is a t'-good path, then

$$\{\boldsymbol{\pi} \text{ present}\} \subseteq \mathcal{E}_{\text{short}}(n, t').$$

Assume $\boldsymbol{\pi}$ is not a t'-good (u, v)-path. Consequently, there is a t'-bad path emanating from either u or v, which is a *subpath* of $\boldsymbol{\pi}$. So, recalling Observation 3.4.5(1) and (2), the first time that this bad subpath is present, i.e.,

$$\tilde{t} := \min\left\{\hat{t} \leqslant t' \;\middle|\; \begin{array}{l} \exists m \leqslant n/2 \text{ s.t. } (u, \pi_1, \ldots, \pi_{\lfloor n/2 \rfloor}) \text{ or} \\ (v, \pi_{n-1}, \ldots, \pi_{n-\lfloor n/2 \rfloor}) \text{ is } (\hat{t}, m)\text{-bad} \end{array}\right\}$$

is well-defined and at most t'. By Observation 3.4.5(3), $\boldsymbol{\pi}$ is bad at $\tilde{t}$, so that for some $m \leqslant n/2$

$$\{\boldsymbol{\pi} \text{ present}\} \subseteq \cup_{q \in \{u,v\}} \mathcal{E}_{\text{bad}}^{(q)}(m, \tilde{t}).$$

Union bound. Having bounded the events between brackets on the lhs in (3.4.8) and (3.4.9), the assertions follow directly from a union bound on the events in (3.4.10). We argue now that the events where one of the indicators in (3.4.8) and (3.4.9) equals zero, happen with probability zero. We start with (3.4.9): $\mathbb{1}_{\{n \geqslant 2 \text{ or } t'=t\}} = 0$ when both $t' > t$ and $n \in \{0, 1\}$. Since no new paths connecting u and v of length one, i.e., a single edge, can be created after time $u \vee v \leqslant t$ we have that for $t' > t$ and $n \in \{0, 1\}$

$$\mathcal{E}_{\text{short}}^{(u,v)}(n, t') = \varnothing,$$

as by its definition in (3.4.7) we require that there was no path of length n before time t'. Similarly, bad paths of length at most one must already be present at time t since $t' \mapsto \ell_{k,t'}^{(t)}$ is nonincreasing and starts at a value at most t. So for $q \in \{u, v\}$, $t' > t$, $k \in \{0, 1\}$

$$\mathcal{E}_{\text{bad}}^{(q)}(k, t') = \varnothing. \qquad \square$$

3.4.2 *Bounding the summands*

The main goal of this section is to prove the following lemma for two suitably chosen sequences $k \mapsto \alpha_{[0,k)}, \beta_{[0,k)}$, defined below in (3.4.22) and (3.4.23). We will obtain bounds on the individual summands in (3.4.8) and (3.4.9) in Lemma 3.4.7.

Lemma 3.4.8. *Let* $k \mapsto \alpha_{[0,k)}, \beta_{[0,k)}$*, as in* (3.4.22) *and* (3.4.23) *below, respectively. Then there exists* $C > 0$ *such that for* $k \geqslant 2, n \geqslant 2$ *and* $t' > t$, $q \in \{u, v\}$

$$\mathbb{P}\big(\mathcal{E}^{(q)}_{\text{bad}}(k,t')\big) \leqslant Ct'^{-1}(k-1)\alpha_{[0,k)} \sum_{x=1}^{\ell^{(t)}_{k,t'}-1} x^{-\gamma}, \tag{3.4.13}$$

$$\mathbb{P}\big(\mathcal{E}^{(u,v)}_{\text{short}}(n,t')\big) \leqslant 2\beta^2_{[0,\lceil n/2\rceil)} t'^{2\gamma-2} + Cnt'^{-1} \sum_{x=\ell_{\lfloor n/2\rfloor,t'}}^{t'-1} \big(\alpha_{[0,\lceil n/2\rceil)} x^{-\gamma} + \beta_{[0,\lceil n/2\rceil)} x^{\gamma-1}\big)^2. \tag{3.4.14}$$

For $t' = t$, $k, n \geqslant 1$ *and* $q \in \{u, v\}$*, it holds that*

$$\mathbb{P}\big(\mathcal{E}^{(q)}_{\text{bad}}(k,t)\big) \leqslant \alpha_{[0,k)} \sum_{x=1}^{\ell^{(t)}_{k,t}-1} x^{-\gamma}, \tag{3.4.15}$$

$$\mathbb{P}\big(\mathcal{E}^{(u,v)}_{\text{short}}(n,t)\big) \leqslant \sum_{x=\ell^{(t)}_{\lfloor n/2\rfloor,t}}^{t} \big(\alpha_{[0,\lceil n/2\rceil)} x^{-\gamma} + \beta_{[0,\lceil n/2\rceil)} x^{\gamma-1}\big)^2. \tag{3.4.16}$$

We prove the lemma at the end of this section after having established the necessary preliminaries and identified the sequences $k \mapsto \alpha_{[0,k)}, \beta_{[0,k)}$. The decomposition method counting paths that traverse the vertex t' (for $t' > t$) yields a bound in (3.4.13) and (3.4.14) that are a factor $1/t'$ smaller than their counterparts with $t' = t$ in (3.4.15) and (3.4.16). By small refinements of the methods in [69] we obtain that the individual sums on the rhs in (3.4.13) and (3.4.14) are of order $1/\log^3(t')$. This is why the error terms are summable in t'. The extra factor $1/t'$ illustrates the necessity of our decomposition method versus previous methods.

In view of (3.4.11) and (3.4.12), it is crucial to understand the probabilities on having self-avoiding paths that are restricted to have specified vertices at some positions. For this we rely on the function p from Proposition 2.4.4, that originally appeared in [69].
For $k > i \geqslant 0$ and a vertex $\pi_i \geqslant \ell^{(t)}_{i,t'}$ and another vertex $\pi_k \in [t']$ we define

$$f^{(t,t')}_{[i,k)}(\pi_i, \pi_k) := \sum_{\substack{(\pi_{i+1},\dots,\pi_{k-1}) \\ \in \mathcal{P}^{\{\pi_i,\pi_k\}}_{(i,k)}}} p(\pi_i, \dots, \pi_k), \tag{3.4.17}$$

where for a vertex set $\mathcal{V} \subset [t']$, $\mathcal{P}^{\mathcal{V}}_{(i,k)}$ denotes the set of pairwise disjoint vertex tuples $(\pi_{i+1}, \dots, \pi_{k-1})$ such that $\pi_j \geqslant \ell^{(t)}_{j,t'}$, $\pi_j \notin \mathcal{V}$ for all $i < j < k$.

Intuitively, $f^{(t,t')}_{[i,k)}(\pi_i, \pi_k)$ is an upper bound for the expected number of t'-good paths on the segment $[i,k)$ from π_i to π_k.

We derive upper bounds for the summands in (3.4.8) and (3.4.9) in terms of $f^{(t,t')}_{[i,k)}$.

Claim 3.4.9. *Consider the preferential attachment model with power-law parameter* $\tau > 2$*. Let* $(\ell^{(t)}_{k,t'})_{k\geqslant 0, t'\geqslant t}$ *be an array satisfying Assumption 3.4.2. Then for* $k, n \geqslant 2$ *and* $t' > t$*,*

$$\mathbb{P}\big(\mathcal{E}^{(q)}_{\mathrm{bad}}(k,t')\big) \leqslant \sum_{x=1}^{\ell^{(t)}_{k,t'}-1} \sum_{i=1}^{k-1} f^{(t,t')}_{[0,i)}(q,t') f^{(t,t')}_{[i,k)}(t',x), \tag{3.4.18}$$

$$\begin{aligned}\mathbb{P}\big(\mathcal{E}^{(u,v)}_{\mathrm{short}}(n,t')\big) \leqslant \sum_{(q_1,q_2)\in\sigma(u,v)} \sum_{x=\ell^{(t)}_{\lfloor n/2\rfloor,t'}}^{t'-1} \sum_{i=1}^{\lfloor n/2\rfloor -1} \Big(& \\ f^{(t,t')}_{[0,\lceil n/2\rceil)}(q_1,x)\cdot f^{(t,t')}_{[0,i)}(q_2,t')\cdot f^{(t,t')}_{[i,k)}(t',x) & \\ + f^{(t,t')}_{[0,\lceil n/2\rceil)}(q_1,t')\cdot f^{(t,t')}_{[0,\lfloor n/2\rfloor)}(q_2,t')\Big), &\end{aligned} \tag{3.4.19}$$

while for any $k, n \geqslant 1$ *and* $t' = t$

$$\mathbb{P}\big(\mathcal{E}^{(q)}_{\mathrm{bad}}(k,t)\big) \leqslant \sum_{x=1}^{\ell^{(t)}_{k,t}-1} f^{(t,t)}_{[0,k)}(q,x), \tag{3.4.20}$$

$$\begin{aligned}\mathbb{P}\big(\mathcal{E}^{(u,v)}_{\mathrm{short}}(n,t)\big) \leqslant \sum_{(q_1,q_2)\in\sigma(u,v)} \sum_{x=\ell^{(t)}_{\lfloor n/2\rfloor,t'}}^{t} \Big(& \\ f^{(t,t)}_{[0,\lceil n/2\rceil)}(q_1,x) f^{(t,t)}_{[0,\lfloor n/2\rfloor)}(q_2,x)\Big). &\end{aligned} \tag{3.4.21}$$

Proof. Recall the set of paths $\{\pi_i \overset{[i,k)}{\rightsquigarrow} \pi_k\}$ from Definition 3.4.6. Then by Markov's inequality, (3.4.17), and Proposition 2.4.4

$$\begin{aligned}\mathbb{P}\Big(\big|\{\pi_i \overset{[i,k)}{\rightsquigarrow} \pi_k\}\big| \geqslant 1\Big) \leqslant \mathbb{E}[|\{x \overset{[i,k)}{\rightsquigarrow} \pi_k\}|] &= \sum_{\substack{(\pi_{i+1},\dots,\pi_{k-1})\\ \in \mathcal{P}^{\{\pi_i,\pi_k\}}_{(i,k)}}} \mathbb{P}(\pi_i \leftrightarrow \cdots \leftrightarrow \pi_k) \\ &\leqslant \sum_{\substack{(\pi_{i+1},\dots,\pi_{k-1})\\ \in \mathcal{P}^{\{\pi_i,\pi_k\}}_{(i,k)}}} p(\pi_i,\dots,\pi_k) \\ &= f^{(t,t')}_{[i,k)}(\pi_i,\pi_k).\end{aligned}$$

Now for concatenated paths, due to the product structure in (2.4.5), and by relaxing the disjointness of sets, we have

$$\begin{aligned}\mathbb{P}\big(\big|\{\pi_0 \overset{[0,i)}{\rightsquigarrow} \pi_i\}\circ\{\pi_i \overset{[i,k)}{\rightsquigarrow} \pi_k\}\big| \geqslant 1\big) &\leqslant \sum_{\substack{(\pi_1,\dots,\pi_{i-1})\\ \in \mathcal{P}^{\{\pi_0,\pi_i,\pi_k\}}_{(0,i)}}} \sum_{\substack{(\pi_{i+1},\dots,\pi_{k-1})\\ \in \mathcal{P}^{\{\pi_0,\dots,\pi_i\}}_{(i,k)}}} p(\pi_0,\dots,\pi_i)p(\pi_i,\dots,\pi_k)\\ &\leqslant f^{(t,t')}_{[0,i)}(\pi_0,\pi_i) f^{(t,t')}_{[i,k)}(\pi_i,\pi_k).\end{aligned}$$

Recall now (3.4.11), so that (3.4.18) follows by a union bound and choosing $\pi_0 = q, \pi_i = t'$, and $\pi_k = x$. Similarly (3.4.19) follows by union bounds over the rhs in (3.4.12). The bounds (3.4.20) and (3.4.21) follow analogously from their definition in (3.4.3) and (3.4.6). □

We establish recursive bounds on $f^{(t,t')}_{[i,k)}$ similar to [69, Lemma 1]. Let $(\ell^{(t)}_{k,t'})_{k\geqslant 0, t'\geqslant t}$ be an array satisfying Assumption 3.4.2 such that $\eta_{j,t'} := (t'/\ell^{(t)}_{j,t'}) \geqslant e$ for all $j \geqslant 0$ and $t' \geqslant t$. Define for $\gamma := 1/(\tau-1)$ and some $c > 1$

$$\alpha^{(t')}_{[0,j)} := \begin{cases} \nu \ell^{\gamma-1}_{0,t'} & j=1,\\ c\big(\alpha^{(t')}_{[0,j-1)}\log(\eta_{j-1,t'}) + \beta^{(t')}_{[0,j-1)} t'^{2\gamma-1}\big) & j>1,\end{cases} \tag{3.4.22}$$

$$\beta^{(t')}_{[0,j)} := \begin{cases} \nu \ell^{-\gamma}_{0,t'} & j=1,\\ c\big(\alpha^{(t')}_{[0,j-1)}\ell^{1-2\gamma}_{j-1,t'} + \beta^{(t')}_{[0,j-1)}\log(\eta_{j-1,t'})\big) & j>1,\end{cases} \tag{3.4.23}$$

similar to the recursions in [69, Lemma 1]. The sequence $\big(\alpha^{(t')}_{[0,j)}\big)_{j\geqslant 1}$ closely is related to the expected number of self-avoiding t'-good paths $(\pi_0,\dots,\pi_j)$ of length j from $\pi_0 \in \{u,v\}$ to π_j such that $\pi_{j-1} > \pi_j$. The sequence $\big(\beta^{(t')}_{[0,j)}\big)_{j\geqslant 1}$ is related to those paths where $\pi_{j-1} < \pi_j$. Observe that since $c > 1$, $\eta_{j,t'} \geqslant e$, and $\alpha^{(t)}_{[0,1)}, \beta^{(t)}_{[0,1)} \geqslant 0$, it follows that $k \mapsto \alpha^{(t)}_{[0,k)}$ and $k \mapsto \beta^{(t)}_{[0,k)}$ are non-decreasing. We define for the same constant $c > 1$ the non-decreasing sequences

$$\phi^{(t')}_{[i,i+j)} := \begin{cases} \nu t'^{\gamma-1} & j=1,\\ c\big(\phi^{(t')}_{[i,i+j-1)}\log(\eta_{i+j-1,t'}) + \psi^{(t')}_{[i,i+j-1)} t'^{2\gamma-1}\big) & j>1,\end{cases} \tag{3.4.24}$$

$$\psi^{(t')}_{[i,i+j)} := \begin{cases} 0 & j=1,\\ c\big(\phi^{(t')}_{[i,i+j-1)}\ell^{1-2\gamma}_{i+j-1,t'} + \psi^{(t')}_{[i,i+j-1)}\log(\eta_{i+j-1,t'})\big) & j>1.\end{cases} \tag{3.4.25}$$

These sequences are related to the t'-good paths emanating from t' that are good on the segment $[i, i+j)$. Observe that the recursions are identical to (3.4.22) and (3.4.23), except that their initial values are different. This is crucial to give summable error bounds in t' later on. Below, we leave out the superscript (t') for notational convenience, but we stress here that these four sequences are dependent on both t' and t.

Claim 3.4.10 (Recursive bounds for number of paths)**.** *Under the same assumptions as Proposition 3.4.1, let* $(\ell^{(t)}_{k,t'})_{k\geqslant 0, t'\geqslant t}$ *be an array satisfying Assumption 3.4.2. Let* $\eta_{j,t'} = t'/\ell^{(t)}_{j,t'}$ *and* $\gamma = 1/(\tau-1)$*. For sufficiently large* $c = c(\tau), \nu = \nu(\tau)$ *in* (3.4.22)*,* (3.4.23)*,* (3.4.24)*, and* (3.4.25)*, it holds that*

$$f^{(t,t')}_{[i,i+j)}(t', x) \leqslant x^{-\gamma}\phi_{[i,i+j)} + \mathbb{1}_{\left\{x>\ell^{(t)}_{i+j-1,t'}\right\}} x^{\gamma-1}\psi_{[i,i+j)}. \tag{3.4.26}$$

Moreover,

$$f^{(t,t')}_{[0,j)}(q, x) \leqslant \mathbb{1}_{\{x<t'\}} x^{-\gamma}\alpha_{[0,j)} + \mathbb{1}_{\left\{x>\ell^{(t)}_{j-1,t'}\right\}} x^{\gamma-1}\beta_{[0,j)}. \tag{3.4.27}$$

We postpone the proof to Section 3.A.1, which follows by induction from arguments analogous to [69, Lemma 1]. As a consequence of (3.4.27), we have for $q \in \{u, v\}$

$$f^{(t,t')}_{[0,i)}(q, t') \leqslant \mathbb{1}_{\{t'<t'\}} t'^{-\gamma}\alpha_{[0,j)} + \mathbb{1}_{\left\{t'>\ell^{(t)}_{j-1,t'}\right\}} t'^{\gamma-1}\beta_{[0,j)} = t'^{\gamma-1}\beta_{[0,j)}.$$

Moreover, since $x < \ell^{(t)}_{k,t'}$ implies that also $x < \ell^{(t)}_{k-1,t'}$ since $k \mapsto \ell^{(t)}_{k,t'}$ is nonincreasing, for $x < \ell^{(t)}_{k,t'}$ it follows from (3.4.26) that

$$f^{(t)}_{[i,k)}(t', x) \leqslant x^{-\gamma}\phi_{[i,i+j)} + \mathbb{1}_{\left\{x>\ell^{(t)}_{k-1,t'}\right\}}\psi_{[i,k)} = x^{-\gamma}\phi_{[i,i+j)}.$$

Hence, we can bound the summands in (3.4.8) using Claim 3.4.10 to obtain for $k \geqslant 2$, $t' > t$,

$$\begin{aligned} \mathbb{P}(\mathcal{E}^{(q)}_{\text{bad}}(k, t')) &\leqslant \sum_{x=1}^{\ell^{(t)}_{k,t'}-1} \sum_{i=1}^{k-1} f^{(t,t')}_{[0,i)}(q, t') f^{(t,t')}_{[i,k)}(t', x) \\ &\leqslant t'^{\gamma-1} \sum_{x=1}^{\ell^{(t)}_{k,t'}-1} x^{-\gamma} \sum_{i=1}^{k-1} \beta_{[0,i)}\phi_{[i,k)}. \end{aligned} \tag{3.4.28}$$

Similarly to (3.4.28) we bound the summands in (3.4.9) from above using (3.4.19) and *replacing* the first sum over the permutation $\sigma(u, v)$ in (3.4.19) by a factor two, i.e., for $n \geqslant 2$ and $t' > t$,

$$\mathbb{P}\big(\mathcal{E}_{\text{short}}^{(u,v)}(n, t')\big) \leqslant 2t'^{\gamma-1} \sum_{x=\ell_{\lfloor n/2 \rfloor, t'}}^{t'} \big(\mathbb{1}_{\{x<t'\}}\alpha_{[0,\lceil n/2 \rceil)}x^{-\gamma} + \beta_{[0,\lceil n/2 \rceil)}x^{\gamma-1}\big)$$
$$\cdot \Big(\mathbb{1}_{\{x<t'\}} \sum_{i=1}^{\lfloor n/2 \rfloor -1} \big(\beta_{[0,i)}\phi_{[i,\lfloor n/2 \rfloor)}x^{-\gamma} + \beta_{[0,i)}\psi_{[i,\lfloor n/2 \rfloor)}x^{\gamma-1}\big)$$
$$+ \mathbb{1}_{\{x=t'\}}\beta_{[0,\lfloor n/2 \rfloor)}\Big). \tag{3.4.29}$$

Both (3.4.28) and (3.4.29) contain convolutions of the sequence $\beta_{[0,i)}$ with $\phi_{[i,k)}$ and $\psi_{[i,k)}$. This motivates to bound these convolutions in terms of the *original* sequences $\alpha_{[0,k)}$ and $\beta_{[0,k)}$.

Claim 3.4.11. *Let* $\phi_{[i,k)}, \psi_{[i,k)}, \alpha_{[0,k)}, \beta_{[0,k)}$ *be as in* (3.4.24), (3.4.25), (3.4.22), (3.4.23), *respectively. Then there exists* $C > 0$ *such that for* $k \geqslant 2$

$$B_k^{\psi} := \sum_{i=1}^{k-1} \beta_{[0,i)}\psi_{[i,k)} \leqslant C\,(k-2)\,\beta_{[0,k)}t'^{-\gamma}, \tag{3.4.30}$$

$$B_k^{\phi} := \sum_{i=1}^{k-1} \beta_{[0,i)}\phi_{[i,k)} \leqslant C(k-1)\alpha_{[0,k)}t'^{-\gamma}. \tag{3.4.31}$$

Proof. We prove by induction. We initialize the induction for $k = 2$, the smallest value of k for which the sums in (3.4.31) and (3.4.30) are non-empty. Indeed, then (3.4.30) holds by the initial value of $\psi_{[i,i+1)} = 0$ in (3.4.25), i.e.,

$$B_2^{\psi} = \beta_{[0,1)}\psi_{[1,2)} = \beta_{[0,1)} \cdot 0 \leqslant C \cdot 0 \cdot \beta_{[0,2)}t'^{-\gamma}.$$

For $k = 2$ in (3.4.31) we substitute the recursion (3.4.22) on $\alpha_{[0,2)}$. Thus, we have to show that

$$B_2^{\phi} = \beta_{[0,1)}\phi_{[1,2)} \leqslant cC\big(\alpha_{[0,1)}\log(\eta_{1,t'}) + \beta_{[0,1)}t'^{2\gamma-1}\big)t'^{-\gamma}.$$

Using the initial values in (3.4.22), (3.4.23), and (3.4.24), this is indeed true for $C \geqslant \nu/c$, i.e.,

$$B_2^{\phi} = \nu\ell_{0,t'}^{-\gamma} \cdot \nu t'^{\gamma-1} \leqslant cC\big(\nu\ell_{0,t'}^{\gamma-1}\log(t'/\ell_{1,t'})t'^{-\gamma} + \nu\ell_{0,t'}^{-\gamma}t'^{\gamma-1}\big).$$

Now, we advance the induction. To this end, one can derive the following recursions using (3.4.24) and (3.4.25):

$$B^{\psi}_{k+1} = \begin{cases} cB^{\phi}_k \ell_{k,t'}^{1-2\gamma} + cB^{\psi}_k \log(\eta_{k,t'}), & \text{if } k > 1 \\ 0, & \text{if } k = 1, \end{cases} \tag{3.4.32}$$

$$B^{\phi}_{k+1} = \begin{cases} cB^{\phi}_k \ell_{k,t'}^{1-2\gamma} + cB^{\psi}_k \log(\eta_{k,t'}), & \text{if } k > 1 \\ \nu^2 t'^{\gamma-1}\ell_{0,t'}^{-\gamma}, & \text{if } k = 1. \end{cases} \tag{3.4.33}$$

The first term in (3.4.33) is a result of the *non-zero* initial value of $\phi_{[i,i+1)}$ in (3.4.24), while $\psi_{[i,i+1)} = 0$, so that there is no such term in (3.4.32). Since the two recursions depend only on each other's previous values, we can carry out the two induction steps simultaneously. By the two induction hypotheses (3.4.30) and (3.4.31), and the definition of $\beta_{[0,k+1)}$ in (3.4.23), we have that

$$\begin{aligned} B^{\psi}_{k+1} &= cB^{\phi}_k \ell_{k,t'}^{1-2\gamma} + cB^{\psi}_k \log(\eta_{k,t'}) \\ &\leqslant Cc\big((k-1)\alpha_{[0,k)}\ell_{k,t'}^{1-2\gamma} + (k-2)\beta_{[0,k)}\log(\eta_{k,t'})\big)t'^{-\gamma} \\ &\leqslant C(k-1)\beta_{[0,k+1)}t'^{-\gamma}, \end{aligned}$$

proving (3.4.32). For (3.4.33), we assume that $cC \geqslant \nu$ so that using the induction hypotheses and (3.4.22) the proof is finished, i.e.,

$$\begin{aligned} B^{\phi}_{k+1} &= c\log(\eta_{k,t'})B^{\phi}_k + cB^{\psi}_k t'^{2\gamma-1} + \nu t'^{\gamma-1}\beta_{[0,k)} \\ &\leqslant c\log(\eta_{k,t'})C(k-1)\alpha_{[0,k)}t'^{-\gamma} \\ &\quad + ct'^{\gamma-1}C\,(k-2)\,\beta_{[0,k)} + \nu t'^{\gamma-1}\beta_{[0,k)} \\ &\leqslant C(k-1)\alpha_{[0,k+1)}t'^{-\gamma}. \end{aligned}$$

□

We combine Claim 3.4.11 with (3.4.28) and (3.4.29) to arrive to the proof of Lemma 3.4.8.

Proof of Lemma 3.4.8. We start with (3.4.13). We recall from (3.4.28) the bound on $\mathbb{P}\big(\mathcal{E}^{(q)}_{\text{bad}}(k,t')\big)$, and observe that (3.4.31) implies (3.4.13), since there is $C > 0$ such that for $k \geqslant 2$

$$\begin{aligned} \mathbb{P}\big(\mathcal{E}^{(q)}_{\text{bad}}(k,t')\big) &\leqslant t'^{\gamma-1}\sum_{i=1}^{k-1}\beta_{[0,i)}\phi_{[i,k)}\sum_{x=1}^{\ell^{(t)}_{k,t'}-1}x^{-\gamma} \\ &\leqslant C(k-1)t'^{-1}\alpha_{[0,k)}\sum_{x=1}^{\ell^{(t)}_{k,t'}-1}x^{-\gamma}. \end{aligned}$$

For (3.4.14), we recall the bound (3.4.29) and bound using (3.4.31) and (3.4.30) the factor on the second line in (3.4.29) by

$$\mathbb{1}_{\{x=t'\}}\beta_{[0,\lfloor n/2\rfloor)} + \mathbb{1}_{\{x<t'\}}Ct'^{-\gamma}\big((\lfloor n/2\rfloor - 1)\alpha_{[0,\lfloor n/2\rfloor)}x^{-\gamma} \\ + (\lfloor n/2\rfloor - 2)\beta_{[0,\lfloor n/2\rfloor)}x^{\gamma-1}\big).$$

Now (3.4.14) follows by distinguishing the summands in (3.4.29) between $x < t'$ and $x = t'$, and using that $j \mapsto \alpha_{[0,j)}$ and $j \mapsto \beta_{[0,j)}$ are non-decreasing so that we may round up their indices to $\lceil n/2\rceil$ to obtain the square. Lastly, the bounds (3.4.15) and (3.4.16) follow directly from (3.4.20), (3.4.21), and (3.4.27), where we again round up the indices to obtain the square. □

3.4.3 *Setting the birth-threshold sequence*

After the event decomposition in Lemma 3.4.7 and the bounds on the individual summands in Lemma 3.4.8, we are ready to choose the birth-threshold array $(\ell_{k,t'}^{(t)})_{k\geqslant 0,t'\geqslant t}$ to ensure that the sums in (3.4.13), (3.4.14), (3.4.15), and (3.4.16) are sufficiently small. Our choice of $(\ell_{k,t'}^{(t)})_{k\geqslant 0,t'\geqslant t}$ will make the error probabilities in (3.4.8) and (3.4.9) arbitrarily small. Fix $\delta' = \delta'(\delta) > 0$ that we choose later to be sufficiently small. We define

$$\ell_{k,t'}^{(t)} := \begin{cases} \lceil \delta' t\rceil & k = 0, \quad (3.4.34a) \\ \underset{x\in\mathbb{N}\setminus\{0,1\}}{\arg\max}\left\{\alpha_{[0,k)}x^{1-\gamma} \overset{(\circledast)}{\leqslant} \big(k\log(t')\big)^{-3}\right\} & k\geqslant 1. \quad (3.4.34b)\end{cases}$$

Since $k \mapsto \alpha_{[0,k)}$ is non-decreasing and $1-\gamma > 0$ by (3.4.34b), $k \mapsto \ell_{k,t'}^{(t)}$ must be nonincreasing in both indices. Using the upper bound on $t'/\ell_{k,t'}^{(t)}$ in the postponed Lemma 3.A.1 in Section 3.A.1, one can verify for t sufficiently large that $\ell_{k,t'}^{(t)} \geqslant 2$ for all $k \leqslant \underline{K}_{t,t'}$. This makes the array $(\ell_{k,t'}^{(t)})_{k\geqslant 0,t'\geqslant t}$ well-defined. The choice of $(\ell_{k,t'}^{(t)})_{k\geqslant 0,t'\geqslant t}$ in (3.4.34) is similar to the choice in [69, Proof of Theorem 2] for $t' = t$. The main difference is the extra $\log^{-3}(t')$ factor on the rhs in (3.4.34b). This factor, in combination with the $1/t'$-factor from Lemma 3.4.8 yields a summable error in t' in (3.4.8) and (3.4.9). We comment that the additional $\log^{-3}(t')$ factor could be changed to another slowly varying function, but the choice has to be $o((t')^{\epsilon})$ for all $\epsilon > 0$, otherwise the entries of $(\ell_{k,t'}^{(t)})$ would not be at least two, whence the array would be ill-defined.

We are ready to prove Proposition 3.4.1.

Proof of Proposition 3.4.1. To prove (3.4.1), due to Corollary 3.4.7, we need to show that the rhs in (3.4.8) and (3.4.9) is at most δ, for t sufficiently large. To keep notation light, we write $\ell_{k,t'} := \ell^{(t)}_{k,t'}$. First, we consider the terms in (3.4.8) where $t' = t$. Recalling the definition of $\mathcal{E}^{(q)}_{\text{bad}}(0,t)$ from (3.4.5) and the upper bound on its probability in (3.4.15), we have for $q \in \{u, v\}$

$$\sum_{k=0}^{\underline{K}_{t,t}} \mathbb{P}\big(\mathcal{E}^{(q)}_{\text{bad}}(k,t)\big) \leqslant \delta' + O(1/t) + \sum_{k=1}^{\underline{K}_{t,t}} \alpha_{[0,k)} \sum_{x=1}^{\ell_{k,t}-1} x^{-\gamma}, \tag{3.4.35}$$

where the term $\delta' + O(1/t)$ comes from the probability that q, the uniform vertex in $[t]$, is born before $\ell_{0,t} = \lceil \delta' t \rceil$. Now approximating the last sum in (3.4.35) by an integral and using ($\circledast$) in (3.4.34b), we have for some $c_1 > 0$, $q \in \{u, v\}$

$$\begin{aligned} \sum_{k=0}^{\underline{K}_{t,t}} \mathbb{P}\big(\mathcal{E}^{(q)}_{\text{bad}}(k,t)\big) &\leqslant \delta' + o(1) + c_1 \sum_{k=1}^{\underline{K}_{t,t}} \alpha_{[0,k)} \ell_{k,t}^{1-\gamma} \\ &= \delta' + o(1) + c_1 \log^{-3}(t) \sum_{k=1}^{\underline{K}_{t,t}} k^{-3} = \delta' + o(1). \end{aligned} \tag{3.4.36}$$

We move on to the terms on the rhs in (3.4.8) for $t' > t$ and show that their sum is of order $O(\delta')$. Recall for $q \in \{u, v\}$ the bound on $\mathbb{P}\big(\mathcal{E}^{(q)}_{\text{bad}}(k,t')\big)$ in (3.4.13), and observe that there is $C' > 0$ such that, approximating the sum over x in (3.4.13) by an integral gives for $t' > t$, $k \geqslant 2$

$$\mathbb{P}\big(\mathcal{E}^{(q)}_{\text{bad}}(k,t')\big) \leqslant C' t'^{-1}(k-1)\alpha_{[0,k)} \ell_{k,t'}^{1-\gamma} \leqslant \frac{C'}{k^2 t' \log^3(t')}.$$

The last inequality follows from ($\circledast$) in (3.4.34b). The rhs is summable in k and t' so that, only considering the tail of the sum,

$$\sum_{t'=t+1}^{\infty} \sum_{k=2}^{\underline{K}_{t,t'}} \mathbb{P}\big(\mathcal{E}^{(q)}_{\text{bad}}(k,t')\big) = O\big(\log^{-2}(t)\big). \tag{3.4.37}$$

Combining (3.4.36) and (3.4.37), this establishes that the rhs in (3.4.8) is at most $2\delta' + o(1)$ for t sufficiently large, when summed over $q \in \{u, v\}$.

We continue by proving that the summed error probability in (3.4.9) is small. First we consider the terms where $t' > t$. Recall (3.4.14). We use now that $(a+b)^2 \leqslant 2(a^2+b^2)$ for $a, b > 0$, so that there exists $C' > 0$ such that

$$\mathbb{P}\big(\mathcal{E}^{(u,v)}_{\text{short}}(n,t')\big) \leqslant 2t'^{2\gamma-2}\beta^2_{[0,\lceil n/2\rceil)} + \frac{C'n}{t'}\beta^2_{[0,\lceil n/2\rceil)} \sum_{x=\ell_{\lfloor n/2\rfloor,t'}}^{t'-1} x^{2\gamma-2}$$
$$+ \frac{C'n}{t'}\alpha^2_{[0,\lceil n/2\rceil)} \sum_{x=\ell_{\lfloor n/2\rfloor,t'}}^{t'-1} x^{-2\gamma}$$
$$=: T_{11}(n,t') + T_{12}(n,t') + T_2(n,t'). \tag{3.4.38}$$

Approximating the sums by integrals and using that $k \mapsto \ell_{k,t'}$ is nonincreasing, there exists a different $C' > 2$ such that, relaxing the first two terms in (3.4.38),

$$T_{11}(n,t') + T_{12}(n,t') \leqslant 2C'n\beta^2_{[0,\lceil n/2\rceil)}t'^{2\gamma-2},$$
$$T_2(n,t') \leqslant C'n\alpha^2_{[0,\lceil n/2\rceil)}\ell^{1-2\gamma}_{\lceil n/2\rceil,t'}t'^{-1}.$$

For $T_1 := T_{11} + T_{12}$, by (3.4.22) we obtain

$$c\beta_{[0,\lceil n/2\rceil)} \leqslant \alpha_{[0,\lceil n/2\rceil+1)}t'^{1-2\gamma},$$

yielding by $(\circledast)$ in (3.4.34b)

$$\begin{aligned} T_1(n,t') &\leqslant 2C'n\alpha^2_{[0,\lceil n/2\rceil+1)}t'^{-2\gamma}/c^2 \\ &= 2C'n\Big(\alpha_{[0,\lceil n/2\rceil+1)}\ell^{1-\gamma}_{\lceil n/2\rceil+1,t'}\Big)^2\big(t'/\ell_{\lceil n/2\rceil+1,t'}\big)^{2-2\gamma}(ct')^{-2} \\ &\leqslant \frac{2C'n}{(\lceil n/2\rceil+1)^6\log^6(t')(ct')^2}\big(t'/\ell_{\lceil n/2\rceil+1,t'}\big)^{2-2\gamma}. \end{aligned} \tag{3.4.39}$$

Rewriting T_2 similarly,

$$\begin{aligned} T_2(n,t') &\leqslant C'n\Big(\alpha_{[0,\lceil n/2\rceil)}\ell^{1-\gamma}_{\lceil n/2\rceil,t'}\Big)^2\big(t'/\ell_{\lceil n/2\rceil,t'}\big)t'^{-2} \\ &\leqslant \frac{C'n}{\lceil n/2\rceil^6\log^6(t')t'^2}\big(t'/\ell_{\lceil n/2\rceil,t'}\big). \end{aligned} \tag{3.4.40}$$

Recall that $\ell_{\lceil n/2\rceil+1,t'} \geqslant 2$ as mentioned after (3.4.34b). Thus, both (3.4.39) and (3.4.40) are summable in t' and n. They tend to zero as t tends to infinity, using for (3.4.39) that $2-2\gamma < 1$.

We are left with verifying that the terms where $t' = t$ in (3.4.9) are of order $O(\delta')$ when summed over n. For this the same reasoning holds as above, starting after (3.4.37), where the initial bound is the one in (3.4.16) instead of (3.4.14). Here, all terms are a factor $t' = t$ larger than before. This yields that

$$\sum_{n=0}^{\underline{K}_{t,t}} \mathbb{P}\big(\mathcal{E}^{(u,v)}_{\mathrm{short}}(n,t)\big) = o(1) + \frac{C'}{\log^6(t)t}(t/\ell_{\underline{K}_{t,t}+1,t'}) \sum_{n=1}^{\underline{K}_{t,t}} \lceil n/2 \rceil^{-6}$$
$$= O(\log^{-6}(t)) = o(1).$$

Recalling the conclusions after (3.4.37) and (3.4.40), we conclude that the error terms in (3.4.8) and (3.4.9) are of order $O(\delta')$, so that (3.4.1) follows by Corollary 3.4.7, when δ' is chosen sufficiently small so that the error probabilities are at most δ, and $\eta_{0,t'} = t/\ell_{0,t} \geqslant e$ as required by the definition of $\alpha_{[0,j)}$ and $\beta_{[0,j)}$ before (3.4.22). □

3.4.4 *Extension to weighted distances*

We extend the result on graph distances from Proposition 3.4.1 to weighted distances, refining the results in Chapter 2. For this we recall $Q_{t,t'}$ from (3.3.2).

Proposition 3.4.12 (Lower bound weighted distance)**.** *Consider the preferential attachment model with power-law exponent* $\tau \in (2,3)$. *Equip every edge upon creation with an i.i.d. copy of the non-negative random variable* L. *Let* u, v *be two typical vertices in* PA_t. *Then for any* $\delta > 0$, *there exists* $M_L > 0$ *such that*

$$\mathbb{P}\big(\exists t' \geqslant t : d^{(L)}_{t'}(u,v) \leqslant 2Q_{t,t'} - 2M_L\big) \leqslant \delta.$$

Proof. Fix δ' sufficiently small. Define

$$\mathcal{E}_{\mathrm{good}}(t) := \bigcap_{t'=t}^{\infty} \left(\bigcap_{k=0}^{\underline{K}_{t,t'}} \bigcap_{q\in\{u,v\}} \neg\mathcal{E}^{(q)}_{\mathrm{bad}}(k,t') \right) \cap \left(\bigcap_{n=0}^{2\underline{K}_{t,t'}} \neg\mathcal{E}^{(u,v)}_{\mathrm{short}}(n,t') \right),$$

for $\underline{K}_{t,t'} := K_{t,t'} - M_G$, where $M_G > 0$ is such that the above event holds with probability at least $1-\delta'$ by the proof of Proposition 3.4.1. Define the conditional probability measure $\mathbb{P}_g(\cdot) := \mathbb{P}(\,\cdot \mid \mathcal{E}_{\mathrm{good}}(t))$. On the event $\mathcal{E}_{\mathrm{good}}(t)$, $d^{(G)}_{t'}(u,v) > 2\underline{K}_{t,t'}$ for all $t' \geqslant t$. Hence, recalling the definition of the graph neighbourhood $\mathcal{B}^{(G)}_t$ and its boundary $\partial\mathcal{B}^{(G)}_t$ from Definition 2.2.1, at all times $t' \geqslant t$ also the graph neighbourhoods of u and v of radius $\underline{K}_{t,t'}$ are disjoint, i.e.,

$$\mathbb{P}_g\left(\bigcap_{t'=t}^{\infty} \left\{ \mathcal{B}^{(G)}_{t'}(u,\underline{K}_{t,t'}) \cap \mathcal{B}^{(G)}_{t'}(v,\underline{K}_{t,t'}) = \varnothing \right\} \right) = 1.$$

Since any path connecting u and v has to pass through the boundary of the graph neighbourhoods, $\mathbb{P}_g$-a.s. for all t',

$$\begin{aligned} d^{(L)}_{t'}(u,v) &\geqslant \sum_{q\in\{u,v\}} d^{(L)}_{t'}\big(q, \partial\mathcal{B}^{(G)}_{t'}(q, \underline{K}_{t,t'})\big) \\ &\geqslant \sum_{q\in\{u,v\}} \sum_{k=0}^{\underline{K}_{t,t'}-1} d^{(L)}_{t'}\big(\partial\mathcal{B}^{(G)}_{t'}(q,k), \partial\mathcal{B}^{(G)}_{t'}(q,k+1)\big), \end{aligned} \tag{3.4.41}$$

where for two sets of vertices $\mathcal{V}, \mathcal{W} \subset [t']$ we define

$$d^{(L)}_{t'}(\mathcal{V},\mathcal{W}) := \min_{v\in\mathcal{V}, w\in\mathcal{W}} d^{(L)}_{t'}(v,w).$$

This leaves to show that, for some $M_L = M_L(\delta')$, $q \in \{u,v\}$, $C > 0$,

$$\mathbb{P}_g\left(\exists t' \geqslant t : \sum_{k=0}^{\underline{K}_{t,t'}-1} d^{(L)}_{t'}\big(\partial\mathcal{B}^{(G)}_{t'}(q,k), \partial\mathcal{B}^{(G)}_{t'}(q,k+1)\big) \leqslant Q_{t,t'} - M_L\right) \leqslant C\delta'. \tag{3.4.42}$$

We argue in three steps: we prove that it is sufficient to consider the error probabilities only along a specific subsequence $(t_i)_{i\geqslant 0}$ of times. This is needed to obtain a summable error bound in t'. Similarly to Proposition 2.4.1, we prove along the subsequence $(t_i)_{i\geqslant 0}$ an upper bound on the sizes of the graph neighbourhood boundaries of u and v up to radius $\underline{K}_{t,t_i}$. This allows to bound the minimal weight on an edge between vertices at distance k and $k+1$ from $q \in \{u,v\}$.

By the definition of $Q_{t,t'}$ and $K_{t,t'}$ in (3.1.1) and (3.3.2), due to the integer part, $\underline{K}_{t,t'}$ and the rhs between brackets in (3.4.42) decrease at the times

$$t_i := \min\{t' : K_{t,t} - K_{t,t'} = 2i\}, \qquad \text{for } i \in \{0,\dots,K_{t,t}/2\}, \tag{3.4.43}$$

while the lhs between brackets in (3.4.42) may decrease for any $t' \geqslant t$. Because the addition of new vertices can create new (shorter) paths, we have for $i \geqslant 1$

$$\begin{aligned} &\big\{\exists t' \in [t_{i-1}, t_i) : d^{(L)}_{t'}\big(q, \partial\mathcal{B}^{(G)}_{t'}(q, \underline{K}_{t,t'})\big) \leqslant Q_{t,t'} - M_L\big\} \\ &\qquad \subseteq \big\{d^{(L)}_{t_i}\big(q, \partial\mathcal{B}^{(G)}_{t_i}(q, \underline{K}_{t,t_i}+2)\big) \leqslant Q_{t,t_{i-1}} - M_L\big\}, \end{aligned}$$

where $K_{t,t'} = K_{t,t_i} + 2$ for $t' \in [t_{i-1}, t_i)$ follows from (3.4.43). By construction of t_i and $Q_{t,t'}$ in (3.3.2), where the summands are nonincreasing, there exists $M_1 > 0$ such that for all t

$$|Q_{t,t_i} - Q_{t,t_{i-1}}| \leqslant M_1.$$

This yields that we can bound (3.4.44) further to obtain

$$\begin{aligned}\big\{\exists t' \in [t_{i-1},t_i) : d^{(L)}_{t'}\big(q, \partial\mathcal{B}^{(G)}_{t'}(q, \underline{K}_{t,t'})\big) \leqslant Q_{t,t'} - M_L\big\} \\ \subseteq \big\{d^{(L)}_{t_i}\big(q, \partial\mathcal{B}^{(G)}_{t_i}(q, \underline{K}_{t,t_i}+2)\big) \leqslant Q_{t,t_i} - M_L + M_1\big\}.\end{aligned}$$

Hence, by a union bound over i, we can bound (3.4.42), i.e.,

$$\begin{aligned}&\mathbb{P}_g\Big(\exists t' \geqslant t : d^{(L)}_{t'}\big(q, \partial\mathcal{B}^{(G)}_{t'}(q, \underline{K}_{t,t'})\big) \leqslant Q_{t,t'} - M_L\Big) \qquad (3.4.44)\\ &\quad\leqslant \sum_{i=1}^{K_{t,t}/2} \mathbb{P}_g\Big(d^{(L)}_{t_i}\big(q, \partial\mathcal{B}^{(G)}_{t_i}(q, \underline{K}_{t,t_i}+2)\big) \leqslant Q_{t,t_i} - M_L + M_1\Big).\end{aligned}$$

Below, in Lemma 3.A.2 in Section 3.A.1, we show that a generalisation of Lemma 2.4.5 gives for B sufficiently large (depending on δ') and $m^{(t)}_{i,k}(B) := \exp\big(2B(1 \vee \log(t_i/t))(\tau-2)^{-k/2}\big)$ that

$$\begin{aligned}&\mathbb{P}_g\Bigg(\bigcup_{k=1}^{\underline{K}_{t,t_i}+2} \Big\{|\partial\mathcal{B}^{(G)}_{t_i}(q,k)| \geqslant m^{(t)}_{i,k}(B)\Big\}\Bigg) \qquad (3.4.45)\\ &\qquad\qquad \leqslant 2\exp\big(-B(1 \vee \log(t_i/t))\big).\end{aligned}$$

We denote the complement of the event inside the $\mathbb{P}$-sign by $\mathcal{E}^{(i)}_{\text{neigh}}(q)$ for a fixed i. Define the conditional probability measure $\mathbb{P}^{(i)}_{g,n}(\cdot) := \mathbb{P}\big(\cdot \mid \mathcal{E}_{\text{good}} \cap \mathcal{E}^{(i)}_{\text{neigh}}(u) \cap \mathcal{E}^{(i)}_{\text{neigh}}(v)\big)$. The number of edges connecting a vertex at distance k from q to a vertex at distance $k+1$ from q can then be bounded for all $k \leqslant \underline{K}_{t,t_i}+2$, i.e., $\mathbb{P}^{(i)}_{g,n}$-a.s.

$$\begin{aligned}|\partial\mathcal{B}^{(G)}_{t_i}(q,k)| \cdot |\partial\mathcal{B}^{(G)}_{t_i}(q,k+1)| &\leqslant m^{(t)}_{i,k}(B) \cdot m^{(t)}_{i,k+1}(B)\\ &\leqslant \exp\big(4B(1 \vee \log(t_i/t))(\tau-2)^{-(k+1)/2}\big)\\ &=: n_{i,k}. \qquad (3.4.46)\end{aligned}$$

Since all edges in the graph are equipped with i.i.d. copies of L, and as the minimum of K i.i.d. random variables is nonincreasing in K, we have by Lemma 2.3.11 for $\xi > 0$ that for $k \leqslant \underline{K}_{t,t_i}+1$

$$\begin{aligned}\mathbb{P}^{(i)}_{g,n}&\Big(d^{(L)}_{t_i}\big(\partial\mathcal{B}^{(G)}_{t_i}(q,k), \partial\mathcal{B}^{(G)}_{t_i}(q,k+1)\big) \leqslant F^{(-1)}_L\big(n_{i,k}^{-1-\xi}\big)\Big)\\ &\leqslant \mathbb{P}^{(i)}_{g,n}\Big(\min_{j\in[n_{i,k}]} L_{j,k} \leqslant F^{(-1)}_L\big(n_{i,k}^{-1-\xi}\big)\Big)\\ &\leqslant \exp\big(-4B\xi(1 \vee \log(t_i/t))(\tau-2)^{-(k+1)/2}\big).\end{aligned}$$

Recall (3.4.41). We apply the inequality in the event in the first row above for $k \leqslant \underline{K}_{t,t_i} + 1$ to obtain a bound on the lhs between brackets in the second row in (3.4.44), i.e., by a union bound

$$\begin{aligned}\mathbb{P}^{(i)}_{g,n}\bigg(d^{(L)}_{t_i}(q,\partial\mathcal{B}^{(G)}_{t_i}(q,\underline{K}_{t,t_i}+2)) &\leqslant \sum_{k=0}^{\underline{K}_{t,t_i}+1} F_L^{(-1)}\big(n_{i,k}^{-1-\xi}\big)\bigg) \\ &\leqslant \sum_{k=0}^{\underline{K}_{t,t_i}+1} \exp\big(-4B\xi(1\vee\log(t_i/t))(\tau-2)^{-(k+1)/2}\big) \\ &\leqslant 2\exp\big(-4B\xi(1\vee\log(t_i/t))(\tau-2)^{-1}\big). \end{aligned} \tag{3.4.47}$$

We now bound the sum in the above event from below to relate it to the rhs between brackets in (3.4.44). We do so by modifying the proof of Proposition 2.4.1. Afterwards, we bound the total error probability by taking a union bound over the times $(t_i)_{i\geqslant 0}$.

To bound $F_L^{(-1)}\big(n_{i,k}^{-1-\xi}\big)$ from below, we need an upper bound on $n_{i,k}$ in (3.4.46) since $z \mapsto F_L^{-(1)}(1/z)$ is nonincreasing. We first establish a lower and upper bound on t_i. Recall the integer $K_{t,t'}$ defined in (3.1.1), so that we may write for $t' \leqslant t_{K_{t,t}/2}$

$$K_{t,t'} = 2\big(\log\log(t) - \log\big(\log(t'/t)\vee 1\big)\big)/|\log(\tau-2)| - a_{t'}$$

for some $a_{t'} \in (0,1)$ being the fractional part of the expression. Using this notation one can verify that for $i \leqslant K_{t,t}/2$

$$t_i \in \big[t\exp\big((\tau-2)^{-i+1}\big), t\exp\big((\tau-2)^{-i-1}\big)\big] =: [\underline{t}_i, \bar{t}_i]. \tag{3.4.48}$$

Substituting the upper bound on t_i into $n_{k,i}$ in (3.4.46), yields that there exists $C = C(\xi, B) > 0$ such that

$$n_{i,k}^{1+\xi} \in \big[\exp\big((\tau-2)^{-i-k/2}/C\big), \exp\big(C(\tau-2)^{-i-k/2}\big)\big],$$

which implies that, recalling $\underline{K}_{t,t_i} = K_{t,t_i} - M_G$ for some constant $M_G > 0$ by (3.4.2),

$$\sum_{k=0}^{\underline{K}_{t,t_i}+1} F_L^{(-1)}\big(n_{i,k}^{-(1+\xi)}\big) \geqslant \sum_{k=0}^{\underline{K}_{t,t_i}-M_G+1} F_L^{(-1)}\big(\exp\big(-C(\tau-2)^{-i-k/2}\big)\big).$$

Observe that for a monotone nonincreasing function g, $g(1) < \infty$

$$\sum_{k=\lceil a\rceil+1}^{\lfloor b\rfloor} g(k) \overset{(\star)}{\leqslant} \int_a^b g(x)dx \overset{(*)}{\leqslant} \sum_{k=\lfloor a\rfloor}^{\lfloor b\rfloor} g(k). \tag{3.4.49}$$

Since $z \mapsto F_L^{-(1)}(1/z)$ is nonincreasing and bounded, we obtain by $(*)$ that

$$\sum_{k=0}^{K_{t,t_i}-M_G+1} F_L^{(-1)}\big(n_{i,k}^{-(1+\xi)}\big) \geqslant \int_{x=0}^{K_{t,t_i}} F_L^{(-1)}\big(\exp\big(-B(\tau-2)^{-i-x/2}\big)\big)dx - M.$$

Applying the change of variables $y = x + 2i + 2\log(B)/|\log(\tau-2)|$ yields for $C = 2\log(B)/|\log(\tau-2)|$ and some constant $M_L \geqslant M$ that

$$\begin{aligned}\sum_{k=0}^{K_{t,t_i}-M_G+1} & F_L^{(-1)}\big(n_{i,k}^{-(1+\xi)}\big) \\ &\geqslant \int_{y=2i+C}^{2i+K_{t,t_i}+C} F_L^{(-1)}\big(\exp\big(-(\tau-2)^{-y/2}\big)\big)dy - M \\ &\geqslant \int_{y=2i}^{2i+K_{t,t_i}} F_L^{(-1)}\big(\exp\big(-(\tau-2)^{-y/2}\big)\big)dy - M_L,\end{aligned}$$

again using that $z \mapsto F_L^{(-1)}(1/z)$ is bounded and nonincreasing. By transforming the integral back to a summation using $(\star)$ in (3.4.49), we obtain by definition of Q_{t,t_i} in (3.3.2), and $K_{t,t} - K_{t,t_i} = 2i$ in (3.5.15)

$$\begin{aligned}\sum_{k=0}^{K_{t,t_i}-M_G+1} & F_L^{(-1)}\big(n_{i,k}^{-(1+\xi)}\big) \\ &\geqslant \sum_{k=2i+1}^{2i+K_{t,t_i}} F_L^{(-1)}\big(\exp\big(-(\tau-2)^{-k/2}\big)\big) - M_L \\ &= \sum_{k=K_{t,t}-K_{t,t_i}+1}^{K_{t,t}} F_L^{(-1)}\big(\exp\big(-(\tau-2)^{-k/2}\big)\big) - M_L \\ &= Q_{t,t_i} - M_L.\end{aligned}$$

Using this lower bound inside the event in (3.4.47), yields that

$$\begin{aligned}\mathbb{P}_{g,n}^{(i)}\big(d_{t_i}^{(L)}&\big(q, \partial\mathcal{B}^{(G)}(q, \underline{K}_{t,t_i}+2)\big) \leqslant Q_{t,t_i} - M_L\big) \\ &\leqslant 2\exp\big(-4B\xi(1 \vee \log(t_i/t))(\tau-2)^{-1}\big).\end{aligned} \tag{3.4.50}$$

Recall that we would like to show (3.4.42). Its proof is accomplished by a union bound over the times $(t_i)_{i \leqslant \underline{K}_{t,t'/2}}$ if we show that there is a B sufficiently large such that the error probabilities on the rhs in (3.4.45) and (3.4.50) are smaller than δ' when summed over $i \leqslant \underline{K}_{t,t'}/2$. For this it is sufficient to show that for any $\hat{\delta} > 0$ and $C' > 0$ there exists $B > 0$ such that

$$\sum_{i=0}^{\infty} \exp\big(-C'B(1 \vee \log(t_i/t))\big) \leqslant \hat{\delta}.$$

This follows from the lower bound on t_i in (3.4.48), since for B large

$$\begin{aligned}\sum_{i=0}^{\infty} \exp\big(-C'B(1 \vee \log(t_i/t))\big) &\leqslant \sum_{i=0}^{\infty} \exp\big(-C'B(1 \vee \log(\underline{t}_i/t))\big)\\ &= \sum_{i=0}^{\infty} \exp\big(-C'B(1 \vee (\tau-2)^{-i+1})\big)\\ &\leqslant 2\exp\big(-C'B(\tau-2)\big) < \hat{\delta}.\end{aligned}$$

□

3.5 PROOF OF THE UPPER BOUND

The upper bound of Theorem 3.3.3 is stated in the following proposition.

Proposition 3.5.1 (Upper bound weighted distance)**.** *Consider the preferential attachment model with power-law exponent* $\tau \in (2,3)$*. Equip every edge upon creation with an i.i.d. copy of the non-negative random variable* L*. Let* u, v *be two typical vertices in* PA_t*. If* $I_2(L) < \infty$*, then for any* $\delta > 0$*, there exists* $M_L > 0$ *such that*

$$\mathbb{P}\Big(\exists t' : d_{t'}^{(L)}(u,v) \geqslant 2Q_{t,t'} + M_L\Big) \leqslant \delta. \tag{3.5.1}$$

Regardless of the value of $I_2(L)$*, for any* $\delta, \varepsilon > 0$*, there exists* $M_L > 0$ *such that*

$$\mathbb{P}\Big(\exists t' : d_{t'}^{(L)}(u,v) \geqslant 2(1+\varepsilon)Q_{t,t'} + M_L\Big) \leqslant \delta. \tag{3.5.2}$$

To prove Proposition 3.5.1, we have to show that for every $t' > t$ there is a t'-present (u,v)-path whose total weight is bounded from above by the rhs between brackets in (3.5.1) and (3.5.2), respectively. We construct a t'-present five-segment path $\boldsymbol{\pi}^{(t')}$ of three segment types. We write it as $\boldsymbol{\pi}^{(t')} = \overrightarrow{\boldsymbol{\pi}}_{u,0}^{(t)} \circ \overrightarrow{\boldsymbol{\pi}}_{u,1}^{(t')} \circ \boldsymbol{\pi}_{\mathrm{core}}^{(t')} \circ \overleftarrow{\boldsymbol{\pi}}_{v,1}^{(t')} \circ \overleftarrow{\boldsymbol{\pi}}_{v,0}^{(t)}$. Here, we denote for a path segment $\overrightarrow{\boldsymbol{\pi}} = (\pi_0, \ldots, \pi_n)$ its reverse by $\overleftarrow{\boldsymbol{\pi}} := (\pi_n, \ldots, \pi_0)$. The path segments are constructed similar to the methods demonstrated in Chapter 2.3. However, now we need stronger error bounds, so that the error terms are also small when *summed* over $t' \geqslant t$. Let $\delta' > 0$ be sufficiently small, and $(M_i)_{i \leqslant 3}$ be suitable positive constants.

Step 1. For $q \in \{u, v\}$, the path segment $\overrightarrow{\pi}^{(t)}_{q,0} := (q, \dots, q_0)$ connects q to a vertex q_0 that has indegree at least $s^{(0)}_0 > 0$ at time $(1-\delta')t$. The path segment uses only vertices that are older than $(1-\delta')t$. The path segments are fixed for all $t' \geqslant t$. The number of edges on $\overrightarrow{\pi}^{(t)}_{q,0}$ is bounded from above by $M_1 = M_1(\delta', s^{(0)}_0)$ with probability close to one.

a) Since the number of edges on $\overrightarrow{\pi}^{(t)}_{q,0}$ is bounded, its total weight can also be bounded by a constant. This is captured by the constant M_L in the statement of Proposition 3.5.1.

b) For the end vertex q_0 of $\overrightarrow{\pi}^{(t)}_{q,1}$, for $q \in \{u, v\}$, we identify the rate of growth of its indegree $(D^{\leftarrow}_{q_0}(t'))_{t' \geqslant t}$. We bound $D^{\leftarrow}_{q_0}(t')$ from below during the *entire interval* $[t, \infty)$ by a sequence that tends to infinity in t' sufficiently fast.

Step 2. For the path segments $\overrightarrow{\pi}^{(t')}_{u,1}, \pi^{(t')}_{\mathrm{core}}$, and $\overleftarrow{\pi}^{(t')}_{v,1}$ we argue, similar to the proof of Proposition 3.4.12, that it is sufficient to construct these path segments along a specific subsequence $(t_i)_{i \geqslant 0}$, as the t_i-present path segments have small enough total weight when compared to $Q_{t,t'}$ for all $t' \in (t_i, t_{i+1}]$. With a slight abuse of notation we abbreviate for the path (segments) $\pi^{(i)} := \pi^{(t_i)}$.

Step 3. For $q \in \{u, v\}$, the path segment $\overrightarrow{\pi}^{(i)}_{q,1}$ consists of at most $K_{t,t_i} + M_2$ edges and connects the vertex q_0 to the so-called i-th inner core, i.e., the set of vertices with a sufficiently large degree at time $(1-\delta')t_i$. These path segments use only edges that arrived *after* time $(1-\delta')t_i$ and the total weight of any such path segment is therefore independent of the total weight on the segments $\overrightarrow{\pi}^{(t)}_{u,0}$ and $\overrightarrow{\pi}^{(t)}_{v,0}$, that use only edges that arrived *before* $(1-\delta')t$. For the weighted distance, we construct the path segment $\overrightarrow{\pi}^{(i)}_{q,1}$ greedily (minimizing the edge weights) to bound the weighted distance between q_0 and the inner core from above by $Q_{t,t_i} + M_3$.

Step 4. Denote the end vertices of $\overrightarrow{\pi}^{(i)}_{u,1}$ and $\overrightarrow{\pi}^{(i)}_{v,1}$ by $w^{(i)}_u$ and $w^{(i)}_v$, respectively. The middle path segment $\pi^{(i)}_{\mathrm{core}}$ connects the two vertices $w^{(i)}_u, w^{(i)}_v$ in the inner core. The number of disjoint paths of bounded length between from $w^{(i)}_u$ to $w^{(i)}_v$ is growing polynomially in t'. This yields that $d^{(L)}_{t_i}(w^{(i)}_u, w^{(i)}_v)$ is bounded by a constant for all i. This weight is captured by M_L in Proposition 3.5.1.

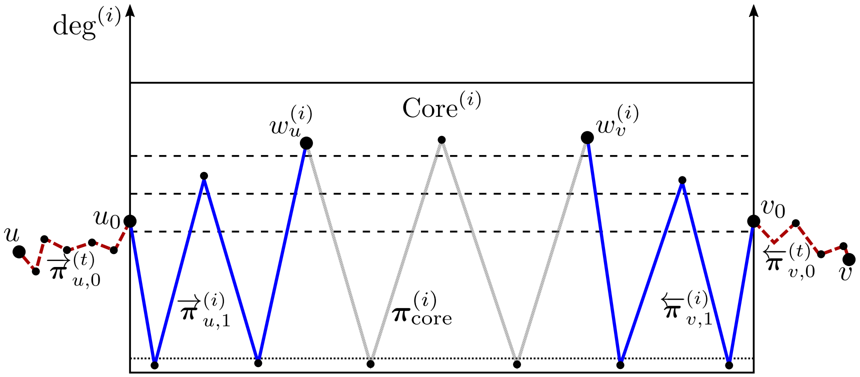


Figure 5: The five-segment path $\pi^{(i)}$ from u to v at a single time t_i. The y-axis represents the degree of the vertices at time t_i. The connected dots form the constructed path $\pi^{(i)}$ from u to v that is present at time t_i. The top continuous black line is the maximal degree in the graph at time t_i, while the dashed horizontal lines represent a degree-threshold sequence $(s_k^{(i)})_{i\geqslant 0, k\geqslant 0}$ defined in the proof of Proposition 3.5.6 for the segments $\overrightarrow{\pi}^{(i)}_{u,1}$ and $\overleftarrow{\pi}^{(i)}_{v,1}$.

Step 5. Eventually we *glue* the different path segments together and obtain the results (3.5.1) and (3.5.2). The path segments $\overrightarrow{\pi}^{(t')}_{u,1}, \pi^{(t')}_{\text{core}}$, and $\overleftarrow{\pi}^{(t')}_{v,1}$ change at the times $(t_i)_{i\geqslant 0}$, while the segments $\overrightarrow{\pi}^{(t)}_{u,0}$ and $\overleftarrow{\pi}^{(t)}_{v,0}$ stay the same for all $t' \geqslant t$.

See Figure 5 for a sketch of the constructed path and Figure 2 in the introduction for a visualization of the construction of the subsequence $(t_i)_{i\geqslant 0}$ and the control of the degree of the vertex q_0. Recall that the proof of the lower bound was based on controlling the dynamically changing *graph neighbourhood* up to distance $\underline{K}_{t,t'} = K_{t,t'} - M_G$. For the proof of the upper bound, the dynamics of the graph are mostly captured by controlling *the degree* of only two vertices, see Step 1b in the outline.

Step 1. Initial segments and degree evolution

Recall $D_v^{\leftarrow}(t')$, the indegree of vertex v at time t'.

Lemma 3.5.2 (Finding a high-degree vertex). *For any* $s_0^{(0)}, \delta' > 0$ *there exists* $M > 0$ *such that for a typical vertex* q *in* PA_t

$$\mathbb{P}\big(\nexists q_0 \in [(1-\delta')t] : d^{(L)}_{(1-\delta')t}(q, q_0) \leqslant M \textit{ and } D^{\leftarrow}_{q_0}\big((1-\delta')t\big) \geqslant s_0^{(0)}\big) \leqslant 3\delta'. \tag{3.5.3}$$

Proof. Let $\mathcal{E}_{\text{old}} := \{q < (1-\delta')t\}$. Since q is chosen uniformly among the first t vertices,

$$\mathbb{P}(\mathcal{E}_{\text{old}}) = 1 - \delta' + O(1/t).$$

Recall $\mathcal{B}_t^{(G)}(x, R)$ from (2.2.1). From minor adaptations of the proofs of [81, Theorem 3.6] for FPA, and [190, Proposition 5.10] for VPA it follows for all $\delta' > 0$ that there exists M_1 such that

$$\mathbb{P}\big(\mathcal{B}^{(G)}_{(1-\delta')t}(q, M_1) \cap \{x : D_x^{\leftarrow}\big((1-\delta')t\big) \geqslant s_0\} \neq \varnothing \mid \mathcal{E}_{\text{old}}\big) \geqslant 1 - \delta'.$$

We refer the reader there for the details. Conditionally on the above event between brackets there is a $(1-\delta')t$-present path segment from q to q_0, whose edges carry i.i.d. weights. Hence, there exists $M_2 > 0$ such that the weight on the segment can be bounded, yielding (3.5.3). □

The following lemma bounds the degree evolution from below. It uses martingale arguments that are inspired by [193]. However, here the statement only refers to the process of a single vertex that has initial degree at least s where [193] considers a set of vertices for any initial degree in FPA. Our statement applies to both FPA *and* VPA. Moreover, we consider the degree during the entire interval $[t, \infty)$. See also [24, Section 5.1] for results on the degree of an early vertex in $\mathrm{FPA}(m, \delta)$, where the considered vertex is born at time $o(t)$.

Lemma 3.5.3 (Indegree lower bound). *Consider the preferential attachment model with power-law parameter $\tau > 2$. Let q_0 be a vertex such that $D_{q_0}^{\leftarrow}(t) \geqslant s \geqslant 2$. There exists a constant $c > 0$, not depending on s, such that for all $\delta' > 0$*

$$\mathbb{P}\left(\exists t' \geqslant t : D_{q_0}^{\leftarrow}(t') \leqslant \delta' s (t'/t)^{1/(\tau-1)}\right) \leqslant c\delta'. \tag{3.5.4}$$

Proof. Let $\gamma := 1/(\tau-1)$. Both for FPA and VPA, it holds by Definitions 1.2.4 and 1.2.5 that

$$\mathbb{P}\left(D_{q_0}^{\leftarrow}(t'+1) \geqslant D_v^{\leftarrow}(t') + 1 \mid \mathrm{PA}_{t'}\right) \geqslant \gamma D_{q_0}^{\leftarrow}(t')/t'$$

for any $t' \geqslant t$. Let $(X_{t,t'})_{t' \geqslant t}$ be a discrete-time pure birth process satisfying $X_{t,t} = s$ and

$$\begin{aligned} &\mathbb{P}\big(X_{t,t'+1} = X_{t,t'} + 1 \mid X_{t,t'} = x\big) \\ &\qquad = 1 - \mathbb{P}\big(X_{t,t'+1} = X_{t,t'} \mid X_{t,t'} = x\big) = \gamma x/t'. \end{aligned} \tag{3.5.5}$$

Then the degree evolution $(D_{q_0}^{\leftarrow}(t'))_{t' \geqslant t}$ and $(X_{t,t'})_{t' \geqslant t}$ can be coupled such that the degree evolution dominates the birth process in the entire interval $[t,\infty)$. We first show that for any $k > -s$ and $\gamma \in (0,1)$, provided that $t' \geqslant t \geqslant k\gamma$,

$$Z^{(k)}_{t,t'} := \frac{\Gamma(t')}{\Gamma(t'+k\gamma)} \frac{\Gamma(t+k\gamma)}{\Gamma(t)} \frac{\Gamma(X_{t,t'}+k)}{\Gamma(X_{t,t'})} \tag{3.5.6}$$

is a non-negative martingale. The result will then follow by an application of the maximal inequality for $k = -1$. Clearly $\mathbb{E}[|Z^{(k)}_{t,t'}|] < \infty$, as the arguments in the Gamma functions in (3.5.6) are bounded away from 0. Moreover, since

$$\Gamma(x) = (x-1)\Gamma(x-1), \tag{3.5.7}$$

and using (3.5.5),

$$\begin{aligned}
\mathbb{E}\Big[\frac{\Gamma\big(X_{t,t'+1}+k\big)}{\Gamma(X_{t,t'+1})} \mid X_{t,t'}\Big] \\
&= \frac{t'-\gamma X_{t,t'}}{t'} \frac{\Gamma(X_{t,t'}+k)}{\Gamma(X_{t,t'})} + \frac{\gamma X_{t,t'}}{t'} \frac{\Gamma(X_{t,t'}+k+1)}{\Gamma(X_{t,t'}+1)} \\
&= \frac{\Gamma(X_{t,t'}+k)}{\Gamma(X_{t,t'})}\Big(1+\frac{k\gamma}{t'}\Big),
\end{aligned}$$

making it straightforward to verify that the martingale property for $Z^{(k)}_{t,t'}$ holds for $t' \geqslant t$. Due to Kolmogorov's maximal inequality, for any $T > t, \lambda > 0$

$$\mathbb{P}\Big(\sup_{t\leqslant t'\leqslant T} Z^{(k)}_{t,t'} \geqslant \lambda\Big) \leqslant \frac{\mathbb{E}[Z^{(k)}_{t,T}]}{\lambda} = \frac{Z^{(k)}_{t,t}}{\lambda} = \frac{\Gamma(X_{t,t}+k)}{\lambda\Gamma(X_{t,t})} = \frac{\Gamma(s+k)}{\lambda\Gamma(s)}. \tag{3.5.8}$$

Substituting (3.5.6) and $k = -1$, we see, using (3.5.7),

$$\begin{aligned}
\Big\{\sup_{t'\geqslant t} Z^{(-1)}_{t,t'} \geqslant \lambda\Big\} &= \Big\{\exists t' \geqslant t : \frac{\Gamma(t')}{\Gamma(t'-\gamma)} \frac{\Gamma(t-\gamma)}{\Gamma(t)} \frac{\Gamma\big(X_{t,t'}-1\big)}{\Gamma\big(X_{t,t'}\big)} \geqslant \lambda\Big\} \\
&= \Big\{\exists t' \geqslant t : X_{t,t'} - 1 \leqslant \frac{\Gamma(t')}{\Gamma(t'-\gamma)} \frac{\Gamma(t-\gamma)}{\Gamma(t)} \frac{1}{\lambda}\Big\} \\
&\supseteq \Big\{\exists t' \geqslant t : X_{t,t'} \leqslant \frac{\Gamma(t')}{\Gamma(t'-\gamma)} \frac{\Gamma(t-\gamma)}{\Gamma(t)} \frac{1}{\lambda}\Big\}.
\end{aligned}$$

Since $\Gamma(x)/\Gamma(x+a) = x^{-a}(1+O(1/x))$, by (3.5.8) there exists c' such that for t sufficiently large

$$\frac{\Gamma(t')}{\Gamma(t'-\gamma)} \frac{\Gamma(t-\gamma)}{\Gamma(t)} \leqslant c'(t'/t)^{\gamma}.$$

Choosing $\lambda = c'/(s\delta')$ yields

$$\mathbb{P}\Big(\exists t' \geqslant t : X_{t,t'} \leqslant \frac{1}{\lambda}c'(t'/t)^{\gamma}\Big) = \mathbb{P}\left(\exists t' \geqslant t : X_{t,t'} \leqslant \delta' s(t'/t)^{\gamma}\right) \leqslant \frac{\delta' s}{c'(s-1)}.$$

Now (3.5.4) follows, since $s \geqslant 2$ and $\big(D_{q_0}(t')\big)_{t'\geqslant t} \overset{d}{\geqslant} \big(X_{t,t'}\big)_{t'\geqslant t}$. □

Remark 3.5.4. *The proof of this lemma can be adapted to obtain an upper bound on the degree evolution, by applying Kolmogorov's maximal inequality to the martingale $Z^{(1)}_{t,t'}$.*

Step 2. Sufficient to construct the path along a subsequence of times

Lemma 3.5.5 (Subsequence of times). *Consider the preferential attachment model under the same conditions as Proposition 3.5.1 and let $(t_i)_{i\geqslant 0}$ be as defined in* (3.4.43). *If there exists $M_L > 0$ such that*

$$\mathbb{P}\big(\exists i \in [K_{t,t}/2] : d^{(L)}_{t_i}(u,v) \geqslant 2Q_{t,t_i} + M_L\big) \leqslant \delta, \tag{3.5.9}$$

then (3.5.1) *holds. Similarly,* (3.5.2) *holds if there exists $M_L > 0$ such that for any $\varepsilon > 0$ and t sufficiently large*

$$\mathbb{P}\big(\exists i \in [K_{t,t}/2] : d^{(L)}_{t_i}(u,v) \geqslant 2(1+\varepsilon)Q_{t,t_i} + M_L\big) \leqslant \delta. \tag{3.5.10}$$

Proof. Recall (3.3.2). The rhs between brackets in (3.5.1) only decreases at the times $(t_i)_{i\geqslant 0}$, while the lhs is nonincreasing in t'. By (3.3.2), for $i > K_{t,t}/2$, $Q_{t,t_i} = Q_{t,t_{K_{t,t}/2}}$ and the asserted statements follow. □

Step 3. Greedy path to the inner core

Define the i-th inner core, for t_i from (3.4.43), and with $\hat{t}_i := (1-\delta')t_i$, as

$$\mathrm{Core}^{(i)} := \{x \in [\hat{t}_i] : D_x\big(\hat{t}_i\big) \geqslant \hat{t}_i^{\frac{1}{2(\tau-1)}} \log^{-\frac{1}{2}}(\hat{t}_i)\}. \tag{3.5.11}$$

Proposition 3.5.6 (Weighted distance to the inner core). *Consider the preferential attachment model under the same conditions as Proposition 3.5.1. There exists $C > 0$ such that for every $\delta' > 0$, there exist $s_0^{(0)}, M > 0$ such that for a vertex q_0 satisfying $D_{q_0}\big((1-\delta')t\big) \geqslant s_0^{(0)}$, when $I_2(L) < \infty$,*

$$\mathbb{P}\Bigg(\bigcup_{i\leqslant K_{t,t}/2} \big\{d^{(L)}_{t_i}(q_0, \mathrm{Core}^{(i)}) \geqslant Q_{t,t_i} + M\big\}\Bigg) \leqslant C\delta'. \tag{3.5.12}$$

Regardless of the value of $I_2(L)$*, there exists* $M > 0$ *such that for every* $\varepsilon > 0$*, there is an* $s_0^{(0)} > 0$ *such that for a vertex* q_0 *satisfying* $D_{q_0}\big((1-\delta')t\big) \geqslant s_0^{(0)}$*,*

$$\mathbb{P}\left(\bigcup_{i \leqslant K_{t,t}/2} \left\{ d_{t_i}^{(L)}(q, \mathrm{Core}^{(i)}) \geqslant (1+\varepsilon)Q_{t,t_i} + M \right\} \right) \leqslant C\delta'. \qquad (3.5.13)$$

The bounds on the weighted distance in Proposition 3.5.6 are realized by constructing the segments $\overrightarrow{\pi}_{u,1}^{(i)}$ and $\overrightarrow{\pi}_{v,1}^{(i)}$, whose total weight we bound from above. For this we follow the same ideas as in Proposition 2.3.4, up to computational differences. Therefore, the proof we give here is not completely self-contained and for some bounds we will refer the reader back to Chapter 2.

Preparations for the proof of Proposition 3.5.6

For some constants $s_0^{(0)}, \delta' > 0$, the sequence $\varepsilon_k := (k+1)^{-2}$, and $\hat{t}_i$ from (3.5.11), define the degree threshold sequence

$$s_k^{(i)} = \begin{cases} \delta' s_0^{(0)} \left(\hat{t}_i/t\right)^{\frac{1}{\tau-1}} & k = 0, \quad (3.5.14a) \\ \min\left\{ \left(s_{k-1}^{(i)}\right)^{(1-\varepsilon_k)/(\tau-2)}, \hat{t}_i^{\frac{1}{2(\tau-1)}} \log^{-\frac{1}{2}}(\hat{t}_i) \right\} & k > 0. \quad (3.5.14b) \end{cases}$$

For each time t_i, the initial value $s_0^{(i)}$ is chosen such that it matches the bound on the degree in (3.5.4). The maximum value of $s_k^{(i)}$, for each fixed $\hat{t}_i$, matches the condition for vertices to be in the i-th inner core, see (3.5.11). Set

$$\kappa^{(i)} := \min\{k : s_{k+1}^{(i)} = s_k^{(i)}\}. \qquad (3.5.15)$$

Denote by $\mathcal{L}_k^{(i)}$ the k-th vertex layer: the set of vertices with degree at least $s_k^{(i)}$ at time $\hat{t}_i$, i.e.,

$$\mathcal{L}_k^{(i)} := \{x \in [\hat{t}_i] : D_x(\hat{t}_i) \geqslant s_k^{(i)}\}.$$

The path segment $\overrightarrow{\pi}_{q,1}^{(i)}$ to the inner core has length $2\kappa^{(i)}$ and uses alternately a *young* vertex $y_k^{(i)} \in [\hat{t}_i, t_i]$ and an *old* vertex $\pi_k^{(i)}$ from the layer $\mathcal{L}_k^{(i)}$. Thus, for $\pi_0 := q_0$, $\overrightarrow{\pi}_{q,1}^{(i)}$ has the form

$$\overrightarrow{\pi}_{q,1}^{(i)} = (\pi_0, y_1^{(i)}, \pi_1^{(i)}, \ldots, y_{\kappa^{(i)}}^{(i)}, \pi_{\kappa^{(i)}}^{(i)}).$$

To keep notation light we omit a subscript q for the individual vertices on the segments $\overrightarrow{\pi}_{q,1}^{(i)}$ for $q \in \{u, v\}$.

In the next lemmas, we show that $\vec{\pi}^{(i)}_{q,1}$ exists for all i with probability close to one, and bound its total weight. We outline the steps briefly. Using the choice of $s^{(i)}_k$, we bound $\kappa^{(i)}$ in terms of K_{t,t_i}. Then, since the number of vertices that have degree at least $s^{(i)}_{k+1}$ at time $\hat{t}_i$ is sufficiently large, it is likely that there are *many* connections from a vertex $\pi^{(i)}_k$ via a connector vertex $y^{(i)}_{k+1}$ to the $(k+1)$-th layer. We denote the set of connectors by $\mathcal{A}^{(i)}_{k+1}(\pi^{(i)}_k)$, i.e., for a vertex $\pi^{(i)}_k \in \mathcal{L}^{(i)}_k$,

$$\mathcal{A}^{(i)}_{k+1}\big(\pi^{(i)}_k\big) := \{y \in (\hat{t}_i, t_i] : \exists x^{(i)}_{k+1} \in \mathcal{L}^{(i)}_{k+1} : \pi^{(i)}_k \leftrightarrow y \leftrightarrow x^{(i)}_{k+1}\}.$$

Given $(\pi_0, \pi^{(i)}_1, \ldots, \pi^{(i)}_k)$, we greedily set, if $\mathcal{A}^{(i)}_{k+1}(\pi^{(i)}_k)$ is non-empty,

$$(y^{(i)}_{k+1}, \pi^{(i)}_{k+1}) := \underset{\substack{(y, x^{(i)}_{k+1}) \in \\ [\hat{t}_i, t_i] \times \mathcal{L}^{(i)}_{k+1}}}{\arg\min} \big\{ L_{(\pi^{(i)}_k, y)} + L_{(y, x^{(i)}_{k+1})} \big\}. \tag{3.5.16}$$

If there exists $k \leqslant \kappa^{(i)}$ such that $\mathcal{A}^{(i)}_{k+1}(\pi^{(i)}_k) = \varnothing$, we say that the construction has failed. When the construction succeeds, we can bound the weighted distance to the inner core, i.e.,

$$\begin{aligned} d^{(L)}_{t_i}(\pi_0, \mathrm{Core}^{(i)}) &\leqslant \sum_{k=0}^{\kappa^{(i)}-1} d^{(L)}_{t_i}(\pi^{(i)}_k, \pi^{(i)}_{k+1}) \\ &\leqslant \sum_{k=0}^{\kappa^{(i)}-1} \Big(L_{(\pi^{(i)}_k, y^{(i)}_{k+1})} + L_{(y^{(i)}_{k+1}, \pi^{(i)}_{k+1})} \Big). \end{aligned} \tag{3.5.17}$$

We show that for all i there exists a sequence $(n^{(i)}_k)_{k \leqslant \kappa^{(i)}}$ such that $|\mathcal{A}^{(i)}_{k+1}\big(\pi^{(i)}_k\big)| \geqslant n^{(i)}_k$ for all $k \leqslant \kappa^{(i)}$ with probability close to one. This allows to bound the minimal weight in the rhs of (3.5.16) from above, so that eventually this yields an upper bound for the rhs in (3.5.17).

We start with a lemma that relates K_{t,t_i} to $\kappa^{(i)}$, half the length of $\vec{\pi}^{(i)}_{q,1}$. Also, we show that $k \mapsto s^{(i)}_k$ is bounded from below by a doubly exponentially growing sequence.

Lemma 3.5.7. *Let* $\big(s^{(i)}_k\big)_{i \leqslant K_{t,t}/2, k \leqslant \kappa^{(i)}}$ *as in* (3.5.14b), *with* $s^{(0)}_0$ *sufficiently large. Then*

$$s^{(i)}_k \geqslant \Big(\delta' s^{(0)}_0 \left(\hat{t}_i / t\right)^{1/(\tau-1)} \Big)^{c'(\tau-2)^{-k}} \tag{3.5.18}$$

for some constant $c' > 0$. *There exists* $M \in \mathbb{N}$ *such that for* $\kappa^{(i)}$ *defined in* (3.5.15) *and* $i \leqslant K_{t,t}/2$

$$\kappa^{(i)} \leqslant K_{t,t_i}/2 + M. \tag{3.5.19}$$

Proof. By our choice $\varepsilon_k = (k+1)^{-2}$ before (3.5.14b), it holds that $\prod_{j=1}^{\infty}(1-\varepsilon_k) > 0$. The bound (3.5.18) follows immediately from the definition of $(s_k^{(i)})$ in (3.5.14b). By Lemma 2.3.7, the bound (3.5.19) immediately follows for $i = 0$, leaving to verify the bound for $i \geqslant 1$. By the choice of $\kappa^{(i)}$ in (3.5.15) and $(s_k^{(i)})$ in (3.5.14b), and the bound (3.5.18), for any $k \geqslant \kappa^{(i)}$ it holds that

$$\left(\delta' s_0^{(0)} \left(\hat{t}_i/t\right)^{1/(\tau-1)}\right)^{c'(\tau-2)^{-k}} \geqslant \hat{t}_i^{\frac{1}{2(\tau-1)}} \log^{-\frac{1}{2}}(\hat{t}_i)$$

Taking logarithms twice, and rearranging gives that for $k \geqslant \kappa^{(i)}$,

$$\begin{aligned} k\log(1/(\tau-2)) + \log\log(\hat{t}_i/t) + \log\left(1 + \frac{(\tau-1)\log\left(\delta' s_0^{(0)}\right)}{\log(\hat{t}_i/t)}\right) \\ \geqslant \log\log(\hat{t}_i) + \log\left(\frac{1}{2c'} - \frac{\tau-1}{2c'}\frac{\log\log(\hat{t}_i)}{\log(\hat{t}_i)}\right). \end{aligned}$$

From (3.5.11) it follows for all $i \leqslant K_{t,t}/2$ that $\hat{t}_i \geqslant (1-\delta')t$. Thus, the last terms on both lines are bounded by a constant for large t. Hence, there is $M = M(\tau)$ such that if $s_0^{(0)} \geqslant 1/\delta'$, $i \geqslant 1$

$$\kappa^{(i)} \leqslant \frac{\log\log(\hat{t}_i) - \log\log(\hat{t}_i/t)}{|\log(\tau-2)|} + M'.$$

By construction of $(t_i)_{i \leqslant K_{t,t}/2}$ in (3.4.48), there exists $b > 0$ such that $\hat{t}_i \leqslant t^b$, for all $i \leqslant K_{t,t}/2$. This relates $\log\log(\hat{t}_i)$ to $\log\log(t)$. Hence, there exists M such that (3.5.19) holds for $i \geqslant 1$, recalling $\hat{t}_i = (1-\delta')t_i$, and the definition of $K_{t,t'}$ in (3.1.1). □

We now prove Proposition 3.5.6. We construct the segment $\overrightarrow{\pi}_{q,1}^{(i)}$ for $i \leqslant \kappa^{(i)}$ and $q \in \{u, v\}$.

Proof of Proposition 3.5.6. Let $\mathcal{E}_{\mathrm{deg}} := \{\forall i \geqslant 0 : D_{q_0}^{\leftarrow}(\hat{t}_i) \geqslant s_0^{(i)}\}$. By Lemma 3.5.3 and the choice of $(s_0^{(i)})_{i\geqslant 0}$ in (3.5.14a) we have that

$$\mathbb{P}(\neg\mathcal{E}_{\mathrm{deg}}) \leqslant c\delta'.$$

We write $\mathbb{P}_{\mathrm{deg}}(\cdot) := \mathbb{P}(\cdot \mid \mathcal{E}_{\mathrm{deg}})$. We will first show that with probability close to one, the sets of connectors are sufficiently large. More precisely, for a set of vertices $\{\pi_k^{(i)}\}_{k \leqslant \kappa^{(i)}, i \leqslant K_{t,t}/2}$, such that $\pi_k^{(i)} \in \mathcal{L}_k^{(i)}$ and setting $\pi_0^{(i)} := \pi_0 = q_0$ for all i, we show that

$$\mathbb{P}_{\mathrm{deg}}\left(\bigcup_{i \leqslant K_{t,t}/2} \bigcup_{k \leqslant \kappa^{(i)}} \left\{|\mathcal{A}_{k+1}^{(i)}(\pi_k^{(i)})| \leqslant n_k^{(i)}\right\}\right) \leqslant \delta_1\left(s_0^{(0)}\right), \qquad (3.5.20)$$

where $\delta_1\big(s_0^{(0)}\big)$ is a function that tends to 0 as $s_0^{(0)}$ tends to infinity and, for $c_1 > 0$ chosen below,

$$n_k^{(i)} := c_1 \delta'\big(s_k^{(i)}\big)^{\varepsilon_k}. \tag{3.5.21}$$

Then, conditioning on the complement of the event in (3.5.20), we will bound the minimal weight of connections to $\mathcal{L}_{k+1}^{(i)}$ via the sets $\mathcal{A}_{k+1}^{(i)}\big(\pi_k^{(i)}\big)$ using (3.5.16) and arrive to (3.5.12) and (3.5.13) using the construction of the greedy path in (3.5.16). We follow the same steps as in Lemma 2.3.10. For notational convenience we leave out the superscript (i) for the various sequences and sets whenever it is clear from the context. For a set of vertices $\mathcal{V} \subset [t']$, define $D_{\mathcal{V}}^{\leftarrow}(t') := \sum_{x\in\mathcal{V}} D_x^{\leftarrow}(t')$. By Lemma 2.3.8, the probability that an arbitrary vertex in $(\hat{t}_i, t_i]$ is in $\mathcal{A}_{k+1}(\pi_k)$, is at least

$$p_k(\pi_k, \mathcal{L}_{k+1}) := \frac{1}{\hat{t}_i^2}\eta D_{\pi_k}^{\leftarrow}(\hat{t}_i) D_{\mathcal{L}_{k+1}}^{\leftarrow}(\hat{t}_i), \tag{3.5.22}$$

for some constant $\eta > 0$, where this event happens independently of the other vertices. Since the set $(\hat{t}_i, t_i]$ contains $\delta' t_i$ vertices, the random variable $|\mathcal{A}_{k+1}(\pi_k)|$ stochastically dominates a binomial random variable, i.e.,

$$|\mathcal{A}_{k+1}(\pi_k)| \overset{d}{\geqslant} \mathrm{Bin}\big(\delta' t_i, p_k(\pi_k, \mathcal{L}_{k+1})\big) =: A_k. \tag{3.5.23}$$

Let $c_2 := c_{2.3.9}$ be the constant from Lemma 2.3.9. Conditioning on $D_{\mathcal{L}_{k+1}}^{\leftarrow}(\hat{t}_i)$ yields that

$$\mathbb{E}[A_k] \geqslant \mathbb{E}\big[A_k \mid D_{\mathcal{L}_{k+1}}^{\leftarrow}(\hat{t}_i) \geqslant c_2 \hat{t}_i s_{k+1}^{2-\tau}\big] \cdot \mathbb{P}\big(D_{\mathcal{L}_{k+1}}^{\leftarrow}(\hat{t}_i) \geqslant c_2 \hat{t}_i s_{k+1}^{2-\tau}\big), \tag{3.5.24}$$

where the latter factor equals $1 - o(1)$ by Lemma 2.3.9. Since $\pi_k \in \mathcal{L}_k$, we have that $D_{\pi_k}^{\leftarrow}(\hat{t}_i) \geqslant s_k$. We substitute this and the conditioned bound on $D_{\mathcal{L}_{k+1}}^{\leftarrow}(\hat{t}_i)$ in (3.5.24) into $p_k(\pi_k, \mathcal{L}_{k+1})$ in (3.5.22). By the recursive definition of $s_k^{(i)}$ in (3.5.14b) and $n_k^{(i)}$ in (3.5.21) we obtain that there exists $c_1 > 0$ such that

$$\mathbb{E}[A_k] \geqslant \delta' t_i \frac{\eta c_2 \hat{t}_i (s_{k+1}^{(i)})^{2-\tau} s_k^{(i)}}{\hat{t}_i^2}(1 - o(1)) \geqslant 2c_1 \delta' (s_k^{(i)})^{\varepsilon_k} = 2n_k^{(i)}.$$

An application of Chernoff's bound and the constructed stochastic domination (3.5.23) yields for $\pi_k^{(i)} \in \mathcal{L}_k^{(i)}$ that

$$\mathbb{P}_{\mathrm{deg}}\big(|\mathcal{A}_{k+1}^{(i)}\big(\pi_k^{(i)}\big)| \leqslant n_k^{(i)}\big) \leqslant \exp\big(-n_k^{(i)}/4\big).$$

For details we refer the reader to the proof of Lemma 2.3.10. We return to (3.5.20). By a union bound over the layers $k \leqslant \kappa^{(i)}$ and times $(t_i)_{i \leqslant K_{t,t}/2}$, it remains to show that

$$\sum_{i=1}^{K_{t,t}/2} \sum_{k=1}^{\kappa^{(i)}} \exp\left(-n_k^{(i)}/4\right) \longrightarrow 0, \qquad \text{as } s_0^{(0)} \to \infty, \tag{3.5.25}$$

with $n_k^{(i)}$ from (3.5.21). We postpone showing this to the end of the proof.

Define the conditional probability measure

$$\mathbb{P}_{d,c}(\cdot) := \mathbb{P}\Big(\cdot \,\Big|\, \mathcal{E}_{\deg} \cap \bigcap_{i=1}^{K_{t,t}/2} \bigcap_{k=1}^{\kappa^{(i)}} \{|\mathcal{A}_{k+1}^{(i)}(\pi_k^{(i)})| > n_k^{(i)}\}\Big).$$

Thus, the path segment $\vec{\pi}_{q,1}$ from q_0 to the inner core exists $\mathbb{P}_{d,c}$-a.s. We greedily choose the vertices $(y_k^{(i)}, \pi_{k+1}^{(i)})$ as in (3.5.16). We bound the weight of the segment, i.e., the rhs of (3.5.17) to prove (3.5.12) and (3.5.13). Let $L_{m,n}^{(i)}$ be i.i.d. copies of L. Since the minimum of N i.i.d. random variables is nonincreasing in N, the weighted distance between $\pi_k^{(i)}$ and $\pi_k^{(i)}$ can be bounded for $k \leqslant \kappa^{(i)} - 1$, i.e., for $k \leqslant \kappa^{(i)}$ and $i \leqslant K_{t,t}/2$, $\mathbb{P}_{d,c}$-a.s.

$$d_{t_i}^{(L)}(\pi_k^{(i)}, \pi_{k+1}^{(i)}) \leqslant \min_{j \in [n_k^{(i)}]} \left(L_{1,j}^{(i)} + L_{2,j}^{(i)}\right).$$

Applying $(\star)$ in (2.3.24) yields for $\xi \in (0,1)$ that

$$\mathbb{P}_{d,c}\Big(d_{t_i}^{(L)}(\pi_k^{(i)}, \pi_{k+1}^{(i)}) \geqslant F_{L_1+L_2}^{(-1)}\Big(\big(n_k^{(i)}\big)^{-1+\xi}\Big)\Big) \leqslant \exp\left(-\big(n_k^{(i)}\big)^{\xi}\right),$$

where $F_{L_1+L_2}^{(-1)}$ denotes the generalised inverse of the distribution of the sum of two i.i.d. copies of L. Recall the bound (3.5.17). By a union bound over the subsegments $(\pi_k^{(i)}, y_k^i, \pi_{k+1}^{(i)})$ for $k \leqslant \kappa^{(i)}$ and the times $(t_i)_{i \leqslant K_{t,t}/2}$,

$$\mathbb{P}_{d,c}\Bigg(\bigcup_{i \leqslant K_{t,t}/2} \Big\{d_{t_i}^{(L)}(q_0, \mathrm{Core}^{(i)}) \geqslant \sum_{k=0}^{\kappa^{(i)}-1} F_{L_1+L_2}^{(-1)}\Big(\big(n_k^{(i)}\big)^{-1+\xi}\Big)\Big\}\Bigg)$$
$$\leqslant \sum_{i=0}^{K_{t,t}/2} \sum_{k=0}^{\kappa^{(i)}-1} \exp\left(-\big(n_k^{(i)}\big)^{\xi}\right). \tag{3.5.26}$$

To bound the error probabilities in (3.5.25) and (3.5.26), and the sum inside the event in (3.5.26), we bound $n_k^{(i)}$ from below. By its definition in (3.5.21), the bound on t_i in (3.4.48) and the bound on $s_k^{(i)}$ in (3.5.18)

$$\begin{aligned} n_k^{(i)} &= c_1\delta'\big(s_k^{(i)}\big)^{\varepsilon_k} \\ &\geqslant c_1\delta'\left(\delta' s_0^{(0)}\left(\hat{t}_i/t\right)^{1/(\tau-1)}\right)^{c'(\tau-2)^{-k}\varepsilon_k} \\ &= c_1\delta'\left(\delta' s_0^{(0)}(1-\delta')^{1/(\tau-1)}\exp\left(\frac{\tau-2}{\tau-1}(\tau-2)^{-i}\right)\right)^{c'(\tau-2)^{-k}\varepsilon_k} \end{aligned}$$

Assuming that $s_0^{(0)} \geqslant \delta'^2(1-\delta')^{2/(\tau-1)}$, we obtain since $\varepsilon_k = (k+1)^{-2}$ by definition above (3.5.14a)

$$\begin{aligned} n_k^{(i)} &\geqslant c_1\delta'\left(\exp\left(\frac{1}{2}\log(s_0^{(0)}) + \frac{\tau-2}{\tau-1}(\tau-2)^{-i}\right)\right)^{c'(\tau-2)^{-k}\varepsilon_k} \\ &= c_1\delta'\exp\Big(\frac{c'}{2}\log(s_0^{(0)})(k+1)^{-2}(\tau-2)^{-k} \\ &\qquad\qquad + \frac{\tau-2}{\tau-1}(k+1)^{-2}(\tau-2)^{-(i+k)}\Big). \end{aligned} \tag{3.5.27}$$

Since $(\tau-2)^{-k}$ grows exponentially for $\tau \in (2,3)$, while $\varepsilon_k = (1+k)^{-2}$ decreases polynomially, there exist $c_3 > 0$, $c_4 > 1$ such that $n_k^{(i)} \geqslant c_1\delta'\exp(\log(s_0)c_3c_4^k + c_3c_4^{k+i})$. Substituting this bound into (3.5.25) and (3.5.26), we observe that the terms are summable in both i and k and tend to zero as $s_0^{(0)}$ tends to infinity.

We are left with relating the sum on the rhs in the event in (3.5.26) to the rhs in (3.5.12) and (3.5.13), respectively, assuming that $s_0^{(0)}$ is large. Recalling (3.5.27), we assume $s_0^{(0)}$ is sufficiently large so that there exists $c_5 > 0$

$$\big(n_k^{(i)}\big)^{(1-\xi)} \geqslant \exp\big(2c_5(k+1)^{-2}(\tau-2)^{-(i+k)}\big).$$

Since $z \mapsto F_L^{(-1)}(1/z)$ is nonincreasing, and $\kappa^{(i)} \leqslant K_{t,t_i}/2 + M$ by (3.5.19), we obtain

$$\begin{aligned} &\sum_{k=0}^{\kappa^{(i)}-1} F_{L_1+L_2}^{(-1)}\Big(\big(n_k^{(i)}\big)^{-1+\xi}\Big) \\ &\qquad \leqslant \sum_{k=0}^{K_{t,t_i}/2+M-1} F_{L_1+L_2}^{(-1)}\Big(\exp\big(-2c_5(k+1)^{-2}(\tau-2)^{-(i+k)}\big)\Big). \end{aligned}$$

Recall $\varepsilon > 0$ from the statement of Proposition 3.5.6. In Claim 3.A.3 we show that for all $\varepsilon > 0$ there exists $M > 0$ such that

$$\sum_{k=0}^{K_{t,t_i}/2+M-1} F^{(-1)}_{L_1+L_2}\Big(\exp\big(-c_5(k+\zeta)^{-2}(\tau-2)^{-(i+k)}\big)\Big) \leqslant (1+\varepsilon\mathbb{1}_{\{I_2(L)=\infty\}})Q_{t,t_i}+M.$$

Substituting this bound inside the event in (3.5.26) and recalling that $s_0^{(0)}$ is chosen sufficiently large so that the total error probability from (3.5.25) and (3.5.26) is at most $C\delta'$ yields Proposition 3.5.6. □

Step 4. Bridging the inner core

We prove a lemma that shows that the path segments $(\boldsymbol{\pi}^{(i)}_{\text{core}})_{i\leqslant K_{t,t}/2}$ exist and their total weight is bounded from above by a constant for all i. Recall $\text{Core}^{(i)}$ from (3.5.11).

Lemma 3.5.8. *Consider the preferential attachment model with power-law exponent $\tau \in (2,3)$. Let $\{w_u^{(i)}, w_v^{(i)}\}_{i\leqslant K_{t,t}/2}$ be a set of vertices such that for all i, $w_u^{(i)}, w_v^{(i)} \in \text{Core}^{(i)}$. Then for every $\delta' > 0$, there exists $M > 0$ such that*

$$\mathbb{P}\Bigg(\bigcup_{i\leqslant K_{t,t}/2}\Big\{d^{(L)}_{t_i}(w_u^{(i)}, w_v^{(i)}) \geqslant M\Big\}\Bigg) \leqslant \delta'. \tag{3.5.28}$$

Proof. From [81, Proposition 3.2] it follows for FPA that for fixed i, whp,

$$\mathbb{P}\Big(\text{diam}_G^{(t_i)}(\text{Core}^{(i)}) \leqslant \frac{2(\tau-1)}{3-\tau}+6\Big) \geqslant 1-o(1/t_i), \tag{3.5.29}$$

where $\text{diam}_G^{(t')}(\mathcal{V}) := \max_{w_1,w_2\in\mathcal{V}} d^{(G)}_{t'}(w_1,w_2)$ for a set of vertices $\mathcal{V} \subset [t']$. The statement (3.5.29) holds also for VPA as explained in the proof of Proposition 2.3.5. A union bound yields that

$$\mathbb{P}\Bigg(\bigcup_{i\leqslant K_{t,t}/2}\Big\{\text{diam}_G(\text{Core}^{(i)}) > \frac{2(\tau-1)}{3-\tau}+6\Big\}\Bigg) = \sum_{i=1}^{K_{t,t}/2} o(1/t_i) = o(1),$$

since $K_{t,t} = O(\log\log(t))$, and t_i is increasing in i. We sketch how to extend this result to weighted distances, using the construction in the proof of [81, Proposition 3.2] which in turn relies on [36, Chapter 10]. In [81, Proposition 3.2] it is shown that the inner core dominates an Erdős-Rényi random graph (ERRG) $G(n_i, p_i)$, where there is an edge between two vertices $x, y \in \mathrm{Core}^{(i)}$ if there is a connector in $[(1-\delta')t_i, t_i]$ in PA_{t_i}, where

$$n_i = \sqrt{t_i}, \qquad p_i = \frac{1}{2} t_i^{\frac{1}{(\tau-1)}-1} \log^{-2}(t_i).$$

The weight on the edge (w_1, w_2) in the ERRG is $L_{(i,y)} + L_{(y,j)}$, where y is a uniformly chosen connector of w_1 and w_2 in PA_{t_i}. Now, for the construction used in [36, Chapter 10], one can embed two r_i-regular trees of depth $\Delta > 0$ in the ERRG, rooted in $w_u^{(i)}$ and $w_v^{(i)}$, respectively, whp. Here $r_i \geqslant t_i^a$, for some $a > 0$, and Δ is a constant such that all vertices at distance Δ from their root are members of both trees. Denote this event by $\mathcal{E}_{\mathrm{tree}}$. On this event, there are at least r_i disjoint paths from $w_u^{(i)}$ to $w_v^{(i)}$ in PA_{t_i} of 4Δ edges, and we can bound

$$d_{t_i}^{(L)}(w_u^{(i)}, w_v^{(i)}) \leqslant \min_{n \leqslant r_i} \sum_{j=1}^{4\Delta} L_j^{(n)},$$

for i.i.d. copies of L. Moreover, for $F_{L_1+\dots L_{4\Delta}}^{(-1)}$ being the distribution of the sum of 4Δ i.i.d. copies of L, for C sufficiently large

$$\mathbb{P}\Bigg(\bigcup_{i \leqslant K_{t,t}/2} \Big\{ \min_{j \leqslant t_i^a} \sum_{j=1}^{4\Delta} L_j^{(n)} \geqslant C \Big\} \Bigg) \leqslant \sum_{i=1}^{K_{t,t}/2} \big(1 - F_{L_1+\dots L_{4\Delta}}(C)\big)^{t_i^a} < \delta',$$

since by choosing C large, but independently of t, $F_{L_1+\dots L_{4\Delta}}(C)$ can be brought arbitrarily close to 1. The asserted bound (3.5.28) follows from a union bound over the above event and $\neg\mathcal{E}_{\mathrm{tree}}$. □

Step 5. Gluing the segments

We are ready to prove the main proposition of this section.

Proof of Proposition 3.5.1. Recall Lemma 3.5.5. We have to show that at the times $(t_i)_{i \geqslant 0}$ there is a path from u to v such that its total weight is bounded from above by the rhs between brackets in (3.5.9) and (3.5.10), with probability at least $1-\delta$. Let $C_{3.5.6}$ be the constant from Proposition 3.5.6. Set $\delta' = \delta/(7+2C_{3.5.6})$. Let $M_{3.5.2}$ and $M_{3.5.6}$ be the constants obtained from applying Lemma 3.5.2 and Proposition 3.5.6 for δ', respectively. Lastly, let $M_{3.5.8}$ be the constant from applying Lemma 3.5.8 for δ'. The existence of the path segments follows now directly from a union bound over the events described in Lemma 3.5.3 and Proposition 3.5.6

for $q \in \{u, v\}$, and Lemma 3.5.8. Hence, the summed error probability is $(2 \cdot 3 + 2C + 1)\delta' = \delta$. The total weight of the constructed paths $(\boldsymbol{\pi}^{(i)})_{i \geqslant 0}$ is bounded from above by $2Q_{t,t_i} + 2(M_{3.5.2} + M_{3.5.6}) + M_{3.5.8}$ for all $i \geqslant 0$. Thus, setting $M_L := 2(M_{3.5.2} + M_{3.5.6}) + M_{3.5.8}$ in the statement of Proposition 3.5.1 finishes the proof. □

3.A REMAINING PROOFS

3.A.1 *Lower bound*

Proof of Lemma 3.4.10. We verify (3.4.26) by induction. Recall the initial values in (3.4.24) and (3.4.25). We initialize the induction for $j = 1$. By (2.4.5), since $x < t'$

$$f^{(t,t')}_{[i,i+1)}(t', x) = p(t', x) = \nu x^{-\gamma} t'^{\gamma-1} = \nu t'^{\gamma-1} \cdot x^{-\gamma} + 0 \cdot x^{\gamma-1},$$

establishing (3.4.26) for $j = 1$. We advance the induction so that we may assume (3.4.26) for $j = k$. Then, using the definition of f in (3.4.17), which counts only the good paths and relies on the product form of p in (2.4.5), we can write

$$f^{(t,t')}_{[i,i+k+1)}(t', x) \leqslant \sum_{z=\ell_{i+k,t'}}^{t'} f^{(t,t')}_{[i,i+k)}(t', z) p(z, x).$$

This bound does not hold with equality, because the first factor on the rhs counts the good self-avoiding paths from t' to z, but the vertex x is not necessarily excluded in these paths, while these paths are excluded on the lhs. Recall from (2.4.5) that $p(z, x) = \nu(x \wedge z)^{-\gamma}(x \vee z)^{\gamma-1}$. Since f only counts the good paths, observe that if $z < x$, then $x > \ell_{i+k,t'}$. Thus, splitting the sum in two, whether $z \geqslant x$ or $z < x$, we obtain that

$$f^{(t,t')}_{[i,i+k+1)}(t', x) \leqslant \nu x^{-\gamma} \sum_{z=\ell_{i+k,t'} \vee x}^{t'} f^{(t,t')}_{[i,i+k)}(t', z) z^{\gamma-1} \qquad (3.A.1)$$

$$+ \mathbb{1}_{\{x > \ell_{i+k,t'}\}} \nu x^{\gamma-1} \sum_{z=\ell_{i+k,t'}}^{x-1} f^{(t,t')}_{[i,i+k)}(t', z) z^{-\gamma}.$$

By the induction hypothesis (3.4.26), we have that

$$\begin{aligned} f^{(t,t')}_{[i,i+k+1)}(t',x) \leqslant \nu x^{-\gamma}\phi_{[i,i+k)} &\sum_{z=\ell_{i+k,t'}\vee x}^{t'} z^{-1} \\ &+ \nu x^{-\gamma}\psi_{[i,i+k)}\sum_{z=\ell_{i+k-1,t'}\vee x}^{t'} z^{2\gamma-2} \\ &+ \mathbb{1}_{\{x>\ell_{i+k,t'}\}}\nu x^{\gamma-1}\phi_{[i,i+k)}\sum_{z=\ell_{i+k,t'}}^{x-1} z^{-2\gamma} \\ &+ \mathbb{1}_{\{x>\ell_{i+k,t'}\}}\nu x^{\gamma-1}\psi_{[i,i+k)}\sum_{z=\ell_{i+k-1,t'}}^{x-1} z^{-1}, \end{aligned}$$

where the lower summation bounds in the second sum on both rows changed as a result of the indicator in (3.4.26). Approximating the sums by integrals yields that there exists $c = c(\nu,\gamma)$ such that

$$\begin{aligned} f^{(t,t')}_{[i,i+k+1)}(t',x) \leqslant x^{-\gamma}c\left(\phi_{[i,i+k)}\log(t'/\ell_{i+k-1,t'}) + \psi_{[i,i+k)}t'^{2\gamma-1}\right) \\ + \mathbb{1}_{\{x>\ell_{i+k,t'}\}}x^{\gamma-1}c\Big(\phi_{[i,i+k)}\ell^{1-2\gamma}_{i+k,t'} \\ + \psi_{[i,i+k)}\log(t'/\ell_{i+k-1,t'})\Big) \end{aligned}$$

and (3.4.26) holds by the definitions (3.4.24) and (3.4.25), as shifting the index of the terms in the logarithm is allowed because $k \mapsto \ell_{k,t'}$ is non-increasing. The bound (3.4.27) follows analogously. The first indicator follows since the sum on the rhs in the analogue of (3.A.1) is equal to zero if $x = t'$, since $p(t',t') = 0$ by definition in Proposition 2.4.4, i.e.,

$$\sum_{z=\ell_{i+k,t'}\vee t'}^{t'} f^{(t,t')}_{[0,k)}(q,z)p(z,t') = f^{(t,t')}_{[0,k)}(q,t')p(t',t') = f^{(t,t')}_{[0,k)}(q,t')\cdot 0 = 0. \qquad \square$$

Lemma 3.A.1 (Upper bound on $t'/\ell^{(t)}_{k,t'}$). *There exists a constant* $B_{3.A.1} = B_{3.A.1}(\gamma,\nu)$ *such that for* $B > B_{3.A.1}$,

$$t'/\ell^{(t)}_{k,t'} \leqslant \exp\Big(B\big(1\vee\log(t'/t)\big)(\tau-2)^{-k/2}\Big).$$

Proof. Let $\gamma = 1/(\tau-1)$ so that $1/(\tau-2) = \gamma/(1-\gamma)$. We omit the superscript (t) of $\ell^{(t)}_{k,t'}$. We prove by induction. For the induction base $k = 0$,

$t'/\ell_{0,t'} \leqslant t'/(\delta' t)$. The advancement of the induction follows from [45, Lemma A.5, after (A.29)], which contains the appendices of [46]. Recall $\alpha_{[0,k)}$, $\beta_{[0,k)}$, $\ell_{k,t'}$ from (3.4.22), (3.4.23), and (3.4.34b), respectively. Write $\eta_{k,t'} := t'/\ell_{k,t'}$, and let $c = c(\gamma,\nu)$ be the constant from Lemma 3.4.10. To use the same calculations as [45, Lemma A.5, after (A.29)], we need to show that there exists $C = C(\gamma,\nu) > 1$, such that

$$\Big(\eta_{k+2,t'}^{-1} + 1/t'\Big)^{\gamma-1} \leqslant C\left(\eta_{k,t'}^{\gamma} + \eta_{k+1,t'}^{1-\gamma}\log(\eta_{k+1,t'})\right). \tag{3.A.2}$$

We start bounding the lhs. Observe that $(\circledast)$ in (3.4.34b) holds by definition of the arg max in the opposite direction when we replace $\ell_{k,t'}$ by $\ell_{k,t'}+1$, i.e.,

$$\alpha_{[0,k)}(\ell_{k,t'}^{(t)}+1)^{1-\gamma} \geqslant \big(k\log(t')\big)^{-3}.$$

Combining this with (3.4.22) yields

$$\begin{aligned}\left(\frac{\ell_{k+2,t'}+1}{t'}\right)^{\gamma-1} &\leqslant \log^3(t')(k+2)^3\alpha_{[0,k+2)}t'^{1-\gamma}\\ &= c\log^3(t')(k+2)^3\alpha_{[0,k+1)}t'^{1-\gamma}\log(\eta_{k+1,t'})\\ &\qquad + c\log^3(t')(k+2)^3\beta_{[0,k+1)}t'^{\gamma}\\ &=: T_1 + T_2. \end{aligned}\tag{3.A.3}$$

Substituting $(\circledast)$ from (3.4.34b) in T_1 yields

$$\begin{aligned}T_1 &:= c\log^3(t')(k+2)^3\alpha_{[0,k+1)}t'^{1-\gamma}\log(\eta_{k+1,t'})\\ &\leqslant c(k+2)^3\frac{1}{(k+1)^3}\ell_{k+1,t'}^{\gamma-1}t'^{1-\gamma}\log(\eta_{k+1,t'})\\ &\leqslant c\left(\frac{k+2}{k+1}\right)^3\eta_{k+1,t'}^{1-\gamma}\log(\eta_{k+1,t'}). \end{aligned}\tag{3.A.4}$$

Hence, T_1 is bounded by the second term on the rhs in (3.A.2) for C sufficiently large. For T_2, we substitute (3.4.23) and (3.4.34b) to get

$$\begin{aligned}T_2 &:= c\log^3(t')(k+2)^3\beta_{[0,k+1)}t'^{\gamma}\\ &= c^2\log^3(t')(k+2)^3\left(t'^{\gamma}\alpha_{[0,k)}\ell_{k,t'}^{1-2\gamma} + t'^{\gamma}\beta_{[0,k)}\log(\eta_{k,t'})\right)\\ &\leqslant c^2\left(\frac{k+2}{k}\right)^3\eta_{k,t'}^{\gamma} + c^2\log^3(t')(k+2)^3t'^{\gamma}\beta_{[0,k)}\log(\eta_{k,t'})\\ &=: cT_{21} + cT_{22}.\end{aligned}$$

The term cT_{21} can be captured by the first term on the rhs in (3.A.2). For T_{22}, by (3.4.22), $\beta_{[0,k)} t'^{\gamma} \leqslant c^{-1}\alpha_{[0,k+1)} t'^{1-\gamma}$. Using (3.4.34b) and that $\eta_{k,t'}$ is nonincreasing, we further bound

$$\begin{aligned} T_{22} &\leqslant \log^3(t')(k+2)^3 \log(\eta_{k,t'})\alpha_{[0,k+1)} t'^{1-\gamma} \\ &\leqslant (k+2)^3 \log(\eta_{k,t'}) \frac{1}{(k+1)^3} \ell_{k+1,t'}^{\gamma-1} t'^{1-\gamma} \\ &\leqslant \left(\frac{k+2}{k+1}\right)^3 \log(\eta_{k+1,t'})\eta_{k+1,t'}^{1-\gamma}. \end{aligned} \tag{3.A.5}$$

Thus, cT_{22} can be captured by the second term on the rhs in (3.A.2) for C sufficiently large. The desired bound (3.A.2) follows by combining (3.A.3), (3.A.4), and (3.A.5) by increasing the constant C in (3.A.2). Now that we have established (3.A.2), the proof is finished by step-by-step following the computations in [45, Lemma A.5, after (A.29)]. □

Recall $\mathcal{E}_{\mathrm{bad}}^{(q)}$ from (3.4.4), and define $\mathcal{E}_{\mathrm{good}}^{(q)} := \bigcap_{k=0}^{\underline{K}_{t,t'}} \neg\mathcal{E}_{\mathrm{bad}}^{(q)}(k,t')$. We define the conditional probability measure

$$\mathbb{P}_{q-\mathrm{good}}(\,\cdot\,) := \mathbb{P}(\,\cdot \mid \mathcal{E}_{\mathrm{good}}^{(q)}),$$

and write $\mathbb{E}_{q-\mathrm{good}}$ for its corresponding expectation.

Lemma 3.A.2 (Upper bound on neighbourhood size). *Consider the preferential attachment model under the same assumptions as Proposition 3.4.12. Let* q *be a typical vertex in* PA_t. *Then for* $t' \geqslant t$

$$\mathbb{P}_{q-\mathrm{good}}\Bigg(\bigcup_{k=1}^{\underline{K}_{t,t'}+2} \big\{|\partial\mathcal{B}_G^{(t')}(q,k)| \geqslant \exp\big(2B(1\vee\log(\tfrac{t'}{t}))(\tau-2)^{-k/2}\big)\big\}\Bigg) \leqslant 2\exp\big(-B(1\vee\log(\tfrac{t'}{t}))\big). \tag{3.A.6}$$

Proof. We first bound $\mathbb{E}_{q-\mathrm{good}}\big[|\partial\mathcal{B}_G^{(t')}(q,k)|\big]$, and let the result follow by a union bound on the events in (3.A.6) and Markov's inequality. Conditionally on $\mathcal{E}_{\mathrm{good}}^{(q)}$, all vertices at distance $k < \underline{K}_{t,t'}$ can only be reached via t'-good q-paths. Recall $f_{[0,k)}^{(t,t')}(q,x)$ from (3.4.17) and its interpretation as an upper bound for the expected number of good paths from q to x of length k. Thus, we have by the law of total probability, and the definition of good paths in Definition 3.4.3,

$$\mathbb{E}_{q-\mathrm{good}}\big[|\partial\mathcal{B}_G^{(t')}(q,k)|\big] \leqslant \frac{1}{\mathbb{P}(\mathcal{E}_{\mathrm{good}}^{(q)})} \sum_{x=\ell_{k,t'}}^{t'} f_{[0,k)}^{(t,t')}(q,x).$$

Recalling (3.4.36), we see that it is sufficient to bound the sum on the rhs. Now, applying the bound in (3.4.27) on f yields for some $c_1, c_2 > 0$,

$$\begin{aligned}\sum_{x=\ell_{k,t'}}^{t'} f_{[0,k)}^{(t,t')}(q,x) &\leqslant \alpha_{[0,k)} \sum_{x=\ell_{k,t'}}^{t'} x^{-\gamma} + \beta_{[0,k)} \sum_{x=\ell_{k-1,t'}}^{t'} x^{\gamma-1} \\ &\leqslant c_1 \left(\alpha_{[0,k)} t'^{1-\gamma} + \beta_{[0,k)} t'^{\gamma}\right) \\ &\leqslant c_2 \alpha_{[0,k+1)} t'^{1-\gamma}. \end{aligned} \tag{3.A.7}$$

The last line follows since by (3.4.22), $t'^{\gamma} \beta_{[0,k)} \leqslant \alpha_{[0,k+1)} t'^{\gamma-1}/c$, and $k \mapsto \alpha_{[0,k)}$ is non-decreasing. We bound the rhs in (3.A.7) in terms of $(t'/\ell_{k,t'})$. By $(\circledast)$ in (3.4.34b) and Lemma 3.A.1

$$\begin{aligned}\alpha_{[0,k+1)} t'^{1-\gamma} &\leqslant ((k+1)\log(t'))^{-3} (t'/\ell_{k+1,t'})^{1-\gamma} \\ &\leqslant ((k+1)\log(t'))^{-3} \\ &\quad \cdot \exp\Big(B(1-\gamma)(1 \vee \log(t'/t))(\tau-2)^{-\frac{k+1}{2}}\Big).\end{aligned}$$

This yields that for B sufficiently large

$$\begin{aligned}\mathbb{E}_{q-\text{good}}\left[|\partial \mathcal{B}_G^{(t')}(q,k)|\right] &\leqslant \frac{2c}{\mathbb{P}(\mathcal{E}_{\text{good}}^{(q)})} \exp\Big(B(1 \vee \log(t'/t))(\tau-2)^{-k/2}\Big) \\ &\leqslant \exp\Big(B'(1 \vee \log(t'/t))(\tau-2)^{-k/2}\Big),\end{aligned}$$

where the last bound follows for some $B' > B$ as

$$\mathbb{P}(\mathcal{E}_{\text{good}}^{(q)}) = 1 - \delta' + o(1)$$

by (3.4.36). The assertion (3.A.6) follows by a union bound over (3.A.6) and summing over k. □

3.A.2 *Upper bound*

Claim 3.A.3. *Recall* K_{t,t_i} *from* (3.1.1), $I_2(L)$ *from* (2.1.2), *and* Q_{t,t_i} *from* (3.3.2). *Let* L_1 *and* L_2 *be two independent copies of the random variable* L. *For all* $c_5, \varepsilon, M > 0$, *there exists* $M_L > 0$ *such that*

$$\begin{aligned}\sum_{k=0}^{K_{t,t_i}/2+M-1} F_{L_1+L_2}^{(-1)}\Big(\exp\left(-2c_5(k+1)^{-2}(\tau-2)^{-(i+k)}\right)\Big) \\ \leqslant (1+\varepsilon\mathbb{1}_{\{I_2(L)=\infty\}}) Q_{t,t_i} + M_L.\end{aligned} \tag{3.A.8}$$

Proof. We proceed along the same lines as in the proof of Proposition 2.3.4. To shorten notation, we define

$$\widetilde{Q}_{t,t_i} := \sum_{k=0}^{K_{t,t_i}/2+M-1} F_{L_1+L_2}^{(-1)}\Big(\exp\big(-2c_5(k+1)^{-2}(\tau-2)^{-(i+k)}\big)\Big)$$

First, we relate the inverse $F_{L_1+L_2}^{(-1)}(\cdot)$ to $F_L^{(-1)}(\cdot)$ by observing that for $x>0$

$$F_{L_1+L_2}(x) = \mathbb{P}(L_1+L_2 \leqslant x) \geqslant \mathbb{P}(\max\{L_1,L_2\} \leqslant x/2) = \big(F_L(x/2)\big)^2.$$

Hence, for any $z>0$, it holds that $F_{L_1+L_2}(z) \leqslant 2F_L^{(-1)}(\sqrt{z})$, so that for some $M_1>0$

$$\begin{aligned}\widetilde{Q}_{t,t_i} &\leqslant 2\sum_{k=0}^{K_{t,t_i}/2+M-1} F_L^{(-1)}\Big(\exp\big(-c_5(k+1)^{-2}(\tau-2)^{-(i+k)}\big)\Big)\\ &\leqslant M_1 + 2\sum_{k=0}^{K_{t,t_i}/2} F_L^{(-1)}\Big(\exp\big(-c_5(k+1)^{-2}(\tau-2)^{-(i+k)}\big)\Big),\end{aligned}$$

since $z\mapsto F_L^{(-1)}(1/z)$ is nonincreasing and bounded. Define $b := \inf\{x : F_L^{(-1)}(x)>0\}$ and $L' := L-b$, so that by $(\star)$ in (3.4.49)

$$\begin{aligned}\widetilde{Q}_{t,t_i} &\leqslant M_1 + bK_{t,t_i}\\ &\quad + 2\int_{x=0}^{\frac{K_{t,t_i}}{2}} F_{L'}^{(-1)}\big(\exp\big(-c_5(x+1)^{-2}(\tau-2)^{-i-x}\big)\big)dx\\ &\leqslant M_2 + bK_{t,t_i}\\ &\quad + 2\int_{x=x_0}^{\frac{K_{t,t_i}}{2}} F_{L'}^{(-1)}\big(\exp\big(-c_5(x+1)^{-2}(\tau-2)^{-i-x}\big)\big)dx \qquad (3.A.9)\end{aligned}$$

for some large constants $x_0, M_2>0$. Apply the change of variables

$$(k+1)^{-2}(\tau-2)^{-x} = (\tau-2)^{-y/2} \quad\Leftrightarrow\quad y = 2x - 4\frac{\log(x+1)}{|\log(\tau-2)|}.$$

Differentiating both sides, rearranging terms, and using $y=\Theta(x)$ yields a bound for dx, i.e.,

$$\Big(1-\frac{2/|\log(\tau-2)|}{x+1}\Big)dx = \frac{1}{2}dy \qquad\Rightarrow\qquad dx \leqslant \frac{1}{2}\Big(1+\frac{C}{y+1}\Big)dy,$$

for some constant $C > 0$ and $x \geqslant x_0$ sufficiently large. Continuing to bound (3.A.9) from above, we obtain if x_0 is sufficiently large

$$\begin{aligned}\widetilde{Q}_{t,t_i} &\leqslant M_2 + bK_{t,t_i}\\ &\quad+\int_{y=2x_0-4\frac{\log(x_0+1)}{|\log(\tau-2)|}}^{K_{t,t_i}-4\frac{\log(1+K_{t,t_i})}{|\log(\tau-2)|}}\Big(1+\frac{C}{y+1}\Big)F_{L'}^{(-1)}\Big(\exp\Big(-c_5(\tau-2)^{-i-\frac{y}{2}}\Big)\Big)dy\\ &\leqslant M_2 + bK_{t,t_i} \qquad (3.A.10)\\ &\quad+\int_{y=x_0}^{K_{t,t_i}}\Big(1+\frac{C}{y+1}\Big)F_{L'}^{(-1)}\big(\exp\big(-c_5(\tau-2)^{-i-y/2}\big)\big)dy.\end{aligned}$$

Recall $I_2(L')$ from (2.1.2). We first assume $I_2(L') < \infty$. In this case, there exists $M_3 > 0$ such that

$$\int_{y=x_0}^{\infty}\frac{C}{y+1}F_{L'}^{(-1)}\big(\exp\big(-(\tau-2)^{-\frac{y}{2}}\big)\big)dy < M_3.$$

Using that the integrand in (3.A.10) is bounded, we obtain for some $M_4 > 0$

$$\begin{aligned}\widetilde{Q}_{t,t_i} &\leqslant M_2 + M_3 + bK_{t,t_i} + \int_{y=x_0}^{K_{t,t_i}}F_{L'}^{(-1)}\big(\exp\big(-c_5(\tau-2)^{-i-y/2}\big)\big)dy\\ &\leqslant M_4 + bK_{t,t_i} + \int_{y=0}^{K_{t,t_i}}F_{L'}^{(-1)}\big(\exp\big(-c_5(\tau-2)^{-i-y/2}\big)\big)dy. \quad (3.A.11)\end{aligned}$$

Since $L' = L - b$ by definition, and using that the integration interval has length K_{t,t_i}, we obtain

$$\begin{aligned}\widetilde{Q}_{t,t_i} &\leqslant M_4 + \int_{y=0}^{K_{t,t_i}}F_{L}^{(-1)}\big(\exp\big(-c_5(\tau-2)^{-(y+2i)/2}\big)\big)dy\\ &= M_4 + \int_{y=2i}^{K_{t,t_i}+2i}F_{L}^{(-1)}\big(\exp\big(-c_5(\tau-2)^{-y/2}\big)\big)dy\end{aligned}$$

by shifting the integration boundaries. Recall $K_{t,t} - K_{t,t_i} = 2i$ by (3.5.15), yielding

$$\widetilde{Q}_{t,t_i} \leqslant M_4 + \int_{y=K_{t,t}-K_{t,t_i}}^{K_{t,t}}F_{L}^{(-1)}\big(\exp\big(-c_5(\tau-2)^{-y/2}\big)\big)dy.$$

We leave it to the reader to verify using another change of variables and $(*)$ in (3.4.49) that, similarly to the proof for the lower bound after (3.4.49), there exists M_5 such that

$$\int_{y=K_{t,t}-K_{t,t_i}}^{K_{t,t}} F_L^{(-1)}\big(\exp\big(-c_5(\tau-2)^{-y/2}\big)\big)dy \\ - \sum_{k=K_{t,t}-K_{t,t_i}+1}^{K_{t,t}} F_L^{(-1)}\big(\exp\big(-(\tau-2)^{-k/2}\big)\big)dy \leqslant M_5.$$

This establishes (3.A.8) when $I_2(L) < \infty$. If $I_2(L) = \infty$, we observe that there exists M_6 such that

$$\int_{y=x_0}^{K_{t,t_i}} \frac{C}{y+1} F_{L'}^{(-1)}\big(\exp\big(-c_5(\tau-2)^{-y/2}\big)\big)dy \\ < M_6 + \varepsilon \int_{y=x_0}^{K_{t,t_i}} F_{L'}^{(-1)}\big(\exp\big(-c_5(\tau-2)^{-y/2}\big)\big)dy.$$

We use this bound in (3.A.10), bound $b \leqslant (1+\varepsilon)b$, and follow the same steps as from (3.A.11) onwards, carrying a factor $(1+\varepsilon)$ for the integrals. □

Part II
Kernel-based spatial random graphs: Component sizes

4

CLUSTER-SIZE DECAY ON KSRGS

4.1 INTRODUCTION

In this part we study component sizes in i-KSRGs, introduced in Section 1.3. For nearest-neighbour Bernoulli percolation on $\mathbb{Z}^d$ [42], it is a result of a sequence of works [5, 9, 47, 110, 158, 171] that, considering a supercritical model, i.e., $p > p_c(\mathbb{Z}^d)$ – the critical percolation probability on $\mathbb{Z}^d$– and $|\mathcal{C}(0)|$ being the number of vertices in the cluster containing the origin,

$$\mathbb{P}(k \leqslant |\mathcal{C}(0)| < \infty) = \exp\big(-\Theta(k^{\zeta})\big), \tag{4.1.1}$$

with $\zeta = (d-1)/d$. Intuitively, the stretched exponential decay with exponent $(d-1)/d$ emanates from the fact that all the $\Omega(k^{(d-1)/d})$ edges on the boundary of a cluster $\mathcal{C}$ with $|\mathcal{C}| \geqslant k$ need to be absent: the tail decay in (4.1.1) is determined by *surface tension*. More recently, these results have been extended to Bernoulli percolation on general classes of transitive graphs [56, 139].

In the next chapters and the accompanying paper [151], we consider $\mathbb{P}(k \leqslant |\mathcal{C}(0)| < \infty)$ for i-KSRGs for which the degree distribution and/or the edge-length distribution are heavy-tailed, and identify when this *structural inhomogeneity changes the surface-tension driven behaviour of finite clusters*. We will see that in those cases, the cluster-size decay in (4.1.1) is still stretched exponential, but the exponent ζ changes. Its value (together with our proof techniques) reflects the structure of the infinite/giant component in spatial graph models: it describes the size and structure of a "backbone" decorated with "traps" – almost isolated peninsulas attached to a well-connected component. This topological description is of independent interest, e.g., it affects the behaviour of random walk [59].

Our starting points are continuum and classical *long-range percolation* (CLRP and LRP) [4, 216], examples of models where the edge-length distribution has a heavy tail, while the degree distribution is light-tailed. Relating to the distribution of small clusters in supercritical settings as in (4.1.1), the autors of [58] gave a polylogarithmic upper bound on the size of the second-largest component $|\mathcal{C}_n^{(2)}|$ of long-range percolation in a box of volume n, with unidentified exponent, which is the only known result in this direction. In long-range percolation, each potential edge $\{u,v\} \in \mathbb{Z}^d \times \mathbb{Z}^d$ is independently present with probability proportional to $\beta^\alpha \|u-v\|^{-d\alpha}$ for some $\alpha > 1$, $\beta > 0$. Our combined results (Chapters 5–6, and [151]) show that, under the additional assumption that β is sufficiently large if $(d-1)/d > 2-\alpha$,

$$\mathbb{P}(k \leqslant |\mathcal{C}(0)| < \infty) = \exp\big(-k^{\max\{2-\alpha+o(1),(d-1)/d\}}\big),$$
$$|\mathcal{C}_n^{(2)}| = (\log n)^{1/\max\{2-\alpha+o(1),(d-1)/d\}},$$

showing stretched-exponential decay similar to (4.1.1), up to a $o(1)$-error term in the exponent when $2-\alpha > (d-1)/d$.

Our main focus is to prove (4.1.1) for models where *both the degree- and the edge-length distribution are heavy-tailed*, such as scale-free percolation (SFP) [63], continuum scale-free percolation (CSFP) [67]; geometric inhomogeneous random graphs (GIRG) [40], hyperbolic random graphs (HRG) [169], the ultra-small scale-free geometric network [228]; scale-free Gilbert model (SGM) [122], the Poisson Boolean model with random radii [103], the age- and the weight-dependent random connection models (ARCM) [105, 107]. Concerning sizes of small clusters in these models, the only known result concerns the asymptotic size of the second-largest component of HRGs [162]. Among the more classical models, for random geometric graphs and Bernoulli percolation on them, the papers [175, 211] identify $|\mathcal{C}_n^{(2)}|$.

In Chapter 5 we prove the stretched exponential decay and identify the values of ζ in (4.1.1) for (regions of the parameter space of) supercritical continuum scale-free percolation, geometric inhomogeneous random graphs, and hyperbolic random graphs (CSFP, GIRG, HRG). We introduce a novel proof technique, the *cover expansion*, and give a quantitative backbone-and-traps description of the giant component; which is novel for these models. Additionally, we unfold the general relation between the cluster-size decay and the size of the second-largest component, which yields a weak law of large numbers for the size of the giant component as a side result. In [151], we establish (4.1.1) for the scale-free Gilbert model, for the age-dependent random connection model, for continuum long-range percolation, and other regions of the parameter space of CSFP, GIRG, and HRG. The results here and in the accompanying paper [151] together also determine the *speed of the lower tail of large deviations* for the size of the giant component $|\mathcal{C}_n^{(1)}|$ for all these models. With the *same* ζ as the one in the cluster-size decay, for all sufficiently small $\rho > 0$,

$$\mathbb{P}\big(|\mathcal{C}_n^{(1)}| < \rho n\big) = \exp\big(-n^{\zeta+o(1)}\big). \tag{4.1.2}$$

In Chapter 5 we present the lower bound and in [151] the upper bound.

4.2 FOUR REGIMES OF ζ IN THE PHASE DIAGRAM

In KSRGs in general, one can determine the value of the exponent ζ of the cluster-size decay via the following back-of-the-envelope calculation, that is based on the lower bound and aims to find a spatially *localized* component of size $\Theta(k)$. Consider two neighboring boxes $\Lambda_k^{(1)}, \Lambda_k^{(2)}$ of volume k in $\mathbb{R}^d$, and compute, as a function of k, the expected number of edges $\mathbb{E}[|\mathcal{E}(\Lambda_k^{(1)}, \Lambda_k^{(2)})|]$ between vertices in these boxes, and also the expected number of vertices in $\Lambda_k^{(1)}$ having *at least one edge* towards a vertex in $\Lambda_k^{(2)}$, $\mathbb{E}[|\mathcal{V}(\Lambda_k^{(1)} \to \Lambda_k^{(2)})|]$. If these two quantities are the same order, and are both $\Theta(k^\zeta)$, then the probability that none of the potential edges is present is $\exp(-\Theta(k^\zeta))$. The non-presence of all these edges is necessary for a large isolated cluster of size $\Theta(k)$ to be confined to its original box. If, however, we find $\mathbb{E}[|\mathcal{V}(\Lambda_k^{(1)} \to \Lambda_k^{(2)})|] = o(\mathbb{E}[|\mathcal{E}(\Lambda_k^{(1)}, \Lambda_k^{(2)})|])$, and the former quantity is $\Theta(k^\zeta)$, then typically there are a few high-mark vertices that have many edges to the neighboring box. The most likely way to have a component of size $\Theta(k)$ is then to simply *not have any of these vertices* present in the two boxes, again happening with probability $\exp(-\Theta(k^\zeta))$. Clearly, large finite clusters can be non-localized, so the rigorous version of this argument only gives lower bounds. Nevertheless, it describes the *most likely way of connectivity* towards and within the infinite/giant component and gives a conjecture for the value of ζ.

Based on this intuition, we distinguish four possible *types of connectivity* that describe how neighboring boxes are connected. We call the type of connectivity that produces the largest contribution to $\mathbb{E}[|\mathcal{V}(\Lambda_k^{(1)} \to \Lambda_k^{(2)})|]$ the *dominant* type. Changes of the dominant type give the phase diagram of ζ in the parameter space. We visualize these phases in Figure 6a for models using κ_{prod}, κ_{pa}, or $\kappa_{\max}$. All arguments below are meant as $k \to \infty$. We say that the model shows dominantly

(ll) *low-low*-type connectivity if the main contribution to $\mathbb{E}[|\mathcal{V}(\Lambda_k^{(1)} \to \Lambda_k^{(1)})|]$ is coming from pairs of *low-mark vertices* (vertices with mark $\Theta(1)$) at distance $\Theta(k^{1/d})$ from each other. We find $\mathbb{E}[|\mathcal{V}(\Lambda_k^{(1)} \to \Lambda_k^{(2)})|] = \Theta(\mathbb{E}[|\mathcal{E}(\Lambda_k^{(1)}, \Lambda_k^{(2)})|)] = \Theta(k^{\zeta_{\mathrm{ll}}})$, where

$$\zeta_{\mathrm{ll}} := 2 - \alpha, \tag{4.2.1}$$

which neither depends on κ nor on τ since low-mark vertices are not affected by those. Models with parameters falling in this regime behave qualitatively similarly to long-range percolation. We present proofs for parameters regimes where ζ_{ll} is dominant in [151]. In KSRGs with kernels in (1.3.4) this regime may occur if both $\alpha < \tau - 1$ and $\alpha < 2$.

(lh) *low-high*-type connectivity if $\mathbb{E}[|\mathcal{E}(\Lambda_k^{(1)}, \Lambda_k^{(2)})|] \gg \mathbb{E}[|\mathcal{V}(\Lambda_k^{(1)} \to \Lambda_k^{(2)})|]$ and the main contributions are coming from edges between a *high-mark* and a low-mark vertex ($\Theta(1)$) at distance $\Theta(k^{1/d})$. A vertex has high-mark here when its mark is $\Omega(k^{\gamma_{lh}})$, where

$$\gamma_{lh} := 1 - 1/\alpha$$

is the *smallest* exponent such that a *linear proportion* of vertices of mark $\Omega(k^{\gamma_{lh}})$ in $\Lambda_k^{(1)}$ is connected to some low-mark vertex in $\Lambda_k^{(2)}$. Vertices of mark $\Omega(k^{\gamma_{lh}})$ have to be absent to find a large but finite component confined to $\Lambda_k^{(1)}$, and since $\mathbb{E}[|\mathcal{V}(\Lambda_k^{(1)} \to \Lambda_k^{(2)})|] = \Theta(k^{1-\gamma_{lh}(\tau-1)})$, we obtain

$$\zeta_{lh} := 1 - \gamma_{lh}(\tau - 1) = 1 - (\tau - 1)(1 - 1/\alpha). \tag{4.2.2}$$

We prove in our accompanying paper [151] that this type of behaviour is dominant in KSRGs with kernels κ_{max}, κ_{sum}, and κ_{pa} (e.g. for the age-dependent random connection model [105] and scale-free Gilbert model [122]) when both $\alpha \in (\tau - 1, (\tau - 1)/(\tau - 2))$, and $\tau \in (2, 3)$ holds.

(hh) *high-high*-type connectivity if $\mathbb{E}[|\mathcal{E}(\Lambda_k^{(1)}, \Lambda_k^{(2)})|] \gg \mathbb{E}[|\mathcal{V}(\Lambda_k^{(1)} \to \Lambda_k^{(2)})|]$, and the main contributions are coming from edges between two *high-mark* vertices at distance $\Theta(k^{1/d})$. A vertex has high mark here when its mark is $\Omega(k^{\gamma_{hh}})$, where γ_{hh} is the smallest exponent such that a *linear proportion* of vertices of mark $\Omega(k^{\gamma_{hh}})$ in $\Lambda_k^{(1)}$ is connected by an edge to a high-mark vertex in $\Lambda_k^{(2)}$. This leads in expectation to $\Theta(k^{\zeta_{hh}})$ many high-mark vertices in $\Lambda_k^{(1)}$ and $\Lambda_k^{(2)}$ (which have to be absent to find a large but finite component), where $\zeta_{hh} := 1 - \gamma_{hh}(\tau - 1)$. The formula of γ_{hh} depends on the choice of the kernel κ. For the interpolating kernel $\kappa_{1,\sigma}$ from (1.3.5) that we use in this chapter we find

$$\gamma_{hh} = \begin{cases} \dfrac{1 - 1/\alpha}{\sigma + 1 - (\tau - 1)/\alpha}, & \text{if } \tau \leqslant 2 + \sigma, \\ \dfrac{1}{\sigma + 1}, & \text{if } \tau > 2 + \sigma, \end{cases} \tag{4.2.3}$$

leading to[1]

$$\zeta_{\mathrm{hh}} = 1 - \gamma_{\mathrm{hh}}(\tau - 1) = \begin{cases} \dfrac{2+\sigma-\tau}{\sigma+1-(\tau-1)/\alpha}, & \text{if } \tau \leqslant 2+\sigma, \\ \dfrac{2+\sigma-\tau}{\sigma+1}, & \text{if } \tau > 2+\sigma. \end{cases} \tag{4.2.4}$$

In Chapter 5 we focus on this phase and show that this type of connectivity can be dominant in KSRGs with kernel κ_{prod}: (continuum) scale-free percolation (SFP) [63, 67], geometric inhomogeneous random graphs (GIRG) [40] and hyperbolic random graphs (HRG) [169]) when $\tau < 3$ and $\alpha \geqslant \tau - 1$.

(nn) *nearest-neighbour*-type connectivity if

$$\mathbb{E}[|\mathcal{E}(\Lambda_k^{(1)}, \Lambda_k^{(2)})|] = \Theta(\mathbb{E}[|\mathcal{V}(\Lambda_k^{(1)} \to \Lambda_k^{(2)})|]),$$

and the main contributions are coming from edges of length $\Theta(1)$ between low-mark vertices close to the shared boundary of the two boxes. The expected number of such edges and vertices is $\Theta(k^{\zeta_{\mathrm{nn}}})$ where

$$\zeta_{\mathrm{nn}} := (d-1)/d. \tag{4.2.5}$$

KSRGs with kernel κ_{triv} and threshold profile show this type of connectivity behaviour in their *entire* parameter space, e.g., random geometric graphs and nearest-neighbour percolation on $\mathbb{Z}^d$. KSRGs with either non-trivial kernel or long-range profile also show this phase in a region of the parameter space, e.g. long-range percolation when $\alpha > 1 + 1/d$. Chapter 6 treats sub-regions of this phase in the (α, τ) parameter-plane.

In general we conjecture that for any model belonging to the KSRG framework, the value of ζ in (4.1.1) is determined by whichever connectivity type yields the largest contribution to $\mathbb{E}[|\mathcal{E}(\Lambda_k^{(1)}, \Lambda_k^{(2)})|]$ and $\mathbb{E}[|\mathcal{V}(\Lambda_k^{(1)} \to \Lambda_k^{(2)})|]$. Formally we may only compute

$$\max\{\zeta_{\mathrm{ll}}, \zeta_{\mathrm{lh}}, \zeta_{\mathrm{hh}}, \zeta_{\mathrm{nn}}\} =: \zeta,$$

yielding the conjectured exponent in (4.1.1). See also the overview in Table 2, and Figure 6a.

1 We will never use the formula for ζ_{hh} when it is negative: then, some other connectivity type is dominant.

Model	**Kernel**	ζ
Bond-percolation on $\mathbb{Z}^d$ [113]	$\kappa_{\mathrm{triv}}, \kappa_{0,0}$	ζ_{nn}
Random geometric graph [210]		ζ_{nn}
Long-range percolation [216]		$\max\{\zeta_{\mathrm{ll}}, \zeta_{\mathrm{nn}}\}$
Continuum long-range percolation [4]		$\max\{\zeta_{\mathrm{ll}}, \zeta_{\mathrm{nn}}\}$
Scale-free percolation [63]	$\kappa_{\mathrm{prod}}, \kappa_{1,1}$	$\max\{\zeta_{\mathrm{hh}}, \zeta_{\mathrm{ll}}, \zeta_{\mathrm{nn}}\}$
Continuum scale-free percolation [67]		$\max\{\zeta_{\mathrm{hh}}, \zeta_{\mathrm{ll}}, \zeta_{\mathrm{nn}}\}$
Geometric inhomogeneous random graph [40]		$\max\{\zeta_{\mathrm{hh}}, \zeta_{\mathrm{ll}}, \zeta_{\mathrm{nn}}\}$
Hyperbolic random graph [169]		ζ_{hh}
Age-dependent random connection model [105]	$\kappa_{\mathrm{pa}}, \kappa_{1,\tau-2}$	$\max\{\zeta_{\mathrm{lh}}, \zeta_{\mathrm{ll}}, \zeta_{\mathrm{nn}}\}$
Scale-free Gilbert graph [122]	$\kappa_{\mathrm{max}}, \kappa_{1,0}$	$\max\{\zeta_{\mathrm{lh}}, \zeta_{\mathrm{ll}}, \zeta_{\mathrm{nn}}\}$
Ultra-small scale-free geometric network [228]	$\kappa_{\mathrm{min}}, \kappa_{0,1}$	$\max\{\zeta_{\mathrm{hh}}, \zeta_{\mathrm{nn}}\}$
Interpolating KSRG	$\kappa_{1,\sigma}$	$\max\{\zeta_{\mathrm{ll}}, \zeta_{\mathrm{lh}}, \zeta_{\mathrm{hh}}, \zeta_{\mathrm{nn}}\}$

Table 2: Models belonging to the KSRG framework, their kernels, and the (conjectured) value of their cluster-size decay exponent ζ. Horizontal lines separate models with different kernels. See Table 1 for the vertex sets and profiles that are used by these models.

The interpolation kernel allows for unified proof techniques, and shows how the possible dominant regimes (ll, lh, hh, and nn) fade into each other as the interpolation parameter σ changes gradually, see Figure 6b. We then state our conjecture, visualized in Figure 6, see also Table 2.

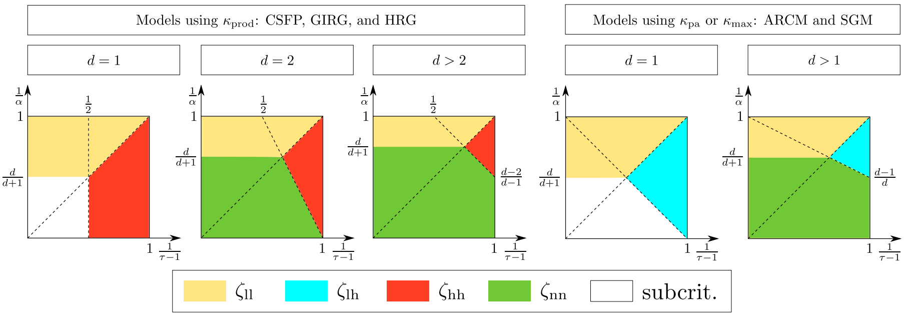
Conjecture 4.2.1. *We conjecture that for supercritical KSRGs in* $\mathbb{R}^d$ *with kernel* $\kappa_{1,\sigma}$*, for all* $(\alpha, \tau) \in (1, \infty] \times (2, \infty]$, (4.1.1) *and* (4.1.2) *hold with*

$$\zeta = \max\{\zeta_{\mathrm{ll}}, \zeta_{\mathrm{lh}}, \zeta_{\mathrm{hh}}, \zeta_{\mathrm{nn}}\}$$

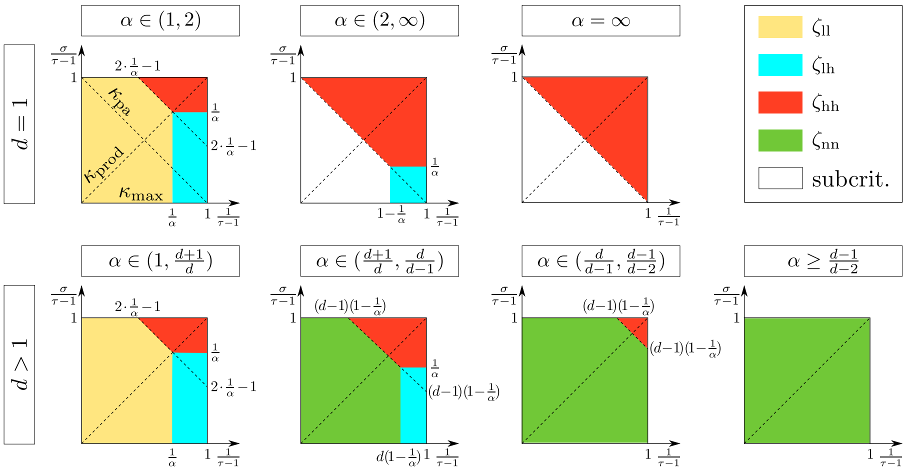
whenever $\zeta > 0$*. When* $\alpha = \infty$ *and/or* $\tau = \infty$*, the value of* ζ *is obtained by taking the corresponding limit of* $\zeta = \zeta(\alpha, \tau)$*.*

4.2.1 *Three related works*

The results in this chapter, combined with [151], prove Conjecture 4.2.1 for i-KSRGs on Poisson point processes (PPP) whenever $\max\{\zeta_{\mathrm{ll}}, \zeta_{\mathrm{lh}}, \zeta_{\mathrm{hh}}\} \geqslant \zeta_{\mathrm{nn}} = \frac{d-1}{d}$ and $\sigma \leqslant \tau - 1$ (the non-green regions in Figure 6). When $\max\{\zeta_{\mathrm{ll}}, \zeta_{\mathrm{lh}}\} > \zeta_{\mathrm{hh}}$ (the blue and yellow regions in Figure 6), or when the maximum is



(a) Phase-diagrams of $\zeta = \zeta(\tau, \alpha)$ for models with kernels κ_{prod}, and κ_{pa} or $\kappa_{\max}$, plotted as a function of $1/(\tau-1)$ and $1/\alpha$. The y-axis (i.e., $1/(\tau-1) = 0$) also describes the phase diagram of (continuum) long-range percolation that has kernel κ_{triv}, while the models on the x-axis ($1/\alpha = 0$) coincide with models using a threshold profile function in (1.3.3). When $1/\alpha > 1$ or $1/(\tau-1) > 1$, then $\mathcal{G}_\infty$ is connected and each vertex has infinite degree almost surely [121]. A white color within the square means that the model is subcritical for each value p, β in (1.3.1) [108].



(b) Phase-diagrams of $\zeta = \zeta(\sigma, \tau)$ for fixed values of α in (1.3.3), plotted as a function of $1/(\tau-1)$ on the x-axis and $\sigma/(\tau-1)$ on the y-axis. The identity line $y = x$ corresponds to models using kernel $\kappa_{\mathrm{prod}} \equiv \kappa_{1,1}$, the x-axis to models using $\kappa_{\max} \equiv \kappa_{1,0}$ and the cross-diagonal $x+y = 1$ to models using $\kappa_{\mathrm{pa}} \equiv \kappa_{1,\tau-2}$. The origin captures models with $\kappa_{\mathrm{triv}} \equiv \kappa_{0,0}$. Observe that ζ_{lh} (blue) is never dominant above the diagonal $y \geqslant x$ (equivalently, $\sigma \geqslant 1$), while ζ_{hh} (red) is never dominant below the cross-diagonal $x+y = 1$ (equivalently, $\sigma \leqslant \tau-2$). In the quadrant $x+y \geqslant 1, y \leqslant x$ all four exponents 'compete' for dominance. Each of the diagrams (except $\alpha \geqslant (d-1)/(d-2)$) can be extended to $y = \sigma/(\tau-1) > 1$, by simply extending the linear boundary lines separating the regions; when $\alpha \geqslant (d-1)/(d-2)$, the line separating the nn- from the hh-regime is given by $x+y = 1+(d-1)(1-1/\alpha)$.

Figure 6: Phase diagrams of the (conjectured) cluster-size decay for kernel-based spatial random graphs. Theorem 5.1.2 proves the upper bound in the red regions, and it gives the lower bounds above the $x+y \geqslant 1$ line on Figure 6b, for all four colors simultaneously, with logarithmic correction terms on phase boundary lines.

non-unique (on region boundaries), we have a $o(1)$ error in the exponent ζ, while in the open red region our bound is sharp. Regarding the speed of the lower tail of large deviations (LTLD) for the size of the giant component (as in (4.1.2)), we prove Conjecture 4.2.1 up to a $o(1)$ error whenever $\max\{\zeta_{ll}, \zeta_{lh}, \zeta_{hh}\} > 0$ and $\sigma \leqslant \tau - 1$ (all parameters for which the model can be supercritical in $d = 1$). Thus, up to $o(1)$-errors in some regions, these results prove Conjecture 4.2.1 fully for locally finite i-KSRGs on PPPs when $d = 1$ and $\sigma \leqslant \tau - 1$. In dimension $d \geqslant 2$, the present chapter and [151] leave open (only) the regime where ζ_{nn} is maximal, but we still obtain partial results in this regime (see the following paragraphs). In Chapter 6, we prove Conjecture 4.2.1 fully for *long-range percolation* on $\mathbb{Z}^d$ whenever the edge density is sufficiently high, i.e., also when ζ_{nn} is maximal.

Upper bounds. In this chapter we prove the upper bound on the cluster-size decay and $|\mathcal{C}_n^{(2)}|$ for $\zeta = \zeta_{hh}$, which is sharp in the red regions in Figure 6, and still non-trivial whenever $\tau < 2 + \sigma$ (above the cross-diagonal in Figure 6b), using a "backbone construction". In [151] we show *upper bounds* on the speed of LTLD for the size of the giant component using a renormalization scheme driven by vertex-marks. This yields an upper bound that matches the conjectured exponent ζ in (4.1.2) up to $o(1)$ when $\max\{\zeta_{ll}, \zeta_{lh}, \zeta_{hh}\} > 0$. In [151] we use this LTLD-type result to construct a backbone (whose construction is different from the construction in this chapter). Using that, we prove the upper bounds on the cluster-size decay and $|\mathcal{C}_n^{(2)}|$ when $\zeta \neq \zeta_{nn}$, obtaining up to $o(1)$-matching upper bounds for $\zeta \in \{\zeta_{ll}, \zeta_{lh}\}$ in (4.1.1) (blue and yellow regions in Figure 6). Only the hh-regime requires $\sigma \leqslant \tau - 1$. All upper bounds for $\zeta = \zeta_{ll}$ extend to i-KSRGs where the vertex locations are given by $\mathbb{Z}^d$. In Chapter 6 we show the upper bounds for long-range percolation on $\mathbb{Z}^d$ when $\zeta_{nn} > \zeta_{ll}$, assuming sufficiently high edge-density. We use combinatorial methods in Chapter 6, since the backbone construction fails for LRP when ζ_{nn} is dominant.

Lower bounds. This chapter proves, for the whole parameter-space, the lower bound for the speed of LTLD for the size of the giant component in (4.1.2). LTLD is related to the lower bound on the cluster-size decay and $|\mathcal{C}_n^{(2)}|$, once we additionally prove that a box with appropriately restricted vertex-marks still contains a linear-sized component. Here, we prove this when $\tau < 2 + \sigma$ using the backbone construction; the (different) backbone construction in [151] yields it for $\tau \geqslant 2 + \sigma$. In Chapter 6, the high edge-density assumption ensures the existence of a giant in a box. The lower bounds do not require $\sigma \leqslant \tau - 1$, extend to i-KSRGs on $\mathbb{Z}^d$, and do not contain $o(1)$-error terms.

5 THE HIGH-HIGH REGIME

Based on [149]:
Cluster-size decay in supercritical kernel-based spatial random graphs,
J. Jorritsma, J. Komjáthy, D. Mitsche,
Preprint arXiv:2303.00712, 2023.

The next section makes definitions and our findings formal for our results in the hh-regime.

5.1 MODEL AND MAIN RESULTS

We start by defining kernel-based random graphs, similar to the definition in [107], but specifically for the interpolating kernel $\kappa_{1,\sigma}$ in (1.3.5). We denote by $\|x-y\|$ the Euclidean distance between $x, y \in \mathbb{R}^d$, and for $x \in \mathbb{R}^d, s > 0$, we denote a box of volume s in $\mathbb{R}^d$ centered at x by

$$\Lambda_s(x) = \Lambda(x, s) := x + [-s^{1/d}/2, s^{1/d}/2]^d \tag{5.1.1}$$

If $x = 0$, we write Λ_s. For a discrete set V we write $\binom{V}{2} := \{\{u, v\}, u \in V, v \in V, u \neq v\}$ for the list of *unordered* pairs in V.

Definition 5.1.1 (Interpolating kernel-based spatial random graph, **i-K-SRG**). Fix $\alpha > 1, \tau > 1, d \in \mathbb{N}$, $\sigma \in \mathbb{R}$. Let Ξ denote an inhomogeneous Poisson-point process on $\mathbb{R}^d \times [1, \infty)$ with intensity measure

$$\mu_\tau(dx \times dw) := \mathrm{Leb} \otimes F_W(dw) := dx \times (\tau - 1)w^{-\tau}dw \tag{5.1.2}$$

For some $p \in (0, 1], \beta > 0$, we define the connectivity function with $\kappa_{1,\sigma}$ from (1.3.5) as

$$p\big((x_u, w_u), (x_v, w_v)\big) := \begin{cases} p \min\left\{1, \left(\beta \dfrac{\kappa_{1,\sigma}(w_u, w_v)}{\|x_v - x_u\|^d}\right)^\alpha\right\}, & \text{if } \alpha < \infty, \\ p\mathbb{1}\left\{\beta \dfrac{\kappa_{1,\sigma}(w_u, w_v)}{\|x_v - x_u\|^d} \geqslant 1\right\}, & \text{if } \alpha = \infty. \end{cases} \tag{5.1.3}$$

Conditionally on Ξ, the infinite random graph $\mathcal{G}_\infty = (\mathcal{V}(\mathcal{G}_\infty), \mathcal{E}(\mathcal{G}_\infty))$ is given by $\mathcal{V}(\mathcal{G}_\infty) = \Xi$ and for each pair $u = (x_u, w_u), v = (x_v, w_v) \in \Xi$ the edge $\{u,v\}$ is in $\mathcal{E}(\mathcal{G}_\infty)$ with probability $p\big((x_u,w_u),(x_v,w_v)\big)$; conditionally independently given Ξ.

Write for $[a,b)\subseteq\mathbb{R}_+$, $\Xi_n[a,b):=\Xi\cap(\Lambda_n\times[a,b))$. For the finite induced subgraphs of $\mathcal{G}_\infty$ on $\Xi_n[a,b)$ and $\Xi_n:=\Xi_n[1,\infty)$, we respectively write

$$\mathcal{G}_n[a,b):=(\mathcal{V}_n[a,b),\mathcal{E}_n[a,b)),\qquad \mathcal{G}_n:=(\mathcal{V}_n,\mathcal{E}_n). \tag{5.1.4}$$

We call Ξ a *realization* of the vertex set. We denote the measure induced by the Palm version of Ξ where a vertex at location x is present (with unknown mark) by $\mathbb{P}^x$.

Both the vertex set $\mathcal{V}$ and the profile function ρ in (5.1.3) can be replaced to obtain more general KSRGs. We expect that our results generalise to KSRGs with lattices as vertex sets, although some of our current proof techniques do not apply, since they make use of the independence property of PPPs. The projection of the vertex set $\mathcal{V}=\Xi$ onto the spatial dimensions is a unit-intensity Poisson point process. The formula in (5.1.3) restricts to models with either a threshold-profile ($\alpha=\infty$) or a specific long-range profile (cf. $\rho_\alpha(t)$ in (1.3.3)), for readability. Our results readily extend to more general regularly varying profiles and fitness distributions, i.e., they allow slowly varying functions in (1.3.3) and (5.1.2), but then the exponent ζ needs to be corrected by an additive $o(1)$ term.

The interpolating KSRG model includes models with kernels $\kappa_{\max}$, κ_{pa}, κ_{prod} as special cases. The parameter p governs Bernoulli percolation of the edges, and a higher β increases the edge density. While p,β do not affect ζ, they may influence supercriticality for certain values of (α,τ,σ,d), see [63, 108, 109] on robustness of some KSRGs. See Section 1.3 above for the role of (α,τ,σ,d).

Main results

We proceed to the main results of this chapter. Define the *multiplicity of the maximum* of a finite set $\mathcal{Z}\subset\mathbb{R}$ as:

$$\mathfrak{m}(\mathcal{Z}):=\mathfrak{m}_{\mathcal{Z}}:=\sum_{\zeta\in\mathcal{Z}}\mathbb{1}\{\zeta=\max(\mathcal{Z})\}. \tag{5.1.5}$$

Recall ζ_{ll}, ζ_{lh}, ζ_{hh}, and ζ_{nn} from (4.2.1), (4.2.2), (4.2.4), and (4.2.5), respectively.

Theorem 5.1.2 (Subexponential decay)**.** *Consider a supercritical i-KSRG model in Definition 5.1.1, with parameters $\alpha>1,\tau\in(2,2+\sigma)$, $\sigma>0$, and $d\in\mathbb{N}$. Then there exists $A>1$ such that, for all $k\geqslant 1$:*

(i) *For* $\mathcal{Z} = \{\zeta_{\mathrm{ll}}, \zeta_{\mathrm{lh}}, \zeta_{\mathrm{hh}}, \zeta_{\mathrm{nn}}\}$ *and* $n \in (Ak, \infty]$

$$\mathbb{P}^0\big(|\mathcal{C}_n(0)| \geqslant k,\ 0 \notin \mathcal{C}_n^{(1)}\big) \geqslant \exp\big(-Ak^{\max(\mathcal{Z})}\log^{m_{\mathcal{Z}}-1}(k)\big), \quad (5.1.6)$$

(ii) *if additionally* $\sigma \leqslant \tau - 1$*, then for all* $n \in (Ak^{1+\zeta_{\mathrm{hh}}/\alpha}, \infty]$

$$\mathbb{P}^0\big(|\mathcal{C}_n(0)| \geqslant k,\ 0 \notin \mathcal{C}_n^{(1)}\big) \leqslant A\exp\big(-(1/A)k^{\zeta_{\mathrm{hh}}}\big). \quad (5.1.7)$$

The results also hold when $\alpha = \infty$ *by taking the appropriate limit in* $\zeta_{\mathrm{hh}}(\alpha)$ *in* (4.2.4).

We will obtain the following statement as a corollary from Theorem 5.1.2.

Corollary 5.1.3 (Law of large numbers for the size of the giant component)**.** *Consider a supercritical i-KSRG model in Definition 5.1.1, with parameters* $\alpha > 1, \tau \in (2, 2+\sigma)$*,* $\sigma \in (0, \tau-1]$*, and* $d \in \mathbb{N}$*. Then,*

$$\frac{|\mathcal{C}_n^{(1)}|}{n} \xrightarrow{\mathbb{P}} \mathbb{P}^0\big(|\mathcal{C}_\infty(0)| = \infty\big), \qquad \text{as } n \to \infty. \quad (5.1.8)$$

The analogue for the size of the second-largest component in $\mathcal{G}_n$ is the following theorem.

Theorem 5.1.4 (Second-largest component)**.** *Consider a supercritical i-KSRG model as in Definition 5.1.1, with parameters* $\alpha > 1, \tau \in (2, 2+\sigma)$*,* $\sigma > 0$*, and* $d \in \mathbb{N}$*. Then the following hold:*

(i) *For* $\mathcal{Z} = \{\zeta_{\mathrm{ll}}, \zeta_{\mathrm{lh}}, \zeta_{\mathrm{hh}}, \zeta_{\mathrm{nn}}\}$*, there exist constants* $A, \delta, n_0 > 0$*, such that for all* $n \in [1, \infty)$

$$\mathbb{P}\Big(|\mathcal{C}_n^{(2)}| \geqslant (1/A)\big(\log(n)/((\log\log(n))^{m_{\mathcal{Z}}-1}\big)^{1/\max(\mathcal{Z})}\Big) \geqslant 1 - n^{-\delta}, \quad (5.1.9)$$

(ii) *If additionally* $\sigma \leqslant \tau - 1$*, then for all* $\delta > 0$*, there exists* $A > 0$ *such that for all* $n \in [1, \infty)$

$$\mathbb{P}\big(|\mathcal{C}_n^{(2)}| \leqslant A\log^{1/\zeta_{\mathrm{hh}}}(n)\big) \geqslant 1 - n^{-\delta}. \quad (5.1.10)$$

The results also hold when $\alpha = \infty$ *by taking the appropriate limit in* $\zeta_{\mathrm{hh}}(\alpha)$ *in* (4.2.4).

We make a few remarks about the two theorems. The assumption $\tau \in (2, 2+\sigma)$, (corresponding to $x + y \geqslant 1$ on Figure 6b) ensures that the model is supercritical for all $p, \beta > 0$ in (5.1.3), see Proposition 5.4.13 below, implying that there exists a unique infinite component $\mathcal{C}_\infty^{(1)}$. The condition $\tau < 2 + \sigma$ also ensures $\zeta_{\mathrm{hh}} > 0$. In part (ii), the condition $\sigma \leqslant \tau - 1$ is automatically satisfied for models with kernels $\kappa_{1,\sigma}$ with $\sigma \leqslant 1$, including $\kappa_{\mathrm{prod}} \equiv \kappa_{1,1}$, because $\tau > 2$. While we believe the restriction $\sigma \leqslant \tau - 1$ is a technical condition, it reflects the change of the tail exponent of the degree distribution at $\sigma = \tau - 1$: the contribution of higher-fitness vertices to the degree of a vertex is negligible compared to the total degree only when $\sigma < \tau - 1$ [179]. The assumption $n > k^{1+\zeta_{\mathrm{hh}}/\alpha}$ in Theorem 5.1.2(ii) is also a technical artifact of our proof, and can be relaxed to $n > k^{1+\varepsilon}$ using results from [151]. The lower bound in (5.1.6) is the same as in Conjecture 4.2.1, here we prove it in the region $\tau \in (2, 2+\sigma)$, and in [151] when $\tau \geqslant 2 + \sigma$.

The next theorem shows that the stretch-exponent ζ is also related to lower large deviations of the largest component. Contrary to the theorems above, it holds also for $\tau \geqslant 2 + \sigma$.

Theorem 5.1.5 (Speed of lower-tail large deviations of the giant)**.** *Consider a supercritical i-KSRG model as in Definition 5.1.1, with parameters* $\alpha > 1, \tau > 2$, $\sigma \geqslant 0$, *and* $d \in \mathbb{N}$. *There exists a constant* $A > 0$ *such that for all* $\rho > 0$, *and* n *sufficiently large, with* $\mathcal{Z} = \{\zeta_{\mathrm{ll}}, \zeta_{\mathrm{lh}}, \zeta_{\mathrm{hh}}, \zeta_{\mathrm{nn}}\}$,

$$\mathbb{P}\big(|\mathcal{C}_n^{(1)}| < \rho n\big) \geqslant \exp\big(-\tfrac{1}{A\rho} \cdot n^{\max(\mathcal{Z})} \log^{m_{\mathcal{Z}}-1}(n)\big). \tag{5.1.11}$$

In [151], we give an almost matching upper bound for the left-hand side in (5.1.11) for some $\rho > 0$, with an error of $o(1)$ in the exponent of n. Next, we interpret our results for κ_{prod} models.

Example 5.1.6 (Matching bounds for product-kernel models)**.** In the *open* region in the $(\alpha, \tau, \sigma, d)$-phase diagram where ζ_{hh} is the unique maximum – corresponding to the (hh)-regime – our bounds in (5.1.6) and (5.1.7), and in (5.1.9) and (5.1.10) are matching. This gives sharp bounds for *scale-free percolation, geometric inhomogeneous random graphs, and hyperbolic random graphs*. In these three models $\sigma = 1$, and whenever $\max\{\zeta_{\mathrm{ll}}, \zeta_{\mathrm{lh}}, \zeta_{\mathrm{hh}}, \zeta_{\mathrm{nn}}\} = \zeta_{\mathrm{hh}} = (3-\tau)/(2-(\tau-1)/\alpha)$ (the red region in Figure 6a), then

$$\mathbb{P}(k \leqslant |\mathcal{C}(0)| < \infty) = \exp(-\Theta(k^{\zeta_{\mathrm{hh}}})),$$

$$\mathbb{P}\big(|\mathcal{C}_n^{(1)}| < \rho n\big) \geqslant \exp\big(-\tfrac{1}{A\rho} n^{\zeta_{\mathrm{hh}}}\big),$$

$$\mathbb{P}\big(A^{-1} \log^{1/\zeta_{\mathrm{hh}}}(n) \leqslant |\mathcal{C}_n^{(2)}| \leqslant A \log^{1/\zeta_{\mathrm{hh}}}(n)\big) \geqslant 1 - n^{-\delta}.$$

In [151], we present similar bounds for these κ_{prod}-KSRGs when $\max(\mathcal{Z}) = \zeta_{\text{ll}} = 2-\alpha$, which occurs when $\alpha \in (1, \min\{\tau-1, 1+1/d\})$ (corresponding to the yellow region in Figure 6a). For κ_{prod} models, ζ_{lh} is never maximal, hence these models present only three phases in their phase diagram for ζ. For (threshold) hyperbolic random graphs (HRG), [162] proved that $|\mathcal{C}_n^{(2)}| = \Theta((\log n)^{2/(3-\tau)})$, which is included in Theorem 5.1.4: there is an isomorphism between HRGs and a 1-dimensional KSRG model with $\sigma{=}1$, $\tau \in (2,3), \alpha = \infty$ (see [40] or [166, Section 9]). Since for HRG we have $\alpha = \infty$, all small components are localized in space in HRGs. However, when $\alpha < \infty$, the long-range edges lead to de-localized small components. This makes proofs of lower bounds somewhat more and upper bounds significantly more complicated than those in [162] (at least up to our judgement).

5.1.1 *Notation*

For a discrete set $\mathcal{S}$, we write $|\mathcal{S}|$ for the size of the set. For a subset $\mathcal{K} \subseteq \mathbb{R}^d$, let $\mathrm{Vol}(\mathcal{K})$ denote the Lebesgue measure of $\mathcal{K}$, $\partial\mathcal{K}$ be its boundary and $\mathcal{K}^c$ its complement. We denote the complement of an event $\mathcal{A}$ by $\neg\mathcal{A}$. Formally we define a vertex v by a pair of location and mark, i.e., $v := (x_v, w_v)$, but we will sometimes still write $v \in \mathcal{K}$ if $x_v \in \mathcal{K}$. We write for $\mathcal{Q} \subseteq \mathbb{R}^d$ and $a \leqslant b$

$$\Xi_{\mathcal{Q}}[a,b) := \Xi \cap \big(\mathcal{Q} \times [a,b)\big), \quad \text{and } \Xi_s[a,b) := \Xi_{\Lambda_s}[a,b) \tag{5.1.12}$$

for the restriction of the vertices with mark in $[a,b)$ and location in $\mathcal{Q}$ and Λ_s from (5.1.1) respectively. By convention, we set $(\alpha-1)/\alpha := 1$ if $\alpha = \infty$ throughout the chapter.

5.2 METHODOLOGY

5.2.1 *Upper bounds*

After sketching the strategy for upper bound on the size of the second-largest component, we explain how to obtain the cluster-size decay from it, then we sketch the lower bound.

5.2.1.1 *Second-largest component*

We aim to show an upper bound of the form

$$\mathbb{P}\big(|\mathcal{C}_n^{(2)}| \geqslant k\big) \leqslant (n/c)\exp(-ck^{\zeta_{hh}}) =: \mathrm{err}_{n,k}. \tag{5.2.1}$$

for arbitrary values of $n \geqslant \mathrm{poly}(k)$ and some constant $c > 0$. Such a bound yields (5.1.10) when one substitutes $k = A\log^{1/\zeta_{hh}}(n)$ for a sufficiently large constant $A = A(\delta) > 0$. Throughout the outline we assume that $(n/k)^{1/d} \in \mathbb{N}$. The proof consists of four revealment stages, illustrated in Figure 7.

Step 1. Building a backbone. We set $w_{hh} := \Theta(k^{\gamma_{hh}})$ with $\gamma_{hh} > 0$ from (4.2.3). We partition the volume-n box Λ_n into n/k smaller sub-boxes of volume k. In this first revealment step we only reveal the location and edges between vertices in $\Xi_n[w_{hh}, 2w_{hh})$, obtaining the graph $\mathcal{G}_{n,1} := \mathcal{G}_n[w_{hh}, 2w_{hh})$. We show that $\mathcal{G}_{n,1}$ contains a connected component $\mathcal{C}_{bb}$ that contains $\Theta(k^{\zeta_{hh}})$ many vertices in each subbox, that we call *backbone* vertices. We show that this event – say $\mathcal{A}_{bb}$ – has probability at least $1 - \mathrm{err}_{n,k}$. We do this by ordering the subboxes so that subboxes with consecutive indices share a $(d-1)$-dimensional face, and by iteratively connecting $\Theta(k^{\zeta_{hh}})$ many vertices in the next subbox to the component we already built, combined with a union bound over all subboxes. The event $\mathcal{A}_{bb}$ ensures us to show that *independently* for all $v \in \Xi_n[2w_{hh}, \infty)$,

$$\mathbb{P}\big(v \leftrightarrow \mathcal{C}_{bb} \mid \mathcal{A}_{bb}, v \in \Xi_n[2w_{hh}, \infty)\big) \geqslant 1/2, \tag{5.2.2}$$

regardless of the location of v. We call vertices in $\Xi_n[2w_{hh}, \infty)$ *connector* vertices.

Step 2: Revealing low-mark vertices. We now also reveal all vertices with mark in $[1, w_{hh})$, and all their incident edges to $\mathcal{G}_{n,1}$ and towards each other, i.e., the graph $\mathcal{G}_{n,2} := \mathcal{G}_n[1, 2w_{hh}) \supseteq \mathcal{G}_{n,1}$.

Step 3: Pre-sampling randomness to avoid merging of small components. To show (5.2.1), in the fourth revealment stage below we must *avoid small-to-large merging*: when the edges to/from some $v \in \Xi_n[2w_{hh}, \infty)$ are revealed, a set of small components, each of size smaller than k, could merge into a component of size at least k without connecting to the giant component. If we simply revealed $\Xi_n[2w_{hh}, \infty)$ after Step 2, (5.2.2) would not be sufficient to show that small-to-large merging occurs with probability at most $1 - \mathrm{err}_{n,k}$. So, we pre-sample randomness: we split $\Xi_n[2w_{hh}, \infty)$ into two PPPs:

$$\Xi_n[2w_{hh}, \infty) = \Xi_n^{(\mathrm{sure})}[2w_{hh}, \infty) \cup \Xi_n^{(\mathrm{unsure})}[2w_{hh}, \infty),$$

where $\Xi_n^{(\mathrm{sure})}[2w_{\mathrm{hh}},\infty), \Xi_n^{(\mathrm{unsure})}[2w_{\mathrm{hh}},\infty)$ are *independent* PPPs with equal intensity: using (5.2.2) and helping random variables that encode the presence of edges, we pre-sample whether a connector vertex connects for *sure* to $\mathcal{C}_{\mathrm{bb}}$ by at least one edge; forming $\Xi_n^{(\mathrm{sure})}[2w_{\mathrm{hh}},\infty)$. Vertices in $\Xi_n^{(\mathrm{unsure})}[2w_{\mathrm{hh}},\infty)$ might still connect to $\mathcal{C}_{\mathrm{bb}}$ since $1/2$ is only a lower bound in (5.2.2), but we ignore that information. We crucially use the property that thinning a PPP yields two *independent* PPPs. The adaptation of our technique to lattices as vertex-set seems non-trivial due to this step. We reveal now $\Xi_n^{(\mathrm{unsure})}[2w_{\mathrm{hh}},\infty)$. Let $\mathcal{G}_{n,3} \supseteq \mathcal{G}_{n,2}$ be the graph induced on the vertex set

$$\mathcal{V}_{n,3} := \Xi_n[1, 2w_{\mathrm{hh}}) \cup \Xi_n^{(\mathrm{unsure})}[2w_{\mathrm{hh}},\infty). \tag{5.2.3}$$

Step 4: Cover expansion, a volume-based argument. We now reveal $\Xi_n^{(\mathrm{sure})}[2w_{\mathrm{hh}},\infty)$ and merge all components of size at least k with the largest component in $\mathcal{G}_{n,3}$ with probability at least $1-\mathrm{err}_{n,k}$. Small-to-large merging cannot happen since vertices in $\Xi_n^{(\mathrm{sure})}[2w_{\mathrm{hh}},\infty)$ all connect to $\mathcal{C}_{\mathrm{bb}}$. We argue how to obtain (5.2.1).

Step 4a: Not too dense components via proper cover. For a component $\mathcal{C} \subseteq \mathcal{G}_{n,3}$, the *proper cover* $\mathcal{K}_n(\mathcal{L}) \subseteq \Lambda_n$ is the *union of volume-*1 *boxes* centered at the vertices of $\mathcal{C}$ (the formal definition below is slightly different). Fixing a constant $\delta > 0$, we say that $\mathcal{C}$ is not *too dense* if

$$\mathrm{Vol}(\mathcal{K}_n(\mathcal{C})) \geqslant \delta|\mathcal{C}|. \tag{5.2.4}$$

Using the connectivity function p in (5.1.3) and $w_{\mathrm{hh}} = \Theta(k^{\gamma_{\mathrm{hh}}})$, there exists k_0 such that for any $k \geqslant k_0$ and *any* pair of vertices $u \in \mathcal{V}_{n,3}$ in (5.2.3) and $v \in \Xi_n^{(\mathrm{sure})}[2w_{\mathrm{hh}},\infty)$ within the same volume-1 box,

$$p(u,v) \geqslant p/2. \tag{5.2.5}$$

Using this bound and that $\Xi_n^{(\mathrm{sure})}[2w_{\mathrm{hh}},\infty)$ is a PPP, when $|\mathcal{C}| \geqslant k$, with probability at least $1-\mathrm{err}_{n,k}$, at least $\Theta(k^{\zeta_{\mathrm{hh}}})$ many vertices of $\Xi_n^{(\mathrm{sure})}[2w_{\mathrm{hh}},\infty)$ fall inside $\mathcal{K}_n(\mathcal{C})$ and at least one of them connects to $\mathcal{C}$ by an edge. Since these vertices belong to $\Xi_n^{(\mathrm{sure})}[2w_{\mathrm{hh}},\infty)$, they connect to $\mathcal{C}_{\mathrm{bb}}$ by construction, merging $\mathcal{C}$ with the component containing $\mathcal{C}_{\mathrm{bb}}$.

Step 4b: Too dense components via cover-expansion. We still need to handle components $\mathcal{C} \subseteq \mathcal{G}_{n,3}$ with $|\mathcal{C}| \geqslant k$ but the opposite of (5.2.4). These may exist (outside the component of $\mathcal{C}_{\mathrm{bb}}$) since the PPP Ξ_n contains dense areas, e.g., volume-one balls with $\Theta(\log(n)/\log\log(n))$ vertices. We introduce a deterministic algorithm, the *'cover-expansion algorithm'*, that outputs for any (deterministic) set $\mathcal{L}$ of at least k vertices a set $\mathcal{K}^{\exp}(\mathcal{L}) \subset \mathbb{R}^d$, called the *expanded cover* of $\mathcal{L}$, that satisfies (5.2.4) and a bound similar to (5.2.5) (provided that $\sigma \leqslant \tau - 1$). In the design of the set $\mathcal{K}^{\exp}(\mathcal{L})$ we quantify the idea that a connector vertex can be *farther away in space* from a too dense subset $\mathcal{L}' \subseteq \mathcal{L}$, while ensuring connection probability at least $p/2$ to $\mathcal{L}'$, as if $\mathcal{L}'$ contains a single vertex. We apply this algorithm with $\mathcal{L} = \mathcal{C}$ for components of size at least k of $\mathcal{G}_{n,3}$ that do not satisfy (5.2.4) and do not contain $\mathcal{C}_{\mathrm{bb}}$. The remainder of the proof is identical to Step 4a. Steps 4a, 4b, and a union bound over all components of size at least k in $\mathcal{G}_{n,3}$ yield (5.2.1).

5.2.1.2 *Subexponential decay, upper bound.*

Consider k fixed. We obtain the cluster-size decay (5.1.7) for any $n \in [\mathrm{poly}(k), n_k]$ with $n_k = \exp(\Theta(k^{\zeta_{\mathrm{hh}}}))$ by substituting n_k into (5.2.1). To extend it to larger n, we first identify the lowest mark $\overline{w}(n)$ such that *all* vertices with mark at least $\overline{w}(n)$ belong to the giant component $\mathcal{C}_n^{(1)} \subseteq \mathcal{G}_n$ with sufficiently high probability (in n). Then we embed Λ_{n_k} in Λ_n and show that

$$\begin{aligned} \mathbb{P}^0\big(|\mathcal{C}_n(0)| \geqslant k,\ 0 \notin \mathcal{C}_n^{(1)}\big) & \\ \leqslant \mathbb{P}^0\big(|\mathcal{C}_{n_k}^{(2)}| \geqslant k\big) &+ \mathbb{P}^0\big(\mathcal{C}_{n_k}^{(1)} \not\subseteq \mathcal{C}_n^{(1)}\big) \\ &+ \mathbb{P}^0\big(|\mathcal{C}_n(0)| \geqslant k,\ 0 \notin \mathcal{C}_n^{(1)}, |\mathcal{C}_{n_k}(0)| < k\big). \end{aligned} \tag{5.2.6}$$

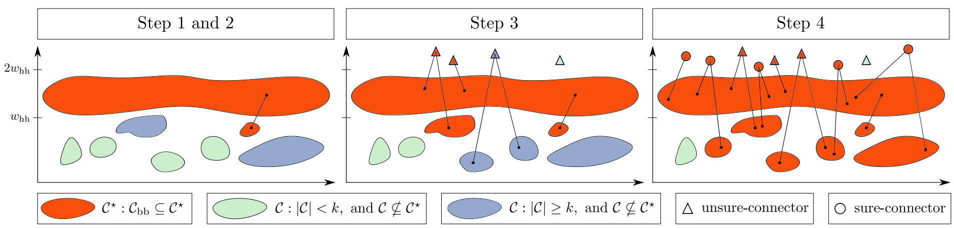


Figure 7: Upper bound. The y-axis represents marks, the x-axis represents space. After Steps 1 and 2 there is a component $\mathcal{C}^\star$ containing the backbone that is connected to some small components from $\mathcal{G}_n[1, w_{\mathrm{hh}})$. After Step 3, the unsure connectors are revealed: there is small-to-large merging; some unsure connectors connect to the backbone. After Step 4, each component of size at least k merged with the largest component via a sure-connector; unmerged small components and unsure-connectors outside $\mathcal{C}_n^{(1)}$ remain.

The first term on the right-hand side has the right error bound by (5.2.1). We relate the second term to the event that for some $\tilde{n} \in (n_k, n]$ there is no polynomially-sized largest component or the second-largest component is too large. The event in the third term implies that one of the at most $k-1$ vertices in $\mathcal{C}_{n_k}(0)$ has an edge of length $\Omega(n_k^{1/d}/k)$, which will have probability at most $\mathrm{err}_{n_k,k}$, since these vertices have mark at most $\overline{w}(n_k)$.

5.2.2 *Lower bound*

For the subexponential decay, we compute the probability of a specific event satisfying $k \leqslant |\mathcal{C}(0)| < \infty$. We draw a ball $\mathcal{B}$ of volume $\Theta(k)$ around the origin, and compute an *optimally suppressed mark-profile*: the PPP Ξ must fall *below* a $(d+1)$-dimensional mark-surface $\mathcal{M} := \{(x, f(x)), x \in \mathbb{R}^d\}$, i.e., $w_v \leqslant f(x_v)$ must hold for all $(x_v, w_v) \in \Xi$. We write $\{\Xi \leqslant \mathcal{M}\}$ for this event. The value of $f(x)$ is increasing in $\|x - \partial\mathcal{B}\|$ since high-mark vertices close to $\partial\mathcal{B}$ are most likely to have edges crossing $\partial\mathcal{B}$. $\mathcal{M}$ is optimized so that $\mathbb{P}(\Xi \leqslant \mathcal{M}) \sim \mathbb{P}(\mathcal{B} \not\leftrightarrow \mathcal{B}^c \mid \Xi \leqslant \mathcal{M})$, where $\{\mathcal{B} \not\leftrightarrow \mathcal{B}^c\}$ is the event that there is no edge present between vertices in $\mathcal{B}$ and those in its complement. Both events occur with probability $\exp(-\Theta(k^{\max\{\zeta_{\mathrm{ll}},\zeta_{\mathrm{lh}},\zeta_{\mathrm{hh}},\zeta_{\mathrm{nn}}\}}))$, up to logarithmic correction terms. We then find an isolated component of size at least k inside $\mathcal{B}$ using a technique that works when $\tau < 2 + \sigma$. We use a boxing argument to extend this argument to the lower bound on the second-largest component of $\mathcal{G}_n$, similar to [162], and to obtain a lower bound on $\mathbb{P}(|\mathcal{C}_n^{(1)}| < \rho n)$.

5.2.3 *Generalisation of results*

Most of our results extend to more general (interpolating) KSRGs than that in Definition 5.1.1. In particular, one can replace the spatial location of vertices to be $\mathbb{Z}^d$ (or any other lattice), called *i-KSRGs on* $\mathbb{Z}^d$. To prove this generalisation, one needs to replace concentration inequalities for Poisson random variables by Chernoff bounds for Binomial random variables and replace integrals over $\mathbb{R}^d$ by summations over $\mathbb{Z}^d$. We believe that Theorems 5.1.2–5.1.5 fully extend to $\sigma > \tau - 1$ and i-KSRGs on $\mathbb{Z}^d$, by pre-sampling more information (similar to Step 3) already before Step 1. We decided to not include this, avoiding the extra amount of required technicalities. We highlight the lemmas and propositions that do *not* immediately generalise by adding a $\star$ to their statement (e.g. Proposition 5.4.1).

Besides that, *all* results extend to i-KSRGs where the number of vertices is fixed to be n, where each vertex u has an iid mark from distribution F_W in (5.1.2) in Definition 5.1.1, and an independent uniform location in the volume-n box Λ_n.

Moreover, all results (even when restricted to the graph in Λ_n) extend to the Palm-version $\mathbb{P}^x$ of $\mathbb{P}$ where the vertex set Ξ is conditioned to contain a point at location $x \in \Lambda_n$ with unknown mark (which is not a priori obvious, since the model restricted to a finite box around the origin is not translation-invariant). We will omit this in our statements and proofs, and only add the superscript x to our notation if we explicitly use the Palm version.

Organisation of the chapter

In Section 5.3 we explain the cover expansion from Step 4 in the outline of the upper bound. This technique is a novel technical contribution and is interesting in its own right. We use it also in [151]. In Section 5.4 we prove the upper bound on the size of the second-largest component. Then in Section 5.5 we extend it to the cluster-size decay, and show Corollary 5.1.3. In Section 5.6 we discuss the lower bounds, including the proof of Theorem 5.1.5. The first propositions in Sections 5.4—5.6 immediately imply the proofs of Theorems 5.1.2 and 5.1.4, as we verify near the end of Section 5.6.

5.3 THE COVER AND ITS EXPANSION

The goal of this section is to develop the cover-expansion technique in Step 4b of Section 5.2.1.1. The statements apply also to KSRGs on vertex sets other than a PPP. First we define a desired property for a set of vertices based on their spatial locations. Recall the definition $\Lambda_s(x) = \Lambda(x, s)$ from (5.1.1).

Definition 5.3.1 (s-expandable point-set)**.** Let $\mathcal{S} \subset \mathbb{R}^d \times [1, \infty)$ be a discrete set of points, and $s > 0$. We call $\mathcal{S}$ s-expandable if for all $x \in \mathbb{Z}^d$ and all $s' \in \{s + \ell/(ed^{d/2}2^{3d}) : \ell \in \mathbb{N}\}$,

$$|\mathcal{S} \cap \Lambda_{s'}(x)|/s' \leqslant e.$$

A discrete set $\mathcal{S} \subset \mathbb{R}^d$ is s-expandable if there are no large boxes, centered at any site $z \in \mathbb{Z}^d$, with too high ratio of number of vertices in $\mathcal{S}$ in the box vs its volume. Further, if $\mathcal{S}$ is s-expandable, then any subset of $\mathcal{S}$ is s-expandable; also, if $\mathcal{S}$ is s-expandable, then for all $\tilde{s} \geqslant s$, $\mathcal{S}$ is also $\tilde{s}$-expandable. The next proposition solves the problem of too-dense components in space, cf. (5.2.4).

Proposition 5.3.2 (Covers and expansions for s-expandable sets)**.** *Consider an i-KSRG with connection probabilities between vertices given by* (5.1.3) *with kernel* $\kappa_{1,\sigma}$ *and profile function with parameter* $\alpha > 1$ *and (arbitrary) marked vertex set* $\mathcal{V} = \{(x_v, w_v)\}_{v \in \mathcal{V}}$*. For a given* $\underline{w} > 2^d d^{d/2}/\beta$*, let* $s(\underline{w}) > 0$ *satisfy*

$$s(\underline{w}) := (2^d \beta \underline{w})^{1/(1-1/\alpha)}. \tag{5.3.1}$$

Then, for any $s(\underline{w})$*-expandable* $\mathcal{L} \subseteq \mathcal{V} \cap \Lambda_n$ *of vertices (for some* $n > 0$*), there is a set* $\mathcal{K}_n(\mathcal{L}) \subseteq \Lambda_n$ *with*

$$\mathrm{Vol}(\mathcal{K}_n(\mathcal{L})) \geqslant \frac{1}{2^{4d+1} e d^{d/2}} |\mathcal{L}|, \tag{5.3.2}$$

such that all vertices $v \in \mathcal{V} \cap \mathcal{K}_n(\mathcal{L})$ *with mark* $w_v \geqslant \underline{w}$ *satisfy independently of each other that*

$$\mathbb{P}\big(v \leftrightarrow \mathcal{L} \mid \mathcal{V}\big) \geqslant p/2. \tag{5.3.3}$$

We use two different constructions for the set $\mathcal{K}_n(\mathcal{L})$. If $\mathcal{L}$ is not too dense (see (5.2.4)), we will use a *proper cover* (see Definition 5.3.6 below). If, however, the points of $\mathcal{L}$ are densely concentrated in small areas, we will use a new (deterministic) algorithm, the *cover-expansion algorithm*, producing an *expanded cover* (see Definition 5.3.7 below) that still satisfies the connection probability in (5.3.3). This will prove Proposition 5.3.2. We start with some preliminaries:

Definition 5.3.3 (Cells in a volume-n box)**.** Let $\widetilde{B}_z$ be a box of volume 1 centered around $z \in \mathbb{Z}^d$. For any two neighbouring boxes $\widetilde{B}_z, \widetilde{B}_{z'}$, allocate the shared boundary $\partial\widetilde{B}_z \cap \partial\widetilde{B}_{z'}$ to precisely one of the boxes. For each $u \in \mathbb{Z}^d$ such that $u \notin \Lambda_n$ but $\widetilde{B}_u \cap \Lambda_n \neq \emptyset$, let $z(u) := \arg\min\{\|u - z\| : z \in \Lambda_n \cap \mathbb{Z}^d\}$, and then define for each $z \in \mathbb{Z}^d \cap \Lambda_n$ the *cell of* z as

$$B_z := \Big(\widetilde{B}_z \cap \Lambda_n\Big) \cup \Big(\bigcup_{u \in \mathbb{Z}^d : z(u) = z} \big(\widetilde{B}_u \cap \Lambda_n\big)\Big).$$

In words, boxes that have their center inside Λ_n but are not fully contained in Λ_n are truncated, while boxes that have their centers outside

Λ_n but intersect it are merged with the closest box with center inside Λ_n. Clearly, at every point of Λ_n at most 2^d cells are merged together, and only $1/2$ of the radius in each coordinate can be truncated. Thus, for each cell B_z

$$\sup\{\|x-y\| : x, y \in B_z\} \leqslant 2\sqrt{d}; \quad \text{and} \qquad 2^{-d} \leqslant \mathrm{Vol}(B_z) \leqslant 2^d. \quad (5.3.4)$$

Definition 5.3.4 (Notation for cells containing vertices)**.** Let $\mathcal{L} \subset \Lambda_n$ be the set of locations of a set of vertices in any given realization $(\mathcal{V}, (w_v)_{v \in \mathcal{V}})$. Let $\{B_{z_i}\}_{i=1}^{m'}$ be the cells with $\mathcal{L} \cap B_{z_i} \neq \emptyset$. Let $\mathcal{L}_i := \mathcal{L} \cap B_{z_i}$, $\ell_i := |\mathcal{L}_i|$ and $L := |\mathcal{L}| = \sum_{i=1}^{m'} \ell_i$.

We will distinguish two cases for the arrangement of the vertices among the cells: either the number of cells is linear in the number of vertices, or there is a positive fraction of all cells that all contain 'many' vertices. To make this more precise, we prove the next combinatorial claim.

Claim 5.3.5 (Pigeon-hole principle for cells)**.** *Let $\delta \in (0,1)$, $\nu \geqslant 1$, and $\ell_1, \dots, \ell_{m'} \geqslant 1$ integers such that $\sum_{i \leqslant m'} \ell_i = L$. If $m' < L(1-\delta)/\nu$ then*

$$\exists\, \mathcal{I} \subseteq [m'] : \quad \forall i \in \mathcal{I} : \ell_i \geqslant \nu, \qquad \textit{and} \qquad \sum_{i \in \mathcal{I}} \ell_i \geqslant \delta L. \qquad (5.3.5)$$

Proof. Assume for contradiction that $\delta, \nu, \ell_1, \dots, \ell_{m'}$ are such that $m' < L(1-\delta)/\nu$ holds but (5.3.5) does not hold. Let $\mathcal{J} := \{j : \ell_j < \nu\} \subseteq [m']$ and let $\mathcal{J}^c := [m'] \backslash \mathcal{J}$. Then $\forall i \in \mathcal{J}^c : \ell_i \geqslant \nu$ and hence, we assumed the opposite of (5.3.5), it holds that $\sum_{j \in \mathcal{J}^c} \ell_j < \delta L$. Since the total sum is L, this implies that $\sum_{j \in \mathcal{J}} \ell_j \geqslant (1-\delta)L$. Moreover, since $\ell_j < \nu$ for $j \in \mathcal{J}$, it must hold that $|\mathcal{J}| \geqslant (1-\delta)L/\nu$, which then gives a contradiction with the assumption in that $m' < L(1-\delta)/\nu$. □

We define the first possibility for the set $\mathcal{K}_n(\mathcal{L})$, which is inspired by Claim 5.3.5 with $\nu = ed^{d/2}2^{3d}$ and $\delta = 1/2$.

Definition 5.3.6 (Proper cover)**.** We say that $\mathcal{L}$ admits a *proper cover* if $m' \geqslant |\mathcal{L}|/(2ed^{d/2}2^{3d})$, in Definition 5.3.4, and we define the cover of $\mathcal{L}$ as

$$\mathcal{K}_n^{(\mathrm{prop.})}(\mathcal{L}) := \bigcup_{i \in [m']} B_{z_i}, \quad \text{satisfying} \quad \mathrm{Vol}(\mathcal{K}_n^{(\mathrm{prop.})}) \geqslant \frac{1}{d^{d/2} e 2^{4d+1}} |\mathcal{L}|.$$

By (5.3.4), $\nu = ed^{d/2}2^{3d}$, and $\delta = 1/2$, hence, we obtain the desired volume bound on the right-hand side above, establishing (5.3.2) for sets admitting a proper cover. Moreover, consider now $(x_v, w_v) \in (B_{z_i} \cap \mathcal{L}) \times [1, \infty)$ and $u := (x_u, w_u) \in B_{z_i} \times [\underline{w}, \infty)$ with $B_{z_i} \subseteq \mathcal{K}_n^{(\text{prop.})}$. Then $\|x_u - x_v\| \leqslant 2\sqrt{d}$ by (5.3.4). Since we assumed $\underline{w} \geqslant 2^d d^{d/2}/\beta$ above (5.3.1), using (5.1.3) and (1.3.5),

$$\begin{aligned} p(u,v) &\geqslant p \min\{1, (\beta\kappa_{1,\sigma}(\underline{w}, 1)/(2\sqrt{d})^d)^\alpha\} \\ &= p \min\{1, (\beta\underline{w}/(2\sqrt{d})^d)^\alpha\} \geqslant p. \end{aligned} \tag{5.3.6}$$

This shows (5.3.3) for sets admitting a proper cover. The argument for $\alpha = \infty$ is similar.

The remainder of this section focuses on sets $\mathcal{L}$ that do not admit a proper cover, i.e., the number of cells that contain vertices of $\mathcal{L}$ is too small. We define an "expanded" cover, obtained after applying a suitable volume-increasing procedure to $\cup_{i \in [m']} B_{z_i}$ that will be explained at the end of the section.

5.3.1 *Cover expansion*

In this section we assume that $\mathcal{L}$ does not admit a proper cover. By Claim 5.3.5, and re-indexing cells in Definition 5.3.4, without loss of generality we may assume that $\mathcal{I} = [m] \subseteq [m']$ satisfies (5.3.5) with $\nu = ed^{d/2}2^{3d}$ and $\delta = 1/2$. We use $\Lambda(x, s)$ in (5.1.1) here for the box of volume s centered at $x \in \mathbb{R}^d$.

Definition 5.3.7 (Cover expansion)**.** Let $\mathcal{L}$ be a set of locations of vertices that does not admit a proper cover in the sense of Definitions 5.3.4 and 5.3.6. Let $[m] := \{j : \ell_j \geqslant ed^{d/2}2^{3d}\} \subseteq [m']$ satisfying (5.3.5) with $\nu = ed^{d/2}2^{3d}$ and $\delta = 1/2$. The *cover expansion* is defined as a subset of labels $\mathcal{J}^{(\star)} \subseteq [m]$ and corresponding boxes $(B_j^{(\star)})_{j \in \mathcal{J}^{(\star)}} \subset \mathbb{R}^d$, centered at $(z_j)_{j \in \mathcal{J}^{(\star)}}$, together with an allocation $\overset{\star}{\mapsto}$ of the cells $B_{z_i} : i \leqslant m$ to these boxes, with

$$\text{Cells}_j^{(\star)} := \bigcup_{i \leqslant m} \{i : B_{z_i} \overset{\star}{\mapsto} B_j^{(\star)}\}, \tag{5.3.7}$$

satisfying the following properties:

(disj.) the boxes $(B_j^{(\star)})_{j \in \mathcal{J}^{(\star)}}$ are pairwise disjoint sets in $\mathbb{R}^d$;

(vol.) for all $j \in \mathcal{J}^{(\star)}$; we have

$$B_j^{(\star)} = \Lambda\Big(z_j, \frac{1}{ed^{d/2}2^{3d}} \sum_{i \in \mathrm{Cells}_j^{(\star)}} \ell_i\Big),$$
$$\mathrm{Vol}(B_j^{(\star)}) = \frac{1}{ed^{d/2}2^{3d}} \sum_{i \in \mathrm{Cells}_j^{(\star)}} \ell_i; \tag{5.3.8}$$

(near) for each $i \in [m]$; $i \in \mathrm{Cells}_j^{(\star)}$

$$\|z_i - z_j\|^d \leqslant d^{d/2}\mathrm{Vol}(B_j^{(\star)}). \tag{5.3.9}$$

We define the *expanded cover* of $\mathcal{L}$ as

$$\mathcal{K}_n^{(\exp)}(\mathcal{L}) := \Lambda_n \cap \Big(\bigcup_{j \in \mathcal{J}^{(\star)}} B_j^{(\star)}\Big) \tag{5.3.10}$$

and call $B_j^{(\star)}$ the expanded boxes.

A few comments about Definition 5.3.7: (disj) and (vol) together ensure that the total volume of the expanded cover is proportional to $|\mathcal{L}|$. Further, (vol) ensures that $\mathrm{Vol}(B_j^{(\star)})$ is proportional to the number of vertices that are in cells allocated to $B_j^{(\star)}$. Finally, (near) ensures that the center z_i of each cell B_{z_i} is relatively close to the center of the box to which it is allocated. Observe that when the final box $B_j^{(\star)}$ is large, (5.3.9) allows for large distances between allocated initial cells and the center of $B_j^{(\star)}$.

Proposition 5.3.8 (Every set has either a proper cover or a cover expansion)**.** *Assume $\mathcal{L}$ does not admit a proper cover defined in Definition 5.3.6. Then there exists a cover expansion for $\mathcal{L}$ in the sense of Definition 5.3.6 and 5.3.7. Further, the total volume of the expanded cover is linear in $|\mathcal{L}|$, i.e.,*

$$\mathrm{Vol}(\mathcal{K}_n^{(\exp)}(\mathcal{L})) \geqslant \frac{1}{2^{4d+1}ed^{d/2}}|\mathcal{L}|. \tag{5.3.11}$$

We defer the proof of existence of the desired cover expansion to the end of the section. Assuming that a cover expansion exists, we show now the linearity of its volume along with some other observations. Afterwards, we show how Proposition 5.3.2 follows from Proposition 5.3.8.

Observation 5.3.9 (Cover-expansion properties)**.** *Consider the cover expansion of a set $\mathcal{L}$ that does not admit a proper cover according to Definition 5.3.6.*

(i) *Every expanded box has volume at least* 1*, i.e., for all* $j \in \mathcal{J}^{(\star)}$, $\mathrm{Vol}(B_j^{(\star)}) \geqslant 1$.

(ii) *For any cell with* $B_{z_i} \overset{\star}{\mapsto} B_j^{(\star)}$,

$$\sup\left\{\|x_u - x_v\| : x_u \in \mathcal{L}_i, x_v \in B_j^{(\star)}\right\} \leqslant 4\sqrt{d}\mathrm{Vol}(B_j^{(\star)})^{1/d}.$$

(iii) *For every box* $B_j^{(\star)}$*, there exists a box* B_j' *centered at* z_j *such that*

$$\mathrm{Vol}(B_j') = d^{d/2}2^{3d}\mathrm{Vol}(B_j^{(\star)}), \quad \textit{and} \quad |\mathcal{L} \cap B_j'| \geqslant e\mathrm{Vol}(B_j'). \quad (5.3.12)$$

(iv) *If* $\mathcal{L}$ *is* s*-expandable, then for all* $j \in \mathcal{J}^{(\star)}$

$$\mathrm{Vol}(B_j^{(\star)}) \leqslant d^{-d/2}2^{-3d}s. \quad (5.3.13)$$

(v) *The total volume of a cover expansion is linear in* $|\mathcal{L}|$*, i.e.,* (5.3.11) *holds.*

Proof. Part (i) is a consequence of Definition 5.3.7: every cell with label at most m has $\ell_i \geqslant ed^{d/2}3^d$, so by (vol), i.e., (5.3.8), Observation (i) follows.

For part (ii) we apply the triangle inequality: since $x_u \in \mathcal{L}_i$, it holds that $x_u \in B_{z_i}$, and so by (5.3.4), $\|x_u - z_i\| \leqslant 2\sqrt{d}$; and by (5.3.9) $\|z_i - z_j\| \leqslant \sqrt{d}\mathrm{Vol}(B_j^{(\star)})^{1/d}$; hence $\|x_u - z_j\| \leqslant 2\sqrt{d} + \sqrt{d}\mathrm{Vol}(B_j^{(\star)})^{1/d}$. Also, for any $x_v \in B_j^{(\star)}$ it holds that $\|z_j - x_v\| \leqslant (\sqrt{d}/2)\mathrm{Vol}(B_j^{(\star)})^{1/d}$ by (5.3.8). Combining these bounds and using $\mathrm{Vol}(B_j^{(\star)})^{1/d} \geqslant 1$ yields

$$\begin{aligned}\|x_u - x_v\| \leqslant 2\sqrt{d} + (3\sqrt{d}/2)\mathrm{Vol}(B_j^{(\star)})^{1/d} &\leqslant (7\sqrt{d}/2)\mathrm{Vol}(B_j^{(\star)})^{1/d} \\ &\leqslant 4\sqrt{d}\mathrm{Vol}(B_j^{(\star)})^{1/d},\end{aligned}$$

and part (ii) is proven. For part (iii), note that part (ii) applied to $u \in \mathcal{L}_i \subset B_{z_i}$ and z_j, yields that

$$\sup_{i \in \mathrm{Cells}_j^{(\star)}} \left\{\|x_u - z_j\| : x_u \in \mathcal{L}_i\right\} \leqslant 4\sqrt{d}\mathrm{Vol}(B_j^{(\star)})^{1/d}.$$

Consequently, the box B_j' centered at z_j of volume

$$\mathrm{Vol}(B_j') = d^{d/2}2^{3d}\mathrm{Vol}(B_j^{(\star)})$$

contains all $u \in \mathcal{L}_i$ with $i \in \mathrm{Cells}_j^{(\star)}$. Hence, using (5.3.8), we obtain

$$|\mathcal{L} \cap B_j'| \geqslant \sum_{i \in \mathrm{Cells}_j^{(\star)}} \ell_i = ed^{d/2}2^{3d}\mathrm{Vol}(B_j^{(\star)}) = e\mathrm{Vol}(B_j'), \quad (5.3.14)$$

and part (iii) follows. For part (iv), by combining (5.3.14) with Definition 5.3.1 we see that $\mathcal{L}$ can only be s-expandable if $\mathrm{Vol}(B_j') \leqslant s$. Rearrangement of the first part of (5.3.12) yields (5.3.13).

Finally, for part (v), by an argument similar to (5.3.4), $\mathrm{Vol}(B_j^{(\star)} \cap \Lambda_n) \geqslant 2^{-d}\mathrm{Vol}(B_j^{(\star)})$ for all $j \in \mathcal{J}^{(\star)}$. Since all boxes of the cover expansion are disjoint, and each cell is allocated once, (5.3.8) and (5.3.10) imply that

$$\begin{aligned}\mathrm{Vol}(\mathcal{K}_n^{(\mathrm{exp})}(\mathcal{L})) &= \sum_{j\in\mathcal{J}^{(\star)}} \mathrm{Vol}(B_j^{(\star)} \cap \Lambda_n) \geqslant 2^{-d} \sum_{j\in\mathcal{J}^{(\star)}} \mathrm{Vol}(B_j^{(\star)}) \\ &= \frac{1}{e2^{4d}d^{d/2}} \sum_{j\in\mathcal{J}^{(\star)}} \sum_{i\in\mathrm{Cells}_j^{(\star)}} \ell_i \\ &= \frac{1}{ed^{d/2}2^{4d}} \sum_{i\leqslant m} \ell_i \geqslant \frac{1}{ed^{d/2}2^{4d+1}} L,\end{aligned}$$

where the last bound follows by the assumption in Definition 5.3.7 that $\ell_i \geqslant ed^{d/2}2^{3d}$ for $i \leqslant m$, and the initial assumption that (5.3.5) in Claim 5.3.5 holds for $[m]$ with $\delta = 1/2$. $\square$

Proof of Proposition 5.3.2 assuming Proposition 5.3.8. For sets $\mathcal{L}$ that admit a proper cover, we recall the reasoning below Definition 5.3.6 (in particular (5.3.6)) which implies both bounds (5.3.2) and (5.3.3) in Proposition 5.3.2. Let $\mathcal{L}$ be an s-expandable set that does not admit a proper cover. Let $\mathcal{K}_n^{(\mathrm{exp})}$ be an expanded cover given by the boxes $(B_j^{(\star)})_{j\in\mathcal{J}^{(\star)}}, \mathcal{J}^{(\star)} \subseteq [m]$ and an allocation $\overset{\star}{\mapsto}$ of the initial cells $(B_{z_i})_{i\in[m]}$ to these boxes. The existence of this cover expansion is guaranteed by Proposition 5.3.8. The volume bound (5.3.2) follows from (5.3.11) in Proposition 5.3.8. Hence, it only remains to verify (5.3.3).

Let $u = (x_u, w_u) \in \mathcal{K}_n^{(\mathrm{exp})} \times [\underline{w}, \infty)$. By (disj), and (5.3.10), there exists $j \in \mathcal{J}^{(\star)}$ such that $x_u \in B_j^{(\star)}$. Recall from (5.3.7) that $\mathrm{Cells}_j^{(\star)}$ are the cells allocated to $B_j^{(\star)}$, and from Definition 5.3.4 that $\mathcal{L}_i = \mathcal{L} \cap B_{z_i}$. Let now $\mathcal{L}_j^{(\star)} := \cup_{i\in\mathrm{Cells}_j^{(\star)}} \mathcal{L}_i$. Recall the formula of the connection probability from (5.1.3).

Case (1): $\alpha < \infty$. By Observation 5.3.9(ii) for any $v = (x_v, w_v) \in \mathcal{L}_j^{(\star)}$, and any $x_u \in B_j^{(\star)}$, using the lower bounds for the marks $w_v \geqslant 1, w_u \geqslant \underline{w}$, and that $\kappa(w_u, w_v) = (w_u \vee w_v)(w_u \wedge w_v)^\sigma$ for some $\sigma \geqslant 0$, we obtain using (5.1.3) that

$$\begin{aligned} p(u,v) &= p \min\Big\{1, \beta^\alpha \frac{\kappa(w_u, w_v)^\alpha}{\|x_v - x_u\|^{d\alpha}}\Big\} \\ &\geqslant p \min\Big\{1, \frac{\beta^\alpha \underline{w}^\alpha}{(4\sqrt{d})^{\alpha d} \mathrm{Vol}(B_j^{(\star)})^\alpha}\Big\} =: r. \end{aligned}$$

By (5.3.8), $|\mathcal{L}_j^{(\star)}| = \sum_{i \in \mathrm{Cells}_j^{(\star)}} \ell_i = ed^{d/2}2^{3d}\mathrm{Vol}(B_j^{(\star)})$. Hence, we have

$$\begin{aligned} &\mathbb{P}\big(\nexists (x_v, w_v) \in \mathcal{L}_j^{(\star)} : (x_u, w_u) \leftrightarrow (x_v, w_v)\big) \\ &\quad\leqslant (1-r)^{ed^{d/2}2^{3d}\mathrm{Vol}(B_j^{(\star)})} \\ &\quad\leqslant \exp\big(-ped^{d/2}2^{3d} \min\big\{\mathrm{Vol}(B_j^{(\star)}), \beta^\alpha \underline{w}^\alpha (4\sqrt{d})^{-\alpha d} \mathrm{Vol}(B_j^{(\star)})^{1-\alpha}\big\}\big). \end{aligned}$$

Since $\alpha \geqslant 1$, and $\mathcal{L}$ is s-expandable, we can use the upper bound in (5.3.13) on $\mathrm{Vol}(B_j^{(\star)})$ to bound the second term in the minimum on the right-hand side of the last row, and we use $\mathrm{Vol}(B_j^{(\star)}) \geqslant 1$ by Observation 5.3.9(i) to bound the first term. We obtain

$$\begin{aligned} &\mathbb{P}\big(\nexists (x_v, w_v) \in \mathcal{L}_j^{(\star)} : (x_u, w_u) \leftrightarrow (x_v, w_v)\big) \\ &\quad\leqslant \exp\big(-ped^{d/2}2^{3d} \min\big\{1, \beta^\alpha \underline{w}^\alpha (4\sqrt{d})^{-\alpha d} s^{1-\alpha} (d^{-d/2}2^{-3d})^{1-\alpha}\big\}\big) \\ &\quad= \exp(-pe \min\{d^{d/2}3^d, (2^d\beta)^\alpha \underline{w}^\alpha s^{1-\alpha}\}) \leqslant \exp(-ep) \leqslant p/2, \end{aligned}$$

where we used in the last row that $d^{d/2}3^d > 1$ and also that the bound on $\underline{w}$ in (5.3.1) ensures that the second term inside the minimum is at least 1, and that $\exp(-ep) \leqslant p/2$ for $p \in [0,1]$. This concludes the proposition for $\alpha < \infty$.

Case (2): $\alpha = \infty$. Using the same bounds as for $\alpha < \infty$ on the distance, mark and volume of boxes, but now (5.1.3) for $\alpha = \infty$, for any $u = (x_u, w_u) \in B_j^{(\star)} \times [\underline{w}, \infty)$ and any $v = (x_v, w_v) \in \mathcal{L}_j^{(\star)}$ that

$$\begin{aligned} p\big((x_u, w_u), (x_v, w_v)\big) &\geqslant p\mathbb{1}\{\beta\underline{w} > (4\sqrt{d})^d \mathrm{Vol}(B_j^{(\star)})\} \\ &\geqslant p\mathbb{1}\{\beta\underline{w} > (4\sqrt{d})^d d^{-d/2}2^{-3d}s\} \\ &= p\mathbb{1}\{\beta\underline{w} > 2^{-d}s\} = p \geqslant p/2, \end{aligned}$$

where in the one-but-last step we used (5.3.1), finishing the proof of $\alpha = \infty$. □

We observe that the case $\alpha = \infty$ does not use of the size of $\mathcal{L}_j^{(\star)}$, and only requires a single vertex in it, which is intuitive considering the threshold nature of p in (5.1.3). It remains to prove Proposition 5.3.8, which is the content of the following subsection.

5.3.2 *The algorithm producing the cover expansion*

Now we give the algorithm producing the expanded cover of any discrete set $\mathcal{L}$ without a proper cover, thence, proving the desired proposition.

Setup for the algorithm.

Recall the notation from Definitions 5.3.4 and 5.3.7. Throughout, we will assume that $\mathcal{L}$ does not admit a proper cover in Definition 5.3.6 and that $(B_{z_i} : i \in [m])$ are the cells satisfying (5.3.5). Contrary to Definition 5.3.7, which allocates the initial cells B_{z_i} *to boxes* $B_j^{(\star)}$, the algorithm allocates the labels i, $i \leqslant m$ of the initial cells B_{z_i} *towards each other* in discrete rounds $r \in \mathbb{N}$. We write $i \overset{r}{\mapsto} j$ to indicate that label i is allocated to label j in the allocation of round r. We also write

$$\overset{r}{\mapsto} := \{(i,j) : i \overset{r}{\mapsto} j\}_{i \leqslant m};$$
$$\mathrm{Cells}_j^{(r)} := \bigcup_{i \leqslant m} \{i : i \overset{r}{\mapsto} j\};$$
$$\mathcal{J}^{(r)} := \{j : \mathrm{Cells}_j^{(r)} \neq \emptyset\}.$$

In each round $r \geqslant 0$, the boxes $\{B_j^{(r)}\}_{j \in \mathcal{J}^{(r)}}$, and the centers of these boxes are completely determined by $\overset{r}{\mapsto}$ by the formula

$$B_j^{(r)} := \Lambda\Big(z_j, \frac{1}{e d^{d/2} 2^{3d}} \sum_{i \in \mathrm{Cells}_j^{(r)}} \ell_i\Big) \quad \textit{for} \quad j \in \mathcal{J}^{(r)}; \tag{5.3.15}$$

where $\Lambda(x, s)$ is a box of volume s centered at $x \in \mathbb{R}^d$ (see (5.1.1)). Since label j corresponds to center z_j across different rounds, by slightly abusing notation we also write $B_{z_i} \overset{r}{\mapsto} B_j^{(r)}$ if and only if $i \overset{r}{\mapsto} j$. We say that $\overset{r}{\mapsto}$ satisfies one (or more) conditions in Definition 5.3.7 if $(B_j^{(r)})_{j \in \mathcal{J}^{(r)}}$ with allocation $\overset{r}{\mapsto}$ satisfies the condition(s).

The algorithm starts with the identity as initial allocation $\overset{0}{\mapsto}$ that induces possibly overlapping boxes $B_1^{(0)}, \dots, B_m^{(0)}$; we will show that $\overset{0}{\mapsto}$ already satisfies (near) and (vol.) of Definition 5.3.7. In each stage the algorithm attempts to remove an overlap – a non-empty intersection – between a pair of boxes by re-allocating a few cell labels, while maintaining properties (near) and (vol.); we achieve (disj) in the last round $r^\star$. The last round $r^\star < \infty$ corresponds to the final output, by setting $\mathcal{J}^{(\star)} := \mathcal{J}^{(r^\star)}; B_j^{(\star)} := B_j^{(r^\star)}$ and defining $B_{z_i} \overset{\star}{\mapsto} B_j^{(\star)}$ if and only if $i \overset{r^\star}{\mapsto} j$.

The cover-expansion algorithm.

(input) $(B_{z_i})_{i\in[m]}$ and $\mathcal{L}_i = \mathcal{L} \cap B_{z_i}$ satisfying (5.3.5) with $\nu = ed^{d/2}2^{3d}$ and $\delta = 1/2$.

(init.) Set $r := 0$, and allocate $j \overset{0}{\mapsto} j$ for all $j \leqslant m$.

(while) If $(B_j^{(r)})_{j\in\mathcal{J}^{(r)}}$ in (5.3.15) are all pairwise disjoint, set $r^\star := r$; and return $\mathcal{J}^{(\star)} := \mathcal{J}^{(r^\star)}; B_j^{(\star)} := B_j^{(r^\star)}$ and $\overset{\star}{\mapsto} := \overset{r^\star}{\mapsto}$.

Otherwise, let $j_1(r) \in \mathcal{J}^{(r)}$ be the label corresponding to the largest box $B_{j_1(r)}^{(r)}$ with an overlap with some other box in round r, and let $j_2(r)$ be the label of the largest box that overlaps with $B_{j_1(r)}^{(r)}$ (using an arbitrary tie-breaking rule). Define

$$\mathcal{I}_1^{(r)} := \mathrm{Cells}_{j_2(r)}^{(r)} \cap \{i : \|z_i - z_{j_1(r)}\| \leqslant \sqrt{d}\mathrm{Vol}(B_{j_1(r)}^{(r)})^{1/d}\}_{i\leqslant m}; \qquad \mathcal{I}_2^{(r)} := \mathrm{Cells}_{j_2(r)}^{(r)} \setminus \mathcal{I}_1^{(r)}. \tag{5.3.16}$$

Then we define $\overset{r+1}{\mapsto}$ by only re-allocating labels in $\mathrm{Cells}_{j_2(r)}^{(r)}$ as follows:

(i) for $i \in \mathcal{I}_1^{(r)}$ we allocate $i \overset{r+1}{\mapsto} j_1(r)$, i.e., the labels of cells in $\mathrm{Cells}_{j_2(r)}^{(r)}$ that are sufficiently close to the center of $B_{j_1(r)}^{(r)}$ in order to satisfy (5.3.9) are re-allocated to $j_1(r)$;

(ii) for $i \in \mathcal{I}_2^{(r)}$ we allocate $i \overset{r+1}{\mapsto} i$, i.e., the labels of cells that are potentially too far away from the center of $B_{j_1(r)}^{(r+1)}$ are re-allocated back to themselves;

(iii) for $i \in [m] \setminus (\mathrm{Cells}_{j_2}^{(r)})$, we set $i \overset{r+1}{\mapsto} k$ if and only if $i \overset{r}{\mapsto} k$ (that is, $\overset{r+1}{\mapsto}$ agrees with $\overset{r}{\mapsto}$ outside labels in $\mathrm{Cells}_{j_2(r)}^{(r)}$).

Increase r by one and repeat (while).

We make the immediate observation:

Observation 5.3.10. *In each iteration of (while), $\mathcal{I}_1^{(r)}$ in (5.3.16) is always non-empty. Moreover,*

$$\mathrm{Vol}\big(B_{j_1(r)}^{(r+1)}\big) - \mathrm{Vol}\big(B_{j_1(r)}^{(r)}\big) \geqslant 1. \tag{5.3.17}$$

Proof. It can be shown inductively that $j \overset{r}{\mapsto} j$ holds for all $j \in \mathcal{J}^{(r)}$. Since the boxes $B_{j_1(r)}^{(r)}$ and $B_{j_2(r)}^{(r)}$ overlap, the distance of their centers $\|z_{j_2(r)} - z_{j_1(r)}\|$ is at most the diameter of $B_{j_1(r)}^{(r)}$, which is $\sqrt{d}\mathrm{Vol}(B_{j_1(r)}^{(r)})^{1/d}$. Hence, $j_2(r) \in \mathcal{I}_1^{(r)}$. Since each cell contains $\ell_i \geqslant ed^{d/2}2^{3d}$ many vertices by the assumption in (input), we obtain by (5.3.15)

$$\mathrm{Vol}\big(B_{j_1(r)}^{(r+1)}\big) - \mathrm{Vol}\big(B_{j_1(r)}^{(r)}\big) \geqslant \frac{\ell_{j_2(r)}}{ed^{d/2}2^{3d}} \geqslant 1. \qquad \square$$

Proof of Proposition 5.3.8. Once having shown that a cover expansion of $\mathcal{L}$ exists, the bound on its volume (5.3.11) holds by Observation 5.3.9(v). So it remains to show that the algorithm produces in finitely many rounds an output satisfying all conditions of a cover expansion given in Definition 5.3.7.

The algorithm stops in finitely many rounds. We argue using a monotonicity argument. We say that a vector $\boldsymbol{a} = (a_1, \dots, a_m) \in \mathbb{R}^m$ is non-increasing if $a_i \geqslant a_{i+1}$ for all $i \leqslant m-1$. We use the lexicographic ordering for non-increasing vectors $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^m$: let $\boldsymbol{a} >_L \boldsymbol{b}$ if there exists a coordinate $j \leqslant m$ such that $a_\ell = b_\ell$ for all $\ell < j$ and $a_\ell > b_\ell$ for $\ell = j$.

For all $r \in \mathbb{N}$, $\mathcal{J}^{(r)} \subseteq [m]$, and hence, $m^{(r)} := |\mathcal{J}^{(r)}| \leqslant m$. Let $\boldsymbol{a}^{(r)} \in \mathbb{R}^m$ be the non-increasing vector of the re-ordered $(\mathrm{Vol}(B_j^{(r)}))_{j \in \mathcal{J}^{(r)}}$ appended with $(m - m^{(r)})$-many zeroes. By Observation 5.3.10, the entry corresponding to $\mathrm{Vol}\big(B_{j_1(r)}^{(r)}\big)$ in $\boldsymbol{a}^{(r)}$ increases in $\boldsymbol{a}^{(r)}$. Moreover, the entry corresponding to $\mathrm{Vol}\big(B_{j_2(r)}^{(r)}\big)$ "crumbles" into smaller volumes (corresponding to boxes with label in $\mathcal{I}_2^{(r)}$) that appear in $\boldsymbol{a}^{(r)}$. Since by definition, $j_1(r)$ corresponds to the largest box among $(B_j^{(r)})_{j \in \mathcal{J}^{(r)}}$ that has an overlap with some other box, so also $\mathrm{Vol}(B_{j_2(r)}^{(r)}) \leqslant \mathrm{Vol}(B_{j_1(r)}^{(r)})$, and the allocation of labels except those in $\mathrm{Cells}_{j_2(r)}^{(r)}$ remains unchanged, these together imply that $\boldsymbol{a}^{(r+1)} >_L \boldsymbol{a}^{(r)}$. Finally, for any r and any $j \in \mathcal{J}^{(r)}$, $\mathrm{Vol}(B_j^{(r)}) \leqslant |\mathcal{L}|/(ed^{d/2}2^{3d}) =: b$ by (5.3.8), implying that for all r, $(b, \dots, b) >_L \boldsymbol{a}^{(r)}$. So, $(\boldsymbol{a}^{(r)})_{r \geqslant 0}$ is an increasing bounded sequence with respect to $>_L$, with an increase of at least 1 per step by (5.3.17). Hence, $(\boldsymbol{a}^{(r)})_{r \geqslant 0}$ converges and attains its limit after finitely many rounds, i.e., $r^\star < \infty$.

The output corresponds to a cover expansion. We now prove that the output $\mathcal{J}^{(\star)}$, $\overset{\star}{\mapsto}$ and the corresponding boxes in (5.3.15) satisfy the conditions of

Definition 5.3.7. By the stopping condition in step (while) of the algorithm, $(B_j^{(\star)})_{j\in\mathcal{J}^{(\star)}}$ satisfy (disj.), and by their definition in (5.3.15), also (vol.). We need to still verify (near). We show this by induction: initially, for $(B_j^{(0)})_{j\in\mathcal{J}^{(0)}}, \stackrel{0}{\mapsto}$, (near) holds, since in (init.) all labels are allocated to themselves, so $\mathrm{Cells}_j^{(0)} = \{j\}$, and thus the left-hand side in (5.3.9) is 0. Assume then $r > 0$. We prove that (near) holds for $\stackrel{r+1}{\mapsto}$, assuming that it holds for $\stackrel{r}{\mapsto}$. Recall from (while) that $j_1(r)$ is the label of the largest box that has an overlap; $j_2(r)$ is the label of the largest box overlapping with $B^{(r)}_{j_1(r)}$; by (5.3.16), $\mathcal{I}_1^{(r)}$ is the set of labels in $\mathrm{Cells}^{(r)}_{j_2(r)}$ re-allocated to $j_1(r)$, and $\mathcal{I}_2^{(r)} = \mathrm{Cells}^{(r)}_{j_2(r)} \setminus \mathcal{I}_1^{(r)}$ is the set labels allocated in round r to $j_2(r)$, and in round $r+1$ to themselves. We distinguish between four cases for the proof of the inductive step:

- Assume $i \notin (\mathrm{Cells}^{(r)}_{j_1(r)} \cup \mathrm{Cells}^{(r)}_{j_2(r)})$ and let k be such that $i \stackrel{r+1}{\mapsto} k$. By (while) part (iii), $\mathrm{Cells}_k^{(r+1)} = \mathrm{Cells}_k^{(r)}$, so by the induction hypothesis, (5.3.9) holds for $\stackrel{r+1}{\mapsto}$.
- Assume $i \in \mathrm{Cells}^{(r)}_{j_1(r)}$. Since $B^{(r)}_{j_1(r)} \subsetneq B^{(r+1)}_{j_1(r)}$ by (5.3.15) and (5.3.17), considering $\|z_i - z_{j_1(r)}\|$ and the left-hand side in (5.3.9) stay the same, while the right-hand side increases, so the inequality required for (near) still holds.
- Assume $i \in \mathcal{I}_1^{(r)} \subseteq \mathrm{Cells}^{(r)}_{j_2(r)}$. The definition of $\mathcal{I}_1^{(r)}$ in (5.3.16) forces that $\|z_i - z_{j_1(r)}\|$ satisfies (5.3.9).
- Assume $i \in \mathcal{I}_2^{(r)} = \mathrm{Cells}^{(r)}_{j_2} \setminus \mathcal{I}_1^{(r)}$: (5.3.9) holds for the same reason as for the base case, i.e., since $i \stackrel{r+1}{\mapsto} i$, $\|z_i - z_i\| = 0$ trivially satisfies (5.3.9).

Having all possible cases covered, this finishes the proof of the induction. Since $r^\star < \infty$, this finishes the proof of Proposition 5.3.8. □

5.3.3 *Poisson point processes are expandable*

We end this section by showing that a Poisson point process is typically s-expandable for s sufficiently large. Recall $\Lambda_n = [-n^{1/d}/2, n^{1/d}/2]^d$.

Lemma 5.3.11 (PPPs are expandable)**.** *Let Γ be a Poisson point process on $\mathbb{R}^d$ equipped with an absolutely continuous intensity measure μ such that $\mu(dx) \leqslant$*

$\mathrm{Leb}(dx)$. *Then there exists a constant* $C_{5.3.11} = C_{5.3.11}(d) > 0$ *such that for any* $s > 0$,

$$\mathbb{P}\big(\Gamma \cap \Lambda_n \textit{ is not s-expandable}\big) \leqslant C_{5.3.11} n \exp(-s).$$

Proof. Using stochastic domination of point processes, without loss of generality we can assume that Γ has intensity measure $\mathrm{Leb}(dx)$. Let us define $R(s) := \{s + \ell/(ed^{d/2}2^{3d}) : \ell \in \mathbb{N}\}$, the range of volumes of boxes that we need to consider in Definition 5.3.1. By a union bound over the at most n possible centers of the boxes in Λ_n, and by translation invariance of Leb, it holds that

$$\begin{aligned}\mathbb{P}\big(&\Gamma \cap \Lambda_n \text{ is not s-expandable}\big)\\ &= \mathbb{P}\big(\exists x \in \mathbb{Z}^d \cap \Lambda_n, \exists s' \in R(s) : |\Lambda_{s'}(x) \cap \Gamma| \geqslant es'\big)\\ &\leqslant n \sum_{s' \in R(s)} \mathbb{P}\big(|\Lambda_{s'} \cap \Gamma| \geqslant es'\big).\end{aligned} \tag{5.3.18}$$

Since the intensity of Γ is equal to one, by Lemma 5.C.1, each summand is at most $\exp(-s')$ on the right-hand side. Hence, using that $\sum_{\ell=a}^{\infty} f(\ell) \leqslant \int_{a-1}^{\infty} f(x)dx$ for a monotone non-increasing function, we obtain for the summation in (5.3.18)

$$\begin{aligned}\mathbb{P}\big(\Gamma \cap \Lambda_n \text{ is not s-expandable}\big) &\leqslant n\exp(-s) \sum_{\ell=0}^{\infty} \exp\big(-\ell/(ed^{d/2}2^{3d})\big)\\ \leqslant n\exp(-s) \int_{-1}^{\infty} &\exp\big(-x/(ed^{d/2}2^{3d})\big)dx\\ = n\exp(-s)ed^{d/2}2^{3d} &\exp\big(1/(ed^{d/2}2^{3d})\big).\end{aligned}$$

The proof of the lemma follows for $C_{5.3.11} = ed^{d/2}2^{3d}\exp\big(1/(ed^{d/2}2^{3d})\big)$. □

5.4 UPPER BOUND: SECOND-LARGEST COMPONENT

The main goal of this section is to prove the following proposition for general values of n and k, which readily implies Theorem 5.1.4(ii), i.e., (5.1.10). Recall $\zeta_{hh} = 1 - \gamma_{hh}(\tau - 1)$ from (4.2.4). We restrict ourselves to $\sigma \in (\tau - 2, \tau - 1]$, since $\sigma \leqslant \tau - 2$ implies that $\zeta_{hh} \leqslant 0$. We recall from Section 5.2.3 that the superscript $\star$ below indicates that the result does not generalise to i-KSRGs on $\mathbb{Z}^d$.

Proposition⋆ 5.4.1. *Consider a supercritical i-KSRG model in Definition 5.1.1, and parameters* $\alpha > 1, \tau > 2$ *and* $d \in \mathbb{N}$. *Assume further that also* $\sigma \in (\tau - 2, \tau - 1]$. *There exists a constant* $c_{5.4.1} > 0$ *such that whenever* $n > k^{1+\zeta_{hh}/\alpha}$ *it holds that*

$$\mathbb{P}\big(|\mathcal{C}_n^{(2)}| \geqslant k\big) \leqslant (n/c_{5.4.1})\exp\big(-c_{5.4.1}k^{\zeta_{hh}}\big). \tag{5.4.1}$$

We follow the steps of the methodology from Section 5.2.1.1. The bulk of the work is to establish Steps 1 and 3, since we already developed the cover expansion of Step 4 in Section 5.3. We first introduce some notation. We aim to partition the box Λ_n into disjoint subboxes of (roughly) volume k. Define

$$n' := k\lfloor (n/k)^{1/d}\rfloor^d. \tag{5.4.2}$$

The box $\Lambda_{n'} \subseteq \Lambda_n$ is the largest box inside Λ_n that can be partitioned into n'/k disjoint subboxes of volume exactly k (boundaries are allocated uniquely similar to Definition 5.3.3). Let the boxes of this partitioning of $\Lambda_{n'}$ be $\mathcal{Q}_1, \dots, \mathcal{Q}_{n'/k}$, labeled so that $\mathcal{Q}_i$ shares a boundary (that is, a $(d-1)$-dimensional face) with $\mathcal{Q}_{i+1}$ for all $i < n'/k$. Define for each $u = (x_u, w_u) \in \Xi_n \subset \Lambda_n$

$$\mathcal{Q}(u) := \arg\min_{\mathcal{Q}_i} \|x_u - \mathcal{Q}_i\|, \tag{5.4.3}$$

with the convention that $\|x_u - \mathcal{Q}_i\| = 0$ if $x_u \in \mathcal{Q}_i$, and take the box with the smallest index if the minimum is non-unique. Similarly to (5.3.4), we observe that for any point $u \in \Xi_n \subset \Lambda_n$

$$\sup_{y\in\mathcal{Q}(u)} \|x_u - y\| \leqslant 2\sqrt{d}k^{1/d}. \tag{5.4.4}$$

Step 1. Construction of the backbone

Recall the definition of $\mathcal{G}_n[a,b)$ from (5.1.4) in Definition 5.1.1. We first show that, for some $w_{hh} = w_{hh}(k)$, the graph $\mathcal{G}_{n,1} := \mathcal{G}_n[w_{hh}, 2w_{hh})$ contains a so-called backbone, a connected component $\mathcal{C}_{bb}$ that contains at least $s_k = \Theta(k^{\zeta_{hh}})$ vertices in every subbox. Using α and β from Definition 5.1.1, define the constant C_1 to be the solution of the equation

$$(p/16)\beta^\alpha C_1^{1-(\sigma+1)\alpha/(\tau-1)}(2\sqrt{d})^{-\alpha d} = \log(2), \qquad \text{if } \alpha < \infty, \tag{5.4.5}$$

$$\beta C_1^{-(1+\sigma)/(\tau-1)} d^{-d/2} 2^{-d-2\sigma} = 1, \qquad \text{if } \alpha = \infty. \tag{5.4.6}$$

We set, with γ_{hh} from (4.2.3),

$$\begin{aligned} w_{hh} &:= w_{hh}(k) := C_1^{-1/(\tau-1)} k^{\gamma_{hh}}, \\ s_k &:= (C_1/16) k^{1-\gamma_{hh}(\tau-1)} = k w_{hh}^{-(\tau-1)}/16. \end{aligned} \tag{5.4.7}$$

To avoid cumbersome notation, we assume that $s_k \in \mathbb{N}$. Recall the notation $\Xi_{\mathcal{Q}}[a, b)$ from (5.1.12). Let

$$\mathcal{A}_{bb} := \mathcal{A}_{bb}(n,k) := \left\{ \begin{array}{l} \mathcal{G}_{n,1} \text{ contains a conn. component } \mathcal{C}_{bb}(n,k) \text{ s.t.} \\ \text{for all } i \leqslant (n'/k) : |\Xi_{\mathcal{Q}_i}[w_{hh}, 2w_{hh}) \cap \mathcal{C}_{bb}| \geqslant s_k \end{array} \right\}. \tag{5.4.8}$$

On $\mathcal{A}_{bb}$, let $\mathcal{C}_{bb} := \mathcal{C}_{bb}(n,k)$, the backbone, be the largest component in $\mathcal{G}_n[w_{hh}, 2w_{hh})$ that satisfies the event $\mathcal{A}_{bb}$. In the following lemma we obtain a lower bound on the probability that there exists a backbone. Observe that $s_k = \Theta(k^{\zeta_{hh}})$ in (5.4.7), so the decay on the right-hand side in (5.4.9) below is of the same order as the desired decay in Proposition 5.4.1.

Lemma 5.4.2 (Backbone construction). *Consider a supercritical i-KSRG model as in Definition 5.1.1, and parameters* $\alpha > 1$, $\tau \in (2, 2+\sigma)$, $\sigma \geqslant 0$, $d \in \mathbb{N}$. *There exist constants* $c_{5.4.2} = c_{5.4.2}(p, \beta, d, \alpha, \tau, \sigma) > 0$ *and* $k_1 \in \mathbb{N}$ *such that for* $k \geqslant k_1$ *and all* n *satisfying* $n \geqslant k^{1+\zeta_{hh}/\alpha}$ *we have*

$$\mathbb{P}\big(\neg \mathcal{A}_{bb}(n,k)\big) \leqslant 3(n/k) \exp\big(-c_{5.4.2} s_k\big). \tag{5.4.9}$$

Proof. Towards proving (5.4.9), we reveal $\Xi_n[w_{hh}, 2w_{hh})$, i.e., *only* the vertex set of $\mathcal{G}_{n,1}$, and define

$$\mathcal{A}_{poi} := \{\forall i \leqslant n'/k : |\Xi_{\mathcal{Q}_i}[w_{hh}, 2w_{hh})| \geqslant 4 s_k\}. \tag{5.4.10}$$

On $\mathcal{A}_{poi}$, every box contains enough vertices in $\mathcal{G}_{n,1}$. Reveal now the edges of $\mathcal{G}_{n,1}$ *only within each of the boxes* $(\mathcal{Q}_i)_{i \leqslant n'/k}$: let $\mathcal{H}_{i,1}$ be the induced subgraph of $\mathcal{G}_{n,1}$ on $\Xi_{\mathcal{Q}_i}[w_{hh}, 2w_{hh})$, and define

$$J := \min\{i : \mathcal{H}_{i,1} \text{ contains a component } \mathcal{C} \text{ with } |\mathcal{C}| \geqslant s_k\}. \tag{5.4.11}$$

We write $J = \infty$ if no such box-index exists. Then

$$\begin{aligned} \mathbb{P}(\neg \mathcal{A}_{bb}) \leqslant\ & \mathbb{P}(\neg \mathcal{A}_{poi}) + \mathbb{P}(J = \infty \mid \mathcal{A}_{poi}) \\ & + \sum_{i=1}^{n'/k} \mathbb{P}(J = i \mid \mathcal{A}_{poi}) \mathbb{P}(\neg \mathcal{A}_{bb} \mid \{J = i\} \cap \mathcal{A}_{poi}). \end{aligned} \tag{5.4.12}$$

We first bound $\mathbb{P}(\neg\mathcal{A}_{\mathrm{poi}})$ from above. The distribution of $|\Xi_{\mathcal{Q}_i}[w_{\mathrm{hh}}, 2w_{\mathrm{hh}})|$ is Poisson[1] with mean $kw_{\mathrm{hh}}^{-(\tau-1)}(1-2^{-(\tau-1)}) = 16(1-2^{-(\tau-1)})s_k \geqslant 8s_k$ by (5.1.2), (5.4.7) and since $\tau \geqslant 2$. Lemma 5.C.1 yields

$$\mathbb{P}\big(|\Xi_{\mathcal{Q}_i}[w_{\mathrm{hh}}, 2w_{\mathrm{hh}})| \leqslant 4s_k\big) \leqslant \mathbb{P}\big(\mathrm{Poi}(8s_k) \leqslant 4s_k\big) \leqslant \exp\big(-4s_k(1-\log(2))\big).$$

Since $1-\log(2) \geqslant 1/4$, by a union bound over the at most $n'/k \leqslant n/k$ subboxes we get

$$\mathbb{P}(\neg\mathcal{A}_{\mathrm{poi}}) \leqslant (n/k)\exp(-s_k). \tag{5.4.13}$$

We will now show an upper bound on the summands in (5.4.12) that holds uniformly in i. For this, we iteratively 'construct' a backbone. The subboxes $\mathcal{Q}_1, \ldots, \mathcal{Q}_{n'/k}$ are ordered so that $\mathcal{Q}_i$ and $\mathcal{Q}_{i+1}$ share a boundary for all i. On $\{J = i\}$, we know that $\mathcal{H}_{i,1}$ inside $\mathcal{Q}_i$ contains a connected component $\mathcal{C}$ with at least s_k many vertices. We now reveal edges between $\mathcal{Q}_i$ and $\mathcal{Q}_{i+1}$, and bound the probability that there are at least s_k many vertices in $\mathcal{Q}_{i+1}$ that are connected by an edge to $\mathcal{C}$: denote this set of vertices by $\widetilde{\mathcal{V}}_{i+1}$. Next, we apply the same bound to show that at least s_k many vertices in $\mathcal{Q}_{i+2}$ connect by an edge to $\widetilde{\mathcal{V}}_{i+1}$, and so on. For $\ell < i$, we proceed similarly. Although $\{J = i\}$ implies that the *induced* graph $\mathcal{H}_{\ell,1} \in \mathcal{Q}_\ell$ does not contain a large enough connected component, we can, thanks to conditioning on $\mathcal{A}_{\mathrm{poi}}$, still ensure that at least s_k many vertices in $\mathcal{Q}_{i-1}$ connect directly by an edge to the connected component $\mathcal{C}$ in $\mathcal{Q}_i$, *irrespective of the vertex positions in* $\mathcal{Q}_{i-1}$. Again, denote these vertices by $\widetilde{\mathcal{V}}_{i-1}$. We repeat this procedure for $\ell \in \{i-2, \ldots, 1\}$. Hence, for $\ell \geqslant i$, we need to analyse the probability that a vertex in $\mathcal{Q}_{\ell+1}$ connects to a vertex in $\widetilde{\mathcal{V}}_\ell$, conditionally on $|\widetilde{\mathcal{V}}_\ell| \geqslant s_k$. Since by assumption $\tau < 2+\sigma$, by definition of γ_{hh} in (4.2.3) for all $\tau < \sigma + 2$ and $\alpha \leqslant \infty$, it holds that

$$1-(1+\sigma)\gamma_{\mathrm{hh}} \geqslant 0, \quad \text{and} \quad 2+\sigma-\tau > 0. \tag{5.4.14}$$

Let c_d be the volume of the unit d-dimensional ball. Then let k_1 be the smallest integer so that, depending on the value of α, the following bounds hold (with $C_1 = C_1(\alpha)$ from (5.4.5)-(5.4.6), respectively):

$$\begin{aligned} &(1-p)^{C_1 k_1^{\zeta_{\mathrm{hh}}}/16} \leqslant 1/2, \quad \text{and} \\ &\quad (C_1/4)k_1^{\zeta_{\mathrm{hh}}} \leqslant \exp\big(p\beta c_d 2^{-d}\big(C_1^{-1/(\tau-1)}k_1\big)^{(2+\sigma-\tau)\gamma_{\mathrm{hh}}}/8\big). \end{aligned} \tag{5.4.15}$$

The Euclidean distance between vertices in neighbouring boxes is at most $2\sqrt{d}k^{1/d}$ (twice the diameter of a single box), and all considered vertices

1 In fact, since Ξ is conditioned to contain 0 in Definition 5.1.1, the density of Ξ in $\mathcal{Q}(0)$ is slightly different. This effect is negligible, and we ignore it in our computations throughout.

have mark at least w_{hh}. When $\alpha = \infty$, we use that $\gamma_{\mathrm{hh}} = 1/(1+\sigma)$ (see (4.2.3)), and so $w_{\mathrm{hh}}^{1+\sigma}/k = \Theta(1)$ by (5.4.7). We obtain using (5.4.6), p in (5.1.3), that for any $u = (x_u, w_u) \in \Xi_{\mathcal{Q}_{\ell+1}}[w_{\mathrm{hh}}, 2w_{\mathrm{hh}})$,

$$\begin{aligned}\mathbb{P}\big((x_u, w_u) \leftrightarrow \widetilde{\mathcal{V}}_\ell \mid |\widetilde{\mathcal{V}}_\ell| \geqslant s_k\big) &\geqslant 1 - \Big(1 - p\mathbb{1}\Big\{\frac{\beta w_{\mathrm{hh}}^{1+\sigma}}{(2\sqrt{d})^d k} \geqslant 1\Big\}\Big)^{s_k} \\ &= 1 - (1-p)^{s_k} \geqslant 1/2,\end{aligned} \tag{5.4.16}$$

for all $k \geqslant k_1$ by the first criterion in (5.4.15). When $\alpha < \infty$, using p in (5.1.3) for $u \in \mathcal{Q}_{\ell+1}$ with $w_u \geqslant w_{\mathrm{hh}}$, either the minimum is at 1 below in (5.4.17) (in which case the right-hand side of (5.4.16) remains valid) or, the minimum in p is attained at the second term below in (5.4.17): then we substitute s_k from (5.4.7),

$$\begin{aligned}&\mathbb{P}\big((x_u, w_u) \leftrightarrow \widetilde{\mathcal{V}}_\ell \mid |\widetilde{\mathcal{V}}_\ell| \geqslant s_k\big) \\ &\qquad \geqslant 1 - \big(1 - p\min\big\{1, \beta^\alpha (2\sqrt{d})^{-\alpha d} w_{\mathrm{hh}}^{(1+\sigma)\alpha} k^{-\alpha}\big\}\big)^{s_k} \qquad (5.4.17)\\ &\qquad = 1 - \big(1 - p\beta^\alpha (2\sqrt{d})^{-\alpha d} w_{\mathrm{hh}}^{(1+\sigma)\alpha} k^{-\alpha}\big)^{k w_{\mathrm{hh}}^{-(\tau-1)}/16} \\ &\qquad \geqslant 1 - \exp\big(-(p/16)\beta^\alpha (2\sqrt{d})^{-\alpha d} w_{\mathrm{hh}}^{(1+\sigma)\alpha-(\tau-1)} k^{1-\alpha}\big).\end{aligned}$$

By choice of w_{hh}, and γ_{hh} and C_1 and as defined in (5.4.7), (5.4.5), and (4.2.3), respectively, factors containing k cancel, and after simplification we arrive at

$$\begin{aligned}\mathbb{P}\big(u \leftrightarrow \widetilde{\mathcal{V}}_\ell \mid |\widetilde{\mathcal{V}}_\ell| \geqslant s_k\big) &\geqslant 1 - \exp\big(-(p/16)\beta^\alpha (2\sqrt{d})^{-\alpha d} C_1^{1-(1+\sigma)\alpha/(\tau-1)}\big) \\ &= 1/2.\end{aligned} \tag{5.4.18}$$

Combining (5.4.18) with (5.4.16), we obtain a lower bound of $1/2$ for all $\alpha > 1$ for any $u \in \Xi_{\mathcal{Q}_{\ell+1}}[w_{\mathrm{hh}}, 2w_{\mathrm{hh}})$. On $\mathcal{A}_{\mathrm{poi}}$ (see (5.4.10)) there are at least $4s_k$ vertices in $\Xi_{\mathcal{Q}_{\ell+1}}[w_{\mathrm{hh}}, 2w_{\mathrm{hh}})$. Each of these vertices connects conditionally independently by an edge to vertices in $\widetilde{\mathcal{V}}_\ell$ with probability at least $1/2$, so for all $\ell \geqslant i$

$$\begin{aligned}\mathbb{P}\big(|\widetilde{\mathcal{V}}_{\ell+1}| \geqslant s_k \mid |\widetilde{\mathcal{V}}_\ell| \geqslant 4s_k, \mathcal{A}_{\mathrm{poi}}\big) &\geqslant \mathbb{P}\big(\mathrm{Bin}(4s_k, 1/2) \geqslant s_k\big) \\ &\geqslant 1 - \exp(-s_k/4),\end{aligned}$$

where the last bound follows by Chernoff's bound, see e.g. [145, Theorem 2.1]. When $J = i$ and $\ell < i$, we can analogously bound the probability that $|\widetilde{\mathcal{V}}_\ell| \geqslant s_k$, conditionally on $|\widetilde{\mathcal{V}}_{\ell+1}| \geqslant s_k$. By a union bound over the at most n'/k subboxes (with indices both smaller as well as larger than i), we obtain

$$\mathbb{P}\big(\neg\mathcal{A}_{bb} \mid \{J = i\} \cap \mathcal{A}_{poi}\big) \leqslant (n'/k)\exp(-s_k/4) \\ \leqslant (n/k)\exp(-s_k/4).$$

The bound holds for all $i \leqslant n'/k$. Using this in the summation on the right-hand side of (5.4.12), and using (5.4.13) to bound $\mathbb{P}(\neg\mathcal{A}_{poi})$, (5.4.12) turns into

$$\begin{aligned}\mathbb{P}\big(\neg\mathcal{A}_{bb}\big) &\leqslant (n/k)\exp(-s_k) + \mathbb{P}(J = \infty \mid \mathcal{A}_{poi}) \\ &\quad + (n/k)\exp(-s_k/4)\mathbb{P}(J \neq \infty \mid \mathcal{A}_{poi}) \\ &\leqslant 2(n/k)\exp(-s_k/4) + \mathbb{P}(J = \infty \mid \mathcal{A}_{poi}).\end{aligned} \tag{5.4.19}$$

It remains to bound $\mathbb{P}(J = \infty \mid \mathcal{A}_{poi})$, with J from (5.4.11). For this we show that the graph $\mathcal{H}_{i,1}$ induced on $\Xi_{\mathcal{Q}_i}[w_{hh}, 2w_{hh})$ stochastically dominates a soft random geometric graph above its connectivity threshold. Indeed, consider (x_u, w_u) and (x_v, w_v) in $\Xi_{\mathcal{Q}_i}[w_{hh}, 2w_{hh})$. Then using (5.1.3), for $\alpha = \infty$,

$$p\big((x_u, w_u), (x_v, w_v)\big) \geqslant p \min\big\{1, \beta^{\alpha} w_{hh}^{(1+\sigma)\alpha}\|x_u - x_v\|^{-d\alpha}\big\} \geqslant p,$$

whenever $\|x_u - x_v\| \leqslant \beta^{1/d} w_{hh}^{(1+\sigma)/d} =: r_{hh}$. The calculation for $\alpha = \infty$ is analogous yielding the same radius r_{hh} and same lower bound p. Hence, $\mathcal{H}_{i,1} \succcurlyeq \mathrm{RG}_i$ where in RG_i is a soft random geometric graph: conditioned on $\mathcal{A}_{poi}$, there are say $N_i \geqslant 4s_k$ many vertices in $\Xi_{\mathcal{Q}_i}[w_{hh}, 2w_{hh})$, and two vertices are connected by an edge with probability p whenever they are within distance $r_{hh} = \beta^{1/d} w_{hh}^{(1+\sigma)/d}$ of each other. Writing B_r for the Euclidean ball of radius r around the origin and for some dimension dependent constant c_d, the expected degree of a vertex at *any* location in RG_i is thus, conditional on N_i:

$$\mathbb{E}[\deg_{\mathrm{RG}_i}(v) \mid N_i] \geqslant pN_i \frac{\mathrm{Vol}(B_{r_{hh}})/2^d}{k} = N_i \cdot pc_d\beta 2^{-d} \cdot w_{hh}^{1+\sigma}/k. \tag{5.4.20}$$

We apply [212, Equation (2.6) below Theorem 2.1] to RG_i, and combine it with $\mathcal{H}_{i,1} \succcurlyeq \mathrm{RG}_i$ to obtain

$$\mathbb{P}(\mathcal{H}_{i,1} \text{ not connected} \mid N_i) \leqslant \mathbb{P}(\mathrm{RG}_i \text{ not connected} \mid N_i) \\ \leqslant N_i \exp\big(-\mathbb{E}[\deg_{\mathrm{RG}_i}(v) \mid N_i]\big).$$

On the event $\mathcal{A}_{\mathrm{poi}}$, $N_i \geqslant 4s_k$ holds for all i. Hence, by (5.4.20), uniformly for all $i \leqslant n'/k$,

$$\begin{aligned}\mathbb{P}(\mathcal{H}_{i,1} \text{ not connected} \mid \mathcal{A}_{\mathrm{poi}}) &\leqslant \sup_{N_i \geqslant 4s_k} N_i \exp\big(-\mathbb{E}[\deg_{\mathrm{RG}_i}(v) \mid N_i]\big) \\ &\leqslant 4s_k \exp\big(-4s_k \cdot pc_d\beta 2^{-d} \cdot w_{\mathrm{hh}}^{1+\sigma}/k\big) \\ &\leqslant (C_1/4)k^{\zeta_{\mathrm{hh}}} \exp\big(-pc_d\beta w_{\mathrm{hh}}^{2+\sigma-\tau}/4\big),\end{aligned}$$

where we used s_k from (5.4.7) to obtain the last row. Using the second criterion in (5.4.15), for all $k \geqslant k_1$,

$$\mathbb{P}(\mathcal{H}_{i,1} \text{ not connected} \mid \mathcal{A}_{\mathrm{poi}}) \leqslant \exp(-pc_d\beta 2^{-d} w_{\mathrm{hh}}^{2+\sigma-\tau}/8) =: p_1.$$

By (5.4.14), the exponent of w_{hh} is positive. Since the induced subgraphs $(\mathcal{H}_{i,1})_{i \leqslant n'/k}$ are independent (being contained in disjoint boxes $\mathcal{Q}_i$), and the number of boxes is $n'/k = \lfloor (n/k)^{1/d} \rfloor^d \geqslant n/(2k)$, using the definition of w_{hh} from (5.4.7), we get

$$\begin{aligned}\mathbb{P}(J=\infty \mid \mathcal{A}_{\mathrm{poi}}) &\leqslant \mathbb{P}(\forall i \leqslant n'/k : \mathcal{H}_{i,1} \text{ not connected} \mid \mathcal{A}_{\mathrm{poi}}) \leqslant p_1^{n/(2k)} \\ &\leqslant \exp\big(-pc_d\beta 2^{-d-4}(n/k)w_{\mathrm{hh}}^{2+\sigma-\tau}\big)\big) \\ &\leqslant \exp\big(-pc_d\beta 2^{-d-4} C_1^{-(2+\sigma-\tau)(\tau-1)}(n/k)k^{\gamma_{\mathrm{hh}}(2+\sigma-\tau)}\big).\end{aligned}$$

Using (4.2.3) and (4.2.4), it is elementary to check that $\gamma_{\mathrm{hh}}(2+\sigma-\tau) = \zeta_{\mathrm{hh}}(1-1/\alpha)$. Hence, whenever $n/k > k^{\zeta_{\mathrm{hh}}/\alpha}$, the exponent of k on the right-hand side is (at least) ζ_{hh}. Since we assumed $n \geqslant k^{1+\zeta_{\mathrm{hh}}/\alpha}$, this is indeed the case. Furthermore, since $s_k = (C_1/16)k^{\zeta_{\mathrm{hh}}}$ by (5.4.7), with $c_2 := (pc_d\beta 2^{-d+1})C_1^{-(2+\sigma-\tau)(\tau-1)-1}$, using C_1 from (5.4.5)), we obtain that $\mathbb{P}(J = \infty \mid \mathcal{A}_{\mathrm{poi}}) \leqslant \exp(-c_2 s_k)$. Combined with (5.4.19), this yields the statement of the lemma in (5.4.9) with $c_{5.4.2} := \min\{c_2, 1/4\}$. □

We will end Step 1 with a claim that shows (5.2.2), using notation for the construction of the graph $\mathcal{G}_n$ that facilitates later steps that we introduce first. We recall the definition of i-KSRG from Definition 5.1.1. Given the vertex set Ξ, it is standard practice to use independent uniform random variables to facilitate couplings with the edge-set. This definition here is more general and allows for other helping random variables as well, leading to different distributions on graphs. This will be useful later.

Definition 5.4.3 (Graph encoding). Let $\xi \subset \mathbb{R}^d \times [1,\infty)$ be a discrete set and assume that $\Psi_\xi = \big\{\varphi_{u,v} : \varphi_{u,v} \in [0,1], \{u,v\} \in \binom{\xi}{2}\big\}$ is a collection of random variables given ξ. For a given connectivity function $p : (\mathbb{R}^d \times$

$[1,\infty))^2 \to [0,1]$, we call $\mathcal{G}' = (\mathcal{V}', \mathcal{E}')$ the (sub)graph *encoded by* (ξ, Ψ_ξ, p) if $\mathcal{V}' = \xi$ and for all $\{u,v\} \in \binom{\xi}{2}$, with $u = (x_u, w_u), v = (x_v, w_v)$,

$$\big\{\{u,v\} \in \mathcal{E}'\big\} \iff \big\{\varphi_{u,v} \leqslant p\big((x_u,w_u),(x_v,w_v)\big)\big\}. \tag{5.4.21}$$

Given Ξ in (5.1.2), and p from (5.1.3), let Ψ_Ξ be a collection of independent Unif[0, 1] random variables given Ξ. $\mathcal{G}_\infty$ in Definition 5.1.1 is then the graph *encoded by* (Ξ, Ψ_Ξ, p). Writing $\Psi_n[a,b) := \big\{\varphi_{u,v} \in \Psi_\Xi : \{u,v\} \in \binom{\Xi_n[a,b)}{2}\big\}$ and $\Psi_n := \Psi_n[1,\infty)$, $\mathcal{G}_n$ in (5.1.4) is then the graph encoded by (Ξ_n, Ψ_n, p).

An immediate corollary is the following:

Corollary 5.4.4. *Assume* $\widetilde{\mathcal{G}}, \widehat{\mathcal{G}}$ *are two random graphs, encoded respectively by* $(\widetilde{\Xi}, \widetilde{\Psi}, p)$, *and* $(\widehat{\Xi}, \widehat{\Psi}, p)$ *for respective point processes* $\widetilde{\Xi}, \widehat{\Xi}$ *on* $\mathbb{R}^d \times [1,\infty)$ *using the* same *connectivity function* p. *If* $(\widetilde{\Xi}, \widetilde{\Psi})$ *and* $(\widehat{\Xi}, \widehat{\Psi})$ *have the same law then the encoded graphs* $\widetilde{\mathcal{G}}$ *and* $\widehat{\mathcal{G}}$ *also have the same law.*

The collection of (conditionally) independent *uniform* variables $\Psi_n = \{\varphi_{u,v} : \{u,v\} \in \binom{\Xi_n}{2}\}$ and the connectivity function p determine the presence of edges in $\mathcal{G}_n$. By (5.4.21), if for some $r > 0$ it holds that $\varphi_{u,v} \leqslant r \leqslant p(u,v)$, then $\{u \leftrightarrow v\}$. Writing $\mathcal{Q}(u)$ for the box containing or closest to $u \in \Xi_n$ (see (5.4.3)), let $v_u(1), v_u(2), \dots, v_u(s_k), \dots$ denote the vertices in $\mathcal{Q}(u) \cap \mathcal{C}_{bb}$, in decreasing order with respect to their marks. Let

$$\mathcal{S}(u) := \big\{v_u(1), \dots, v_u(s_k)\big\}. \tag{5.4.22}$$

Claim 5.4.5 (Connections to the backbone)**.** *Consider an i-KSRG satisfying the conditions of Proposition 5.4.1. Fix* n *and* k *and assume* $\mathcal{G}_{n,1}$ *satisfies the event* $\mathcal{A}_{bb}(n,k)$. *Let* $\Psi_n = \{\varphi_{u,v} : \{u,v\} \in \binom{\Xi_n}{2}\}$ *be a collection of i.i.d.* Unif[0, 1] *random variables and* $r_k := 1 - 2^{-1/s_k}$. *Then, for all* $u \in \Xi_n[2w_{hh}(k), \infty)$ *and* $v \in \mathcal{S}(u)$, $p(u,v) \geqslant r_k$ *and*

$$\begin{aligned}\mathbb{P}\big(\forall v \in \mathcal{S}(u) : \varphi_{u,v} > r_k \mid \mathcal{G}_{n,1}, \mathcal{A}_{bb}\big)& \\ = \mathbb{P}\big(\exists v \in \mathcal{S}(u) : \varphi_{u,v} \leqslant r_k \mid \mathcal{G}_{n,1}, \mathcal{A}_{bb}\big) &= 1/2.\end{aligned} \tag{5.4.23}$$

Proof. On the event $\mathcal{A}_{bb}$, $\mathcal{C}_{bb} \subseteq \mathcal{G}_{n,1}$ satisfies (5.4.8) and in particular $\mathcal{S}(u)$ in (5.4.22) is well-defined and has size s_k. Since $\{\varphi_{u,v}\}$ is a collection of iid Unif[0, 1] random variables, (cf. Definition 5.4.3), one *must* set $r_k := 1 - 2^{-1/s_k}$ for (5.4.23) to hold. Hence, it only remains to show $p(u,v) \geqslant r_k$ in the statement.

With $\mathcal{Q}(u)$ and $\mathcal{S}(u)$ from (5.4.3) and (5.4.22), respectively, by (5.4.4), every $u \in \Xi_n[2w_{hh}, \infty)$ is at distance at most $2\sqrt{d}k$ from any vertex in

$v \in \mathcal{S}(u)$. Then since $w_u \geqslant 2w_{hh} \geqslant w_{hh}$, and $|\mathcal{S}(u)| = s_k$, the computations (5.4.16) – (5.4.18) word-by-word carry through with $\widetilde{\mathcal{V}}_\ell$ replaced by $\mathcal{S}(u)$, obtaining

$$\mathbb{P}(u \leftrightarrow \mathcal{S}(u) \mid \mathcal{G}_{n,1}, \Xi_n, \mathcal{A}_{bb}) = 1 - \prod_{v \in \mathcal{S}(u)} (1 - p(u,v)) \geqslant 1 - (1 - z_k)^{s_k} \geqslant 1/2,$$

with z_k either equaling p in the right-hand side of (5.4.16) or the appropriate expression in the right-hand side of (5.4.17), that bound *individually* each $p(u,v)$ from below. Following now the calculations towards (5.4.18) ensures that in both cases $z_k \geqslant 1 - 2^{-1/s_k}$. The assumptions on $k \geqslant k_1$ in (5.4.15) are also needed for this (for instance that $p \geqslant r_k$). □

Step 2. Revealing low-mark vertices

Having established that $\mathcal{G}_{n,1}$ contains a backbone with sufficiently high probability, we define $\mathcal{G}_{n,2} := \mathcal{G}_n[1, 2w_{hh}) \supseteq \mathcal{G}_{n,1}$.

Step 3. Presampling the vertices connecting to the backbone

We make Step 3 of Section 5.2.1.1 precise now. Step 3 ensures that during Step 4 below no *small-to-large* merging occurs when revealing the connector vertices of $\Xi_n[2w_{hh}, \infty)$. That is, components of size smaller than k do not merge into a larger component via edges towards a connector vertex $v \in \Xi_n[2w_{hh}, \infty)$ that is not connected to the backbone $\mathcal{C}_{bb}$ ($\mathcal{C}_{bb}$ will be contained in the giant component of $\mathcal{G}_n$). So, we partially pre-sample some randomness that encodes the presence of some edges.

For a pair n, k, we now present the alternative graph-encoding $\widehat{\mathcal{G}}_n$ of KSRGs (cf. Definitions 5.1.1 and 5.4.3) and verify that $\widehat{\mathcal{G}}_n$ and $\mathcal{G}_n$ in Definition 5.1.1 have the same law. The difference between the encoding in Definition 5.4.3 and the construction of $\widehat{\mathcal{G}}_n$ is that in the latter the edge-variables $\varphi_{u,v}$ are no longer independent Unif$[0,1]$ random variables, but are sampled from a suitable (conditional) joint distribution, whenever $u \in \Xi_n[2w_{hh}(k), \infty)$ and $v \in \mathcal{S}(u)$ from (5.4.22). Recall $r_k = 1 - 2^{-1/s_k}$ from Claim 5.4.5, with $w_{hh}(k) := w_{hh}$ and s_k defined in (5.4.7).

Definition 5.4.6 (Alternative graph construction)**.** Fix n and k. Let $\mathcal{G}_{n,2} = \mathcal{G}_n[1, 2w_{hh}(k))$ from Definition 5.1.1 be the graph encoded by the tuple $(\Xi_n[1, 2w_{hh}), \Psi_n[1, 2w_{hh}), p)$. Let $\widehat{\Xi}_n^{(\mathrm{unsure})}[2w_{hh}, \infty)$ and $\widehat{\Xi}_n^{(\mathrm{sure})}[2w_{hh}, \infty)$ be two *independent* Poisson point processes on $\Lambda_n \times [2w_{hh}, \infty)$ with intensity $(1/2)\mathrm{Leb} \otimes F_W(dw)$, with F_W as in (5.1.2), and let

$$\widehat{\Xi}_n[2w_{hh}, \infty) := \widehat{\Xi}_n^{(\text{unsure})}[2w_{hh}, \infty) \cup \widehat{\Xi}_n^{(\text{sure})}[2w_{hh}, \infty). \tag{5.4.24}$$

Let $\Sigma_n := \{U_{u,v} : u \in \widehat{\Xi}_n[2w_{hh}, \infty), v \in \Xi_n[1, 2w_{hh}) \cup \widehat{\Xi}_n[2w_{hh}, \infty)\}$ be a collection of i.i.d. Unif[0, 1] random variables (conditionally on these PPPs).

(i) If $\mathcal{G}_{n,1} = \mathcal{G}_n[w_{hh}, 2w_{hh}) \subseteq \mathcal{G}_{n,2}$ does not satisfy the event $\mathcal{A}_{bb}$ in (5.4.8), then set $\widehat{\Psi}_n := \Psi_n[1, 2w_{hh}) \cup \Sigma_n$ in Definition 5.4.3 to construct $\widehat{\mathcal{G}}_n \supseteq \mathcal{G}_{n,2}$ on $\Xi_n[1, 2w_{hh}) \cup \widehat{\Xi}_n[2w_{hh}, \infty)$, i.e.,

$$\widehat{\mathcal{G}}_n := (\Xi_n[1, 2w_{hh}) \cup \widehat{\Xi}_n[2w_{hh}, \infty), \Psi_n[1, 2w_{hh}) \cup \Sigma_n, p).$$

(ii) If $\mathcal{G}_{n,1} \subseteq \mathcal{G}_{n,2}$ satisfies the event $\mathcal{A}_{bb}$, then we construct $\widehat{\mathcal{G}}_n \supseteq \mathcal{G}_{n,2}$ *conditionally on* $\mathcal{G}_{n,2}$ as follows. For each $u \in \widehat{\Xi}_n[2w_{hh}, \infty)$ in (5.4.24), the set of vertices $\mathcal{S}(u) \subseteq \Xi_n[1, 2w_{hh})$ is a deterministic function of $\mathcal{G}_{n,1} \subseteq \mathcal{G}_{n,2}$, given by (5.4.22). Let

$$\begin{aligned}
\widehat{\Psi}_n^{(\text{iid,unsure})} &:= \big\{U_{u,v} : u \in \widehat{\Xi}_n^{(\text{unsure})}[2w_{hh}, \infty), \\
&\qquad v \in \widehat{\Xi}_n^{(\text{unsure})}[2w_{hh}, \infty) \cup \Xi_n[1, 2w_{hh}) \setminus \mathcal{S}(u)\big\}, \\
\widehat{\Psi}_n^{(\text{iid,sure})} &:= \big\{U_{u,v} : u \in \widehat{\Xi}_n^{(\text{sure})}[2w_{hh}, \infty), \\
&\qquad v \in \widehat{\Xi}_n^{(\text{sure})}[2w_{hh}, \infty) \cup \Xi_n[1, 2w_{hh}) \setminus \mathcal{S}(u)\big\}, \\
\widehat{\Psi}_n^{(\text{iid,both})} &:= \big\{U_{u,v} : u \in \widehat{\Xi}_n^{(\text{sure})}[2w_{hh}, \infty), \\
&\qquad v \in \widehat{\Xi}_n^{(\text{unsure})}[2w_{hh}, \infty)\big\}
\end{aligned} \tag{5.4.25}$$

be disjoint subsets of Σ_n, and write $\widehat{\Psi}_n^{(\text{iid})} := \widehat{\Psi}_n^{(\text{iid,unsure})} \cup \widehat{\Psi}_n^{(\text{iid,sure})} \cup \widehat{\Psi}_n^{(\text{iid,both})}$ for the union. Conditionally on $\widehat{\Xi}_n^{(\text{unsure})}[2w_{hh}, \infty), \widehat{\Xi}_n^{(\text{sure})}[2w_{hh}, \infty)$ and the graph $\mathcal{G}_n[1, 2w_{hh})$, define also the collections of random variables

$$\widehat{\Psi}_n^{(\text{cond,unsure})} := \big\{\widehat{\varphi}_{u,v} : u \in \widehat{\Xi}_n^{(\text{unsure})}[2w_{hh}, \infty), v \in \mathcal{S}(u)\big\}, \tag{5.4.26}$$

$$\widehat{\Psi}_n^{(\text{cond,sure})} := \big\{\widehat{\varphi}_{u,v} : u \in \widehat{\Xi}_n^{(\text{sure})}[2w_{hh}, \infty), v \in \mathcal{S}(u)\big\}, \tag{5.4.27}$$

so that for different vertices $u_1, u_2 \in \widehat{\Xi}_n[2w_{hh}, \infty)$, the sets $\{\widehat{\varphi}_{u_1,v}\}_{v \in \mathcal{S}(u_1)}$ and $\{\widehat{\varphi}_{u_2,v'}\}_{v' \in \mathcal{S}(u_2)}$ are *independent*. The joint distribution of $\{\widehat{\varphi}_{u,v}\}_{v \in \mathcal{S}(u)}$ for a single $u \in \widehat{\Xi}_n^{(\text{unsure})}[2w_{hh}, \infty)$ is as follows: for *any* $(z_{u,v})_{v \in \mathcal{S}(u)} \in [0, 1]^{\mathcal{S}(u)}$ of length s_k, and with $r_k = 1 - 2^{-1/s_k}$,

$$\begin{aligned}
&\mathbb{P}\big(\forall v \in \mathcal{S}(u) : \widehat{\varphi}_{u,v} \leqslant z_{u,v} \mid u \in \widehat{\Xi}_n^{(\text{unsure})}[2w_{hh}, \infty)\big) \\
&\qquad := \mathbb{P}\big(\forall v \in \mathcal{S}(u) : U_{u,v} \leqslant z_{u,v} \mid \forall v \in \mathcal{S}(u) : U_{u,v} > r_k\big).
\end{aligned} \tag{5.4.28}$$

Similarly we define the joint distribution of $\{\widehat{\varphi}_{u,v}\}_{v\in\mathcal{S}(u)}$ for a single $u \in \widehat{\Xi}_n^{(\mathrm{sure})}[2w_{\mathrm{hh}},\infty)$ as follows: for any sequence $(z_{u,v})_{v\in\mathcal{S}(u)} \in [0,1]^{\mathcal{S}(u)}$ of length s_k,

$$\begin{aligned}\mathbb{P}\big(\forall v \in \mathcal{S}(u) : \widehat{\varphi}_{u,v} \leqslant z_{u,v} \mid u \in \widehat{\Xi}_n^{(\mathrm{sure})}[2w_{\mathrm{hh}},\infty)\big)\\ := \mathbb{P}\big(\forall v \in \mathcal{S}(u) : U_{u,v} \leqslant z_{u,v} \mid \exists v \in \mathcal{S}(u) : U_{u,v} \leqslant r_k\big).\end{aligned} \tag{5.4.29}$$

We define $\widehat{\mathcal{G}}_n$ as the graph encoded by $(\widehat{\Xi}_n, \widehat{\Psi}_n, p)$, where

$$\begin{aligned}\widehat{\Xi}_n &:= \Xi_n[1,2w_{\mathrm{hh}}) \cup \widehat{\Xi}_n^{(\mathrm{unsure})}[2w_{\mathrm{hh}},\infty) \cup \widehat{\Xi}_n^{(\mathrm{sure})}[2w_{\mathrm{hh}},\infty),\\ \widehat{\Psi}_n &:= \Psi_n[1,2w_{\mathrm{hh}}) \cup \widehat{\Psi}_n^{(\mathrm{iid})} \cup \widehat{\Psi}_n^{(\mathrm{cond,unsure})} \cup \widehat{\Psi}_n^{(\mathrm{cond,sure})}.\end{aligned} \tag{5.4.30}$$

An immediate corollary is the following statement.

Corollary 5.4.7. *Consider an i-KSRG $\widehat{\mathcal{G}}_n$ following Definition 5.4.6 for some n, k. On the event $\mathcal{A}_{\mathrm{bb}}(n,k)$, every vertex in $\widehat{\Xi}_n^{(\mathrm{sure})}[2w_{\mathrm{hh}},\infty)$ is connected by an edge to $\mathcal{C}_{\mathrm{bb}}(n,k)$ in $\widehat{\mathcal{G}}_n$.*

Proof. The conditioning in (5.4.29) guarantees that for all vertices $u \in \widehat{\Xi}_n^{(\mathrm{sure})}[2w_{\mathrm{hh}},\infty)$ at least one $\widehat{\varphi}_{u,v} \leqslant r_k$ occurs among the edge-variables $\{\widehat{\varphi}_{u,v} : v \in \mathcal{S}(u)\}$, where $\mathcal{S}(u) \subseteq \mathcal{C}_{\mathrm{bb}}$ (cf. (5.4.22)). Since $\widehat{\varphi}_{u,v} \leqslant r_k \leqslant p(u,v)$ holds by Claim 5.4.5, this ensures that $\{u,v\}$ is in the edge set of $\widehat{\mathcal{G}}_n$ by the graph-encoding in Definition 5.4.3. □

Proposition 5.4.8. *Fix a connectivity function* p. *The law of the random graph $\widehat{\mathcal{G}}_n$ formed by Definition 5.4.6 is identical to the law of the random graph $\mathcal{G}_n$ formed by Definition 5.4.3.*

Proof. By Corollary 5.4.4 it is sufficient to show that $(\widehat{\Xi}_n, \widehat{\Psi}_n)$ defined in (5.4.30) has the same distribution as (Ξ_n, Ψ_n) from Definition 5.4.6. By standard properties of PPPs, Ξ_n can be written as two independent PPPs defined on the same space, each having half the intensity measure of Ξ_n. Thus, by the construction of $\widehat{\Xi}_n[1,2w_{\mathrm{hh}})$ in Definition 5.4.6, it holds that the PPPs Ξ_n and $\widehat{\Xi}_n$ have the same law. Hence, we may couple the vertex sets so that a.s. $\Xi_n = \widehat{\Xi}_n$, and only need to show that the collection of variables encoding the edges, Ψ_n and $\widehat{\Psi}_n$, share the same law, conditionally under the given vertex set realization (say) $\widehat{\Xi}_n$. By (5.4.30) in Definition 5.4.6, the graph $\mathcal{G}_{n,2}$ spanned on $\Xi_n[1,2w_{\mathrm{hh}}) \subseteq \widehat{\Xi}_n$ is determined by $\Psi_n[1,2w_{\mathrm{hh}}) = \{\varphi_{u,v} : u,v \in \Xi_n[1,2w_{\mathrm{hh}})\}$ in Definition 5.4.3. Thus $\mathcal{G}_{n,2}$ has the same distribution both in Definition 5.4.3 and in Definition 5.4.6.

(i) If now $\Psi_n[1,2w_{\mathrm{hh}})$ is such that the graph $\mathcal{G}_{n,2}$ does *not* satisfy the event $\mathcal{A}_{\mathrm{bb}}$, by (i) of Definition 5.4.6, the statement holds since both

$\{\varphi_{u,v}\}$ and $\{U_{u,v}\}$ are i.i.d. uniforms whenever $u \in \widehat{\Xi}_n[2w_{hh}, \infty)$, i.e., $\widehat{\Psi}_n \setminus \Psi_n[1, 2w_{hh}) = \Sigma_n$ and $\Psi_n \setminus \Psi_n[1, 2w_{hh})$ have the same distribution.

(ii) If $\Psi_n[1, 2w_{hh})$ is such that the graph $\mathcal{G}_{n,2}$ does satisfy the event $\mathcal{A}_{bb}$, then we work conditionally on a realization of the graph $\mathcal{G}_{n,2} = (\Xi_n[1, 2w_{hh}), \Psi_n[1, 2w_{hh}), p)$, and also on the coupled realization of the PPPs $\Xi_n[2w_{hh}, \infty) = \widehat{\Xi}_n[2w_{hh}, \infty)$. Let us define the conditional probability measure (of the edges) under the coupling by

$$\mathbb{P}^\star(\cdot) := \mathbb{P}(\,\cdot \mid \mathcal{G}_{n,2}, \Xi_n[2w_{hh}, \infty)) = \mathbb{P}(\,\cdot \mid \mathcal{G}_{n,2}, \text{unlabeled } \widehat{\Xi}_n[2w_{hh}, \infty)), \tag{5.4.31}$$

where in the conditioning we do *not* reveal to which sub-PPP (either $\widehat{\Xi}_n^{(\text{sure})}[2w_{hh}, \infty)$ or $\widehat{\Xi}_n^{(\text{unsure})}[2w_{hh}, \infty)$) a vertex in $\widehat{\Xi}_n[2w_{hh}, \infty)$ belongs to. Using $\widehat{\Psi}_n$ from (5.4.30) and $\widehat{\Psi}_n^{(\text{iid})} \subseteq \Sigma_n$ from (5.4.25) (containing *independent* copies $U_{u,v}$ of $\text{Unif}[0,1]$ random variables, like Ψ in Definition 5.4.3), we see that variables in $\Psi_n \setminus \Psi_n[1, 2w_{hh})$ and $\widehat{\Psi}_n \setminus \Psi_n[1, 2w_{hh})$ also share the same (joint) law of i.i.d. $\text{Unif}[0,1]$ whenever u and v are such $u \in \widehat{\Xi}_n[2w_{hh,\infty})$ and that $v \notin \mathcal{S}(u)$. Moreover, in (5.4.26)-(5.4.27), the collections $\{\widehat{\varphi}_{u,v}\}_{v \in \mathcal{S}(u)}$ are independent across u for different vertices $u \in \widehat{\Xi}_n[2w_{hh}, \infty)$. So for $\widehat{\mathcal{G}}_n \overset{d}{=} \mathcal{G}_n$ we are left with showing that for any fixed $u \in \widehat{\Xi}_n[2w_{hh}, \infty) = \Xi_n[2w_{hh}, \infty)$, under the measure $\mathbb{P}^\star$,

$$\big\{\varphi_{u,v}, v \in \mathcal{S}(u)\big\} \overset{d}{=} \big\{\widehat{\varphi}_{u,v}, v \in \mathcal{S}(u)\big\}. \tag{5.4.32}$$

We first analyse the distribution of the left-hand side, i.e., $\varphi_{u,v}$ being i.i.d. from Definition 5.4.3. Let $(z_{u,v})_{v \in \mathcal{S}(u)} \in [0,1]^{\mathcal{S}(u)}$ be any sequence of length s_k. By Claim 5.4.5, and the law of total probability

$$\begin{aligned}
&\mathbb{P}^\star\big(\forall v \in \mathcal{S}(u) : \varphi_{u,v} \leqslant z_{u,v}\big) \qquad (5.4.33)\\
&\quad = (1/2)\mathbb{P}^\star\big(\forall v \in \mathcal{S}(u) : \varphi_{u,v} \leqslant z_{u,v} \mid \forall v \in \mathcal{S}(u) : \varphi_{u,v} > r_k\big)\\
&\qquad + (1/2)\mathbb{P}^\star\big(\forall v \in \mathcal{S}(u) : \varphi_{u,v} \leqslant z_{u,v} \mid \exists v \in \mathcal{S}(u) : \varphi_{u,v} \leqslant r_k\big).
\end{aligned}$$

We now analyse the right-hand side in (5.4.32). By the construction in (5.4.24), $\widehat{\Xi}_n[2w_{hh}, \infty)$ is the union of two i.i.d. sub-PPPs. Under $\mathbb{P}^\star$ in (5.4.31) we did not reveal to which sub-PPP vertices belong to. Hence, for each $u \in \widehat{\Xi}_n[2w_{hh}, \infty)$, independently of each other

$$\begin{aligned}
&\mathbb{P}^\star(u \in \widehat{\Xi}_n^{(\text{unsure})}[2w_{hh}, \infty) \mid u \in \widehat{\Xi}_n[2w_{hh}, \infty))\\
&\qquad = \mathbb{P}^\star(u \in \widehat{\Xi}_n^{(\text{sure})}[2w_{hh}, \infty) \mid u \in \widehat{\Xi}_n[2w_{hh}, \infty)) = 1/2.
\end{aligned}$$

Thus, by the law of total probability, and using the distributions of $(\varphi_{u,v})_{v\in\mathcal{S}(u)}$ given by (5.4.28), (5.4.29),

$$\begin{aligned}
&\mathbb{P}^\star\big(\forall v \in \mathcal{S}(u) : \widehat{\varphi}_{u,v} \leqslant z_{u,v}\big) &(5.4.34)\\
&\quad= (1/2)\mathbb{P}^\star\big(\forall v \in \mathcal{S}(u) : \widehat{\varphi}_{u,v} \leqslant z_{u,v} \mid u \in \widehat{\Xi}_n^{(\text{unsure})}[2w_{\text{hh}},\infty)\big)\\
&\qquad+ (1/2)\mathbb{P}^\star\big(\forall v \in \mathcal{S}(u) : \widehat{\varphi}_{u,v} \leqslant z_{u,v} \mid u \in \widehat{\Xi}_n^{(\text{sure})}[2w_{\text{hh}},\infty)\big)\\
&\quad= (1/2)\mathbb{P}^\star\big(\forall v \in \mathcal{S}(u) : U_{u,v} \leqslant z_{u,v} \mid \forall v \in \mathcal{S}(u) : U_{u,v} > r_k\big)\\
&\qquad+ (1/2)\mathbb{P}^\star\big(\forall v \in \mathcal{S}(u) : U_{u,v} \leqslant z_{u,v} \mid \exists v \in \mathcal{S}(u) : U_{u,v} \leqslant r_k\big).
\end{aligned}$$

Note that $\{U_{u,v}\}_{u,v}$ and $\{\varphi_{u,v}\}_{u,v}$ are both sets of independent Unif$[0,1]$ random variables by Definitions 5.4.6 and 5.4.3, respectively. Hence, (5.4.32) follows by combining (5.4.33) and (5.4.34). □

For the remainder of this section, we assume that we construct $\mathcal{G}_n$ following Definition 5.4.6 and write

$$\Xi_n[2w_{\text{hh}},\infty) := \Xi_n^{(\text{unsure})}[2_{\text{hh}},\infty) \cup \Xi_n^{(\text{sure})}[2w_{\text{hh}},\infty)$$

as the union of two independent Poisson point processes of equal intensity, such that if $\mathcal{G}_{n,2} = \mathcal{G}[1, 2w_{\text{hh}})$ satisfies $\mathcal{A}_{\text{bb}}$ in (5.4.8), every vertex in $\Xi_n^{(\text{sure})}[2w_{\text{hh}},\infty)$ connects by an edge to $\mathcal{C}_{\text{bb}}$, by Corollary 5.4.7.

To finish Step 3, on the event $\mathcal{A}_{\text{bb}}$, we define $\mathcal{G}_{n,3} := (\mathcal{V}_{n,3}, \Psi_{n,3}, p)$, with

$$\begin{aligned}
\mathcal{V}_{n,3} &:= \Xi_n[1, 2w_{\text{hh}}) \cup \Xi_n^{(\text{unsure})}[2_{\text{hh}},\infty),\\
\Psi_{n,3} &:= \Psi_n[1, 2w_{\text{hh}}) \cup \widehat{\Psi}_n^{(\text{iid,unsure})} \cup \widehat{\Psi}_n^{(\text{cond,unsure})},
\end{aligned} \tag{5.4.35}$$

i.e., the graph spanned on $\mathcal{V}_{n,3}$. We call the vertices in $\Xi_n^{(\text{sure})}[2w_{\text{hh}},\infty)$ *sure-connector* vertices. If the event A_{bb} does not hold then we say that the construction failed and we leave $\mathcal{G}_{n,3}$ undefined.

Step 4. Cover expansion

In this step, we ensure that all components of size at least k of $\mathcal{G}_{n,3}$ merge with the giant component of $\mathcal{G}_n$ via edges towards sure-connector vertices, with error probability $\text{err}_{n,k}$ from (5.2.1). The next lemma proves this using the cover-expansion technique of Section 5.3. The notion of expandability is from Definition 5.3.1, and s_k is from (5.4.7). Let for some $c'_{5.4.9} > 0$

$$\mathcal{A}_{\text{exp}}(c'_{5.4.9}) := \mathcal{A}_{\text{exp}}(n, k, c'_{5.4.9}) := \big\{\mathcal{V}_{n,3} \text{ is } (c'_{5.4.9}s_k)\text{-expandable}\big\}. \tag{5.4.36}$$

Lemma⋆ 5.4.9 (Cover-expansion and $\sigma \leqslant \tau - 1$)**.** *Consider an i-KSRG satisfying the conditions of Proposition 5.4.1. For any* $c'_{5.4.9} > 0$

$$\mathbb{P}\big(\neg\mathcal{A}_{\mathrm{exp}}(c'_{5.4.9})\big) \leqslant C_{5.3.11}\, n \exp(-c'_{5.4.9} s_k). \tag{5.4.37}$$

Moreover, if $\sigma \leqslant \tau - 1$*, there exist* $k_2, c_{5.4.9}, c'_{5.4.9} > 0$*, such that conditionally on any realization of* $\mathcal{G}_{n,3}$ *satisfying* $\mathcal{A}_{\mathrm{bb}} \cap \mathcal{A}_{\mathrm{exp}}(c'_{5.4.9})$*, for all* $k \geqslant k_2$ *and any connected component* $\mathcal{C}$ *of* $\mathcal{G}_{n,3}$ *with* $|\mathcal{C}| \geqslant k$*,*

$$\mathbb{P}\big(\mathcal{C} \not\leftrightarrow \Xi_n^{(\mathrm{sure})}[2w_{\mathrm{hh}}, \infty) \mid \mathcal{G}_{n,3}, \mathcal{A}_{\mathrm{bb}} \cap \mathcal{A}_{\mathrm{exp}}(c'_{5.4.9})\big) \leqslant \exp\big(-c_{5.4.9} s_k\big). \tag{5.4.38}$$

Proof. The statement (5.4.37) follows directly for any $c'_{5.4.9} > 0$ from its definition in (5.4.36) and Lemma 5.3.11. In Proposition 5.3.2, for a given mark $\underline{w}$, the function $s(\underline{w})$ in (5.3.1) describes the necessary "expandability parameter", such that all vertices with mark at least $\underline{w}$ in $\mathcal{K}_n(\mathcal{L})$ connect to any $s(\underline{w})$-expandable set $\mathcal{L}$ of vertices with probability at least $p/2$. We shall take $\underline{w} := 2w_{\mathrm{hh}}$, the lowest possible mark in $\Xi_n^{(\mathrm{sure})}$. By Definition 5.3.1 of expandability, if $\mathcal{V}_{n,3}$ is $(c'_{5.4.9} s_k)$-expandable, it is also $s(2w_{\mathrm{hh}})$-expandable whenever

$$s(2w_{\mathrm{hh}}) \geqslant c'_{5.4.9} s_k. \tag{5.4.39}$$

We compute the left-hand side using the value of w_{hh} from (5.4.7) and the function $s(\underline{w})$ from (5.3.1):

$$s(2w_{\mathrm{hh}}) = \big(2^{d+1}\beta C_1^{-1/(\tau-1)}\big)^{1/(1-1/\alpha)} k^{\gamma_{\mathrm{hh}}/(1-1/\alpha)}.$$

By (5.4.7) and (4.2.4), it holds that $s_k = \Theta(k^{1-\gamma_{\mathrm{hh}}(\tau-1)}) = \Theta(k^{\zeta_{\mathrm{hh}}})$. Using (4.2.4), it is elementary to verify that $\gamma_{\mathrm{hh}}/(1-1/\alpha) \geqslant \zeta_{\mathrm{hh}}$ if and only if $\sigma \leqslant \tau - 1$. Thus, if $\sigma \leqslant \tau - 1$, (5.4.39) holds for some $c'_{5.4.9} > 0$, whenever $k > k_2$ for some sufficiently large $k_2 \geqslant 1$, that ensures the condition $\underline{w} > 2^d d^{d/2}/\beta$ in Proposition 5.3.2. On the event $\mathcal{A}_{\mathrm{exp}}(c'_{5.4.9})$, $\mathcal{V}_{n,3}$ is thus also $s(2w_{\mathrm{hh}})$-expandable (since it is $c'_{5.4.9} s_k$ expandable). Since expandability carries through for subsets of $\mathcal{V}_{n,3}$ (below Definition 5.3.1), *any* subset of $\mathcal{V}_{n,3}$ is $s(2w_{\mathrm{hh}})$-expandable. Hence, Proposition 5.3.2 is applicable for any set $\mathcal{L} \subseteq \mathcal{V}_{n,3}$ and $\underline{w} := 2w_{\mathrm{hh}}$, and guarantees the existence of a set $\mathcal{K}_n(\mathcal{L}) \subseteq \Lambda_n$ satisfying (5.3.2) and (5.3.3).

Consider an arbitrary connected component $\mathcal{C}$ of $\mathcal{G}_{n,3}$ that satisfies $|\mathcal{C}| \geqslant k$. With $\mathcal{K}_n(\mathcal{C})$ from Proposition 5.3.2, we define the set of sure-connector vertices inside $\mathcal{K}_n(\mathcal{C})$ connected to $\mathcal{C}$ as

$$\mathcal{H}_{\mathcal{C}} := \{v \in \mathcal{K}_n(\mathcal{C}) \cap \Xi_n^{(\mathrm{sure})}[2w_{\mathrm{hh}}, \infty)\big) : v \leftrightarrow \mathcal{C}\}. \tag{5.4.40}$$

Since $\Xi^{(\mathrm{sure})}[2w_{\mathrm{hh}},\infty)$ is a Poisson process, the number of its points in $\mathcal{K}_n(\mathcal{C})$ follows a Poisson distribution. Since each of these points connects by an edge *independently* to $\mathcal{C}$ with probability at least $p/2$ by (5.3.3), and an independent thinning of a PPP is another PPP; we obtain using the intensity measure in Definition 5.4.6 by the volume bound (5.3.2) on $\mathcal{K}_n(\mathcal{C})$ for $|\mathcal{C}|\geqslant k$ and $c:=(p/4)2^{-(4d+1)}2^{-(\tau-1)}d^{-d/2}/e>0$

$$\begin{aligned}\mathbb{P}\big(|\mathcal{H}_{\mathcal{C}}|=0 \mid \mathcal{G}_{n,3},\mathcal{A}_{\mathrm{bb}}\cap\mathcal{A}_{\exp}(c'_{5.4.9})\big)&\\ &\leqslant \mathbb{P}\big(\mathrm{Poi}\big((p/2)\cdot(1/2)\cdot\mathrm{Vol}(\mathcal{K}_n(\mathcal{C}))\cdot(2w_{\mathrm{hh}})^{-(\tau-1)}\big)=0\big)\\ &\leqslant \exp\big(-(p/4)\mathrm{Vol}(\mathcal{K}_n(\mathcal{C}))(2w_{\mathrm{hh}})^{-(\tau-1)}\big)\big)\\ &\leqslant \exp\big(-c\cdot kw_{\mathrm{hh}}^{-(\tau-1)}\big)=\exp\big(-16c\cdot s_k\big),\end{aligned}$$

where we used s_k from (5.4.7) in the last step. Since $\{|\mathcal{H}_{\mathcal{C}}|>0\}$ in (5.4.40) implies $\{\mathcal{C}\leftrightarrow\Xi_n^{(\mathrm{sure})}[2w_{\mathrm{hh}},\infty)\}$, this finishes the proof of (5.4.38) with $c_{5.4.9}:=16(p/4)2^{-(4d+1)}2^{-(\tau-1)}d^{-d/2}/e$. □

Combining everything: preventing too large components

Proof of Proposition 5.4.1. Assume that $k\geqslant\max\{k_1,k_2\}$ defined in Lemmas 5.4.2 and 5.4.9, respectively, and that $n>k^{1+\zeta_{\mathrm{hh}}/\alpha}$, i.e., n satisfies the bound in Lemma 5.4.2. We construct $\mathcal{G}_n\supseteq\mathcal{G}_{n,3}$ following Definition 5.4.6, where $\mathcal{G}_{n,3}$ from (5.4.35) is the subgraph of $\mathcal{G}_n$ induced on $\mathcal{V}_{n,3}=\Xi_n[1,2w_{\mathrm{hh}})\cup\Xi_n^{(\mathrm{unsure})}[2w_{\mathrm{hh}},\infty)$. The events $\mathcal{A}_{\mathrm{bb}}$ in (5.4.8) and $\mathcal{A}_{\exp}(c'_{5.4.9}):=\mathcal{A}_{\exp}$ in (5.4.36) are measurable with respect to $\mathcal{G}_{n,3}$. By the law of total probability (taking expectation over realizations of $\mathcal{G}_{n,3}$), we obtain

$$\begin{aligned}\mathbb{P}\big(|\mathcal{C}_n^{(2)}|\geqslant k\big)\leqslant \mathbb{E}\big[\mathbb{1}_{\{\mathcal{A}_{\mathrm{bb}}\cap\mathcal{A}_{\exp}\}}\mathbb{P}\big(|\mathcal{C}_n^{(2)}|\geqslant k\mid\mathcal{G}_{n,3},\mathcal{A}_{\exp}\cap\mathcal{A}_{\mathrm{bb}}\big)\big]&\\ +\mathbb{P}\big(\neg\mathcal{A}_{\mathrm{bb}}\big)+\mathbb{P}\big(\neg\mathcal{A}_{\exp}\big).&\end{aligned}\tag{5.4.41}$$

The not-yet revealed vertices after Step 3 are $\mathcal{V}_n\setminus\mathcal{V}_{n,3}=\Xi_n^{(\mathrm{sure})}[2w_{\mathrm{hh}},\infty)$, and by Corollary 5.4.7 each vertex in $\Xi_n^{(\mathrm{sure})}[2w_{\mathrm{hh}},\infty)$ connects by an edge to $\mathcal{C}_{\mathrm{bb}}$. Thus each component $\mathcal{C}\not\supseteq\mathcal{C}_{\mathrm{bb}}$ of $\mathcal{G}_{n,3}$ either remains the same in $\mathcal{G}_n$ or it merges with the component containing $\mathcal{C}_{\mathrm{bb}}$ via a vertex in $\mathcal{V}\setminus\mathcal{V}_{n,3}$. Hence, conditionally on $\mathcal{A}_{\mathrm{bb}}$ and $\mathcal{G}_{n,3}$,

$$\big\{|\mathcal{C}_n^{(2)}|\geqslant k\big\}\subseteq\{\exists\text{ a component }\mathcal{C}\text{ of }\mathcal{G}_{n,3}:|\mathcal{C}|\geqslant k,\mathcal{C}\not\leftrightarrow\Xi_n^{(\mathrm{sure})}[2w_{\mathrm{hh}},\infty)\}.\tag{5.4.42}$$

By a union bound over the at most $|\mathcal{V}_{n,3}|/k$ components of size at least k, (5.4.38) of Lemma 5.4.9 yields

$$\mathbb{P}\big(|\mathcal{C}_n^{(2)}|\geqslant k\mid\mathcal{G}_{n,3},\mathcal{A}_{\exp}\cap\mathcal{A}_{\mathrm{bb}}\big)\leqslant(|\mathcal{V}_{n,3}|/k)\exp\big(-c_{5.4.9}s_k\big).$$

Substituting this bound into (5.4.41), and using Lemmas 5.4.2 and 5.4.9 to bound the last two terms yields

$$\begin{aligned}\mathbb{P}\big(|\mathcal{C}_n^{(2)}| \geqslant k\big) \leqslant (\mathbb{E}\big[|\mathcal{V}_{n,3}|\big]/k)\exp\big(-c_{5.4.9}s_k\big) + (3n/k)\exp\big(-c_{5.4.2}s_k\big)\\ + C_{5.3.11}n\exp\big(-c'_{5.4.9}s_k\big).\end{aligned}$$

Since $\mathcal{V}_{n,3} \subseteq \Xi_n$ by construction, $\mathbb{E}[|\Xi_n|] = n$ by (5.1.2), and $s_k = \Theta(k^{\zeta_{hh}})$ by (5.4.7), this finishes the proof of Proposition 5.4.1 for $k \geqslant \max\{k_1, k_2\}$ and $n > k^{1+\zeta_{hh}/\alpha}$. For $k < \max\{k_1, k_2\}$, (5.4.1) is trivially satisfied for $c_{5.4.1} > 0$ sufficiently small. □

The backbone: intermediate results

We state two corollaries of the proof of Proposition 5.4.1, and two propositions based on the backbone constructions for later use. We defer the detailed proofs to the appendix and only give a sketch here. We start with a corollary of the proof of Proposition 5.4.1.

Corollary⋆ 5.4.10 (Backbone becoming part of the giant)**.** *Consider an i-KSRG satisfying the conditions of Proposition 5.4.1. Then conditionally on the graph $\mathcal{G}_{n,2} = \mathcal{G}_n[1, 2w_{hh})$ satisfying $\mathcal{A}_{bb}(n,k)$ in* (5.4.8)*,*

$$\mathbb{P}\big(\mathcal{C}_{bb}(n,k) \not\subseteq \mathcal{C}_n^{(1)} \mid \mathcal{G}_{n,2}, \mathcal{A}_{bb}(n,k)\big) \leqslant (n/c_{5.4.1})\exp\big(-c_{5.4.1}k^{\zeta_{hh}}\big). \tag{5.4.43}$$

Proof sketch. By definition of the backbone in (5.4.8), the backbone contains at least k vertices. The proof of Proposition 5.4.1 “merges” each component of size at least k with the backbone, with error probability given on the right-hand side of (5.4.43). Hence, the component containing the backbone is the only remaining component of size at least k. □

The next corollary follows from Lemma 5.4.2. It is not sharp but it yields a useful temporary estimate.

Corollary 5.4.11 (Lower bound on largest component)**.** *Consider a supercritical interpolating KSRG model as in Definition 5.1.1 with parameters $\alpha > 1$, $\tau \in (2, 2+\sigma)$, $\sigma \geqslant 0$, $d \in \mathbb{N}$. For each $\delta > 0$, there exists a constant $A > 0$ such that for all n sufficiently large*

$$\mathbb{P}\big(|\mathcal{C}_n^{(1)}| \leqslant n(A\log(n))^{-1/\zeta_{hh}}\big) \leqslant n^{-\delta}.$$

Proof. If $\mathcal{A}_{\mathrm{bb}}(n, k_n)$ holds for some k_n, then the largest component $\mathcal{C}_n^{(1)}$ must have at least the size of the backbone $\mathcal{C}_{\mathrm{bb}}(n, k_n)$. Setting $k = k_n = (A\log(n))^{1/\zeta_{\mathrm{hh}}}$ in (5.4.9), the backbone exists with probability at least $1 - n^{-\delta}$ for $A = A(\delta)$ sufficiently large, since $s_k = s_{k_n} = (C_1/16)A\log(n)$ by (5.4.7). Then its size is at least $(n'/k)s_k \geqslant (n'/k) = \Theta(nk_n^{-1/\zeta_{\mathrm{hh}}})$ by definition of n' in (5.4.2), finishing the proof. □

The next proposition shows that all vertices with a sufficiently high mark (simultaneously) belong to the largest component $\mathcal{C}_n^{(1)}$ with polynomially small error probability.

Proposition⋆ 5.4.12 (Controlling marks of non-giant vertices)**.** *Consider an i-KSRG in the setting of Theorem 5.1.2(ii). For* all *$\delta > 0$, there exists $M_\delta > 0$ such that for $\overline{w}(n, \delta) = (M_\delta \log(n))^{(1-\sigma\gamma_{\mathrm{hh}})/\zeta_{\mathrm{hh}}}$*

$$\mathbb{P}\big(\exists u \in \Xi_n[\overline{w}(n, \delta), \infty) : u \notin \mathcal{C}_n^{(1)}\big) \leqslant n^{-\delta}. \tag{5.4.44}$$

Proof sketch. We give the detailed proof in the Appendix on page 197, and here a sketch. We consider k as a free parameter, so using Lemma 5.4.2 with $k = k_n = \Theta\big(\log^{1/\zeta_{\mathrm{hh}}}(n)\big)$, a backbone $\mathcal{C}_{\mathrm{bb}}(n, k_n)$ exists and satisfies $\mathcal{C}_{\mathrm{bb}}(n, k_n) \subseteq \mathcal{C}_n^{(1)}$, with probability at least $1 - n^{-\delta}$ by Corollary 5.4.10 and Lemma 5.4.2 (the same calculation as the proof of Corollary 5.4.11). The choice of $\overline{w} = \overline{w}(n, \delta)$ is so that a vertex u with mark $w_u \geqslant \overline{w}(n, \delta)$ connects to each backbone-vertex in its own subbox with probability at least p in (5.1.3). For u to not be contained in $\mathcal{C}_n^{(1)}$, these s_{k_n} many edges must be all absent. Then we apply concentration inequalities to obtain the result. □

Remark. Combined with the proof of the lower bound of Theorem 5.1.4 below in Section 5.6, one may show that Proposition 5.4.12 is sharp up to a constant factor, i.e., there exist constants $\delta, m_w > 0$ such that for all n sufficiently large

$$\mathbb{P}\big(\exists v \in \Xi_n[(m_w \log(n))^{(1-\sigma\gamma_{\mathrm{hh}})/\zeta_{\mathrm{hh}}}, \infty) : v \notin \mathcal{C}_n^{(1)}\big) \geqslant 1 - n^{-\delta}.$$

We state a proposition that 0 is with constant probability in a linear-sized component in the induced subgraph $\mathcal{G}_{n,2} = \mathcal{G}_n[1, 2w_{\mathrm{hh}}(k))$ with $k = k_n = \mathrm{polylog}(n)$ and $w_{\mathrm{hh}}(k)$ from (5.4.7). Let $\mathcal{C}_{n,2}(0)$ be the component containing 0 in $\mathcal{G}_{n,2} \subseteq \mathcal{G}_n$, by setting $\mathcal{C}_{n,2}(0)$ to be the empty set if $w_0 \geqslant 2w_{\mathrm{hh}}(k)$.

Proposition 5.4.13 (Existence of a large component)**.** *Consider a supercritical i-KSRG model as in Definition 5.1.1, with parameters* $\alpha > 1$*,* $\tau \in (2, 2+\sigma)$*,* $\sigma \geqslant 0$*,* $d \in \mathbb{N}$*. There exist constants* $\rho, m > 0$ *such that for all* n *sufficiently large, when* $k = k_n = (m \log(n))^{1/\zeta_{hh}}$*,*

$$\mathbb{P}^0\big(|\mathcal{C}_{n,2}(0)| \geqslant \rho n\big) \geqslant \rho, \qquad \textit{and} \qquad \mathbb{P}^0\big(|\mathcal{C}_\infty(0)| = \infty\big) \geqslant \rho. \tag{5.4.45}$$

Proof sketch. For the first inequality in (5.4.45), we build a connected backbone on vertices with mark in $[w_{hh}(k_n), 2w_{hh}(k_n))$ using Lemma 5.4.2. Then we use a second-moment method to show that the origin and linearly many other vertices are connected to this backbone via paths along which the vertex marks are increasing. The second statement follows similarly, forming an infinite path along which the marks are increasing. The detailed proof can be found in Appendix 5.A.1. □

5.5 UPPER BOUND: SUBEXPONENTIAL DECAY

In this section we prove Theorem 5.1.2(ii). We carry out the plan in Section 5.2.1.2 in detail. Instead of arguing directly for KSRGs with parameters described in Theorem 5.1.2(ii), we derive general conditions that ensure that a bound on the size of the second-largest component as in Proposition 5.4.1 readily implies subexponential decay. Recall from Definition 5.1.1 that $\mathbb{P}^x$ denotes the conditional measure that $\mathcal{V}$ contains a vertex at location $x \in \Lambda_n$, that has an unknown mark from distribution F_W.

Proposition 5.5.1 (Prerequisites for subexponential decay)**.** *Consider a supercritical i-KSRG model as in Definition 5.1.1, with parameters* $\alpha > 1, \tau > 2$ *and* $d \in \mathbb{N}$*. Assume that there exist* $\zeta, \eta, c, c', \eta, M_c > 0$*, and a function* $n_0(k) = O(k^{1+c'})$*, such that for all* k *sufficiently large constant and whenever* $n \in [n_0(k), \infty)$*, with* $\overline{w}(n, c) := M_c \log^\eta(n)$ *it holds that for any* $x \in \Lambda_n$

$$\mathbb{P}^x\big(|\mathcal{C}_n^{(2)}| \geqslant k\big) \leqslant n^{c'} \exp\big(-ck^\zeta\big), \tag{5.5.1}$$

$$\mathbb{P}^x\big(|\mathcal{C}_n^{(1)}| \leqslant n^c\big) \leqslant n^{-1-c}, \tag{5.5.2}$$

$$\mathbb{P}^x\big(\exists v \in \Xi_n[\overline{w}(n,c), \infty) : v \notin \mathcal{C}_n^{(1)}\big) \leqslant n^{-c}. \tag{5.5.3}$$

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

$$\mathbb{P}^0\big(|\mathcal{C}_n(0)| \geqslant k, 0 \notin \mathcal{C}_n^{(1)}\big) \leqslant A \exp\big(-(1/A)k^\zeta\big), \tag{5.5.4}$$

and

$$\frac{|\mathcal{C}_n^{(1)}|}{n} \xrightarrow{\mathbb{P}} \mathbb{P}^0\big(|\mathcal{C}_\infty(0)| = \infty\big), \qquad \textit{as } n \to \infty. \tag{5.5.5}$$

Observe that (5.5.4) does not follow from a naive application of (5.5.1), since we sharpened the polynomial prefactor on the right-hand side of (5.5.1) to a universal constant A in (5.5.4), and in (5.5.4) $n = \infty$ is also allowed in (5.5.4). The inequalities (5.5.1)–(5.5.3) are satisfied when $\tau \in (2, 2+\sigma)$ and $\sigma \leqslant \tau - 1$ by Propositions 5.4.1 and 5.4.12 and Corollary 5.4.11 (we leave it to the reader to verify that the results hold also for the Palm-version $\mathbb{P}^x$ of $\mathbb{P}$). Thus, Theorem 5.1.2(ii) follows immediately after we prove Proposition 5.5.1. We state and prove an intermediate claim that we need for Proposition 5.5.1. We write $\mathcal{C}_{\mathcal{Q}}^{(1)}$ for the largest component in the graph induced on vertices in $\mathcal{Q} \subseteq \mathbb{R}^d$.

Claim 5.5.2 (Leaving the giant)**.** *Consider a supercritical i-KSRG model as in Definition 5.1.1, under the same setting as Proposition 5.5.1. Assume that* (5.5.1)–(5.5.3) *hold,* n *is sufficiently large and let* $N \in [n, \infty]$. *Then there exists* $\delta > 0$ *such that for any box* $\mathcal{Q}_n \subseteq \Lambda_N$ *with* $\mathrm{Vol}(\mathcal{Q}_n) = n$*, any* $x \in \mathcal{Q}_n$*,* n *sufficiently large, and all* $N \geqslant n$*, it holds for* $u := (x, w_u)$ *that*

$$\mathbb{P}^x\big(u \in \mathcal{C}_{\mathcal{Q}_n}^{(1)}, u \notin \mathcal{C}_N^{(1)}\big) \leqslant n^{-\delta}. \tag{5.5.6}$$

Proof. We will first prove the following bound that holds generally for a sequence of increasing graphs $G_n \subseteq G_{n+1} \subseteq \dots$, whose largest and second-largest components we denote by $C_n^{(1)}$ and $C_n^{(2)}$, respectively. Let $(k_n)_{n\geqslant 0}$ and $(K_n)_{n\geqslant 0}$ be two non-negative sequences such that $k_{n+1} < K_n$ for all $n \geqslant 0$. Then, for all $0 < n \leqslant N \leqslant \infty$,

$$\mathbb{P}\big(u \in C_n^{(1)}, u \notin C_N^{(1)}\big) \leqslant \sum_{\tilde{n}=n}^{N} \Big(\mathbb{P}\big(|C_{\tilde{n}}^{(1)}| < K_{\tilde{n}}\big) + \mathbb{P}\big(|C_{\tilde{n}}^{(2)}| > k_{\tilde{n}}\big)\Big). \tag{5.5.7}$$

We verify the bound using an inductive argument. We define for $\tilde{n} \geqslant n$ the events

$$\mathcal{A}(\tilde{n}) := \big\{|C_{\tilde{n}}^{(1)}| \geqslant K_{\tilde{n}}\big\} \cap \big\{|C_{\tilde{n}+1}^{(2)}| \leqslant k_{\tilde{n}+1}\big\}.$$

Since by assumption $k_{\tilde{n}+1} < K_{\tilde{n}}$, the event $\mathcal{A}(\tilde{n})$ ensures that $|C_{\tilde{n}}^{(1)}|$ is already larger than $|C_{\tilde{n}+1}^{(2)}|$, hence $\mathcal{A}(\tilde{n})$ implies that $C_{\tilde{n}}^{(1)} \subseteq C_{\tilde{n}+1}^{(1)}$. Iteratively applying this argument yields that $\cap_{\tilde{n}\in[n,N]}\mathcal{A}(\tilde{n})$ implies that $\{C_n^{(1)} \subseteq C_N^{(1)}\}$.

We combine this with the observation that given that $u \in V_n$, it holds that $\{u \in C_n^{(1)}, u \notin C_N^{(1)}\} \subseteq \{C_n^{(1)} \nsubseteq C_N^{(1)}\}$. This yields

$$\begin{aligned}\mathbb{P}\big(u \in C_n^{(1)}, u \notin C_N^{(1)}\big) &\leqslant \mathbb{P}\big(C_n^{(1)} \nsubseteq C_N^{(1)}\big) \\ &\leqslant \mathbb{P}\Big(\{C_n^{(1)} \nsubseteq C_N^{(1)}\} \cap \bigcap_{\tilde{n}=n}^{N-1} \mathcal{A}(\tilde{n})\Big) + \sum_{\tilde{n}=n}^{N-1} \mathbb{P}\big(\neg\mathcal{A}(\tilde{n})\big) \\ &\leqslant 0 + \sum_{\tilde{n}=n}^{N} \Big(\mathbb{P}\big(|C_{\tilde{n}}^{(1)}| < K_{\tilde{n}}\big) + \mathbb{P}\big(|C_{\tilde{n}}^{(2)}| > k_{\tilde{n}}\big)\Big),\end{aligned} \tag{5.5.8}$$

showing (5.5.7). We move on to (5.5.6) for which we have to define the increasing sequence of graphs. Consider any sequence of boxes $(\mathcal{Q}_{\tilde{n}})_{\tilde{n} \geqslant n}$ such that $\mathcal{Q}_n \subseteq \mathcal{Q}_{n+1} \subseteq \cdots \subseteq \mathcal{Q}_N := \Lambda_N$ and $\mathrm{Vol}(\mathcal{Q}_{\tilde{n}}) = \tilde{n}$, and let $G_{\tilde{n}}$ denote the induced subgraph of $\mathcal{G}$ on $\mathcal{Q}_{\tilde{n}}$ for $\tilde{n} \in [n, N]$, that is conditioned to contain a vertex at location $x \in \mathcal{Q}_n$. Using the assumed lower bound on $|\mathcal{C}_{\tilde{n}}^{(1)}|$ in (5.5.2), we set $K_{\tilde{n}} := \tilde{n}^c$, and using the assumed upper bound on $|\mathcal{C}_{\tilde{n}}^{(2)}|$ in (5.5.1), there exists a large constant A so that whenever n is sufficiently large, and $\tilde{n} \geqslant n$, by setting $k_{\tilde{n}} := (A\log(\tilde{n}))^{1/\zeta_{\mathrm{hh}}}$, it holds that

$$\mathbb{P}^x\big(|\mathcal{C}_{\mathcal{Q}_{\tilde{n}}}^{(1)}| < K_{\tilde{n}}\big) \leqslant \tilde{n}^{-c-1}, \qquad \mathbb{P}^x\big(|\mathcal{C}_{\mathcal{Q}_{\tilde{n}}}^{(2)}| > k_{\tilde{n}}\big) \leqslant \tilde{n}^{-c-1}.$$

Clearly $k_{\tilde{n}+1} < K_{\tilde{n}}$ for $\tilde{n} \geqslant n$, so that substituting the bounds into (5.5.8) and summing over $\tilde{n} \geqslant n$ yields the assertion (5.5.6) for any $\delta < c$ and n sufficiently large. □

We continue to prove Proposition 5.5.1, starting with some notation. For some $\varepsilon \in (0, \min\{\alpha-1, \tau-2\})$, and using c, c', η, ζ, and $\overline{w}(n, c)$ from the statement of Proposition 5.5.1

$$\begin{aligned} N_k &:= \exp(k^\zeta c/(2c')), & n_k &:= N_k^\varepsilon, \\ \overline{w}_{N_k} &:= \overline{w}(N_k, c) = M_c(k^\zeta)^\eta(\tfrac{c}{2c'})^\eta, & t_k &:= n_k^{1/d}/(2k). \end{aligned} \tag{5.5.9}$$

Note that $N_k, n_k = \exp(\Theta(k^\zeta))$. For $n \leqslant N_k$, the statement (5.5.4) follows directly from (5.5.1), since when $n = N_k$, the right-hand side of (5.5.1) becomes $k^{dc'}\exp(-k^\zeta c/2) \leqslant \exp(-k^\zeta c/4)$, so we may set any A such that $1/A \leqslant c/2$ in (5.5.4). So we may assume in the remainder of the section that

$n > N_k$. We write for $\mathcal{C}_{\Lambda(x,n)}(u)$ the component of vertex $u := (x, w_u) \in \mathcal{V}$ in the graph $\mathcal{G}_\infty$ restricted to $\Lambda(x,n)$. Define for $x \in \mathbb{R}^d$ the two events

$$\begin{aligned} &\mathcal{A}_{\text{low-edge}}(x, n_k, N_k, \overline{w}_{N_k}) \\ &\quad := \begin{Bmatrix} |\mathcal{C}_{\Lambda(x,n_k)}(u)| < k, \exists v_1, v_2 \in \Xi_{\Lambda(x,N_k)}[1, \overline{w}_{N_k}) \text{ such that} \\ v_1 \in \mathcal{C}_{\Lambda(x,n_k)}(u), \|x_{v_1} - x_{v_2}\| \geqslant t_k, \text{ and } v_1 \leftrightarrow v_2 \end{Bmatrix}, \end{aligned} \tag{5.5.10}$$

$$\begin{aligned} &\mathcal{A}_{\text{long-edge}}(x, n_k, N_k, \overline{w}_{N_k}) \\ &\quad := \big\{\exists v_1 \in \Xi_{\Lambda(x,n_k)}[1, \overline{w}_{N_k}), \exists v_2 \in \Xi \setminus \Xi_{\Lambda(x,N_k)}, v_1 \leftrightarrow v_2\big\}. \end{aligned} \tag{5.5.11}$$

The next lemma relates the probability of the event $\{|\mathcal{C}_n(u)| \geqslant k, u \notin \mathcal{C}_n^{(1)}\}$ to the events $\mathcal{A}_{\text{low-edge}}$ and $\mathcal{A}_{\text{long-edge}}$ using the assumed bounds in Proposition 5.5.1.

Lemma 5.5.3 (Extending the box-sizes)**.** *Consider a supercritical i-KSRG model as in Definition 5.1.1, under the same setting as Proposition 5.5.1. Assume that* (5.5.1)–(5.5.3) *hold. Then there exists* $A' > 0$ *such that for all* n *with* $n \in [N_k, \infty]$, *whenever* $u = (x, w_u) \in \mathcal{V}$ *is a vertex with* $\|x - \partial\Lambda_n\| \geqslant N_k^{1/d}/2$,

$$\begin{aligned} &\mathbb{P}^x\big(|\mathcal{C}_n(u)| \geqslant k, u \notin \mathcal{C}_n^{(1)}\big) \\ &\quad \leqslant A' \exp\big(-(1/A')k^\zeta\big) + \mathbb{P}^0\big(\mathcal{A}_{\text{low-edge}}(0, n_k, N_k, \overline{w}_{N_k})\big) \\ &\qquad + \mathbb{P}^0\big(\mathcal{A}_{\text{long-edge}}(0, n_k, N_k, \overline{w}_{N_k})\big). \end{aligned} \tag{5.5.12}$$

Proof. Let $\tilde{n} \geqslant 1$, and denote by $\mathcal{C}^{(1)}_{\Lambda(x,\tilde{n})}$ the largest connected component in the induced subgraph of $\mathcal{G}_\infty$ inside the box $\Lambda(x, \tilde{n}) \subseteq \Lambda_n$. For a vertex $u = (x, w_u) \in \mathcal{V}$ define

$$\mathcal{A}_{\text{leave-giant}}(x, \tilde{n}) := \{u \in \mathcal{C}^{(1)}_{\Lambda(x,\tilde{n})}, u \notin \mathcal{C}_n^{(1)}\}, \tag{5.5.13}$$

$$\mathcal{A}_{\text{mark-giant}}(x, N_k, \overline{w}_{N_k}) := \big\{\forall v \in \Xi_{\Lambda(x,N_k)}[\overline{w}_{N_k}, \infty) : v \in \mathcal{C}^{(1)}_{\Lambda(x,N_k)}\big\}. \tag{5.5.14}$$

The first events relates to (5.5.6) in Claim 5.5.2, while the second one to (5.5.3) of Proposition 5.5.1. The values of $n_k < N_k \leqslant n$ from (5.5.9) and that we assumed $\|x - \partial\Lambda_n\| \geqslant N_k^{1/d}/2$ ensure that $\Lambda(x, n_k) \subseteq \Lambda(x, N_k) \subseteq \Lambda_n$. Then we bound

$$\begin{aligned}
\big\{|\mathcal{C}_n(u)| \geqslant k, u \notin \mathcal{C}_n^{(1)}\big\} \\
&\subseteq \big\{|\mathcal{C}_n(u)| \geqslant k, u \notin \mathcal{C}_n^{(1)}, u \notin \mathcal{C}^{(1)}_{\Lambda(x,n_k)}, |\mathcal{C}^{(2)}_{\Lambda(x,n_k)}| \geqslant k\big\} \\
&\quad \cup \big\{u \in \mathcal{C}^{(1)}_{\Lambda(x,n_k)}, u \notin \mathcal{C}_n^{(1)}\big\} \\
&\quad \cup \big\{|\mathcal{C}_n(u)| \geqslant k, u \notin \mathcal{C}_n^{(1)}, u \notin \mathcal{C}^{(1)}_{\Lambda(x,n_k)}, |\mathcal{C}^{(2)}_{\Lambda(x,n_k)}| < k\big\} \\
&\subseteq \big\{|\mathcal{C}^{(2)}_{\Lambda(x,n_k)}| \geqslant k\big\} \cup \mathcal{A}_{\text{leave}-\text{giant}}(x, n_k) \qquad (5.5.15) \\
&\quad \cup \big\{|\mathcal{C}_n(u)| \geqslant k, u \notin \mathcal{C}_n^{(1)}, u \notin \mathcal{C}^{(1)}_{\Lambda(x,n_k)}, |\mathcal{C}_{\Lambda(x,n_k)}(u)| < k\big\}.
\end{aligned}$$

Applying probabilities on both sides we obtain the inequality stated in (5.2.6) for $x=0$. We introduce a shorthand notation for the third event on the right-hand side of (5.5.15), i.e.,

$$\mathcal{A}_{\text{goal}} := \big\{|\mathcal{C}_n(u)| \geqslant k, u \notin \mathcal{C}_n^{(1)}, u \notin \mathcal{C}^{(1)}_{\Lambda(x,n_k)}, |\mathcal{C}_{\Lambda(x,n_k)}(u)| < k\big\}$$

Define the auxiliary events

$$\begin{aligned}
\mathcal{A}_{\text{becomes}-\text{large}} &:= \{|\mathcal{C}_{\Lambda(x,n_k)}(u)| < k, |\mathcal{C}_n(u)| \geqslant k\}, \\
\mathcal{A}_{\text{outof}-\text{giant}}(n_k, n) &:= \{u \notin \mathcal{C}_n^{(1)}, u \notin \mathcal{C}^{(1)}_{\Lambda(x,n_k)}\},
\end{aligned} \qquad (5.5.16)$$

and observe that $\mathcal{A}_{\text{goal}} = \mathcal{A}_{\text{becomes}-\text{large}} \cap \mathcal{A}_{\text{outof}-\text{giant}}(n_k, n)$. In order to bound $\mathbb{P}(\mathcal{A}_{\text{goal}})$, we distinguish whether u enters the giant at the intermediate box of size $N_k \in (n_k, n)$ or not:

$$\begin{aligned}
\mathcal{A}_{\text{goal}} &\subseteq \{u \in \mathcal{C}^{(1)}_{\Lambda(x,N_k)}, u \notin \mathcal{C}_n^{(1)}\} \\
&\quad \cup \{|\mathcal{C}_n(u)| \geqslant k, u \notin \mathcal{C}^{(1)}_{\Lambda(x,N_k)}, u \notin \mathcal{C}^{(1)}_{\Lambda(x,n_k)}, |\mathcal{C}_{\Lambda(x,n_k)}(u)| < k\} \\
&= \mathcal{A}_{\text{leave}-\text{giant}}(x, N_k) \cup \big(\mathcal{A}_{\text{becomes}-\text{large}} \cap \mathcal{A}_{\text{outof}-\text{giant}}(n_k, N_k)\big),
\end{aligned} \qquad (5.5.17)$$

with $\mathcal{A}_{\text{leave}-\text{giant}}(x, \tilde{n})$ defined in (5.5.14). We observe that $\mathcal{A}_{\text{becomes}-\text{large}}$ in (5.5.16) implies that at least one of the at most $k-1$ vertices in $\mathcal{C}_{\Lambda(x,n_k)}(u)$ has an incident edge crossing the boundary of $\Lambda(x, n_k)$, and so there must exist a “fairly” long edge either inside the cluster $\mathcal{C}_{\Lambda(x,n_k)}(u)$ or between a vertex in $\mathcal{C}_{\Lambda(x,n_k)}(u)$ and a vertex in $\mathcal{C}_n(u) \setminus \mathcal{C}_{\Lambda(x,n_k)}(u)$. More precisely, recalling $t_k = n_k^{1/d}/(2k)$, define

$$\mathcal{A}_{\text{edge}} := \left\{ \begin{matrix} |\mathcal{C}_{\Lambda(x,n_k)}(u)| < k, \exists v_1 \in \mathcal{C}_{\Lambda(x,n_k)}(u), \\ v_2 \in \Xi : v_1 \leftrightarrow v_2, \|v_1 - v_2\| \geqslant t_k \end{matrix} \right\}. \qquad (5.5.18)$$

We argue that $\mathcal{A}_{\mathrm{becomes-large}} \subseteq \mathcal{A}_{\mathrm{edge}}$. By the pigeon-hole principle, if all edges adjacent to all vertices in $\mathcal{C}_{\Lambda(x,n_k)}(u)$ were shorter than t_k, the furthest point that could be reached from x with at most $k-1$ edges has Euclidean norm at most $(k-1)t_k < n_k^{1/d}/2$, and thus its location would be inside $\Lambda(x, n_k)$, contradicting the definition of $\mathcal{A}_{\mathrm{becomes-large}}$ in (5.5.16). Returning to (5.5.17), we obtain that

$$\begin{aligned}\mathcal{A}_{\mathrm{goal}} &\subseteq \mathcal{A}_{\mathrm{leave-giant}}(x, N_k) \cup \big(\mathcal{A}_{\mathrm{edge}} \cap \mathcal{A}_{\mathrm{outof-giant}}(n_k, N_k)\big)\\ &\subseteq \mathcal{A}_{\mathrm{leave-giant}}(x, N_k) \cup \big(\mathcal{A}_{\mathrm{edge}} \cap \{u \notin \mathcal{C}^{(1)}_{\Lambda(x,N_k)}\}\big).\end{aligned}$$

In order to bound the existence of long edges, we put restrictions on the marks: we distinguish whether all vertices in $\Xi_{\Lambda(x,N_k)} \setminus \mathcal{C}^{(1)}_{\Lambda(x,N_k)}$ have mark at most $\overline{w}_{N_k}$ – this is the event $\mathcal{A}_{\mathrm{mark-giant}}(x, N_k, \overline{w}_{N_k})$ in (5.5.14) – or not. We obtain

$$\begin{aligned}\mathcal{A}_{\mathrm{goal}} \subseteq\; &\mathcal{A}_{\mathrm{leave-giant}}(x, N_k) \cup \big(\neg\mathcal{A}_{\mathrm{mark-giant}}(x, N_k, \overline{w})\big)\\ &\cup \big(\mathcal{A}_{\mathrm{edge}} \cap \{u \notin \mathcal{C}^{(1)}_{\Lambda(x,N_k)} \cap \mathcal{A}_{\mathrm{mark-giant}}(x, N_k, \overline{w})\}\big).\end{aligned} \tag{5.5.19}$$

The intersection with $\{u \notin \mathcal{C}^{(1)}_{\Lambda(x,N_k)}\} \cap \mathcal{A}_{\mathrm{mark-giant}}(x, N_k, \overline{w}_{N_k})$ in the last event ensures that *all* vertices in the cluster of u with location in $\Lambda(x, N_k) \supseteq \Lambda(x, n_k)$ have mark at most $\overline{w}_k$. We make another case distinction, with respect to the locations of the vertices the edge $e_{\geqslant t_k}$ of length at least t_k that exists on the event $\mathcal{A}_{\mathrm{edge}}$ in (5.5.18). Namely, $e_{\geqslant t_k}$ either has both endpoints in Λ_{N_k} or it has one endpoint inside Λ_{n_k} and the other one outside Λ_{N_k}. For the first event, we obtain the event $\mathcal{A}_{\mathrm{low-edge}}(x, n_k, N_k, \overline{w}_k)$, and for the latter $\mathcal{A}_{\mathrm{long-edge}}(x, n_k, N_k, \overline{w}_k)$, respectively (defined in (5.5.10)—(5.5.11)). Hence,

$$\begin{aligned}&\mathcal{A}_{\mathrm{edge}} \cap \{u \notin \mathcal{C}^{(1)}_{\Lambda(x,N_k)}\} \cap \mathcal{A}_{\mathrm{mark-giant}}(x, N_k, \overline{w}_{N_k})\\ &\qquad\subseteq \mathcal{A}_{\mathrm{low-edge}}(x, n_k, N_k, \overline{w}_k) \cup \mathcal{A}_{\mathrm{long-edge}}(x, n_k, N_k, \overline{w}_k).\end{aligned}$$

Using this in (5.5.19), then substituting (5.5.19) back into (5.5.15), and then taking probabilities yields

$$\begin{aligned}&\mathbb{P}^x\big(|\mathcal{C}_n(u)| \geqslant k, u \notin \mathcal{C}^{(1)}_n\big)\\ &\quad\leqslant \mathbb{P}^x\big(|\mathcal{C}^{(2)}_{\Lambda(x,n_k)}| \geqslant k\big) + \mathbb{P}^x\big(\neg\mathcal{A}_{\mathrm{mark-giant}}(x, N_k, \overline{w})\big)\\ &\qquad+ \sum_{\tilde{n}\in\{n_k,N_k\}} \mathbb{P}^x\big(\mathcal{A}_{\mathrm{leave-giant}}(x, \tilde{n})\big)\\ &\qquad+ \mathbb{P}^x\big(\mathcal{A}_{\mathrm{low-edge}}(x, n_k, N_k, \overline{w})\big) + \mathbb{P}^x\big(\mathcal{A}_{\mathrm{long-edge}}(x, n_k, N_k, \overline{w})\big).\end{aligned} \tag{5.5.20}$$

The event $\mathcal{A}_{\text{leave-giant}}(x, \tilde{n}) = \{u \in \mathcal{C}^{(1)}_{\Lambda(x,\tilde{n})}, u \notin \mathcal{C}^{(1)}_n\}$ considers the graph in the box Λ_n (which is centered at the origin) and therefore does not necessarily have the same probability for all $x \in \Lambda_n$. The four other events consider the graph in boxes centered at x. Hence, we translate those events (and the Palm measure $\mathbb{P}^x$) by $-x$ to obtain

$$\begin{aligned}\mathbb{P}^x&\big(|\mathcal{C}_n(u)| \geqslant k, u \notin \mathcal{C}^{(1)}_n\big)\\ &\leqslant \mathbb{P}^0\big(|\mathcal{C}^{(2)}_{n_k}| \geqslant k\big) + \mathbb{P}^0\big(\neg\mathcal{A}_{\text{mark-giant}}(0, N_k, \overline{w})\big)\\ &\quad + \sum_{\tilde{n}\in\{n_k,N_k\}} \mathbb{P}^x\big(u \in \mathcal{C}^{(1)}_{\Lambda(x,\tilde{n})}, u \notin \mathcal{C}^{(1)}_n\big)\\ &\quad + \mathbb{P}^0\big(\mathcal{A}_{\text{low-edge}}(0, n_k, N_k, \overline{w})\big) + \mathbb{P}^0\big(\mathcal{A}_{\text{long-edge}}(0, n_k, N_k, \overline{w})\big),\end{aligned}$$

The first two terms can be bounded by substituting the definitions $n_k, N_k = \exp\big(\Theta(k^\zeta)\big)$ in (5.5.9) into the assumed bounds on the probabilities in Proposition 5.5.1. The terms in the sum are bounded from above by $n_k^{-\delta} = \exp\big(-\Theta(k^\zeta)\big)$ by Claim 5.5.2. This finishes the proof of (5.5.12).

□

We move on to bounding $\mathbb{P}^0(\mathcal{A}_{\text{low-edge}})$ on the right-hand side of (5.5.12) in Proposition 5.5.3, with $\mathcal{A}_{\text{low-edge}}$ from (5.5.10). To do so, we need an auxiliary claim that controls the probability that for every point in $\Xi_{N_k}[1, \overline{w}_{N_k})$ there are not "too many" points at distance at least t_k, with $t_k = n_k^{1/d}/(2k)$. We define first for $i \geqslant 1$, and $u = (x_u, w_u) \in \Xi_{n_k}[0, \overline{w})$ the annuli

$$\mathcal{R}_i(x_u) := \big(\Lambda\big(x_u, (2^i t_k)^d\big) \setminus \Lambda\big(x_u, (2^{i-1} t_k)^d\big)\big) \times [1, \infty). \qquad (5.5.21)$$

With the measure μ_τ from (5.1.2), we then define the *bad* events

$$\mathcal{A}_{\text{dense}} := \big\{\exists i \geqslant 1, u \in \Xi_{n_k}[1, \overline{w}_{N_k}) : |\Xi_{N_k} \cap \mathcal{R}_i(x_u)| > 2 \cdot \mu_\tau\big(\mathcal{R}_i(x_u)\big)\big\}. \qquad (5.5.22)$$

In the following auxiliary claim we give an upper bound on $\mathbb{P}\big(\mathcal{A}_{\text{dense}}\big)$. Its proof is standard, based on Palm theory and Chernoff bounds, see page 206 of Appendix 5.B.

Claim 5.5.4. *For all* $c, \delta > 0$, *there exists* n_0 *such that* $\mathbb{P}^0\big(\mathcal{A}_{\text{dense}}\big) \leqslant n_k^{-c}$ *for all* $n_k \geqslant \max\{n_0, k^{d+\delta}\}$.

We can now analyse $\mathbb{P}^0(\mathcal{A}_{\text{low-edge}})$ in Lemma 5.5.3. The next claim holds for the specific choices of $n_k, N_k, \overline{w}_{N_k}$ and t_k in (5.5.9) for all k is sufficiently large, but extends to more general settings, e.g. KSRGs defined on any vertex set that satisfies $\neg\mathcal{A}_{\text{dense}}$ in (5.5.22).

Claim 5.5.5 (No *low*-mark edge from a small component)**.** *Consider an i-KSRG that satisfies the conditions in Proposition 5.5.1. For any* $\varepsilon \in (0, \min\{\alpha - 1, \tau - 2\})$ *in* (5.5.9)*, there exists a constant* $A' > 0$*, such that for* k *sufficiently large*

$$\mathbb{P}^0\big(\mathcal{A}_{\mathrm{low-edge}}(0, n_k, N_k, \overline{w}) \mid \neg\mathcal{A}_{\mathrm{dense}}\big) \leqslant \begin{cases} A' \exp(-(1/A')k^{\zeta}), & \text{if } \alpha < \infty, \\ 0, & \text{if } \alpha = \infty. \end{cases} \tag{5.5.23}$$

Proof. Assume first $\alpha = \infty$. The event $\mathcal{A}_{\mathrm{low-edge}}(0, n_k, N_k, \overline{w}_{N_k})$ is by definition in (5.5.10) restricted to vertices of mark at most $\overline{w}_{N_k} = \overline{w}(N_k, c)$ in (5.5.9). We write t_k in (5.5.9) as $t_k = \exp((\varepsilon/d)k^{\zeta})/(2k)$, which is larger than $\beta\overline{w}_{N_k}^{1+\sigma} = \beta M_c^{1+\sigma}((c/(2c'))k^{\zeta})^{\eta(\sigma+1)}$ for k sufficiently large. Hence, the indicator in $p(u,v)$ is then 0 by (5.1.3), so a connection between u, v can not occur.

Assume $\alpha < \infty$. To obtain an upper bound on the left-hand side of (5.5.23), we condition on the full realization Ξ satisfying the event $\neg\mathcal{A}_{\mathrm{dense}}$ containing 0:

$$\begin{aligned} &\mathbb{P}^0\big(\mathcal{A}_{\mathrm{low-edge}}(0, n_k, N_k, \overline{w}) \mid \neg\mathcal{A}_{\mathrm{dense}}\big) \\ &\qquad = \mathbb{E}^0\big[\mathbb{P}^0\big(\mathcal{A}_{\mathrm{low-edge}}(0, n_k, N_k, \overline{w}_{N_k}) \mid \Xi, \neg\mathcal{A}_{\mathrm{dense}}\big)\big]. \end{aligned} \tag{5.5.24}$$

Let us denote the subgraph of $\mathcal{G}_{n_k}$ with edges of length at most $t_k = n_k^{1/d}/(2k)$ by $\mathcal{G}_{n_k}(\leqslant t_k)$ and write $\mathcal{C}_{n_k}(0, \leqslant t_k)$ for the component in this graph containing the origin. Clearly,

$$\begin{aligned} &\left\{\begin{matrix} |\mathcal{C}_{n_k}(0)| < k, \exists v_1, v_2 \in \Xi_{N_k}[1, \overline{w}_{N_k}) \text{ s.t.} \\ v_1 \in \mathcal{C}_{n_k}(0), \|x_{v_1} - x_{v_2}\| \geqslant t_k, \text{ and } v_1 \leftrightarrow v_2 \end{matrix}\right\} \\ &\qquad \subseteq \left\{\begin{matrix} |\mathcal{C}_{n_k}(0, \leqslant t_k)| < k, \exists v_1, v_2 \in \Xi_{N_k}[1, \overline{w}_{N_k}) \text{ such that} \\ v_1 \in \mathcal{C}_{n_k}(0, \leqslant t_k), \|x_{v_1} - x_{v_2}\| \geqslant t_k, \text{ and } v_1 \leftrightarrow v_2 \end{matrix}\right\}, \end{aligned}$$

where the left-hand side is the definition of $\mathcal{A}_{\mathrm{low-edge}}$ in (5.5.10). Conditionally on Ξ, all edges of length at least t_k are present independently of edges shorter than t_k. We obtain by a union bound over all vertices in $\Xi_{n_k}[1, \overline{w}_{N_k}) \subseteq \Xi$,

$$\begin{aligned}
&\mathbb{P}^0\big(\neg\mathcal{A}_{\mathrm{low-edge}}(0, n_k, N_k, \overline{w}_{N_k}) \mid \Xi, \neg\mathcal{A}_{\mathrm{dense}}\big) \\
&\quad\leqslant \sum_{v_1 \in \Xi_{n_k}[1,\overline{w}_{N_k})} \mathbb{P}^0\big(v_1 \in \mathcal{C}_{n_k}(0, \leqslant t_k), |\mathcal{C}_{n_k}(0, \leqslant t_k)| < k \mid \Xi, \neg\mathcal{A}_{\mathrm{dense}}\big) \\
&\qquad\qquad \underbrace{\sum_{\substack{v_2 \in \Xi_{N_k}[1,\overline{w}_{N_k}): \\ \|x_{v_1} - x_{v_2}\| \geqslant t_k}} p(v_1, v_2)}_{:= T(v_1)}. \qquad (5.5.25)
\end{aligned}$$

Using the definitions of p in (5.1.3), $\kappa_{1,\sigma}$ from (1.3.5), the upper bound on $|\mathcal{R}_i(v_1)|$ in $\mathcal{A}_{\mathrm{dense}}$ in (5.5.22), and mark bounds $w_{v_1}, w_{v_2} \leqslant \overline{w}_{N_k}$, and the distance bound $\|x_{v_1} - x_{v_2}\| \geqslant 2^{-(i-1)} t_k$ when $x_{v_2} \in \mathcal{R}_i(x_{v_1})$ in (5.5.21), the following bound holds uniformly for all $v_1 \in \Xi_{N_k}[1, \overline{w}_k)$:

$$\begin{aligned}
T(v_1) &\leqslant \sum_{i \geqslant 1} \sum_{\substack{v_2 \in \Xi_{N_k}[1,\overline{w}_k) \\ v_2 \in \mathcal{R}_i(v_1)}} p(v_1, v_2) \\
&\leqslant \sum_{i \geqslant 1} \sum_{\substack{v_2 \in \Xi_{N_k}[1,\overline{w}_k) \\ v_2 \in \mathcal{R}_i(v_1)}} p\beta^\alpha \kappa_{1,\sigma}^\alpha(\overline{w}_{N_k}, \overline{w}_{N_k}) 2^{-(i-1)\alpha d} t_k^{-\alpha d} \\
&\leqslant 2 \sum_{i \geqslant 1} t_k^d (2^{id} - 2^{(i-1)d}) p\beta^\alpha \kappa_{1,\sigma}^\alpha(\overline{w}_{N_k}, \overline{w}_{N_k}) 2^{-(i-1)\alpha d} t_k^{-\alpha d} \\
&= 2(2^d - 1) p\beta^\alpha \overline{w}_{N_k}^{\alpha(\sigma+1)} t_k^{(1-\alpha)d} \sum_{i \geqslant 1} 2^{-(\alpha-1)(i-1)d}. \qquad (5.5.26)
\end{aligned}$$

Since $\alpha > 1$ by assumption in Theorem 5.1.2, the sum on the right-hand side is finite. This gives a bound on $T(v_1)$ in (5.5.25) that does not depend on v_1. Hence, returning to (5.5.25),

$$\begin{aligned}
&\sum_{v_1 \in \Xi_{n_k}[1,\overline{w}_{N_k})} \mathbb{P}^0\big(v_1 \in \mathcal{C}_{n_k}(0, \leqslant t_k), |\mathcal{C}_{n_k}(0, \leqslant t_k)| < k \mid \Xi, \neg\mathcal{A}_{\mathrm{dense}}\big) \\
&\qquad = \mathbb{E}^0\Bigg[\mathbb{1}_{\{|\mathcal{C}_{n_k}(0,\leqslant t_k)| < k\}} \sum_{v_1 \in \Xi_{n_k}[1,\overline{w}_{N_k})} \mathbb{1}_{\{v_1 \in \mathcal{C}_{n_k}(0,\leqslant t_k)\}} \mid \Xi, \neg\mathcal{A}_{\mathrm{dense}}\Bigg] < k,
\end{aligned}$$

since on realizations of the graph satisfying $\{\mathcal{C}_{n_k}(0, \leqslant t_k) \leqslant k\}$, the sum that follows is at most $k-1$. Substituting this with (5.5.26) into (5.5.25) yields for (5.5.24), that for some constant $C > 0$

$$\mathbb{P}^0\big(\mathcal{A}_{\mathrm{low-edge}}(0, n_k, N_k, \overline{w}_{N_k}) \mid \neg\mathcal{A}_{\mathrm{dense}}\big) \leqslant Ck\overline{w}_{N_k}^{\alpha(\sigma+1)} t_k^{(1-\alpha)d}.$$

Substituting on the right-hand side the choices of $t_k = n_k^{1/d}/(2k)$, and $n_k = \exp(\varepsilon(c/2)k^\zeta)$ from (5.5.9), yield (5.5.23) for any $\varepsilon > 0$ (using $\overline{w}_{N_k}$ is polynomial in k). □

The last claim bounds the last term in Lemma 5.5.3. Recall $\mathcal{A}_{\mathrm{long-edge}}$ from (5.5.11).

Claim 5.5.6 (No *long* edge from a small component)**.** *Consider a supercritical i-KSRG model in Definition 5.1.1 with parameters* $\alpha > 1, \tau > 2$ *and* $d \in \mathbb{N}$. *Assume* $N \geqslant n \geqslant 1$, *and* $\overline{w} \geqslant 1$ *such that*

$$(N^{1/d} - n^{1/d})/2 \geqslant \max\big\{\sqrt{d}n^{1/d}, (\beta\overline{w}^{1+\sigma})^{1/d}, N^{1/d}/4\big\}. \tag{5.5.27}$$

There exists a constant $C_{5.5.6} > 0$ *such that*

$$\begin{aligned}\mathbb{P}^0\big(\exists v_1 \in \Xi_n[1,\overline{w}), \exists v_2 \in \Xi \setminus \Xi_N : v_1 \leftrightarrow v_2\big) \\ \leqslant C_{5.5.6}\overline{w}^{c_{5.5.6}} n N^{-\min\{\alpha-1,\tau-2\}}(1 + \mathbb{1}_{\{\alpha=\tau-1\}}\log(N)).\end{aligned} \tag{5.5.28}$$

In particular, for n_k, N_k, $\overline{w}_{N_k}$ *as in* (5.5.9), *if* $\varepsilon = \varepsilon(M_c, \eta) \in (0, \min\{\alpha - 1, \tau - 2\})$ *is sufficiently small, then for* k *sufficiently large, with* $\mathcal{A}_{\mathrm{long-edge}}$ *defined in* (5.5.11),

$$\mathbb{P}^0\big(\mathcal{A}_{\mathrm{long-edge}}(0, n_k, N_k, \overline{w}_{N_k})\big) \leqslant \exp\big(-\varepsilon k^{\zeta_{\mathrm{hh}}}\big). \tag{5.5.29}$$

Proof. We defer the proof of (5.5.28) (based on a first-moment method) to Appendix 5.B on page 207. The bound (5.5.29) follows directly from (5.5.28) by substituting n_k, N_k and $\overline{w}_{N_k}$ from (5.5.9) to (5.5.28), then using that $\overline{w}_{N_k}$ and $\log(N_k)$ are polynomial in k and of much smaller order than n_k and N_k. □

Having bounded all terms on the right-hand side in (5.5.12), we finish the section:

Proof of Proposition 5.5.1. For $n \leqslant N_k$, using that $N_k = \exp\big((c/2)k^\zeta\big)$, Proposition 5.5.1 follows directly from (5.5.1), since

$$\begin{aligned}\mathbb{P}^0\big(|\mathcal{C}_n(0)| \geqslant k, 0 \notin \mathcal{C}_n^{(1)}\big) \leqslant \mathbb{P}^0\big(|\mathcal{C}_n^{(2)}| \geqslant k\big) \leqslant N_k \exp\big(-ck^\zeta\big) \\ = \exp\big(-(c/2)k^\zeta\big).\end{aligned}$$

We now consider $n > N_k$. Recall the values of $n_k, \overline{w}_{N_k}$, and t_k from (5.5.9). Lemma 5.5.3 and Claims 5.5.4–5.5.6 directly imply (5.5.4) in Proposition 5.5.1. We will now prove the law of large numbers (5.5.5). In [134] it is shown that finite i-KSRGs $\mathcal{G}_n = (\mathcal{V}_n, \mathcal{E}_n)$ rooted at a vertex at the origin (see Definition 5.1.1) converge locally to their infinite rooted version $(\mathcal{G}_\infty, 0)$ as $n \to \infty$. We refer to [127] and its references for an introduction to local limits. We use the concept of local limits as a black box and verify a necessary and sufficient condition for the law of large numbers for the size of the giant component for graphs that have a local limit by Van der Hofstad [128, Theorem 2.2]. We state the condition: let $(\mathcal{G}_n)_{n\geqslant 1}$ be a sequence of graphs with $|\mathcal{V}_n| = n$ that converges locally in probability to $(\mathcal{G}_\infty, \varnothing)$ (here $\varnothing$ denotes the root of the graph). Assume that

$$\lim_{k\to\infty} \limsup_{n\to\infty} \frac{1}{n^2} \mathbb{E}\Big[\sum_{u,v\in\mathcal{V}_n} \mathbb{1}_{\{|\mathcal{C}_n(u)|\geqslant k, |\mathcal{C}_n(v)|\geqslant k, \mathcal{C}(u)\neq\mathcal{C}(v)\}}\Big] = 0, \quad (5.5.30)$$

then, as $n \to \infty$,

$$\frac{|\mathcal{C}_n^{(1)}|}{n} \xrightarrow{\mathbb{P}} \mathbb{P}\big(|\mathcal{C}_\infty(\varnothing)| = \infty\big).$$

To verify condition (5.5.30), we have to consider a model satisfying $|\mathcal{V}_n| = n$. Therefore, we consider i-KSRGs on Λ_n that are conditioned to have $|\mathcal{V}_n| = n$. All our previous results extend to this model, as remarked in Section 5.2.3. We analyze the indicator random variables in (5.5.30) for a fixed pair $u, v \in \mathcal{V}_n$. We distinguish whether one of them is part of the largest component to obtain

$$\begin{aligned}\{|\mathcal{C}_n(u)| \geqslant k, |\mathcal{C}_n(v)| \geqslant k,\ \mathcal{C}_n(u) \neq \mathcal{C}_n(v)\}& \\ \subseteq \{|\mathcal{C}_n(u)| \geqslant k, u \notin \mathcal{C}_n^{(1)}\} &\cup \{|\mathcal{C}_n(v)| \geqslant k, v \notin \mathcal{C}_n^{(1)}\}.\end{aligned}$$

Using linearity of expectation over the n^2 vertex pairs in the sum of (5.5.30), and that all n vertices are identically distributed, we obtain that

$$\begin{aligned}\lim_{k\to\infty} \limsup_{n\to\infty} \frac{1}{n^2}\mathbb{E}\Big[\sum_{u,v\in\mathcal{V}_n} \mathbb{1}_{\{|\mathcal{C}_n(u)|\geqslant k, |\mathcal{C}_n(v)|\geqslant k, \mathcal{C}(u)\neq\mathcal{C}(v)\}}\Big]& \\ \leqslant \lim_{k\to\infty} \limsup_{n\to\infty} 2\mathbb{P}\big(|\mathcal{C}_n(U)| \geqslant k, U \notin \mathcal{C}_n^{(1)}\big)&,\end{aligned}$$

where U is a vertex with a uniform location $X \in \Lambda_n$ and mark w_u following distribution F_W. We condition on the location $X = x$, to obtain that

$$\begin{aligned}\mathbb{P}\big(|\mathcal{C}_n(U)| \geqslant k, U \notin \mathcal{C}_n^{(1)}\big) &\leqslant \mathbb{P}\big(\|X - \partial\Lambda_n\| < N_k^{1/d}/2\big) \\ &\quad + \frac{1}{n}\int_{x\in\Lambda_n: \|x-\partial\Lambda_n\|\geqslant N_k^{1/d}/2} \mathbb{P}^x\big(|\mathcal{C}_n(u)| \geqslant k, u \notin \mathcal{C}_n^{(1)}\big)\mathrm{d}x.\end{aligned}$$

The first term tends to zero as $n \to \infty$, while the second term is of order $\exp\big(-\Omega(k^\zeta)\big)$ by Lemma 5.5.3 and Claims 5.5.4–5.5.6 that apply for all

$n \geqslant N_k$ and x such that $\|x - \partial\Lambda_n\| \geqslant N_k^{1/d}/2$. Thus, the second term tends to zero as $k \to \infty$. This finishes the proof of the condition (5.5.30). The law of large numbers (5.5.5) follows. $\square$

5.6 LOWER BOUNDS

In this section we prove Theorem 5.1.2(i). We first show that certain prerequisite inequalities imply the lower bound on the cluster-size decay, and afterwards we verify these inequalities for $\tau \in (2, 2+\sigma)$. Eventually, we prove Theorem 5.1.5. We define for some constant $\eta > 0$ the function and set

$$\begin{aligned} w_\eta(\ell) &:= (\log(\ell))^\eta, \qquad \Lambda_{\ell,\eta} := \Lambda_\ell \times [1, w_\eta(\ell)), \\ \mathcal{C}_{\ell,\eta}(0) &:= \text{the component containing } (0, w_0) \text{ in } \mathcal{G}_\ell[1, w_\eta(\ell)), \end{aligned} \tag{5.6.1}$$

i.e., $\mathcal{C}_{\ell,\eta}(0)$ is the component of $(0, w_0)$ in the graph induced on vertices in $\Lambda_{\ell,\eta}$ – and let $\mathcal{C}_{\ell,\eta}(0) := \emptyset$ if $w_0 \geqslant w_\eta(\ell)$. We recall $\mathfrak{m}(\mathcal{Z})$, the multiplicity of the maximum of $\mathcal{Z}$, from (5.1.5); and the exponents $\zeta_{ll}, \zeta_{lh}, \zeta_{hh}, \zeta_{nn}$ from (4.2.1), (4.2.2), (4.2.4), and (4.2.5). Define for some small constant $\varepsilon > 0$ to be defined later, and $\mathcal{Z} = \{\zeta_{ll}, \zeta_{lh}, \zeta_{hh}, \zeta_{nn}\}$

$$\underline{k}_{n,\varepsilon} := \begin{cases} \left(\varepsilon \log(n)/(\log\log(n))^{\mathfrak{m}_{\mathcal{Z}}-1}\right)^{1/\max(\mathcal{Z})}, & \text{if } \max(\mathcal{Z}) > 0, \\ \exp\left((\varepsilon \log(n))^{1/(\mathfrak{m}_{\mathcal{Z}}-1)}\right), & \text{if } \max(\mathcal{Z}) = 0, \mathfrak{m}_{\mathcal{Z}} > 1. \end{cases} \tag{5.6.2}$$

Note that $\zeta_{nn} = \frac{d-1}{d}$, which is equal to 0 only if $d = 1$, and positive otherwise. Thus, only in dimension $d = 1$ one can have $\underline{k}_{n,\varepsilon}$ different from $\Theta(\text{polylog}(n))$. More precisely, for $d = 1$, $\underline{k}_{n,\varepsilon} = \Theta(\text{poly}(n))$ if only one out of $\{\zeta_{ll}, \zeta_{lh}, \zeta_{hh}\}$ is zero, and the others are negative; and it is stretched exponential in the logarithm if at least two elements out of $\{\zeta_{ll}, \zeta_{lh}, \zeta_{hh}\}$ are zero, and none of them is positive. If the three values $\zeta_{ll}, \zeta_{lh}, \zeta_{hh}$ are all negative, then the model is always subcritical when $d = 1$ [108].

By an "i-KSRG on $\mathbb{Z}^d$" we mean the model mentioned in Section 5.2.3, i.e., when Ξ is replaced by vertices in $\mathbb{Z}^d$ having iid marks from distribution F_W in (5.1.2) in Definition 5.1.1.

Proposition 5.6.1 (Lower bound holds when linear-sized giant on truncated weights exists). *Consider a supercritical i-KSRG model in Definition 5.1.1 with parameters* $\alpha \in (1, \infty]$, $\tau \in (2, \infty]$, $\sigma \geqslant 0$, $d \in \mathbb{N}$, *or a supercritical i-KSRG on* $\mathbb{Z}^d$ *with additionally* $\min\{p, p\beta^\alpha\} < 1$ *in* (5.1.3). *Assume that there exist* *constants* $\eta, \rho > 0$ *such that for all* ℓ *sufficiently large, and with* $\mathcal{C}_{\ell,\eta}(0)$ *from* (5.6.1),

$$\mathbb{P}^0\big(|\mathcal{C}_{\ell,\eta}(0)| \geqslant \rho\ell\big) \geqslant \rho. \tag{5.6.3}$$

Then there exists $A > 0$ *such that for all* $n \in [Ak, \infty]$, *with* $\mathcal{Z} = \{\zeta_{ll}, \zeta_{lh}, \zeta_{hh}, \zeta_{nn}\}$,

$$\mathbb{P}^0\big(|\mathcal{C}_n(0)| \geqslant k, 0 \notin \mathcal{C}_n^{(1)}\big) \geqslant \exp\big(-Ak^{\max(\mathcal{Z})}\log^{m_{\mathcal{Z}}-1}(k)\big). \tag{5.6.4}$$

Moreover, there exists $\delta, \varepsilon > 0$, *such that for all* n *sufficiently large, with* $k_{n,\varepsilon}$ *from* (5.6.2),

$$\mathbb{P}\big(|\mathcal{C}_n^{(2)}| > \underline{k}_{n,\varepsilon}\big) \geqslant 1 - n^{-\delta}. \tag{5.6.5}$$

By Proposition 5.4.13, the condition (5.6.3) is satisfied when $\tau \in (2, 2+\sigma)$, implying Theorem 5.1.2(i). In [151], we show that (5.6.3) holds whenever $\max\{\zeta_{ll}, \zeta_{lh}, \zeta_{hh}\} > 0$ and the model is supercritical. We prove Proposition 5.6.1 by formalizing the reasoning at the beginning of Section 4.2. First, we assume that the vertex set is given by the PPP Ξ in Definition 5.1.1. Then we explain how to extend to i-KSRGs on $\mathbb{Z}^d$.

5.6.1 *Finding a localized component*

The goal of this section is to prove Proposition 5.6.1. First we aim to bound the probability of $\{\mathcal{C}_n(0) \geqslant k, 0 \notin \mathcal{C}_n^{(1)}\}$ in (5.6.4) from below. To do so, we write it as the intersection of "almost independent" events, for which we introduce some notation now:

Two large components.

We write $\Lambda(x, s) = \Lambda_s(x)$ for a box of volume s centered at x, see (5.1.1). We define for a constant $M_{in} > 0$ the radius and Euclidean ball

$$r_k := (kM_{in})^{1/d}, \qquad \mathcal{B}_{kM_{in}} := \{x \in \mathbb{R}^d : \|x\| \leqslant r_k\}, \tag{5.6.6}$$

where $M_{in} > 1/\rho$ is implicitly defined given ρ from Proposition 5.6.1, such that $\|\partial\Lambda_{k/\rho} - \partial\mathcal{B}_{kM_{in}}\| = r_k/2$, implying also that $\Lambda_{k/\rho} \subseteq \mathcal{B}_{kM_{in}}$. We would like to constrain $\mathcal{C}_n(0)$ to the ball $\mathcal{B}_{kM_{in}}$, and we would like to find a component outside $\mathcal{B}_{kM_{in}}$ that is larger than $|\mathcal{C}_n(0)|$. We 'construct' these

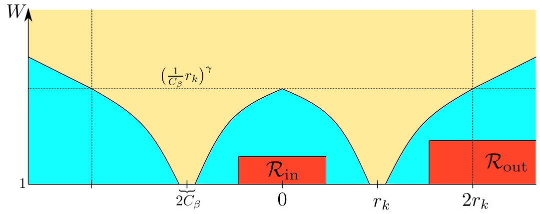


Figure 8: A visualization of the γ-suppressed profile $\mathcal{M}_\gamma$. The horizontal axis represents space, the vertical axis represents marks. $\{\Xi \leqslant \mathcal{M}_\gamma\}$ demands no vertices in the yellow region. $\mathcal{A}_{\text{noedge}}(\gamma)$ demands that there is no edge between vertices in the inner and vertices in the outer blue region, $\mathcal{A}_{\text{components}}$ requires that the two red areas contain large components; $\mathcal{A}_{\text{regular}}(\eta)$ ensures that Ξ is 'close to being typical' in the red areas $\mathcal{R}_{\text{in}}$ and $\mathcal{R}_{\text{out}}$.

two components, with the constant $M_{\text{out}} := 2^{d+2} M_{\text{in}}$, on vertices in the (hyper)rectangles (see also (5.6.1))

$$\begin{aligned} \mathcal{R}_{\text{in}} &:= \Lambda_{k/\rho,\eta} = \Lambda(0, k/\rho) \times \left[1, w_\eta(k/\rho)\right), \\ \mathcal{R}_{\text{out}} &:= \Lambda(x_{\text{out}}, kM_{\text{out}}/\rho) \times \left[1, w_\eta(kM_{\text{out}}/\rho)\right), \\ \Lambda_{\text{in}} &:= \Lambda(0, k/\rho), \\ \Lambda_{\text{out}} &:= \Lambda(x_{\text{out}}, (kM_{\text{out}}/\rho)), \end{aligned} \tag{5.6.7}$$

where $x_{\text{out}} := (x_1^{\text{out}}, 0, \ldots, 0) \in \mathbb{R}^d$ is defined implicitly given ρ such that $\|\partial\Lambda(x_{\text{out}}, kM_{\text{out}}/\rho) - \partial\mathcal{B}_{kM_{\text{in}}}\| := r_k/2$. We assume that n is sufficiently large so that $\Lambda_{\text{in}} \cup \Lambda_{\text{out}} \subseteq \Lambda_n$. Let $\mathcal{C}_{\text{in}}(0)$ be the component of $(0, w_0)$ in the subgraph of $\mathcal{G}_n$ induced on vertices in $\mathcal{R}_{\text{in}}$, where $\mathcal{C}_{\text{in}}(0) = \emptyset$ if $w_0 \geqslant w_\eta(k/\rho)$. Since $\mathcal{R}_{\text{in}} \subseteq \mathcal{B}_{kM_{\text{in}}}$, it is immediate that $\mathcal{C}_{\text{in}}(0) \subseteq \Xi_{\mathcal{B}_{kM_{\text{in}}}}$. Let $\mathcal{C}_{\text{out}}^{(1)}$ be the largest component in the subgraph of $\mathcal{G}_n$ induced on vertices in $\mathcal{R}_{\text{out}}$. Define the events

$$\begin{aligned} \mathcal{A}_{\text{ingiant}}^{(k)} &:= \{|\mathcal{C}_{\text{in}}(0)| \geqslant k\}, & \mathcal{A}_{\text{outgiant}}^{(k)} &:= \{|\mathcal{C}_{\text{out}}^{(1)}| \geqslant kM_{\text{out}}\}, \\ \mathcal{A}_{\text{insmall}}^{(k)} &:= \{|\Xi_{\mathcal{B}_{kM_{\text{in}}}}| \leqslant kM_{\text{out}}/2\}, & \mathcal{A}_{\text{components}}^{(k)} &:= \mathcal{A}_{\text{ingiant}}^{(k)} \cap \mathcal{A}_{\text{outgiant}}^{(k)}. \end{aligned} \tag{5.6.8}$$

Isolation

On $\mathcal{A}_{\text{components}}^{(k)}$, $\mathcal{C}_n(0) \supseteq \mathcal{C}_{\text{in}}(0)$ could still connect to the giant/infinite component. To prevent this, we will ban edges that cross the boundary of $\mathcal{B}_{kM_{\text{in}}}$ as follows. In Section 5.2.2 we informally described an "optimally-suppressed mark-profile". We first define a suppressed mark-profile, then optimise its shape. Let $\gamma \in (0, 1/(\sigma+1)]$ be a constant. We define, for $C_\beta := (2\beta)^{1/d}$, and for $x \in \mathbb{R}^d$ with $\|x - \partial\mathcal{B}_{kM_{\text{in}}}\| = |\|x\| - r_k| =: z$, the γ-suppressed profile by

$$f_\gamma(z) := \begin{cases} 1 & \text{if } z \leqslant C_\beta, \\ (z/C_\beta)^{\gamma d} & \text{if } z \in (C_\beta, r_k], \\ (z/C_\beta)^d (r_k/C_\beta)^{-d(1-\gamma)} & \text{if } z > r_k, \end{cases} \tag{5.6.9}$$

$$\mathcal{M}_\gamma := \{(x_v, f_\gamma(|\|x_v\| - r_k|)) : x_v \in \mathbb{R}^d\}. \tag{5.6.10}$$

We say that v is *below, on, or above* $\mathcal{M}_\gamma$ if w_v is at most, equal to, or strictly larger than $f_\gamma(|\|x_v\| - r_k|)$, respectively. We say $\Xi \leqslant \mathcal{M}_\gamma$ if all points in Ξ are below $\mathcal{M}_\gamma$, see Figure 8. We split the PPP Ξ into four independent PPPs, whether points fall below or above $\mathcal{M}_\gamma$, and inside or outside $\mathcal{B}_{kM_{\text{in}}}$:

$$\begin{aligned} \Xi^{\text{in}}_{\leqslant\mathcal{M}_\gamma} &:= \{(x_u, w_u) \in \Xi : x_u \in \mathcal{B}_{kM_{\text{in}}}, w_u \leqslant f_\gamma(|\|x_u\| - r_k|)\}, \\ \Xi^{\text{out}}_{\leqslant\mathcal{M}_\gamma} &:= \{(x_v, w_v) \in \Xi : x_v \notin \mathcal{B}_{kM_{\text{in}}}, w_v \leqslant f_\gamma(|\|x_v\| - r_k|)\}, \\ \Xi^{\text{in}}_{>\mathcal{M}_\gamma} &:= \{(x_u, w_u) \in \Xi : x_u \in \mathcal{B}_{kM_{\text{in}}}, w_u > f_\gamma(|\|x_u\| - r_k|)\}, \\ \Xi^{\text{out}}_{>\mathcal{M}_\gamma} &:= \{(x_v, w_v) \in \Xi : x_v \notin \mathcal{B}_{kM_{\text{in}}}, w_v > f_\gamma(|\|x_v\| - r_k|)\}. \end{aligned} \tag{5.6.11}$$

For $A, B \subseteq \Xi$ we denote by $|\mathcal{E}(A, B)|$ the number of edges between vertices in A and B. Define

$$\{\Xi \leqslant \mathcal{M}_\gamma\} = \{|\Xi^{\text{in}}_{>\mathcal{M}_\gamma} \cup \Xi^{\text{out}}_{>\mathcal{M}_\gamma}| = 0\}, \quad \mathcal{A}^{(k)}_{\text{noedge}}(\gamma) := \{|\mathcal{E}(\Xi^{\text{in}}_{\leqslant\mathcal{M}_\gamma}, \Xi^{\text{out}}_{\leqslant\mathcal{M}_\gamma})| = 0\}. \tag{5.6.12}$$

Under $\{\Xi \leqslant \mathcal{M}_\gamma\} \cap \mathcal{A}^{(k)}_{\text{noedge}}(\gamma)$, the vertices in $\mathcal{B}_{kM_{\text{in}}}$ are not connected to the unique infinite component when $n = \infty$ and are isolated from the rest of $\mathcal{G}_n$ when $n < \infty$. We comment on the profile function f_γ in (5.6.9). The event $\{\Xi \leqslant \mathcal{M}_\gamma\}$ demands no vertices within distance C_β from $\partial\mathcal{B}_{M_{\text{in}}}$, since $f_\gamma(|\|x\| - r_k|) = 1$ for $\|x - \partial\mathcal{B}_{kM_{\text{in}}}\| \leqslant C_\beta$, and vertex marks are above 1 a.s. The function f_γ is continuous and increasing in z: the closer a point x is to the boundary of $\mathcal{B}_{kM_{\text{in}}}$, the stronger the restriction on its mark in $\Xi \leqslant \mathcal{M}_\gamma$. This is natural since vertices with higher mark close to $\partial\mathcal{B}_{kM_{\text{in}}}$ are more likely to have an edge crossing this boundary, which we want to prevent. While the event $\{\Xi \leqslant \mathcal{M}_\gamma\}$ becomes less likely when γ is small, $\mathcal{A}^{(k)}_{\text{noedge}}(\gamma)$ becomes more likely. This naturally leads to an optimization problem, that we set up next. Let $\mathcal{Y} = \{0, \alpha - (\tau - 1), (\sigma + 1)\alpha - 2(\tau - 1)\}$, and define

$$\gamma_{\neg\mathrm{nn}} := \begin{cases} \max\limits_{0<\gamma\leqslant 1/(\sigma+1)} \{\gamma : 1 - \gamma(\tau - 1) \overset{(\star)}{\geqslant} 2 - \alpha + \gamma \max(\mathcal{Y})\}, & \text{if } \alpha < \infty, \\ 1/(\sigma + 1), & \text{if } \alpha = \infty, \end{cases}$$
$$\zeta_{\neg\mathrm{nn}} := 1 - \gamma_{\neg\mathrm{nn}}(\tau - 1), \tag{5.6.13}$$

The inequality $(\star)$ holds for $\gamma = 0$ since $\alpha > 1$ by assumption, therefore $\gamma_{\neg\mathrm{nn}} > 0$ and is well-defined. We call $f_\star := f_{\gamma_{\neg\mathrm{nn}}}$ the *optimally-suppressed mark-profile* (and show below that it is indeed "optimal") and write $\mathcal{M}_\star := \mathcal{M}_{\gamma_{\neg\mathrm{nn}}}$. We define the isolation event using the events in (5.6.12), see also Figure 8:

$$\mathcal{A}^{(k)}_{\mathrm{isolation}} := \{\Xi \leqslant \mathcal{M}_\star\} \cap \mathcal{A}^{(k)}_{\mathrm{noedge}}(\gamma_{\neg\mathrm{nn}}). \tag{5.6.14}$$

Ensuring almost independence

The events $\mathcal{A}^{(k)}_{\mathrm{isolation}}$ and $\{|\mathcal{C}_{\mathrm{in}}(0)| \geqslant k\}$ in (5.6.8) are correlated, since $\{|\mathcal{C}_{\mathrm{in}}(0)| \geqslant k\}$ may push up the number of high-mark vertices in $\mathcal{R}_{\mathrm{in}}$, making $\mathcal{A}^{(k)}_{\mathrm{isolation}}$ less likely. To overcome the dependence, we introduce two auxiliary events that ensure regularity of Ξ in the hyperrectangles $\mathcal{R}_{\mathrm{in}}, \mathcal{R}_{\mathrm{out}}$ from (5.6.7). Using w_η from (5.6.1), define, using $c_{\mathrm{in}} := 1/\rho, c_{\mathrm{out}} := M_{\mathrm{out}}/\rho$:

$$\begin{aligned} j^\star_{\mathrm{in}} &:= \min\{j \in \mathbb{N} : w_\eta(kc_{\mathrm{in}})/2^j < 1\}, \\ j^\star_{\mathrm{out}} &:= \min\{j \in \mathbb{N} : w_\eta(kc_{\mathrm{out}})/2^j < 1)\}. \end{aligned} \tag{5.6.15}$$

and writing $\mathrm{loc} \in \{\mathrm{in}, \mathrm{out}\}$, define

$$I^{\mathrm{loc}}_j := [w_\eta(kc_{\mathrm{loc}})/2^j, w_\eta(kc_{\mathrm{loc}})/2^{j-1}), \qquad 1 \leqslant j < j^\star_{\mathrm{loc}}. \tag{5.6.16}$$

Using $\Lambda_{\mathrm{in}}, \Lambda_{\mathrm{out}}$ in (5.6.7), the intensity measure μ_τ of Ξ in (5.1.2), and $\Xi_{\mathrm{loc}}(I^{\mathrm{loc}}_j)$ for the vertices in $\Xi \cap (\Lambda_{\mathrm{loc}} \times I^{\mathrm{loc}}_j)$, consider the following events (see also Figure 8) for $\mathrm{loc} \in \{\mathrm{in}, \mathrm{out}\}$:

$$\begin{aligned} \mathcal{A}^{(k,\mathrm{loc})}_{\mathrm{regular}}(\eta) &:= \{\forall j \leqslant j^\star_{\mathrm{loc}} : |\Xi_{\mathrm{loc}}(I^{\mathrm{loc}}_j)| \leqslant 2\mu_\tau(\Lambda_{\mathrm{loc}} \times I^{\mathrm{loc}}_j)\}, \\ \mathcal{A}^{(k)}_{\mathrm{regular}}(\eta) &:= \mathcal{A}^{(k,\mathrm{in})}_{\mathrm{regular}}(\eta) \cap \mathcal{A}^{(k,\mathrm{out})}_{\mathrm{regular}}(\eta). \end{aligned} \tag{5.6.17}$$

Finally, fix a realization of the induced subgraphs

$$\mathcal{G}_{\mathcal{R}_{\mathrm{in}}} \cup \mathcal{G}_{\mathcal{R}_{\mathrm{out}}} = (\Xi_{\mathcal{R}_{\mathrm{in}}}, \mathcal{E}(\mathcal{G}_{\mathcal{R}_{\mathrm{in}}})) \cup (\Xi_{\mathcal{R}_{\mathrm{out}}}, \mathcal{E}(\mathcal{G}_{\mathcal{R}_{\mathrm{out}}}))$$

so that the vertex set $\Xi_{\mathcal{R}_{\text{in}}} \cup \Xi_{\mathcal{R}_{\text{out}}}$ satisfies the event $\mathcal{A}^{(k)}_{\text{regular}}(\eta)$ for some $\eta > 0$, and define the conditional probability measure and expectation by

$$\begin{aligned}\widetilde{\mathbb{P}}_{\text{io}}(\,\cdot\,) &:= \mathbb{P}\big(\cdot \mid \mathcal{G}_{\mathcal{R}_{\text{in}}} \cup \mathcal{G}_{\mathcal{R}_{\text{out}}}, \mathcal{A}^{(k)}_{\text{regular}}(\eta)\big),\\ \widetilde{\mathbb{E}}_{\text{io}}[\,\cdot\,] &:= \widetilde{\mathbb{E}}_{\text{io}}\big[\cdot \mid \mathcal{G}_{\mathcal{R}_{\text{in}}} \cup \mathcal{G}_{\mathcal{R}_{\text{out}}}, \mathcal{A}^{(k)}_{\text{regular}}(\eta)\big].\end{aligned} \tag{5.6.18}$$

In the conditioning we reveal both the vertex and edge sets in the disjoint boxes $\mathcal{R}_{\text{in}}, \mathcal{R}_{\text{out}}$. The event $\mathcal{A}^{(k)}_{\text{regular}}(\eta)$ checks the number of vertices in hyperrectangles inside $\mathcal{R}_{\text{in}}, \mathcal{R}_{\text{out}}$, hence $\mathcal{A}^{(k)}_{\text{regular}}(\eta)$ is measurable with respect to $\mathcal{G}_{\mathcal{R}_{\text{in}}}, \mathcal{G}_{\mathcal{R}_{\text{out}}}$ and in principle can be left out of the conditioning in (5.6.18). We state a lemma that will imply Proposition 5.6.1.

Lemma 5.6.2 (Lower bound for isolation)**.** *Consider a supercritical i-KSRG model in Definition 5.1.1 with parameters* $\alpha \in (1, \infty]$, $\tau \in (2, \infty]$, $\sigma \geqslant 0$, *and* $d \in \mathbb{N}$, *or a supercritical i-KSRG on* $\mathbb{Z}^d$ *with additionally* $p < 1$ *in* (5.1.3). *For any constant* $\eta > 0$ *in* (5.6.1), *there exists* $A > 0$, *such that for any realization of* $\mathcal{G}_{\mathcal{R}_{\text{in}}} \cup \mathcal{G}_{\mathcal{R}_{\text{out}}}$ *that satisfies* $\mathcal{A}^{(k)}_{\text{regular}}(\eta)$, *with* $\mathcal{Z} = \{\zeta_{\text{ll}}, \zeta_{\text{lh}}, \zeta_{\text{hh}}, \zeta_{\text{nn}}\}$, *with* $\widetilde{\mathbb{P}}_{\text{io}}$ *from* (5.6.18)

$$\widetilde{\mathbb{P}}_{\text{io}}\big(\mathcal{A}^{(k)}_{\text{isolation}}\big) \geqslant \exp\big(-A r_k^{d\max(\mathcal{Z})} \log^{m_{\mathcal{Z}}-1}(r_k)\big). \tag{5.6.19}$$

The same bound holds for the Palm-version $\widetilde{\mathbb{P}}^0_{\text{io}}$ *of* $\widetilde{\mathbb{P}}_{\text{io}}$.

We now show that Proposition 5.6.1 follows directly from Lemma 5.6.2.

Proof of Proposition 5.6.1, assuming Lemma 5.6.2. We first show (5.6.4). Recall $\mathcal{A}^{(k)}_{\text{components}}, \mathcal{A}^{(k)}_{\text{insmall}}$ from (5.6.8), and $\mathcal{A}^{(k)}_{\text{isolation}}, \mathcal{A}^{(k)}_{\text{regular}}(\eta)$ from (5.6.14), (5.6.17) respectively. The intersection of these events implies $\{|\mathcal{C}_n(0)| \geqslant k, 0 \notin \mathcal{C}^{(1)}_n\}$, since $|\mathcal{C}_n(0)| \geqslant |\mathcal{C}_{\text{in}}(0)| \geqslant k$, and $\mathcal{A}^{(k)}_{\text{isolation}}$ ensures that $\mathcal{C}_n(0)$ is fully contained in $\mathcal{B}_{kM_{\text{in}}}$. Hence, $|\mathcal{C}_n(0)| \leqslant |\Xi_{\mathcal{B}_{kM_{\text{in}}}}| \leqslant kM_{\text{out}}/2$ and $|\mathcal{C}_{\text{out}}| \geqslant kM_{\text{out}}$ ensure that $\mathcal{C}_n(0)$ is not the largest component of $\mathcal{G}_n$. So, by the law of total probability

$$\begin{aligned}&\mathbb{P}^0\big(|\mathcal{C}_n(0)| \geqslant k, 0 \notin \mathcal{C}^{(1)}_n\big)\\ &\quad\geqslant \mathbb{P}^0\big(\mathcal{A}^{(k)}_{\text{components}} \cap \mathcal{A}^{(k)}_{\text{isolation}} \cap \mathcal{A}^{(k)}_{\text{insmall}} \cap \mathcal{A}^{(k)}_{\text{regular}}(\eta)\big)\\ &\quad\geqslant \mathbb{P}^0\big(\mathcal{A}^{(k)}_{\text{components}} \cap \mathcal{A}^{(k)}_{\text{isolation}} \cap \mathcal{A}^{(k)}_{\text{regular}}(\eta)\big) - \mathbb{P}^0\big(\neg\mathcal{A}^{(k)}_{\text{insmall}}\big)\\ &\quad= \mathbb{P}^0\big(\mathcal{A}^{(k)}_{\text{components}} \cap \mathcal{A}^{(k)}_{\text{regular}}(\eta)\big)\mathbb{P}^0\big(\mathcal{A}^{(k)}_{\text{isolation}} \mid \mathcal{A}^{(k)}_{\text{components}} \cap \mathcal{A}^{(k)}_{\text{regular}}(\eta)\big)\\ &\qquad - \mathbb{P}^0\big(\neg\mathcal{A}^{(k)}_{\text{insmall}}\big).\end{aligned} \tag{5.6.20}$$

Recall $\mathcal{A}^{(k)}_{\text{insmall}} = \{|\Xi_{\mathcal{B}_{kM_{\text{in}}}}| \leqslant kM_{\text{out}}/2\}$ from (5.6.8), and $M_{\text{out}} = 2^{d+2}M_{\text{in}}$ above (5.6.7). The box with side-length $2r_k = 2(kM_{\text{in}})^{1/d}$ (by definition in

(5.6.6)) centered at the origin is the smallest box that contains $\mathcal{B}_{kM_{\text{in}}}$. Using the intensity measure μ_τ from (5.1.2), and writing $\mathcal{B}_{kM_{\text{in}}}^{\leqslant} := \{(x, w_x) \in \mathbb{R}^{d+1} : x \in \mathcal{B}_{kM_{\text{in}}}; (x, w_v) \leqslant \mathcal{M}_\gamma\}$, we have

$$\mu_\tau(\mathcal{B}_{kM_{\text{in}}}^{\leqslant}) \leqslant 2^d kM_{\text{in}} = 2^{d+2}kM_{\text{in}}/4 = M_{\text{out}}k/4,$$

and by a standard concentration inequality for Poisson random variables (see Lemma 5.C.1 for $x = 2$), there exist $c', c > 0$ such that, since $r_k = \Theta(k^{1/d})$,

$$\mathbb{P}^0\big(\neg\mathcal{A}_{\text{insmall}}^{(k)}\big) \leqslant \exp(-c' r_k^d) = \exp(-ck). \tag{5.6.21}$$

Returning to (5.6.20), the event $\mathcal{A}_{\text{regular}}^{(k)}(\eta) = \mathcal{A}_{\text{regular}}^{(k,\text{in})}(\eta) \cap \mathcal{A}_{\text{regular}}^{(k,\text{out})}(\eta)$, defined in (5.6.17), holds with probability tending to 1 as $k \to \infty$, again by concentration inequalities for Poisson random variables (see Lemma 5.C.1 for $x = 2$). Hence,

$$\begin{aligned}\mathbb{P}^0\big(\mathcal{A}_{\text{components}}^{(k)} \cap \mathcal{A}_{\text{regular}}^{(k)}(\eta)\big) &\geqslant \mathbb{P}^0\big(\mathcal{A}_{\text{components}}^{(k)}\big) - \mathbb{P}^0\big(\neg(\mathcal{A}_{\text{regular}}^{(k)}(\eta)\big) \\ &= \mathbb{P}^0\big(\mathcal{A}_{\text{components}}^{(k)}\big) - o_k(1). \end{aligned}\tag{5.6.22}$$

We recall from (5.6.8) that $\mathcal{A}_{\text{components}}^{(k)} = \{|\mathcal{C}_{\text{in}}(0)| \geqslant k\} \cap \{|\mathcal{C}_{\text{out}}^{(1)}| \geqslant kM_{\text{out}}\}$. Translate the hyperrectangle $\mathcal{R}_{\text{out}}$ in (5.6.7) containing $\mathcal{C}_{\text{out}}^{(1)}$ to the origin of $\mathbb{R}^d$ (using $w_\eta(\ell)$) in (5.6.1)):

$$\mathcal{R}_{\text{out}}' := \Lambda(0, kM_{\text{out}}/\rho) \times [1, w_\eta(kM_{\text{out}}/\rho)) = \Lambda_{kM_{\text{out}}/\rho,\eta}, \tag{5.6.23}$$

and write $\mathcal{C}_{\text{out}'}^{(1)}, \mathcal{C}_{\text{out}'}(0)$ for the largest component and for the component containing $(0, w_0)$ in the subgraph of $\mathcal{G}_n$ induced by vertices in $\mathcal{R}_{\text{out}}'$. As before, we ignore the conditioning $(0, w_0) \in \Xi$ in Definition 5.1.1 in our computations. By the translation invariance of the annealed measure and that the events $\{|\mathcal{C}_{\text{in}}(0)| \geqslant k\}$ and $\{|\mathcal{C}_{\text{out}}^{(1)}| \geqslant M_{\text{out}}k\}$ are independent because they are induced subgraphs of the disjoint hyperrectangles $\mathcal{R}_{\text{in}}$ and $\mathcal{R}_{\text{out}}$ in (5.6.7),

$$\begin{aligned}\mathbb{P}^0\big(\mathcal{A}_{\text{components}}^{(k)}\big) &= \mathbb{P}^0\big(|\mathcal{C}_{\text{in}}(0)| \geqslant k\big)\mathbb{P}\big(|\mathcal{C}_{\text{out}'}^{(1)}| \geqslant M_{\text{out}}k\big) \\ &\geqslant \mathbb{P}^0\big(|\mathcal{C}_{\text{in}}(0)| \geqslant k\big)\mathbb{P}^0\big(|\mathcal{C}_{\text{out}'}(0)| \geqslant M_{\text{out}}k\big).\end{aligned}$$

The bound $\mathbb{P}^0\big(|\mathcal{C}_{\ell,\eta}(0)| \geqslant \rho\ell\big) \geqslant \rho$ in (5.6.3) in Proposition 5.6.1 holds for all ℓ sufficiently large. In particular, since $\mathcal{C}_{\text{in}}(0) = \mathcal{C}_{k/\rho,\eta}(0)$ by definition of $\mathcal{R}_{\text{in}}$ in (5.6.7), and $\mathcal{C}_{\text{out}'}(0) = \mathcal{C}_{kM_{\text{out}}/\rho,\eta}(0)$ by definition of $\mathcal{R}_{\text{out}}'$ in (5.6.23), we obtain for k sufficiently large

$$\begin{aligned}\mathbb{P}^0\big(\mathcal{A}_{\text{components}}^{(k)}\big) &\geqslant \mathbb{P}^0\big(|\mathcal{C}_{k/\rho,\eta}(0)| \geqslant k\big)\mathbb{P}^0\big(|\mathcal{C}_{kM_{\text{out}}/\rho,\eta}(0)| \geqslant M_{\text{out}}k\big) \\ &\geqslant \rho^2,\end{aligned}\tag{5.6.24}$$

implying that $\mathbb{P}^0\big(\mathcal{A}^{(k)}_{\text{components}} \cap \mathcal{A}^{(k)}_{\text{regular}}(\eta)\big) \geqslant \rho^2 - o_k(1) > 3\rho^2/4$ in (5.6.22). Since the event $\mathcal{A}^{(k)}_{\text{components}} \cap \mathcal{A}^{(k)}_{\text{regular}}(\eta)$ is measurable with respect to the σ-algebra generated by the subgraph $\mathcal{G}_{\mathcal{R}_{\text{in}}} \cup \mathcal{G}_{\mathcal{R}_{\text{out}}}$, we take expectation over all possible realizations of the latter, and recalling the measure $\widetilde{\mathbb{P}}_{\text{io}}$ from (5.6.18),

$$\begin{aligned}\mathbb{P}^0\big(\mathcal{A}^{(k)}_{\text{isolation}} \mid \mathcal{A}^{(k)}_{\text{components}} \cap \mathcal{A}^{(k)}_{\text{regular}}(\eta)\big) &\\ = \mathbb{E}^0\Big[\mathbb{P}^0\big(\mathcal{A}^{(k)}_{\text{isolation}} \mid \mathcal{G}_{\mathcal{R}_{\text{in}}} \cup \mathcal{G}_{\mathcal{R}_{\text{out}}}, \mathcal{A}^{(k)}_{\text{regular}}(\eta)\big)\Big]&\\ = \mathbb{E}^0\Big[\widetilde{\mathbb{P}}^0_{\text{io}}(\mathcal{A}^{(k)}_{\text{isolation}})\Big].&\end{aligned}$$

Apply now Lemma 5.6.2 on the right-hand side, and substitute the bound $3\rho^2/4$ below (5.6.24) into (5.6.22) and then in turn into (5.6.20) and (5.6.21), to obtain for k sufficiently large

$$\begin{aligned}\mathbb{P}^0\big(\mathcal{A}^{(k)}_{\text{components}} \cap \mathcal{A}^{(k)}_{\text{isolation}} \cap \mathcal{A}^{(k)}_{\text{insmall}} \cap \mathcal{A}^{(k)}_{\text{regular}}(\eta)\big)&\\ \geqslant (\rho^2/2)\exp\big(-Ar_k^{d\max(\mathcal{Z})}\log^{m_{\mathcal{Z}}-1}(r_k)\big)&\\ \geqslant (\rho^2/2)\exp\big(-A'k^{\max(\mathcal{Z})}\log^{m_{\mathcal{Z}}-1}(k)\big).&\end{aligned} \tag{5.6.25}$$

We obtained the second row by substituting $r_k = (kM_{\text{in}})^{1/d}$ in (5.6.6) and setting $A' := AM_{\text{in}}^{\max(\mathcal{Z})}/2$ that also compensates for the constants from the log-correction term. Observe that $\max(\mathcal{Z}) < 1$, so that the right-hand side in (5.6.21) is of smaller order than the right-hand side in (5.6.19). By (5.6.20) this finishes the proof of (5.6.4). We turn to the proof of (5.6.5).

Lower bound on second-largest component. We generalise an argument from [162]. We have to bound $\mathbb{P}\big(|\mathcal{C}_n^{(2)}| \leqslant \underline{k}_{n,\varepsilon}\big)$ from above for a suitably chosen ε in the definition of $\underline{k}_{n,\varepsilon}$ in (5.6.2). To do so, we fix $\vartheta > 0$ to be specified later, and assume for simplicity that $n^{(1-\vartheta)/d} \in \mathbb{N}$. We then partition Λ_n into $m_n := n^{1-\vartheta}$ many subboxes $\Lambda^{(1)}_{\tilde n}, \dots, \Lambda^{(m_n)}_{\tilde n}$, centered respectively at $x_1, \dots, x_{m_n}$, each of volume $\tilde n := n^{\vartheta}$. By disjointness, the induced subgraphs $\mathcal{G}^{(1)}_{\tilde n}, \dots, \mathcal{G}^{(m_n)}_{\tilde n}$ in these boxes are independent realizations of $\mathcal{G}_{\tilde n}$, translated to $x_1, \dots, x_{m_n}$. We write $\widetilde{\mathcal{G}}^{(1)}_{\tilde n}, \dots, \widetilde{\mathcal{G}}^{(m_n)}_{\tilde n}$ for the induced subgraphs, translated back to the origin; $\widetilde{\Xi}^{(\text{in},i)}_{\leqslant \mathcal{M}_\star}$ for the vertex set in $\widetilde{\mathcal{G}}^{(i)}_{\tilde n}$ that is below $\mathcal{M}_\star$ and inside $\mathcal{B}_{kM_{\text{in}}}$ after the translation, see (5.6.11); and write $\Xi^{(\text{in},i)}_{\leqslant \mathcal{M}_\star}$ for the same vertex set before the translation. For the translated subgraphs $\big(\widetilde{\mathcal{G}}^{(i)}_{\tilde n}\big)_{i \leqslant m_n}$, we define for $k = \underline{k}_{n,\varepsilon}$ the same events as in (5.6.8), (5.6.14), (5.6.17),

$$\mathcal{A}^{(i)}_{\text{good}} := \mathcal{A}^{(\underline{k}_{n,\varepsilon}),i}_{\text{components}} \cap \mathcal{A}^{(\underline{k}_{n,\varepsilon}),i}_{\text{isolation}} \cap \mathcal{A}^{(\underline{k}_{n,\varepsilon}),i}_{\text{insmall}} \cap \mathcal{A}^{(\underline{k}_{n,\varepsilon}),i}_{\text{regular}}(\eta),$$

where now in the definition of these events we replace $\mathcal{C}_\square(0)$ with the component containing the point of Ξ closest to the origin $0 \in \mathbb{R}^d$ for $\square \in \{\text{in}, \tilde{n}\}$. We also assume that $\tilde{n} = n^\vartheta$ is sufficiently large compared to $\underline{k}_{n,\varepsilon}$ in (5.6.2) so that the spatial projection of the box $\mathcal{R}_{\text{out}}$ still fits within $\Lambda_{\tilde{n}}$. This can be ensured even if $k_{n,\varepsilon} = \Theta(n^\varepsilon)$ is maximal in (5.6.2) by choosing $\varepsilon < \vartheta$. If $\mathcal{A}^{(i)}_{\text{good}}$ holds for some $i \leqslant m_n$, then the induced graph $\mathcal{G}^{(i)}_{\tilde{n}}$ in subbox $\Lambda^{(i)}_{\tilde{n}}$ contains a component $\mathcal{C}^{(i)}_{\tilde{n}}$ in $\Xi^{(\text{in},i)}_{\leqslant \mathcal{M}_\star}$ (which we call a 'candidate' second-largest component of $\mathcal{G}_n$) with size at least $\underline{k}_{n,\varepsilon}$ that is not the largest component in its own box, and all vertices in $\Lambda^{(i)}_{\tilde{n}}$ are below $\mathcal{M}_\star(x_i)$, i.e., $\mathcal{M}_\star$ shifted to x_i. If additionally for all i there is no edge between $\Xi^{(\text{in},i)}_{\leqslant \mathcal{M}_\star}$ to a vertex outside $\Lambda^{(i)}_{\tilde{n}}$, then the component $\mathcal{C}^{(i)}_{\tilde{n}}$ is not the giant and is larger than $k_{n,\varepsilon}$. Taking complements we obtain that

$$\big\{|\mathcal{C}^{(2)}_n| \leqslant \underline{k}_{n,\varepsilon}\big\} \subseteq \{\exists i \leqslant m_n : \Xi^{(\text{in},i)}_{\leqslant \mathcal{M}_\star} \leftrightarrow \Xi_n \setminus \Xi^{(i)}_{\tilde{n}}\} \cup \Bigg(\bigcap_{i \leqslant m_n} \big(\neg \mathcal{A}^{(i)}_{\text{good}}\big)\Bigg).$$

By translation invariance, a union bound, and the independence of the graphs $(\mathcal{G}^{(i)}_{\tilde{n}})_{i \leqslant m_n}$ in disjoint boxes,

$$\begin{aligned}\mathbb{P}\big(|\mathcal{C}^{(2)}_n| \leqslant \underline{k}_{n,\varepsilon}\big) &\leqslant m_n \mathbb{P}\big(\Xi^{\text{in}}_{\leqslant \mathcal{M}_\star} \leftrightarrow \Xi_n \setminus \Xi_{\tilde{n}}\}\big) + \big(1 - \mathbb{P}\big(\mathcal{A}^{(1)}_{\text{good}}\big)\big)^{m_n} \\ &=: T_1 + T_2. \end{aligned} \tag{5.6.26}$$

By the definitions in (5.6.11) and (5.6.6), each $u \in \Xi^{\text{in}}_{\leqslant \mathcal{M}_\star}$ has $\|x_u\| \leqslant r_k$ with $k = k_{n,\varepsilon}$, and mark $w_u \leqslant f_\star(r_{\underline{k}_{n,\varepsilon}})$, ($f_\star = f_{\gamma_{\neg\text{nn}}}$ is from below (5.6.13)), hence $\Xi^{\text{in}}_{\leqslant \mathcal{M}_\star} \subseteq \Xi_{(2r_{\underline{k}_{n,\varepsilon}})^d}[1, f_\star(r_{\underline{k}_{n,\varepsilon}})) \subseteq \Xi_{\tilde{n}}$, whenever $(2r_{\underline{k}_{n,\varepsilon}})^d = k_{n,\varepsilon} M_{\text{in}} 2^d < n^\vartheta$, which holds whenever $\varepsilon < \vartheta$ by (5.6.2). Hence we can bound T_1 as

$$T_1 \leqslant m_n \mathbb{P}\big(\Xi_{(2r_{\underline{k}_{n,\varepsilon}})^d}[1, f_\star(r_{\underline{k}_{n,\varepsilon}})) \leftrightarrow \Xi_n \setminus \Xi_{\tilde{n}}\}\big).$$

We can directly apply Claim 5.5.6 on the right-hand side, i.e., setting there $N := \tilde{n} = n^\vartheta$ and $n := (2r_{\underline{k}_{n,\varepsilon}})^d = k_{n,\varepsilon} M_{\text{in}} 2^d$ (by (5.6.6)) and $\overline{w} := f_\star(r_{\underline{k}_{n,\varepsilon}})$. The profile $f_\star = f_{\gamma_{\neg\text{nn}}}$ is defined below (5.6.13), using (5.6.9) with exponent $\gamma_{\neg\text{nn}}$ and $r_k = (M_{\text{in}} k)^{1/d}$ in (5.6.6), and finally $k_{n,\epsilon}$ from (5.6.2) we obtain

$$\overline{w} := f_\star(r_{\underline{k}_{n,\varepsilon}}) = C_\beta^{-\gamma_{\neg\text{nn}} d} r_{\underline{k}_{n,\varepsilon}}^{\gamma_{\neg\text{nn}} d} = C_\beta^{-\gamma_{\neg\text{nn}} d} M_{\text{in}}^{\gamma_{\neg\text{nn}}} \underline{k}_{n,\varepsilon}^{\gamma_{\neg\text{nn}}}$$

Condition (5.5.27) holds since $N = n^\vartheta \gg k_{n,\varepsilon} \gg \overline{w}$. Then Claim 5.5.6 yields for some $C > 0$

$$\begin{aligned} T_1 &\leqslant n^{1-\vartheta} C_{5.5.6} f_\star(r_{\underline{k}_{n,\varepsilon}})^{c_{5.5.6}} (2r_{\underline{k}_{n,\varepsilon}})^d n^{-\vartheta \min\{\alpha-1, \tau-2\}} \big(1 + \mathbb{1}_{\{\alpha=\tau-1\}} \log(n^\vartheta)\big) \\ &\leqslant C \log(n) \cdot \underline{k}_{n,\varepsilon}^{1+\gamma_{\neg\text{nn}} c_{5.5.6}} \cdot n^{1-\vartheta \min\{\alpha, \tau-1\}}. \end{aligned}$$

Since $\underline{k}_{n,\varepsilon}$ in (5.6.2) is at most n^{ε}, as long as $1-\vartheta\min(\alpha,\tau-1) < 0$, we can choose $\varepsilon > 0$ in (5.6.2) sufficiently small such that for any $\delta \in (0, \vartheta\min\{\alpha,\tau-1\}-1)$, for all n sufficiently large,

$$T_1 \leqslant n^{-\delta}. \tag{5.6.27}$$

We turn to bound T_2 in (5.6.26) using $(1-x)^{m_n} \leqslant \exp(-m_n x)$, where we apply (5.6.25) on $x = \mathbb{P}(\mathcal{A}^{(1)}_{\text{good}})$ to obtain a lower bound on the exponent

$$\begin{aligned} m_n\mathbb{P}\big(\mathcal{A}^{(1)}_{\text{good}}\big) &\geqslant (\rho^2/2)n^{1-\vartheta}\exp\big(-A'\underline{k}_{n,\varepsilon}^{\max(\mathcal{Z})}\log^{m_{\mathcal{Z}}-1}(\underline{k}_{n,\varepsilon})\big) \\ &= (\rho^2/2)\exp\big((1-\vartheta)\log(n) - A'\underline{k}_{n,\varepsilon}^{\max(\mathcal{Z})}\log^{m_{\mathcal{Z}}-1}(\underline{k}_{n,\varepsilon})\big). \end{aligned}$$

In order to show $T_2 \leqslant n^{-\delta}$ in (5.6.26), it is much stronger to show that with $m(\mathcal{Z})-1 = m'$

$$\begin{aligned} &\forall\varepsilon' > 0, \text{ there exists } \varepsilon_1 > 0 \text{ such that for all } \varepsilon < \varepsilon_1: \\ &\qquad A'\underline{k}_{n,\varepsilon}^{\max(\mathcal{Z})}\log^{m'}(\underline{k}_{n,\varepsilon}) < \varepsilon'\log(n). \end{aligned} \tag{5.6.28}$$

We recall the definition of $\underline{k}_{n,\varepsilon}$ in (5.6.2). While $k_{n,\varepsilon}$ may be as large as n^{ε}, only when $\max(\mathcal{Z}) = 0$, so effectively the expression in (5.6.28) is still small. We formally check:

Case 1. $\max(\mathcal{Z}) > 0$. We substitute $\underline{k}_{n,\varepsilon}$ in the first row of (5.6.2) to (5.6.28)

$$A'\underline{k}_{n,\varepsilon}^{\max(\mathcal{Z})}\log^{m'}(\underline{k}_{n,\varepsilon}) = A'\frac{\varepsilon\log(n)}{\log\log(n))^{m'}}\log^{m'}\Big(\big(\frac{\varepsilon\log(n)}{(\log\log(n))^{m'}}\big)^{1/\max(\mathcal{Z})}\Big).$$

The last factor is at most $(\max\mathcal{Z})^{-m'}(\log\log(n))^{m'}$, and (5.6.28) follows whenever $\varepsilon < \varepsilon'(\max\mathcal{Z})^{m'}/A'$.

Case 2. $\max(\mathcal{Z}) = 0$. We substitute $\underline{k}_{n,\varepsilon}$ in the second row of (5.6.2) to (5.6.28)

$$A'\underline{k}_{n,\varepsilon}^{\max(\mathcal{Z})}\log^{m'}(\underline{k}_{n,\varepsilon}) = A'\big(\log\big(\exp\big[(\varepsilon\log(n))^{1/m'}\big]\big)\big)^{m'} = A'\varepsilon\log(n),$$

and (5.6.28) again follows. Choose now any $\vartheta < 1/\min\{\alpha,\tau-1\}$, and then combine (5.6.27) with $T_2 \leqslant n^{-\delta}$ to bound (5.6.26). This finishes the proof of (5.6.5) and hence Proposition 5.6.1 subject to Lemma 5.6.2. □

5.6.2 *Isolation*

We aim to prove Lemma 5.6.2. We suppress the superscripts (k) of events, and leave it to the reader to verify that the results extend to the Palm-version $\widetilde{\mathbb{P}}^{o}_{\text{io}}$ of $\widetilde{\mathbb{P}}_{\text{io}}$, which is defined in (5.6.18). We work towards bounding the event in (5.6.19) via a few lemmas/claims, also related to the back-of-the-envelope reasoning of Section 4.2. Recall f_γ, M_γ from (5.6.9), (5.6.10), the PPPs in (5.6.11), and $\zeta_{\mathrm{nn}} = \frac{d-1}{d}$ from (4.2.5).

Lemma 5.6.3 (Vertices above $\mathcal{M}_\gamma$)**.** *Consider a supercritical i-KSRG under the conditions of Lemma 5.6.2. For each $\gamma \in (0, 1/(\sigma+1)]$, there exists a constant $C_{5.6.3} > 0$ such that*

$$\widetilde{\mathbb{E}}_{\mathrm{io}}\big[|\Xi^{\mathrm{in}}_{>\mathcal{M}_\gamma} \cup \Xi^{\mathrm{out}}_{>\mathcal{M}_\gamma}|\big] \leqslant \begin{cases} C_{5.6.3} r_k^{d(1-\gamma(\tau-1))}, & \textit{if } 1-\gamma(\tau-1) > \zeta_{\mathrm{nn}}, \\ C_{5.6.3} r_k^{d\zeta_{\mathrm{nn}}} \log(r_k), & \textit{if } 1-\gamma(\tau-1) = \zeta_{\mathrm{nn}}, \\ C_{5.6.3} r_k^{d\zeta_{\mathrm{nn}}}, & \textit{if } 1-\gamma(\tau-1) < \zeta_{\mathrm{nn}}. \end{cases} \tag{5.6.29}$$

Recall the multiplicity $\mathfrak{m}$ of the maximum from (5.1.5).

Lemma 5.6.4 (Edges between vertices below $\mathcal{M}_\gamma$)**.** *Consider a supercritical i-KSRG under the conditions of Lemma 5.6.2 with $\alpha < \infty$. Let $\mathcal{Y} = \{0, \alpha-(\tau-1), (\sigma+1)\alpha - 2(\tau-1)\}$. For each $\gamma \in (0, 1/(\sigma+1)]$ there exists a constant $C_{5.6.4} = C_{5.6.4}(\rho) > 0$ such that for any realization of $\Xi_{\mathcal{R}_{\mathrm{in}}} \cup \Xi_{\mathcal{R}_{\mathrm{out}}}$ that satisfies $\mathcal{A}_{\mathrm{regular}}(\eta)$ in (5.6.17) for some $\eta > 0$,*

$$\begin{aligned} &\widetilde{\mathbb{E}}_{\mathrm{io}}\big[|\mathcal{E}\big(\Xi^{\mathrm{in}}_{\leqslant\mathcal{M}_\gamma}, \Xi^{\mathrm{out}}_{\leqslant\mathcal{M}_\gamma}\big)|\big] \\ &\leqslant \begin{cases} C_{5.6.4} r_k^{d(2-\alpha+\gamma\max(\mathcal{Y}))} \log^{\mathfrak{m}(\mathcal{Y})-1}(r_k), & \textit{if } 2-\alpha+\gamma\max(\mathcal{Y}) > \zeta_{\mathrm{nn}}, \\ C_{5.6.4} r_k^{d\zeta_{\mathrm{nn}}} \log^{\mathfrak{m}(\mathcal{Y})}(r_k), & \textit{if } 2-\alpha+\gamma\max(\mathcal{Y}) = \zeta_{\mathrm{nn}}, \\ C_{5.6.4} r_k^{d\zeta_{\mathrm{nn}}}, & \textit{if } 2-\alpha+\gamma\max(\mathcal{Y}) < \zeta_{\mathrm{nn}}. \end{cases} \end{aligned} \tag{5.6.30}$$

Proof sketch of Lemmas 5.6.3 and 5.6.4. We defer the lengthy integrals to the appendix on page 210 but we give some intuition. The expectation $\widetilde{\mathbb{E}}_{\mathrm{io}}$ in (5.6.18) is conditional on $\Xi \cap (\mathcal{R}_{\mathrm{in}} \cup \mathcal{R}_{\mathrm{out}})$ where $\mathcal{R}_{\mathrm{in}}, \mathcal{R}_{\mathrm{out}}$ in (5.6.7) are hyperrectangles below $\mathcal{M}_\gamma$ for all η, since $w_\eta(\ell)$ is poly-logarithmic in ℓ in (5.6.1). Hence $\Xi_{>\mathcal{M}_\gamma}$ is independent of the conditioning in $\widetilde{\mathbb{E}}_{\mathrm{io}}$, so

$$\widetilde{\mathbb{E}}_{\mathrm{io}}\big[|\Xi^{\mathrm{in}}_{>\mathcal{M}_\gamma} \cup \Xi^{\mathrm{out}}_{>\mathcal{M}_\gamma}|\big] = \mathbb{E}\big[|\Xi^{\mathrm{in}}_{>\mathcal{M}_\gamma} \cup \Xi^{\mathrm{out}}_{>\mathcal{M}_\gamma}|\big] = \mu_\tau\big(\mathbb{R}^{d+1}_{>\mathcal{M}_\gamma}\big), \tag{5.6.31}$$

with the intensity μ_τ in (5.1.2), and where we denote $\mathbb{R}^{d+1}_{>\mathcal{M}_\gamma}$ the points of $\mathbb{R}^{d+1}$ above the d-dimensional surface $\mathcal{M}_\gamma$, see (5.6.10). Define the hyperrectangle $\mathcal{R}^\uparrow := [-2r_k, 2r_k]^d \times [r_k^\gamma, \infty)$ and $A_\beta := \{x \in \mathbb{R}^d, \|x\| \in [r_k - C_\beta, r_k + C_\beta]\} \times [1, \infty)$, an annulus in $\mathbb{R}^d$ times all mark-coordinates. Then by definition of f_γ in (5.6.9), $(\mathcal{R}^\uparrow \cup A_\beta) \subseteq \mathbb{R}^{d+1}_{>\mathcal{M}_\gamma}$, and $\mu_\tau(\mathcal{R}^\uparrow \cup A_\beta) =$

$\Theta(r_k^{d(1-\gamma(\tau-1))} + r_k^{d-1})$. Careful integration shows that the right-hand side of (5.6.31) is of the same order, unless $1-\gamma(\tau-1) = (d-1)/d = \zeta_{\mathrm{nn}}$, when we get an extra $\log(r_k)$ factor.

For Lemma 5.6.4, we use independence of Ξ in disjoint sets: inside and outside of $\mathcal{R}_{\mathrm{in}}, \mathcal{R}_{\mathrm{out}}$. When counting edges with at least one point in $\mathcal{R}_{\mathrm{in}} \cup \mathcal{R}_{\mathrm{out}}$, we use that Ξ in these sets is regular, i.e., $\mathcal{A}^{(k)}_{\mathrm{regular}}$ holds, and outside these sets we integrate using the intensity μ_τ in (5.6.29). We explain now how the exponents of r_k in (5.6.30) arise: the expected number of edges between vertices of constant mark within constant distance of $\partial\mathcal{B}_{kM_{\mathrm{in}}}$ is $\Theta(r_k^{d-1}) = \Theta(r_k^{d\zeta_{\mathrm{nn}}})$, which yields lower bounds for all cases in (5.6.30). Let $0 \leqslant \gamma_1 \leqslant \gamma_2 \leqslant \gamma$ such that $\sigma\gamma_1 + \gamma_2 \leqslant \sigma + 1$. Using μ_τ in (5.1.2), the expected number of pairs $(u,v) \in \Xi^{\mathrm{in}}_{\leqslant \mathcal{M}_\gamma} \times \Xi^{\mathrm{out}}_{\leqslant \mathcal{M}_\gamma}$ within distance $\Theta(r_k)$ from $\partial\mathcal{B}_{kM_{\mathrm{in}}}$, and marks $w_u = \Theta(k^{\gamma_1}), w_v = \Theta(k^{\gamma_2})$ is

$$\begin{aligned}\mathbb{E}[\mathrm{Pairs}(\gamma_1,\gamma_2)] &:= \Theta\big(r_k^{d(1-\gamma_1(\tau-1))} \cdot r_k^{d(1-\gamma_2(\tau-1))}\big)\\ &= \Theta\big(r_k^{d(2-(\gamma_1+\gamma_2)(\tau-1))}\big).\end{aligned}$$

The typical Euclidean distance between such vertices is $\Theta(r_k)$. Therefore, by the connection probability p in (5.1.3), a pair of such vertices are connected with probability roughly $\Theta(r_k^{d\alpha(\sigma\gamma_1+\gamma_2-1)})$, and hence there are

$$\begin{aligned}\mathbb{E}[\mathrm{Edges}(\gamma_1,\gamma_2)] &:= \mathbb{E}[\mathrm{Pairs}(\gamma_1,\gamma_2)] \cdot \Theta(r_k^{d\alpha(\sigma\gamma_1+\gamma_2-1)})\\ &= \Theta\big(r_k^{d(2-\alpha+\gamma_1(\sigma\alpha-(\tau-1))+\gamma_2(\alpha-(\tau-1)))}\big).\end{aligned} \tag{5.6.32}$$

such edges in expectation. The proof below on page 211 reveals that the expectation of $|\mathcal{E}(\Xi^{\mathrm{in}}_{\leqslant \mathcal{M}_\gamma}, \Xi^{\mathrm{out}}_{\leqslant \mathcal{M}_\gamma})|$ can be bounded by maximizing (5.6.32) with respect to γ_1, γ_2, i.e., edges between vertices of mark $\Theta(k^{\gamma_1^\star})$ and $\Theta(k^{\gamma_2^\star})$, with the optimal pair $(\gamma_1^\star, \gamma_2^\star)$ dominate the whole expectation. Depending on (σ, α, τ), the optimal $(\gamma_1^\star, \gamma_2^\star)$ is either

(ll) $(0,0)$, pairs of low-mark vertices, which, when substituted in (5.6.32) yields $0 \in \mathcal{Y}$ in (5.6.30);

(lh) $(0,\gamma)$, pairs of one low-mark vertex and one high-mark vertex, which, when substituted in (5.6.32) yields $\alpha-(\tau-1) \in \mathcal{Y}$ in (5.6.30);

(hh) (γ,γ), pairs of high-mark vertices, which, when substituted in (5.6.32) yields $(\sigma+1)\alpha - 2(\tau-1) \in \mathcal{Y}$ in (5.6.30);

or the convex combinations of these if the maximum in $\mathcal{Y}$ is non-unique, leading to the poly-logarithmic correction terms in (5.6.30). These three options of $(\gamma_1^\star, \gamma_2^\star)$, combined with the $\Theta(r_k^{d-1})$ many low-mark edges close to the boundary explain what we call the *dominant types of connectivity* in Section 4.2. We remark that without the assumption that $\gamma \leqslant 1/(\sigma+1)$, the optimal solution in (hh) would change to $(\gamma \wedge 1/(\sigma+1), \gamma \wedge 1/(\sigma+1))$. □

Lemmas 5.6.3 and 5.6.4 hold for any $\gamma \leqslant 1/(\sigma+1)$. Now we show that setting $\gamma = \gamma_{\neg\mathrm{nn}}$ from (5.6.13) minimizes the sum of (5.6.29) (increasing in γ) and (5.6.30) (decreasing in γ).

Claim 5.6.5 (The optimally-suppressed mark-profile)**.** *Consider a supercritical i-KSRG under the conditions of Lemma 5.6.2. Let* $\mathcal{Z} = \{\zeta_{\mathrm{ll}}, \zeta_{\mathrm{lh}}, \zeta_{\mathrm{hh}}, \zeta_{\mathrm{nn}}\}$. *There exists a constant* $C_{5.6.5} > 0$ *such that when* $\alpha < \infty$

$$\begin{aligned} \widetilde{\mathbb{E}}_{\mathrm{io}}\big[|\Xi^{\mathrm{in}}_{>\mathcal{M}_\star} \cup \Xi^{\mathrm{out}}_{>\mathcal{M}_\star}|\big] &\leqslant C_{5.6.5}\widetilde{\mathbb{E}}_{\mathrm{io}}\big[|\mathcal{E}\big(\Xi^{\mathrm{in}}_{\leqslant\mathcal{M}_\star}, \Xi^{\mathrm{out}}_{\leqslant\mathcal{M}_\star}\big)|\big] \\ &\leqslant C_{5.6.5}C_{5.6.4}r_k^{d\max(\mathcal{Z})}\log^{m_{\mathcal{Z}}-1}(r_k), \end{aligned} \tag{5.6.33}$$

and when $\alpha = \infty$

$$\widetilde{\mathbb{E}}_{\mathrm{io}}\big[|\Xi^{\mathrm{in}}_{>\mathcal{M}_\star} \cup \Xi^{\mathrm{out}}_{>\mathcal{M}_\star}|\big] \leqslant C_{5.6.5}r_k^{d\max(\mathcal{Z})}\log^{m_{\mathcal{Z}}-1}(r_k). \tag{5.6.34}$$

We postpone the proof, based on elementary rearrangements of the formulas of $\zeta_{\mathrm{ll}}, \zeta_{\mathrm{lh}}, \zeta_{\mathrm{hh}}$, and ζ_{hh}, to the appendix on page 217. We proceed with a claim towards bounding the event $\mathcal{A}^{(k)}_{\mathrm{noedge}}$ in (5.6.12), its proof explaining the restriction $\gamma \leqslant 1/(\sigma+1)$ in $\gamma_{\neg\mathrm{nn}}$ in (5.6.13). We recall $p(u,v)$ defined in (5.1.3).

Claim 5.6.6 (Cross-edge probability bounds)**.** *Consider a supercritical i-KSRG under conditions of Lemma 5.6.2, and assume* $\gamma \leqslant 1/(\sigma+1)$. *For any* $(u,v) \in \Xi^{\mathrm{in}}_{\leqslant\mathcal{M}_\gamma} \times \Xi^{\mathrm{out}}_{\leqslant\mathcal{M}_\gamma}$

$$p(u,v) \leqslant \begin{cases} 2^{-\alpha}, & \text{if } \alpha < \infty, \\ 0, & \text{if } \alpha = \infty. \end{cases} \tag{5.6.35}$$

Proof. We show that whenever u, v are below $\mathcal{M}_\gamma$ and on different sides of $\partial\mathcal{B}_{kM_{\mathrm{in}}}$, then $\beta\kappa_{1,\sigma}(w_u, w_v)/\ \|x_u - x_v\|^d \leqslant 1/2$. By definition of p in (5.1.3), this directly implies (5.6.35). To see this bound, for $\sigma \geqslant 0$ the connection probability p is increasing in the marks. Therefore, without loss of generality we will assume that $u \in \mathcal{B}_{kM_{\mathrm{in}}}$ and $v \notin \mathcal{B}_{kM_{\mathrm{in}}}$ fall exactly on $\mathcal{M}_\gamma$, and that k is large enough that $r_k \geqslant C_\beta$. Since f_γ in (5.6.9) switches formula at r_k outside $\partial\mathcal{B}_{kM_{\mathrm{in}}}$, we distinguish two cases.

Case 1. Assume $|\|x_v\| - r_k| \leqslant r_k$. Since f_γ in (5.6.9) and the explanation below (5.6.12) both u, v are at least C_β distance from $\partial\mathcal{B}_{kM_{\text{in}}}$. Thus, for any (u, v), there exist $t \geqslant 2C_\beta$, $\nu \in (0,1)$ such that $\|x_u - \partial\mathcal{B}_{kM_{\text{in}}}\| = r_k - \|x_u\| = (1-\nu)t$, and $\|x_v - \partial\mathcal{B}_{kM_{\text{in}}}\| = \|x_v\| - r_k = \nu t$. By the triangle inequality $\|x_u - x_v\| \geqslant t$. Hence, $w_v = C_\beta^{-\gamma d}(\nu t)^{\gamma d}, w_u = C_\beta^{-\gamma d}((1-\nu)t)^{\gamma d}$ by assuming u, v being on $\mathcal{M}_\gamma$ and f_γ in (5.6.9). Using $\kappa_{1,\sigma}$ in (1.3.5), and $\nu \in (0,1)$

$$\beta\frac{\kappa_{1,\sigma}(w_u, w_v)}{\|x_u - x_v\|^d} \leqslant \beta C_\beta^{-\gamma d(\sigma+1)}\frac{\max\{(1-\nu)t, \nu t\}^{\gamma d}\min\{(1-\nu)t, \nu t\}^{\sigma\gamma d}}{t^d}$$
$$\leqslant \beta C_\beta^{-\gamma d(\sigma+1)} t^{d(\gamma(\sigma+1)-1)}.$$

The right-hand side is non-increasing in t *whenever* $\gamma \leqslant \frac{1}{\sigma+1}$. Since $t \geqslant 2C_\beta \geqslant C_\beta$, and $C_\beta = (2\beta)^{1/d}$

$$\beta\frac{\kappa_{1,\sigma}(w_u, w_v)}{\|x_u - x_v\|^d} \leqslant \beta C_\beta^{-d\gamma(\sigma+1)} C_\beta^{d(\gamma(\sigma+1)-1)} = \beta C_\beta^{-d} = 1/2.$$

Case 2. Assume $|\|x_v\| - r_k| > r_k$. In this case $\|x_v - \partial\mathcal{B}_{kM_{\text{in}}}\| \geqslant \|x_u - \partial\mathcal{B}_{kM_{\text{in}}}\|$, and $\|x_v - x_u\| \geqslant \|x_v - \partial\mathcal{B}_{kM_{\text{in}}}\| = |\|x_v\| - r_k|$. Since $f_\gamma(z)$ is increasing, $w_v = f_\gamma(|\|x_v\| - r_k|) \geqslant f_\gamma(r_k) \geqslant f_\gamma(|\|x_u\| - r_k|) = w_u$. We obtain by definition of $\kappa_{1,\sigma}$ in (1.3.5) and (5.6.9)

$$\beta\frac{\kappa_{1,\sigma}(w_u, w_v)}{\|x_u - x_v\|^d} \leqslant \beta\frac{f_\gamma(|\|x_v\| - r_k|) f_\gamma(|\|x_u\| - r_k|)^\sigma}{|\|x_v\| - r_k|^d}$$
$$\leqslant \beta C_\beta^{-\gamma d(\sigma+1)} |\|x_v\| - r_k|^d r_k^{-d(1-\gamma)} r_k^{\sigma\gamma d} |\|x_v\| - r_k|^{-d}$$
$$= \beta C_\beta^{-\gamma d(\sigma+1)} r_k^{-d(1-\gamma(\sigma+1))} \leqslant 1/2.$$

where to obtain the second inequality we used that $f_\gamma(|\|x_u\| - r_k|)^\sigma \leqslant r_k^{\sigma\gamma d}$, and to obtain the second row the power of $|\|x_v\| - r_k|$ canceled. The bound $1/2$ follows because $r_k \geqslant C_\beta$, $\gamma \leqslant 1/(\sigma+1)$ and $C_\beta = (2\beta)^{1/d}$. □

We are ready to prove Lemma 5.6.2.

Proof of Lemma 5.6.2. We will obtain a lower bound on $\widetilde{\mathbb{P}}_{\text{io}}\big(\{\Xi \leqslant \mathcal{M}_\gamma\} \cap \mathcal{A}^{(k)}_{\text{noedge}}(\gamma)\big)$ for *arbitrary* $\gamma \in (0, 1/(\sigma+1)]$. Then we maximise this bound by setting $\gamma = \gamma_{\neg\text{nn}}$, yielding the event $\mathcal{A}^{(k)}_{\text{isolation}}$ in (5.6.14) and the bound in Lemma 5.6.2. The events $\{\Xi \leqslant \mathcal{M}_\gamma\}$ and $\mathcal{A}^{(k)}_{\text{noedge}}$ (defined in (5.6.12)) are independent of each other under $\widetilde{\mathbb{P}}_{\text{io}}$ in (5.6.18), since having no points above $\mathcal{M}_\gamma$ is independent of the conditioning in $\widetilde{\mathbb{P}}_{\text{io}}$ (since each point in

$\mathcal{R}_{\mathrm{in}} \cup \mathcal{R}_{\mathrm{out}}$ is below $\mathcal{M}_\gamma$ if k is sufficiently large), and $\mathcal{A}^{(k)}_{\mathrm{noedge}}$ only depends on points of Ξ below $\mathcal{M}_\gamma$. Hence

$$\widetilde{\mathbb{P}}_{\mathrm{io}}\big(\{\Xi \leqslant \mathcal{M}_\gamma\} \cap \mathcal{A}^{(k)}_{\mathrm{noedge}}(\gamma)\big) = \widetilde{\mathbb{P}}_{\mathrm{io}}\big(\Xi \leqslant \mathcal{M}_\gamma\big)\widetilde{\mathbb{P}}_{\mathrm{io}}\big(\mathcal{A}^{(k)}_{\mathrm{noedge}}(\gamma)\big). \tag{5.6.36}$$

We analyse the two probabilities separately. For the first term, by the above independence,

$$\widetilde{\mathbb{P}}_{\mathrm{io}}\big(\Xi \leqslant \mathcal{M}_\gamma\big) = \widetilde{\mathbb{P}}_{\mathrm{io}}\big(|\Xi^{\mathrm{in}}_{>\mathcal{M}_\gamma} \cup \Xi^{\mathrm{out}}_{>\mathcal{M}_\gamma}| = 0\big) = \exp\big(-\widetilde{\mathbb{E}}_{\mathrm{io}}\big[|\Xi^{\mathrm{in}}_{>\mathcal{M}_\gamma} \cup \Xi^{\mathrm{out}}_{>\mathcal{M}_\gamma}|\big]\big), \tag{5.6.37}$$

for which we will use Lemma 5.6.3 shortly. We now turn to the second factor in (5.6.36). By definition of $\mathcal{A}^{(k)}_{\mathrm{noedge}}$ in (5.6.12), and using the conditional independence of edges,

$$\widetilde{\mathbb{P}}_{\mathrm{io}}\big(\mathcal{A}^{(k)}_{\mathrm{noedge}}(\gamma)\big) = \widetilde{\mathbb{E}}_{\mathrm{io}}\Bigg[\prod_{u \in \Xi^{\mathrm{in}}_{\leqslant \mathcal{M}_\gamma}, v \in \Xi^{\mathrm{out}}_{\leqslant \mathcal{M}_\gamma}} \big(1 - p(u,v)\big)\Bigg]. \tag{5.6.38}$$

We distinguish now the two cases $\alpha = \infty$ and $\alpha < \infty$.

Case $\alpha = \infty$. By Claim 5.6.6, since $\gamma \leqslant 1/(\sigma+1)$ by assumption, $p(u,v) = 0$ for each factor, hence $\widetilde{\mathbb{P}}_{\mathrm{io}}\big(\mathcal{A}^{(k)}_{\mathrm{noedge}}(\gamma)\big) = 1$. Since the right-hand side of (5.6.37) is increasing in γ, we set $\gamma = 1/(\sigma+1) = \gamma_{\neg\mathrm{nn}}$ when $\alpha = \infty$ by (5.6.13). Combining then (5.6.36) with (5.6.37) and Claim 5.6.5, it follows that

$$\widetilde{\mathbb{P}}_{\mathrm{io}}\big(\{\Xi \leqslant \mathcal{M}_\star\} \cap \mathcal{A}^{(k)}_{\mathrm{noedge}}(\gamma_{\neg\mathrm{nn}})\big) \geqslant \exp\big(-C_{5.6.3} r_k^{d\max(\mathcal{Z})} \log^{m_{\mathcal{Z}}-1}(r_k)\big),$$

finishing the proof of the lemma for $\alpha = \infty$.

Case $\alpha < \infty$. By Claim 5.6.6, for all $(u,v) \in \Xi^{\mathrm{in}}_{\leqslant \mathcal{M}_\gamma} \times \Xi^{\mathrm{out}}_{\leqslant \mathcal{M}_\gamma}$ it holds that $1 - p(u,v) \geqslant 1 - 2^{-\alpha}$ since $\gamma \leqslant 1/(\sigma+1)$ by assumption. Hence, there exists a constant $c > 0$ such that for all such (u,v) we have $1 - p(u,v) \geqslant \exp(-c \cdot p(u,v))$. Using this in (5.6.38) and that $s \mapsto \exp(-s)$ is a convex function, Jensen's inequality gives a lower bound in terms of the expected number of edges between vertices below $\mathcal{M}_\gamma$, i.e.,

$$\begin{aligned}
\widetilde{\mathbb{P}}_{\mathrm{io}}\big(\mathcal{A}^{(k)}_{\mathrm{noedge}}(\gamma)\big) &\geqslant \widetilde{\mathbb{E}}_{\mathrm{io}}\Bigg[\exp\Bigg(-c \sum_{\substack{u \in \Xi^{\mathrm{in}}_{\leqslant \mathcal{M}_\gamma},\\ v \in \Xi^{\mathrm{out}}_{\leqslant \mathcal{M}_\gamma}}} p(u,v)\Bigg)\Bigg] \\
&\geqslant \exp\Bigg(-c\widetilde{\mathbb{E}}_{\mathrm{io}}\Bigg[\sum_{\substack{u \in \Xi^{\mathrm{in}}_{\leqslant \mathcal{M}_\gamma},\\ v \in \Xi^{\mathrm{out}}_{\leqslant \mathcal{M}_\gamma}}} p(u,v)\Bigg]\Bigg) \\
&= \exp\big(-c\widetilde{\mathbb{E}}_{\mathrm{io}}\big[|\mathcal{E}\big(\Xi^{\mathrm{in}}_{\leqslant \mathcal{M}_\gamma}, \Xi^{\mathrm{out}}_{\leqslant \mathcal{M}_\gamma}\big)|\big]\big).
\end{aligned} \tag{5.6.39}$$

Together with (5.6.36) and (5.6.37), we obtain that

$$\widetilde{\mathbb{P}}_{\mathrm{io}}\big(\{\Xi \leqslant \mathcal{M}_\gamma\} \cap \mathcal{A}^{(k)}_{\mathrm{noedge}}(\gamma)\big)$$
$$\geqslant \exp\Big(-\big(\widetilde{\mathbb{E}}_{\mathrm{io}}\big[|\Xi^{\mathrm{in}}_{>\mathcal{M}_\gamma} \cup \Xi^{\mathrm{out}}_{>\mathcal{M}_\gamma}|\big] + c\widetilde{\mathbb{E}}_{\mathrm{io}}\big[|\mathcal{E}\big(\Xi^{\mathrm{in}}_{\leqslant\mathcal{M}_\gamma}, \Xi^{\mathrm{out}}_{\leqslant\mathcal{M}_\gamma}\big)|\big]\big)\Big). \quad (5.6.40)$$

By Lemmas 5.6.3 and 5.6.4, the first expectation is decreasing, while the second one is increasing in γ on the right-hand side. By (5.6.33) in Lemma 5.6.5, when we set $\gamma = \gamma_{\neg\mathrm{nn}}$ in (5.6.13), we obtain that (5.6.40) turns into (5.6.19), finishing the proof of Lemma 5.6.2 when $\mathcal{V} = \Xi$. □

Proof of Lemma 5.6.2 for i-KSRGs on $\mathbb{Z}^d$. We explain how to adapt the proof for i-KSRGs on $\mathbb{Z}^d$ using the extra assumption $\min\{p, p\beta^\alpha\} < 1$ in Lemma 5.6.2. Since the vertex set $\mathcal{V} = \mathbb{Z}^d$, the event $\Xi \leqslant \mathcal{M}_\gamma$ as defined in (5.6.10) would never hold, since $\mathbb{Z}^d$ does have points within distance C_β from $\partial\mathcal{B}_{kM_{\mathrm{in}}}$ (see f_γ in (5.6.9) and the reasoning below (5.6.12)). Hence, we must adjust the definition of f_γ within distance C_β of $\partial\mathcal{B}_{kM_{\mathrm{in}}}$ to be any constant $c > 1$. Then the upper bound $2^{-\alpha}$ on $p(u, v)$ in Claim 5.6.6 can be replaced by $\min\{p, p\beta^\alpha\} < 1$ (the edge-retention probability for vertices at distance 1 in (5.1.3)), that only affects constant prefactors in (5.6.39) and (5.6.40).

The proofs of the preliminary Lemmas 5.6.3 and 5.6.4 remain valid by replacing concentration bounds for Poisson random variables by concentration bounds for Binomial random variables, and replacing integrals over $\mathbb{R}^d$ by summations over $\mathbb{Z}^d$. The proof of Claim 5.6.5 remains verbatim valid. □

Proofs of Theorems 5.1.2, 5.1.4. Proposition 5.6.1 yields the lower bounds for both theorems, since its condition (5.6.3) on having a large enough component on restricted marks occurs with positive probability by Proposition 5.4.13, under the assumption $\tau < 2 + \sigma$. The assumption $\tau > 2$ is necessary to have a locally finite graph. Propositions 5.4.1 and 5.5.1 imply the upper bounds (i.e., part (ii)) in Theorems 5.1.4 and 5.1.2, respectively. The conditions in Proposition 5.5.1 are satisfied by Propositions 5.4.1, 5.4.12 and Claim 5.5.2. The condition $\tau < 2 + \sigma$ is needed for Claim 5.4.5, which yields that a constant proportion of high-mark vertices, i.e., vertices having mark $\Omega(k^{\gamma_{\mathrm{hh}}})$, is connected by an edge to another high-mark vertex. This allows to construct a backbone of high-mark vertices (Lemma 5.4.2), and to merge components of size at least k with the backbone via a high-mark vertex in (5.4.42). The additional condition $\sigma \leqslant \tau - 1$ in Theorems 5.1.2

and 5.1.4 is needed in the cover-expansion step (Lemma 5.4.9) in the proof of Proposition 5.4.1. □

It remains to prove Theorem 5.1.5, whose proof is based on Lemma 5.6.2.

Proof of Theorem 5.1.5. For $\rho \geqslant 1$ the statement is trivial. There exists a constant $C > 0$ such that for any $\rho \in (0, 1)$ and $n \geqslant 1$ a box of volume n is contained in the union of $\lceil C/\rho \rceil$ (partially overlapping) balls of volume $n\rho/2$. Fix $\rho \in (0, 1)$, and write $\Xi^{(i)}$ for the vertices in the i-th ball of such a cover of balls of volume $n\rho/2$. Recall that $|\mathcal{E}(A, B)|$ denotes the number of edges between the sets A, B. Then

$$\{|\mathcal{C}_n^{(1)}| < \rho n\} \supseteq \bigcap_{i \leqslant \lceil C/\rho \rceil} \{|\mathcal{E}(\Xi^{(i)}, \Xi \setminus \Xi^{(i)})| = 0\} \cap \{|\Xi^{(i)}| < \rho n\} \tag{5.6.41}$$

Indeed, on the event on the right-hand side, each connected component of $\mathcal{G}_n$ is fully contained in some ball (or the intersection of some balls) with at most ρn vertices. We apply an FKG inequality to bound the probability of intersection from below.

We give a (natural) definition of increasing events, using the collection Ψ from Definition 5.4.3 that encodes the presence of edges using a set of uniform random variables $\Psi_\Xi = \{\varphi_{u,v} : u, v \in \Xi\}$. We say that a function $f(\Xi, \Psi_\Xi)$ defined on the marked vertex set Ξ and edge-variable set Ψ_Ξ is *increasing* if it is non-decreasing in Ξ with respect to set inclusion (formally, if $\Xi' \supseteq \Xi, \Psi_{\Xi'} \supseteq \Psi_\Xi$, then $f(\Xi', \Psi_{\Xi'}) \geqslant f(\Xi, \Psi_\Xi)$ holds), as well as coordinate-wise non-increasing with respect to the edge variables (formally, if Ψ'_Ξ satisfies $\varphi'_{u,v} \leqslant \varphi_{u,v}$ for all $u, v \in \binom{\Xi}{2}$, then $f(\Xi, \Psi'_\Xi) \geqslant f(\Xi, \Psi_\Xi)$ holds). Intuitively this means that more vertices and edges increase the value of f. Similarly to [79], we obtain that for two increasing functions f_1, f_2,

$$\begin{aligned}
\mathbb{E}[f_1(\Xi, \Psi_\Xi) \cdot f_2(\Xi, \Psi_\Xi)] &= \mathbb{E}\big[\mathbb{E}[f_1(\Xi, \Psi_\Xi) \cdot f_2(\Xi, \Psi_\Xi) \mid \Xi]\big] \\
&\geqslant \mathbb{E}\big[\mathbb{E}[f_1(\Xi, \Psi_\Xi) \mid \Xi] \cdot \mathbb{E}[f_2(\Xi, \Psi_\Xi) \mid \Xi]\big] \\
&\geqslant \mathbb{E}[f_1(\Xi, \Psi_\Xi)\big] \cdot \mathbb{E}[f_2(\Xi, \Psi_\Xi)],
\end{aligned}$$

by applying FKG to the random graph conditioned to have Ξ as its vertex set for the first inequality, and then FKG for point processes for the second inequality [173, Theorem 20.4]. We say that an event $\mathcal{A}$ is decreasing iff the function $-\mathbb{1}_\mathcal{A}$ is increasing. It follows that for decreasing events $\mathcal{A}, \mathcal{A}'$

$$\begin{aligned}
\mathbb{P}\big(\mathcal{A} \cap \mathcal{A}'\big) &= \mathbb{E}[(-\mathbb{1}_\mathcal{A}(\Xi, \Psi_\Xi)) \cdot (-\mathbb{1}_{\mathcal{A}'}(\Xi, \Psi_\Xi))] \\
&\geqslant \mathbb{E}[\mathbb{1}_\mathcal{A}(\Xi, \Psi_\Xi)] \cdot \mathbb{E}[\mathbb{1}_{\mathcal{A}'}(\Xi, \Psi_\Xi)] = \mathbb{P}\big(\mathcal{A}\big) \cdot \mathbb{P}\big(\mathcal{A}'\big).
\end{aligned} \tag{5.6.42}$$

Observe that the events on the right-hand side in (5.6.41) are all decreasing (adding vertices/edges make the events less likely to occur) so that (5.6.42) applies. Hence,

$$\mathbb{P}\big(|\mathcal{C}_n^{(1)}| < \rho n\big) \geqslant \prod_{i \leqslant \lceil C/\rho \rceil} \mathbb{P}\big(|\mathcal{E}(\Xi^{(i)}, \Xi \setminus \Xi^{(i)})| = 0\big) \cdot \mathbb{P}\big(|\Xi^{(i)}| < \rho n\big). \quad (5.6.43)$$

Since each ball has volume $n\rho/2$, the event $\{|\Xi^{(i)}| < \rho n\}$ holds with probability at least $1/2$ by concentration inequalities for Poisson random variables (Lemma 5.C.1 for $x = 2$). To bound $\mathbb{P}\big(|\mathcal{E}(\Xi^{(i)}, \Xi \setminus \Xi^{(i)})| = 0\big)$, we consider the optimally-suppressed mark profile translated to the center of the i-th ball, with kM_{in} replaced by $n\rho/2$. We restrict Ξ to be below the mark profile, and to have no edges between $\Xi^{(i)}$ and $\Xi \setminus \Xi^{(i)}$. We apply Lemma 5.6.2, integrate over all realisations of $\mathcal{G}_{\text{in}}, \mathcal{G}_{\text{out}}$ satisfying $\mathcal{A}_{\text{regular}}$, and use that the event $\mathcal{A}_{\text{regular}}$ in the conditioning in $\widetilde{\mathbb{P}}_{\text{io}}$ in (5.6.18) holds with high probability by Poisson concentration (Lemma 5.C.1 for $x = 2$), see the argument below (5.6.21). We obtain that for all $i \leqslant \lceil C/\rho \rceil$,

$$\mathbb{P}\big(|\mathcal{E}(\Xi^{(i)}, \Xi \setminus \Xi^{(i)})| = 0\big) \cdot \mathbb{P}\big(|\Xi^{(i)}| < \rho n\big) \geqslant \tfrac{1}{2} \exp\big(-\Theta(n^{\max(\mathcal{Z})} \log^{m_{\mathcal{Z}}-1}(n))\big),$$

which yields the desired statement in (5.1.11) when taking the product over $\lceil C/\rho \rceil$ balls in (5.6.43). □

5.A PROOFS BASED ON BACKBONE CONSTRUCTION

We present the proofs of the propositions at the end of Section 5.4.

Proof of Proposition 5.4.12. We will first derive a bound on

$$\mathbb{P}\big(\exists v \in \Xi_n[\overline{w}, \infty) : v \notin \mathcal{C}_n^{(1)}\big)$$

for arbitrary $\overline{w} \geqslant 1$. We make use of the backbone construction from Section 5.4: we will show that vertices with mark at least $\overline{w}$ are likely to connect to the backbone $\mathcal{C}_{\text{bb}}$, which will be a subset of the giant component. Observe that the event in (5.4.44) allows us to choose the size of the boxes when we build the backbone, i.e., the value of k is not yet defined with respect to $\overline{w}$. We define $k = k(\overline{w})$ implicitly by $\overline{w} =: A_1 k^{1-\sigma\gamma_{\text{hh}}}$, where A_1 is a large enough constant to be determined later. We aim to show that for some $A_2 > 0$,

$$\begin{aligned}\mathbb{P}\big(\neg\mathcal{A}_{\text{mark-giant}}(n, \overline{w})\big) &:= \mathbb{P}\big(\exists v \in \Xi_n[A_1 k^{1-\sigma\gamma_{\text{hh}}}, \infty) : v \notin \mathcal{C}_n^{(1)}\big) \\ &\leqslant n \exp\big(-A_2 k^{\zeta_{\text{hh}}}\big).\end{aligned} \quad (5.A.1)$$

If this bound holds, then substituting $\overline{w} = \overline{w}_n = (M_w \log(n))^{(1-\sigma\gamma_{hh})/\zeta_{hh}}$ yields

$$k(\overline{w}_n) = A_1^{-1/(1-\sigma\gamma_{hh})}\overline{w}_n^{1/(1-\sigma\gamma_{hh})} = A_1^{-1/(1-\sigma\gamma_{hh})}(M_w \log n)^{1/\zeta_{hh}},$$

which, when substituted back to (5.A.1) yields (5.4.44) if M_w is chosen sufficiently large.

Recall $\mathcal{A}_{bb}(n,k)$ and $\mathcal{C}_{bb}(n,k)$ from (5.4.8). Distinguishing two cases depending on whether $\mathcal{A}_{bb}(n,k)$ holds for $\mathcal{G}_{n,2} = \mathcal{G}_n[1, 2w_{hh})$ or not (with $w_{hh}(k)$ in (5.4.7)), by Lemma 5.4.2,

$$\begin{aligned}\mathbb{P}\big(\neg\mathcal{A}_{mark-giant}(n,\overline{w})\big) &\leqslant \mathbb{P}\big(\neg\mathcal{A}_{bb}\big) \\ &\qquad + \mathbb{E}\big[\mathbb{1}_{\{\mathcal{A}_{bb}\}}\mathbb{P}\big(\neg\mathcal{A}_{mark-giant}(n,\overline{w}) \mid \mathcal{G}_{n,2}, \mathcal{A}_{bb}\big)\big] \\ &\leqslant 3n\exp(-c_{5.4.2}s_k) \\ &\qquad + \mathbb{E}\big[\mathbb{1}_{\{\mathcal{A}_{bb}\}}\mathbb{P}\big(\neg\mathcal{A}_{mark-giant}(n,\overline{w}) \mid \mathcal{G}_{n,2}, \mathcal{A}_{bb}\big)\big].\end{aligned} \tag{5.A.2}$$

On the event $\mathcal{A}_{bb}$, there is a backbone $\mathcal{C}_{bb}$. This backbone is either not part of the giant component, or if it is, then a vertex with mark at least $\overline{w}$ outside the giant has *no connection* to any of the vertices in the backbone. Hence, conditionally on the event $\mathcal{A}_{bb}$,

$$\neg\mathcal{A}_{mark-giant}(n,\overline{w}) \subseteq \{\mathcal{C}_{bb} \not\subseteq \mathcal{C}_n^{(1)}\} \cup \{\exists v \in \Xi_n[\overline{w},\infty) : v \not\leftrightarrow \mathcal{C}_{bb}, \mathcal{C}_{bb} \subseteq \mathcal{C}_n^{(1)}\}.$$

By a union bound and Corollary 5.4.10, this implies that

$$\begin{aligned}&\mathbb{P}\big(\neg\mathcal{A}_{mark-giant}(n,\overline{w}) \mid \mathcal{G}_{n,2}, \mathcal{A}_{bb}\big) \\ &\quad \leqslant (n/c_{5.4.1})\exp\big(-c_{5.4.1}k^{\zeta_{hh}}\big) + \mathbb{P}\big(\exists v \in \Xi_n[\overline{w},\infty) : v \not\leftrightarrow \mathcal{C}_{bb} \mid \mathcal{G}_{n,2}, \mathcal{A}_{bb}\big).\end{aligned} \tag{5.A.3}$$

Recall that $\mathcal{G}_{n,2}$ is the graph spanned on vertices with mark in $[1, 2w_{hh})$, see Definition 5.4.6. With C_1 from (5.4.5)–(5.4.6), we may assume $A_1 \geqslant 2C_1^{-1/(\tau-1)}$. Since $w_{hh} = C_1^{-1/(\tau-1)}k^{\gamma_{hh}}$ defined in (5.4.7), and since $1 - (1+\sigma)\gamma_{hh} \geqslant 0$ (see (5.4.14)), this implies that

$$\overline{w} = A_1 k^{1-\sigma\gamma_{hh}} \geqslant 2w_{hh} = 2C_1^{-1/(\tau-1)}k^{\gamma_{hh}}. \tag{5.A.4}$$

Hence, vertices of mark at least $\overline{w}$ are part of $\Xi_n[2w_{hh},\infty)$ and are not revealed in $\mathcal{G}_{n,2}$. Conditioning on the number of vertices $|\Xi_n[\overline{w},\infty)|$, the location of each vertex is independent and uniform in Λ_n. Taking a union bound over these vertices in $\Xi_n[\overline{w},\infty)$, yields

$$\begin{aligned}\mathbb{P}\big(\exists v\in\Xi_n[\overline{w},\infty):u\not\leftrightarrow\mathcal{C}_{\mathrm{bb}}\mid\mathcal{G}_{n,2},\mathcal{A}_{\mathrm{bb}}\big)\\ &\leqslant\mathbb{E}[|\Xi_n[\overline{w},\infty)|]\cdot\mathbb{P}\big(v\not\leftrightarrow\mathcal{C}_{\mathrm{bb}}\mid\mathcal{G}_{n,2},\mathcal{A}_{\mathrm{bb}},v\in\Xi_n[\overline{w},\infty)\big)\\ &\leqslant n\,\mathbb{P}\big(v\not\leftrightarrow\mathcal{C}_{\mathrm{bb}}\mid\mathcal{G}_{n,2},\mathcal{A}_{\mathrm{bb}},v\in\Xi_n[\overline{w},\infty)\big).\end{aligned}\tag{5.A.5}$$

Let $\mathcal{Q}(v)$ be as in (5.4.3). Conditionally on $\mathcal{A}_{\mathrm{bb}}$, $\mathcal{Q}(v)$ contains at least $s_k=\Theta(k^{\zeta_{\mathrm{hh}}})$ vertices in $\mathcal{C}_{\mathrm{bb}}$ with mark in $[w_{\mathrm{hh}},2w_{\mathrm{hh}})$, where w_{hh} is defined in (5.4.7), yielding the set of vertices $\mathcal{S}(v)$ in (5.4.22). We use the distance bound in (5.4.4), and p, $\kappa_{1,\sigma}$ defined in (5.1.3), and (1.3.5), respectively, and the value $\overline{w}$ in (5.A.4), to obtain that for any $v\in\Xi_n[\overline{w},\infty)$ and $u\in\mathcal{S}(v)$, when $\alpha<\infty$,

$$\begin{aligned}p(u,v)&\geqslant p\min\big\{1,\big(\beta\kappa_{1,\sigma}(w_{\mathrm{hh}},A_1k^{1-\sigma\gamma_{\mathrm{hh}}})(2\sqrt{d})^{-d}k^{-1}\big)^{\alpha}\big\}\\ &=p\min\big\{1,\big(\beta C_1^{-\sigma/(\tau-1)}k^{\sigma\gamma_{\mathrm{hh}}}A_1k^{1-\sigma\gamma_{\mathrm{hh}}}(2\sqrt{d})^{-d}k^{-1}\big)^{\alpha}\big\}=p,\end{aligned}$$

whenever $A_1\geqslant(2\sqrt{d})^dC_1^{\sigma/(\tau-1)}/\beta$, since the exponent of k in the second term of the minimum is 0. The same bound holds when $\alpha=\infty$. Since v connects by an edge to each of the $s_k=\Theta(k^{\zeta_{\mathrm{hh}}})$ many backbone vertices in $\mathcal{S}(u)$ with probability at least p, conditionally independently of each other, we bound (5.A.5) by

$$\mathbb{P}\big(\exists v\in\Xi_n[\overline{w},\infty):v\not\leftrightarrow\mathcal{C}_{\mathrm{bb}}\mid\mathcal{G}_{n,2},\mathcal{A}_{\mathrm{bb}}\big)\leqslant n(1-p)^{s_k}.$$

Since $s_k=\Theta(k^{\zeta_{hh}})$ in (5.4.7), combining this with (5.A.2) and (5.A.3) yields (5.A.1) for A_2 sufficiently small. As argued below (5.A.1), this yields (5.4.44). □

5.A.1 *Construction of a linear-sized component*

This section presents the proof of the first statement of Proposition 5.4.13. At the end of the section, we comment how the proof can be adjusted to obtain the second statement considering the infinite model. Throughout the proof, we will consider the Palm version of $\mathbb{P}$, conditioning Ξ to contain a vertex at location 0. We will leave it out in the notation. We will show using a second-moment method that linearly many vertices connect to the backbone $\mathcal{C}_{\mathrm{bb}}(n,k)$ for a properly chosen k. First we introduce some notation. Recall C_1 from (5.4.5)-(5.4.6). We implicitly define $k=k_n$ as the solution of the equation

$$(C_1/16)k_n^{\zeta_{\mathrm{hh}}}:=(2/c_{5.4.1})\log(n),\tag{5.A.6}$$

yielding $m = (2/c_{5.4.1})^{1/\zeta_{hh}} 16/C_1$ in the statement of Proposition 5.4.13, and the mark-truncation value in the definition of $\mathcal{G}_{n,2}$ at $2w_{hh}(k_n) = 2C_1^{-1/(\tau-1)} k_n^{\gamma_{hh}}$ from (5.4.7), with $w_{hh}(k_n) = (M_w \log(n))^{\gamma_{hh}/\zeta_{hh}}$ for some constant M_w. We reveal the graph $\mathcal{G}_{n,1} = \mathcal{G}_n[w_{hh}(k_n), 2w_{hh}(k_n))$ (defined above (5.4.5)), conditioned to satisfy the event $\mathcal{A}_{bb}(n, k_n)$ in (5.4.8). Recalling the intensity measure of the Poisson vertex set μ_τ from (5.1.2), for any constant $m_w \geqslant 1$ we define the event

$$\mathcal{A}'_{reg} := \big\{ |\Xi_n[m_w, 2m_w)| / \mu_\tau\big(\Lambda_n \times [m_w, 2m_w)\big) \in [1/2, 2] \big\}, \qquad (5.A.7)$$

that is, that the PPP in Λ_n is regular in the sense that the number of constant-mark vertices is roughly as expected. Writing w_0 for the mark of 0, we define the conditional probability measure

$$\widetilde{\mathbb{P}}_{bb}(\,\cdot\,) := \mathbb{P}\big(\,\cdot \mid \mathcal{A}'_{reg}, \mathcal{A}_{bb}(n, k_n), \mathcal{G}_{n,1}, w_0 \in [m_w, 2m_w)\big), \qquad (5.A.8)$$

with corresponding expectation $\widetilde{\mathbb{E}}_{bb}$. We state a lemma that implies Proposition 5.4.13.

Lemma 5.A.1 (Constructing a component). *Consider a supercritical i-KSRG under the same conditions as Proposition 5.4.13. Take* k_n *as in* (5.A.6)*. For any* $m_w \geqslant 1$*, for all sufficiently large* n*,*

$$\mathbb{P}\big(\mathcal{A}'_{reg}, \mathcal{A}_{bb}(n, k_n), w_0 \in [m_w, 2m_w)\big) \geqslant m_w^{-(\tau-1)}/2. \qquad (5.A.9)$$

Moreover, there exist constants $m_w \geqslant 1, \rho > 0$*, such that for all sufficiently large* n

$$\widetilde{\mathbb{P}}_{bb}\big(|\mathcal{C}_{n,2}(0)| \geqslant \rho_{5.A.1} n\big) \geqslant \rho_{5.A.1}. \qquad (5.A.10)$$

To prove the lemma (in particular the second statement), we need to define auxiliary notation and an auxiliary claim: we define for $u := (x_u, w_u) \in \Xi_n[m_w, 2m_w)$ the event

$$\{u \xrightarrow{\pi} \mathcal{C}_{bb}\} := \left\{ \begin{array}{l} \exists \text{ a path in } \mathcal{G}_{n,2} \text{ from } u \text{ to } \mathcal{C}_{bb}(n, k_n)\text{, s.t.} \\ \text{vertices (except } u\text{) having mark in } \mathbb{R}^+ \setminus [m_w, 2m_w) \end{array} \right\}. \qquad (5.A.11)$$

Recall $\Lambda(0, s)$ from (5.1.1) the box centered at 0 of volume s, and that $u = (x_u, w_u)$ denotes the location and mark of a vertex $u \in \Xi$. The next claim states that if u and 0 are both vertices with mark in $[m_w, 2m_w)$, falling into different subboxes $\mathcal{Q}(0) \neq \mathcal{Q}(u)$ ((5.4.3)) with respect to the boxing defined by k_n above (5.4.3), then the event that both 0 and u connect to the backbone $\mathcal{C}_{bb}$ happens with constant probability.

Claim 5.A.2 (Paths to the backbone)**.** *Consider a supercritical i-KSRG under the conditions of Proposition 5.4.13. There exist positive constants* m_w*,* $q_{5.A.2}$*,* $C_{5.A.2} > 0$ *such that for all* $u \in \Xi_n[m_w, 2m_w)$ *with* $x_u \notin \Lambda(0, C_{5.A.2}k_n)$*, and* n *sufficiently large*

$$\widetilde{\mathbb{P}}_{bb}\left(\begin{matrix}\{0 \xrightarrow{\pi} \mathcal{C}_{bb}(n,k_n)\} \\ \cap\{u \xrightarrow{\pi} \mathcal{C}_{bb}(n,k_n)\}\end{matrix} \,\middle|\, \begin{matrix} u \in \Xi_n[m_w, 2m_w), \\ x_u \notin \Lambda(0, C_{5.A.2}k_n)\end{matrix}\right) \geqslant q_{5.A.2}.$$

Proof. We will build what we call "mark-increasing paths". Recall $k_n = (m\log(n))^{1/\zeta_{hh}}$ in Prop. 5.4.13 and that $w_{hh}(k_n) = (M_w\log(n))^{\gamma_{hh}/\zeta_{hh}}$ computed below (5.A.6). Define

$$j_* := \max\{j : 2^{j+1}m_w < w_{hh}(k_n)\}, \tag{5.A.12}$$

and define for $0 \leqslant j \leqslant j_*$ and $x \in \Lambda_n$ the following boxes and disjoint mark intervals

$$\begin{aligned} Q_j(x) &:= \Lambda\big(x, \beta 2^{-\sigma} d^{-d/2}(2^j m_w)^{\sigma+1}\big) \cap \Lambda_n, \\ I_j &:= [2^j m_w, 2^{j+1}m_w), \end{aligned} \tag{5.A.13}$$

and write $Q_{j_*+1} := \mathcal{Q}(x)$, i.e., the volume-$k_n$ subbox containing x in the partitioning of Λ_n for the backbone construction given in (5.4.3), and define also $I_{j_*+1} := [w_{hh}(k_n), 2w_{hh}(k_n))$. Even with the truncation by Λ_n in (5.A.13), the volume bound

$$\beta 2^{-\sigma-d} d^{-d/2}(2^j m_w)^{\sigma+1} \leqslant \mathrm{Vol}(Q_j(x)) \leqslant \beta 2^{-\sigma} d^{-d/2}(2^j m_w)^{\sigma+1} \tag{5.A.14}$$

holds for all $x \in \Lambda_n$ and $j \leqslant j_*$ similarly to (5.3.4). By (5.1.2), the number of vertices of Ξ in $Q_j(x_u) \times I_j$ has Poisson distribution with mean

$$\mu_\tau\big(Q_j(x_u) \times I_j\big) \geqslant 2^{-1}\beta(2^j m_w)^{-(\tau-1)}2^{-\sigma-d}d^{-d/2}(2^j m_w)^{\sigma+1}, \tag{5.A.15}$$

where we used that $1 - 2^{-(\tau-1)} \geqslant 2^{-1}$ for $\tau > 2$. Moreover, (5.A.12) implies $2^j m_w \leqslant w_{hh}(k_n)$ for all $j \leqslant j_\star$ and substituting this in (5.A.13) with $w_{hh}(k_n) = C_1^{-1/(\tau-1)}k_n^{\gamma_{hh}}$ from (5.4.7) yields

$$\begin{aligned} Q_j(x) &\subseteq \Lambda\big(x, \beta 2^{-\sigma} d^{-d/2}(w_{hh}(n))^{\sigma+1}\big) \\ &= \Lambda\big(x, \beta 2^{-\sigma} d^{-d/2} C_1^{-(\sigma+1)/(\tau-1)} k_n^{\gamma_{hh}(\sigma+1)}\big) \\ &\subseteq \Lambda\big(x, \beta 2^{-\sigma} d^{-d/2} C_1^{-(\sigma+1)/(\tau-1)} k_n\big) =: Q^\star(x), \end{aligned}$$

where the last inclusion follows from $\gamma_{hh} \leqslant 1/(\sigma+1)$ by (5.4.14). Let $\mathrm{diam}(Q^\star(x)) =: Ck_n^{1/d}$ denote the diameter of $Q^\star(x)$, and let $Q^\diamond :=$

$\Lambda(0, 2^{1/d}Ck_n)$ denote the box centered at 0, so $\mathrm{diam}(Q^\diamond) = 2\mathrm{diam}(Q^\star)$. For any $x_u \notin \mathcal{Q}^\diamond$, and all pairs $j, j' \leqslant j_*$, it holds that

$$Q_j(x_u) \cap Q_{j'}(0) = \emptyset, \tag{5.A.16}$$

and thus the PPPs restricted to $Q_j(x_u) \times I_j$ are independent of each other and of $Q_{j'}(0) \times I_{j'}$ for all $j, j' \leqslant j_*$. On the conditional measure $\widetilde{\mathbb{P}}_{\mathrm{bb}}$, defined in (5.A.8), we fixed (revealed) the realization of $\Xi_n[w_{\mathrm{hh}}(k_n), 2w_{\mathrm{hh}}(k_n))$. Edges among $\Xi_n[w_{\mathrm{hh}}(k_n), 2w_{\mathrm{hh}}(k_n))$ and $\Xi \cap (Q_{j_*}(x_u) \times I_{j_*})$ are thus also present conditionally independently. We define for $u = (x_u, w_u) =: u_0$ the event of having a "mark-increasing path" (a subevent of $\{u \xrightarrow{\pi} \mathcal{C}_{\mathrm{bb}}\}$ defined in (5.A.11)):

$$\{u \rightsquigarrow \mathcal{C}_{\mathrm{bb}}\} := \begin{Bmatrix} \exists(u_1, \ldots, u_{j_*+1}), u_{j_*+1} \in \mathcal{C}_{\mathrm{bb}} : \\ \forall j \in [j_* + 1] : u_j \in Q_j(x_u) \times I_j, u_{j-1} \leftrightarrow u_j \end{Bmatrix}. \tag{5.A.17}$$

on which there is a path from u to the backbone, where the jth vertex on the path is in $Q_j(x_u) \times I_j$ (that are disjoint across j). The mark of u_{j_*+1} is in the right range by definition I_{j_*+1} of below (5.A.13). By this disjointness and (5.A.16), the events $\{0 \rightsquigarrow \mathcal{C}_{\mathrm{bb}}\}$ and $\{u \rightsquigarrow \mathcal{C}_{\mathrm{bb}}\}$ are independent conditionally on $\Xi_n[w_{\mathrm{hh}}(k_n), 2w_{\mathrm{hh}}(k_n))$.

To bound $\widetilde{\mathbb{P}}_{\mathrm{bb}}(u \rightsquigarrow \mathcal{C}_{\mathrm{bb}})$ from below, we greedily 'construct' a path from $u = u_0$ to the backbone. By assumption $w_u \in [m_w, 2m_w)$ hence $u_0 \in Q_0(x_u) \times I_0$. We first bound the probability that u_0 connects by an edge to a vertex $u_1 \in Q_1(x_u) \times I_1$. Then, if there is such a connection, we choose u_1 to be an arbitrary vertex connected to u_0, and give a uniform lower bound (over the possible u_1) on the probability that it connects by an edge to a vertex $u_2 \in Q_2(x_u) \times I_2$. We continue this process until we reach u_{j_*} that has mark just smaller than the minimal mark of vertices in the backbone, by definition of j_* in (5.6.15). Then we find a connection from u_{j_*} to the backbone. We now bound the probability that two vertices u_{j-1} and u_j are connected by an edge.

By construction, $w_{u_{j-1}} \geqslant 2^{j-1}m_w$, $w_{u_j} \geqslant 2^j m_w$. Hence, with the kernel $\kappa_{1,\sigma}$ from (1.3.5), and volume bound (5.A.14), we obtain

$$\begin{aligned}\beta\kappa_{1,\sigma}(w_{u_{j-1}}, w_{u_j}) = \beta w_{u_{j-1}}^{\sigma} w_{u_j} \geqslant \beta m_w^{\sigma+1} 2^{(\sigma+1)j-\sigma} &\geqslant \beta 2^{-\sigma}(2^j m_w)^{\sigma+1} \\ &\geqslant d^{d/2}\mathrm{Vol}(Q_j(x_u)).\end{aligned}$$

Further, $\|x_{u_{j-1}} - x_{u_j}\|^d \leqslant d^{d/2}\mathrm{Vol}(Q_j(x_u)) = \beta 2^{-\sigma}(2^j m_w)^{\sigma+1}$ by (5.A.13), hence

$$p(u_{j-1}, u_j) \geqslant p \min\Big\{1, \frac{(\beta\kappa_{1,\sigma}(w_{u_j}, w_{u_{j+1}}))^\alpha}{\|x_{u_j} - x_{u_{j+1}}\|^{\alpha d}}\Big\} = p.$$

The same computation (without α in the exponent) is valid for $\alpha = \infty$. Having already chosen u_{j-1} on the path, each $v \in \Xi_n \cap (Q_j(x_u) \times I_j)$ connects independently by an edge to u_{j-1} with probability p. The conditioning in the measure $\widetilde{\mathbb{P}}_{bb}$ in (5.A.8) only affects the number of vertices with mark in $[m_w, 2m_w)$ and in $[w_{hh}, 2w_{hh})$. Due to independence of the number of points of PPPs in disjoint sets, for $j \leqslant j_*$, the number of candidate vertices for the role of u_j is thus stochastically dominated from below by a $\mathrm{Poi}\big(p\mu_\tau(Q_j(x_u) \times I_j)\big)$ random variable. The mean is at least $p\beta 2^{-\sigma-d-1} d^{-d/2} (2^j m_w)^{2+\sigma-\tau}$ by (5.1.2) and (5.A.15). For $j = j_* + 1$ we use that the vertices in the backbone in $Q_{j_*+1} \times I_{j_*+1}$ are exactly those in $\mathcal{Q}(u) \times [w_{hh}, 2w_{hh})$, that we denoted by $\mathcal{S}(u)$ in (5.4.22):

$$\begin{aligned} \widetilde{\mathbb{P}}_{bb}\big(\neg\{u \rightsquigarrow \mathcal{C}_{bb}\}\big) &\leqslant \widetilde{\mathbb{P}}_{bb}\big(u_{j_*} \not\leftrightarrow \mathcal{S}(u)\big) \\ &\quad + \sum_{j=1}^{j_*} \mathbb{P}\big(\mathrm{Poi}(p\beta 2^{-\sigma-d-1} d^{-d/2} (2^j m_w)^{2+\sigma-\tau}) = 0\big) \\ &\leqslant \widetilde{\mathbb{P}}_{bb}\big(u_{j_*} \not\leftrightarrow \mathcal{S}(u)\big) \\ &\quad + \sum_{j=1}^{\infty} \exp\big(-p\beta 2^{-\sigma-d-1} d^{-d/2} (2^j m_w)^{2+\sigma-\tau}\big). \end{aligned} \tag{5.A.18}$$

Since $\tau < 2 + \sigma$ by assumption, the sum is finite and decreasing in m_w. By definition of j_* and I_{j_*} in (5.A.12), $w_{u_{j_*}} \geqslant 2^{j_*} m_w \geqslant w_{hh}/4$. We recall that $\mathcal{Q}(x_u)$ is the subbox of volume k_n in the partitioning for the backbone containing x_u ((5.4.3)), that contains at least s_{k_n} backbone vertices of mark at least $w_{hh}(k_n)$, both defined in (5.4.7). When $\alpha < \infty$, we follow the computations in (5.4.17), (that are also valid under the conditional measure $\widetilde{\mathbb{P}}_{bb}$ from (5.A.8), ensuring that $\mathcal{C}_{bb}(n, k_n)$ exists and $|\mathcal{S}(u)| \geqslant s_{k_n}$ by $\mathcal{A}_{bb}$ in (5.4.8)), and use that u_{j_*} is at distance at most $2\sqrt{d}k_n^{1/d}$ from any vertex in $\mathcal{S}(u)$, to obtain

$$\begin{aligned} &\widetilde{\mathbb{P}}_{bb}\big(u_{j_*} \not\leftrightarrow \mathcal{S}(u)\big) \\ &\quad\geqslant 1 - \big(1 - p\min\{1, \beta^\alpha d^{-\alpha d/2} 2^{-(2\sigma+d)\alpha} w_{hh}^{(1+\sigma)\alpha} k_n^{-\alpha}\}\big)^{s_{k_n}} \\ &\quad\geqslant 1 - \exp\big(p\min\{s_{k_n}, \beta^\alpha d^{-\alpha d/2} 2^{-(2\sigma+d)\alpha} w_{hh}^{(1+\sigma)\alpha} k_n^{-\alpha} s_{k_n}\}\big). \end{aligned} \tag{5.A.19}$$

Using that $s_{k_n} = k_n w_{hh}^{-(\tau-1)}/16$, $w_{hh} = C_1^{-1/(\tau-1)} k_n^{\gamma_{hh}}$, and $\gamma_{hh} = (\alpha - 1)/((\sigma+1)\alpha - (\tau-1))$, we have

$$w_{hh}^{(1+\sigma)\alpha} k_n^{-\alpha} s_{k_n} \geqslant 2^{-4} w_{hh}^{(1+\sigma)\alpha-(\tau-1)} k_n^{1-\alpha} = 2^{-4} C_1^{1-(1+\sigma)\alpha/(\tau-1)}.$$

Using this bound on the right-hand side in (5.A.19), this yields combined with (5.A.18) that there exists a constant $q > 0$ such that if m_w is sufficiently large $\widetilde{\mathbb{P}}_{bb}(u \rightsquigarrow \mathcal{C}_{bb}) \geqslant q$ establishing Claim 5.A.2 when $\alpha < \infty$, by the reasoning about independence below (5.A.17).

When $\alpha = \infty$, the choice of C_1 in (5.4.6) ensures that for any vertex $u_{bb} \in \mathcal{C}_{bb} \cap \mathcal{Q}(u)$ that

$$p(u_{j_*}, u_{bb}) \geqslant p\mathbb{1}\Big\{\beta \frac{\kappa_{1,\sigma}(w_{hh}/4, w_{hh})}{2^d d^{d/2} k_n} \geqslant 1\Big\} = p,$$

establishing Lemma 5.A.2 for $\alpha = \infty$ when combined with (5.A.18) for m_w sufficiently large. □

We are now ready to prove Lemma 5.A.1.

Proof of Lemma 5.A.1. We first show (5.A.9). By a union bound, concentration inequalities for Poisson random variables (Lemma 5.C.1 for $x \in \{1/2, 2\}$) and $F_W(dw_0) = (\tau-1)w_0^{-\tau}dw$ in Definition 5.1.1, it follows that $\mathbb{P}(\neg\mathcal{A}'_{reg}) = \exp(-\Theta(nm_w^{-(\tau-1)})) = o(1)$, and thus

$$\begin{aligned}
&\mathbb{P}\big(\neg\big(\mathcal{A}'_{reg} \cap \mathcal{A}_{bb} \cap \{w_0 \in [m_w, 2m_w)\}\big)\big) \\
&\quad \leqslant \mathbb{P}\big(\neg\mathcal{A}_{bb}\big) + \mathbb{P}\big(w_0 \notin [m_w, 2m_w)\big) + o(1) \\
&\quad \leqslant \mathbb{P}\big(\neg\mathcal{A}_{bb}\big) + 1 - (1 - 2^{-(\tau-1)})m_w^{-(\tau-1)} + o(1).
\end{aligned}$$

By the choice of $k = k_n$ in (5.A.6), the first term tends to zero by Lemma 5.4.2 as n tends to infinity. Since $1 - 2^{-(\tau-1)} > 1/2$ for $\tau > 2$, for n sufficiently large (depending also on the constant m_w) it follows that

$$\begin{aligned}
\mathbb{P}\big(\neg\big(\mathcal{A}'_{reg} \cap \mathcal{A}_{bb} \cap \{w_0 \in [m_w, 2m_w)\}\big)\big) &\leqslant 1 - m_w^{-(\tau-1)}(1 - 2^{-(\tau-1)}) + o(1) \\
&\leqslant 1 - m_w^{-(\tau-1)}/2,
\end{aligned}$$

and (5.A.9) follows. We proceed to (5.A.10). Conditionally on the realization of $\mathcal{G}_{n,1}$ satisfying $\mathcal{A}_{bb}$ (present in the conditioning in $\widetilde{\mathbb{P}}_{bb}$ in (5.A.8)), we define the following set and random variable:

$$\begin{aligned}
\mathcal{U} &:= \{u \in \Xi_n[m_w, 2m_w) : u \xrightarrow{\pi} \mathcal{C}_{bb}\}, \\
X &:= \mathbb{1}_{\{0 \xrightarrow{\pi} \mathcal{C}_{bb}\}} \sum_{\substack{u \in \Xi_n[m_w, 2m_w): \\ x_u \notin \Lambda(0, C_{5.A.2}k_n)}} \mathbb{1}_{\{u \xrightarrow{\pi} \mathcal{C}_{bb}\}},
\end{aligned} \tag{5.A.20}$$

with $\xrightarrow{\pi}$ from (5.A.11). The measure $\widetilde{\mathbb{P}}_{\mathrm{bb}}$ is a conditional measure where $\mathcal{A}'_{\mathrm{reg}}$ (defined in (5.A.7)) holds and so

$$|\Xi_n[m_w, 2m_w)| \leqslant 2\mu_\tau(\Lambda_n \times [m_w, 2m_w)).$$

Using μ_τ in (5.1.2), we obtain that deterministically under $\widetilde{\mathbb{P}}_{\mathrm{bb}}$:

$$X^2 \leqslant 4\big(\mu_\tau(\Lambda_n \times [m_w, 2m_w)\big)^2 \leqslant 4(m_w^{-(\tau-1)}n)^2.$$

When $0 \in \mathcal{U}$ holds, then $|\mathcal{C}_{n,2}(0)| \geqslant |\mathcal{U}| \geqslant X$, and so we apply Paley-Zygmund's inequality to X under the measure $\widetilde{\mathbb{P}}_{\mathrm{bb}}$, which yields for $\rho' := \widetilde{\mathbb{E}}_{\mathrm{bb}}[X]/(2n)$ that

$$\begin{aligned}\widetilde{\mathbb{P}}_{\mathrm{bb}}\big(|\mathcal{C}_{n,2}(0)| \geqslant \rho' n\big) &\geqslant \widetilde{\mathbb{P}}_{\mathrm{bb}}\big(|\mathcal{U}| \geqslant \rho' n,\ 0 \in \mathcal{U}\big) \geqslant \widetilde{\mathbb{P}}_{\mathrm{bb}}\big(X \geqslant \widetilde{\mathbb{E}}_{\mathrm{bb}}[X]/2\big) \\ &\geqslant (1/4)\frac{\widetilde{\mathbb{E}}_{\mathrm{bb}}[X]^2}{\widetilde{\mathbb{E}}_{\mathrm{bb}}[X^2]} \geqslant \frac{\widetilde{\mathbb{E}}_{\mathrm{bb}}[X]^2}{16n^2 m_w^{-2(\tau-1)}}.\end{aligned} \tag{5.A.21}$$

We now bound the numerator on the right-hand side from below. Conditionally on $|\Xi_n[m_w, 2m_w)|$, the vertices have a uniform location in Λ_n, so

$$\begin{aligned}&\widetilde{\mathbb{P}}_{\mathrm{bb}}\big(x_u \notin \Lambda(0, C_{5.A.2}k_n) \mid u \in \Xi_n[m_w, 2m_w), |\Xi_n[m_w, 2m_w)|\big) \\ &\qquad = \big(n - C_{5.A.2}k_n\big)/n \geqslant 1/2,\end{aligned}$$

where the last inequality follows from assuming that n is sufficiently large (recall $k_n = \Theta\big(\log^{1/\zeta_{\mathrm{hh}}}(n)\big)$ by (5.A.6)). The conditioning on $\mathcal{A}'_{\mathrm{reg}}$ implies that $|\Xi_n[m_w, 2m_w)| \geqslant \mu_\tau(\Lambda_n \times [m_w, 2m_w))/2$. Using linearity of expectation of X in (5.A.20), and the tower rule, (by first conditioning on $|\Xi_n[m_w, 2m_w)|$) we obtain for n sufficiently large

$$\begin{aligned}&\widetilde{\mathbb{E}}_{\mathrm{bb}}\big[X\big] \geqslant (\mu_\tau(\Lambda_n \times [m_w, 2m_w))/2) \\ &\cdot \widetilde{\mathbb{E}}_{\mathrm{bb}}\Big[\widetilde{\mathbb{P}}_{\mathrm{bb}}\big(x_u \notin \Lambda(0, C_{5.A.2}k_n) \mid u \in \Xi_n[m_w, 2m_w), |\Xi_n[m_w, 2m_w)|\big)\Big] \\ &\cdot \widetilde{\mathbb{P}}_{\mathrm{bb}}\big(\{0 \xrightarrow{\pi} \mathcal{C}_{\mathrm{bb}}\} \cap \{u \xrightarrow{\pi} \mathcal{C}_{\mathrm{bb}}\} \mid u \in \Xi_n[m_w, 2m_w), x_u \notin \Lambda(0, C_{5.A.2}k_n)\big) \\ &\qquad \geqslant n m_w^{-(\tau-1)}(1 - 2^{-(\tau-1)})(1/2) \cdot (1/2) \cdot q_{5.A.2} \\ &\qquad \geqslant n q_{5.A.2} m_w^{-(\tau-1)} 2^{-3},\end{aligned}$$

where the second bound follows if m_w is chosen as in Claim 5.A.2, and from the definition of μ_τ in (5.1.2); the last bound holds since $2^{-(\tau-1)} \leqslant$

1/2 for $\tau > 2$. Substituting the last bound (5.A.22) in the numerator on the right-hand side to (5.A.21), we have obtained that

$$\widetilde{\mathbb{P}}_{\mathrm{bb}}\big(|\mathcal{C}_{n,2}(0)| \geqslant \rho' n\big) \geqslant 2^{-10} q_{5.A.2}^2,$$

holds with $\rho' = \widetilde{\mathbb{E}}_{\mathrm{bb}}[X]/(2n) \geqslant q_{5.A.2} m_w^{-(\tau-1)} 2^{-4}$, which yields which yields the statement of Lemma 5.A.1 for

$$\rho_{5.A.1} = \min\{q_{5.A.2} m_w^{-(\tau-1)} 2^{-4}, 2^{-10} q_{5.A.2}^2\}. \qquad \square$$

Proof of Proposition 5.4.13. We start with the first inequality in (5.4.45). Using $\widetilde{\mathbb{P}}_{\mathrm{bb}}$ in (5.A.8), we observe that the bound in Lemma 5.A.1 holds uniformly over all realizations of $\mathcal{G}_{n,1}$ satisfying $\mathcal{A}_{\mathrm{bb}}$. Hence, by first taking expectation over these possible realizations, we obtain that

$$\mathbb{P}(\mathcal{C}_{n,2}(0) \mid \mathcal{A}'_{\mathrm{reg}}, \mathcal{A}_{\mathrm{bb}}(n, k_n), w_0 \in [m_w, 2m_w]) \geqslant \rho_{5.A.1}$$

also holds. The statement now follows with $\rho := \rho_{5.A.1} m_w^{-(\tau-1)}/2$ by the law of total probability combining (5.A.9) and (5.A.10) of Lemma 5.A.1. The second inequality in (5.4.45) for $n = \infty$ follows from the same construction as the greedy path in the proof of Claim 5.A.2 below (5.A.17), making the greedy path infinitely long. We leave it to the reader to fill in the details. $\square$

5.B PROOFS USING FIRST-MOMENT METHOD

We start with the proof of Claim 5.5.4.

Proof of Claim 5.5.4. By Palm theory and symmetry of the PPP μ_τ on $\mathbb{R}^d$, recalling that we conditioned on $0 \in \Xi$ in Definition 5.1.1, we consider $\mathcal{R}_i(0)$ in (5.5.21) for some $i \geqslant 1$. We estimate the mean from below using $t_k = n_k^{1/d}/(2k)$

$$\begin{aligned}\lambda_i := \mu_\tau(\mathcal{R}_i(0)) = (2^i t_k)^d - (2^{i-1} t_k)^d &= (2^{i-1} t_k)^d (2^d - 1) \\ &\geqslant 2^{d(i-1)} (2k)^{-d} n_k \geqslant 2^{d(i-2)} n_k^\delta.\end{aligned}$$

Hence, since $|\Xi_{N_k} \cap \mathcal{R}_i(0)|$ is distributed as $\mathrm{Poi}(\lambda_i)$ with $\lambda_i > 2^{d(i-2)} n_k^\delta$, by Chernoff bounds for Poisson random variables (see Lemma 5.C.1 applied with $x = 2$, using $1 + 2\log(2) - 2 \geqslant 1/4$),

$$\begin{aligned}\mathbb{P}\big(|\Xi_{N_k} \cap \mathcal{R}_i(0)| > 2 \cdot \mu_\tau\big(\mathcal{R}_i(0)\big)\big) &\leqslant \exp(-\lambda_i(1 + 2\log 2 - 2)) \\ &\leqslant \exp\big(-2^{d(i-2)-2} n_k^\delta\big),\end{aligned}$$

and the statement follows for every $c > 0$, i.e.,

$$\begin{aligned}\mathbb{P}\big(\mathcal{A}_{\text{dense}}\big) &= \mathbb{P}\big(\exists u \in \Xi_{n_k}[1,\overline{w}), i \geqslant 1 : |\Xi_{N_k} \cap \mathcal{R}_i(x_u)| > 2 \cdot \mu_\tau\big(\mathcal{R}_i(x_u)\big)\big)\\ &\leqslant \mathbb{E}[|\Xi_{n_k}[1,\overline{w})|] \cdot \mathbb{P}\big(\exists i \geqslant 1 : |\Xi_{N_k} \cap \mathcal{R}_i(0)| > 2 \cdot \mu_\tau\big(\mathcal{R}_i(0)\big)\big)\\ &\leqslant n_k \sum_{i \geqslant 1} \exp\big(-2^{d(i-2)-2} n_k^\delta\big)\\ &\leqslant 2n_k \exp\big(-2^{-d-2} n_k^\delta\big) = o(n_k^{-c}). \qquad \square\end{aligned}$$

Now we prove Claim 5.5.6 used for the upper bound of subexponential decay.

Proof of Claim 5.5.6. For compact sets $\mathcal{K}_1, \mathcal{K}_2 \subseteq \mathbb{R}^d$, we define $\|\mathcal{K}_1 - \mathcal{K}_2\| := \min\{\|x - y\|, x \in \mathcal{K}_1, y \in \mathcal{K}_2\}$. We define

$$\widetilde{t}_k := \|\partial\Lambda_N - \partial\Lambda_n\| = (N^{1/d} - n^{1/d})/2.$$

The definition of $\mathcal{A}_{\text{long-edge}}$ in (5.5.11) implies that

$$\mathcal{A}_{\text{long-edge}}(n, N, \overline{w}) \subseteq \big\{\exists u \in \Xi_n[1,\overline{w}), v \in \Xi : \|x_u - x_v\| \geqslant \widetilde{t}_k, u \leftrightarrow v\big\},$$

so that after conditioning on $\Xi_n[1,\overline{w})$ it follows by a union bound that

$$\begin{aligned}&\mathbb{P}\big(\mathcal{A}_{\text{long-edge}}(n, N, \overline{w})\big)\\ &\quad\leqslant \mathbb{E}\Bigg[\sum_{u \in \Xi_n[1,\overline{w})} \mathbb{P}\big(\exists v \in \Xi : \|x_u - x_v\| \geqslant \tilde{t}_k, u \leftrightarrow v \mid u \in \Xi_n[1,\overline{w})\big)\Bigg]\\ &\quad=: \mathbb{E}\Bigg[\sum_{u \in \Xi_n[1,\overline{w})} q(u)\Bigg] \qquad (5.B.1)\end{aligned}$$

Assume $\alpha < \infty$. Since the diameter of Λ_n is $\sqrt{d}n^{1/d}$, the lower bound on N in the statement of Claim 5.5.6 implies that $\|x_u - x_v\| \leqslant \widetilde{t}_k$ for all $x_u, x_v \in \Lambda_n$. Hence, $v \notin \Xi_n$ whenever $\|x_u - x_v\| > \widetilde{t}_k$. This implies by Markov's bound, using the connection probability in (5.1.3), and that

$w_u \leqslant \overline{w}$ and the intensity μ_τ and the translation invariance of the intensity of Ξ in (5.1.2), that for all $u \in \Xi_n$,

$$q(u) \leqslant \mathbb{E}\Bigg[\sum_{v\in\Xi:\|x_u-x_v\|>\widetilde{t}_k} p\min\Big\{1,\Big(\beta\frac{\kappa_{1,\sigma}(\overline{w},w_v)}{\|x_u-x_v\|^d}\Big)^{\alpha}\Big\}\Bigg]$$

$$= p(\tau-1)\int_{w_v=1}^{\infty}\int_{x_v:\|x_u-x_v\|\geqslant\widetilde{t}_k}\min\Big\{1,\Big(\beta\frac{\kappa_{1,\sigma}(\overline{w},w_v)}{\|x_u-x_v\|^d}\Big)^{\alpha}\Big\}w_v^{-\tau}dw_v dx_v \tag{5.B.2}$$

$$\begin{aligned}&= p(\tau-1)\int_{w_v=1}^{\infty} w_v^{-\tau}\int_{x_v:\|x_v\|\geqslant\widetilde{t}_k,\|x_v\|^d\leqslant\beta\kappa_{1,\sigma}(\overline{w},w_v)} dw_v dx_v\\ &\quad + p(\tau-1)\int_{w_v=1}^{\infty}(\beta\kappa_{1,\sigma}(\overline{w},w_v))^{\alpha}w_v^{-\tau}\\ &\qquad\qquad \int_{x_v:\|x_v\|\geqslant\widetilde{t}_k,\|x_v\|^d\geqslant\beta\kappa_{1,\sigma}(\overline{w},w_v)}\|x_v\|^{-\alpha d}dw_v dx_v\\ &=: T_1+T_2,\end{aligned} \tag{5.B.3}$$

where we cut the integral into two based on the value on the minimum. We analyse the two integrals separately. Analyzing T_1, the integration with respect to x_v gives the Lebesgue measure of the set $\{x_v : \widetilde{t}_k^d \leqslant \|x_v^d\| \leqslant \beta\kappa_{1,\sigma}(\overline{w},w_v)\}$, which is nonzero only if this set is nonempty, and then can be bounded from above by $c_d\beta\kappa_{1,\sigma}(\overline{w},w_v)$ for some constant depending only on d. So we obtain

$$\begin{aligned}T_1 &\leqslant c_d p\beta(\tau-1)\int_{w_v=1}^{\infty}\mathbb{1}\{\widetilde{t}_k^d\leqslant\beta\kappa_{1,\sigma}(\overline{w},w_v)\}\kappa_{1,\sigma}(\overline{w},w_v)w_v^{-\tau}dw_v\\ &= c_d p\beta(\tau-1)\Big(\int_{w_v=1}^{\overline{w}}\mathbb{1}\{\widetilde{t}_k^d\leqslant\beta\overline{w}w_v^{\sigma}\}\overline{w}w_v^{\sigma-\tau}dw_v\\ &\qquad\qquad + \int_{w_v=\overline{w}}^{\infty}\mathbb{1}\{\widetilde{t}_k^d\leqslant\beta\overline{w}^{\sigma}w_v\}\overline{w}^{\sigma}w_v^{1-\tau}dw_v\Big),\end{aligned}$$

where in the last step we used the definition of $\kappa_{1,\sigma}$ in (1.3.5) and cut the integration into two based on the minimum of $\overline{w}$ and w_v. Since we assumed $\widetilde{t}_k \geqslant \beta\overline{w}^{\sigma+1}$ in the statement of the lemma, and $w_v^{\sigma} \leqslant \overline{w}^{\sigma}$, the indicator in the first integral is 0. Moving the indicator in the second integral into the integration boundary yields

$$T_1 \leqslant c_d p\beta\overline{w}^{\sigma}(\widetilde{t}_k^d/(\beta\overline{w}^{\sigma}))^{-(\tau-2)} = c_d p\beta^{\tau-1}\overline{w}^{\sigma(\tau-1)}\widetilde{t}_k^{-d(\tau-2)}. \tag{5.B.4}$$

We turn to T_2 in (5.B.3). For some d-dependent constant c'_d, using again $\kappa_{1,\sigma}$ in (1.3.5)

$$\begin{aligned} T_2 &\leqslant pc'_d(\tau-1)\int_{w_v=1}^{\infty}(\beta\kappa_{1,\sigma}(\overline{w},w_v))^{\alpha}w_v^{-\tau}\max\{\widetilde{t}_k^d,\beta\kappa_{1,\sigma}(\overline{w},w_v)\}^{1-\alpha}\mathrm{d}w_v \\ &= pc'_d(\tau-1)\int_{w_v=1}^{\overline{w}}(\beta\overline{w}w_v^{\sigma})^{\alpha}w_v^{-\tau}\max\{\widetilde{t}_k^d,\beta\overline{w}w_v^{\sigma}\}^{1-\alpha}\mathrm{d}w_v \qquad (5.B.5) \\ &\quad + pc'_d(\tau-1)\int_{w_v=\overline{w}}^{\infty}(\beta\overline{w}^{\sigma}w_v)^{\alpha}w_v^{-\tau}\max\{\widetilde{t}_k^d,\beta\overline{w}^{\sigma}w_v\}^{1-\alpha}\mathrm{d}w_v. \qquad (5.B.6) \end{aligned}$$

For the first integral, we bound $w_v^{\sigma} \leqslant \overline{w}^{\sigma}$, and observe that by the assumption $\widetilde{t}_k \geqslant \beta\overline{w}^{\sigma+1}$ in the statement of the lemma, the maximum is always attained at $\widetilde{t}_k^d$. Hence,

$$(5.B.5) \leqslant pc'_d(\beta\overline{w}^{\sigma+1})^{\alpha}\widetilde{t}_k^{-(\alpha-1)d}. \qquad (5.B.7)$$

We split the integral in (5.B.6) according to where the maximum is attained, i.e.,

$$\begin{aligned} (5.B.6) &= pc'_d\beta\overline{w}^{\sigma}(\tau-1)\int_{w_v=\max\{\overline{w},\widetilde{t}_k^d/(\overline{w}^{\sigma}\beta)\}}^{\infty}w_v^{-(\tau-1)}\mathrm{d}w_v \\ &\quad + \mathbb{1}_{\{\widetilde{t}_k^d/(\overline{w}^{\sigma}\beta)\geqslant\overline{w}\}}pc'_d(\tau-1)\widetilde{t}_k^{-(\alpha-1)d}\int_{w_v=\overline{w}}^{\widetilde{t}_k^d/(\overline{w}^{\sigma}\beta)}(\beta\overline{w}^{\sigma}w_v)^{\alpha}w_v^{-\tau}\mathrm{d}w_v \\ &= pc'_d\beta\overline{w}^{\sigma}(\tau-1)\int_{w_v=\widetilde{t}_k^d/(\overline{w}^{\sigma}\beta)}^{\infty}w_v^{-(\tau-1)}\mathrm{d}w_v \\ &\quad + pc'_d(\tau-1)\widetilde{t}_k^{-(\alpha-1)d}(\beta\overline{w}^{\sigma})^{\alpha}\int_{w_v=\overline{w}}^{\widetilde{t}_k^d/(\overline{w}^{\sigma}\beta)}w_v^{\alpha-\tau}\mathrm{d}w_v \\ &= \frac{pc'_d(\beta\overline{w}^{\sigma})^{\tau-1}(\tau-1)}{\tau-2}\widetilde{t}_k^{-d(\tau-2)} \\ &\quad + pc'_d(\tau-1)\widetilde{t}_k^{-(\alpha-1)d}(\beta\overline{w}^{\sigma})^{\alpha}\int_{w_v=\overline{w}}^{\widetilde{t}_k^d/(\overline{w}^{\sigma}\beta)}w_v^{\alpha-\tau}\mathrm{d}w_v \qquad (5.B.8) \end{aligned}$$

where the second step follows from the assumption that $\widetilde{t}_k \geqslant \beta\overline{w}^{\sigma+1}$ in the statement of the lemma. For the remaining integral on the right-hand side of (5.B.8), say T_{22}, we consider three cases, i.e., for some $C > 0$

$$T_{22} \leqslant \begin{cases} C\widetilde{t}_k^{-(\alpha-1)d}(\beta\overline{w}^{\sigma})^{\alpha}\big(\widetilde{t}_k^d/(\overline{w}^{\sigma}\beta)\big)^{\alpha-(\tau-1)} & \text{if } \alpha > \tau-1, \\ C\widetilde{t}_k^{-(\alpha-1)d}(\beta\overline{w}^{\sigma})^{\alpha}\log\big(\widetilde{t}_k^d/(\overline{w}^{\sigma}\beta)\big) & \text{if } \alpha = \tau-1, \\ C\overline{w}^{\alpha-(\tau-1)}\widetilde{t}_k^{-(\alpha-1)d}(\beta\overline{w}^{\sigma})^{\alpha} & \text{if } \alpha < \tau-1, \end{cases}$$

Elementary rewriting of the first and third case yields

$$T_{22} \leqslant \begin{cases} C\widetilde{t}_k^{-d(\tau-2)}(\beta\overline{w}^{\sigma})^{\tau-1} & \text{if } \alpha > \tau-1, \\ C\widetilde{t}_k^{-d(\alpha-1)}(\beta\overline{w}^{\sigma})^{\alpha}\log\big(\widetilde{t}_k^d/(\overline{w}^{\sigma}\beta)\big) & \text{if } \alpha = \tau-1, \\ C\widetilde{t}_k^{-d(\alpha-1)}\beta^{\alpha}\overline{w}^{\alpha(\sigma+1)-(\tau-1)} & \text{if } \alpha > \tau-1. \end{cases}$$

Combining this bound with (5.B.8), then (5.B.7) and (5.B.4), gives in (5.B.3) and (5.B.1) that for some $C > 0, b > 0$

$$\begin{aligned} &\mathbb{P}\big(\mathcal{A}_{\text{long-edge}}(n, N, \overline{w})\big) \\ &\quad\leqslant C'\overline{w}^b\widetilde{t}_k^{-d\min\{\alpha-1,\tau-2\}}(1+\mathbb{1}_{\{\alpha=\tau-1\}}\log(t_k))\mathbb{E}[|\Xi_n[1,\overline{w})]|] \\ &\quad\leqslant C\overline{w}^b N^{-d\min\{\alpha-1,\tau-2\}}(1+\mathbb{1}_{\{\alpha=\tau-1\}})\log(N))n, \end{aligned}$$

where the last bound follows by the assumed bound in (5.5.27), since $\widetilde{t}_k = (N^{1/d} - n^{1/d})/2$, and the intensity of $\Xi_n[1, \overline{w})$ in (5.1.2). This finishes the proof of (5.5.28) when $\alpha < \infty$.

It remains to show the bound for the case $\alpha = \infty$. In this case, the same calculations hold, with only T_1 in (5.B.3) present, since the connection probability is 0 when the minimum in (5.B.2) is not at 1. □

We proceed with the proofs of two lemmas for the lower bound.

Proof of Lemma 5.6.3. We recall from (5.6.31) that it is sufficient to bound $\mu_\tau(\mathbb{R}^{d+1}_{>\mathcal{M}_\gamma}) = \mathbb{E}\big[\big|\Xi^{\text{in}}_{>\mathcal{M}_\gamma}\big|\big] + \mathbb{E}\big[\big|\Xi^{\text{out}}_{>\mathcal{M}_\gamma}\big|\big]$. We introduce some notation: for two functions $g(k), h(k)$, we write $g \lesssim h$ if $g = O(h)$. Since $f_\gamma(x)$ is symmetric around the boundary of $\partial\mathcal{B}_{kM_{\text{in}}}$ (see its definition in (5.6.9)), it is easy to see that $\mathbb{E}\big[\big|\Xi^{\text{in}}_{>\mathcal{M}_\gamma}\big|\big] \leqslant \mathbb{E}\big[\big|\Xi^{\text{out}}_{>\mathcal{M}_\gamma}\big|\big]$. Since $\zeta_{\text{nn}} = (d-1)/d$ by (4.2.5), for (5.6.29) it is sufficient to show that

$$\mathbb{E}\big[\big|\Xi^{\text{out}}_{>\mathcal{M}_\gamma}\big|\big] \lesssim r^{d(1-\gamma(\tau-1))} + (1+\mathbb{1}_{\{1-\gamma(\tau-1)=1-1/d\}}\log(r))r^{d-1}. \quad (5.B.9)$$

Using the intensity measure of Ξ in (5.1.2), switching to polar coordinates in the first d directions, and integrating with respect to the mark-coordinate, we obtain using the exact form of f_γ in (5.6.9)

$$\begin{aligned} \mathbb{E}\big[\big|\Xi^{\text{out}}_{>\mathcal{M}_\gamma}\big|\big] &\lesssim \int_{z=0}^{\infty}(z+r_k)^{d-1}\int_{f_\gamma(z)}^{\infty}(\tau-1)w^{-\tau}\,dw\,dz \qquad (5.B.10) \\ &\lesssim \int_{z=0}^{C_\beta}(z+r_k)^{d-1}dz + \int_{z=C_\beta}^{r_k}(z+r_k)^{d-1}z^{-d\gamma(\tau-1)}dz \\ &\quad + \int_{z=r_k}^{\infty}(z+r_k)^{d-1}(z^d r_k^{-d(1-\gamma)})^{-(\tau-1)}dz =: I_1 + I_2 + I_3. \end{aligned}$$

The integration length of I_1 is a constant, so $I_1 \lesssim r_k^{d-1}$. For I_2 we apply the binomial theorem, i.e.,

$$I_2 \lesssim \sum_{j=0}^{d-1} r_k^j \int_{C_\beta}^{r_k} z^{(1-\gamma(\tau-1))d-1-j} dz. \tag{5.B.11}$$

Analyzing the summands separately, we obtain for $j \leqslant d-1$

$$r_k^j \int_{C_\beta}^{r_k} z^{(1-\gamma(\tau-1))d-1-j} dz \lesssim \begin{cases} r_k^{d(1-\gamma(\tau-1))} & \text{if } d(1-\gamma(\tau-1)) > j, \\ \log(r_k) r_k^j & \text{if } d(1-\gamma(\tau-1)) = j, \\ r_k^j & \text{if } d(1-\gamma(\tau-1)) < j. \end{cases}$$

Using these bounds, which are increasing for $j \in [d-1]$, in (5.B.11), we obtain

$$I_2 \lesssim (1 + \mathbb{1}_{\{1-\gamma(\tau-1)=1-1/d\}} \log(r_k)) r_k^{d-1} + r_k^{d(1-\gamma(\tau-1))}. \tag{5.B.12}$$

It remains to bound I_3 in (5.B.10). Using that $\tau > 2$ by assumption, and $z + r_k \leqslant 2z$,

$$\begin{aligned} I_3 &\lesssim r_k^{d(1-\gamma)(\tau-1)} \int_{z=r_k}^{\infty} z^{-d(\tau-2)-1} dz \\ &\lesssim r_k^{d((1-\gamma)(\tau-1)-(\tau-2))} = r_k^{d(1-\gamma(\tau-1))}. \end{aligned}$$

Together with the bound on I_1 below the definitions of I_1, I_2, and I_3 in (5.B.11), and on I_2 in (5.B.12), this proves (5.B.9) and also finishes the proof of (5.6.29). □

Proof of Lemma 5.6.4. We split the expected number of edges depending on the locations of the endpoints of the vertices.

$$\begin{aligned} \widetilde{\mathbb{E}}_{\mathrm{io}}\Big[\big|\mathcal{E}\big(\Xi^{\mathrm{in}}_{\leqslant \mathcal{M}_\gamma}, \Xi^{\mathrm{out}}_{\leqslant \mathcal{M}_\gamma}\big)\big|\Big] = {}& \widetilde{\mathbb{E}}_{\mathrm{io}}\Big[\big|\mathcal{E}\big(\Xi^{\mathrm{in}}_{\leqslant \mathcal{M}_\gamma \setminus \mathcal{R}_{\mathrm{in}}}, \Xi^{\mathrm{out}}_{\leqslant \mathcal{M}_\gamma \setminus \mathcal{R}_{\mathrm{out}}}\big)\big|\Big] \\ &+ \widetilde{\mathbb{E}}_{\mathrm{io}}\Big[\big|\mathcal{E}\big(\Xi_{\mathcal{R}_{\mathrm{in}}}, \Xi^{\mathrm{out}}_{\leqslant \mathcal{M}_\gamma \setminus \mathcal{R}_{\mathrm{out}}}\big)\big|\Big] \\ &+ \widetilde{\mathbb{E}}_{\mathrm{io}}\Big[\big|\mathcal{E}\big(\Xi_{\mathcal{R}_{\mathrm{in}}}, \Xi_{\mathcal{R}_{\mathrm{out}}}\big)\big|\Big] \\ &+ \widetilde{\mathbb{E}}_{\mathrm{io}}\Big[\big|\mathcal{E}\big(\Xi^{\mathrm{in}}_{\leqslant \mathcal{M}_\gamma \setminus \mathcal{R}_{\mathrm{in}}}, \Xi_{\mathcal{R}_{\mathrm{out}}}\big)\big|\Big] \end{aligned} \tag{5.B.13}$$

We analyse the first term on the right-hand side and at the end we sketch how the bounds could be adapted for the other three terms. Since $\mathcal{A}_{\mathrm{regular}}(\eta)$ is measurable with respect to $\Xi_{\mathcal{R}_{\mathrm{in}}} \cup \Xi_{\mathcal{R}_{\mathrm{out}}}$, it can be left out of the conditioning. Further, points of Ξ in disjoint sets are independently present, hence

$$\begin{aligned} \mathbb{E}\Big[\big|\mathcal{E}\big(\Xi^{\text{in}}_{\leqslant \mathcal{M}_\gamma\setminus\mathcal{R}_{\text{in}}}, \Xi^{\text{out}}_{\leqslant \mathcal{M}_\gamma\setminus\mathcal{R}_{\text{out}}}\big)\big| \;\Big|\; \Xi_{\mathcal{R}_{\text{in}}}\cup\Xi_{\mathcal{R}_{\text{out}}}, \mathcal{A}_{\text{regular}}(\eta)\Big] \\ = \mathbb{E}\big[\big|\mathcal{E}\big(\Xi^{\text{in}}_{\leqslant \mathcal{M}_\gamma\setminus\mathcal{R}_{\text{in}}}, \Xi^{\text{out}}_{\leqslant \mathcal{M}_\gamma\setminus\mathcal{R}_{\text{out}}}\big)\big|\big] \\ \leqslant \mathbb{E}\big[\big|\mathcal{E}\big(\Xi^{\text{in}}_{\leqslant \mathcal{M}_\gamma}, \Xi^{\text{out}}_{\leqslant \mathcal{M}_\gamma}\big)\big|\big]. \end{aligned} \tag{5.B.14}$$

We use the notation $g \lesssim h$ if $g = O(h)$. We integrate over the locations and marks of the vertices in $\Xi^{\text{in}}_{\leqslant \mathcal{M}_\gamma} \cup \Xi^{\text{out}}_{\leqslant \mathcal{M}_\gamma}$ by writing $z(x) = \|x - \partial\mathcal{B}_{kM_{\text{in}}}\|$, and upper bound the connectivity function p in (5.1.3) to obtain

$$\begin{aligned} &\mathbb{E}\big[\big|\mathcal{E}\big(\Xi^{\text{in}}_{\leqslant \mathcal{M}_\gamma}, \Xi^{\text{out}}_{\leqslant \mathcal{M}_\gamma}\big)\big|\big] \\ &\lesssim \int_{x_u:\|x_u\|\leqslant r_k}\int_{x_v:\|x_v\|\geqslant r_k}\int_{w_u=1}^{f_\gamma(z(x_u))}\int_{w_v=1}^{f_\gamma(z(x_v))} \\ &\qquad \frac{(\kappa_{1,\sigma}(w_u,w_v))^\alpha}{\|x_u-x_v\|^{\alpha d}} w_u^{-\tau}w_v^{-\tau}\,\mathrm{d}w_v\mathrm{d}w_u\mathrm{d}x_v\mathrm{d}x_u. \end{aligned} \tag{5.B.15}$$

We analyse the double integral over the marks. Define

$$g_1(z_1,z_2) := \int_{w_1=1}^{f_\gamma(z_1)}\int_{w_2=1}^{f_\gamma(z_2)} \kappa_{1,\sigma}(w_1,w_2)w_1^{-\tau}w_2^{-\tau}\mathrm{d}w_1\mathrm{d}w_2.$$

Using the definition and symmetry of $\kappa_{1,\sigma}$ in (1.3.5), we reparametrise by $w \leqslant \tilde{w}$ and using that f_γ is increasing

$$\begin{aligned} g_1(z_1,z_2) &\lesssim \int_{w=1}^{f_\gamma(z_1\wedge z_2)}\int_{\widetilde{w}=w}^{f_\gamma(z_1\vee z_2)} \kappa_{1,\sigma}(w,\widetilde{w})(w\widetilde{w})^{-\tau}\mathrm{d}\widetilde{w}\mathrm{d}w \\ &= \int_{w=1}^{f_\gamma(z_1\wedge z_2)} w^{\sigma\alpha-\tau}\int_{\widetilde{w}=w}^{f_\gamma(z_1\vee z_2)} \widetilde{w}^{\alpha-\tau}\mathrm{d}\widetilde{w}\mathrm{d}w. \end{aligned}$$

The definition f_γ in (5.6.9) undergoes a change at $z = r_k$. When integrating, we have nine cases depending on whether the exponents are below, equal to, or above -1 each. This yields for $z_2 \leqslant \min\{z_1, r_k\}$ (so $f_\gamma(z_2) = 1 \vee (z_2/C_\beta)^{\gamma d}$) that $g_1(z_1, z_2) \lesssim g(z_1, z_2)$, with

$$g(z_1, z_2) := \begin{cases} f_\gamma(z_1)^{\alpha-(\tau-1)} z_2^{\gamma d(\sigma\alpha-(\tau-1))}, & \text{if } \alpha > \tau-1, \sigma\alpha > \tau-1, \\ f_\gamma(z_1)^{\alpha-(\tau-1)} \log(z_2), & \text{if } \alpha > \tau-1, \sigma\alpha = \tau-1, \\ f_\gamma(z_1)^{\alpha-(\tau-1)}, & \text{if } \alpha > \tau-1, \sigma\alpha < \tau-1, \\ \big(1+\log\big(\frac{f_\gamma(z_1)}{f_\gamma(z_2)}\big)\big) z_2^{\gamma d((\sigma+1)\alpha-2(\tau-1))}, & \text{if } \alpha = \tau-1, \sigma\alpha > \tau-1, \\ \log(f_\gamma(z_1)) \log(z_2), & \text{if } \alpha = \tau-1, \sigma\alpha = \tau-1, \\ \log(f_\gamma(z_1)), & \text{if } \alpha = \tau-1, \sigma\alpha < \tau-1, \\ z_2^{\gamma d((\sigma+1)\alpha-2(\tau-1))}, & \text{if } \alpha < \tau-1, (\sigma+1)\alpha > 2(\tau-1), \\ \log(z_2), & \text{if } \alpha < \tau-1, (\sigma+1)\alpha = 2(\tau-1), \\ 1, & \text{if } \alpha < \tau-1, (\sigma+1)\alpha < 2(\tau-1). \end{cases} \tag{5.B.16}$$

Returning to (5.B.15), we use $g(z_1, z_2)$ to bound the inner two integrals from above. We take a case-distinction on whether the vertex u (inside) or v (outside) is closer to the boundary $\partial B_{kM_{\text{in}}}$. Then we obtain (since $f_\gamma(x) = 1$ when $z(x) \leqslant C_\beta$ by (5.6.9)),

$$\begin{aligned} &\mathbb{E}\big[\big|\mathcal{E}\big(\Xi^{\text{in}}_{\leqslant \mathcal{M}_\gamma}, \Xi^{\text{out}}_{\leqslant \mathcal{M}_\gamma}\big)\big|\big] \\ &\quad\lesssim \int_{x_u: \|x_u\| \leqslant r_k - C_\beta} \int_{x_v: \|x_v\| > 2r_k} \|x_u - x_v\|^{-\alpha d} g(z(x_u), z(x_v)) dx_v dx_u \\ &\qquad+ \int_{x_u: \|x_u\| \leqslant r_k - C_\beta} \int_{x_v: z(x_v) \leqslant z(x_u)} \|x_u - x_v\|^{-\alpha d} g(z(x_u), z(x_v)) dx_v dx_u \\ &\qquad+ \int_{x_v: C_\beta \leqslant \|x_v\| - r_k \leqslant r_k} \int_{x_u: z(x_u) \leqslant z(x_v)} \|x_u - x_v\|^{-\alpha d} g(z(x_u), z(x_v)) dx_v dx_u \\ &\quad=: I_1 + I_{2a} + I_{2b} =: I_1 + I_2. \end{aligned} \tag{5.B.17}$$

To evaluate the integrals we change variables. For I_1 we use $z(x_u) \leqslant r_k$, and g is increasing in both its arguments, and that $\|x_u - x_v\| \geqslant |\|x_v\| - r_k| \geqslant z(x_v) := t$ and use polar coordinates in the second row below.

Then $\|x_v\| = t + r_k$. Thus, since there are $\Theta((t+r_k)^{d-1})$ points outside at distance t from $\partial\mathcal{B}_{kM_{\text{in}}}$,

$$\begin{aligned} I_1 &\lesssim \int_{x_u:\|x_u\|\leqslant r_k - C_\beta} \int_{x_v:\|x_v\|>2r_k} |\|x_v\| - r_k|^{-\alpha d} g(r_k, z(x_v)) dx_v dx_u \\ &\lesssim r_k^d \int_{t>r_k} (t+r_k)^{d-1} t^{-\alpha d} g(r_k, t) dt \lesssim r_k^d \int_{t\geqslant r_k} t^{-d(\alpha-1)-1} g(r_k, t) dt. \end{aligned} \tag{5.B.18}$$

Before substituting the definition of g into the bound, we recall that $f_\gamma(t) = C_\beta^{-\gamma d} t^d r_k^{-d(1-\gamma)}$ for $t > r_k$ by (5.6.9). The following elementary integration inequalities will be helpful soon:

$$\begin{aligned} r_k^d &\int_{t\geqslant r_k} t^{-d(\alpha-1)-1} f_\gamma(t)^{\alpha-(\tau-1)} dt \\ &\lesssim r_k^{d-d(1-\gamma)(\alpha-(\tau-1))} \int_{t\geqslant r_k} t^{-d(\tau-2)-1} dt \\ &\lesssim r_k^{d(2-\alpha+\gamma(\alpha-(\tau-1)))}, \end{aligned} \tag{5.B.19}$$

$$r_k^d \int_{t\geqslant r_k} t^{-d(\alpha-1)-1} dt \lesssim r_k^{d(2-\alpha)}. \tag{5.B.20}$$

Substituting the definition of g in (5.B.16) into (5.B.18), since $r_k < t$, we must set $z_2 = r_k$ in g in (5.B.16). We obtain then by elementary integration on (5.B.18) and the bounds in (5.B.19)-(5.B.20) that:

$$I_1 \lesssim \begin{cases} r_k^{d(2-\alpha+\gamma((\sigma+1)\alpha-2(\tau-1)))}, & \text{if } \alpha > \tau-1, \sigma\alpha > \tau-1, \\ r_k^{d(2-\alpha+\gamma(\alpha-(\tau-1)))} \log(r_k), & \text{if } \alpha > \tau-1, \sigma\alpha = \tau-1, \\ r_k^{d(2-\alpha+\gamma(\alpha-(\tau-1)))}, & \text{if } \alpha > \tau-1, \sigma\alpha < \tau-1, \\ r_k^{d(2-\alpha+\gamma((\sigma+1)\alpha-2(\tau-1))}, & \text{if } \alpha = \tau-1, \sigma\alpha > \tau-1, \\ r_k^{d(2-\alpha)} \log^2(r_k), & \text{if } \alpha = \tau-1, \sigma\alpha = \tau-1, \\ r_k^{d(2-\alpha)} \log(r_k), & \text{if } \alpha = \tau-1, \sigma\alpha < \tau-1, \\ r_k^{d(2-\alpha+\gamma((\sigma+1)\alpha-2(\tau-1)))}, & \text{if } \alpha < \tau-1, (\sigma+1)\alpha > 2(\tau-1), \\ r_k^{d(2-\alpha)} \log(r_k), & \text{if } \alpha < \tau-1, (\sigma+1)\alpha = 2(\tau-1), \\ r_k^{d(2-\alpha)}, & \text{if } \alpha < \tau-1, (\sigma+1)\alpha < 2(\tau-1). \end{cases} \tag{5.B.21}$$

We turn to I_{2a} in (5.B.17), handling the case when the outside vertex v is closer to the boundary $\partial\mathcal{B}_{kM_{\text{in}}}$ than u, implying $z(x_v) \leqslant z(x_u)$. We reparametrise this integral based on the distance z_u from $\partial\mathcal{B}_{kM_{in}}$ of the inside vertex u. Indeed, when $z_u \in [C_\beta, r_k]$ then $\|x_u\| = r_k - z_u$. Since v is closer, we must also have that $z_v = \|x_v\| - r_k \in [C_\beta, z_u]$, and hence $t := \|x_u - x_v\| \in [z_u + C_\beta, 2r_k]$. Hence,

$$I_{2a} \lesssim \int_{z_u=C_\beta}^{r_k} \int_{x_u: r_k - \|x_u\| = z_u} \int_{t=z_u+C_\beta}^{2r_k} \int_{z_v=C_\beta}^{z_u} \int_{\substack{x_v: z(x_v)=z_v,\\ \|x_u-x_v\|=t}} t^{-\alpha d} g(z_u, z_v) dx_v dz_v dt dx_u dz_u.$$

The integrand does not depend on x_v anymore, hence the most inside integral, over x_v, can be bounded from above by maximizing the Lebesgue measure of where x_v may fall: x_v has distance t from x_u and distance z_v from the boundary. Some geometry shows that x_v is then on the intersection of two spheres with radii t and $r_k + z_v$, respectively, with Lebesgue measure then at most $\Theta(t^{d-2})$. We can also integrate over all the potential locations x_u, giving a factor $\Theta((r_k - z_u)^{d-1})$, so we obtain

$$\begin{aligned} I_{2a} &\lesssim \int_{z_u=C_\beta}^{r_k} (r_k - z_u)^{d-1} \int_{t=z_u+C_\beta}^{2r_k} t^{d-2-\alpha d} \int_{z_v=C_\beta}^{z_u} g(z_u, z_v) dz_v dt dz_u \\ &\lesssim \int_{z_u=C_\beta}^{r_k} (r_k - z_u)^{d-1} z_u^{-d(\alpha-1)-1} \int_{z=C_\beta}^{z_u} g(z_u, z_v) dz_v dz_u, \qquad (5.B.22) \end{aligned}$$

where we integrated over t to obtain the second row. Treating I_{2b} in (5.B.17) is very similar, but now we reparametrise the integral based on the distance z_v of v from the boundary and the distance $t = \|x_u - x_v\|$. We obtain

$$\begin{aligned} I_{2b} &\lesssim \int_{z_v=C_\beta}^{r_k} \int_{x_v: \|x_v\| - r_k = z_v} \int_{t=z_v+C_\beta}^{3r_k} \int_{z_u=C_\beta}^{z_v} \int_{\substack{x_u: z_u = z(x_u),\\ \|x_u-x_v\|=t}} t^{-\alpha d} g(z_u, z_v) dx_u dz_u dt dx_v dz_v \\ &\lesssim \int_{z_v=C_\beta}^{r_k} (r_k + z_v)^{d-1} z_v^{-d(\alpha-1)-1} \int_{z_u=C_\beta}^{z_v} g(z_u, z_v) dz_u dz_v. \end{aligned}$$

This bound dominates the bound on I_{2a} in (5.B.22). Applying the binomial theorem on $(r_k + z_v)^{d-1}$, we obtain

$$I_2 = I_{2a} + I_{2b} \lesssim \sum_{j=0}^{d-1} r_k^j \int_{z_v=C_\beta}^{r_k} z_v^{d(2-\alpha)-2-j} \int_{z_u=C_\beta}^{z_v} g(z_u, z_v) dz_u dz_v. \qquad (5.B.23)$$

We evaluate the inner integral using the definition of g in (5.B.16), and since $z_u \leqslant z_v$ we set $z_1 = z_v, z_2 = z_u$ in (5.B.16), and $f_\gamma(z) = (z/C_\beta)^{\gamma d}$, and obtain the nine cases:

$$\int_{z_u=C_\beta}^{z_v} g(z_u, z_v) dz_u \lesssim \begin{cases} z_v^{\gamma d((\sigma+1)\alpha-2(\tau-1))+1}, & \text{if } \alpha > \tau-1, \sigma\alpha > \tau-1, \\ z_v^{\gamma d(\alpha-(\tau-1))+1} \log(z_v), & \text{if } \alpha > \tau-1, \sigma\alpha = \tau-1, \\ z_v^{\gamma d(\alpha-(\tau-1))+1}, & \text{if } \alpha > \tau-1, \sigma\alpha < \tau-1, \\ z_v^{\gamma d((\sigma+1)\alpha-2(\tau-1))+1}, & \text{if } \alpha = \tau-1, \sigma\alpha > \tau-1, \\ z_v \log^2(z_v), & \text{if } \alpha = \tau-1, \sigma\alpha = \tau-1, \\ z_v \log(z_v), & \text{if } \alpha = \tau-1, \sigma\alpha < \tau-1, \\ z_v^{\gamma d((\sigma+1)\alpha-2(\tau-1))+1}, & \text{if } \alpha < \tau-1, (\sigma+1)\alpha > 2(\tau-1), \\ z_v \log(z_v), & \text{if } \alpha < \tau-1, (\sigma+1)\alpha = 2(\tau-1), \\ z_v, & \text{if } \alpha < \tau-1, (\sigma+1)\alpha < 2(\tau-1). \end{cases} \tag{5.B.24}$$

For $\mathcal{Y} = \{0, \alpha-(\tau-1), (\sigma+1)\alpha-2(\tau-1)\}$, recall that $m(\mathcal{Y})$ counts the multiplicity of the maximum in $\mathcal{Y}$. Substituting (5.B.24) into (5.B.23), the nine cases can be summarized as obtaining the integrand of

$$z_v^{d(2-\alpha+\gamma\max(\mathcal{Y}))-1-j} \log^{m(\mathcal{Y})-1}(z_v),$$

so that following the similar reasoning as from (5.B.11) to (5.B.12),

$$I_2 \lesssim \begin{cases} r_k^{d(2-\alpha+\gamma\max(\mathcal{Y}))} \log^{m(\mathcal{Y})-1}(r_k), & \text{if } d(2-\alpha+\gamma\max(\mathcal{Y}))) > d-1, \\ r_k^{d(2-\alpha+\gamma\max(\mathcal{Y}))} \log^{m(\mathcal{Y})}(r_k), & \text{if } d(2-\alpha+\gamma\max(\mathcal{Y}))) = d-1, \\ r_k^{d-1}, & \text{if } d(2-\alpha+\gamma\max(\mathcal{Y}))) < d-1, \end{cases}$$

where the second bound follows from similar reasoning as in (5.B.11) leading to (5.B.12). The presence of a $(d-1)$ term and the additional log-factors ensure that the bound on I_2 dominates the bound on I_1 in (5.B.21). Recalling that $I_1 + I_2$ dominates the expected number of edges below the γ-suppressed profile from (5.B.17), this yields by (5.B.14) and $\zeta_{nn} = (d-1)/d$ that

$$\mathbb{E}\Big[\big|\mathcal{E}\big(\Xi^{in}_{\leqslant \mathcal{M}_\gamma \setminus \mathcal{R}_{in}}, \Xi^{out}_{\leqslant \mathcal{M}_\gamma \setminus \mathcal{R}_{out}}\big)\big| \,\Big|\, \Xi_{\mathcal{R}_{in}} \cup \Xi_{\mathcal{R}_{out}}, \mathcal{A}_{regular}(\eta)\Big] \lesssim \begin{cases} r_k^{d(2-\alpha+\gamma\max(\mathcal{Y}))} \log^{m(\mathcal{Y})-1}(r_k), & \text{if } 2-\alpha+\gamma\max(\mathcal{Y})) > \zeta_{nn}, \\ r_k^{d(2-\alpha+\gamma\max(\mathcal{Y}))} \log^{m(\mathcal{Y})}(r_k), & \text{if } 2-\alpha+\gamma\max(\mathcal{Y})) = \zeta_{nn}, \\ r_k^{d-1}, & \text{if } 2-\alpha+\gamma\max(\mathcal{Y})) < \zeta_{nn}. \end{cases}$$

To obtain bounds on the other three expectations in (5.B.13), one can replace the integrals over the marks in (5.B.15) by summing over the mark intervals I_j defined in (5.6.15): the upper bounds on the number of vertices using the definitions of $\mathcal{A}^{(k,in)}_{regular}(\eta)$ and $\mathcal{A}^{(k,out)}_{regular}(\eta)$ in (5.6.17) ensure that the total number of points in each interval only differs from its expectation by a constant factor. Then one can use an upper bound on the mark of each vertex in I^{loc}_j in (5.6.16), given by $I^{loc}_j = [w_\eta(kc_{loc})/2^j, w_\eta(kc_{loc})/2^{j-1})$, and thus also the mark is at most a factor two larger than the mark of a typical vertex in I^{loc}_j. Lastly, the distance between vertices in $\mathcal{R}_{in}$ and outside $\mathcal{B}_{kM_{in}}$ (but within distance r_k of $\mathcal{B}_{kM_{in}}$) can be bounded from below by $r_k/2$ by Lemma 5.A.1, and analogously we can bound the distance between vertices in $\mathcal{R}_{out}$ and vertices inside $\mathcal{B}_{kM_{in}}$. We leave it to the reader to fill in the details. □

5.C AUXILIARY PROOFS

It remains to prove Claim 5.6.5.

Proof of Claim 5.6.5. We first assume $\alpha < \infty$. We set $\gamma := \gamma_{\neg nn}$ defined in (5.6.13) in Lemmas 5.6.3 and 5.6.4, and compare the exponents of r_k^d on the right-hand sides of (5.6.29) and (5.6.30), respectively. We will show that the three cases there can be 'merged' by taking the exponent of $r_k^d = \Theta(k)$ to be $\max(\mathcal{Z})$, and the exponent of the log-factors in (5.6.29) and (5.6.30) is then at most $m_{\mathcal{Z}} - 1$. We distinguish whether $(\star)$ in (5.6.13) holds with equality or inequality. Let $\mathcal{Y} = \{0, \alpha-(\tau-1), (\sigma+1)\alpha - 2(\tau-1)\}$ as defined above (5.6.13) and $\mathcal{Z} = \{\zeta_{ll}, \zeta_{lh}, \zeta_{hh}, \zeta_{nn}\}$ as usual.

Case 1A. Assume $1-\gamma_{\neg nn}(\tau-1) > 2-\alpha+\gamma_{\neg nn}\max(\mathcal{Y})$, *and* $\alpha < \infty$. This, by (5.6.13) implies that $\gamma_{\neg nn} = 1/(\sigma+1)$. Considering that $\mathcal{Y} =$

$\{0, \alpha-(\tau-1), (\sigma+1)\alpha-2(\tau-1)\}$, the assumed inequality of Case 1A turns then into the following three inequalities after elementary rearrangements:

$$\begin{aligned} (\sigma+1)(\alpha-1) &> \tau-1, \\ \alpha &< (\sigma+1)(\alpha-1), \\ (\sigma+1)\alpha-(\tau-1) &< (\sigma+1)(\alpha-1). \end{aligned} \tag{5.C.1}$$

Solving the last inequality yields $\sigma < \tau-2$, which readily implies $\zeta_{hh} < 0$ by its definition in (4.2.4). The second inequality implies that $\alpha > 1+1/\sigma > 1+1/(\tau-2)$, which is equivalent to $\zeta_{lh} < 0$. We are left with showing that $\zeta_{ll} = 2-\alpha < 0$. The first inequality, and $(\sigma+1)/(\tau-1) < 1$ (coming from the third inequality), yield

$$\frac{1}{\alpha-1} < \frac{\sigma+1}{\tau-1} < 1,$$

which is equivalent to $\alpha > 2$. Summarizing, we obtained that under the assumption of Case 1A, $\max\{\zeta_{ll}, \zeta_{lh}, \zeta_{hh}\} < 0 \leqslant \zeta_{nn}$, hence $\max(\mathcal{Z}) = \zeta_{nn}$ and $\mathfrak{m}(\mathcal{Z}) = 1$. We also check that in this case the exponents of r_k^d in (5.6.29) and (5.6.30) are respectively at most $\zeta_{nn} = (d-1)/d \geqslant 0$. This is true since

$$2-\alpha+\gamma_{\neg nn}\max(\mathcal{Y}) < 1-\gamma_{\neg nn}(\tau-1) = 1-\frac{\tau-1}{\sigma+1} < 0,$$

where the first inequality follows by the assumption of Case 1A, and the second one by the last inequality in (5.C.1). Hence, under Case 1A, the third case holds in (5.6.29) and (5.6.30), the exponent of r_k^d is $\zeta_{nn} = \max(\mathcal{Z})$ in this case, and there are no logarithmic factors, which is also reflected in $\mathfrak{m}(\mathcal{Z})-1 = 0$. Hence, both inequalities in the statement of (5.6.33) hold in Case 1A.

Case 1B) Assume $\zeta_{\neg nn} = 1-\gamma_{\neg nn}(\tau-1) = 2-\alpha+\gamma_{\neg nn}\max(\mathcal{Y})$ *and* $\alpha < \infty$. The exponents of r_k^d in (5.6.29) and (5.6.30) match by assumption, and the three cases can be summarized as this exponent being $\max\{\zeta_{\neg nn}, \zeta_{nn}\}$. By the equality assumption for Case 1B, we compute that

$$\gamma_{\neg nn} = \frac{\alpha-1}{\max(\mathcal{Y})+\tau-1} = \min\left\{\frac{\alpha-1}{\tau-1}, \frac{\alpha-1}{\alpha}, \frac{\alpha-1}{(\sigma+1)\alpha-(\tau-1)}\right\}. \tag{5.C.2}$$

Since $\gamma_{\neg nn} \leqslant 1/(\sigma+1)$ by (5.6.13), it is elementary to verify that the third term may be minimal only when $\tau \leqslant 2+\sigma$.

Assume $\tau \leqslant 2+\sigma$. Using (5.C.2), (4.2.1), (4.2.2), (4.2.4), it is elementary to verify that

$$\begin{aligned}\zeta_{\neg nn} &= 1-\min\Big\{\frac{\alpha-1}{\tau-1}, \frac{\alpha-1}{\alpha}, \frac{\alpha-1}{(\sigma+1)\alpha-(\tau-1)}\Big\}(\tau-1)\\ &= \max\{\zeta_{ll}, \zeta_{lh}, \zeta_{hh}\}.\end{aligned} \tag{5.C.3}$$

It immediately follows that $\max\{\zeta_{\neg nn}, \zeta_{nn}\} = \max(\mathcal{Z})$. Hence, the exponents of r_k^d in (5.6.29) and (5.6.30) equal to $\max(\mathcal{Z})$, which is what we aimed for in (5.6.33). We now treat the exponent of the logarithm in (5.6.30) and in (5.6.29). Recall that $\mathfrak{m}$ in (5.1.5) denotes the multiplicity of the maximum of a set. From (5.C.2) and (5.C.3) it follows that $\mathfrak{m}(\mathcal{Y}) = \mathfrak{m}(\{\zeta_{ll}, \zeta_{lh}, \zeta_{hh}\})$. Hence, using that $\zeta_{\neg nn} = \max\{\zeta_{ll}, \zeta_{lh}, \zeta_{hh}\}$,

$$\begin{aligned}\mathfrak{m}(\mathcal{Z}) &= \mathfrak{m}(\{\zeta_{ll}, \zeta_{lh}, \zeta_{hh}, \zeta_{nn}\})\\ &= \mathfrak{m}(\mathcal{Y})\mathbb{1}_{\{\zeta_{\neg nn}>\zeta_{nn}\}} + (\mathfrak{m}(\mathcal{Y})+1)\mathbb{1}_{\{\zeta_{\neg nn}=\zeta_{nn}\}} + \mathbb{1}_{\{\zeta_{\neg nn}<\zeta_{nn}\}}\end{aligned} \tag{5.C.4}$$

The relations between ζ_{nn} and $1-\gamma_{\neg nn}(\tau-1)$ on the one hand, and ζ_{nn} and $2-\alpha+\gamma_{\neg nn}\max(\mathcal{Y})$ on the other hand, are the same by our equality assumption in Case 1B. This confirms that the log-factor in (5.6.29) is at most the log-factor in (5.6.30), showing therefore the first inequality of (5.6.33), and that the log-factor in (5.6.30) equals what we aimed for in the second inequality of (5.6.33), thereby finishing Case 1B when $\tau \leqslant 2+\sigma$.

Assume $\tau > 2+\sigma$. The minimum in (5.C.2) is never attained at the third term (and equivalently the maximum in $\mathcal{Y}$ is never attained at $(\sigma+1)\alpha-2(\tau-1)$). Similarly to (5.C.3), we obtain that $\zeta_{\neg nn} = \max\{\zeta_{ll}, \zeta_{lh}\}$. Moreover, $\zeta_{hh} < 0 \leqslant \zeta_{nn}$ by the formula for ζ_{hh} in (4.2.4), yielding $\max(\mathcal{Z}) = \max\{\zeta_{ll}, \zeta_{lh}, \zeta_{nn}\} = \max\{\zeta_{\neg nn}, \zeta_{nn}\}$. Hence, the exponents of r_k^d in (5.6.29) and (5.6.30) equal to $\max(\mathcal{Z})$, which is what we aimed for in (5.6.33). For the log-factor we argue similarly as in (5.C.4) and below, with the additional restriction that the maximum in $\mathcal{Y}$ is never attained at $(\sigma+1)\alpha-2(\tau-1)$, and the maximum in $\mathcal{Z}$ is never attained at ζ_{hh}.

Case 2. Assume $\alpha = \infty$. Since $\zeta_{ll} = 2-\alpha = -\infty$, $\zeta_{lh} = -(\tau-2) < 0$ ($\tau > 2$ by assumption in Lemma 5.6.2), and $\zeta_{nn} = (d-1)/d \geqslant 0$, we obtain $\max\{\zeta_{ll}, \zeta_{lh}, \zeta_{hh}, \zeta_{nn}\} = \max\{\zeta_{hh}, \zeta_{nn}\} = \max\{1-\gamma_{hh}(\tau-1), \zeta_{nn}\}$ by the formula of ζ_{hh} in (4.2.4). Since $\gamma_{hh} = \gamma_{\neg nn}$ when $\alpha = \infty$, it follows that the exponents of r_k^d and $\log(r_k)$ in (5.6.29) and (5.6.34) coincide. This finishes the proof. □

Lastly, we state a Poisson concentration bound (without proof) that we often rely on in the chapter.

Lemma 5.C.1 (Poisson bound [186]). *For* $x > 1$,

$$\mathbb{P}\big(\mathrm{Poi}(\lambda) \geqslant x\lambda\big) \leqslant \exp(-\lambda(1 + x\log(x) - x)),$$

and for $x < 1$

$$\mathbb{P}\big(\mathrm{Poi}(\lambda) \leqslant x\lambda\big) \leqslant \exp(-\lambda(1 - x - x\log(1/x)).$$

6 THE NEAREST-NEIGHBOUR REGIME

Based on [150]:
Cluster-size decay in supercritical long-range percolation,
J. Jorritsma, J. Komjáthy, D. Mitsche,
Preprint arXiv:2303.00724, 2023.

6.1 MODEL AND MAIN RESULTS

In the previous chapter and the accompanying paper [151], we investigate(d) the influence of inhomogeneity in the graph on the cluster-size distribution in i-KSRGS, see Section 1.3, where both the degree distribution and the edge-length distribution have heavy tails. We show(ed) that if the tails are "heavy enough", that $\mathbb{P}(k \leqslant |\mathcal{C}(0)| < \infty)$ decays stretched exponentially with exponent at least $(d-1)/d$. We leave the part of the phase diagram *open* where the model parameters are such that the conjectured exponent in (4.1.1) stays $(d-1)/d$.

In particular, the paper [151] also considers long-range percolation (LRP) [4, 216] (see Theorem 6.1.3 below for the result), and leaves open a region of the phase diagram of cluster-size decay. This missing region for LRP is the main focus in this chapter.

Definition 6.1.1 (Long-range percolation (LRP))**.** *Fix constants* $d \in \mathbb{N}, \alpha > 1$, $p \in (0,1)$*, and* $\beta > 0$*. We consider the random graph* $\mathcal{G}_\infty = (V(\mathcal{G}_\infty), E(\mathcal{G}_\infty))$ *with* $V(\mathcal{G}_\infty) = \mathbb{Z}^d$ *such that each edge* $\{x,y\}$ *is included in* $E(\mathcal{G}_\infty)$*, independently of all the other edges, with probability*

$$p(\|x-y\|) := p \cdot \left(\min\left\{1, \frac{\beta}{\|x-y\|^d}\right\}\right)^\alpha, \tag{6.1.1}$$

where $\|\cdot\| := \|\cdot\|_2$ *denotes the (Euclidean) 2-norm throughout this chapter. We set* $\Lambda_n := \mathbb{Z}^d \cap [-n^{-1/d}/2, n^{1/d}/2)^d$ *and* $E_n := \{\{x,y\} \in E_\infty : \{x,y\} \subseteq \Lambda_n\}$*, and write* $\mathcal{G}_n := (\Lambda_n, E_n)$ *for the induced subgraph of* $\mathcal{G}_\infty$ *on* Λ_n*. We write* $\mathcal{C}(0)$ *and* $\mathcal{C}_n(0)$ *for the connected component containing the origin in* $\mathcal{G}_\infty$ *and* $\mathcal{G}_n$*, respectively.*

Throughout the chapter we assume that the parameters are such that the graph is supercritical, i.e.,

$$\mathbb{P}\big(|\mathcal{C}(0)| = \infty\big) > 0,$$

implying that there exists a unique infinite component almost surely [4, 96]. We write $\mathcal{C}_n^{(i)}$ for the i-th largest component in $\mathcal{G}_n$, and for $n = \infty$, $\mathcal{C}_\infty^{(1)}$ for the unique infinite component in $\mathcal{G}_\infty$. We will also assume that $\min\{p, p\beta^\alpha\} < 1$, so that the graph is not connected almost surely. We state our main result.

Theorem 6.1.2 (Second-largest component and cluster-size decay)**.** *Consider supercritical long-range percolation on* $\mathbb{Z}^d$ *for* $d \geqslant 2$ *and* $\alpha > 1 + 1/d$*. If* β *in* (6.1.1) *is sufficiently large (depending on* p, α, d*), or if* $\beta \geqslant 1$ *and* p *is sufficiently close to* 1*, then there exist constants* $A, \delta > 0$ *such that for all* n *sufficiently large,*

$$\mathbb{P}\big(\tfrac{1}{A}(\log(n))^{d/(d-1)} \leqslant |\mathcal{C}_n^{(2)}| \leqslant A(\log(n))^{d/(d-1)}\big) \geqslant 1 - n^{-\delta}.$$

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

$$-k^{-(d-1)/d} \log\big(\mathbb{P}\big(|\mathcal{C}_n(0)| \geqslant k, 0 \notin \mathcal{C}_n^{(1)}\big)\big) \in [1/A, A]. \tag{6.1.2}$$

Lastly, under the same assumptions,

$$\frac{|\mathcal{C}_n^{(1)}|}{n} \xrightarrow{\mathbb{P}} \mathbb{P}\big(|\mathcal{C}(0)| = \infty\big), \qquad \textit{as } n \to \infty. \tag{6.1.3}$$

Note that (6.1.2) allows for $n = \infty$. Our results can be extended to hold when $\beta < 1$ and $p\beta^\alpha$ are sufficiently close to one. In Remark 6.1 we discuss a further generalisation. Theorem 6.1.2 complements the result of [151] applied to long-range percolation that we state here for completeness.

Theorem 6.1.3 (Complementary result for $\alpha < 1 + 1/d$ [151])**.** *Consider supercritical long-range percolation on* $\mathbb{Z}^d$ *for* $d \geqslant 1$ *and* $\alpha < 1 + 1/d$*. There exists constants* $A, \delta > 0$ *such that for all* $\varepsilon > 0$ *and* n *sufficiently large,*

$$\mathbb{P}\big(\tfrac{1}{A}(\log(n))^{1/(2-\alpha)} \leqslant |\mathcal{C}_n^{(2)}| \leqslant A(\log(n))^{1/(2-\alpha-\varepsilon)}\big) \geqslant 1 - n^{-\delta}. \tag{6.1.4}$$

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

$$-k^{-(2-\alpha)} \log\big(\mathbb{P}\big(|\mathcal{C}_n(0)| \geqslant k, 0 \notin \mathcal{C}_n^{(1)}\big)\big) \in [1/A, Ak^\varepsilon]. \tag{6.1.5}$$

Lastly, under the same assumptions,

$$\frac{|\mathcal{C}_n^{(1)}|}{n} \xrightarrow{\mathbb{P}} \mathbb{P}\big(|\mathcal{C}(0)| = \infty\big), \qquad \textit{as } n \to \infty.$$

In Theorem 6.1.3 we do not require β or p sufficiently large, and also allow one-dimensional models: when $d = 1$, LRP is supercritical when $\alpha \leqslant 2 = 1 + 1/d$ and when p, β are sufficiently large [86, 216]. When $d = 1$ and $\alpha > 2$, LRP is subcritical for any $p, \beta > 0$ such that $p(1) < 1$ [216], so Theorems 6.1.2 and 6.1.3 together give a complete picture for the cluster-size decay for supercritical long-range percolation (under the additional assumption that β or p is sufficiently large when $\alpha > 1 + 1/d$). In [151], we also study the phase boundary $\alpha = (d-1)/d$. In that case the lower bounds (6.1.4) and (6.1.5) contain lower order correction factors, that we conjecture to be sharp. We omit further details here. To the extent of our knowledge, for LRP, the only related results regarding the distribution of smaller clusters in *supercritical* LRP is an upper bound on the second-largest component with unidentified exponent by Crawford and Sly [58] for $\alpha \in (1, 2)$ in dimension 1 and $\alpha \in (1, 1 + 2/d)$ in dimensions 2 and higher. For *(sub)critical* LRP with $\alpha \in (1, 2)$, a polynomial upper bound on $\mathbb{P}(|\mathcal{C}(0)| \geqslant n)$ is established in [138].

Before proceeding to the technical contributions, we remark that our results could be generalised to a more general class of random graph models on $\mathbb{Z}^d$. Theorem 6.1.2 extends to random graph models on $\mathbb{Z}^d$ with independent edges for any connectivity function that has a lighter tail than p in Theorem 6.1.2, provided that the probability of 'short-range' edges is still sufficiently large. In particular, it extends to spread-out percolation, in which two vertices within distance R are connected independently with probability p, or long-range percolation models in which the connection probability decays subpolynomially. We refrain from proving the result in this generality, since it would require many technically involved changes in the already technical Chapter 5.

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

$$\int_{x: x \in \mathbb{R}^d} \|x\| J(\|x\|) < \infty. \tag{6.1.6}$$

Then we have the following two cases:

1) If the connectivity function p is of the form

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

and there is an $\varepsilon > 0$ such that $J(x) > \varepsilon$ whenever $x < \varepsilon$, then (6.1.2)–(6.1.3) hold for all sufficiently large β depending on ε.

2) If the connectivity function is of the form

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

then (6.1.2)–(6.1.3) hold for all p sufficiently close to 1.

The integral in the condition (6.1.6) represents the order of the expected number of edges $\{x, y\}$ for which the line-segment (x, y) crosses a fixed box of volume one. If this number is finite, Theorem 6.1.2 holds in more generality. The connectivity function p from Definition 6.1.1 satisfies the integrability condition (6.1.6) if and only if $\alpha > 1 + 1/d$. We conjecture that the *upper* bounds in (6.1.2)–(6.1.2) remain valid if (6.1.6) is violated (but no longer match the lower bounds). However, this would require a different proof technique.

We state a proposition that contains the main technical contribution of this paper. Together with statements from Chapter 5, where we establish the relation between the second-largest component and the cluster-size decay for spatial random graph models more generally, this proposition will readily imply Theorem 6.1.2.

Proposition 6.1.4 (Second-largest component, upper bound)**.** *Consider supercritical long-range percolation on $\mathbb{Z}^d$ for $\alpha > 1 + 1/d$, $d \geqslant 2$. If β in (5.1.3) is sufficiently large (depending on p, α, d), or if $\beta \geqslant 1$ and p is sufficiently close to 1, then there exists a constant $A > 0$ such that for all k sufficiently large and for all n satisfying $n(\log(n))^{-2d/(d-1)} \geqslant k$:*

$$\mathbb{P}\big(|\mathcal{C}_n^{(2)}| \geqslant k\big) \leqslant n \log(n) \exp\big(-k^{(d-1)/d}\big).$$

We believe that Proposition 6.1.4, and hence Theorem 6.1.2, should hold for any values of β, p that lead to a supercritical graph: however, this would require non-trivial adaptations of our proof techniques. This proposition also can be extended to $\beta < 1$ such that $p\beta^\alpha$ is sufficiently close to 1.

6.1.1 *Idea of proof*

The proof of Proposition 6.1.4 relies on a careful first-moment analysis in which we count all possible candidates of isolated components of size at least k. The starting point is the classic isoperimetric inequality that says that any set $\mathcal{S}$ of at least k vertices has an edge-boundary of size $|\partial\mathcal{S}| = \Omega(k^{(d-1)/d})$. These edges need to be absent when $\mathcal{S}$ is a connected component (or simply component below), i.e., detached from the rest of the graph. The combinatorial difficulty arises when we account for all possible candidate components $\mathcal{S}$: the structure of $\mathcal{S}$ is more complex than for nearest-neighbour bond percolation in $\mathbb{Z}^d$, since $\mathcal{S}$ can be "delocalized" in space. The second difficulty arises in the finite box $\Lambda_n \subseteq \mathbb{Z}^d$, where we need to take boundary effects into account caused by possibly shared boundary of $\partial\mathcal{S}$ and $\partial\Lambda_n$.

To resolve these two complications, we distinguish two types of components: the first type consists of several "blocks" connected by long edges: each block is a connected subset of $\mathbb{Z}^d$ (with respect to nearest-neighbour relation in $\mathbb{Z}^d$). We consider each possible combination of blocks with fixed total outer edge-boundary size m, and give an upper bound on the probability that these blocks form a connected component by counting all possible spanning trees on these blocks. We show that the combinatorial factor arising from counting all potential components with boundary m is at most exponential in m. We then use the large value of β or p in our favor to prove that the probability that such a component is formed and *isolated* is sufficiently small.

The second type of potential component $\mathcal{S}$ contains a large block that has a large overlap with the boundary of Λ_n, and consequently $|\partial\mathcal{S}|$ inside Λ_n may be small. In this case we enumerate all such blocks and use an adapted isoperimetric inequality to still ensure that many edges need to be absent, thereby showing the right error probability.

Organization

In Section 6.2, we derive an intermediate upper bound for $\mathbb{P}\big(|\mathcal{C}_n^{(2)}| \geqslant k\big)$, defining the two types of components formally. Then, we state two lemmata and show that they imply Proposition 6.1.4. We prove the two lemmata in separate sections. In the last section we use the result of Proposition 6.1.4 to prove Theorem 6.1.2.

Notation

For $A \subseteq V_H$, we write $H[A]$ for the induced subgraph of H on vertices in A. Denote by $\mathbb{Z}^d_\infty$ the graph on the vertex set $\mathbb{Z}^d$ and an edge between

$x, y \in \mathbb{Z}^d$ if and only if $\|x-y\|_\infty = 1$. Similarly, let $\mathbb{Z}^d_1$ be the graph on the vertex set $\mathbb{Z}^d$ and an edge between $x, y \in \mathbb{Z}^d$ if and only if $\|x-y\|_1 = 1$. As already mentioned, we write $\|\cdot\| := \|\cdot\|_2$. For two sets $A, B \subseteq \mathbb{Z}^d$, denote by $\|A-B\|_p = \min\{\|x-y\|_p \mid x \in A, y \in B\}$. We say that a path $\pi = (v_1, v_2, v_3, \dots)$ is *self-avoiding* if its vertices are all distinct.

6.2 PRELIMINARIES AND SETUP

Throughout the rest of the paper, we assume that $d \geqslant 2$, $\alpha > 1 + 1/d$, and $n^{1/d} \in \mathbb{N}$. We will now formalise the concepts from the proof outline in Section 6.1.1 that eventually lead to two lemmas for each of the two described types of components. To ensure that the upcoming definitions naturally follow each other, we will postpone the (sometimes standard) proofs of intermediate claims to the appendix. We start with a definition to describe sets of Λ_n that form (subsets of) the second-largest component.

Definition 6.2.1 (Connected sets and blocks)**.** We call a set $A \subseteq \mathbb{Z}^d$ of vertices 1*-connected* or a *block*, if the graph $\mathbb{Z}^d_1[A]$ consists of a single connected component. We similarly define A being $*$*-connected* if the graph $\mathbb{Z}^d_\infty[A]$ consists of a single connected component. We write

$$\begin{aligned} \mathcal{A} &:= \big\{A \subseteq \Lambda_n \mid A \text{ is 1-connected}\big\}, \\ \mathcal{A}_* &:= \big\{A \subseteq \Lambda_n \mid A \text{ is } *\text{-connected}\big\}. \end{aligned} \tag{6.2.1}$$

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

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

$$\bar{A} = A \cup \{x \in \mathbb{Z}^d : x \text{ surrounded by } A\}. \tag{6.2.2}$$

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

We make a few comments. The closures of blocks will be used for the first type of components described in Section 6.1.1. Take now a block $A \subseteq \Lambda_n$. Then x can only be surrounded by A if $x \in \Lambda_n$ too, hence $A \subseteq \Lambda_n$ implies that $\bar{A} \subseteq \Lambda_n$.

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

Definition 6.2.2 (Block graph)**.** Let $A_1, \dots, A_b \in \mathcal{A}$ be a sequence of blocks, and consider a graph G on vertices $V_G \supseteq \cup_{i \leqslant b} A_i$. The *block graph* $\mathcal{H}_G((A_i)_{i \leqslant b}) = (V_{\mathcal{H}_G}, E_{\mathcal{H}_G})$ of G on blocks $A_1, \dots, A_b$ is defined as

$$V_{\mathcal{H}_G} := \{1, \dots, b\}, \qquad E_{\mathcal{H}_G} := \big\{\{i, j\} : A_i \leftrightarrow_G A_j\big\}.$$

In words, the vertices of each block are contracted to a single vertex in the block-graph, and two corresponding vertices for the blocks i and j are connected in the block graph if and only if there exists an edge in the original graph G between (some vertices in) the two blocks. We continue with a simple claims, with proof in the appendix on page 261.

Claim 6.2.3 (Unique block-decomposition of components)**.** *Let $\mathcal{C}$ be a finite component of a graph $\mathcal{G}$ with vertex set either Λ_n or $\mathbb{Z}^d$. Then $\mathcal{C}$ can be uniquely decomposed into a 1-disconnected sequence $(A_i)_{i \leqslant b}$ of blocks, with $b < \infty$, so that he block graph $\mathcal{H}_{\mathcal{G}}((A_i)_{i \leqslant b})$ is connected.*

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

Definition 6.2.4 (Boundaries)**.** Let $A \subseteq \Lambda_n$. We define the *exterior boundary* of A with respect to Λ_n and $\mathbb{Z}^d$, respectively, as

$$\begin{aligned} \partial_{\text{ext}} A &:= \{x \in \Lambda_n : \|x - A\|_1 = 1\}, \\ \widetilde{\partial}_{\text{ext}} A &:= \{x \in \mathbb{Z}^d : \|x - A\|_1 = 1\}. \end{aligned} \tag{6.2.3}$$

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

$$\begin{aligned} \partial_{\text{int}} A &:= \{x \in A : \|x - \partial_{\text{ext}} A\|_1 = 1\}, \\ \widetilde{\partial}_{\text{int}} A &:= \{x \in A : \|x - \widetilde{\partial}_{\text{ext}} A\|_1 = 1\}. \end{aligned} \tag{6.2.4}$$

If, in words, we mention the exterior, interior, or outer boundary of A then – unless explicitly specified differently – we mean with respect to Λ_n.

We mention that $\widetilde{\partial}_{\text{int}}\Lambda_n$ is the 'usual' vertex boundary of Λ_n. The boundary $\widetilde{\partial}_{\text{ext}}A$ may contain vertices outside Λ_n, and will be useful in the enumeration of subsets forming isolated components below. It may happen that a block A contains (many) vertices of $\widetilde{\partial}_{\text{int}}\Lambda_n$. On such regions, A may not have external boundary vertices, implying that there $\partial_{\text{int}}A$ is also empty. The next claim contains basic properties of blocks, their closures, and their boundaries, with proof in the Appendix on page 262.

Claim 6.2.5 (Blocks, their closures and their boundaries). *The following four statements hold:*

(i) For any block B, $\widetilde{\partial}_{\text{int}}\bar{B} \subseteq \widetilde{\partial}_{\text{int}}B$.

(ii) Let B_1, B_2 be 1-disconnected blocks such that $\bar{B}_1 \cap \bar{B}_2 \neq \emptyset$. Then either $\bar{B}_1 \subseteq \bar{B}_2$ or $\bar{B}_2 \subseteq \bar{B}_1$.

(iii) Let B_1, B_2 be 1-disconnected blocks such that $\bar{B}_1 \cap \bar{B}_2 = \emptyset$. Then $\bar{B}_1, \bar{B}_2$ are also 1-disconnected from each other.

(iv) For any block B, $\widetilde{\partial}_{\text{int}}\bar{B}$ and $\widetilde{\partial}_{\text{ext}}\bar{B}$ are $*$-connected.

(v) For any hole H of a block B, we have that $H = \bar{H}$, so $\partial_{\text{int}}H$ and $\partial_{\text{ext}}H$ are $*$-connected.

We point out that the fourth statement of the preceding claim is [76, Lemma 2.1]. The next claim shows that the sizes of the boundaries with respect to Λ_n and $\mathbb{Z}^d$ are of the same order, provided that the set is smaller than $3n/4$. Moreover, it contains an *isoperimetric inequality* that we extensively use below. The proof is given in the appendix on page 263.

Claim 6.2.6 (Boundary bounds and isoperimetry). *There exists $\delta > 0$ such that for all $A \subseteq \Lambda_n$ with $|A| \leqslant 3n/4$ or $A \cap \widetilde{\partial}_{\text{int}}\Lambda_n = \emptyset$,*

$$\begin{aligned} |\partial_{\text{int}}A| &\geqslant \delta|\widetilde{\partial}_{\text{int}}A| \overset{(\star)}{\geqslant} \delta|A|^{(d-1)/d}, \\ |\partial_{\text{ext}}A| &\geqslant \delta|\widetilde{\partial}_{\text{ext}}A| \overset{(\star)}{\geqslant} \delta|A|^{(d-1)/d}. \end{aligned} \tag{6.2.5}$$

The inequalities with $(\star)$ hold for any $A \subseteq \Lambda_n$ without conditions on A.

The next lemma is due to Peierls (its proof is given in the appendix on page 266) and is crucial for the enumeration of blocks that satisfy $A = \bar{A}$.

Lemma 6.2.7 (Peierls' argument). *There exists a constant $c_{\text{pei}} > 0$ such that for all $x \in \mathbb{Z}^d$ and $m \in \mathbb{N}$,*

$$|\{A \in \mathcal{A} : A \ni x, A = \bar{A}, |\widetilde{\partial}_{\text{int}}A| = m\}| \leqslant \exp(c_{\text{pei}}m). \tag{6.2.6}$$

We remark that the proof relies on the fact that $\widetilde{\partial}_{\text{int}}A$ is $*$-connected [76, Lemma 2.1] (which would not hold if A contained holes, or may not hold if one replaces $\widetilde{\partial}_{\text{int}}A$ by $\partial_{\text{int}}A$).

In (6.2.5), we would like to replace $\partial_{\text{int}}A$ by $\partial_{\text{int}}\bar{A}$ from (6.2.2) (enumeration of sets that are equal to their closure would allow us to use Peierls' argument). However, then the isoperimetric inequality (6.2.5) may not hold anymore if the total boundary size of the holes in A is too large compared to $\partial_{\text{int}}\bar{A}$, which could happen if $|\bar{A}| > 3n/4$. We define two types of blocks, based on the size of the closures of the blocks (that is, whether Claim 6.2.6 applies to $\bar{A}$ or not), i.e.,

$$\begin{aligned} \mathcal{A}_{\text{small}} &:= \{A \in \mathcal{A} : |\bar{A}| \leqslant 3n/4, \ \text{ and } \ \bar{A} = A\}, \\ \mathcal{A}_{\text{large}} &:= \big\{A \in \mathcal{A} : |\bar{A}| > 3n/4, \ \text{ and } \ |A| \leqslant n/2\big\}. \end{aligned} \tag{6.2.7}$$

For each of the sets we define an event, i.e.,

$$\mathcal{E}_1(b) := \left\{ \exists\, 1\text{-disconn. } (A_i)_{i \leqslant b} \in \mathcal{A}_{\text{small}} \;\middle|\; \begin{array}{c} (\cup_i \widetilde{\partial}_{\text{int}} A_i) \not\leftrightarrow_{\mathcal{G}_n} (\Lambda_n \setminus \cup_i A_i), \\ |\cup_i A_i| \geqslant k, \\ \mathcal{H}_{\mathcal{G}_n}((A_i)_{i\leqslant b}) \text{ connected} \end{array} \right\}, \tag{6.2.8}$$

$$\mathcal{E}_2 := \{\exists\, A \in \mathcal{A}_{\text{large}} : A \not\leftrightarrow_{\mathcal{G}_n} \partial_{\text{ext}}A\}. \tag{6.2.9}$$

The following deterministic claim holds for any graph on vertices in Λ_n. It shows that the union of these events contains the event $\{|\mathcal{C}_n^{(2)}| \geqslant k\}$. In particular, the proof reveals why we could restrict to sets with $A = \bar{A}$ in the definition of $\mathcal{A}_{\text{small}}$ in (6.2.7).

Claim 6.2.8. *Consider the graph $\mathcal{G}_n$ from Definition 6.1.1, with $\mathcal{C}_n^{(2)}$ the second largest component of $\mathcal{G}_n$. Then*

$$\big\{|\mathcal{C}_n^{(2)}| \geqslant k\big\} \subseteq \mathcal{E}_2 \cup \Big(\bigcup_{b=1}^{\lfloor n/2 \rfloor} \mathcal{E}_1(b) \Big).$$

Proof of Claim 6.2.8. Clearly $\big\{|\mathcal{C}_n^{(2)}| \geqslant k\big\} \subseteq \{\exists \text{ a component } \mathcal{C} \text{ of } \mathcal{G}_n : |\mathcal{C}| \in [k, \lfloor n/2 \rfloor]\}$. The size restriction $n/2$ is needed since otherwise $\mathcal{C}_n^{(2)}$ would be the largest component. Take any such $\mathcal{C}$. We use Claim 6.2.3 to first uniquely decompose $\mathcal{C}$ into 1-disconnected (hence disjoint) blocks $A_1, \ldots, A_b$ for some $b \geqslant 1$. Since $k \leqslant |\mathcal{C}| \leqslant n/2$, $|A_i| \leqslant n/2$ also holds for all $i \leqslant b$, and also $b \leqslant n/2$, and $\sum_{i \leqslant b} |A_i| \geqslant k$.

We distinguish two cases. Either (1) there is at least one block that is in $\mathcal{A}_{\text{large}}$ or (2) all the blocks are in $\mathcal{A} \setminus \mathcal{A}_{\text{large}}$. In the first case the event $\mathcal{E}_2$

in (6.2.9) holds: the block of $\mathcal{C}$ satisfying $\mathcal{A}_{\text{large}}$ (say A_1) is per definition 1-disconnected from the other blocks of $\mathcal{C}$, and since $\mathcal{C}$ is a component of $\mathcal{G}_n$, A_1 is $\mathcal{G}_n$-disconnected from its exterior boundary, hence $\mathcal{E}_2$ holds for $A := A_1$ in (6.2.9).

In case (2) all the blocks $(A_i)_{i\leqslant b}$ are in $\mathcal{A}\setminus\mathcal{A}_{\text{large}}$, and, since they form the component $\mathcal{C}$, the graph $\mathcal{G}_n$ spanned on $\cup_{i\leqslant b}A_i$ is connected, while their union is disconnected in $\mathcal{G}_n$ from the rest of the graph. Finally their disjointness and $|\mathcal{C}|\in[k,n/2]$ ensures that $|\cup_{i\leqslant b}A_i|\in[k,n/2]$. Formally we describe this event as (changing the sets to be denoted by B_i to avoid clash of notation later):

$$\widetilde{\mathcal{E}}_1(b):=\left\{\exists\text{1-disconn.}(B_i)_{i\leqslant b}\in\mathcal{A}\setminus\mathcal{A}_{\text{large}}\;\middle|\;\begin{array}{l}\mathcal{G}_n[\cup_{i\leqslant b}B_i]\text{ connected},\\ |\cup_{i\leqslant b}B_i|\in[k,n/2],\\ (\cup_{i\leqslant b}B_i)\not\leftrightarrow_{\mathcal{G}_n}(\Lambda_n\setminus\cup_{i\leqslant b}B_i)\end{array}\right\}.$$

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

$$\begin{aligned}&\left\{|\mathcal{C}_n^{(2)}|\geqslant k\right\}\\ &\subseteq\{\exists\text{ a component }\mathcal{C}\text{ of }\mathcal{G}_n:|\mathcal{C}|\in[k,n/2]\}\subseteq\mathcal{E}_2\cup\left(\cup_{b=1}^{\lfloor n/2\rfloor}\widetilde{\mathcal{E}}_1(b)\right).\end{aligned}$$

We will show that

$$\bigcup_{b=1}^{\lfloor n/2\rfloor}\widetilde{\mathcal{E}}_1(b)\subseteq\bigcup_{b'=1}^{\lfloor n/2\rfloor}\mathcal{E}_1(b'). \tag{6.2.10}$$

Take now any 1-disconnected blocks $(B_1,\dots,B_b)$ for which $\widetilde{\mathcal{E}}_1(b)$ holds. The conditions of Claim 6.2.5 are satisfied for any pair B_i,B_j with $i\neq j$, hence for each pair, $\bar{B}_i$ and $\bar{B}_j$ are either 1−disconnected disjoint sets, or one contains fully the other one. Choose now those sets in $\{\bar{B}_1,\dots,\bar{B}_b\}$ that are not contained in any other set in the same list. We then obtain an integer $b'\leqslant b$ and a 1-disconnected subset $\{\bar{B}_{i_1},\dots,\bar{B}_{i_{b'}}\}\subseteq\{\bar{B}_1,\dots,\bar{B}_b\}$ such that

$$\bigcup_{i=1}^{b}B_i\subseteq\bigcup_{i=1}^{b}\bar{B}_i=\bigcup_{j=1}^{b'}\bar{B}_{i_j}. \tag{6.2.11}$$

Let $B_1,\dots,B_b$ in $\mathcal{A}\setminus\mathcal{A}_{\text{large}}$ be an arbitrary sequence of 1-disconnected blocks satisfying $\widetilde{\mathcal{E}}_1(b)$, and assume without loss of generality that the indices $i_1,\dots,i_{b'}$ correspond to indices $1,\dots,b'$, and we may thus assume that the sets $\bar{B}_1,\dots,\bar{B}_{b'}$ satisfy (6.2.11). Since $(\bar{B}_1,\dots,\bar{B}_{b'})$ is a 1-

disconnected sequence of blocks that are equal to their own closure, (6.2.10) follows if

$$\{\mathcal{G}_n[\cup_{i\leqslant b}B_i] \text{ connected}\} \subseteq \{\mathcal{H}_{\mathcal{G}_n}((\bar{B}_i)_{i\leqslant b'}) \text{ connected}\}, \quad (6.2.12)$$

$$\{|\cup_{i\leqslant b} B_i| \in [k, n/2]\} \subseteq \{|\cup_{i\leqslant b} \bar{B}_i| \geqslant k\}, \quad (6.2.13)$$

$$\{(\cup_{i\leqslant b}B_i) \not\leftrightarrow_{\mathcal{G}_n} (\Lambda_n \setminus \cup_{i\leqslant b}\bar{B}_i)\} \subseteq \{(\cup_{i\leqslant b'}\widetilde{\partial}_{\text{int}}\bar{B}_i) \not\leftrightarrow_{\mathcal{G}_n} (\Lambda_n \setminus \cup_{i\leqslant b'}\bar{B}_i)\}, \quad (6.2.14)$$

since then all conditions for $\mathcal{E}_1(b')$ defined in (6.2.8) are satisfied by setting $A_i = \bar{B}_i$ for all $i \leqslant b'$.

For the first inclusion (6.2.12) we observe that

$$\begin{aligned}\big\{\mathcal{G}_n[\cup_{i\leqslant b}B_i] \text{ connected}\big\} &\subseteq \big\{\mathcal{H}_{\mathcal{G}_n}((B_i)_{i\leqslant b}) \text{ connected}\big\}\\ &\subseteq \big\{\mathcal{H}_{\mathcal{G}_n}((\bar{B}_i)_{i\leqslant b'}) \text{ connected}\big\},\end{aligned}$$

since the left-hand side was assumed in $\widetilde{\mathcal{E}}_1(b)$, the block graph being connected is a less demanding event than the actual spanned graph being connected, and the second containment follows since each set of edges in $\mathcal{G}_n$ that ensures that $\mathcal{H}_{\mathcal{G}_n}((B_i)_{i\leqslant b})$ is connected, also ensures that the block graph on the closures of $(B_i)_{i\leqslant b}$ is connected.

The second inclusion (6.2.13) is trivial since $\cup_{i\leqslant b}B_i \subseteq \cup_{i\leqslant b'}\bar{B}_i$ by (6.2.11). For the third inclusion (6.2.14) we have to argue that the set of edges that is excluded on the right-hand side is smaller than the set of excluded edges on the left-hand side. Clearly $(\Lambda_n \setminus \cup_{i\leqslant b}B_i) \supseteq (\Lambda_n \setminus \cup_{i\leqslant b'}\bar{B}_i)$, and by part (i) in Claim 6.2.5 it follows that $\partial_{\text{int}}\bar{B}_i \subseteq B_i$. □

We state two lemmas that together with Claim 6.2.8 prove Proposition 6.1.4.

Lemma 6.2.9 (Unlikely block graphs). *Let $\mathcal{G}_n$ be long-range percolation on Λ_n as in Definition 6.1.1 with $d \geqslant 2$, $\alpha > 1 + 1/d$. There exists a constant $c_{6.2.9} > 0$, such that for all $p \in (0, 1)$, there exists $\beta_\star = \beta_\star(p, d, \alpha) > 0$ such that for all $\beta \geqslant \beta_\star$, and k, n sufficiently large,*

$$\mathbb{P}\bigg(\bigcup_{b=1}^{\lfloor n/2 \rfloor} \mathcal{E}_1(b)\bigg) \leqslant n\log(n)\exp\Big(-c_{6.2.9}\log\big(\tfrac{1}{1-p}\big)\beta^{1/d}k^{(d-1)/d}\Big). \quad (6.2.15)$$

Moreover, there exists $p_d < 1$ such that $\beta_\star \leqslant 1$ for all $p \in (p_d, 1)$.

Lemma 6.2.10 (No large isolated component). *Let $\mathcal{G}_n$ be long-range percolation on Λ_n as in Definition 6.1.1 with $d \geqslant 2$, $\alpha > 1 + 1/d$. There exists a constant*

$c_{6.2.10} > 0$ *such that for all* $p \in (0,1)$, *there exists* $\beta_\star = \beta_\star(p,d,\alpha) > 0$ *so that for all* $\beta \geqslant \beta_\star$, *and* n *sufficiently large*

$$\mathbb{P}(\mathcal{E}_2) \leqslant \exp\Big(-c_{6.2.10}\log\big(\tfrac{1}{1-p}\big)\beta^{(d-1)/d^2}\frac{n^{(d-1)/d}}{\log^2(n)}\Big). \qquad (6.2.16)$$

Moreover, there exists $p_d < 1$ *such that* $\beta_\star \leqslant 1$ *for all* $p \in (p_d, 1)$.

We prove the two lemmata in the following sections.

6.3 SPANNING TREES ON BLOCK GRAPHS

We work towards proving Lemma 6.2.9. Recall the event $\mathcal{E}_1(b)$ from (6.2.8), and $\mathcal{A}_{\text{small}}$ for the blocks of Λ_n without holes and closure of size at most $3n/4$ from (6.2.7) on which there should be a connected block graph $\mathcal{H}_{\mathcal{G}_n}((A_i)_{i\leqslant b})$ (see Definition 6.2.2). By a union bound over all possible 1-disconnected sequences of blocks $(A_1,\dots,A_b) \subseteq \mathcal{A}_{\text{small}}$ whose total size is at least k, we obtain

$$\mathbb{P}(\mathcal{E}_1(b)) \leqslant \frac{1}{b!}\sum_{A_1}\cdots\sum_{A_b}\mathbb{P}\begin{pmatrix}\mathcal{H}_{\mathcal{G}_n}((A_i)_{i\leqslant b}) \text{ is connected,}\\ (\cup_{i\leqslant b}\widetilde{\partial}_{\text{int}}A_i) \nleftrightarrow_{\mathcal{G}_n} (\Lambda_n\setminus\cup_{i\leqslant b}A_i)\end{pmatrix},$$

where the factor $1/b!$ corrects for the permutations of $(A_1,\dots,A_b)$ yielding the same blocks, but ordered differently. Using the independence of edges in long-range percolation in Definition 6.1.1, we obtain

$$\begin{aligned}\mathbb{P}(\mathcal{E}_1(b)) \leqslant \frac{1}{b!}\sum_{A_1}\cdots\sum_{A_b}&\mathbb{P}\big(\mathcal{H}_{\mathcal{G}_n}((A_i)_{i\leqslant b}) \text{ is connected}\big)\\ &\cdot\mathbb{P}\big((\cup_{i\leqslant b}\widetilde{\partial}_{\text{int}}A_i) \nleftrightarrow_{\mathcal{G}_n} (\Lambda_n\setminus\cup_{i\leqslant b}A_i)\big).\end{aligned} \qquad (6.3.1)$$

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

Definition 6.3.1 (**f**-tree)**.** Let $\mathcal{F}_b$ be the set of vectors $\mathbf{f} = (f_1,\dots,f_b) \in \mathbb{N}_0^b$ satisfying $\sum_{i\in[b]} f_i = b-1$, $f_1+\dots+f_j \geqslant j$ for all $j \in [b-1]$. We call **f** the vector of *forward degrees*. A rooted labeled tree on b vertices is an **f**-tree if the root has label 1, and it has an outgoing edge to each of the vertices with labels $2,\dots,f_1+1$, vertex 2 has an outgoing edge to each of the vertices with labels $f_1+2,\dots,f_1+f_2+1$, and so on, the vertex with label j has an outgoing edge to each of the vertices with labels $2+\sum_{i=1}^{j-1} f_i, \dots, 1+\sum_{i=1}^{j} f_i$. If (i,j) is a directed edge in an $\mathbf{f}$-tree, then we say that i is the parent of j and j is the child of i. We say that the labeled block graph $\mathcal{H}_{\mathcal{G}_n}((A_i)_{i\leqslant b})$ is $\mathbf{f}$-*connected*, if it contains an $\mathbf{f}$-tree on its vertices $(1,2,\dots,b)$.

Given a forward-degree vector $\mathbf{f}$ and a labeled set of vertices, the $\mathbf{f}$-tree is uniquely determined. The construction ensures that the block with label b must be a leaf, i.e., it has forward degree $f_b = 0$, and its parent corresponds to the label of the last nonzero entry of $\mathbf{f}$. Further, given a tree T with labeling $\mathbf{f} \in \mathcal{F}_b$, upon removing the leaf with label b, we obtain a tree $T \setminus \{b\}$ on $\{1,\dots,b-1\}$ with a labeling in $\mathcal{F}_{b-1}$.

Further, an $\mathbf{f}$-tree always has vertex 1 as its root, and the forward neighbours of any vertex have consecutive labels. Hence, not all the b! labelings of a tree T are valid labelings, i.e., no vector $\mathbf{f} \in \mathcal{F}_b$ can be associated to some labelings. However, for a fixed tree T on a connected block graph $\mathcal{H}_{\mathcal{G}_n}((A_i)_{i\leqslant b})$, there is at least one permutation σ of $(1,2,\dots,b)$ with $\sigma(1)=1$ and a vector $\mathbf{f} \in \mathcal{F}_b$ such that $\mathcal{H}_{\mathcal{G}_n}((A_{\sigma(i)})_{i\leqslant b})$ is $\mathbf{f}$-connected. In other words, we can relabel the blocks so that the new labeling $(1,\sigma(2),\dots\sigma(b))$ is a proper labeling of T, for some $\mathbf{f} \in \mathcal{F}_b$ in Definition 6.3.1. We denote the set of permutations of $(1,2,\dots,b)$ with 1 a fixed point by $\mathcal{S}_b^1$. Note that the choice of the spanning tree T may not be unique if $\mathcal{H}_{\mathcal{G}_n}((A_i)_i)$ is connected. We obtain on the first factor inside the sum in (6.3.1) that

$$\begin{aligned}&\mathbb{P}\big(\mathcal{H}_{\mathcal{G}_n}((A_i)_{i\leqslant b}) \text{ connected}\big)\\&\qquad = \mathbb{P}\Big(\bigcup_{\mathbf{f}\in\mathcal{F}_b}\bigcup_{\sigma\in\mathcal{S}_b^1}\big\{\mathcal{H}_{\mathcal{G}_n}((A_{\sigma(i)})_{i\leqslant b}) \text{ is } \mathbf{f}\text{-connected}\big\}\Big). \qquad (6.3.2)\end{aligned}$$

If, for a given $(\mathbf{f},\sigma)$ the block graph is $\mathcal{H}_{\mathcal{G}_n}((A_{\sigma(i)})_{i\leqslant b})$ is $\mathbf{f}$-connected, then there are $\prod_i f_i!$ other pairs $(\mathbf{f}',\sigma')$ such that $\mathcal{H}_{\mathcal{G}_n}((A_{\sigma'(i)})_{i\leqslant b})$ is $\mathbf{f}'$-connected, counting the isomorphisms of rooted trees: namely, the (consecutive) labels of the forward neighbours of any vertex v may be permuted (yielding the factor $f_v!$ for each vertex), resulting in permuting the labels in the forward-subtrees of v accordingly. For any such $(\mathbf{f},\sigma)$ and $(\mathbf{f}',\sigma')$, we then also have that $\prod_{i=1}^b f_i! = \prod_{i=1}^b f_i'!$. Hence, in the above union each rooted tree T (with root fixed) on $\mathcal{H}_{\mathcal{G}_n}((A_{\sigma'(i)})_{i\leqslant b})$ is counted $\prod_{i=1} f_i!$ times. Thus, we obtain from (6.3.2) that

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

We substitute this into (6.3.1), and use that the second factor inside the sum in (6.3.1) is invariant under label permutations. So we arrive at

$$\mathbb{P}\big(\mathcal{E}_1(b)\big) \leqslant \frac{1}{b}\sum_{\mathbf{f}\in\mathcal{F}_b}\left(\prod_{i=1}^{b}\frac{1}{f_i!}\right)\sum_{A_1}\frac{1}{(b-1)!}\sum_{\sigma\in\mathcal{S}_b^1}\sum_{A_2}\cdots\sum_{A_b}\Big(\mathbb{P}\big(\mathcal{H}_{\mathcal{G}_n}((A_{\sigma(i)})_{i\leqslant b}) \text{ is } \mathbf{f}\text{-connected}\big)\cdot\mathbb{P}\Big((\cup_i\widetilde{\partial}_{\text{int}}A_{\sigma(i)}) \not\leftrightarrow_{\mathcal{G}_n} (\Lambda_n\backslash\cup_i A_{\sigma(i)})\Big)\Big).$$

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

$$\mathbb{P}\big(\mathcal{E}_1(b)\big) \leqslant \frac{1}{b}\sum_{\mathbf{f}\in\mathcal{F}_b}\left(\prod_{i=1}^{b}\frac{1}{f_i!}\right)\sum_{A_1,\dots,A_b}\mathbb{P}\big(\mathcal{H}_{\mathcal{G}_n}((A_i)_{i\leqslant b}) \ \mathbf{f}\text{-connected}\big)\cdot\mathbb{P}\big((\cup_{i\leqslant b}\widetilde{\partial}_{\text{int}}A_i) \not\leftrightarrow_{\mathcal{G}_n} (\Lambda_n\backslash\cup_{i\leqslant b}A_i)\big),$$

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

$$\mathcal{M}_b(k) := \left\{\mathbf{m}=(m_i)_{i=1}^b\subseteq\mathbb{N}^b : \exists(A_i)_{i=1}^b\subseteq\mathcal{A}_{\text{small}}\ \middle|\ \begin{matrix}|\widetilde{\partial}_{\text{int}}A_i| = m_i,\\ |\cup_{i\leqslant b}A_i|\geqslant k\end{matrix}\right\}. \tag{6.3.3}$$

Then,

$$\mathbb{P}\big(\mathcal{E}_1(b)\big) \leqslant \frac{1}{b} \sum_{\mathbf{m}\in\mathcal{M}_b(k)} \sum_{\mathbf{f}\in\mathcal{F}_b} \Big(\prod_{i\in[b]} \frac{1}{f_i!}\Big) \sum_{\substack{A_1,\dots,A_b:\\ |\widetilde{\partial}_{\mathrm{int}}A_i|=m_i}} \Big(\mathbb{P}\big(\mathcal{H}_{\mathcal{G}_n}((A_i)_i) \ \mathbf{f}\text{-conn.}\big) \cdot \mathbb{P}\big(\cup_i \partial_{\mathrm{int}}A_i \not\leftrightarrow_{\mathcal{G}_n} \Lambda_n \setminus \cup_i A_i\big)\Big), \tag{6.3.4}$$

where we omitted under the summation that the blocks are 1-disconnected from each other and in $\mathcal{A}_{\mathrm{small}}$. In what follows we we omit these descriptions under the sum for readability. The next two statements will imply Lemma 6.2.9.

Statement 6.3.2 (Counting spanning trees). *Let $\mathcal{G}_n$ be long-range percolation on Λ_n as in Definition 6.1.1 with $d \geqslant 2$, $\alpha > 1 + 1/d$. There exists $c_{6.3.2} > 0$ such that for all fixed $\mathbf{m} \in \mathcal{M}_b(1)$*

$$\sum_{\mathbf{f}\in\mathcal{F}_b} \Big(\prod_{i\in[b]} \frac{1}{f_i!}\Big) \sum_{\substack{A_1,\dots,A_b\in\mathcal{A}_{\mathrm{small}},\\ |\widetilde{\partial}_{\mathrm{int}}A_i|=m_i\,\forall i\leqslant b}} \mathbb{P}\big(\mathcal{H}_{\mathcal{G}_n}((A_i)_i) \text{ is } \mathbf{f}\text{-connected}\big) \leqslant n \exp\Big(\big(c_{6.3.2} + \log(p\beta^{\alpha})\big) \sum_{i\in[b]} m_i\Big). \tag{6.3.5}$$

Statement 6.3.3 (Isolation). *Let $\mathcal{G}_n$ be long-range percolation on Λ_n as in Definition 6.1.1 with $d \geqslant 2$, $\alpha > 1 + 1/d$. There exists $c_{6.3.3} > 0$ such that for each $\beta \geqslant 1$, any $\mathbf{m} \in \mathcal{M}_b(1)$, and any 1-disconnected blocks $A_1, \dots, A_b \in \mathcal{A}_{\mathrm{small}}$ with $|\widetilde{\partial}_{\mathrm{int}}A_i| = m_i$ for all i,*

$$\mathbb{P}\Big((\cup_i \widetilde{\partial}_{\mathrm{int}}A_i) \not\leftrightarrow_{\mathcal{G}_n} (\Lambda_n \setminus \cup_i A_i)\Big) \leqslant \exp\Big(-c_{6.3.3} \log\big(\tfrac{1}{1-p}\big)\beta^{1/d} \sum_{i\in[b]} m_i\Big). \tag{6.3.6}$$

We prove the two statements below and show first that Lemma 6.2.9 follows from them.

Proof of Lemma 6.2.9, assuming Statements 6.3.2 and 6.3.3. We substitute the bounds from Statements 6.3.2 and 6.3.3 into the right-hand side of (6.3.4) and obtain

$$\mathbb{P}\big(\mathcal{E}_1(b)\big) \leqslant \frac{1}{b} \sum_{\mathbf{m}\in\mathcal{M}_b(k)} n \exp\Big(-\big(c_{6.3.3} \log\big(\tfrac{1}{1-p}\big)\beta^{1/d} - c_{6.3.2} - \log(p\beta^{\alpha})\big) \sum_{i\in[b]} m_i\Big). \tag{6.3.7}$$

In what follows, we evaluate the summation over the vectors $\mathbf{m} \in \mathcal{M}_b(k)$. We recall from (6.3.3) that $\mathbf{m}$ represents the vector of interior boundary sizes of 1-connected sets $(A_i)_{i\leqslant b} \in \mathcal{A}_{\text{small}}$ with total size at least k, and $A_i = \bar{A}_i \leqslant 3n/4$ for all $i \leqslant b$ by the definition of $\mathcal{A}_{\text{small}}$ in (6.2.7). In (6.3.3), the boundary is taken with respect to $\mathbb{Z}^d$, i.e., *not* with respect to Λ_n. By the isoperimetric inequality in Claim 6.2.6, for all blocks $(A_i)_{i\leqslant b}$ it simultaneously holds that $|\widetilde{\partial}_{\text{int}}A_i| \geqslant |A_i|^{(d-1)/d}$. Since the function $g(k) = k^{(d-1)/d}$ is concave, we obtain for all $\mathbf{m} \in \mathcal{M}_b(k)$ and any $A_1, \dots, A_b \in \mathcal{A}_{\text{small}}$ that

$$\begin{aligned} m_1 + \ldots + m_b &= |\widetilde{\partial}_{\text{int}}A_1| + \ldots + |\widetilde{\partial}_{\text{int}}A_b| \\ &\geqslant |A_1|^{(d-1)/d} + \ldots + |A_b|^{(d-1)/d} \geqslant k^{(d-1)/d}. \end{aligned}$$

We define the set $\mathcal{M}_b(k,\ell) := \{\mathbf{m} \in \mathcal{M}_b(k) : m_1 + \ldots + m_b = \ell\}$. By standard estimates (using that each summand is at least one), we bound $|\mathcal{M}_b(k,\ell)| \leqslant \binom{\ell+b}{b} \leqslant \binom{2\ell}{\ell} \leqslant 2^{2\ell} \leqslant e^{2\ell}$. Hence, separating the summation in (6.3.7) according to the possible values of $\sum_i m_i = \ell \geqslant k^{(d-1)/d}$, we arrive at

$$\begin{aligned} &\mathbb{P}\big(\mathcal{E}_1(b)\big) \qquad (6.3.8)\\ &\leqslant \frac{1}{b}\sum_{\ell=k^{(d-1)/d}}^{\infty} n \exp\Big(-\ell\big(c_{6.3.3}\log\big(\tfrac{1}{1-p}\big)\beta^{1/d} - c_{6.3.2} - \log(p\beta^{\alpha}) - 2\big)\Big). \end{aligned}$$

Since $b \leqslant \lfloor n/2 \rfloor$, we obtain by a union bound over the number of blocks that

$$\begin{aligned} &\mathbb{P}\Big(\cup_{b\leqslant\lfloor n/2\rfloor}\mathcal{E}_1(b)\Big) \qquad (6.3.9)\\ &\leqslant \sum_{b=1}^{\lfloor n/2\rfloor}\frac{1}{b}\sum_{\ell=k^{(d-1)/d}}^{\infty} n \exp\Big(-\ell\big(c_{6.3.3}\log\big(\tfrac{1}{1-p}\big)\beta^{1/d} - c - \log(p\beta^{\alpha})\big)\Big). \end{aligned}$$

We investigate the factor after ℓ in the exponent. For fixed $p \in (0,1)$, this factor is positive whenever β is sufficiently large, depending on p, d, α, yielding $\beta_\star(p,d,\alpha)$ in Lemma 6.2.9. Evaluating the geometric summation yields (6.2.15), where the factor $\log n$ comes from the first summation in (6.3.9). Whenever p is sufficiently close to 1 (larger than p_d, for some dimension-dependent $p_d < 1$), $\beta^\star = 1$ can be also chosen since in this case the factor $\log(\frac{1}{1-p})\beta^{1/d}$ dominates the term $-\log(p\beta^{\alpha})$ for all $\beta \geqslant 1$, and the statement holds with $\beta^\star = 1$. □

6.3.1 *Proof of Statement 6.3.2*

We start with a geometric claim.

Claim 6.3.4. *There exists a constant* $C_{6.3.4} > 0$ *such that for each block* $A \in \mathcal{A}_{\text{small}}$ *in* Λ_n *and all* $r \in \mathbb{N}$

$$\left|\left\{(x,y) \in \mathbb{Z}^d \times \mathbb{Z}^d : x \in A, y \notin A, \|x-y\| \in (r, r+1]\right\}\right| \leqslant C_{6.3.4} r^d |\widetilde{\partial}_{\text{int}} A|. \tag{6.3.10}$$

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

Let $B_0 := [-1/2, 1/2]^d$. For two vertices $x, y \in \mathbb{Z}^d$, let $\mathcal{L}_{x,y}$ denote the segment between x and y on the unique line connecting x and y. Define

$$\text{Cross}(r) := \{(x,y) \in \mathbb{Z}^d \times \mathbb{Z}^d : \|x-y\|_2 \in (r, r+1], \mathcal{L}_{x,y} \cap B_0 \neq \emptyset\}.$$

We will show that there exists a constant $C_{6.3.4} > 0$ such that

$$|\text{Cross}(r)| \leqslant C_{6.3.4} r^d. \tag{6.3.11}$$

Indeed, for each pair $(x,y) \in \text{Cross}(r)$, at least one of the inequalities $\|x\| \geqslant r/2$ and $\|y\| \geqslant r/2$ is satisfied. Without loss of generality we may assume that $\|x\| \geqslant r/2$, and then also $\|x\| \leqslant r+1$. Fix then such a vertex x. Let $\mathcal{S}_x$ denote the smallest spherical cone with apex at x that completely contains B_0. This cone has then radius between $\|x\| + 1/2$ and $\|x\| + \sqrt{d}/2$. Let $\mathcal{S}_x(r)$ denote a cone with apex x that has the same boundary lines (and the same angle) as $\mathcal{S}_x$, but radius exactly r. Then, every $y \in \mathbb{Z}^d$ such that $(x,y) \in \text{Cross}(r)$ must be contained in $\mathcal{S}_x(r+1) \setminus \mathcal{S}_x(r)$, since all half-lines emanating from x that cross B_0 are contained in $\mathcal{S}_x(\infty)$, and $\|x-y\| \in (r, r+1]$. Since the radius of $\mathcal{S}_x(r+1)$ is at most by a factor two larger than the radius of $\mathcal{S}_x$ for all $r \geqslant 1$, by homothety of the cones, $|(\mathcal{S}_{r+1} \setminus \mathcal{S}_r) \cap \mathbb{Z}^d|$ is bounded from above by a dimension-dependent constant, and so for each x with $\|x\| \in [r/2, r+1]$, the number of pairs $(x,y) \in \text{Cross}(r)$ is bounded from above by a dimension-dependent constant. Summing over all the at most $O(r^d)$ many such x, we obtain (6.3.11) for some $C_{6.3.4} > 0$.

To arrive to (6.3.10), the block $A \in \mathcal{A}_{\text{small}}$ in (6.2.7) ensures that $A = \bar{A}$. Its inner boundary $\widetilde{\partial}_{\text{int}} A$ is then $*$-connected by Claim 6.2.5 Part (iv) ([76, Lemma 2.1]). This implies that there exists a continuous surface fully contained in $\widetilde{\partial}_{\text{int}} A$ separating vertices in $A \setminus \widetilde{\partial}_{\text{int}} A$ from vertices in $\Lambda_n \setminus A$. Hence, for each pair $x \in A$ and $y \notin A$, there exists (at least one) vertex

$z \in \widetilde{\partial}_{\mathrm{int}}A$ such that $\mathcal{L}_{x,y}$ intersects the axis-parallel box $z + B_0$. Here $x = z$ may occur. The statement of the claim now follows by (6.3.11) when summing the at most $C_{6.3.4}r^d$ such pairs for each vertex z in the boundary of A. □

We continue with a lemma treating the connectedness of the block graphs, i.e., the inner summation on the left-hand side of (6.3.5) in Statement 6.3.2. We point out that in this lemma we only bound the event that the block graph $\mathcal{H}_{\mathcal{G}_n}$ is connected, not the event that the actual graph is connected.

Lemma 6.3.5. *Let $\mathcal{G}_n$ be long-range percolation on Λ_n as in Definition 6.1.1 with $d \geqslant 2$, $\alpha > 1 + 1/d$. There exists a constant $C_{6.3.5} > 0$ such that for all $\mathbf{m} \in \mathcal{M}_b(1)$, $\mathbf{f} \in \mathcal{F}_b$,*

$$\sum_{\substack{A_1\dots,A_b \in \mathcal{A}_{\mathrm{small}},\\ |\widetilde{\partial}_{\mathrm{int}}A_i| = m_i\ \forall i \leqslant b}} \mathbb{P}\big(\mathcal{H}_{\mathcal{G}_n}((A_i)_{i\in[b]})\ \mathbf{f}\text{-}conn.\big) \leqslant n(C_{6.3.5}p\beta^{\alpha})^{b-1} \prod_{i\in[b]} e^{c_{\mathrm{pei}}m_i} m_i^{f_i}, \tag{6.3.12}$$

where we omitted from the summation that $A_1, \dots A_b$ are 1-disconnected from each other.

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

Proof. We will prove the statement by induction on b. We first define the finite constant, using that $\alpha > 1 + 1/d$ as follows:

$$C_{6.3.5} := C_{6.3.4} \sum_{r=1}^{\infty} r^{-(\alpha-1)d}, \tag{6.3.13}$$

We start with the initialization. Assume first that $b = 1$, which corresponds to a tree on a single vertex (representing the block A_1), so its forward degree $f_1 = 0$. A tree on a single vertex is connected by convention. We obtain

$$\sum_{\substack{A_1 \in \mathcal{A}_{\mathrm{small}}:\\ |\widetilde{\partial}_{\mathrm{int}}A_1| = m_1}} \mathbb{P}\big(\mathcal{H}_{\mathcal{G}_n}((A_1))\ \mathbf{f}\text{-conn.}\big) \leqslant \sum_{x\in\Lambda_n} |\{A_1 \in \mathcal{A}_{\mathrm{small}} : A_1 \ni x, |\widetilde{\partial}_{\mathrm{int}}A| = m_1\}|.$$

Since $A = \bar{A}$ for all $A \in \mathcal{A}_{\mathrm{small}}$ by definition in (6.2.7), we can apply Peierls' argument in Lemma 6.2.7, that yields, since $|\Lambda_n| = n$,

$$\sum_{\substack{A_1 \in \mathcal{A}_{\mathrm{small}}:\\ |\widetilde{\partial}_{\mathrm{int}}A_1| = m_1}} \mathbb{P}\big(\mathcal{H}_{\mathcal{G}_n}((A_1)))\text{ is }\mathbf{f}\text{-connected}\big) \leqslant \sum_{x\in\Lambda_n} e^{c_{\mathrm{pei}}m_1} = ne^{c_{\mathrm{pei}}m_1}.$$

Since $m_1^{f_1} = m_1^0 = 1$, this finishes the induction base for (6.3.12). We now advance the induction. Assume (6.3.12) holds up to $b-1$. Let $\mathbf{f} \in \mathcal{F}_b$ and consider the summation over $A_b \in \mathcal{A}_{\mathrm{small}}$ on the left-hand side in (6.3.12). By construction of the $\mathbf{f}$-tree in Definition 6.3.1, the b-th block is a leaf in the $\mathbf{f}$-tree, and $f_b = 0$. Its parent in the $\mathbf{f}$-tree is the largest vertex-label ℓ in $\mathbf{f}$ that is nonzero, and the remaining labeled graph upon removing b is a tree, with a labeling in $\mathcal{F}_{b-1}$ (see the comment below Definition 6.3.1). Then, the forward degrees of this new tree are given by $\mathbf{f}' := (f_1, \dots, f_{\ell-1}, f_\ell - 1, f_{\ell+1}, \dots, f_{b-1}) \in \mathcal{F}_{b-1}$, since the forward degree of the vertex ℓ decreased by one upon removing the leaf b. With this notation at hand,

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

Independence of edges in $\mathcal{G}_n$ by Definition 6.1.1 yields

$$\sum_{A_1,\dots,A_b} \mathbb{P}\big(\mathcal{H}_{\mathcal{G}_n}((A_i)_{i\in[b]}) \ \mathbf{f}\text{-conn.}\big) \tag{6.3.14}$$
$$\leqslant \sum_{A_1,\dots,A_{b-1}} \mathbb{P}\big(\mathcal{H}_{\mathcal{G}_n}((A_i)_{i\in[b-1]}) \ \mathbf{f}'\text{-conn.}\big) \sum_{A_b} \mathbb{P}\big(A_\ell \leftrightarrow_{\mathcal{G}_n} A_b\big)$$

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

We focus on the summation over A_b. Here, $A_b \in \mathcal{A}_{\mathrm{small}}$ and A_b is 1-disconnected from A_ℓ. We decompose the sum according to the length r of an edge (x, y) connecting $x \in A_\ell$ and $y \in A_b$, and use that $\|x - y\| > 1$ by the 1-disconnectedness of A_ℓ, A_b. By the connection probability (5.1.3) and a union bound, it follows that

$$\begin{aligned}\sum_{A_b} \mathbb{P}\big(A_\ell \leftrightarrow_{\mathcal{G}_n} A_b\big) &\leqslant \sum_{r=1}^{\infty} \sum_{x\in A_\ell, y \in \mathbb{Z}^d \setminus A_\ell} \mathbb{1}_{\{\|x-y\| \in (r, r+1]\}} \sum_{A_b \ni y} \mathbb{P}\big(x \leftrightarrow_{\mathcal{G}_n} y\big) \\ &\leqslant p\beta^\alpha \sum_{r=1}^{\infty} r^{-\alpha d} \sum_{x\in A_\ell, y \in \mathbb{Z}^d \setminus A_\ell} \mathbb{1}_{\{\|x-y\| \in (r, r+1]\}} \sum_{A_b \ni y} 1.\end{aligned} \tag{6.3.15}$$

In the last row the sum over y runs over a larger set than the actual allowed set of vertices: we did not exclude vertices from the other blocks $\{A_1, \dots, A_{\ell-1}, A_{\ell+1}, \dots, A_{b-1}\}$. Using (6.2.6), we bound the last sum over A_b from above by $\exp(c_{\mathrm{pei}} m_b)$. Next, we can apply Claim 6.3.4 to evaluate the summation over $x \in A_\ell, y \in \mathbb{Z}^d \setminus A_\ell$, since this sum equals the set described in (6.3.10) with $A = A_\ell$. The conditions of the claim are satisfied since $A_\ell = \bar{A}_\ell$ by assuming $A_\ell \in \mathcal{A}_{\text{small}}$. Hence,

$$\sum_{x\in A_\ell,\, y\in\mathbb{Z}^d\setminus A_\ell} \mathbb{1}_{\{\|x-y\|\in(r,r+1]\}} \sum_{A_b\ni y} 1 = e^{c_{\text{pei}} m_b} C_{6.3.4} r^d |\widetilde{\partial}_{\text{int}} A_\ell| \\ = e^{c_{\text{pei}} m_b} C_{6.3.4} r^d m_\ell.$$

Substituting this back into (6.3.15) yields with the constant $C_{6.3.5}$ from (6.3.13),

$$\sum_{A_b} \mathbb{P}\big(A_\ell \leftrightarrow_{\mathcal{G}_n} A_b\big) \leqslant p\beta^\alpha m_\ell e^{c_{\text{pei}} m_b} C_{6.3.4} \sum_{r=1}^{\infty} r^{(1-\alpha)d} \\ = C_{6.3.5} p\beta^\alpha m_\ell e^{c_{\text{pei}} m_b}.$$

We substitute this bound back into (6.3.14), and use the induction hypothesis:

$$\sum_{A_1,\dots,A_b} \mathbb{P}\big(\mathcal{H}_{\mathcal{G}_n}((A_i)_{i\in[b]})\ \mathbf{f}\text{-conn.}\big) \\ \leqslant C_{6.3.5} p\beta^\alpha m_\ell e^{c_{\text{pei}} m_b} \sum_{A_1,\dots,A_{b-1}} \mathbb{P}\big(\mathcal{H}_{\mathcal{G}_n}((A_i)_{i\in[b-1]})\ \mathbf{f}'\text{-conn.}\big) \\ \leqslant C_{6.3.5} p\beta^\alpha m_\ell e^{c_{\text{pei}} m_b} n (C_{6.3.5} p\beta^\alpha)^{b-2} \frac{1}{m_\ell} \prod_{i\in[b-1]} e^{c_{\text{pei}} m_i} m_i^{f_i}.$$

We used that $f_i' = f_i$ for all $i \neq \ell, i \leqslant b-1$, and $f_\ell' = f_\ell - 1$ by construction, yielding the $1/m_\ell$ factor. We can rearrange the expression and obtain (6.3.12), using that $f_b = 0$ (the last block is a leaf). This finishes the proof. □

We are ready to prove Statement 6.3.2.

Proof of Statement 6.3.2. Using Lemma 6.2.9 in (6.3.5) of Statement (6.3.2), we arrive at

$$\sum_{\mathbf{f}\in\mathcal{F}_b} \Big(\prod_{i\in[b]} \frac{1}{f_i!}\Big) \sum_{A_1,\dots,A_b} \mathbb{P}\big(\mathcal{H}_{\mathcal{G}_n}((A_i)_i) \text{ is } \mathbf{f}\text{-connected}\big) \tag{6.3.16}$$
$$\leqslant n \sum_{\mathbf{f}\in\mathcal{F}_b} (C_{6.3.5} p\beta^\alpha)^{b-1} \prod_{i\in[b]} \frac{e^{c_{\text{pei}} m_i} m_i^{f_i}}{f_i!}.$$

We first analyse a single summand, i.e., the value for a fixed $\mathbf{f} \in \mathcal{F}_b$. By Stirling's approximation, there exists a constant $c > 1$ such that $f_i! \geqslant (f_i/e)^{f_i}/c$. Thus,

$$n(C_{6.3.5}p\beta^{\alpha})^{b-1}\prod_{i\in[b]}\frac{e^{c_{\text{pei}}m_i}m_i^{f_i}}{f_i!}$$
$$\leqslant n(C_{6.3.5}p\beta^{\alpha})^{b-1}\exp\Big(c_{\text{pei}}\sum_{i\in[b]}m_i\Big)c^b\prod_{i\in[b]}\Big(\frac{m_i\cdot e}{f_i}\Big)^{f_i}.$$

It follows from standard differentiation techniques that for any $a, x \geqslant 1$, the function $g_a(x) = (ae/x)^x$ is maximized at $x = a$. Maximizing all factors $(m_i e/f_i)^{f_i}$ at $f_i = m_i$ yields that $(m_i e/f_i)^{f_i} \leqslant e^{m_i}$ for all $i \leqslant b$. Since by definition of $\mathcal{F}_b$ in Definition 6.3.1, we have $f_1 + \ldots + f_b = b - 1$ for all $\mathbf{f} \in \mathcal{F}_b$, we have

$$n(C_{6.3.5}p\beta^{\alpha})^{b-1}\prod_{i\in[b]}\frac{e^{c_{\text{pei}}m_i}m_i^{f_i}}{f_i!}$$
$$\leqslant n(C_{6.3.5}p\beta^{\alpha})^{b-1}\exp\Big((c_{\text{pei}}+1)\sum_{i\in[b]}m_i\Big)c^b e^{b-1}$$
$$\leqslant n\exp\Big(\big(\log(p\beta^{\alpha})+2+c_{\text{pei}}+c'\big)\sum_{i\in[b]}m_i\Big)$$

for $c' = c'(c, C_{6.3.5})$, where to obtain the last row we used that $b-1 \leqslant b \leqslant m_1 + \ldots + m_b$ by the fact that each block has at least one inner boundary vertex. Substituting this back into (6.3.16), we obtain with $c'' := 2 + c_{\text{pei}} + c'$

$$n\sum_{\mathbf{f}\in\mathcal{F}_b}(C_{6.3.5}p\beta^{\alpha})^{b-1}\prod_{i\in[b]}\frac{e^{c_{\text{pei}}m_i}m_i^{f_i}}{f_i!}$$
$$\leqslant n\exp\Big((c''+\log(p\beta^{\alpha}))\sum_{i\in[b]}m_i\Big)\sum_{\mathbf{f}\in\mathcal{F}_b}1.$$

Using again that $f_1 + \ldots + f_b = b - 1$ for all $\mathbf{f} \in \mathcal{F}_b$, and using the same combinatorial bounds as for $\mathbf{m}$ above (6.3.8), we obtain $|\mathcal{F}_b| \leqslant \binom{2b}{b} \leqslant 2^{2b} \leqslant \exp\big(2\sum_{i\in[b]} m_i\big)$, finishing the proof with $c_{6.3.2} := c'' + 2$. □

Proof of Statement 6.3.3

We start with a geometric claim. Recall $\mathcal{A}_{\text{small}}$ from (6.2.7), and holes from Definition 6.2.1.

Claim 6.3.6. *Let* $(A_1, \dots, A_b) \subseteq \mathcal{A}$ *be a* 1*-disconnected sequence of blocks without holes (i.e.,* $A_i = \bar{A}_i$ *for all* $i \leqslant b$*), with* $A := \cup_{i \leqslant b} A_i \subseteq \Lambda_n$*. Then,* $(\Lambda_n \setminus A) \cup (\cup_{i \leqslant b} \widetilde{\partial}_{\mathrm{int}} A_i) \supseteq \widetilde{\partial}_{\mathrm{int}} \Lambda_n$*. Moreover,* $(\Lambda_n \setminus A) \cup (\cup_{i \leqslant b} \widetilde{\partial}_{\mathrm{int}} A_i)$ *is* $*$*-connected.*

Proof. We first show that that $(\Lambda_n \setminus A) \cup (\cup_{i \leqslant b} \widetilde{\partial}_{\mathrm{int}} A_i) \supseteq \widetilde{\partial}_{\mathrm{int}} \Lambda_n$. When $x \in (\widetilde{\partial}_{\mathrm{int}} \Lambda_n \setminus A)$, then $x \in \Lambda_n \setminus A$. The (only) other case is when $x \in \widetilde{\partial}_{\mathrm{int}} \Lambda_n \cap A$. Then there must exist $A_i \subseteq \Lambda_n$ such that $x \in A_i$. Since $x \in \widetilde{\partial}_{\mathrm{int}} \Lambda_n$ and $A \subseteq \Lambda_n$, it follows from (6.2.3) in Definition 6.2.4 that there is a vertex y in $\mathbb{Z}^d \setminus \Lambda_n$ neighbouring x. Since $x \in A_i \subseteq \Lambda_n$, we evidently have $y \notin A_i$, hence $y \in \widetilde{\partial}_{\mathrm{ext}} A_i$. As a result of (6.2.4), $x \in \widetilde{\partial}_{\mathrm{int}} A_i$, establishing the statement.

We turn to prove $*$-connectedness of $(\Lambda_n \setminus A) \cup (\cup_{i \leqslant b} \widetilde{\partial}_{\mathrm{int}} A_i)$. Using Claim 6.2.3, we decompose the set $\Lambda_n \setminus A$ into a 1-disconnected sequence of 1-connected blocks $(B_j)_{j \leqslant b'}$ for some $b' \geqslant 1$:

$$\cup_{j \leqslant b'} B_j = \Lambda_n \setminus A. \tag{6.3.17}$$

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

By the first statement of the claim and (6.3.17), $\widetilde{\partial}_{\mathrm{int}} \Lambda_n$ is completely contained in $(\cup_{i \leqslant b} \widetilde{\partial}_{\mathrm{int}} A_i) \cup (\cup_{j \leqslant b'} B_j)$, hence $\mathcal{I}_t \cup \mathcal{I}_a$ is non-empty. Since the indices of the sets/blocks that have an intersection with $\widetilde{\partial}_{\mathrm{int}} \Lambda_n$ are all collected in $\mathcal{I}_a$ and $\mathcal{I}_t$, respectively, we have

$$\widetilde{\partial}_{\mathrm{int}} \Lambda_n \subseteq \Big(\bigcup_{i : a_i \in \mathcal{I}_a} \widetilde{\partial}_{\mathrm{int}} A_i \Big) \cup \Big(\bigcup_{j : t_j \in \mathcal{I}_t} B_j \Big) =: D. \tag{6.3.18}$$

Since we assumed that $A_i = \bar{A}_i$ for all $i \leqslant b$, Claim 6.2.5(iv) is applicable and hence $\widetilde{\partial}_{\mathrm{int}} A_i$ is $*$-connected for each $i \leqslant b$. Further, $(B_j)_{j \leqslant b'}$ are blocks, i.e., 1-connected and also $*$-connected. So, because $\widetilde{\partial}_{\mathrm{int}} \Lambda_n$ is 1-connected in dimensions 2 and higher, it is also $*$-connected, and hence the set D on the right-hand side of (6.3.18) is also $*$-connected.

We now decompose the set $(\cup_{j \leqslant b'} B_j) \cup (\cup_{i \leqslant b} \widetilde{\partial}_{\mathrm{int}} A_i)$ (using the edges of graph $\mathbb{Z}^d_\infty$) into $*$-connected *components*, in other words, sets that are $*$-connected themselves but they are $*$-disconnected from each other. Our goal is then to show that there is only a single component, namely, the one containing vertices of D. Assume now that there exists another $*$-connected *component* $\mathcal{C}$ of $(\cup_{j\leqslant b'}B_j)\cup(\cup_{i\leqslant b}\widetilde{\partial}_{\text{int}}A_i)$ that does not contain any vertices of D in (6.3.18). Being a component in $\mathbb{Z}^d_\infty$ also means $\mathcal{C}$ is then *not* $*$-connected to D. By the $*$-connectedness of each set in $(B_j)_{j\leqslant b'}$ and $(\widetilde{\partial}_{\text{int}}A_i)_{i\leqslant b}$, this is only possible if $\mathcal{C}$ is a union of some of $(B_j)_{j\leqslant b'}$ and $(\widetilde{\partial}_{\text{int}}A_i)_{i\leqslant b}$. Define then $\mathcal{I}_\mathcal{C}\subset\{a_1,\dots,a_b\}\cup\{t_1,\dots t'_b\}$ so that $a_i\in\mathcal{I}_\mathcal{C}$ if and only if $\widetilde{\partial}_{\text{int}}A_i\subseteq\mathcal{C}$, and similarly $t_j\in\mathcal{I}_\mathcal{C}$ if and only if $B_j\subseteq\mathcal{C}$. With formulas

$$\begin{aligned}&\mathcal{C}=(\cup_{i:a_i\in\mathcal{I}_\mathcal{C}}\widetilde{\partial}_{\text{int}}A_i)\cup(\cup_{j:t_j\in\mathcal{I}_\mathcal{C}}B_j)\ *\text{-connected component},\\&\mathcal{I}_\mathcal{C}\cap(\mathcal{I}_a\cup\mathcal{I}_t)=\emptyset,\end{aligned}\qquad(6.3.19)$$

where the second statement is equivalent to $\mathcal{C}\cap D=\emptyset$. Take now the closest vertex in $\cup_{j:t_j\in\mathcal{I}_\mathcal{C}}B_j$ to $\widetilde{\partial}_{\text{int}}\Lambda_n$, i.e., let (with arbitrary tie-breaking rule)

$$x_\star:=\operatorname*{arg\,min}_{x\in B_j:\ t_j\in\mathcal{I}_\mathcal{C}}\|x-\widetilde{\partial}_{\text{int}}\Lambda_n\|_1,\quad\text{and}\qquad x_\star\in B_{j^\star}\text{ with }t_{j_\star}\in\mathcal{I}_\mathcal{C}.\quad(6.3.20)$$

Since $t_{j_\star}\in\mathcal{I}_\mathcal{C}$, by our assumption in (6.3.19) the block $B_{j_\star}$ does not intersect $\widetilde{\partial}_{\text{int}}\Lambda_n$. So

$$\|x_\star-\widetilde{\partial}_{\text{int}}\Lambda_n\|_1=\|B_{j_\star}-\widetilde{\partial}_{\text{int}}\Lambda_n\|_1\geqslant 1.\qquad(6.3.21)$$

Hence, there exists a vertex $y_\star\in\Lambda_n\setminus B_{j_\star}$ such that $\|y_\star-\widetilde{\partial}_{\text{int}}\Lambda_n\|_1=\|x_\star-\widetilde{\partial}_{\text{int}}\Lambda_n\|_1-1$ and $\|x_\star-y_\star\|_1=1$. This vertex $y_\star$ cannot be part of $\cup_{j:t_j\in\mathcal{C}}B_j$ since $x_\star$ was minimal, and $y_\star$ cannot be part of $(\cup_{j\leqslant b'}B_j)\setminus(\cup_{j:t_j\in\mathcal{C}}B_j)$ by 1-disconnectedness of the sets $(B_j)_{j\leqslant b'}$, see above (6.3.17). So, there exists a block A_ℓ such that $y_\star\in A_\ell$ for some $\ell\leqslant b$. Then immediately also $y_\star\in\widetilde{\partial}_{\text{int}}A_\ell$, since $x_\star\in B_{j_\star}$ serves as an external boundary vertex for $y_\star$ in A_ℓ.

We also observe that via the pair $(x_\star,y_\star)$ the block $B_{j_\star}$ is 1-connected (and $*$-connected) to A_ℓ. Hence, since we assumed $\mathcal{C}\supseteq B_{j_\star}$ is a $*$-connected component, $A_\ell\subseteq\mathcal{C}$, equivalently, $a_\ell\in\mathcal{I}_\mathcal{C}$ and so by (6.3.19), $a_\ell\notin\mathcal{I}_a$. This then implies that $A_\ell\cap\widetilde{\partial}_{\text{int}}\Lambda_n=\emptyset$ by definition of $\mathcal{I}_a$ above (6.3.18). Hence, $\|A_\ell-\widetilde{\partial}_{\text{int}}\Lambda_n\|_1\geqslant 1$ and in turn $\widetilde{\partial}_{\text{ext}}A_\ell\subseteq\Lambda_n$. By the same reasoning as for the existence of $x_\star,y_\star$ below (6.3.21), we find a vertex $y_\circ\in\widetilde{\partial}_{\text{int}}A_\ell$ that is closest to $\widetilde{\partial}_{\text{int}}\Lambda_n$ within $\widetilde{\partial}_{\text{int}}A_\ell$, and a vertex $x_\circ\in\widetilde{\partial}_{\text{ext}}A_\ell$ that is strictly closer to $\widetilde{\partial}_{\text{int}}\Lambda_n$ than $y_\circ$, with $\|x_\circ-y_\circ\|_1=1$. We arrive at the sequence of (in)equalities:

$$\begin{aligned}\|x_\circ-\widetilde{\partial}_{\text{int}}\Lambda_n\|_1+1=\|y_\circ-\widetilde{\partial}_{\text{int}}\Lambda_n\|_1&=\|\widetilde{\partial}_{\text{int}}A_\ell-\widetilde{\partial}_{\text{int}}\Lambda_n\|_1\\&\leqslant\|y_\star-\widetilde{\partial}_{\text{int}}\Lambda_n\|_1\\&=\|x_\star-\widetilde{\partial}_{\text{int}}\Lambda_n\|_1-1,\end{aligned}\qquad(6.3.22)$$

where the inequality in the middle follows since $y_\star \in \widetilde{\partial}_{\mathrm{int}} A_\ell$ by construction, and the equalities follow by the choices of $y_\circ, x_\circ$ and $x_\star, y_\star$ below (6.3.21).

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

The block $B_{j_\circ}$ is then 1-connected via the pair $(x_\circ, y_\circ)$ to $\widetilde{\partial}_{\mathrm{int}} A_\ell$, which is itself $*$-connected, and via the pair $(x_\star, y_\star)$ the set $\widetilde{\partial}_{\mathrm{int}} A_\ell$ is 1-connected to $B_{j_\star}$. Hence, $B_{j_\circ}, A_\ell, B_{j_\star}$ are all in the same $*$-connected component $\mathcal{C}$.

However, $\|x_\circ - \widetilde{\partial}_{\mathrm{int}} \Lambda_n\|_1 < \|x_\star - \widetilde{\partial}_{\mathrm{int}} \Lambda_n\|_1$ by the inequality (6.3.22), which contradicts that $x_\star$ was a vertex in $(\cup_{j : t_j \in \mathcal{I}_\mathcal{C}} B_j)$ that minimized $\|x - \widetilde{\partial}_{\mathrm{int}} \Lambda_n\|_1$. So, the $*$-connected component that contains D in (6.3.18) contains all blocks $\cup_{j \leqslant b'} B_j = \Lambda_n \setminus A$.

Starting from (6.3.20) with reversing the role of the blocks $(B_j)_{j \leqslant b'}$ and the sets $(\widetilde{\partial}_{\mathrm{int}} A_i)_{i \leqslant b}$, the same reasoning yields that $\mathcal{I}_\mathcal{C} \cap \{a_1, \dots, a_b\} = \emptyset$, and hence the $*$-connected component that contains D in (6.3.18) contains also all sets $\cup_{i \leqslant b} \widetilde{\partial}_{\mathrm{int}} A_i$. Together, we conclude thus that $\cup_{i \leqslant b} \widetilde{\partial}_{\mathrm{int}} A_i \cup (\Lambda_n \setminus A)$ is $*$-connected. □

Proof of Statement 6.3.3. We assume $\beta \geqslant 1$. Let $A_1, \dots, A_b \in \mathcal{A}_{\mathrm{small}}$, and denote $A := \cup_{i \leqslant b} A_i$. We define the set of potential edges between the interior boundary of A with respect to $\mathbb{Z}^d$ and the set of vertices outside A within distance $\beta^{1/d}$ as

$$\Delta(A) := \big\{ \{x, y\} \mid x \in \cup_{i \leqslant b} \widetilde{\partial}_{\mathrm{int}} A_i, y \in (\Lambda_n \setminus A) : \|y - x\| \in [1, \beta^{1/d}] \big\}.$$

Considering the event on the left-hand side of (6.3.6) in Statement 6.3.3, we would like $(\cup_{i \leqslant b} \widetilde{\partial}_{\mathrm{int}} A_i)$ to be not $\mathcal{G}_n$-connected to the rest of the graph. In order to achieve this, in particular, all edges in $\Delta(A)$ must be absent. Hence,

$$\begin{aligned} \mathbb{P}\Big((\cup_{i \leqslant b} \widetilde{\partial}_{\mathrm{int}} A_i) \not\leftrightarrow_{\mathcal{G}_n} (\Lambda_n \setminus \cup_{i \leqslant b} A_i) \Big) &\leqslant (1-p)^{|\Delta(A)|} \qquad (6.3.23) \\ &= \exp\Big(-\log(\tfrac{1}{1-p}) |\Delta(A)| \Big). \end{aligned}$$

Our goal is to show that for some constant $c > 0$,

$$|\Delta(A)| \geqslant c\beta^{1/d} \sum_{i \in [b]} |\widetilde{\partial}_{\mathrm{int}} A_i| = c\beta^{1/d} \sum_{i \in [b]} m_i, \qquad (6.3.24)$$

which then immediately yields (6.3.6) in combination with (6.3.23). In what follows we estimate $|\Delta(A)|$. In order to do so, we will make use of the boundary $\partial_{\text{int}} A_i$, i.e., the interior boundary with respect to the box Λ_n. Using that all blocks in $\mathcal{A}_{\text{small}}$ have size at most $3n/4$ by definition in (6.2.7), the conditions of the isoperimetric inequality in Claim 6.2.6 are satisfied, and hence $\delta m_i \leqslant |\partial_{\text{int}} A_i| \leqslant m_i$ for all $i \leqslant b$. Hence, (6.3.24) is equivalent to showing that that there exists $c' > 0$ such that for any 1-disconnected blocks $A_1, \dots, A_b \in \mathcal{A}_{\text{small}}$, with $A = \cup_{i \leqslant b} A_i$

$$|\Delta(A)| \geqslant c' \beta^{1/d} \sum_{i \in [b]} |\partial_{\text{int}} A_i|, \tag{6.3.25}$$

since then (6.3.24) holds with $c = c'\delta$.

In order to show (6.3.25), our goal is to find enough pairs of vertices in $\Delta(A)$ around a linear fraction of vertices in $\cup_{i \leqslant b} \partial_{\text{int}} A_i$. For this, we claim that a set $\mathcal{T} := \{(x_\ell, y_\ell)\}_{\ell \geqslant 1}$ with the following properties exists:

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

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

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

Note that requirement (i) implies that all $(x_\ell, y_\ell) \in \mathcal{T}$ are elements of $\Lambda_n \times \Lambda_n$. We now show that a set $\mathcal{T}$ exists. Consider the following greedy algorithm: order the vertices in $\cup_i \partial_{\text{int}} A_i$ in an arbitrary order, to obtain the list $(v_1, v_2, \dots, v_M)$ with $M = \sum_{i \in [b]} |\partial_{\text{int}} A_i|$. Since each v_j is in $\cup_i \partial_{\text{int}} A_i$, for each v_j there is at least one vertex $y_j \in \cup_i \partial_{\text{ext}} A_i \subseteq (\Lambda_n \setminus A)$ with $\|v_j - y_j\| = 1$ by Definition 6.2.4 (recall that the sets $A_1, \dots, A_b$ are 1-disconnected). Starting with $\mathcal{T}_1 := \{(v_1, y_1)\}$, going through the ordering of $(v_j)_j$ one-by-one, append the pair (v_j, y_j) to the list $\mathcal{T}_{j-1}$, if and only if y_j has not been contained in any pair of $\mathcal{T}_{j-1}$ yet and so obtain $\mathcal{T}_j$. Set then $\mathcal{T} := \mathcal{T}_M$. Since any $y \in \cup_i \partial_{\text{ext}} A_i$ neighbours at most $2d$ many interior boundary vertices, adding a certain pair (v_j, y_j) only affects at most $2d - 1$ other indices where a pair may not be added later. Hence,

$$|\mathcal{T}| \geqslant \frac{1}{2d} \sum_{i \in [b]} |\partial_{\text{int}} A_i|. \tag{6.3.26}$$

Next, assume that $\beta^{1/d} \geqslant 2\sqrt{d} + 2$ and set $R := \lfloor (\beta^{1/d} - 1)/\sqrt{d} \rfloor \geqslant 1$. Take any pair $(x_\ell, y_\ell) \in \mathcal{T}$. Since $\cup_{i \leqslant b} \widetilde{\partial}_{\text{int}} A_i \cup (\Lambda_n \setminus A)$ is $*$-connected by Lemma 6.3.6, there exists a self-avoiding path

$$\pi(x_\ell) = (x_\ell, z_1^{(x_\ell)}, \dots, z_R^{(x_\ell)}) \subseteq \cup_{i \leqslant b} \widetilde{\partial}_{\text{int}} A_i \cup (\Lambda_n \setminus A) \tag{6.3.27}$$

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

$$\|x_\ell - z_j^{(x_\ell)}\| \leqslant \sqrt{d}R \leqslant \beta^{1/d}, \qquad \|y_\ell - z_j^{(x_\ell)}\| \leqslant \sqrt{d}R + 1 \leqslant \beta^{1/d}. \quad (6.3.28)$$

Define now the type $\mathrm{typ}(z_j^{(x_\ell)}) := x_\ell$ if $z_j^{(x_\ell)} \in (\Lambda_n \setminus A)$ and set $\mathrm{typ}(z_j^{(x_\ell)}) := y_\ell$ if $z_j^{(x_\ell)} \in \cup_{i \leqslant b} \widetilde{\partial}_{\mathrm{int}} A_i$. Then define the set of (unordered) pairs representing potential edges in $\mathcal{G}_n$

$$\Delta(x_\ell, y_\ell) := \Big\{ \{z_1^{(x_\ell)}, \mathrm{typ}(z_1^{(x_\ell)})\}, \ldots, \{z_R^{(x_\ell)}, \mathrm{typ}(z_R^{(x_\ell)})\} \Big\} \subseteq \Delta(A). \quad (6.3.29)$$

The inclusion holds since for each of these pairs, exactly one element is in $\Lambda_n \setminus A$ and the other one is in $\cup_{i \leqslant b} \widetilde{\partial}_{\mathrm{int}} A_i$, and the distance between the two vertices of each pair is at most $\beta^{1/d}$ by (6.3.28). We claim that

$$|\Delta(A)| \geqslant \Big| \bigcup_{(x_\ell, y_\ell) \in \mathcal{T}} \Delta(x_\ell, y_\ell) \Big| \overset{\diamond}{\geqslant} (1/2) \sum_{\ell=1}^{|\mathcal{T}|} |\Delta(x_\ell, y_\ell)| = |\mathcal{T}| \cdot R/2. \quad (6.3.30)$$

To see the inequality with $\diamond$, we show that each potential edge $\{z, z'\} \in \Lambda_n \times \Lambda_n$ appears at most twice in a set in the union in the middle. Consider $\{z, z'\} \in \Lambda_n \times \Lambda_n$. First, assume that there exists ℓ such that $(z, z') = (x_\ell, y_\ell) \in \mathcal{T}$ or $(z', z) = (x_\ell, y_\ell) \in \mathcal{T}$. Without loss of generality, we assume that the pair is ordered such that $(z, z') = (x_\ell, y_\ell) \in \mathcal{T}$. Then there is no $(x_j, y_j) \in \mathcal{T}$ different from (x_ℓ, y_ℓ) such that $\{z, z'\} \in \Delta(x_j, y_j)$, since each element in $\Delta(x_j, y_j)$ contains either x_j or y_j, which are different from x_ℓ and from y_ℓ by requirement (ii) in the construction of $\mathcal{T}$. Moreover, the element $\{z, z'\} = \{x_\ell, y_\ell\}$ is contained at most once in the set $\Delta(x_\ell, y_\ell)$, since the first coordinates in (6.3.29) are all different as they form a self-avoiding path, and the first coordinates do no not contain $x_\ell = z$ by (6.3.27).

Next, assume that $(z, z'), (z', z) \notin \mathcal{T}$, but $\{z, z'\}$ is contained in some $\Delta(x_\ell, y_\ell)$. Then, either z or z' must be equal to either x_ℓ or to y_ℓ by (6.3.29). Assume without loss of generality that $z \in \{x_\ell, y_\ell\}$, and therefore $z' \notin \{x_\ell, y_\ell\}$. Thus, $\{z, z'\}$ is contained exactly once in $\Delta(x_\ell, y_\ell)$. The only way that $\{z, z'\}$ could be in a set $\Delta(x_{\ell'}, y_{\ell'})$ for some $(x_{\ell'}, y_{\ell'}) \neq (x_\ell, y_\ell)$, is when $z' \in \{x_{\ell'}, y_{\ell'}\}$ and $(x_{\ell'}, y_{\ell'}) \in \mathcal{T}$ for some $\ell' \neq \ell$. This implies that the element $\{z, z'\}$ can be contained at most twice in a set in the union in (6.3.30), namely in $\Delta(x_\ell, y_\ell)$ and $\Delta(x_{\ell'}, y_{\ell'})$, and the inequality $\diamond$ in (6.3.30) holds.

Combining (6.3.30) with (6.3.26), $R = \lfloor(\beta^{1/d}-1)/\sqrt{d}\rfloor$, and the assumption that $\beta^{1/d} \geqslant 2\sqrt{d}+2$, (see before (6.3.28)), we arrive at

$$|\Delta(A)| \geqslant \frac{R}{4d}\sum_{i\in[b]}|\partial_{\text{int}}A_i| = \frac{1}{4d}\left\lfloor\frac{\beta^{1/d}-1}{\sqrt{d}}\right\rfloor\sum_{i\in[b]}|\partial_{\text{int}}A_i|$$
$$\geqslant \frac{\beta^{1/d}}{8d\sqrt{d}}\sum_{i\in[b]}|\partial_{\text{int}}A_i|,$$

since whenever $x \geqslant 2\sqrt{d}+2$, then $\lfloor(x-1)/\sqrt{d}\rfloor \geqslant x/2\sqrt{d}$.

This proves (6.3.25) whenever $\beta^{1/d} \geqslant 2\sqrt{d}+2$. For the case $1 \leqslant \beta^{1/d} \leqslant 2\sqrt{d}+2$, we use that each vertex on the interior boundary is within distance one from a vertex on the exterior boundary, hence

$$|\Delta(A)| \geqslant \sum_{i\in[b]}|\partial_{\text{int}}A_i| \geqslant \frac{\beta^{1/d}}{2\sqrt{d}+2}\sum_{i\in[b]}|\partial_{\text{int}}A_i|,$$

and so (6.3.25) holds for both cases with $c := 1/\max\{8d\sqrt{d}, 2\sqrt{d}+2\}$. This finishes the proof of Statement 6.3.3. □

6.4 COUNTING HOLES

We turn to the proof of Lemma 6.2.10. We set up a few preliminaries about holes. Recall that $\mathcal{A}$ denotes the 1-connected blocks in Λ_n from (6.2.1), and that $\mathcal{A}_{\text{large}} = \{A \in \mathcal{A} : |\bar{A}| > 3n/4,\ |A| \leqslant n/2\}$ from (6.2.7). Recall also from Definition 6.2.1 that the holes $\mathfrak{H}_A$ of a 1-connected set $A \in \mathcal{A}$ are the 1-connected subsets of $\bar{A} \setminus A$. By Definition 6.2.1, each hole $H \in \mathfrak{H}_A$ is surrounded by A. This implies that H does not intersect the boundary of the box $\partial_{\text{int}}\Lambda_n$. Hence, it follows by Definition 6.2.4 of the boundaries that for all $H \in \mathfrak{H}_A$

$$\partial_{\text{int}}H = \widetilde{\partial}_{\text{int}}H \subseteq \partial_{\text{ext}}A. \tag{6.4.1}$$

Hence, comparing this to $\mathcal{E}_2 = \{\exists A \in \mathcal{A}_{\text{large}} : A \not\leftrightarrow_{\mathcal{G}_n} \partial_{\text{ext}}A\}$ from (6.2.9), we obtain that

$$\mathbb{P}(\mathcal{E}_2) \leqslant \mathbb{P}\big(\exists A \in \mathcal{A}_{\text{large}} : A \not\leftrightarrow_{\mathcal{G}_n} (\cup_{H\in\mathfrak{H}_A}\partial_{\text{int}}H)\big). \tag{6.4.2}$$

The following definition (and claim) of principal holes will ensure that the total size of the boundaries of the holes of A in (6.4.1) is sufficiently large compared to the combinatorial factor arising from the number of possible sets A there.

Definition 6.4.1 (Principal holes)**.** Let $A \subseteq \Lambda_n$. A hole $H \in \mathfrak{H}_A$ has type $i \in \mathbb{N}$ if $2^{i-1} < |H| \leqslant 2^i$. We write $\mathfrak{H}_A(i) \subseteq \mathfrak{H}_A$ for the set of holes of A of type i. A hole-type i is called *principal* for a set A if

$$|\mathfrak{H}_A(i)| = |\{H : H \in \mathfrak{H}_A(i)\}| \geqslant 2^{-i-3} i^{-2} n =: h_n(i). \tag{6.4.3}$$

Since $|\mathfrak{H}_A(i)|$ is an integer for all i, the inequality $|\mathfrak{H}_A(i)| \geqslant \lceil h_n(i) \rceil$ also holds whenever the inequality in (6.4.3) holds. Hence, we define for $i \in \mathbb{N}$

$$\mathcal{A}_{\text{large}}(i) := \{A \in \mathcal{A}_{\text{large}} : |\mathfrak{H}_A(i)| \geqslant \lceil h_n(i) \rceil\}, \tag{6.4.4}$$

and observe that A may appear in both $\mathcal{A}_{\text{large}}(i)$ and $\mathcal{A}_{\text{large}}(j)$ if both type i and type j are principal for A. Define the following β-dependent constants

$$\begin{aligned} R_2 = R_2(\beta) &:= \max\left\{ \left\lfloor \frac{\beta^{1/d}}{\sqrt{d}} \right\rfloor, 1 \right\}, \\ i_\star = i_\star(\beta) &:= 1 + \lceil \log_2(R_2) \rceil . \end{aligned} \tag{6.4.5}$$

Claim 6.4.2 (Large blocks have a principal hole-type)**.** *For all $A \in \mathcal{A}_{\text{large}}$, there exists $i_A \leqslant \lceil \log_2(n) \rceil$ such that hole-type i is principal, i.e.,*

$$\mathcal{A}_{\text{large}} \subseteq \bigcup_{i \leqslant \lceil \log_2(n) \rceil} \mathcal{A}_{\text{large}}(i). \tag{6.4.6}$$

There exists a constant $c_{6.4.3} > 0$ such that for all $A \subseteq \Lambda_n$, with i_A a principal hole-type for A,

$$\sum_{H \in \mathfrak{H}_A(i_A)} |\partial_{\text{int}} H_i| \geqslant c_{6.4.2}\, i_A^{-2}\, 2^{-i_A/d}\, n. \tag{6.4.7}$$

Moreover, for each hole H with type $i \geqslant i_\star$ in (6.4.5)*,*

$$|\partial_{\text{ext}} H| \geqslant R_2^{(d-1)/d}. \tag{6.4.8}$$

Proof. We argue by contradiction for the first part. By definition of $\mathcal{A}_{\text{large}}$ in (6.2.7), $|\bar{A}| \geqslant 3n/4$, and also $|A| \leqslant n/2$. Hence, the total size of the holes is at least $n/4$, i.e., $|\cup_{H \in \mathfrak{H}_A} H| = \sum_{H \in \mathfrak{H}_A} |H| \geqslant n/4$ hold. Suppose (6.4.3) holds in the opposite direction for all $i \geqslant 1$. Since the holes are 1-disconnected and form together the complement of $\bar{A} \setminus A$, it follows from the size requirement in Definition 6.4.1 that

$$\Big| \bigcup_{H \in \mathfrak{H}_A} H \Big| = \sum_{H \in \mathfrak{H}_A} |H| \leqslant \sum_{i \geqslant 1} |\mathfrak{H}_A(i)| 2^i \leqslant \frac{n}{8} \sum_{i \geqslant 1} i^{-2} < n/4,$$

since the sum converges increasingly to $\pi^2/6 = 1.64... < 2$. This contradicts the assumption that the total size is at least $n/4$, so there must be at least one principal hole-type, say i_A. The restriction $i \leqslant \lceil \log_2(n) \rceil$ follows since the number of vertices in Λ_n is n, and $H > 2^{(\lceil \log_2(n) \rceil + 1) - 1}$ thus can never be satisfied. This shows (6.4.6).

We turn to (6.4.7). As argued before (6.4.1), holes do not intersect the boundary of the box $\partial_{\text{int}}\Lambda_n$. So, $\partial_{\text{ext}}H = \widetilde{\partial}_{\text{ext}}H$, and $|\partial_{\text{ext}}H| \geqslant |H|^{(d-1)/d}$ by Claim 6.2.6 for each hole. Combined with $|H| > 2^{i_A - 1}$ for all $H \in \mathfrak{H}_A(i_A)$ and the lower bound $|\mathfrak{H}_A(i_A)| \geqslant 2^{-i_A-3} i^{-2} n$ in (6.4.3), this yields that

$$\begin{aligned}\sum_{H \in \mathfrak{H}_A(i_A)} |\partial_{\text{int}}H| &\geqslant \sum_{H \in \mathfrak{H}_A(i_A)} |H|^{(d-1)/d} \\ &\geqslant 2^{i_A(d-1)/d} 2^{(d-1)/d} |\mathfrak{H}_A(i_A)| \\ &\geqslant 2^{-(d-1)/d} 2^{i_A(d-1)/d} \cdot 2^{-i_A-3} i_A^{-2} n \\ &\geqslant c_{6.4.2}\, i_A^{-2}\, 2^{-i_A/d}\, n,\end{aligned}$$

for some constant $c_{6.4.2} > 0$. Lastly, we prove (6.4.8). Using again that $\partial_{\text{ext}}H \geqslant |H|^{(d-1)/d}$ for each hole, we obtain for any hole H with type $i \geqslant i_\star$,

$$|\partial_{\text{ext}}H| \geqslant |H|^{(d-1)/d} > 2^{(i_\star-1)(d-1)/d} \geqslant \left(2^{\log_2(R_2)}\right)^{(d-1)/d} = R_2^{(d-1)/d}.$$

This finishes the proof of the claim. □

We use a union bound on (6.4.6) first in (6.4.2) (with the convention that the empty sum from 1 to $i_\star$ is 0). We arrive at

$$\mathbb{P}(\mathcal{E}_2) \leqslant \sum_{i=1}^{i_\star-1} \mathbb{P}\big(\exists A \in \mathcal{A}_{\text{large}}(i) : A \not\leftrightarrow_{\mathcal{G}_n} \cup_{H \in \mathfrak{H}_A} \partial_{\text{int}}H\big) \tag{6.4.9}$$

$$+ \sum_{i=i_\star}^{\lceil \log_2(n) \rceil} \mathbb{P}\Big(\exists A \in \mathcal{A}_{\text{large}}(i) : A \not\leftrightarrow_{\mathcal{G}_n} \cup_{H \in \mathfrak{H}_A} \partial_{\text{int}}H\Big). \tag{6.4.10}$$

We will now bound these two sums, the first one corresponding to *small principal hole types*, the second one corresponding to *large principal hole types*.

Excluding small principal hole types

We bound the two sums on the right-hand side of (6.4.9) separately and start with the first one.

Claim 6.4.3 (Sets with small principal hole-types are unlikely components)**.** *Let* $\mathcal{G}_n$ *be long-range percolation on* Λ_n *as in Definition 6.1.1 with* $d \geqslant 2$, $\alpha > 1$, *and* $i_\star(\beta)$ *from* (6.4.5). *Then there exists* $c_{6.4.3} > 0$ *such that for all* $\beta \geqslant 1$

$$\begin{aligned}\mathrm{Err}_{\mathrm{small}} &:= \sum_{i=1}^{i_\star-1} \mathbb{P}\big(\exists A \in \mathcal{A}_{\mathrm{large}}(i) : A \not\leftrightarrow_{\mathcal{G}_n} \cup_{H\in\mathfrak{H}_A} \partial_{\mathrm{int}} H\big) \qquad (6.4.11)\\ &\leqslant i_\star 2^n \exp\big(-n c_{6.4.3}\big(\log\big(\tfrac{1}{1-p}\big)\beta^{(d-1)/d^2}/(1+\log_2(\beta^{1/d}))^2\big)\big).\end{aligned}$$

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

Proof. We may assume that $i_\star(\beta) > 0$ in (6.4.5), since otherwise the sum would be empty and the bound holds trivially. We start estimating a single summand on the left-hand side of (6.4.11). Consider some $A \in \mathcal{A}_{\mathrm{large}}(i)$. By Definition 6.4.1,

$$h_A := |\mathfrak{H}_A| \geqslant |\mathfrak{H}_A(i)| \geqslant h_n(i) = 2^{-i-3} i^{-2} n.$$

We now find enough potential edges that all must be absent in order for the event $\{A \not\leftrightarrow_{\mathcal{G}_n} \cup_{H\in\mathfrak{H}_A} \partial_{\mathrm{int}} H\}$ in (6.4.11) to occur. By (6.4.7) in Claim 6.4.2, since i is a principal hole-type of $A \in \mathcal{A}_{\mathrm{large}}(i)$

$$|\cup_{H\in\mathfrak{H}_A} \partial_{\mathrm{int}} H| = \sum_{H\in\mathfrak{H}_A(i)} |\partial_{\mathrm{int}} H| \geqslant c_{6.4.2} i^{-2} 2^{-i/d} n =: \ell_i. \qquad (6.4.12)$$

We now obtain a lower bound on $|A|$ using the isoperimetric inequality of $\mathbb{Z}^d$ in Claim 6.2.6. By definition of $\widetilde{\partial}_{\mathrm{int}} A$ in Definition 6.2.4, and since $\widetilde{\partial}_{\mathrm{int}} A \subseteq \widetilde{\partial}_{\mathrm{int}} \bar{A}$ by Claim 6.2.5(i), it follows from Claim 6.2.6 applied to $\bar{A}$ that for all $A \in \mathcal{A}_{\mathrm{large}}$,

$$|A| \geqslant |\widetilde{\partial}_{\mathrm{int}} A| \geqslant |\widetilde{\partial}_{\mathrm{int}} \bar{A}| \geqslant |\bar{A}|^{(d-1)/d} \geqslant (3/4)^{(d-1)/d} n^{(d-1)/d}.$$

Take now a vertex $x \in \partial_{\mathrm{int}} H \subseteq \partial_{\mathrm{ext}} A$ and recall $R_2(\beta)$ from (6.4.5). Then, since $|A|$ diverges with n and A is 1-connected, whenever n is sufficiently large compared to β,

$$|\{y \in A : \|y-x\| \leqslant R_2(\beta)\}| \geqslant R_2(\beta), \qquad \forall x \in \cup_{H\in\mathfrak{H}_A} \partial_{\mathrm{int}} H.$$

Hence, by (6.4.12),

$$\begin{aligned}\big|\big\{\{x,y\} : x \in \cup_{H\in\mathfrak{H}_A} \partial_{\mathrm{int}} H, y \in A : \|y-x\| \leqslant R_2(\beta)\big\}\big| &\geqslant R_2 \cdot |\cup_{H\in\mathfrak{H}_A} \partial_{\mathrm{int}} H| \\ &\geqslant R_2 \ell_i.\end{aligned}$$

These edges must be all absent in order for $\{A \not\leftrightarrow_{\mathcal{G}_n} \cup_{H\in\mathfrak{H}_A}\partial_{\mathrm{int}}H\}$ to occur for $A \in \mathcal{A}_{\mathrm{large}}(i)$ in (6.4.11). Using a union bound and the independence of edges, we obtain

$$\begin{aligned}\mathbb{P}\big(\exists A \in \mathcal{A}_{\mathrm{large}}(i) : A \not\leftrightarrow_{\mathcal{G}_n} \cup_{H\in\mathfrak{H}_A}\partial_{\mathrm{int}}H\big) \\ &\leqslant \sum_{A\in\mathcal{A}_{\mathrm{large}}(i)} \mathbb{P}\big(A \not\leftrightarrow_{\mathcal{G}_n} \cup_{H\in\mathfrak{H}_A}\partial_{\mathrm{int}}H\big) \\ &\leqslant \sum_{A\in\mathcal{A}_{\mathrm{large}}(i)} (1-p)^{R_2\ell_i} \leqslant 2^n(1-p)^{R_2\ell_i},\end{aligned}$$

where we used that $\mathcal{A}_{\mathrm{large}}(i)$ counts subsets of Λ_n, and the number of subsets of Λ_n is at most 2^n. This bounds a single summand in (6.4.11). To evaluate the sum, recalling that $\ell_i = c_{6.4.2}i^{-2}2^{-i/d}n$ from (6.4.12), we have

$$\begin{aligned}\mathrm{Err}_{\mathrm{small}} &\leqslant \sum_{i=1}^{i_\star(\beta)-1} 2^n(1-p)^{R_2\ell_i} \\ &= 2^n \sum_{i=1}^{i_\star(\beta)} \exp\Big(-n\log(\tfrac{1}{1-p})R_2c_{6.4.2}i^{-2}2^{-i/d}\Big) \\ &\leqslant i_\star 2^n \exp\Big(-nc_{6.4.2}2^{-1/d}\log\big(\tfrac{1}{1-p}\big)R_2^{(d-1)/d}/(\log_2(R_2)+1)^2\Big),\end{aligned}$$

where for the last inequality we used that for all $i \leqslant i_\star - 1 = \lceil\log_2(R_2)\rceil$, we have that $2^{-i/d}i^{-2} \geqslant 2^{-1/d}R_2^{-1/d}/(\log_2(R_2)+1)^2$. Recalling that $R_2 = \max\{\lfloor\beta^{1/d}/\sqrt{d}\rfloor, 1\}$ from (6.4.5), the statement in (6.4.11) follows by changing the constant factor in the exponent to obtain $c_{6.4.3}$. □

Excluding large principal holes

We turn to the second sum in (6.4.10).

Claim 6.4.4 (Sets with large principal hole-types are unlikely components)**.** *Let $\mathcal{G}_n$ be long-range percolation on Λ_n as in Definition 6.1.1 with $d \geqslant 2$, $\alpha > 1$, and $i_\star(\beta)$ from (6.4.5). Then there exists $c_{6.4.3} > 0$ such that*

$$\begin{aligned}\mathrm{Err}_{\mathrm{large}} &:= \sum_{i=i_\star}^{\lceil\log_2(n)\rceil} \mathbb{P}\Big(\exists A \in \mathcal{A}_{\mathrm{large}}(i) :\, A \not\leftrightarrow_{\mathcal{G}_n} \cup_{H\in\mathfrak{H}_A}\partial_{\mathrm{int}}H\Big) \\ &\leqslant \exp\Big(-c_{6.4.4}\log\big(\tfrac{1}{1-p}\big)\beta^{(d-1)/d^2}\log^{-2}(n)n^{(d-1)/d}\Big).\end{aligned} \qquad (6.4.13)$$

when $\beta^{(d-1)/d^2}$ is sufficiently large or p is sufficiently close to one.

Proof. Similarly to the small principal hole-types, we will find enough potential edges that all must be absent in order for the events on the left-hand-side in (6.4.13) to occur. Take a fixed ordering L of vertices in Λ_n so that $x_1 <_L x_2 <_L \cdots <_L x_n$ with respect to this ordering (e.g., the lexicographic ordering). For a block $A \in \mathcal{A}_{\text{large}}(i)$ (which has at least $\lceil h_n(i) \rceil$ holes of type i by (6.4.3)), we order its holes $\mathfrak{H}_A$ in such a way that the holes of type i are $H_A^{(1)}, \dots, H_A^{|\mathfrak{H}_A(i)|}$, and that for all $r < s \leqslant |\mathfrak{H}_A(i)|$ the vertices smallest in the ordering within $H_A^{(r)}$ and $H_A^{(s)}$ – say $x_r \in H_A^{(r)}$ and $x_s \in H_A^{(s)}$ – satisfy $x_r <_L x_s$. We obtain, when excluding edges from A towards only its first $\lceil h_n(i) \rceil$ holes of type i

$$\begin{aligned}
&\mathbb{P}\big(\exists A \in \mathcal{A}_{\text{large}}(i) : A \not\leftrightarrow_{\mathcal{G}_n} \cup_{H \in \mathfrak{H}_A} \partial_{\text{int}} H\big) \\
&\qquad \leqslant \mathbb{P}\big(\exists A \in \mathcal{A}_{\text{large}}(i) : A \not\leftrightarrow_{\mathcal{G}_n} \cup_{j \leqslant \lceil h_n(i) \rceil} \partial_{\text{int}} H_A^{(j)}\big), && (6.4.14) \\
&\qquad \leqslant \mathbb{P}\Big(\exists A \in \mathcal{A}_{\text{large}}(i) : \cup_{j \leqslant \lceil h_n(i) \rceil} \partial_{\text{ext}} H_A^{(j)} \not\leftrightarrow_{\mathcal{G}_n} \cup_{j \leqslant \lceil h_n(i) \rceil} \partial_{\text{int}} H_A^{(j)}\Big), && (6.4.15)
\end{aligned}$$

where to get the second row we only look at edges emanating from A that are on the exterior boundaries of the holes. This is an upper bound since $\partial_{\text{ext}} H_A^{(j)} \subseteq A$ by (6.4.1) for all $j \leqslant \lceil h_n(i) \rceil$. If for two blocks $A, A' \in \mathcal{A}_{\text{large}}(i)$, the first $\lceil h_n(i) \rceil$ holes coincide, also the exterior boundaries of these first $\lceil h_n(i) \rceil$ holes coincide, and the event in (6.4.15) excludes the exact same edges. So, a simple union bound over A in (6.4.15) would overcount the non-presence of those edges too many times. Instead, we carry out a union bound over all possible lists of the first $\lceil h_n(i) \rceil$ holes. To this end, we consider for all $A \in \mathcal{A}_{\text{large}}(i)$ the following:

$$\begin{aligned}
D(A) &:= \Lambda_n \setminus \{\cup_{j \leqslant \lceil h_n(i) \rceil} H_A^{(j)}\}, \\
\mathcal{D}(i) &:= \{D : \exists A \in \mathcal{A}_{\text{large}}(i), D = D(A)\}.
\end{aligned} \qquad (6.4.16)$$

Since $D(A)$ shares the first $\lceil h_n(i) \rceil$ holes with A,

$$\begin{aligned}
&\{\cup_{j \leqslant \lceil h_n(i) \rceil} \partial_{\text{ext}} H_{D(A)}^{(j)} \not\leftrightarrow_{\mathcal{G}_n} \cup_{j \leqslant \lceil h_n(i) \rceil} \partial_{\text{int}} H_{D(A)}^{(j)}\} \\
&\qquad = \{\cup_{j \leqslant \lceil h_n(i) \rceil} \partial_{\text{ext}} H_A^{(j)} \not\leftrightarrow_{\mathcal{G}_n} \cup_{j \leqslant \lceil h_n(i) \rceil} \partial_{\text{int}} H_A^{(j)}\}.
\end{aligned}$$

Hence, in (6.4.15) we can group the blocks in $\mathcal{A}_{\text{large}}(i)$ that all map to the same $D \in \mathcal{D}(i)$, and obtain that

$$\begin{aligned}
&\mathbb{P}\Big(\exists A \in \mathcal{A}_{\text{large}}(i) : \cup_{j \leqslant \lceil h_n(i) \rceil} \partial_{\text{ext}} H_A^{(j)} \not\leftrightarrow_{\mathcal{G}_n} \cup_{j \leqslant \lceil h_n(i) \rceil} \partial_{\text{int}} H_A^{(j)}\Big), \\
&\qquad = \mathbb{P}\Big(\exists D \in \mathcal{D}(i) : \cup_{j \leqslant \lceil h_n(i) \rceil} \partial_{\text{ext}} H_D^{(j)} \not\leftrightarrow_{\mathcal{G}_n} \cup_{j \leqslant \lceil h_n(i) \rceil} \partial_{\text{int}} H_D^{(j)}\Big) \\
&\qquad \leqslant \sum_{D \in \mathcal{D}(i)} \mathbb{P}\Big(\cup_{j \leqslant \lceil h_n(i) \rceil} \partial_{\text{ext}} H_D^{(j)} \not\leftrightarrow_{\mathcal{G}_n} \cup_{j \leqslant \lceil h_n(i) \rceil} \partial_{\text{int}} H_D^{(j)}\Big). && (6.4.17)
\end{aligned}$$

We combine the following three observations to bound a single summand in the last row. First, each vertex $x \in \partial_{\text{int}} H_D^{(j)}$ is at distance one from at least one vertex $y_x \in \partial_{\text{ext}} H_D^{(j)} \subseteq D \subseteq \Lambda_n$ (by definition, a hole $H_D^{(j)}$ does not intersect $\widetilde{\partial}_{\text{int}}\Lambda_n$). Second, for all $j \leqslant \lceil h_n(i) \rceil$, $|\partial_{\text{ext}} H_D^{(j)}| \geqslant R_2^{(d-1)/d}$ by (6.4.8) and the fact that $i \geqslant i_\star$. Third, the exterior boundary of a hole $\partial_{\text{ext}} H_D^{(j)}$ is $*$-connected by Claim 6.2.5(v).

Hence, for each vertex $x \in \partial_{\text{int}} H_D^{(j)}$, starting from $y_x \in \partial_{\text{ext}} H_D^{(j)}$, one can find a $*$-connected set of vertices $\mathcal{B}_x \subseteq \partial_{\text{ext}} H_D^{(j)}$ that satisfies

$$|\mathcal{B}_x| \geqslant R_2^{(d-1)/d}, \quad \text{and} \quad \forall z \in \mathcal{B}_x : \|x - z\| \leqslant \sqrt{d} R_2^{(d-1)/d}.$$

Then, the edges $\{\{x, z\} : x \in \partial_{\text{int}} H_D^{(j)}, z \in \mathcal{B}_x\}$ all need to be absent for the event in (6.4.17) to occur, and the distance bound on $\|x - z\|$ ensures, using (6.1.1) and that $R_2 = \max\{\beta^{1/d}/\sqrt{d}, 1\}$ in (6.4.5), that all these edges are present with probability p whenever $\beta \geqslant 1$. Combining then (6.4.14), (6.4.15) with (6.4.17), it follows by the independence of edges in $\mathcal{G}_n$ that

$$\begin{aligned}\mathbb{P}\big(\exists A \in \mathcal{A}_{\text{large}}(i) : A \not\leftrightarrow_{\mathcal{G}_n} \cup_{H \in \mathfrak{H}_A} \partial_{\text{int}} H\big) \\ \leqslant \sum_{D \in \mathcal{D}(i)} \prod_{j \leqslant \lceil h_n(i) \rceil} (1-p)^{|\partial_{\text{int}} H_D^{(j)}| \cdot R_2^{(d-1)/d}}.\end{aligned}$$

We will now encode the holes of $D \in \mathcal{D}(i)$ similar to the encoding of the blocks in Section 6.3. We write $\mathbf{x}_D := (x_1, \dots, x_{\lceil h_n(i) \rceil})$ for the vertices with the smallest label in the L-ordering within the respective holes $H_D^{(1)}, \dots, H_D^{\lceil h_n(i) \rceil}$. Let us write $\Lambda_n^{\lceil h_n(i) \rceil, <}$ for the vectors $\mathbf{x} \in \Lambda_n^{\lceil h_n(i) \rceil}$ with $x_r <_L x_s$ for all $r < s$. By the initial ordering of the holes above (6.4.14), $\mathbf{x}_D \in \Lambda_n^{\lceil h_n(i) \rceil, <}$. Let then $m_j := |\partial_{\text{int}} H_D^{(j)}|$ for all $j \leqslant \lceil h_n(i) \rceil$, and write $\mathbf{m}_D := (m_1, \dots, m_{\lceil h_n(i) \rceil})$. Define then for all $\mathbf{x} \in \Lambda_n^{\lceil h_n(i) \rceil, <}$, and $\mathbf{m} \in \mathbb{N}^{\lceil h_n(i) \rceil}$:

$$\mathcal{D}(i, \mathbf{x}, \mathbf{m}) := \{D \in \mathcal{D}(i) : \mathbf{x}_D = \mathbf{x}, \mathbf{m}_D = \mathbf{m}\}.$$

The set $D \in \mathcal{D}(i)$ has the first $\lceil h_n(i) \rceil$ holes of some $A \in \mathcal{A}_{\text{large}}(i)$ where $D(A) = D$, and for that A, hole-type i was principal in terms of Definition 6.4.1 and (6.4.4). Definition 6.4.1 readily implies that Claim 6.4.2 is applicable to already the first $\lceil h_n(i) \rceil$ many holes of A, and in turn, D. Hence, the total interior boundary size $m := \sum_{j=1}^{\lceil h_n(i) \rceil} m_j$ satisfies that $m \geqslant c_{6.4.2} i^{-2} 2^{-i/d} n$. So for $m \geqslant c_{6.4.2} i^{-2} 2^{-i/d} n$ we introduce the possible boundary-length vectors with total size m:

$$\mathcal{M}_i(m) := \big\{\mathbf{m} \in \mathbb{N}^{\lceil h_n(i) \rceil} : m_1 + \ldots + m_{\lceil h_n(i) \rceil} = m\big\}. \tag{6.4.18}$$

Returning to (6.4.17), we decompose the summation on the right hand side as follows:

$$
\begin{aligned}
&\mathbb{P}\big(\exists A \in \mathcal{A}_{\text{large}}(i) : A \not\leftrightarrow_{\mathcal{G}_n} \cup_{H\in\mathfrak{H}_A} \partial_{\text{int}} H\big) && (6.4.19)\\
&\quad\leqslant \sum_{m\geqslant c_{6.4.2} i^{-2}2^{-i/d} n} \ \sum_{\mathbf{m}\in\mathcal{M}_i(m)} \ \sum_{\mathbf{x}\in\Lambda_n^{\lceil h_n(i)\rceil,<}} \ \sum_{D\in\mathcal{D}(i,\mathbf{x},\mathbf{m})} && (6.4.20)\\
&\qquad\qquad \prod_{j\leqslant\lceil h_n(i)\rceil} (1-p)^{|\partial_{\text{int}} H_D^{(j)}|\cdot R_2^{(d-1)/d}}\\
&\quad= \sum_{m\geqslant c_{6.4.2} i^{-2}2^{-i/d} n} (1-p)^{m R_2^{(d-1)/d}} \sum_{\mathbf{m}\in\mathcal{M}_i(m)} \ \sum_{\mathbf{x}\in\Lambda_n^{\lceil h_n(i)\rceil,<}} \ \sum_{D\in\mathcal{D}(i,\mathbf{x},\mathbf{m})} 1.
\end{aligned}
$$

Now we evaluate the number of terms of the last three summations. Each block $D \in \mathcal{D}(i, \mathbf{x}, \mathbf{m})$ is uniquely characterized by its $\lceil h_n(i)\rceil$ holes by (6.4.16). Having fixed the vectors $\mathbf{x}$ and $\mathbf{m}$, we apply Lemma 6.2.7 to each hole $H_D^{(j)}$ with $H_D^{(j)} \ni x_j$ and $|\partial_{\text{int}} H_D^{(j)}| = m_j$ to count the size of $\mathcal{D}(i, \mathbf{x}, \mathbf{m})$. The lemma is applicable since for each hole H, we have $H = \bar{H}$ by Claim 6.2.5(v). Hence, there are at most $\exp(c_{\text{pei}} m_j)$ possible holes of interior boundary size m_j containing x_j. So, for all $\mathbf{x} \in \Lambda_n^{\lceil h_n(i)\rceil}$ and $\mathbf{m} \in \mathcal{M}_i(m)$,

$$
|\mathcal{D}(i, \mathbf{x}, \mathbf{m})| \leqslant \prod_{j\leqslant\lceil h_n(i)\rceil} \exp\left(c_{\text{pei}} m_j\right) = \exp(c_{\text{pei}} m).
$$

Moreover, by (6.4.18), $m \geqslant \lceil h_n(i)\rceil$, and so $|\mathcal{M}_i(m)| \leqslant \binom{m+\lceil h_n(i)\rceil}{m} \leqslant 2^{2m} \leqslant e^{2m}$. Next, since the vertices in $\mathbf{x}$ are ordered, there are at most $\binom{n}{\lceil h_n(i)\rceil}$ many choices for the vector $\mathbf{x} \in \Lambda_n^{\lceil h_n(i)\rceil,<}$. Using these bounds in (6.4.19) and evaluating the geometric sum in m it follows for some $C > 0$ that

$$
\begin{aligned}
&\mathbb{P}\big(\exists A \in \mathcal{A}_{\text{large}}(i) : A \not\leftrightarrow_{\mathcal{G}_n} \cup_{H\in\mathfrak{H}_A} \partial_{\text{int}} H\big) && (6.4.21)\\
&\leqslant \binom{n}{\lceil h_n(i)\rceil} \sum_{m\geqslant c_{6.4.2} i^{-2}2^{-i/d} n} (1-p)^{m R_2^{(d-1)/d}} \exp\left((c_{\text{pei}}+2)m\right)\\
&\leqslant C\binom{n}{\lceil h_n(i)\rceil} \exp\Big(\big(c_{\text{pei}} + 2 - \log\big(\tfrac{1}{1-p}\big) R_2^{(d-1)/d}\big) c_{6.4.2} i^{-2} 2^{-i/d} n\Big),
\end{aligned}
$$

whenever $(1-p)^{R_2^{(d-1)/d}} e^{c_{\text{pei}}+2} < 1$ (see $R_2(\beta)$ in (6.4.5)). We recall from (6.4.3) that $h_n(i) = 2^{-i-3} i^{-2} n$, and using $\binom{n}{h} \leqslant (e \cdot n/h)^h$ it follows

$$
\begin{aligned}
\binom{n}{\lceil h_n(i)\rceil} &\leqslant (e 2^{i+3} i^2)^{i^{-2}2^{-i-3} n + 1}\\
&\leqslant \exp\Big(\big(i + 4 + 2\log(i)\big)\big(i^{-2} 2^{-i-3} n + 1\big)\Big), \qquad (6.4.22)
\end{aligned}
$$

where we also used that $2 < e$ to obtain the right-hand-side. Using this bound in the right-hand side of (6.4.21), we may compare the exponents. Let $i_\circ$ be the smallest $i \in \mathbb{N}$ such that for all $n \geqslant 1$ and all $i \in [i_\circ, \lceil \log_2(n) \rceil]$,

$$(i + 4 + 2\log(i))(i^{-2}2^{-i-3}n + 1) < c_{6.4.2} i^{-2} 2^{-i/d} n.$$

Then for $i \geqslant i_\circ$,

$$\begin{aligned} &\mathbb{P}\big(\exists A \in \mathcal{A}_{\text{large}}(i) : A \not\leftrightarrow_{\mathcal{G}_n} \cup_{H \in \mathfrak{H}_A} \partial_{\text{int}} H\big) \\ &\quad \leqslant C \exp\Big(\big(c_{\text{pei}} + 3 - \log\big(\tfrac{1}{1-p}\big) R_2^{(d-1)/d}\big) c_{6.4.2} i^{-2} 2^{-i/d} n \Big). \end{aligned} \tag{6.4.23}$$

Using that $i_\circ$ is a constant that only depends on d, (comparing the coefficients of $c_{6.4.2}n$ in (6.4.21) to (6.4.22)) we require that $R_2^{(d-1)/d} \log\big(\frac{1}{1-p}\big)$ is large enough so that for all $i \leqslant i_\circ$,

$$\begin{aligned} &(c_{\text{pei}} + 2) i^{-2} 2^{-i/d} + (i + 4 + 2\log(i))\big(i^{-2}2^{-i-3} + 1\big)/c_{6.4.2} \\ &\qquad \leqslant \tfrac{1}{2} R_2^{(d-1)/d} \log\big(\tfrac{1}{1-p}\big) i^{-2} 2^{-i/d}. \end{aligned}$$

In this case we obtain that for all $i \leqslant i_\circ$

$$\begin{aligned} &\mathbb{P}\big(\exists A \in \mathcal{A}_{\text{large}}(i) : A \not\leftrightarrow_{\mathcal{G}_n} \cup_{H \in \mathfrak{H}_A} \partial_{\text{int}} H\big) \\ &\quad \leqslant C \exp\big(-\tfrac{1}{2} \log\big(\tfrac{1}{1-p}\big) R_2^{(d-1)/d} c_{6.4.2} i^{-2} 2^{-i/d} n \big). \end{aligned} \tag{6.4.24}$$

We recall that $R_2 = \Theta(\beta^{1/d})$ from (6.4.5), so by combining (6.4.23) and (6.4.24) it follows for some $c' > 0$ that, whenever $\beta^{(d-1)/d^2} \log\big(\frac{1}{1-p}\big)$ is sufficiently large,

$$\begin{aligned} &\mathbb{P}\big(\exists A \in \mathcal{A}_{\text{large}}(i) : \partial_{\text{int}} A \not\leftrightarrow_{\mathcal{G}_n} \cup_{H \in \mathfrak{H}_A(i)} \partial_{\text{int}} H\big) \\ &\quad \leqslant C \exp\big(-c' \log\big(\tfrac{1}{1-p}\big) \beta^{(d-1)/d^2} i^{-2} 2^{-i/d} n \big). \end{aligned}$$

We apply this bound to all summands in (6.4.13), use that the terms are increasing in i, that the last term is $i = \lceil \log_2(n) \rceil$ (so $2^{-i/d} i^{-2} n \geqslant 2^{-1/d} (\log_2(n) + 1)^{-2} n^{(d-1)/d}$ for all i), and obtain for some $c_{6.4.4} > 0$ and n sufficiently large

$$\begin{aligned} &\sum_{i=i_\star+1}^{\lceil \log_2(n) \rceil} \mathbb{P}\big(\exists A \in \mathcal{A}_{\text{large}}(i) : \partial_{\text{int}} A \not\leftrightarrow_{\mathcal{G}_n} \cup_{H \in \mathfrak{H}_A(i)} \partial_{\text{int}} H\big) \\ &\qquad \leqslant \exp\big(-c_{6.4.4} \log\big(\tfrac{1}{1-p}\big) \beta^{(d-1)/d^2} (\log(n))^{-2} n^{(d-1)/d} \big). \end{aligned}$$

This finishes the proof of Lemma 6.2.10. □

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

Proof of Lemma 6.2.10. We recall the bound (6.4.9) – (6.4.10) on the probability on the event $\mathcal{E}_2 = \{\exists A \in \mathcal{A}_{\text{large}} : A \not\leftrightarrow_{\mathcal{G}_n} (\cup_{H \in \mathfrak{H}_A} \partial_{\text{int}} H)\}$, splitting it into two sums: one sum (6.4.9) for small principal holes, and one sum (6.4.10) for large principal holes. As an immediate corollary of Claims 6.4.3 and 6.4.4, we obtain for the constant $i_\star$ from (6.4.5) that

$$\mathbb{P}(\mathcal{E}_2) \leqslant i_\star 2^n \exp\big(-nc_{6.4.3}\big(\log\big(\tfrac{1}{1-p}\big)\beta^{(d-1)/d^2}/(1+\log_2(\beta^{1/d}))^2\big)\big) + \exp\Big(-\frac{n^{(d-1)/d}}{\log^{-2}(n)}c_{6.4.4}\log\big(\tfrac{1}{1-p}\big)\beta^{(d-1)/d^2}\Big).$$

If $\log\big(\frac{1}{1-p}\big)\beta^{(d-1)/d^2}$ is sufficiently large, the first term decays exponentially in n. Under this assumption, the second term on the right-hand side dominates the expression. The existence of $\beta_\star(p,d,\alpha)$ the follows for fixed p immediately. Similarly, whenever p is sufficiently close to 1, then $\log\big(\frac{1}{1-p}\big)\beta^{(d-1)/d^2}$ is already sufficiently large for $\beta = 1$, and hence $\beta_\star(p,d,\alpha) = 1$ can be set in this interval, finishing the proof of the final statement. □

Proof of Proposition 6.1.4. The proof is immediate from Claim 6.2.8 and Lemmata 6.2.9 and 6.2.10. The condition $n(\log n)^{-2d/(d-1)} \geqslant k$ implies that $k^{(d-1)/d} \leqslant n^{(d-1)/d}/(\log n)^2$, so that the error bound from (6.2.15) dominates the error bound from (6.2.16). The coefficient of $-k^{(d-1)/d}$ in the exponent can be made at least 1 by choosing either β sufficiently large or p sufficiently close to 1 and $\beta \geqslant 1$ so that the total coefficient $c_{6.2.9}\log(\frac{1}{1-p})\beta^{1/d} + c_{6.2.10}\log(\frac{1}{1-p})\beta^{(d-1)/d^2} > 1$ holds. □

6.5 PROOF OF THEOREM 6.1.2

We will verify the statements in Theorem 6.1.2, based on Proposition 6.1.4.

Upper bounds

The upper bound on the second-largest component in (6.1.2) follows immediately from Proposition 6.1.4, by substituting $k = A(\log(n))^{d/(d-1)}$ for some large constant $A = A(\delta)$. For the upper bound on the cluster-size decay and the lower bounds, we recall two statements from Chapter 5 which considers a more general class of percolation models where vertices have associated vertex marks. The model long-range percolation in Definition 6.1.1 hence forms a subclass of the model in Chapter 5 in which the vertex set is $\mathbb{Z}^d$ and all the vertex marks are identical to 1. Due to this, conditions that regard vertices with high vertex-marks in Chapter 5 are automatically satisfied for the long-range percolation model in Definition 6.1.1. We rephrase Proposition 5.5.1 to the setting of long-range percolation of Definition 6.1.1 by setting all vertex marks identical to 1 in Chapter 5.

Proposition 6.5.1 (Prerequisites for the upper bound)**.** *Consider supercritical long-range percolation with parameters* $\alpha > 1$, *and* $d \in \mathbb{N}$. *Assume that there exist* $\zeta, c, c' > 0$ *and a function* $g(k) = O(k^{1+c'})$ *such that for all* n, k *sufficiently large, whenever* $n \geqslant g(k)$, *it holds that*

$$\mathbb{P}\big(|\mathcal{C}_n^{(2)}| \geqslant k\big) \leqslant n^{c'} \exp\big(-ck^{\zeta}\big), \tag{6.5.1}$$

$$\mathbb{P}\big(|\mathcal{C}_n^{(1)}| \leqslant n^{c}\big) \leqslant n^{-1-c}. \tag{6.5.2}$$

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

$$\mathbb{P}\big(|\mathcal{C}_n(0)| \geqslant k, 0 \notin \mathcal{C}_n^{(1)}\big) \leqslant \exp\big(-(1/A)k^{\zeta}\big),$$

and

$$\frac{|\mathcal{C}_n^{(1)}|}{n} \xrightarrow{\mathbb{P}} \mathbb{P}\big(|\mathcal{C}(0)| = \infty\big), \qquad \text{as } n \to \infty.$$

We have just proved the prerequisite (6.5.1) for our case in Proposition 6.1.4 with $\zeta = (d-1)/d$, $c = 1$ and $c' = 2$. The other prerequisite (6.5.2) is a consequence of the following lemma ((6.5.3) below in particular). The second statement of the lemma, (6.5.4) will be needed for the lower bound of Theorem 6.1.2 shortly.

Lemma 6.5.2. *Consider long-range percolation in Definition 6.1.1 with* $\alpha > 1$ *and* $d \geqslant 2$. *For all* $p \in (0,1)$, *there exists* $\beta_\star = \beta_\star(p, d, \alpha) > 0$ *such that for all* $\beta \geqslant \beta_\star$ *there exists* $\rho > 0$ *such that for* n *sufficiently large*

$$\mathbb{P}\big(|\mathcal{C}_n^{(1)}| \geqslant \rho n\big) \geqslant 1 - \exp\big(-\rho n^{(d-1)/d}\big), \tag{6.5.3}$$

$$\mathbb{P}\big(|\mathcal{C}_n(0)| \geqslant \rho n\big) \geqslant \rho. \tag{6.5.4}$$

For each $d \geqslant 2$, *there exists* $p_d < 1$ *such that* $\beta_\star \leqslant 1$ *for all* $p \in (p_d, 1)$.

Proof. We start with showing (6.5.3). If p is sufficiently close to one, then we can do the following. Taking G_n as a realization of LRP in Λ_n, and retaining only the edges of G_n that are between nearest neighbor vertices in $\mathbb{Z}_1^d$, we obtain the classical iid nearest-neighbor Bernoulli percolation in Λ_n. Denote this graph by $G_n^{(nn)}$ and its largest component by $\mathcal{C}_n^{(1)}(G_n^{(nn)})$. Then since $\mathcal{E}(G_n^{(nn)}) \subseteq \mathcal{E}(G_n)$, for the sizes of the largest components it holds that $|\mathcal{C}_n^{(1)}| \geqslant |\mathcal{C}_n^{(1)}(G_n^{(nn)})|$. Since p is sufficiently close to 1, the surface order large devation result of [76, Theorem 1.1] applies to $\mathcal{C}_n^{(1)}(G_n^{(nn)})$, and (6.5.3) immediately follows.

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

$$m(n) := \left\lceil \frac{n^{1/d}}{\beta^{1/d}/(2\sqrt{d})} \right\rceil^d. \tag{6.5.5}$$

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

$$C(\beta) := \mathrm{Vol}(Q_1)/2 = r(n)^d/2 \in \left[\frac{\beta}{4(2\sqrt{d})^d}, \frac{\beta}{2(2\sqrt{d})^d}\right]. \tag{6.5.6}$$

Let us say that Q_i, Q_j are adjacent boxes if they share a $(d-1)$-dimensional face. Since the diameter of each box is at most $\beta^{1/d}/2$, by (5.1.3), if Q_i, Q_j are adjacent boxes,

$$\mathbb{P}\left(\{x, y\} \in \mathcal{E}(\mathcal{G}_n)\right) = p \qquad \text{for all } x \in Q_i \cap \mathbb{Z}^d, y \in Q_j \cap \mathbb{Z}^d, \tag{6.5.7}$$

independently of other edges. The same is true when $x, y \in Q_i \cap \mathbb{Z}^d$ both.

Let us denote the subgraph of $\mathcal{G}_n$ induced by the vertices in the box Q_i by $\mathcal{G}_n(Q_i)$. $\mathcal{G}_n(Q_i)$ then stochastically dominates an Erdős-Rényi random graph with $n_i \in [C(\beta), 4C(\beta)]$ vertices and edge probability p. Since p is constant, for any fixed $\varepsilon > 0$, by choosing β (and hence also $C(\beta)$) large enough depending on ε, the probability that $\mathcal{G}_n(Q_i)$ is connected is at least $1-\varepsilon$ by [35]. Further, the graphs $(\mathcal{G}_n(Q_i))_{i\leqslant m(n)}$ are *independent* since they are induced subgraphs of long-range percolation on vertices in disjoint boxes, and edges are present independently in $\mathcal{G}_n$ by Definition 6.1.1.

We define an auxiliary graph G. First we define the vertex set. Every box Q_i corresponds to a vertex v_i, for each $i \leqslant m(n)$, and two vertices v_i, v_j in G are adjacent if the corresponding boxes Q_i, Q_j are adjacent, i.e., they share a $(d-1)$-dimensional face. Similarly, one can define the 1-distance between any two vertices v_i, v_j by the length of the shortest path between

v_i, v_j via adjacent vertices. Hence, the vertices of G then form a box $\widetilde{\Lambda}_{m(n)}$ of volume $m(n)$ of $\mathbb{Z}_1^d$. This we call the re-normalised lattice.

Now we define the edge-set of G. We declare a vertex v_i of G *active* when $\mathcal{G}_n(Q_i)$ is connected. Edges of G will be only present between active (and adjacent) vertices. Assuming that two vertices v_i, $v_j \in G$ are adjacent and both active, we declare the edge between v_1 and v_2 *open*, equivalently, present in $\mathcal{E}(G)$, if there exist vertices $x \in Q_i \cap \mathbb{Z}^d, y \in Q_j \cap \mathbb{Z}^d$ with the edge $\{u, v\} \in \mathcal{E}(\mathcal{G}_n)$. Conditional on v_i, v_j being active, by (6.5.7), the nearest-neighbor edge between v_1 and v_2 is open with probability $1-(1-p)^{n_i n_j} \geqslant 1-(1-p)^{C(\beta)^2} \geqslant 1-\varepsilon$, where the last inequality holds for arbitrarily small $\varepsilon > 0$ by making $C = C(\beta)$ in (6.5.6) large enough. Different edges of $\mathcal{E}(G)$ are present conditionally independently given that the end-vertices are active.

Let us denote by H the induced graph obtained from G on active vertices and open edges $\mathcal{E}(G)$. By the observation above, the vertices $(v_i)_{i \leqslant m(n)}$ form a box $\widetilde{\Lambda}_{m(n)}$ of volume $m(n)$ of $\mathbb{Z}_1^d$. Then H stochastically dominates a site-bond percolation of $\mathbb{Z}_d^1$ in $\widetilde{\Lambda}_{m(n)}$.

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

Since ε can be chosen arbitrarily small, by [176, Theorem 0.0], the graph H therefore stochastically dominates iid nearest-neighbor bond percolation $G^\star$ on $\widetilde{\Lambda}_{m(n)}$ with parameter $p^\star$ that can also be made arbitrarily close to 1. Hence for the sizes of the largest connected components $|\mathcal{C}_n^{(1)}(H)| \geqslant |\mathcal{C}_n^{(1)}(G^\star)|$ holds. Thus, [76, Theorem 1.1] applies to $|\mathcal{C}_n^{(1)}(G^\star)|$, and so for some $c(\beta) > 0$ we obtain that using (6.5.5)

$$\mathbb{P}\left(|\mathcal{C}_n^{(1)}(H)| \geqslant \rho m(n)\right) \geqslant e^{-c m(n)^{(d-1)/d}} \geqslant 1 - e^{-c(\beta) n^{(d-1)/d}} \tag{6.5.8}$$

Since in each box Q_i that an active $v_i \in G$ corresponds to contains at least $C(\beta)$ vertices and the graphs $\mathcal{G}(Q_i)$ are connected, it holds deterministically that $|\mathcal{C}_n^{(1)}(\mathcal{G}_n)| \geqslant C(\beta)|\mathcal{C}_n^{(1)}(H)|$. This, combined with (6.5.8) implies (6.5.3).

We turn now to prove (6.5.4). Consider a smaller box $\Lambda_{2^{-d}n}$. Define then

$$Z_\ell := \sum_{x \in \Lambda_{2^{-d}n}} \mathbb{1}_{\{|\mathcal{C}_{2^{-d}n}(x)| \geqslant \ell\}}.$$

We argue that $\{\mathcal{C}_n^{(1)} \geqslant \ell\} \subseteq \{Z_\ell \geqslant \ell\}$. Indeed, if the largest component is at least of size ℓ then in Z_ℓ at least ℓ many indicators are 1. Then applying a Markov's inequality with $\ell = \rho 2^{-d} n$ followed by a union bound yields that

$$\begin{aligned}\mathbb{P}(\mathcal{C}_n^{(1)} \geqslant \rho 2^{-d} n) &\leqslant \mathbb{P}(Z_{\rho 2^{-d} n} \geqslant \rho 2^{-d} n) \\ &\leqslant \frac{\mathbb{E}[Z_{\rho 2^{-d} n}]}{\rho 2^{-d} n} \leqslant \frac{1}{\rho 2^{-d} n} \sum_{x \in \Lambda_{2^{-d} n}} \mathbb{P}\big(|\mathcal{C}_{2^{-d} n}(x)| \geqslant \rho 2^{-d} n\big).\end{aligned}$$

If for all $x \in \Lambda_{2^{-d} n}$ it would hold that $\mathbb{P}\big(|\mathcal{C}_{2^{-d} n}(x)| \geqslant \rho 2^{-d} n\big) \leqslant \rho/2$, then the right hand side would be at most $1/2$. This would then contradict (6.5.3) for $2^{-d} n$ in place of n.

Hence, there must exist $x \in \Lambda_{2^{-d} n}$ such that $\mathbb{P}\big(|\mathcal{C}_{2^{-d} n}(x)| \geqslant \rho 2^{-d} n\big) \geqslant \rho/2$. Let $x \in \Lambda_{2^{-d} n}$ be such a vertex. Then, by the translation invariance of the infinite model $\mathcal{G}_\infty$, looking at the component of the origin $\mathcal{C}_{2^{-d} n}^{(-x)}(0)$ inside the box $\Lambda_{2^{-d} n}(-x)$, it holds that

$$\mathbb{P}\big(|\mathcal{C}_{2^{-d} n}^{(-x)}(0)| \geqslant \rho 2^{-d} n\big) = \mathbb{P}\big(|\mathcal{C}_{2^{-d} n}(x)| \geqslant \rho 2^{-d} n\big).$$

However, the shifted box $\Lambda_{2^{-d} n}(-x) \subseteq \Lambda_n$ for any $x \in \Lambda_{2^{-d} n}$, and hence $\mathcal{C}_{2^{-d} n}^{(-x)}(0) \subseteq \mathcal{C}_n(0)$. Hence, we obtain

$$\mathbb{P}\big(|\mathcal{C}_n(0)| \geqslant \rho 2^{-d} n\big) \geqslant \mathbb{P}\big(|\mathcal{C}_{2^{-d} n}(x)| \geqslant \rho 2^{-d} n\big) \geqslant \rho/2.$$

Hence, (6.5.4) follows by adapting the constant ρ. □

Since both prerequisites of Proposition 6.5.1 are satisfied, this finishes the proof of the upper bounds of Theorem 6.1.2.

Lower bounds

For the lower bound we adapt the lower bound from Chapter 5, rephrased to the model of long-range percolation of Definition 6.1.1, by setting the vertex set to $\mathbb{Z}^d$ and all vertex marks to 1. Proposition 5.6.1 –the lower bound of cluster-size decay and second-largest component– requires that in a box of volume ℓ a linear sized (at least $\rho\ell$) giant component on vertices with marks in the interval $[1, \mathrm{polylog}(\ell)]$ exists, with probability at least $\rho > 0$. Since in long-range percolation all vertex marks are identical to 1, this requirement (5.6.3) turns into the requirement (6.5.9) below.

Proposition 6.5.3 (Prerequisites for the lower bound)**.** *Consider supercritical long-range percolation with parameters* $\alpha > 1 + 1/d$, $d \geqslant 2$, *and assume that* $\min\{p, p\beta^{\alpha}\} \in (0, 1)$. *Assume that there exists a constant* $\rho > 0$ *such that for all* n *sufficiently large,*

$$\mathbb{P}\big(|\mathcal{C}_n(0)| \geqslant \rho n\big) \geqslant \rho. \tag{6.5.9}$$

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

$$\mathbb{P}\big(|\mathcal{C}_n(0)| \geqslant k, 0 \notin \mathcal{C}_n^{(1)}\big) \geqslant \exp\big(-Ak^{(d-1)/d}\big). \tag{6.5.10}$$

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

$$\mathbb{P}\big(|\mathcal{C}_n^{(2)}| \leqslant \varepsilon(\log(n))^{d/(d-1)}\big) \leqslant n^{-\delta}. \tag{6.5.11}$$

Since Lemma 6.5.2 has just proved in (6.5.4) the requirement (6.5.9), and in Theorem 6.1.2 we have also assumed that $\min\{p, p\beta^{\alpha}\} < 1$, the lower bounds in Theorem 6.1.2 follow from Proposition 6.5.3, in particular, (6.5.11) implies the lower bound in (6.1.2), and after taking logarithm of both sides, (6.5.10) implies the lower bound in (6.1.2).

6.A PROOFS OF PRELIMINARY CLAIMS

Proof of Claim 6.2.3. Identify a path on $\mathbb{Z}_1^d$ with its vertex set. We define an equivalence class $\sim_{\mathcal{C},1}$ on the vertices of $\mathcal{C}$, where $x \sim_{\mathcal{C},1} y$ if and only if there is a 1-connected path π consisting of vertices of $\mathcal{C}$ that connects x and y (i.e., the edges of this path are not necessarily part of the edges of $\mathcal{G}$). We then define the blocks $A_1, A_2, \ldots, A_b$ as the equivalence classes of $\sim_{\mathcal{C},1}$. In other words, start from any vertex $x \in \mathcal{C}$ and define its block as all vertices that x is 1-connected to using paths only vertices of $\mathcal{C}$ (but the edges of $\mathbb{Z}_1^d$), and then we iterate this over all $x \in \mathcal{C}$, yielding the (different) blocks $A_1, A_2, \ldots, A_b$.

Each A_i is 1-connected since every pair of vertices in A_i is connected by a 1-connected path by the definition of $\sim_{\mathcal{C},1}$, i.e., A_i is a block. Further, if $i \neq j$ then $\|A_i - A_j\|_1 > 1$ must hold, since otherwise there would be a 1-connected path from some $x \in A_i$ to some $y \in A_j$, and that would contradict $x \not\sim_{\mathcal{C},1} y$. Uniqueness of this decomposition follows because $\sim_{\mathcal{C},1}$ is an equivalence relation.

We show that the block graph $\mathcal{H}_{\mathcal{G}}((A_i)_{i \leqslant b})$ is connected. Suppose otherwise. This means that there is a subset of blocks whose union is not connected to the union of all the other blocks (w.l.o.g. we may assume that for the first k blocks for some $k \in [1, b-1]$ this happens, i.e.,

$(\cup_{i\leqslant k}A_i) \not\leftrightarrow_{\mathcal{G}} (\cup_{i>k}A_i)$. However, this contradicts that $\mathcal{C}$ is a component of $\mathcal{G}$. □

Proof of Claim 6.2.5. We show that $\widetilde{\partial}_{\text{int}}\bar{B} \subseteq \widetilde{\partial}_{\text{int}}B$. We argue by contradiction. Assume that there exists $x \in \widetilde{\partial}_{\text{int}}\bar{B} \setminus \widetilde{\partial}_{\text{int}}B$. Since $\widetilde{\partial}_{\text{int}}\bar{B} \subseteq \bar{B} = B \cup (\cup_{H\in\mathfrak{H}_B} H)$ and B is disjoint from $\bar{B} \setminus B$, there are two cases. Either $x \in B \setminus \widetilde{\partial}_{\text{int}}B$ or $x \in (\cup_{H\in\mathfrak{H}_B} H) \setminus \widetilde{\partial}_{\text{int}}B$.

For the first case assume that $x \in B \setminus \widetilde{\partial}_{\text{int}}B$. Then all its $\mathbb{Z}_1^d$-neighbouring vertices were also in B by Definition 6.2.4 of the interior boundary. Thus x is surrounded by B, and hence $x \in \bar{B}$. Similarly, the neighbouring vertices are also in $\bar{B}$, contradicting that $x \in \widetilde{\partial}_{\text{int}}\bar{B}$.

For the second case assume that $x \in (\cup_{H\in\mathfrak{H}_B} H)$. Then x was surrounded by B, but then also all its $\mathbb{Z}_1^d$-neighbouring vertices were either a member of B or surrounded by B, contradicting again that $x \in \widetilde{\partial}_{\text{int}}\bar{B}$.

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

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

$$\begin{aligned} \|\bar{B}_1 - \bar{B}_2\|_1 = \min_{x_1\in\bar{B}_1\setminus B_1, x_2\in\bar{B}_2\setminus B_2} \big\{ &\|B_1 - B_2\|_1, \\ &\|x_1 - B_2\|_1, \|B_1 - x_2\|_1, \\ &\|x_1 - x_2\|_1 \big\}. \end{aligned} \tag{6.A.1}$$

Each 1-connected path from $x_i \in \bar{B}_i \setminus B_i$ to any $y \notin \bar{B}_i$ must cross a vertex in $\widetilde{\partial}_{\text{int}}\bar{B}_i$. Consequently,

$$\begin{aligned} \|x_1 - B_2\|_1 &\geqslant \|x_1 - \widetilde{\partial}_{\text{int}}\bar{B}_1\|_1 + \|\widetilde{\partial}_{\text{int}}\bar{B}_1 - B_2\|_1 \\ &\geqslant \|x_1 - \widetilde{\partial}_{\text{int}}\bar{B}_1\|_1 + \|\widetilde{\partial}_{\text{int}}B_1 - B_2\|_1 \geqslant 2, \end{aligned}$$

where the second inequality follows since $\widetilde{\partial}_{\text{int}}\bar{B}_1 \subseteq \widetilde{\partial}_{\text{int}}B_1$ by part (i), and the third inequality since by 1-disconnectedness of B_1 and B_2 the second term on the right-hand side is at least two. The third and fourth term in the minimum in (6.A.1) can be bounded similarly. It follows that $\bar{B}_1$ and $\bar{B}_2$ are 1-disconnected.

The fourth statement is immediate from [76, Lemma 2.1] which states that the interior and exterior boundaries with respect to $\mathbb{Z}^d$ of any $*$-connected set are $*$-connected as well.

We turn to the last statement, and show that for any hole H of a block B, $H = \bar{H}$, i.e., that H does not contain holes. This is true since H was formed as a maximal 1-connected subset of the vertices in $\Lambda_n \setminus B$ surrounded by B, see below (6.2.2) in Definition 6.2.1. So if there were a hole J inside H, then J must be part of B, which would then contradict the 1-connectedness of B, since $\|J - \partial_{\text{ext}} H\| \geqslant 2$ as H fully surrounds J. The fact that the interior and exterior boundaries of a hole are $*$-connected follows now from Part (iv). □

Proof of Claim 6.2.6. The proof is inspired by an argument by Deuschel and Pisztora [76, Proof of (A.3)]. The inequalities with $(\star)$ in (6.2.5) follow by standard isoperimetric inequalities, but we will also derive them below.

We will first show the bounds for ∂_{int} and $\widetilde{\partial}_{\text{int}}$. At the end of the proof we adjust it to ∂_{ext} and $\widetilde{\partial}_{\text{ext}}$. We start by showing that there exists $\delta' > 0$ such that $\partial_{\text{int}} A \geqslant \delta' \widetilde{\partial}_{\text{int}} A$ for all $A \subseteq \Lambda_n$ of size at most $3n/4$. Fix such a set $A \subseteq \Lambda_n$. We recall an inequality related to the isoperimetric inequality by Loomis and Whitney [178, Theorem 2]. For a set $A \subseteq \Lambda_n$, let S_i denote the projection of A onto the i-th coordinate hyperplane. That is, for a vertex with coordinates $x = (x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_d)$ we define $\pi_i x := (x_1, \dots, x_{i-1}, 0, x_{i+1}, \dots, x_d)$. Then

$$|A|^{d-1} \leqslant \prod_{i \in [d]} |S_i|. \tag{6.A.2}$$

Let $i_\star$ be the dimension that contains the largest projected set $S_{i_\star}$ (ties broken arbitrarily), so that as a result of (6.A.2),

$$|S_{i_\star}| \geqslant |A|^{(d-1)/d}. \tag{6.A.3}$$

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

$$\pi_\star^\uparrow := \{y \in A : \pi_\star(y) = s\}$$

We now call a vertex $s \in S$ a *fiber* if there is no vertex in $\partial_{\text{int}} A$ that projects to s via $\pi_\star$. Formally we define the *fibers* of S as

$$F := \{s \in S : \nexists y \in \partial_{\text{int}} A \text{ with } \pi_\star(y) = s\}. \tag{6.A.4}$$

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

$$|\pi_\star^\uparrow(s) \cap \widetilde{\partial}_{\text{int}}\Lambda_n| = |\pi_\star^\uparrow(s) \cap \widetilde{\partial}_{\text{int}}A| = 2, \qquad \forall s \in F.$$

The pre-image of vertices in F does not contribute to $\partial_{\text{int}}A$ by their definition in (6.A.4). By the same definition, the pre-image of each vertex $z \in S \setminus F$ contains at least one vertex in $\partial_{\text{int}}A$. A similar argument as the one for fibers shows that the pre-image of each vertex $z \in S \setminus F$ contains at least two vertices in $\widetilde{\partial}_{\text{int}}A$. Formally

$$|\pi_\star^\uparrow(z) \cap \partial_{\text{int}}A| \geqslant 1 \quad \text{and} \quad |\pi_\star^\uparrow(z) \cap \widetilde{\partial}_{\text{int}}A| \geqslant 2 \qquad \forall z \in S \setminus F.$$

The difference between the above two intersections is the number of vertices that $A \cap \mathcal{L}_s \cap \widetilde{\partial}_{\text{int}}\Lambda_n$ contains. We obtain

$$|\pi_\star^\uparrow(z) \cap \widetilde{\partial}_{\text{int}}A| - |\pi_\star^\uparrow(z) \cap \partial_{\text{int}}A| \leqslant 2 \qquad \forall z \in S \setminus F.$$

We characterise $z \in S \setminus F$ according to this difference. For $i \in \{0, 1, 2\}$ we define

$$(S \setminus F)_i := \big\{ z \in S \setminus F : |\pi_\star^\uparrow(z) \cap \widetilde{\partial}_{\text{int}}A| - |\pi_\star^\uparrow(z) \cap \partial_{\text{int}}A| = i \big\}.$$

Then

$$\begin{aligned} |\widetilde{\partial}_{\text{int}}A| &= \sum_{s \in F} |\pi_\star^\uparrow(s) \cap \widetilde{\partial}_{\text{int}}A| + \sum_{i \in \{0,1,2\}} \sum_{z \in (S \setminus F)_i} |\pi_\star^\uparrow(z) \cap \widetilde{\partial}_{\text{int}}A| \\ &= 2|F| + \sum_{i \in \{0,1,2\}} \sum_{z \in (S \setminus F)_i} |\pi_\star^\uparrow(z) \cap \widetilde{\partial}_{\text{int}}A| \\ &= 2|F| + \sum_{i \in \{0,1,2\}} \sum_{z \in (S \setminus F)_i} (|\pi_\star^\uparrow(z) \cap \partial_{\text{int}}A| + i). \end{aligned}$$

Now we consider the ratio of the boundaries, i.e.,

$$\begin{aligned} &\frac{|\partial_{\text{int}}A|}{|\widetilde{\partial}_{\text{int}}A|} \\ &= \frac{\sum_{i \in \{0,1,2\}} \sum_{z \in (S \setminus F)_i} |\pi_\star^\uparrow(z) \cap \partial_{\text{int}}A|}{2|F| + |(S \setminus F)_1| + 2|(S \setminus F)_2| + \sum_{i \in \{0,1,2\}} \sum_{z \in (S \setminus F)_i} |\pi_\star^\uparrow(z) \cap \partial_{\text{int}}A|}. \end{aligned}$$

Taking the set sizes $|(S \setminus F)_i|, |F|$ fixed, it is elementary to see that the ratio is increasing in the summands of the double sum, i.e., its minimal value is attained when all summands are minimal. Now we use that each of the summands is at least 1, and obtain

$$\frac{|\partial_{\text{int}}A|}{|\widetilde{\partial}_{\text{int}}A|} \geqslant \frac{|S \setminus F|}{2|F| + |(S \setminus F)_1| + 2|(S \setminus F)_2| + |S \setminus F|}.$$

We bound the denominator from above, i.e.,

$$|F| + |S \setminus F| + |F| + |(S \setminus F)_1| + |(S \setminus F)_2| + |(S \setminus F)_2| \leqslant 3|S|,$$

to obtain

$$\frac{|\partial_{\text{int}}A|}{|\widetilde{\partial}_{\text{int}}A|} \geqslant \frac{|S \setminus F|}{3|S|}.$$

We focus on the ratio on the right-hand side. For each vertex $s \in F$, there are $n^{1/d}$-many vertices in A that are projected onto it (namely $\mathcal{L}_s$). Using also $|S| \geqslant |A|^{(d-1)/d}$ by (6.A.3) and $n \geqslant (4/3)|A|$, we get

$$|A| \geqslant |F|n^{1/d} = \frac{|F|}{|S|}|S|n^{1/d} \geqslant \frac{|F|}{|S|}|A|^{(d-1)/d}n^{1/d} \geqslant (4/3)^{1/d}\frac{|F|}{|S|}|A|.$$

After rearranging we obtain $|F| \leqslant (3/4)^{1/d}|S|$, and

$$|\partial_{\text{int}}A| \geqslant \big(1 - (3/4)^{1/d}\big)|\widetilde{\partial}_{\text{int}}A|/3$$

follows. The second inequality in (6.2.5) follows immediately from (6.A.3), since each projected vertex corresponds to at least two boundary vertices with respect to $\mathbb{Z}^d$, i.e., $|\widetilde{\partial}_{\text{int}}A| \geqslant 2|S| \geqslant 2|A|^{(d-1)/d}$.

We turn to the inequality concerning $\partial_{\text{ext}}A$ and $\widetilde{\partial}_{\text{ext}}A$ in (6.2.5). The inequality with $(\star)$ in (6.2.5) holds for the same reason as for $\partial_{\text{int}}A$. To obtain a lower bound on the ratio $|\partial_{\text{ext}}A|/|\widetilde{\partial}_{\text{ext}}A|$, we use that each exterior boundary vertex is within distance one from an interior boundary vertex, which holds for both for ∂ and $\widetilde{\partial}$. Since each vertex has at most $2d$ vertices within distance one, it follows that

$$\begin{aligned} |\partial_{\text{ext}}A| \geqslant \frac{1}{2d} \cdot |\partial_{\text{int}}A| &\geqslant \frac{1}{2d} \cdot \frac{1 - (3/4)^{1/d}}{3} \cdot |\widetilde{\partial}_{\text{int}}A| \\ &\geqslant \frac{1}{2d} \cdot \frac{1 - (3/4)^{1/d}}{3} \cdot \frac{1}{2d} \cdot |\widetilde{\partial}_{\text{ext}}A|, \end{aligned}$$

and the proof is finished for $\delta = (2d)^{-2}\big(1 - (3/4)^{1/d}\big)/3$. □

Proof of Lemma 6.2.7. We first show that there exists $c'_{\mathrm{pei}} > 0$ such that for all $x \in \mathbb{Z}^d$ and $m \in \mathbb{N}$

$$|\{A \subseteq \Lambda_n : A \ni x, |A| = m, A \text{ is } *\text{-connected}\}| \leqslant \exp(c'_{\mathrm{pei}} m). \qquad (6.A.5)$$

Let A be in the set on the left-hand side. Since A is $*$-connected, the induced subgraph $\mathbb{Z}^d_\infty[A]$ contains a spanning tree containing x, which can be associated to a walk on the spanning tree (for example, walking through the tree in depth-first order), visiting each vertex in A at most twice. Since the degree of any vertex is $3^d - 1$ in $\mathbb{Z}^d_\infty$, the walk has at most $3^d - 1$ options at each step for its next vertex, and has length at most $2m$. This shows (6.A.5).

For (6.2.6), we observe that each set A without holes ($A = \bar{A}$) can be uniquely reconstructed from its interior boundary $\widetilde{\partial}_{\mathrm{int}} A$ (a vertex is in $\bar{A} \setminus \widetilde{\partial}_{\mathrm{int}} A$ iff it is surrounded by $\widetilde{\partial}_{\mathrm{int}} A$), which is $*$-connected [76, Lemma 2.1]. Since we assume $|\widetilde{\partial}_{\mathrm{int}} A| = m$, the isoperimetric inequality (6.2.5) ensures that $|A| \leqslant C_1 m^{d/(d-1)}$ for some $C_1 > 0$ for all $m \in \mathbb{N}$. This interior boundary must either contain x or surround x as defined in Definition 6.2.4.

We claim that there is a constant $C > 0$ such that $\|x - \widetilde{\partial}_{\mathrm{int}} A\|_2 \leqslant Cm^{1/(d-1)}$ for all x, A with $A \ni x$ and $A = \bar{A}$. Indeed, suppose otherwise. Then, on $\mathbb{Z}^d$, vertices in the Euclidean ball of radius $Cm^{1/(d-1)}$ around x would be contained fully in A (without containing a vertex of $\widetilde{\partial}_{\mathrm{int}} A$). This would mean, for some dimension-dependent constant c_d, that $|A| \geqslant c_d (Cm^{1/(d-1)})^d$, which contradicts that $|A| \leqslant m^{d/(d-1)}$ by Claim 6.2.6 when C is chosen sufficiently large.

Hence, we may find a vertex $y \in \widetilde{\partial}_{\mathrm{int}} A \cap \mathrm{Ball}(Cm^{1/(d-1)}, x)$ where the latter set denotes the Euclidean ball of radius $Cm^{1/(d-1)}$ around x. Then, since $\widetilde{\partial}_{\mathrm{int}} A$ is a $*$-connected set of size m, (6.A.5) ensures that the number of possible sets S that may form $\widetilde{\partial}_{\mathrm{int}} A$ is $\exp(c_{\mathrm{pei}} m)$. Summing over the possible choices of $y \in \mathrm{Ball}(Cm^{1/(d-1)}, x)$, we arrive at

$$|\{A \in \mathcal{A} : A \ni x, A = \bar{A}, |\widetilde{\partial}_{\mathrm{int}} A| = m\}| \leqslant c_d C^d m^{d/(d-1)} \exp(c'_{\mathrm{pei}} m).$$

The result follows by absorbing the factor $c_d C m^{d/(d-1)}$ into the constant c_{pei}. $\square$

Part III

Agent-based modelling: Infection spreading

7 NOT ALL INTERVENTIONS ARE EQUAL FOR THE SECOND PEAK

Based on [152]
Not all interventions are equal for the second peak,
J. Jorritsma, T. Hulshof, J. Komjáthy,
Chaos, Solitons & Fractals, 2020 (139).

7.1 INTRODUCTION

When recovering from a disease grants temporary immunity against it, it can happen that an epidemic dies out locally, but survives elsewhere, returning at a later point in time. We observe a "second peak". A second peak can also happen when interventions are effectively applied to slow down the spread of a disease locally, but are then lifted. This phenomenon has a clear geometric component. Standard compartmental models for epidemic curves are inherently a-geometric, because they assume a perfectly mixed population. Although there are variants of standard compartmental models that do allow for geometric influences [80, 120, 224], we focus here on agent-based epidemiological models that allow for a natural embedding of the population in a geometric space. As a result of our modelling choices, the underlying contact network cannot be approximated by branching processes as in [17, 18, 19, 93], and the epidemic cannot be approximated well by a system of differential equations, making explicit calculations less tractable than in the mentioned papers.

In this chapter, we study this geometric effect empirically. We conduct a simulation study about *the effect of temporary immunity* on the spread of a disease on various agent-based models. Immunity being only temporary has been observed for various types of coronaviruses, see [95] and references therein. When this simulation study was conducted, it was still debated whether the COVID-19 epidemic provides lasting or temporal immunity, see e.g. [159]. Besides temporary immunity, we investigate the effect of three different interventions:

- Intervention $\mathcal{I}_{\text{soc}}$: social distancing,

- Intervention $\mathcal{I}_{\text{trav}}$: traveling restrictions, and
- Intervention $\mathcal{I}_{\text{deg}}$: limiting the number of social contacts.

Our simulation results may serve as a qualitative indication of possible outcomes of epidemic spread and intervention strategies for highly infectious diseases, such as the COVID-19 pandemic. It also gives qualitative predictions about the ways an epidemic might (or might not) return, depending on the average duration and level of temporary immunity. The focus is twofold: first, on understanding how the underlying space affects the outcome of the simulations, and second, to see and compare the effect of the three intervention strategies on the pandemic. For the first goal, we compare four scenarios with various underlying geometries:

- *Scenario* $\mathcal{S}_{\text{grid}}$*:* The underlying network is "purely geometric", ignoring long-range connections. For this scenario we use the square nearest-neighbour torus $\mathbb{Z}_n^2$ as the base graph.
- *Scenario* $\mathcal{S}_{\text{cm}}$*:* The underlying network is a "mean-field network", ignoring the spatial component. For this scenario we use the *configuration model,* which can mimic the local statistical properties of real human contact networks, such as degree distributions.
- *Scenario* $\mathcal{S}_{\text{girg}}$*:* The underlying network is a mixture between a geometric and mean-field network. For this we use *Geometric Inhomogeneous Random Graphs*, which possess geometric features and can match the statistical properties of real human contact networks.
- *Scenario* $\mathcal{S}_{\text{ode}}$*:* The epidemic is modeled in a "mean-field continuous space", ignoring the spatial component *and* approximating the discrete population by a continuum. For this we use *systems of ordinary differential equations* (also called: a compartmental model), which are currently the most popular tool for epidemic curve modelling.

Our aim is to compare these models, and see what the effect is of considering more realistic representations of the underlying space and spreading mechanisms. We emphasize that this chapter provides qualitative estimates, not quantitative ones. As a result, the study here might underpin or support certain intervention strategies more than others, but we refrain from (and see no justification for) using these simplified models to make numerical predictions with respect to the COVID-19 outbreak.

Organisation.

We first introduce the epidemic model and define the four underlying networks (corresponding to the four scenarios). In Section 7.3 we present the results of the first experiment that compares the four different scenarios. After that, in Section 7.4, we elaborate on the intervention methods and present the results of a second experiment which compares these intervention methods. We remark that this chapter focuses on the main results from the paper [152]. For a more detailed and quantitative discussion of the simulations and results, including extra figures that confirm the results, we refer to the original paper [152].

7.2 MODEL DESCRIPTION: EPIDEMICS WITH TEMPORARY IMMUNITY

We choose the simplest possible model that shows the behaviour that we would like to observe: the effect of temporary immunity. Other states could easily be added, such as an incubation period (exposed state) (as e.g. in [170]), deaths, asymptomatic cases, etc. The model is similar to the one studied in [62]. For simplicity, we describe the dynamics in *discrete time*. A continuous-time version is analogous and shows the same qualitative behaviour. The spreading process changes at discrete time steps, $t = \{0, 1, 2, \dots\}$. We fix the network G in advance and think of nodes in the network as individuals. Each node in the network can be in three possible states: *susceptible (S), infected (I) or temporarily immune (R)*. The temporarily immune state is often called temporarily *removed*, hence the abbreviation (R). The neighbours of a node u are nodes with a direct connection (also called link, or edge) to u. A connection may correspond to a friendship or an acquaintance, or simply a contact event. The discrete time dynamics between the three states are described as follows (see also Figure 9):

- *Infecting:* Each infected node, while being infectious, *infects* each of its neighbours within the network with probability β at every time-step. Infections to different neighbours happen independently.
- *Healing:* When infectious, each node *heals* with probability γ at every time-step, independently of other nodes, and independently of infecting other nodes. The average infectious period of an infected node is $1/\gamma$ time-steps. Upon healing, the node becomes temporarily immune.

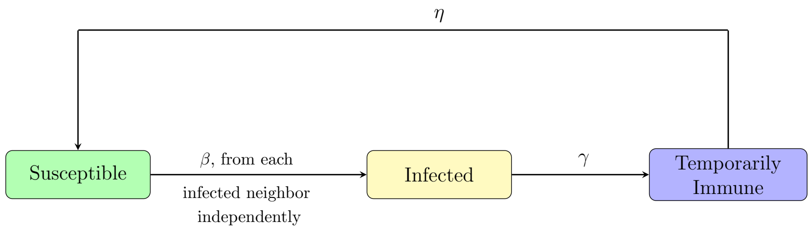


Figure 9: Schematic diagram of node states and their transition probabilities. A susceptible node u may only become infected when it has at least one infected neighbour. Each of the infected neighbours infects u with probability β, independently, in each step.

- *Losing immunity:* Each temporarily immune node loses its immunity with probability η at every time-step, independently of other nodes. The average immune period of a node is thus $1/\eta$ time-steps. After losing immunity, the node becomes susceptible again.

We run these dynamics on three different networks, obtaining the scenarios $\mathcal{S}_{\mathrm{cm}}$, $\mathcal{S}_{\mathrm{girg}}$, and $\mathcal{S}_{\mathrm{grid}}$, and adapt the dynamics to obtain the continuous model $\mathcal{S}_{\mathrm{ode}}$. We introduce the four scenarios, and afterwards explain the intervention methods.

7.2.1 *Scenario $\mathcal{S}_{\mathrm{grid}}$: Lattice network models*

Arguably the simplest geometric network model is the nearest-neighbour lattice. To avoid boundary effects, and make the network completely homogeneous, we shall study epidemics on the torus $\widetilde{\mathbb{Z}}_n^2$. For simplicity, we assume that $\sqrt{n} \in \mathbb{N}$.

Definition 7.2.1 (Nearest-neighbour tori)**.** Arrange nodes on a square grid, and label them with two coordinates $u = (x, y)$, for $x, y \in \{1, 2 \ldots, \sqrt{n}\}$. To obtain $\mathbb{Z}_n^2$, connect $u = (x_1, y_1)$ to $v = (x_2, y_2)$ when either $x_1 = x_2$ and $|y_1 - y_2| = 1 \mod \sqrt{n}$ or $y_1 = y_2$ and $|x_1 - x_2| = 1 \mod \sqrt{n}$. To obtain $\widetilde{\mathbb{Z}}_n^2$, connect $u = (x_1, y_1)$ to $v = (x_2, y_2)$ when $|x_1 - x_2| \leqslant 1 \mod n$ and $|y_1 - y_2| \leqslant 1 \mod \sqrt{n}$. The networks have n nodes.

Lattice models are homogeneous and the typical graph distance within the network is large, it grows polynomially with the number of nodes. We recall that the typical graph distance $d_n^{(G)}(u, v)$ denotes the number of connections on the shortest path between nodes u and v that are sampled uniformly at random from the graph on n vertices. We have

$$d_n^{(G)}(u,v) = \Theta(\sqrt{n}). \tag{7.2.1}$$

7.2.2 *Scenario $\mathcal{S}_{\mathrm{cm}}$: The configuration model and other a-geometric network models*

A-geometric random network models are often used as null-models to compare network data to purely random networks. In Section 1.2.1, we described a commonly used model, the configuration model, and mention rank-1 inhomogeneous random graphs as alternative that behaves qualitatively similarly. The main advantage of these simple models is that they can mimic the degree distribution of real-life networks. We choose the *erased* configuration model for baseline comparison, which is obtained from the Configuration Model from Definition 1.2.1 by removing any possible multi-edges and self-loops to obtain a simple graph.

While the configuration model easily accommodates power-law node degrees, and has the small-world property, it neither contains communities nor clustering [82, 217]. Communities are parts of the network that have significantly more connections towards each other than the same number of randomly selected nodes. One could argue that clustering and communities in human contact networks often arise due to spatial effects: people living nearby tend to know each other with higher probability. The following network accommodates for clustering and communities, while also keeping node degree variability high.

7.2.3 *Scenario $\mathcal{S}_{\mathrm{girg}}$: Geometric Inhomogeneous Random Graphs*

For our last class of networks we run the SIRS epidemic on a mixture of pure geometric and purely random network models. For this we use a general Geometric Inhomogeneous Random Graph (GIRG) model, that is a state-of-the-art model for real-world social and technological networks, embedded in geometric space. The presence of underlying geometry underpins the model: individuals are embedded in space, just like in real life, allowing for local community structures to be present in the contact network, leading to strong clustering [40, 217]. The GIRG incorporates edges bridging spatial distance on all scales, ranging from short to long-range edges, as well as a high variability of node degree: in fact, the model is *scale-free* in two respects: both in spatial distance that edges cover, and in node-degree variability [40].

Contact- or activity networks of humans have been found to show similar behaviour, including heavy-tailed degree distributions, strong clustering, and community structures [7, 32, 195, 200, 202], as well as heavy-tailed distance-distribution for edges, see references in [140].

The kernel-based spatial random graphs described in Section 1.3 were amongst others developed to incorporate these features and contain geometric inhomogeneous random graphs as a special case [7, 32, 83, 200, 198]. We formally define GIRGs. We denote by $x \wedge y$ the minimum of two numbers x, y.

Definition 7.2.2 (Geometric Inhomogeneous Random Graph (GIRG))**.** Fix $n \geqslant 1$ the number of nodes, and a dimension $d \geqslant 1$. Assign to each node $u \in \{1, 2, \ldots, n\}$ an i.i.d. *fitness* $w_u \geqslant 1$ such that $\mathbb{P}(w_u \geqslant w) = w^{-(\tau-1)}$ for some $\tau > 1$, and a *spatial location* x_u chosen uniformly in a d-dimensional box of volume n. Fix $\alpha > 0$. For any pair of nodes u, v with fixed w_u, w_v, x_u, x_v, connect them by an edge with probability

$$\mathbb{P}(u \text{ is connected by an edge to } v) \asymp \left(\frac{w_u w_v}{\|x_u - x_v\|^d}\right)^{\alpha} \wedge 1, \qquad (7.2.2)$$

where $\|\cdot\|$ denotes the Euclidean norm, independently of other edges. Throughout the chapter, we set the dimension $d = 2$ for all considered graphs.

GIRGs have a natural interpretation: the fitness expresses the ability of nodes to have many connections, and α is the *long-range* parameter (the smaller α is, the more the model favors longer connections). This gives an intuitive way of modelling travel restrictions: increasing α, decreases the probabilities of long edges. Indeed, in (7.2.2), for two nodes u, v with locations x_u, x_v, the ratio

$$R_{u,v} := \frac{w_u w_v}{\|x_u - x_v\|^2}$$

is raised to the power α, so as long $R_{u,v} < 1$, increasing α reduces the chance of a long connection being present. When $\tau > 3$, the node fitnesses are, typically, not high enough[1] so that, typically, $R_{u,v} < 1$, hence the intervention of increasing α is very effective. The typical distance increases from poly-logarithmic to polynomial in the network size. On the other

1 Node fitnesses are less than $\sqrt{n}$ for all nodes, with high probability.

hand, when $\tau \in (2,3)$, node pairs with ratio $R_{u,v} > 1$ happen frequently enough (even on global scale, nodes with fitness $\Omega(\sqrt{n})$ are present), hence, increasing α turns out to be ineffective in reducing the typical distance in the network, since many long-range edges remain.

The parameter space of GIRG is rich enough to model many desired features observed in real networks.

(a) Extreme variability of the number of neighbours (degrees). In networks this corresponds to the presence of a few individuals with extreme influence on spreading processes, the *hubs or superspreaders*. Mathematically, this corresponds to the empirical distribution of node degrees following a *power-law*, as in (1.1.1) above. Hence, we choose a power-law fitness distribution for the fitnesses.

(b) Connections present on all length-scales. The abundance of long-range connections is tuned by α in (7.2.2): the smaller α, the more likely are long-range connections.

(c) Small and ultra-small distances. The power-law exponent τ and the long-range parameter α tune the typical graph distance $d_n^{(G)}(u,v)$, see [30, 39, 63, 66, 116]:

$$d_n^{(G)}(U_n, V_n) = \begin{cases} \Theta(\log\log n) & \text{when } \tau \in (2,3), \alpha > 1 \\ O\big((\log n)^{\zeta}\big) & \text{when } \tau > 3, \alpha \in (1,2) \\ \Omega(\sqrt{n}) & \text{when } \tau > 3, \alpha > 2. \end{cases} \tag{7.2.3}$$

Comparing this to the typical distance in the configuration model and to the lattice models in (7.2.1), one sees that geometry and long-range connections play a role with respect to distances when $\tau > 3$, and the model interpolates between the small-world configuration model and the lattice.

(d) Strong clustering and local communities. Empirically, clustering quantifies the effect commonly known as "a friend of a friend is also likely to be my friend". Mathematically, clustering means the presence of *triangles* in the network. Communities and clustering are naturally present in GIRG, because the model favors connections between nodes that are close to each other in space.

For the spread of information or infections in such networks, it is known that large-degree nodes and many triangles have opposing effects. On the

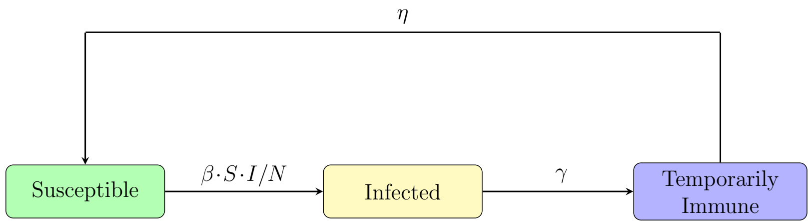


Figure 10: Schematic diagram of states and their transition rates in the continuous compartmental model. Infected individuals make contact at rate β to randomly chosen individuals, as a result, the number of susceptibles increases at rate β·S·I/N. This is a mean-field approximation of the graph model in Section 7.2, see Figure 9.

.

one hand, nodes of large degree (also called hubs, super-spreaders, or influencers) contribute to fast dissemination, and foster explosive propagation of information or infections [84, 112, 207, 209]. On the other hand, clustering and community structures provide natural barriers that slow down the process, while long-range edges accelerate the spread [16, 141, 148, 157, 184].

7.2.4 *Scenario $\mathcal{S}_{\text{ode}}$: SIRS compartmental model*

For baseline comparison to continuous epidemic models, we use a continuous SIRS model, where there are only three possible states, *susceptible (S), infected (I) or temporarily immune (R)*, in a total population of N individuals. Susceptible individuals may become infected via a contact to infectious individuals, while infectious individuals heal and thus become temporarily immune. The assumption is that individuals make contact with a random individual at rate β. Immune individuals loose their immunity at a certain rate η and become susceptible again. The deterministic ordinary differential equation (ODE) model is governed by the three main parameters:

- β: rate of infecting a susceptible individual when being infected,
- γ: rate of healing, and becoming temporarily immune,
- η: rate of losing temporal immunity and becoming susceptible again.

See Figure 10 for a schematic diagram. We assume the population size is N and that each infected individual can infect any susceptible person (the underlying graph is the complete graph), denote by $S(t), I(t), R(t)$ the number of susceptible, infected and temporarily immune individuals at time t, then the corresponding ordinary differential equation becomes:

$$\frac{dS}{dt} = -\frac{\beta IS}{N} + \eta R$$
$$\frac{dI}{dt} = +\frac{\beta IS}{N} - \gamma I$$
$$\frac{dR}{dt} = +\gamma I - \eta R$$

For the infection to survive, $\beta > \gamma$ is necessary, otherwise $\lim_{t\to\infty} I(t) = 0$. When $\beta > \gamma$ holds, this ODE has a stationary state, which is explicitly computable:

$$S(\infty) = N\frac{\gamma}{\beta}, \quad I(\infty) = N\frac{\eta}{\eta+\gamma}\Big(1-\frac{\gamma}{\beta}\Big), \quad R(\infty) = N\frac{\gamma}{\eta+\gamma}\Big(1-\frac{\gamma}{\beta}\Big).$$

Observe that the equilibrium proportion of infected is roughly linear in η, as long as $\eta \ll \gamma$. One can also compute the basic reproduction number, $\mathcal{R}_0 = \beta/\gamma$, and observe that $R(\infty) = I(\infty) \cdot (\gamma/\eta)$. Qualitatively, the ODE exhibits either (sub)critical behaviour, implying extinction, or supercritical behaviour, as described in Section 7.3.2 below.

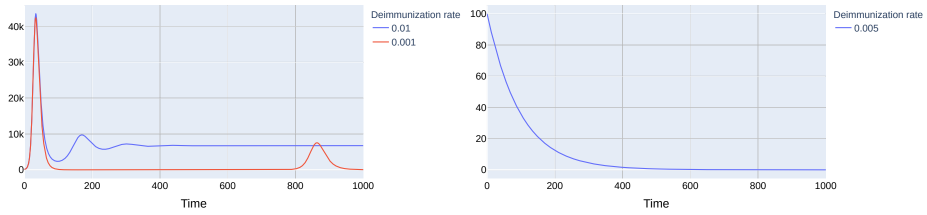
We remark that the above model can be generalised to incorporate spatial structures by adding extra compartments [80, 120, 224] that represent the spatial locations. For such models, the value of $\mathcal{R}_0$ remains explicitly computable and is influenced by the spectrum of the adjacency matrix that describes the underlying network [224].

7.3 EXPERIMENT 1: NEW PHENOMENA IN GEOMETRIC NETWORKS

We highlight the most important results of our first experiment, and illustrate these results using figures. A complete overview of the simulations can be found in [152], which considers a larger set of parameters, and states more quantitative results.

7.3.1 *Simulation setup*

Throughout this chapter, the underlying networks for the different scenarios are sampled once per parameter setting and kept constant over the simulation runs. On every graph, for every described parameter setting of



(a) Supercritical: $\beta = 0.38$.

(b) Subcritical: $\beta = 0.13$, $\eta = 0.005$.

Figure 11: Solutions of the ODE when $N = 160000$, $\gamma = 0.14$ with 100 initially infected nodes. The infection almost seemingly disappears from the system before the second peak appears when $\eta = 0.001$. The infection is subcritical when $\beta < \gamma$.

the SIRS model, the results are based on 100 simulations of the epidemic spreading starting from a single node. This node is sampled uniformly at random at the beginning of every simulation, and differs per run. All networks have the same number of $N = 160000$ nodes, all with the *same average degree* $\mathbb{E}[\deg(u)] = 8$: two configuration models with $\tau = 2.5$ and $\tau = 3.3$, respectively, three geometric inhomogeneous random graphs (GIRG) with $(\tau, \alpha) = (2.5, 2.3), (3.3, 1.3)$ and $(3.3, 2.3)$ respectively, and the modified lattice $\widetilde{\mathbb{Z}}^2$. The continuous line on the figures, see e.g. Figure 12 shows the median of infected individuals, the shaded area covers 95% of all runs for which the number of infected nodes was positive.

7.3.2 *Phases of SIRS on continuous compartmental models*

For Scenario $\mathcal{S}_{\text{ode}}$, the parameters may be set such that the epidemic is either subcritical, critical, or supercritical, see Figure 11. We briefly describe these so-called phases for comparison.

1) *Subcritical and critical phase*: *Immediate extinction*. Whenever the basic reproduction number, $\mathcal{R}_0 = \beta/\gamma \leqslant 1$, the epidemic dies out without producing a peak. For $\beta/\gamma < 1$, it dies out quickly (logarithmic in the initial number of infected), while for $\beta/\gamma = 1$ it dies out slowly (polynomial in the initial number of infected).

2) *Supercritical phase*: *Peaks of decreasing magnitude towards a limiting stationary proportion*. Whenever $\beta/\gamma > 1$, the proportion of infected population stabilizes at $\frac{\eta}{\eta+\gamma}\left(1 - \frac{\gamma}{\beta}\right)$. There are several larger peaks

before the equilibrium is reached. The incline and decline of peaks are *exponential*.

7.3.3 *New supercritical phases in SIRS epidemics on network models*

The presence of the network structure, as well as the underlying space, combined with temporary immunity has a surprising effect on the qualitative behaviour of the epidemic. More specifically, new phases arise in the phase diagram compared to those in Section 7.3.2. The duration of the immune period plays a more profound role than in the continuous compartmental model of Section 7.2.4, where the single supercritical phase is entirely determined by the ratio $\beta/\gamma = \mathcal{R}_0$ of the infection rate and the healing rate.

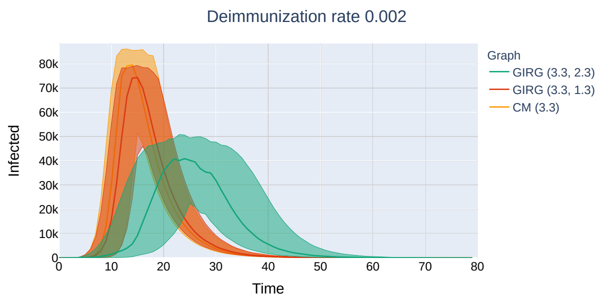
The final size of the epidemic is, at least for the most commonly known SIR model, often computable even for network models. We refer to the works [17, 18, 19, 93, 146, 192, 197, 208, 223] for SIR models in mean-field type networks, and the phase diagram for SIRS on the lattice [62, Figure 1]. Many of the cited papers assume that the degree distribution of the underlying has *finite* variance, while we consider infinite-variance degree distributions as well.

For the GIRGs that we consider, the branching process approximations used in the cited papers are no longer valid, and there is no explicit formula known for the value of $\mathcal{R}_0$ or an alternative threshold parameter that describes whether a first peak occurs with probability bounded away from 0.

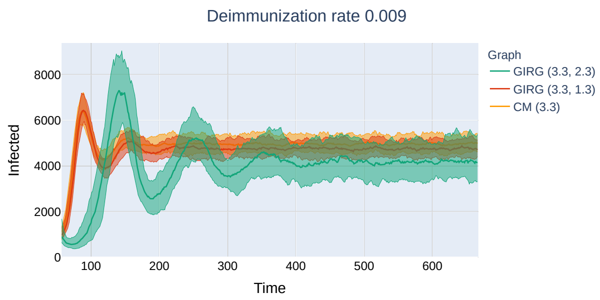
For the SIRS epidemic spread on all three network models (Scenarios $\mathcal{S}_{\text{grid}}$, $\mathcal{S}_{\text{cm}}$, and $\mathcal{S}_{\text{girg}}$) we observe the following phases[2], with 2S below being new compared to Scenario $\mathcal{S}_{\text{ode}}$.

1) *Subcritical and critical phase*: *immediate extinction.* The epidemic goes extinct almost immediately (for all runs), and the number of infected decays exponentially in this case. When $\mathcal{R}_0 \approx 1$, the epidemic may die out more slowly. We emphasize that the critical and subcritical

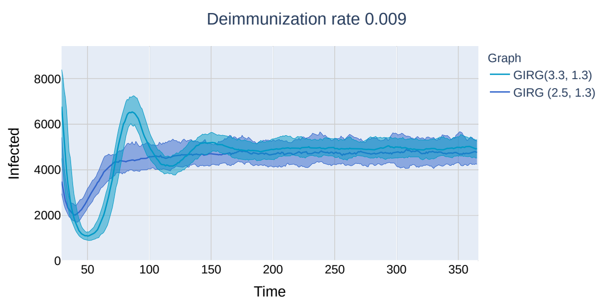
2 Due to the stochastic nature of the model, the phases 1, 2S, 2M are, (at least on a finite but large network) intertwined. This means that for a given fixed set of parameter β, γ, η, a single run of the SIRS epidemic may enter any of the three phases, with different probabilities. However, the probability of entering a given phase undergoes a sharp transition in the parameter values: for fixed β, γ, a small change in η results in a shift from almost always seeing phase 2S to almost always seeing phase 2M. We believe that the long-survival probabilities are monotone in η.



(a) Single peak when immunity is long-lasting ($\eta = 0.002$). Fewer long-range connections leads to slower incline and decline of the curves.



(b) Later peaks when immunity is short-lasting ($\eta = 0.009$); time starts at 50 days. Fewer long-range connections yield heavy oscillations after the first peak.



(c) Later peaks when immunity is short-lasting ($\eta = 0.009$); fewer hubs yield heavy oscillations after the first peak.

Figure 12: Supercritical phases for network-based models. The network CM(3.3) corresponds to Scenario $\mathcal{S}_{cm}$ (without intervention), GIRG(3.3,1.3) and GIRG(2.5, 1.3) to Scenario $\mathcal{S}_{girg}$ (without intervention), and GIRG(3.3,2.3) to Scenario $\mathcal{S}_{girg}$ (*with* intervention $\mathcal{I}_{trav}$, see Section 7.4 below). GIRG(2.5, 1.3) corresponds to a network where the degree distribution asymptotically has infinite variance (corresponding to many hubs being present in the network), while GIRG(3.3, 1.3) corresponds to finite variance (fewer hubs).

.

phase do not separate so clearly on the stochastic SIRS model on networks: due to the stochastic nature, the epidemic might die out quickly even when $\mathcal{R}_0 > 1$. For the extinction time (logarithmic vs polynomial of the network size), the initial number of infected individuals plays a more important role. This is, however, out of the scope of our current study and we do not pursue this direction further.

2) *Supercritical phases*: *possible large outbreak.* With positive probability, there is a large outbreak. For network-based SIRS, the supercritical phase separates into two sub-phases:

   2S) *Extinction after a Single peak*: when the duration of the immune period is long, the epidemic has a single peak, after which it immediately goes extinct. Heuristically, the reason for the existence of this new phase is that after the first epidemic peak, earlier infected nodes become immune and maintain their immunity long-enough to provide barriers in the network that the infection cannot pass through. This phenomenon is known as 'depletion of susceptibles' [218] and visualized in Figure 12a.

   2M) *Long-time survival, Many peaks:* when the duration of the immunity period is relatively short, the epidemic follows a (qualitatively) similar curve to the one observed in the ODE in Section 7.2.4. There is a first major outbreak, followed by smaller second, third, etc., peaks, decreasing in magnitude, and eventually settling on a *(meta-stable) stationary proportion* of infected and immune population. This is visualized in Figure 12b. We mention that this equilibrium is a meta-stable state, since the all-susceptible configuration is an absorbing state. The time to reach that, however, is exponentially long in the network size, see [48]. Figure 13 visualizes when the epidemic can exhibit phase 2M.

Contrary to networks, the $\mathcal{S}_{\text{ode}}$ scenario discussed above is a continuous system. As a result, the second peak is always present, which explains the absence of phase 2S in this compartmental model.

The first peak. The height of the first peak, in all models, is insensitive to η, just like for standard continuous compartmental models, visualized below in Figure 17 for Scenario $\mathcal{S}_{\text{girg}}$. A possible explanation for this insensitivity

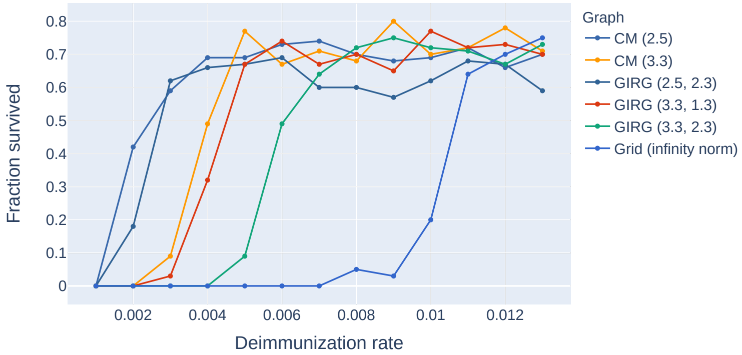


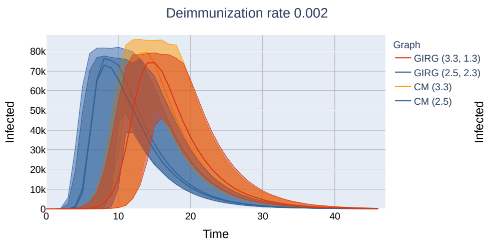
Figure 13: A comparison of the frequency of entering phase 2M during 100 simulation runs on eight different networks with the same average degree 8 and number of nodes $N = 160000$, as a function of the η. The frequencies were computed using a 100 runs for all parameter values (η) and models, with $\beta = 0.225, \gamma = 0.2$. For all networks we see a sharp transition at a critical value of η where the system moves from phase 2S (single peak followed by extinction) to 2M (survival with multiple peaks).

is that during the initial spread, the SIRS process is closely approximated by an SIR process, where temporarily immune nodes are simply removed.

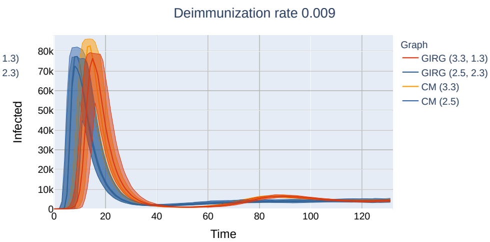
The configuration model with the smallest τ has the smallest threshold η_c, followed closely by GIRG with $(\tau, \alpha) = (2.5, 1.3)$. The grid $\widetilde{\mathbb{Z}}_n^2$ and GIRG with $\tau = 3.3$ and $\alpha = 2.3$ have the highest η_c (≈ 0.085 and ≈ 0.01), these are networks where the typical distance is linear.

Shape and location of the first peak: For Scenario $\mathcal{S}_{\mathrm{cm}}$, the shape of the first and later peaks are similar to that of Scenario $\mathcal{S}_{\mathrm{ode}}$, i.e., exponential incline followed by exponential decline, see Figure 12. Scenario $\mathcal{S}_{\mathrm{grid}}$ (not present in the figures) is different: it has a *linear* incline followed by linear decline when η is large. The shape of the curve for Scenario $\mathcal{S}_{\mathrm{girg}}$ depends on the parameter values: as long as there are many long-range connections, and hubs in the network (nodes with very high degree), the curve resembles that of Scenario $\mathcal{S}_{\mathrm{cm}}$. When this is not the case, Scenario $\mathcal{S}_{\mathrm{girg}}$ starts to gradually resemble the epidemic curve of the grid (Scenario $\mathcal{S}_{\mathrm{grid}}$): *linear* incline followed by linear decline, see Figure 12a.

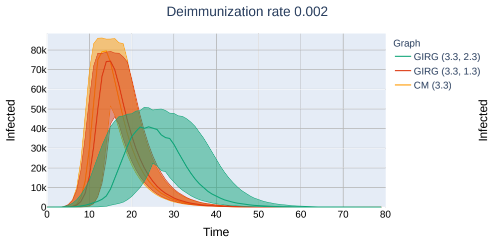
Besides that, we find that the epidemic curves for $\mathcal{S}_{\mathrm{girg}}$ and the a-geometric scenario $\mathcal{S}_{\mathrm{cm}}$ are in good match, as long as the long range parameter α of GIRG is in the interval $(1, 2)$ when $\tau > 3$, see Figure 14.



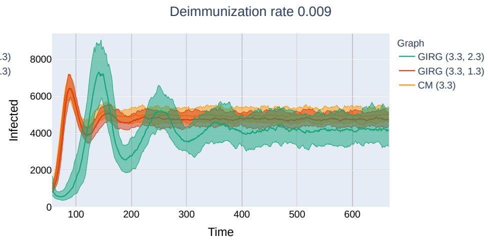
(a) Long immunity ($\eta = 0.002$), many hubs ($\tau = 2.5$): single peak.



(b) Shorter immunity ($\eta = 0.009$), many hubs ($\tau = 2.5$): many peaks.



(c) Long immunity ($\eta = 0.002$), few hubs ($\tau = 3.3$): single peak.



(d) Shorter immunity ($\eta = 0.009$), few hubs ($\tau = 3.3$): many peaks. Time starts at 50 days.

Figure 14: A comparison of the number of infected on the configuration model versus GIRG with matching parameters. The two scenarios match unless $\tau > 3$ and $\alpha > 2$, corresponding to the green curves. Figure (c) and (d) coincide with Figure 12(a-b) and included again for comparison to $\tau = 2.5$.

For $\tau \in (2, 3)$, hubs dominate the network, and the role of the parameter α is insignificant. This is in accordance with theoretical results (7.2.3), where distances are doubly-logarithmic regardless of the value α as long as $\tau \in (2, 3)$.

When $\tau > 3$, the effect of the long range parameter α plays a crucial role. When $\alpha > 2$, even when keeping the average degree and the degree distribution the same, the curve is flattened significantly. In this parameter range, even though hubs are present in the network, they mostly connect to nearby nodes. As a result, the typical distance is significantly larger (see (7.2.3)) and the infection needs more time to spread.

The linear growth was also investigated in a similar setting recently in [221]. There, the authors proposed to model the contact network using the related (one-dimensional) Newman-Watts model [199]. The mathematical explanation for seeing a linear growth curve is as follows: these models are embedded in geometry, and the parameters are such that long edges do not dominate yet, and hence the typical ball-growth is polynomial. This is in accordance with theoretical results related to the typical distance in the network being short (poly-logarithmic) when there are many long-connections but long (polynomial) when these are scarce, see (7.2.3). See [170] and references therein for the role of long-connections.

After the first peak: topology-dependent fluctuations. The fluctuations after the first peak are highly dependent on the network topology as can be seen in Figure 12c. For power-law networks with infinite asymptotic variance ($\tau < 3$), there is no second peak and the system reaches equilibrium very quickly after the first peak. For finite variance networks ($\tau > 3$), several further peaks occur. The transition is even more profound when interventions change the topology, see Section 7.4.2 below.

We mention that for Scenario $\mathcal{S}_{\text{girg}}$ the way the second peak happens is entirely different from the first peak: the initial spread is local, emanating from the source, (with some long connections causing non-local new infection centers), while the second peak is spread out, infections occur everywhere in space. This is quite natural, since during the initial wave, the epidemic fills the space, and thus the second peak happens roughly when the immunity arising from the first wave starts to wear off, at which time the epidemic is endemic. This is why we never see the epidemic to die out after a second peak. It either dies out after the first peak, or many peaks occur and the system reaches equilibrium.

We conclude that epidemics modeled on networks can exhibit phenomena, mostly related to the later behaviour of the epidemic, that continuous ODE approximations or a-geometric networks fail to capture.

7.4 EXPERIMENT 2: MODELLING INTERVENTION METHODS ON GIRGS

We turn to the second experiment, which compares intervention methods under the SIRS epidemic. These intervention methods we model only on GIRGs, since GIRGs offer an intuitive way of modifying the original network to model interventions. We introduce the methods, and explain how we modify the topology of the 'original' network.

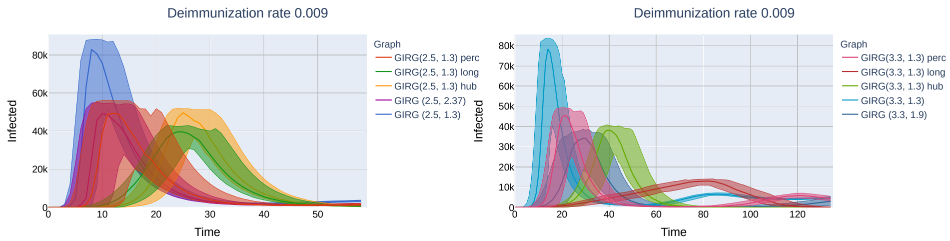
($\mathcal{I}_{\text{soc}}$) During the COVID-19 outbreak, many governments issued a collection of social distancing measures: keeping 1.5 meter distance from each other, wearing face masks and gloves, washing hands frequently, and so on. Such measures reduce the number of contact moments between individuals and thus decrease the chance for the virus to spread. When the social distancing measures are not combined with travel restrictions, this results in the underlying contact network to lose its connections roughly randomly, see also [75]. Hence, we model the effect of *social distancing* by *randomly removing* a certain proportion of edges in the network, such that each edge is removed independently of other edges. The same method is used to model social distancing in e.g. [206].

($\mathcal{I}_{\text{trav}}$) During the COVID-19 outbreak, there was restricted travel between countries, and even within countries. As a result, the underlying contact network loses some of its long connections. Such travel restrictions are modeled in two different ways. First, it is modeled by drawing for each present edge a random exponential threshold with a common mean L, and cutting the edge if it exceeds its threshold value, resulting in a connection-length distribution that decays exponentially above the threshold L. We call this intervention the *strong* travel restriction. Secondly, it is also modeled by increasing the long-range parameter α to α^{new} (see Section 7.2.3 for a discussion of the model parameters of GIRGs).

($\mathcal{I}_{\text{deg}}$) Another intervention measure is to limit the maximal number of contact a person can have. On the contact level, this rule barely affects individuals with a low number of contacts, but it aims to decrease the number of contact of 'superspreaders' in the network: nodes that have a relatively high degree. Limiting the maximal number of contacts per person is modeled by prescribing a maximal node degree M. For each node u with degree higher than M, randomly chosen connections of u are cut until at most M connections remain.

7.4.1 *Simulation setup*

We choose the parameters $M, L, \alpha^{\text{new}}$ so that the average degree is the same after all interventions to enable comparison between the interventions. We model the interventions on two GIRGs, corresponding to the first two universality classes with respect to the typical distance from (7.2.3). The GIRG with $\tau = 2.5$ has average degree 9.6 and contains $N = 160000$ nodes. The networks after intervention have average degree ≈ 4.9. The GIRG with $\tau = 3.3$ has average degree 8.7 and contains $N = 160000$ nodes. The networks after intervention have average degree ≈ 4.7. As before, the



(a) GIRGs with many hubs: $\tau = 2.5$.

(b) GIRGs with few hubs: $\tau = 3.3$.

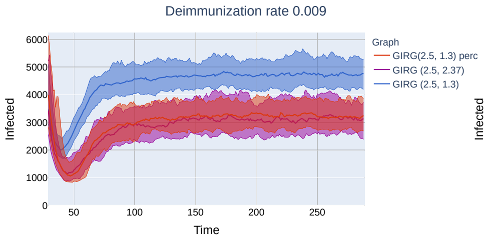
Figure 15: The first peak under interventions. The hard no-travel rule is most effective in flattening the curve: both in its height as well as the day of the peak. The first peak disappears before day 60 in all cases. The curves for the weak no-travel rule and social distancing are almost identical. Explanation of the legend: social distancing ("GIRG$(2.5, 1.3)$ perc" and "GIRG$(3.3, 1.3)$ perc"), the hard no-travel rule ("GIRG$(2.5, 1.3)$ long" and "GIRG$(3.3, 1.3)$ long"), limiting the maximal number of contact ("GIRG$(2.5, 1.3)$ hub" and "GIRG$(3.3, 1.3)$ hub"), and the weak no-travel rule ("GIRG$(2.5, 2.37)$" and "GIRG$(3.3, 1.9)$").

.

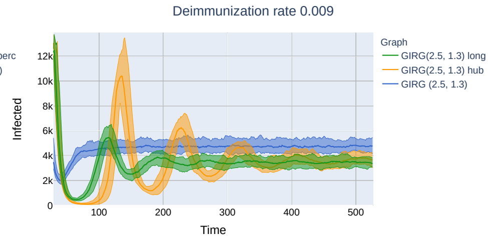
results are based on 100 simulations of the epidemic spreading starting from a single node. This node is sampled uniformly at random at the beginning of every simulation, and differs per run. The continuous line on the figures shows the median of infected individuals, the shaded area is covers 95% of all runs for which the number of infected nodes was positive.

7.4.2 *Comparison of intervention methods $\mathcal{I}_{soc}$, $\mathcal{I}_{trav}$, $\mathcal{I}_{deg}$ on GIRGs*

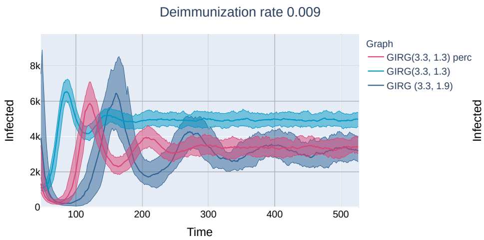
We state the most important, qualitative findings of the intervention methods on Scenario $\mathcal{S}_{girg}$, the geometric scale-free network model. For additional quantitative values and additional figures we refer to [152]. *Long-survival probabilities drop.* Each intervention pushes the probability of the system entering phase (2M) lower for a fixed parameter setting β, γ, η. When increasing η (thus decreasing the immunity duration), at a critical η_c the system abruptly goes from dominantly phase 2S from dominantly phase 2M, (a sharp threshold). The threshold η_c of long survival is increased under all interventions, see Figure 17b. In this respect, interventions $\mathcal{I}_{trav}$ and $\mathcal{I}_{deg}$ perform best.



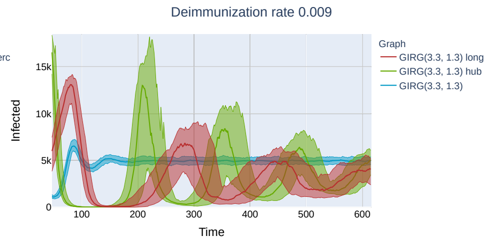
(a) Social distancing ("GIRG(2.5, 1.3) perc"), as well as the weak no-travel rule ("GIRG(2.5, 2.37)") on GIRGs with many hubs ($\tau = 2.5$). No second peak can be observed, just like on the original network.



(b) Limiting the maximal number of contact ("GIRG(2.5, 1.3) hub"), and the hard no-travel rule ("GIRG(2.5, 1.3) long'"): second and further peaks are periodically present, with decreasing magnitude and period roughly the average immunity duration.



(c) Later effects of interventions on GIRGs with less hubs ($\tau = 3.3$. Social distancing ("GIRG(3.3, 1.3) perc"), as well as the weak no-travel rule ("GIRG(3.3, 1.9)") compared to the original network ("GIRG(3.3, 1.3)"). Second and further peaks are present, equilibrium is reached within 600 days.

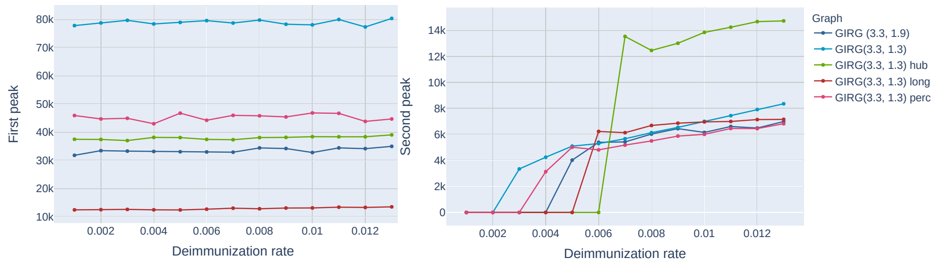


(d) Limiting the maximal number of contact ("GIRG(3.3, 1.3) hub"), and the hard no-travel rule ("GIRG(3.3, 1.3) long'"): second and further peaks are more profound, equilibrium is only reached around 1200 days. The first peak of the red curve (corresponding to truncating long edges) is the first peak of the epidemic, contrary to the others.

Figure 16: Effect of interventions on GIRGs.

The height of the first peak drops. In all interventions, the height of the first peak is dropping, by at least as much as the shrinkage in average node degree (intervention $\mathcal{I}_{soc}$) and even more with other interventions. The most effective intervention in this respect is the strong no-travel rule, see Figure 17a. The first peak is insensitive to the immunity rate η.

Elongated first peak. The time and duration of the first peak is later/longer than without intervention. For keeping physical distance, the time shift is



(a) First peak under interventions.

(b) Second peak under interventions.

Figure 17: The effect of interventions on the height of the first and second peak, as a function of η (the inverse of the average immunity period). The height of the first peak is insensitive to the duration of immunity , and cutting long edges is most effective in reducing the first peak. For the second peak, limiting node degree pushes the critical η for appearance of the second peak from 0.002 to 0.006. However, for larger η the second peak is twice as high as for other interventions. Abbreviations in the legend: 'hub': limiting maximal node degree, 'long': strong restriction on travel, 'perc': social distancing. GIRG$(3.3, 1.3)$ is the original network, GIRG(3.3, 1.9) corresponds a weak restriction on travel.

the least apparent (a few time steps), while for intervention $\mathcal{I}_{\text{trav}}$ it is the most profound, a factor of 10, see Figure 15b.

Higher second peak. Intervention $\mathcal{I}_{\text{trav}}$ and $\mathcal{I}_{\text{deg}}$ result in a higher second peak, or even make a second peak appear where originally it was not present. The worst intervention in this respect is intervention $\mathcal{I}_{\text{deg}}$ (limiting the maximum degree), where the second peak can be as high as $1/3$ of the first peak, see Figure 17b. For infinite variance degree GIRG, the strong travel restriction also causes a second peak, while social distancing and increasing α do not, see [152, Figure 21].

We give a possible explanation for a higher second peak. Under intervention $\mathcal{I}_{\text{deg}}$ (limiting number of contacts via limiting node degree) and Intervention $\mathcal{I}_{\text{trav}}$ (cutting long-edges), there are no more hubs in the network: nodes with very high degree. In networks that contain hubs, the hubs infect a large proportion of their neighbours roughly. Moreover, hubs are much closer to each other (in terms of graph distance) than in networks without hubs, thus the infection can travel quickly between hubs. The randomness ensures that the infection times are not synchronized, explaining the absence of oscillations in the epidemic curve. To summarize, hubs *stabilise* the system.

We give a possible explanation for a higher second peak. Under intervention $\mathcal{I}_{\text{deg}}$ (limiting number of contacts via limiting node degree) and Intervention $\mathcal{I}_{\text{trav}}$ (cutting long-edges), there are no more hubs in the network: nodes with very high degree. In networks that contain hubs, the typical graph distance is much smaller than in networks without hubs, and there are typically many short paths from low-degree vertices to hubs. As a result, once a node loses its immunity, it is almost immediately infected afterwards via an infected hub, shortening the the time spent in the susceptible state. This happens for almost all nodes in the network, and the randomness ensures that the infection times are not synchronized, explaining the absence of oscillations in the epidemic curve. To summarize, hubs *stabilise* the system.

Once the hubs are removed, the typical graph distance increases, nodes stay susceptible for a longer time, which leads to oscillations in the system since vertices close to each other in graph distance are infected around the same time. Limiting node degree without travel restrictions has the strongest effect on the removal of hubs. Limiting travel reduces the degree of hubs, but does not remove them completely: nodes that were hubs before the intervention, have typically many local connections as well.

More oscillations to reach equilibrium. Interventions $\mathcal{I}_{\text{deg}}$ and $\mathcal{I}_{\text{trav}}$, while they are most effective regarding the first peak, introduce high oscillations in the system, and the time to reach equilibrium can be much longer than in the original model (from a few days to more than a thousand days). Again, these oscillations are explained by the lack of stabilisation. See Figure 16 for these later effects.

To summarize, we conclude that the effect of interventions vary to a high extent and depend on the precise way the intervention changes the network topology. Non-geometric models such as the configuration model fail to capture these various effects.

8 INCREASING EFFICACY OF CONTACT-TRACING APPLICATIONS BY USER REFERRALS AND STRICTER QUARANTINING

Based on [100]:
Increasing efficacy of contact-tracing applications by user referrals and stricter quarantining,
L.A. Goldberg, J. Jorritsma, J. Komjáthy, J. Lapinskas
PLoS One, 2021, 16(5): e0250435.

8.1 INTRODUCTION

Agent-based epidemiological models allow a population to be embedded in a geometric space to capture the effect of the "local" and "long-distance" contacts that arise in real populations. We use a stochastic susceptible-exposed-infected-removed (SEIR) type agent-based model to study the efficacy of contact-tracing applications such as the mobile phone contact-tracing applications that have been introduced during the COVID-19 epidemic. Our simulation study compares the efficacy of four different uptake scenarios for the contact-tracing application (CTA), combined with the effect of quarantining measures of various strengths.

We compare four uptake scenarios of the CTA given a probability $p \in [0, 1]$:

- $\mathcal{S}_{\mathrm{rand}}$: each individual uses the CTA independently of the rest with probability p;
- the users are selected by one of two different recommender scenarios: each individual is an initial user with probabability p_{init} and recommends the CTA to some of their contacts:
  - $\mathcal{S}_{\mathrm{basic}}$: a uniformly chosen neighbour;
  - $\mathcal{S}_{\mathrm{ring}}$: roughly half of its neighbours (each neighbour uses the app with probability $1/2$ independently of the rest);

- $\mathcal{S}_{\text{deg}}$: a (hypothetical) degree-targetted uptake scenario, in which CTA-users are the $(p \cdot 100)\%$ individuals with the highest number of connections.

For the recommender-based scenarios we choose p_{init} such that the final uptake percentage is $(p \cdot 100)\%$. In each of these CTA-uptake scenarios, we assume that individuals in the population will quarantine themselves as soon as their symptoms are sufficiently strong (exceeding a certain threshold). At this point, users of the CTA also notify their CTA-using contacts, who also quarantine themselves.

The ring recommendation method has some connections with a certain vaccination protocol, called ring vaccination, hence the name. Ring vaccination has been successful in the past to eradicate small pox [219] and has also been used against Ebola [119]. Such recommender scenarios have previously been shown to be effective in the context of vaccines [137] on synthetic models.

Basic recommendation also has connections to a vaccination protocol called acquaintance vaccination which typically outperforms random vaccination in network models [41, 54]. The main difference between the two scenarios is that in acquaintance vaccination, only the neighbours of the randomly-chosen set of “initial users” receive the vaccine rather than the users themselves. This makes sense in the context of trying to allocate a limited supply of vaccines, but not in our context — it is reasonable to assume that anyone recommending the app to a friend will also install it themselves.

We remark that degree-targetted uptake is an unrealistic scenario to consider, since it requires a complete knowledge of the contact network. However, studying the performance of this hypothetical scenario gives a baseline comparison for determining the potential performance of a given uptake percentage.

We examine the combined effect of differing CTA-uptake scenarios and different symptom thresholds for quarantine. We decouple the effects of CTA-uptake scenarios and symptom thresholds from the effects of testing delays, which have been studied in [167]. In order to focus on the former, we work in an idealised hypothetical scenario where testing is immediate upon symptom onset (or where there is only a single virus present in the population that could cause the symptoms) and where the symptom-severity does not influence the probability of a virus transmission between two individuals. We emphasise that our study is qualitative, rather than quantitative, and we aim for universally valid observations in terms of the performance of CTA-uptake scenarios.

Main observations

In our first experiment (Section 8.3) we qualitatively compare the four uptake scenarios. In all uptake scenarios, we find that contact-tracing applications (CTAs) are effective in decreasing the size of an epidemic (the total number of people that get infected) and in decreasing the maximum number of people who are simultaneously hospitalised, and typically this effectiveness increases linearly or super-linearly with the percentage of the population that use CTA. In these respects, we show that CTAs are effective even with low uptake rates. This confirms the results of a recent study modelling CTAs in Washington state [1].

The novelty of our study is to compare recommendation-based uptake with random uptake, where we find that recommending can significantly improve the efficacy of the CTA. In brief, we find that in scenarios where the CTA is recommended to acquaintances the epidemic size and maximum hospital load decrease at a much higher rate than the rate that would be achieved by randomly selecting the same number of CTA users. One might expect these advantages of recommendation-based uptake to come at the cost of increased quarantine. This is true, however in networks with fewer super-spreaders the quarantine load decreases with uptake percentage (after an initial increase). Once the uptake percentage is sufficiently high, recommendation-based uptake leads to less quarantine, rather than more. Finally, we emphasise that the uptake percentage necessary to completely eradicate the epidemic from the population is generally very high, as was also found in [172], though this is not the focus of this chapter.

In our second experiment (Section 8.4) we study the impact of the severity of quarantine measures, both on the epidemic itself and on the economic disruption that it may cause (measured in terms of the time that people spend in quarantine over the course of the epidemic). Already under "mild" quarantine measures, the maximum hospital load is drastically reduced and the size of the epidemic is somewhat reduced. However, the economic disruption caused by quarantine is very high. Perhaps surprisingly, imposing stricter quarantine measures (i.e., quarantining individuals already with less severe symptoms), ensures not only that the epidemic is better contained, but also that the economic disruption is lower (under strict enough social distancing measures, travel restrictions, and sufficiently high app uptake). The critical point on the 'quarantine-strictness scale' above which the social disruption starts decreasing comes earlier in recommender scenarios and degree-targetted uptake than in random uptake, and even earlier if individuals are more restricted in their movements, i.e., the underlying contact network does not have many long-range connection, ranging over large spatial distances.

As a disclaimer, we recall that we study a case where a *single* SEIR (susceptible-exposed-infected-removed) epidemic is present in the population (see Section 8.2.2) and individuals who are not infected do not show any symptoms similar to those that are caused by the virus under investigation. This might be unrealistic for certain types of diseases, but might be a better approximation of reality for others. Finally, we emphasise that we aim for qualitative, comparative results that are robust against changes in the dynamics and the underlying contact network. We refrain from making numerical predictions.

Organisation

In the following section we describe the underlying contact networks that model the population and the epidemic model. We then explain and interpret our main findings regarding the CTA-uptake scenarios and the quarantining measures. For additional baseline comparisons and the robustness of parameters, we refer to [100]. The software used for the simulations is available at [155]. Besides our own code, we used [220] to run simulations in parallel, and [31] to sample the four underlying networks for the population (geometric inhomogeneous random graphs).

8.2 MODELLING THE EPIDEMIC

We use a SEIR-type (susceptible-exposed-infected-removed) agent-based models to model the spread of the epidemic, where individuals may or may not show symptoms. The model is defined precisely in Section 8.2.2. Roughly, each agent goes through the following stages upon infection: exposed, infected, either symptomatic or asymptomatic, and finally removed. Upon showing symptoms, individuals using the CTA send notifications to the other CTA users that they are in contact with, who then move to quarantine for a fixed duration of time. The model involves a *quarantining* scheme: agents who show sufficiently severe symptoms and agents notified via the CTA both move into quarantine.

We start by introducing the underlying network.

8.2.1 *Networks to model the population*

To be able to carry out our study, we represent the underlying contact network of individuals (nodes, or agents) as a network. Based on the universality results in Chapter 7, our primary choice for the underlying contact network is a mixture of pure geometric and purely random network models, called Geometric Inhomogeneous Random Graphs, which possess geometric features and can match the statistical properties of real human contact networks, such as degree distributions and clustering, as well as long-range connections.

We recall that there are two robust parameters of GIRGs that control the qualitative features of the network: a parameter τ controls the variability in the number of neighbours that individual nodes have (the number of neighbours of a node is called its "degree"). Smaller values of τ correspond to more variability, while keeping the average the same.

A second parameter α controls the number of long-range edges: long-range edges would tend to be present in populations without travel restrictions, (as observed in real-life contact networks, see e.g. [61, 102, 225]). A value of α close to 1 corresponds to having many long-range edges, in this case the network resembles an ageometric (or mean-field) network model; while increasing α reduces the number of long-range contacts, thus transforming the network into one where the underlying geometry is intrisically more apparent. For more information, we refer to Section 7.2.3 above. The following summary will help the reader to understand the experimental results reported below.

- $\tau = 2.3$: a GIRG with plenty of "super-spreaders" (nodes with many connections) due to degree variability,
- $\tau = 3.3$: a GIRG with fewer super-spreaders,
- $\alpha = 1.3$: a GIRG with many long-range edges, hence the underlying geometry is "less apparent",
- $\alpha = 2.3$ a "more geometric" GIRG, i.e., a network that resembles a lattice better, since it does not have too many long-range connections.

The parameter values (τ, α) are chosen to represent *each universality class* of GIRGs with respect to typical distance in the network, see (7.2.3).

For baseline comparison, in the paper [100], we also study the case when the underlying network is a "mean-field network", ignoring the spatial component. For this scenario we used the *configuration model*, which can mimic the local statistical properties of real human contact networks, such as degree distributions. We refer to [100] for these results, that are all similar to the results on GIRGs that we present in the current chapter.

8.2.2 *SEIR with quarantining: The epidemic model*

The epidemic model that we use is a refinement of the agent-based version of the discrete-time SEIR model. The spreading process changes at discrete time steps, $t = \{0, 1, 2, \dots\}$, each time step corresponds to, say, a day. We fix the network G in advance. We think of nodes in the network as individuals, and denote the set of nodes by $\mathcal{V}$. The neighbours of a node u are nodes with a direct connection (also called link, or edge) to u. A connection may correspond to a friendship or an acquaintance, or simply a contact event.

The model has a parameter $q \in [0, 1]$ corresponding to quarantine-strictness. The quantity q is the fraction of individuals whose symptoms would be so severe that they would go into home-quarantine if infected. Determining the value of q is a social/economic decision. The model reflects this decision as follows. For each node, independently, and in advance of the epidemic simulation, it is determined randomly, with probability q, whether the individual corresponding to this node will, upon infection, develop symptoms exceeding the socially-determined symptom severity threshold, and then go into home-quarantine. The set of nodes which will do so is called $\mathcal{V}_{\text{sev}}$ and the set of nodes which will not is called $\mathcal{V}_{\text{mild}}$. During the epidemic, nodes in $\mathcal{V}_{\text{sev}}$ that become infected will go into home-quarantine. One may worry about uncooperative individuals who do not self-isolate, despite having sufficiently severe symptoms. We do not need to adjust the model to account for these, since the quarantine-strictness q can be viewed as the fraction of individuals that would develop sufficiently-severe symptoms, and would be willing to self-isolate. Note that the actual distribution of the symptom-strength in the population is irrelevant to the problems that we study; the only relevant parameter is the final proportion q. The assumption that we do make, however, is that each individual has an independent symptom-strength variable, making $\mathcal{V}_{\text{sev}}$ a random set of nodes.

Definition 8.2.1 (SEISeR model)**.** Determine the node sets $\mathcal{V}_{\text{sev}}$ and $\mathcal{V}_{\text{mild}}$. Nodes in $\mathcal{V}_{\text{sev}}$ can be in five possible states: *susceptible (S), exposed (E), infectious but not yet showing severe symptoms (I), infectious and showing severe* *symptoms (Se), or removed (R)*, while nodes in $\mathcal{V}_{\text{mild}}$ can only be in state $(S), (E), (I)$ or (R). At $t = 0$, a subset of nodes $\mathcal{E}_0 \subset \mathcal{V}$ is put in state (E), all other nodes are in state (S) (susceptible). The discrete time dynamics between the states are described as follows (see also Figure 18):

- *Infecting:* Each infectious node (in state (I)), *infects* each of its susceptible neighbours within the network with probability β at every time step. Infections to different neighbours happen independently. These neighbour nodes enter state (E), exposed to the virus.
- *Exposed $\rightarrow$ Infectious (not (yet) severely symptomatic):* When exposed to the virus, each node becomes infectious (transitions to state (I)) with probability γ at every time step, independently of other nodes.
- *Infectious $\rightarrow$ Severely Symptomatic:* When a node in $\mathcal{V}_{\text{sev}}$ is in state (I), (infectious but does not show severe symptoms yet), it transitions to state (Se) with probability $\gamma_{\text{sev}} = 1/2.5$ at each time step.
- *Infectious $\rightarrow$ Removed:* When a node in $\mathcal{V}_{\text{mild}}$ is in state (I), infectious but does not show severe symptoms, it heals and transitions to state (R) with probability $\eta_{\text{mild}} = 1/7$ at each time step.
- *Severe Symptomatic $\rightarrow$ Removed:* When a node is showing severe symptoms, it *heals* with probability $\eta_{\text{sev}} = 1/4.5$ at every time step, independently of other nodes. Upon healing, the node enters state R (removed).

We emphasise that in our model, severely symptomatic nodes do not infect anymore. We assume that they self-quarantine until being removed. In reality, some infectious individuals may choose not to self-quarantine, but we have already incorporated this into the model: $\mathcal{V}_{\text{sev}}$ contains only nodes that would develop sufficiently severe symptoms, and would go into home-quarantine.

To keep the parameter space tractable, we set $\gamma = 1/3$, $\eta_{\text{mild}} = 1/7$, $\gamma_{\text{sev}} = 1/2.5$, and $\eta_{\text{sev}} = 1/4.5$. Thus, in the model, after being exposed to the virus, it takes on average 3 days to become infectious and a further 7 days to heal completely for each individual. Nodes in $\mathcal{V}_{\text{sev}}$ infect for an average of 2.5 days before showing symptoms and moving to isolation (that lasts until their perfect healing) while nodes in $\mathcal{V}_{\text{mild}}$ are infectious for an average of 7 days before removal. These values are based on empirical findings e.g. in [115, 167, 168, 174], where we emphasise that our results are robust in these parameters. Experiment 2 varies the value of q, which is the expectation of $|\mathcal{V}_{\text{sev}}|/|\mathcal{V}|$.

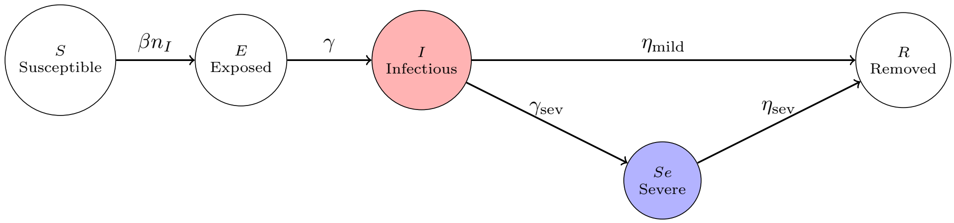


Figure 18: An illustration of the dynamical change of states. On the first transition arrow, n_I denotes the number of neighbours in state (I) of a susceptible node. Only nodes in $\mathcal{V}_{\text{sev}}$ transition to state (Se). By setting $1/\eta_{\text{mild}} = 1/\gamma_{\text{sev}} + 1/\eta_{\text{sev}}$ the average duration of having the virus in an individual's system is $1/\eta_{\text{mild}}$ for each node. Only nodes in the red circle can infect, while nodes in the blue circle stay in home isolation.

Next, we describe our model for the CTA and define the necessary modifications to the dynamics of the SEISeR model above to incorporate the presence of a CTA. We recall the four uptake scenarios described in Section 8.1.

Definition 8.2.2 (Random, basic-recommender, ring-recommender, and degree-targetted CTA uptake)**.** Before the epidemic starts, we determine the sets $\mathcal{V}_u, \mathcal{V}_{nu}$ of users and non-users of the application. We denote by $p = 100|\mathcal{V}_u|/|\mathcal{V}|$ the empirical uptake percentage after determining the CTA-users according to one of the scenarios in Section 8.1.

Note that, to achieve the same uptake ratio p, the initial ratio of users for the recommender-based scenarios p_{init} is *smaller* than p, see Figure 20.
Next, we describe the necessary modifications to the SEISeR model in the presence of a CTA. Informally, a CTA user, upon either being tested positive or simply showing sufficient symptoms, will notify its contacts to stay in quarantine for a duration of T days ($T = 14$ in most countries), regardless of whether these contacts themselves show symptoms.

Definition 8.2.3 (SEISeR with a contact-tracing application (CTA))**.** In addition to the original states (S), (E), (I), (Se) and (R) of the SEISeR model, there are five new states (NS), (NE), (NI), (NSe) and (NR), corresponding to "notified" versions of the original states — we refer to these as "N-states" and we refer to the original states as "O-states". Only CTA-users (nodes in $\mathcal{V}_u$) will ever enter the N-states. They enter the N-states when they receive a notification, via the CTA. They stay in the N-states until $T = 14$ days have elapsed since the last notification received. While in N-states, they self-quarantine, so cannot spread infection.

The transitions between the N-states are exactly the same as the transitions between the original versions of these states, except that there is no transition from (NS) to (NE), because the notified nodes in state (NS) will be quarantining, so will not become exposed. See the black arrows denoting transitions in Figure 19.

All nodes start in O-states, exactly as in the SEISeR model. After each time step t, nodes move between the O-states and the N-states as follows:

- *Notification:* If a node in $\mathcal{V}_u$ starts to show sufficient symptoms at time t (i.e., either it transitions from (I) to (Se) at time t or it transitions from (NI) to (NSe) at time t) then it sends a notification to all of its network-neighbours that are in $\mathcal{V}_u$.
- *Moving to* N*-states:* If a node in an O-state receives a notification, it moves to its corresponding N-state (following the corresponding red arrow in Figure 19).
- *Moving to* O*-states:* If a node is in an N-state and it has not received a notification for $T = 14$ time steps, then it moves to its corresponding O-state (following the corresponding blue arrow in Figure 19).

8.2.3 *Justification of modelling choices and robustness*

We make a few comments to justify our modelling choices. Nodes in the N-states and in state (Se) are assumed to be self-quarantining, so they cannot spread infection. These are the blue nodes in Figure 19). Only nodes in state (I) (coloured red) can spread infection.

Sending Notifications: The reason that nodes entering state (NSe) send notifications (even though they are already self-quarantining) is that they may have spread infection before entering the N-states (and starting to quarantine): their neighbours may be in exposed or infectious states when receiving the notification. In fact, while this model does not include testing, such a scenario also happens in real life when an individual, already part of a contact tracing chain, receives a positive test. In this case, the rest of the contacts of this individual need to be notified as well.

Removed nodes in quarantine: The underlying SEISeR model assumes that individuals can only be infected once (so “removed” nodes do not become susceptible again). Despite this, in light of possible re-infection, many countries (e.g. the Netherlands, where two of the authors are located) do ask individuals who are notified to self-isolate, even if they have already have Covid-19. While transitioning from (R) to (NR) has no effect on the course of the epidemic, it does have an effect on the total population's effective workforce (which we also study). Indeed, the self-isolating blue nodes represent a social cost, since all of these nodes are self-isolating, and cannot go to work. A node who is say, exposed when it receives a notification via a CTA may go through all the phases (NE), (NI), (NSe), (NR) before the T days are over. In this case it cannot become simply removed and go back to work immediately, but has to wait out the T days and then transition to state (R). For the same reason we make a distinction between nodes in (Se) and (NSe): nodes in (Se) may leave home

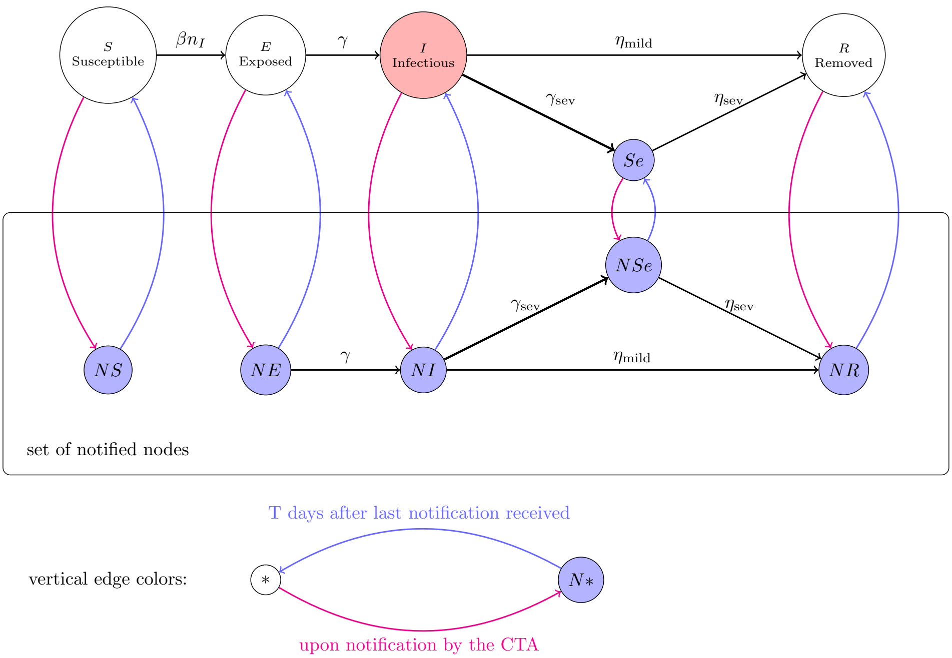


Figure 19: An illustration of the dynamical state changes in the presence of a contract-tracing application. The black arrows represent the transitions of the SEISeR model and the corresponding transitions in the notified versions of the states. After each time step, nodes move between the O-states (S), (E), (I), (Se) and (R) and the N-states (NS), (NE), (NI), (NSe) and (NR). Moves to N-states (depicted by red arrows) follow from CTA notifications. CTA users (nodes in $\mathcal{V}_u$) receive such notifications when neighbouring CTA users show sufficient symptoms, thus transitioning from (I) to (Se) or from (NI) to (NSe). Nodes in blue states are required to self-quarantine so cannot spread infection. Moves to O-states (depicted by blue arrows) occur when a node has received no new notifications for $T = 14$ days.

isolation immediately after healing, while nodes in (NSe) have to wait for the necessary total T days even after healing to be able to leave isolation.

Multiple notifications extend the quarantine time: We make a further comment on why we choose to extend the quarantine by T days after an additional notification is sent to a node already in an N-state. The reason for this is again based on real life experience: since not all contacts and hence notifications lead to infection, a second notification should not be ignored. Whether individuals are actually willing to comply with such rules belongs to behavioural science and is out of the scope of the current study. We assume an idealised scenario where each node in $\mathcal{V}_u$ complies with the rules. Varying the size of V_u captures the effect of less compliance as well.

Robustness of choices: To conclude the justification, we emphasise that we conduct a *qualitative and comparative* study. The model is sufficiently robust that the conclusions of our study also apply under small changes to the model. For example, if an uptake scenario performs better than another with regard to this particular model, it will also perform better in a model with slight changes, e.g. a model that does not extend the time of home isolation when a second notification is received, or a model that does not require already-removed nodes to go into quarantine.

We note that since each infectious node infects each of its susceptible neighbours with probability β at each time step, this model is a reactive process in the sense of [101].

8.2.4 *Key performance indicators (KPIs)*

To compare the performance of the four uptake scenarios as well as the strictness of quarantining, we study three KPIs as a function of the parameters.

- Size: The size of the epidemic, meaning the total number of individuals ever infected during the whole course of the epidemic.
- HMax: The (approximated) maximum hospital load. In our study, we define the hospital load at any time to be 5% of the number of agents in infected states (symptomatic or asymptomatic) individuals. Of course, the 5% is just a scaling, and changing this factor does not change the *shape* of the curve. HMax is then the maximum of this hospital load, over the course of the epidemic. This concept (without scaling) is also known as 'peak prevalence' [181], which is arguably

a better name: recovery after hospitalization can take longer, and, moreover, there is a time lag between the onset of symptoms and the moment of actual hospitalization. However, to keep the notation and figures in line with [100], we will refer to HMax throughout the chapter.

- Quar: The average number of days spend in quarantine per person, over the course of the epidemic. We obtain this quantity as the accumulated number of days spent in quarantine by the population, divided by the total population size.

8.3 EXPERIMENT 1: APPLICATION-UPTAKE SCENARIOS

An important aspect of our study is a comparison of the efficacy of four different CTA-uptake scenarios. Suppose that p% of the population use and also comply with a CTA, in the sense of reporting symptoms and also going into self-quarantine when instructed to do so. We compare the effectiveness of the CTA in terms of reducing the total size of the epidemic (which we denote Size) and the maximum hospital load HMax, when this p% uptake is achieved by the four uptake scenarios described in Section 8.1. In terms of controlling the epidemic, one would expect that it is desirable if the p% most influential members of the population use the CTA, rather than a randomly-chosen p%. However, *targetting* the application in this way incurs cost (it is necessary to determine which members of the population have the most influence, and to persuade them to use the application). The recommender scenarios provide a less-costly, easier-to-implement solution.

8.3.1 *Setup*

We varied the CTA-uptake percentage for the four uptake scenarios on six networks, each with $500,000$ individuals and average degree 13 (average degree 13 roughly corresponds to empirical findings concerning the number of contacts per individual [194]), using the network models from Section 8.2. In each case, we studied the effect of the CTA-uptake percentage and the uptake scenario (randomly chosen, degree-targetted, basic recommender or ring recommender) on the three KPIs from Section 8.2.4. We explain our results using the GIRGs from Section 8.2 with $\tau \in \{2.3, 3.3\}$ and $\alpha = \{1.3, 2.3\}$.

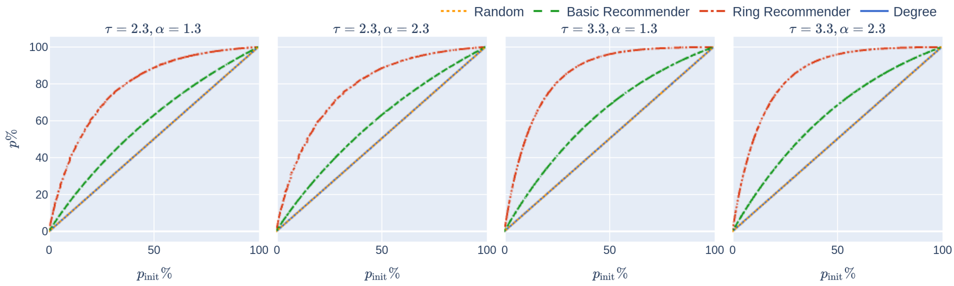


Figure 20: The initial app uptake $p_{\text{init}}\%$ in relation to the final app uptake $p\%$ on four networks and the different uptake scenarios. The curves for random uptake degree-targetted uptake coincide and are straight lines, since it holds that $p\% = p_{\text{init}}\%$, while for recommender scenarios $p\% > p_{\text{init}}\%$. Increasing $p\%$ increases the probability that a node is recommended by multiple persons, explaining the concave curves corresponding to the recommender scenarios.

The figures should be read as follows (unless mentioned otherwise). For each parameter value, the plotted result is the median over 10 runs. The shaded region around the plot covers the results of all 10 simulations. The epidemic is started by infecting 100 individuals, chosen uniformly at random. Simulation is halted when there are no more exposed or infected vertices.

8.3.2 *Hospital load and Epidemic Size: recommenders perform very well, ring recommender best*

The main message indicated by our simulations (row 2 of Figure 21) is that introducing recommender scenarios strongly increases the efficacy of a CTA in terms of reducing hospital load: having achieved uptake $p\%$ via any recommender scenario is significantly more desirable (in terms of reducing the maximum hospital load) than having achieved $p\%$ uptake by randomly chosen members of the population. Furthermore, the degree-targetted scenario works best, and the ring-recommender scenario is almost as good as this. Adding the subscripts rand, deg, basic, and ring to indicate the uptake scenarios, we find typically that

$$\text{HMax}_{\text{deg}}(p) < \text{HMax}_{\text{ring}}(p) < \text{HMax}_{\text{basic}}(p) < \text{HMax}_{\text{rand}}(p). \qquad (8.3.1)$$

We explain what we mean by “typically” and in what sense Inequality (8.3.1) is intended. Obviously, it is not strictly true — for example if the uptake percentage $p\%$ is 0% or 100%, then all uptake scenarios are

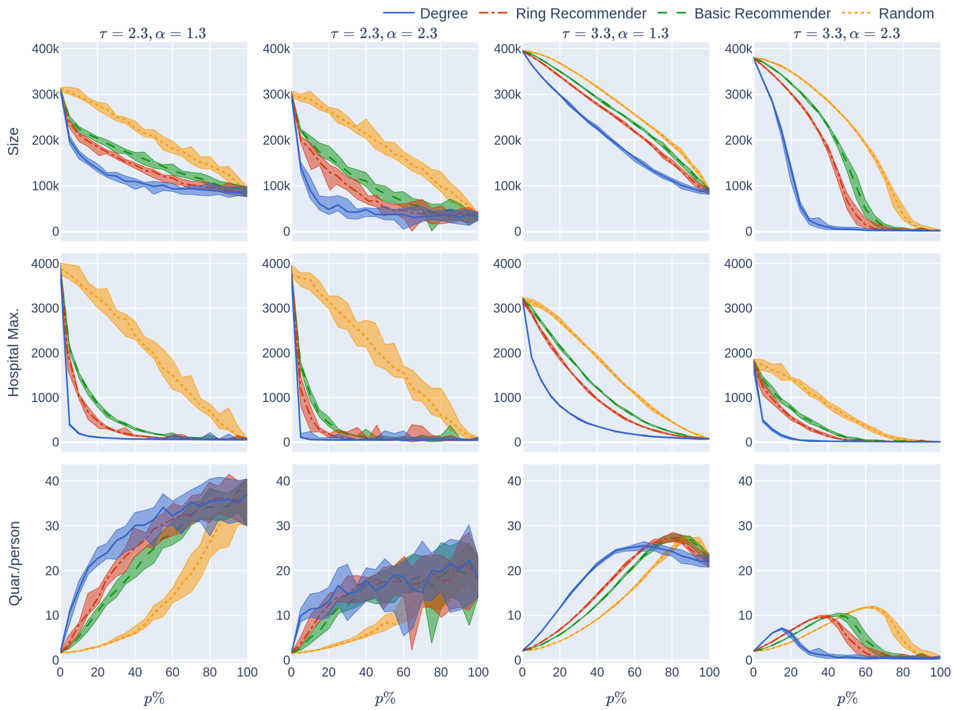


Figure 21: The KPIs (Size, HMax, Quar) as a function of the CTA-uptake percentage $p\%$ for the four CTA-uptake scenarios can be found in the rows. The population is modeled using four GIRGs (arranged in the four columns): $\tau = 2.3$ has more super-spreaders and $\tau = 3.3$ has fewer super-spreaders, $\alpha = 1.3$ is less geometric and $\alpha = 2.3$ is more geometric. The x-axis shows the uptake percentage $p\%$ varying from 0 to 100% at step size 5%. Simulations are done for the 21 values of $p\%$ corresponding to these steps. The y-axis shows the corresponding value of the KPI. The four uptake scenarios correspond to the four curves on each figure. The simulations use $\beta = 0.05$ and $q = 0.6$.

equivalent. Furthermore, our results are numerical, and hence have some inaccuracies. As can be seen from the second column of row 2 of the figure, if the CTA works well enough to suppress the epidemic then the various uptake scenarios are roughly equivalent, and cannot be compared. What we mean by "typically" is that, as is apparent from row 2, when the CTA uptake does matter, our data strongly suggest that the hospital max is smallest under the degree-based scenario, next smallest under the ring scenario, and largest (by some measure!) under the random scenario.

The same inequalities typically hold true for the size of the epidemic (first row of Figure 21):

$$\mathrm{Size}_{\mathrm{deg}}(p) < \mathrm{Size}_{\mathrm{ring}}(p) < \mathrm{Size}_{\mathrm{basic}}(p) < \mathrm{Size}_{\mathrm{rand}}(p).$$

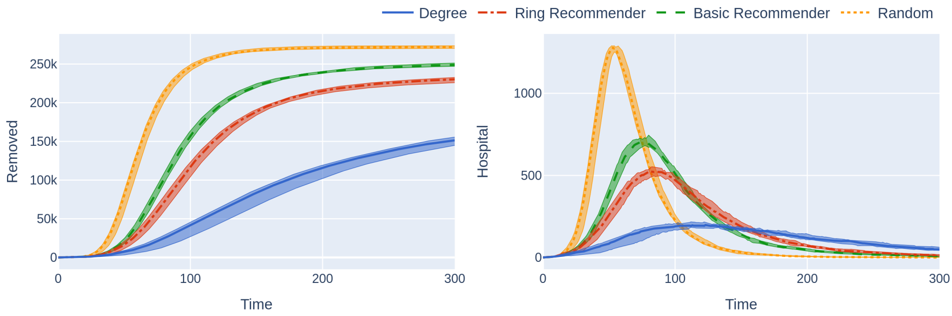


Figure 22: Simulations of an epidemic on two GIRGs with few super-spreaders ($\tau = 3.3$) and less geometry ($\alpha = 1.3$), with uptake percentage $p\% = 55\%$. The x-axis shows the time steps over which the epidemic is simulated. The y-axis of the left figure shows the number of removed individuals over the course of the epidemic. The y-axis of the right figure shows the number of hospitalised individuals over time (in the same simulation). The plotted curves correspond to the median over 20 runs (in the four scenarios); the shaded area covers the 25th-75th percentile of these runs over time. The left figure shows that the random uptake scenario is outperformed by all other scenarios and has a steeper incline, leading to a situation where more people get infected. The right figure shows that the number of hospitalised individuals is also higher under the random uptake scenario. This figure corresponds to the $p = 55\%$ value in the first two rows of the leftmost column of Figure 21.

Figure 22 shows a different depiction of this phenomenon at $p = 5\%$. There, the course of the epidemic is plotted against *time*. The epidemic curve on the left represents the total number of infected individuals (during the course of the epidemic). The epidemic curve on the right represents the number of individuals that are currently infected. The figure demonstrates how the random-uptake scenario is outperformed by the other uptake scenarios.

The reason that recommender scenarios always perform better than the random uptake scenario is explained by the so-called friendship-paradox [90]: a friend of a random user is likely to have a higher number of contacts than the random user itself, so a recommender scenario intrinsically finds users with high numbers of connections. This effect is more pronounced when the node-degrees are highly varying ($\tau = 2.3$) and it is even more pronounced in the ring-recommendation scenario. The reason for this is that ring recommendation introduces highly connected clusters of CTA users, and it is hard for the epidemic to thrive in such an environment: the CTA clusters block its spreading.

8.3.3 *Improving CTA efficacy by slightly increasing uptake rates*

We study the shape of the curves $\mathrm{Size}_{\square}(p)$ and $\mathrm{HMax}_{\square}(p)$ as p varies and $\square = \mathrm{rand}, \mathrm{basic}, \mathrm{ring}, \mathrm{deg}$. Intuitively, we examine what happens when we slightly increase the CTA-uptake percentage, especially in the (at least in Europe) more realistic low-uptake regimes ($< 50\%$).

Our first finding (see the second row of Figure 21) is that random uptake decreases HMax roughly *linearly* in the uptake percentage $p\%$. Here we do not mean that the plotted yellow curves in the second row (which are numerical approximations to $\mathrm{HMax}_{\mathrm{rand}}(p)$) are precisely linear — instead we mean that the rate at which the hospital maximum decreases is *roughly* constant, as the uptake percentage $p\%$ increases. This rate of decrease depends on the underlying network (the four columns). It appears that the epidemic size $\mathrm{Size}_{\mathrm{rand}}(p)$ when $\tau = 2.3$ (see the first row of Figure 21) decreases concavely as the uptake percentage $p\%$ increases, meaning that the size decreases more quickly for larger percentages $p\%$.

A similar phenomenon has been observed for HMax for higher values of the infection rate, β: the function HMax decreases concavely as $p\%$ increases [100] for all networks (also when $\tau = 3.3$).

We explain why the approximately linear decline is intuitively surprising, where it appears. For an infection to be prevented via the CTA, both the infector and the infected individual have to use the CTA, and the chance of this by random uptake is $(p/100)^2 \ll p/100$. Thus one's first guess would be that $\mathrm{HMax}_{\mathrm{rand}}(p)$ and $\mathrm{Size}_{\mathrm{rand}}(p)$ would be concave in p, not linear, over all infectiousness parameters. Yet, our results combined with [100] show that the curve is concave only in scenarios with high infectiousness and in contact networks with strong underlying geometry. The conclusion is that, even with random uptake, the CTA works better than might be expected, decreasing Size and HMax roughly linearly for a large range of parameter values (of infectiousness, and of the underlying network), rather than concavely.

The recommender scenarios react even better to small increases in the uptake percentage (for small values of p). Our experiments show (see the first row of Figure 21) that all recommender scenarios (and especially the ring-recommender scenario) react more strongly (than the random-uptake scenario) to a slight increase in the uptake percentage $p\%$. That is, the epidemic size curves decrease at a faster rate under basic and ring recommendation, as compared to under the random uptake scenario. Roughly, we find that for $p\% < 50\%$ and a small positive number dp,

$$\mathrm{Size}_{\mathrm{ring}}(p) - \mathrm{Size}_{\mathrm{ring}}(p + dp) \gg \mathrm{Size}_{\mathrm{rand}}(p) - \mathrm{Size}_{\mathrm{rand}}(p + dp) > 0,$$
$$\mathrm{Size}_{\mathrm{basic}}(p) - \mathrm{Size}_{\mathrm{basic}}(p + dp) \gg \mathrm{Size}_{\mathrm{rand}}(p) - \mathrm{Size}_{\mathrm{rand}}(p + dp) > 0,$$

meaning that (apart from perhaps for $\tau = 3.3, \alpha = 1.3$), the total size of the epidemic as a function of p decreases substantially more steeply for ring-recommender and basic-recommender uptake than for the random uptake. While this is true for both $\tau = 2.3$ and $\tau = 3.3$, in the situation where there are many super-spreaders ($\tau = 2.3$), the curves seem in addition to be convex (meaning that the rate of decay is highest when the uptake percentage is small).

For the maximum hospital load HMax, the effect is similar, but stronger (see row 2 of Figure 21). Here the HMax curve is a convex curve for all recommender scenarios.

In summary, we find that for many underlying contact networks, the functions $\mathrm{Size}_{\mathrm{ring}}(p)$, $\mathrm{Size}_{\mathrm{basic}}(p)$, $\mathrm{HMax}_{\mathrm{ring}}(p)$, and $\mathrm{HMax}_{\mathrm{basic}}(p)$ are convex, having a steep decline when $p\% < 50\%$, while $\mathrm{Size}_{\mathrm{rand}}(p)$ and $\mathrm{HMax}_{\mathrm{rand}}(p)$ are linear or concave functions. This is illustrated further in [100, Figures 4–5] where the level of infectiousness is varied: the phenomenon is robust in the "supercritical" regime where there is a large outbreak, apart from in graphs with few super-spreaders in situations with very high infection rates. In this latter case the curves might coincide or are more concave.

8.3.4 *Geometry helps the CTA*

A less obvious observation that we make is that (considering networks with the same average number of connections per node), the more geometric the underlying contact network is, the better the performance of the CTA in terms of reducing the size of the epidemic and the maximum hospital load, across the four uptake scenarios. In other words, lack of long-range connections helps CTAs. This agrees with existing work on classical contact tracing [87, 160], despite the increased heterogeneity introduced by random app uptake; see [196] for a detailed survey.

Moreover, the positive effect of recommending becomes more exaggerated in geometric networks, see Figure 23. Emphasising in notation by adding a superscript 'geo' and 'ageo' to emphasise whether the underlying

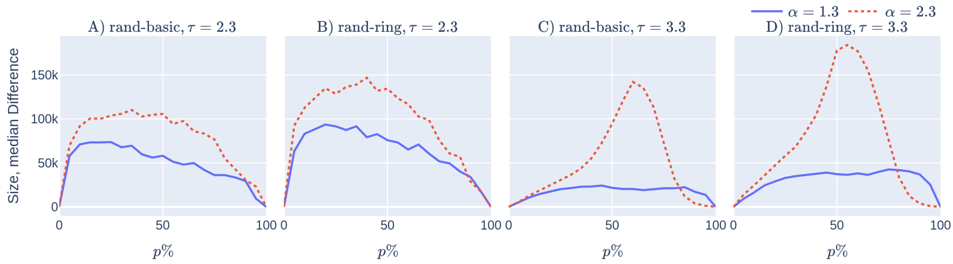


Figure 23: The reduction in epidemic size achieved by CTAs with user recommendation improves when the network has stronger geometry (for all values of $p\% < 85\%$). Subfigures (A) and (C): The red dashed curve shows $\mathrm{Size}^{\mathrm{geo}}_{\mathrm{rand}}(p) - \mathrm{Size}^{\mathrm{geo}}_{\mathrm{basic}}(p)$, while the blue continuous curve shows $\mathrm{Size}^{\mathrm{ageo}}_{\mathrm{rand}}(p) - \mathrm{Size}^{\mathrm{ageo}}_{\mathrm{basic}}(p)$ as a function of the uptake percentage $p\%$. In subfigures (B), and (D), the red dashed curve shows $\mathrm{Size}^{\mathrm{geo}}_{\mathrm{rand}}(p) - \mathrm{Size}^{\mathrm{geo}}_{\mathrm{ring}}(p)$, while the blue continuous curve shows $\mathrm{Size}^{\mathrm{ageo}}_{\mathrm{rand}}(p) - \mathrm{Size}^{\mathrm{ageo}}_{\mathrm{ring}}(p)$ as a function of the uptake percentage $p\%$. Observe that the red dashed curves are (typically) above the blue curves. In subfigures (A) and (B) the underlying network contains many super-spreaders, while in subfigures (C) and (D) the network contains fewer: observe that the presence of fewer super-spreaders makes the red dashed curve, the epidemic size difference in geometric networks between recommender and random uptake, more spiky. The simulation details are the same as those in Figure 21. For each parameter value, the plotted result is the median over 10 runs.

network has strong geometry ($\alpha = 2.3$) or less geometry ($\alpha = 1.3$), we typically find that (for fixed τ and reasonable uptake percentages)

$$\begin{aligned} \mathrm{Size}^{\mathrm{geo}}_{\mathrm{rand}}(p) - \mathrm{Size}^{\mathrm{geo}}_{\mathrm{basic}}(p) &\gg \mathrm{Size}^{\mathrm{ageo}}_{\mathrm{rand}}(p) - \mathrm{Size}^{\mathrm{ageo}}_{\mathrm{basic}}(p), \\ \mathrm{Size}^{\mathrm{geo}}_{\mathrm{rand}}(p) - \mathrm{Size}^{\mathrm{geo}}_{\mathrm{ring}}(p) &\gg \mathrm{Size}^{\mathrm{ageo}}_{\mathrm{rand}}(p) - \mathrm{Size}^{\mathrm{ageo}}_{\mathrm{ring}}(p). \end{aligned} \tag{8.3.2}$$

This is illustrated on Figure 23, where the epidemic size difference (corresponding to the number of infections that are prevented by the recommendation) is plotted. Observe that the dashed red curve, corresponding to the epidemic size difference in geometric networks stays above the blue curve, corresponding to the epidemic size difference in ageometric networks. We emphasise that both curves are positive, i.e., recommendation performs better than random uptake in both geometric and ageometric networks. The intuitive explanation for Inequality (8.3.2) is the same as what we gave in Section 8.3.2: recommending has two main effects. The first effect is finding the super-spreaders via the friendship-paradox effect. The second effect of recommending is only present in geometric networks: recommending also forms geometric barriers of CTA-users around a central node. These barriers are hard for the infection to pass through.

On the empirical side, we mention that a way to introduce more geometry into the underlying contact network is via travel restrictions as in Chapter 7.

8.3.5 *How costly is quarantine?*

One might be afraid that using a CTA might cause a high quarantine load on the population. However, row 3 of Figure 21 demonstrates that, at least in the case of fewer super-spreaders ($\tau = 3.3$) the average number of days that an individual has to quarantine does not rise very steeply with the CTA-uptake percentage $p\%$. Note that the date for the case with more superspreaders ($\tau = 2.3$) has very high variance, so we do not draw conclusions about this case. Note also that the conclusions are relevant in situations where social-distancing measures make the underlying infection rate not too high, this corresponds to the choice of $\beta = 0.05$ here. Though it is beyond the scope of this study, note that quarantine time can be reduced by providing prompt testing to people who are notified to quarantine by the CTA.

Here we note a surprising feature that applies to underlying networks with relatively few super-spreaders ($\tau = 3.3$) in Figure 21. In this case, $\mathrm{Quar}(p)$ is roughly unimodal, meaning that when the uptake percentage $p\%$ is small, the amount of quarantine time per person increases as p increases. However, roughly, the amount of quarantine is maximised for some $p^\star\% < 100\%$, and decreases above $p^\star$.

This means that using CTAs only increases the average quarantine time below uptake $p^\star$ but not above it. We mention that $p^\star$ *depends on the uptake scenario*, and comes earlier with recommendation scenarios than with random uptake. In [100, Figure 7] the same phenomenon is observed for various levels of infectiousness of the disease: this rough unimodality in $\mathrm{Quar}(p)$ is robust for these underlying networks. However, it disappears for networks where either there are more super-spreaders (so the data is too noisy to draw conclusions) or the geometry is less apparent.

We conclude that, in this particular setting, using CTAs at high uptake rates is effective not only for reducing the epidemic size and maximum hospital load of the epidemic but also in reducing the average quarantine length.

8.3.6 *Other parameter values*

The experiments in Figure 21 assumed an epidemic model (see Section 8.2.2) with a low rate of infection $\beta = 0.05$. This means that each infectious node infects each of its susceptible neighbours with probability β at every time step (e.g., every day). Low rates of infection can be achieved by social-distancing measures. The results that we have mentioned so far are robust over other infection rates, see [100, Figure 4–7].

8.4 EXPERIMENT 2: STRICTNESS OF QUARANTINE MEASURES

It has been observed that the severity of symptoms varies with the individual [53, 183, 205, 226]. In our model, we assume that there is an underlying probability distribution that determines whether an individual would be symptomatic in the case that this individual is infected, and also that there is an underlying probability distribution determining the severity of these symptoms. Policy makers may then impose quarantine or home isolation on individuals that have symptoms above a certain severity-threshold; setting the threshold is a political/economic decision. In our model, the strictness of quarantine measures is represented by a value $q \in [0, 1]$. This quantity q is the fraction of individuals whose symptoms would be so severe that they would go into home-quarantine if infected. Strict quarantine measures correspond to large values of q – in the extreme, taking $q = 1$ means that all infected individuals would go into home-quarantine. We recall that we work under the assumption that all relevant symptoms are caused by this virus, which is important for interpreting the results in this section.

For any given epidemic, there is an underlying probability q_{symp}, which is the proportion of infected individuals that experiences any symptoms upon infection (mathematically, the probability of showing symptoms upon infection). Implementing a quarantine-strictness $q > q_{symp}$ would require a program of mass testing. The implementation of such a testing program is beyond the scope of this chapter, however we make two remarks:

(1) For many epidemics, the value of q_{symp} is fairly high (for COVID-19, the literature is varied, [214] finds a $q_{symp} \approx 0.3$ over all population, varying in age, [187, 203] find it ≈ 0.6 and ≈ 0.7, respectively, while [118] finds it as high as ≈ 0.85).

(2) our results indicate that setting quarantine-strictness even below q_{symp} can still be effective for decreasing the epidemic size and maximum hospital load.

The social benefits of quarantining measures are nuanced. While quarantining clearly helps to reduce the size of the epidemic (Size) and the maximum hospital load (HMax), it also comes at the high societal cost of reduced work capacity. As a policy maker, one might like to balance the reduction in Size and HMax against Quar, the average number of days spent in quarantine per person, over the course of the epidemic. In Experiment 1 we took the fixed value $q = 0.6$. In Experiment 2, we study how the setting of the quarantine-strictness q affects all of these KPIs. In this experiment, we vary the quarantining strength q and the CTA-uptake percentage p% for the four CTA-uptake scenarios. Our results are summarised in Figure 24, that we elaborate below.

8.4.1 *Reducing epidemic size & hospital load by stricter quarantining*

Increasing the quarantine-strictness q (i.e., requiring individuals to quarantine even with mild symptoms) reduces the number of individuals who are able to spread the virus, and hence reduces the epidemic size and maximum hospital load of the epidemic. This is a very natural observation, valid across all parameter settings and underlying contact networks. The top row of Figure 24 shows how the size of the epidemic decreases, for various CTA-uptake percentages p%, as the quarantine-strictness q increases. We observe that for $q_1 > q_2$, for any fixed p and uptake scenario $\square = \text{rand}, \text{basic}, \text{ring}, \text{deg}$,

$$\text{Size}_{\square,p}(q_1) \leqslant \text{Size}_{\square,p}(q_2), \quad \text{HMax}_{\square,p}(q_1) \leqslant \text{HMax}_{\square,p}(q_2).$$

However, the rate of decrease varies across uptake scenarios. Figure 25 shows that in the case with fewer super-spreaders ($\tau = 3.3$) typically

$$\text{Size}_{\text{deg},p}(q) < \text{Size}_{\text{ring},p}(q) < \text{Size}_{\text{basic},p}(q) < \text{Size}_{\text{rand},p}(q), \tag{8.4.1}$$

by plotting the various uptake scenarios on the same diagram. To some extent, Inequality (8.4.1) also holds with more super-spreaders ($\tau = 2.3$) but the curves for the recommendation scenarios are close, or overlapping, in these cases. Similar results hold for HMax, as shown in [100].

Going back to Figure 24, we find that the Size and HMax curves are steepest under degree-targetted uptake, ring recommendation is second

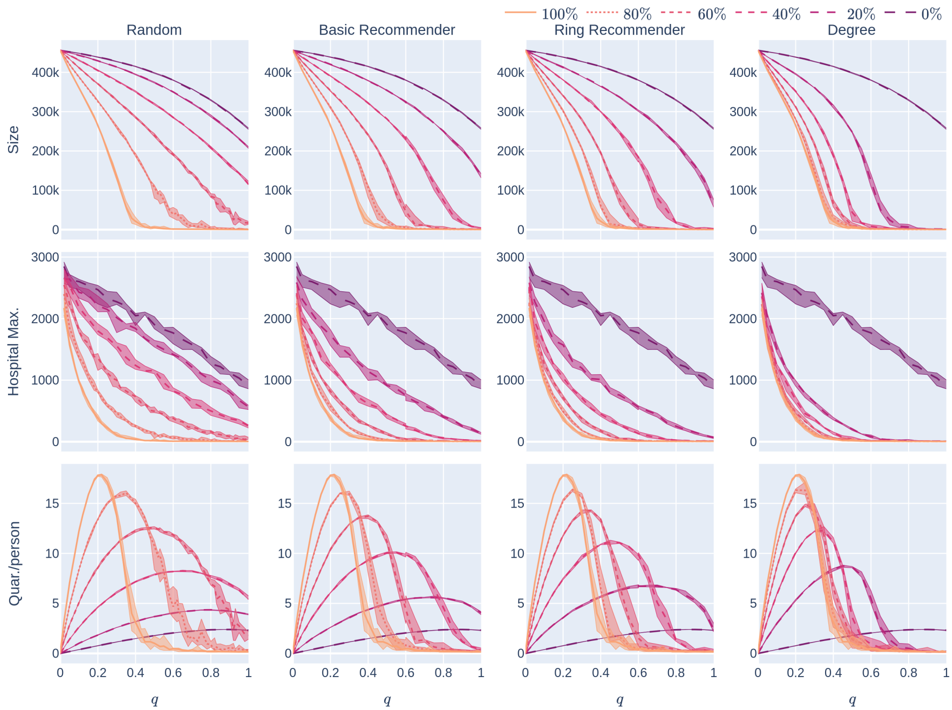
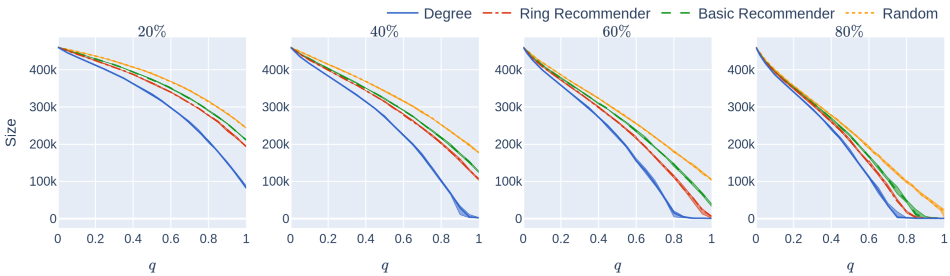
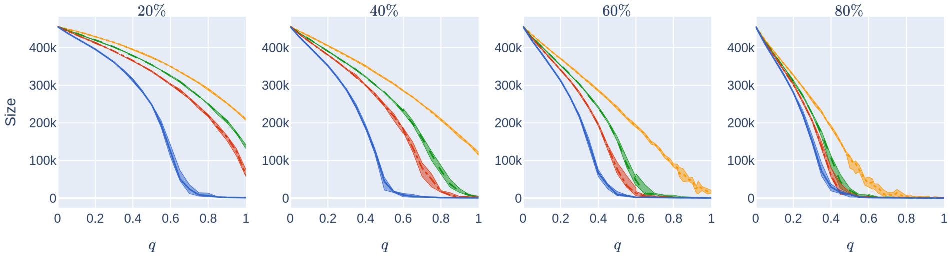


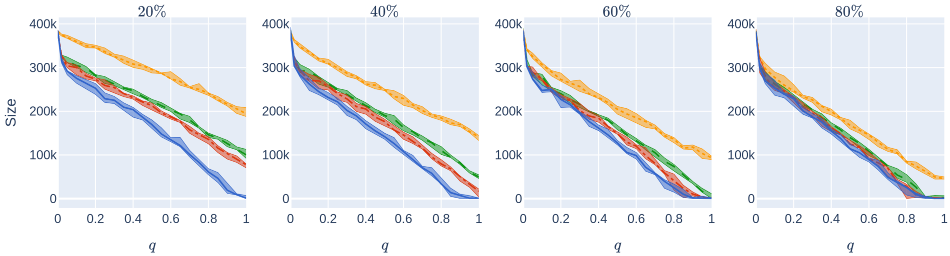
Figure 24: Influence of the quarantine strictness q for various uptake percentages p% and the four uptake scenarios: the KPIs are plotted against q, the quarantine strictness. We work here with a fixed underlying GIRG with $\tau = 3.3$ (fewer super-spreaders) and $\alpha = 2.3$ (more geometric). Each column represents a given uptake scenario (random, basic-recommender and ring-recommender, and degree-targetted). The x-axis shows the quarantine strictness q varying from 0 to 1 at step size 0.05. Simulations are done for the 21 values of q corresponding to these steps. There is an additional simulation point at 0.02 (in order to avoid division by 0 in the computation for HMax and still have a point close to 0). The y-axis shows the corresponding value of the KPI. The curves on each figure correspond to the different uptake percentages, as labelled. For each parameter value, the plotted result is the median over 5 runs. The shaded region around the plot covers the results of all 5 simulations. The infection rate is $\beta = 0.05$.



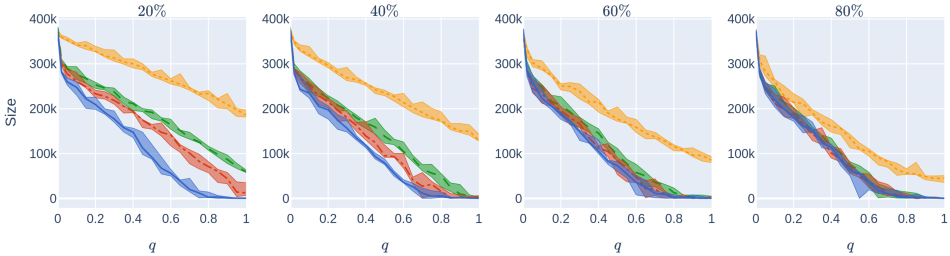
(a) Less geometric ($\alpha = 1.3$), few super-spreaders ($\tau = 3.3$).



(b) More geometric ($\alpha = 2.3$), few super-spreaders ($\tau = 3.3$).



(c) Less geometric ($\alpha = 1.3$), many super-spreaders ($\tau = 2.3$).



(d) More geometric ($\alpha = 2.3$), many super-spreaders ($\tau = 2.3$).

Figure 25: Experiment 2 on four GIRGs. Each column represents an uptake percentage. The x-axis shows the quarantine strictness q varying from 0 to 1 at step size 0.05. Simulations are done for the 21 values of q corresponding to these steps. The y-axis shows the corresponding value of the KPI. The curves in each figure correspond to the different uptake scenarios, as labelled. For each parameter value, the plotted result is the median over 5 runs. The shaded region around the plot covers the results of all 5 simulations.

best, while basic recommendation still performs better (declines more steeply) than random uptake. This is similar to what we have already observed in Section 8.3.

Figures 13–15 in [100] are qualitatively similar to Figure 24 and consider GIRGs with other parameters. For network with more super-spreaders, the effects are even more exaggerated. The observation that recommending helps significantly already at very low uptake rates remains valid.

8.4.2 *Improving CTA efficacy by slightly increasing quarantine strength*

We next study the shape of the curves $\mathrm{Size}_{\square,p}(q)$ and $\mathrm{HMax}_{\square,p}(q)$ for fixed p as q varies and $\square = \text{rand}, \text{basic}, \text{ring}, \text{deg}$. Intuitively, we examine what happens when we slightly increase the quarantine strength q, especially in the mild-quarantine scenario where $q < 0.5$. We first note (From Figure 24) that without a CTA present ($p = 0\%$), HMax decreases roughly linearly as a function of q. Whenever a CTA is present, ($p\% > 0\%$) HMax becomes a *steeply decreasing convex* function of q for low q.

The main message is: a somewhat stricter quarantine rule can already lower the maximum hospital load and thus "*flatten the curve*" very well in the presence of a CTA. A similar observation is valid for $\mathrm{Size}_p(q)$: the curve is concave for $p\% = 0\%$, while it becomes linear or convex when p increases.

When the value of q is small (so the quarantine measures are not strict), Figure 24 also shows that another way to flatten the curve is to increase the uptake percentage: the rate at which Size and HMax decrease is much higher for large CTA-uptake percentages: our numerical simulations indicate that (for the values of p that we have tested) when $p_1 > p_2$ and dq is a small positive number,

$$\mathrm{Size}_{p_1}(q) - \mathrm{Size}_{p_1}(q + dq) \gg \mathrm{Size}_{p_2}(q) - \mathrm{Size}_{p_2}(q + dq),$$
$$\mathrm{HMax}_{p_1}(q) - \mathrm{HMax}_{p_1}(q + dq) \gg \mathrm{HMax}_{p_2}(q) - \mathrm{HMax}_{p_2}(q + dq).$$

Thus a high(er) CTA uptake rate makes it possible to impose less strict quarantine measures and yet keep the epidemic size and hospital load under control.

We mention that these findings are robust across underlying networks. Figures 13–15 in [100] illustrate qualitatively similar behaviour for networks with more long connections and super-spreaders.

8.4.3 *Reducing quarantining by quarantining*

Although it is valuable to increase quarantine-strictness in order to reduce Size and HMax, one may wonder whether stricter quarantine measures require more social sacrifice. Our experiments demonstrate that the picture is not this simple: Row 3 of Figure 24 shows that, for every positive CTA-uptake percentage $p\%$ and every uptake scenario $\square = \text{rand}, \text{basic}, \text{ring}, \text{deg}$, the curve $\text{Quar}_{\square,p}(q)$ is roughly unimodal. Thus, once q is sufficiently large (larger than some value $q^*_{\square,p}$ depending on the uptake scenario $\square$ and the uptake percentage $p\%$), increasing q actually decreases the average number of days that people have to quarantine. One reason for this might be that the quarantining decreases the spread of the epidemic (so other individuals do not get infected and cause further quarantining).

This observation is most useful if the CTA-uptake percentage $p\%$ is high enough that $q^*_{\square,p} < q_{\text{symp}}$, so that quarantine-strictness quantities $q > q^*_{\square,p}$ are easy to implement in reality. Note that $q^\star_{\square,p}$ also depends on the infectiousness of the disease. Stricter social distancing, corresponding to lower reproduction number, leads to lower value of $q^\star_{\square,p}$. Thus, when social-distancing measures are in place, it will be more possible to reduce the average time that people quarantine by setting $q > q^*_{\square,p}$.

8.5 CONCLUSION AND OUTLOOK

The research leading to Chapters 7 and 8 was performed in the early stage of the COVID-19 pandemic (first six months of 2020), during which there were many uncertainties about the evolution of the epidemic and the impact of intervention measures (including contact-tracing applications). Some of the results were presented to local governments. Although the addressed research questions are less pressing at the moment, some of the insights could still be of future importance from an epidemiological perspective. Besides that, the simulation-based results show interesting phenomena, which would be interesting to study from a mathematically rigorous perspective.

We have used scale-free networks embedded in a geometric space to study interventions methods. The geometry allows for intuitive modelling of several intervention methods as seen in Chapter 7 for an SIRS epidemic: social distancing, travel restrictions, and meeting a limited number of people. We compared the effect of these interventions by 'permanently' modifying the underlying contact network, while maintaining the same average node-degree across interventions. We found that the strong travel restrictions are most effective in elongating and diminishing the first peak: the shape changes from exponential to linear. However, travel restrictions and meeting a limited number of people result in a higher second peak, where the latter restriction yields the highest second peak. We could explain the effects of the intervention measures on an intuitive level, by relating it to different universality classes for the typical graph distance. Our main conclusion was that the behaviour of the epidemic highly depends on the topology of the network.

An interesting topic for future research is to study the fluctuations of the epidemic curve immediately after the first peak. In models with infinite variance degrees ($\tau < 3$, similar to the preferential attachment models considered in Part i), the oscillations seem to be less significant than in models with $\tau > 3$. This result is even more exaggerated if the typical graph distance is not of poly-logarithmic order.

One of the main goals of Chapter 7 was to compare the effect of an underlying geometry influences the spread to two inherently non-geometric network models: the configuration model and standard continuous compartmental models. It would be of interest to involve spatial variants of the compartmental models into this comparison [80, 120, 224].

Several adjustments can be made to make the model more realistic. Other compartments could be added, or one could study a scenario where interventions are only applied once certain thresholds in the number of infected nodes are exceeded, or are being lifted.

In Chapter 8 we considered a different epidemic model: instead of temporary immunity we considered long-lasting immunity; besides that, the presence of a contact-tracing application that induces quarantining changed the dynamics significantly. While some of the observations (i.e., concavity or convexity of the size of the epidemic as the percentage of CTA-users/quarantining-strictness is varied) are possibly hard to prove rigorously, it would be of interest to prove that stricter quarantining can lead to less quarantining for large enough values of q (possibly for a simpler model than GIRGs), i.e., to show theoretically that the curve $\mathrm{Quar}_p(q)$ described in Section 8.4.3 is indeed unimodal.

Discussion and open problems

9 DISCUSSION AND OPEN PROBLEMS

In this thesis, we studied graph and weighted distances, and component sizes in various scale-free random graphs. Our theoretical results from Part i on distances in random graphs whose degree distribution has infinite variance, gave (together with related results from literature) intuition for the possible spread of COVID-19 as illustrated in our simulation-based studies in Part iii. These studies showed that the underlying (possibly spatial) structure of networks heavily influences epidemics. In Part ii we studied the component structure in a large class of spatial random graph models, which contains geometric inhomogeneous random graphs (the random graph model used in Part iii) as a special case.

In Section 8.5 we have already mentioned possible directions for future research related to Part iii. As Parts i and ii form the main contribution of this thesis to the mathematical literature, we conclude the thesis by discussing some possible avenues for future research that are related to these parts.

9.1 FIRST-PASSAGE PERCOLATION IN SCALE-FREE RANDOM GRAPHS

In Chapter 2 we established the first results on weighted distances in the three preferential attachment models (PAMs) defined in Section 1.2.2: fixed preferential attachment (FPA), variable preferential attachment (VPA), and generalised variable preferential attachment (GVPA). The same universality classes of weight distributions from Definition 2.1.3 appear for static models as the Configuration Model (CM) [2, 22, 21], and for the spatial models scale-free percolation (SFP), geometric inhomogeneous random graphs (GIRG), hyperbolic random graphs (HRG) [136, 166], when $\tau \in (2,3)$. Our technique developed for a *dynamic* model can be adapted to obtain similar results on static models for weight distributions in the conservative class. In particular, the improved recursion on $(s_k)_{k\geqslant 0}$ in (2.3.9) above can be used to prove tightness of typical weighted distances in CM for $\tau \in (2,3)$ around the main term under condition (2.2.3), proving part of [2, Problem 2.10].

For a fixed $\tau \in (2,3)$, the main difference between CM and the preferential attachment models studied in Part i is that all distances in the three PAMs are roughly twice the distance in CM. For CM, the main term of the *graph* distance is $2\log\log(t)/|\log(\tau-2)|$ [132], compared to $4\log\log(t)/|\log(\tau-2)|$ in PA. Combining the results on weighted distances in CM [2] with our results, after a variable transformation on the sum in Q_t in (2.1.3), the factor two extends to *weighted* distances, i.e.,

$$d^{(L)}_{\mathrm{CM}_t(\tau)}(u,v)\Big/d^{(L)}_{\mathrm{PA}_t(\tau)}(u,v) \longrightarrow 2\,, \qquad \text{as } t \to \infty.$$

This is due to the difference of the geometry of typical shortest paths: in CM high-degree vertices are often directly connected via an edge, while in the PAMs we need two edges to connect two high-degree vertices. This factor two is studied in [72], where it is illustrated how this factor two vanishes in GVPA(f) (see Definition 1.2.5) with power-law exponent $\tau = 3$ and a function f that has logarithmic corrections.

For spatial preferential attachment as introduced in [142], an upper bound on the graph distance is established in [123]: it uses a similar *two-connector* procedure that we also use in Chapters 2 and 3. In the proof for the upper bound, the authors show that two vertices with high degree are connected via two edges, similarly to the non-spatial models. To the extent of our knowledge, no matching lower bound for spatial preferential attachment is known, although for the age-dependent random connection model (a variant of the spatial preferential attachment), matching upper and lower bounds are known when the typical graph distance scales as $\Theta(\log\log(t))$ [106]. Thus, it would be interesting to translate our results on the typical weighted distance to these models.

Another interesting property is the hopcount, the number of edges on the least-weighted path. For Erdős-Rényi random graphs (ERRG) with i.i.d. exponential weights on the edges, it is known that the hopcount is much larger than the graph distance [27]. We conjecture that a similar result should hold for preferential attachment models with explosive edge weights, as these models contain a subgraph on at least $\sqrt{t}$ vertices (called the inner core in Section 2.3) that dominates a dense ERRG. Moreover, we expect that the hopcount in variable preferential attachment and fixed preferential attachment is tight around the sequence $2K_t^*$ for any weight distribution of the form $L = 1 + X$ with $I(X) = \infty$, improving Theorem 2.2.6. To show this, however, a better upper bound is necessary. For the conservative class, extending Theorem 2.2.6 on the hopcount to weights that are not bounded away from zero is more difficult, and requires a detailed study of the hopcount within the dense inner core. We also believe that (2.2.3) in Theorem 2.2.5 is not a necessary condition. However, we were not able to remove it in our proof of the upper bound.

Lastly, it would be interesting to study the geodesic (i.e., the least-weighted path) in the neighbourhood of u and v in more detail. Would it be possible to prove local weak limits of the geodesic of the parts close to u and v? For CM, local weak limit theorems are established in [74]. We conjecture that using Theorem 2.2.9 in the present chapter and results from [24, 71, 97], similar results can be derived for PAMs. Another interesting question would be to analyse the age distribution of the vertices on the geodesic beyond the local neighbourhood of u and v.

9.2 EVOLVING PROPERTIES IN PAMS AND RELATED MODELS

Chapter 3 commences a research line by studying an evolving graph property (other than the degree of fixed vertices [70, 193]). Statements involving the evolution of a property describe the structure of the graph during a *time interval*, rather than at a *single time*. We consider the evolution of a *global* graph property: the distance evolution in two classical PAMs, FPA and VPA. Studying a global graph property requires a more fine-grained control of the *entire graph* than required for a *local* graph property, such as the degree evolution of a *fixed vertex*, and yields more insight on the evolution of the structure of the graph.

One of the main reasons to consider distances for these PAMs is that they display a notable change over time. The growth terms decrease from $\log\log$-order to constant order as the graph grows. This is in contrary to, for example, the local clustering coefficient, a graph property related to the number of triangles which a typical vertex is a member of. The local clustering coefficient of a typical vertex is of constant order and tends to zero for typical vertices due to the locally tree-like structure in classical PAMs. However, one could analyse the local clustering coefficient in versions of PAMs that are not locally tree-like. Static analyses of the local clustering coefficient on spatial variants of PAMs have been done in [105, 142]. The local clustering coefficient of the age-dependent random connection model, a spatial variant of PAMs, has been done in [135].

A natural extension would be to study the distance evolution in PAMs where the asymptotic degree distribution has finite variance. For this regime, it is known that the *static* typical graph distance is of order $\Theta(\log(t))$, but the precise constant has not been determined [81]. We expect that in this regime the time-scaling of the growth is different from the scaling of the hydrodynamic limit in Corollary 3.1.2 and to see the distance drop by a constant factor when t'/t is of polynomial order, rather than stretched exponential in the logarithm.

As mentioned above, matching lower and upper bounds were established for graph distances for some regimes in the age-dependent random connection model [106]. Parts of our techniques should help to establish results on the distance evolution for this spatial model. The age-dependent random connection model is a special case of the kernel-based spatial random graphs (KSRGs) studied in Part ii (it uses the kernel κ_{pa}, and the marks should be interpreted as a rescaled age). Relating Parts i and ii of this thesis, it would be of interest to understand the evolution of the component-size distribution in the age-dependent random connection model.

In most PAMs the graph and its edge set are increasing over time. In [57, 65] variations of PAMs are introduced where edges can be deleted. As a result the distance evolution is no longer monotone and other behaviour may be expected. The variations of PAMs mentioned in Section 1.2.3 all have properties that can be considered from a non-static perspective.

9.3 CLUSTER-SIZE DISTRIBUTION IN SPATIAL RANDOM GRAPHS

In Part ii and the accompanying paper [151] we study the cluster-size decay of kernel-based spatial random graphs (KSRGs), and consider the interpolating kernel $\kappa_{1,\sigma}$ in (1.3.5) to obtain the interpolation KSRG (i-KSRG) that contains many well-known models as special cases. These works open up several avenues for future research beyond proving Conjecture 4.2.1 for the entire parameter regime (without $o(1)$-errors, and also for i-KSRGs on $\mathbb{Z}^d$).

Monotonicity of ζ in the nearest-neighbour regime

In the regimes where we have proofs for the cluster-size decay (6.1.2), the stretch-exponent $\zeta = \zeta(\alpha, \tau, \sigma, d)$ is monotone in the parameters. For parameters such that $\zeta_{\mathrm{nn}} = \max\{\zeta_{\mathrm{ll}}, \zeta_{\mathrm{lh}}, \zeta_{\mathrm{hh}}\}$ (see (4.2.4), (4.2.2), (4.2.1), (4.2.5)) we can still show that the stretch-exponent ζ equals ζ_{nn}. Moreover, for $\alpha, \tau = \infty$ (corresponding to nearest-neighbour bond percolation or the soft random geometric graph), the cluster-size decay is also known to be stretched exponential with parameter $\zeta = (d-1)/d$ [110, 158] (for the soft random geometric graph it is only proven for $d = 2$ in [175, 211]).

This naturally raises the question whether ζ exists for any parameters α, τ, σ, d, p, β that give rise to a supercritical i-KSRG. The non-monotonicity of the second-largest component under inclusion of edges makes this highly non-trivial. A positive answer to the question would prove Conjecture 4.2.1 additionally for the parameter regime where $\zeta_{nn} > \max\{\zeta_{ll}, \zeta_{lh}, \zeta_{hh}\}$.

Deviation principles for the size of the giant component.

Our results characterise the speed of the lower tail bound of large deviations (LTLD) for the size of the largest component (up to $o(1)$ error in the exponent). It would be a natural (but far from trivial) extension to remove the $o(1)$-error and determine also *rate* of the LTLD, thus obtaining a large-deviation principle for the lower tail.

Instead of considering the lower tail bound for large deviations (of which we show that it has *stretched-exponential* decay in many cases), an alternative is to consider the upper tail bound. We expect that this bound decays as a *polynomial* in n, rather than stretched exponentially. A related work in progress is [133], which studies the upper tail bound for the total number of edges in KSRGs.

As a side result of our study of *small* components, we established a weak law of large numbers for the size of the *largest* component. Besides studying large deviations of the size of the largest component, it would be of interest to study lower-order deviations for the size of the giant component, and to see when a central limit theorem (CLT) is satisfied. In [92], it is shown that the number of isolated vertices in hyperbolic random graphs, a 1-dimensional KSRG using the product kernel, satisfies a CLT if $\tau > 3$ (when the degree distribution has finite variance) and does not satisfy a CLT when $\tau \in (2, 3)$ (corresponding to a degree distribution that asymptotically has infinite variance). Of course, limit theorems could also be proven for other graph properties than the size of the giant component, e.g. independence number, domination number, set- and vertex-cover, similar to the laws of large numbers established for the random geometric graph [185].

Scaling limit of random walks.

The understanding of the geometry of clusters (both the largest/infinite component and isolated component) could help in studying other properties of i-KSRGs, in particular simple random walks on these graphs. Transience and recurrence of simple random walks on (special cases of) i-KSRGs with non-trivial kernels have been studied in [107, 121], while more recently the mixing time, and cover and hitting times have been studied in [52, 161, 163].

Related to the study of evolutional properties in Chapter 3, it would be of interest to study the evolution of the mixing time of a random walk in the age-dependent random connection model.

(Sub)critical and dense i-KSRGs.

While supercritical KSRGs with non-trivial kernels have been studied extensively in the past years, the (sub)critical counterparts have received less attention. The existence of subcritical and supercritical phases by varying β are studied in [108, 109, 144], and first steps to study properties of subcritical KSRGs are made in [79] for higher dimensions. Yet, further steps need to be taken before results for the cluster-size decay can be obtained similar to [6] for nearest-neighbour percolation (in which case the decay is exponential). For (sub)critical long-range percolation, a polynomial upper bound on $\mathbb{P}\big(|\mathcal{C}(0)| \geqslant n\big)$ is obtained in [138], but no lower bounds are proven (for subcritical long-range percolation it is expected that the decay is exponential, rather than polynomial).

For some parameter regimes of i-KSRGs, i.e., when $\tau < 2 + \sigma$ or $\alpha < 1/(\tau - 2)$, the model is supercritical for every value of $\beta \in (0, \infty)$ and $p \in (0, 1)$ and the model is called robust [109]. For those parameters it would be of interest to study sequences $(\beta_n)_{n \geqslant 1}$ tending to zero so that the sequence of KSRGs with parameters $(\alpha, \tau, \sigma, d, \beta_n)$ is in a *critical* window, or chosen *barely supercritical*. Defining and identifying such regimes – and finding a potential relation to ζ– would already be interesting, and would allow for comparison to (non-spatial) analogues of the models: the configuration model and rank-1 inhomogeneous random graph with infinite variance degrees. For such models and parameters, the critical window, and component sizes have been studied in a sequence of papers [25, 77, 78].

On the contrary, instead of considering sequences $(\beta_n)_{n\geqslant 1}$ tending to zero, one may consider sequences $(\beta_n)_{n\geqslant 1}$ that tend to infinity, so that the average degree is increasing in n. To the extent of our knowledge, no results on KSRGs are known for such dense models.

Variations and generalisations of KSRGs

We highlight a few variations of KSRGs that could be interesting from both a mathematical and (more) applied perspective. For instance, the vertex set could be embedded in other metric spaces, e.g. a Riemannian manifolds, or be formed by other point processes, or the edges may be directed.

Alternatively, the role of the marks could be changed and could lead to interesting behavior: instead of assigning an i.i.d. non-negative marks W_u to each vertex u, one could assign i.i.d. vectors $\mathbf{W}_u := (W_u^{(1)}, \dots, W_u^{(r)})$ whose entries are dependent. The kernel $\kappa(\mathbf{W_u}, \mathbf{W_v})$, described near (1.3.4) for the traditional KSRG, should still be symmetric and map to $\mathbb{R}_+$. The simplest kernel to work with, corresponding to κ_{prod} in (1.3.4), is to take the inner product of the weights. The obtained graph with this inner-product kernel could be viewed as a superposition of *dependent* KSRGs. Thus, by choosing the distribution of $\mathbf{W}$ such that the entries are negatively (respectively positively) correlated, one could model disassortative (respectively assortative) communities, which could lead to graphs where the giant component is non-unique. For non-spatial random graphs, this model is known as the random dot product graph [227], which has gained significant attention in the field of statistical inference (see [13] for an overview).

BIBLIOGRAPHY

[1] M. Abueg et al. Modeling the effect of exposure notification and non-pharmaceutical interventions on COVID-19 transmission in Washington state. *npj Digital Medicine*, 4(1), 2021.

[2] E. Adriaans and J. Komjáthy. Weighted distances in scale-free configuration models. *Journal of Statistical Physics*, 173(3):1082–1109, 2018.

[3] W. Aiello, A. Bonato, C. Cooper, J. Janssen, and P. Prałat. A spatial web graph model with local influence regions. *Internet Mathematics*, 5(1-2):175–196, 2008.

[4] M. Aizenman, H. Kesten, and C. M. Newman. Uniqueness of the infinite cluster and continuity of connectivity functions for short and long range percolation. *Communications in Mathematical Physics*, 111(4):505 –531, 1987.

[5] M. Aizenman, F. Delyon, and B. Souillard. Lower bounds on the cluster size distribution. *Journal of Statistical Physics*, 23(3):267–280, 1980.

[6] M. Aizenman and C. M. Newman. Tree graph inequalities and critical behavior in percolation models. *Journal of Statistical Physics*, 36(1):107–143, 1984.

[7] R Albert and A. Barabási. Statistical mechanics of complex networks. *Reviews Of Modern Physics*, 74(1):47–97, 2002.

[8] R. Albert, H. Jeong, and A.-L. Barabási. Internet: diameter of the world-wide web. *Nature*, 401(6749):130, 1999.

[9] K Alexander, J. Chayes, and L Chayes. The Wulff construction and asymptotics of the finite cluster distribution for two-dimensional Bernoulli percolation. *Communications in Mathematical Physics*, 131(1):1–50, 1990.

[10] C. Alves, R. Ribeiro, and R. Sanchis. Preferential attachment random graphs with edge-step functions. *Journal of Theoretical Probability*:1–39, 2019.

[11] M. Amin Abdullah and N. Fountoulakis. A phase transition in the evolution of bootstrap percolation processes on preferential attachment graphs. *Random Structures & Algorithms*, 52(3):379–418, 2018.

[12] O. Amini, L. Devroye, S. Griffiths, and N. Olver. On explosions in heavy-tailed branching random walks. *The Annals of Probability*, 41(3B):1864–1899, 2013.

[13] A. Athreya, D. E. Fishkind, M. Tang, C. E. Priebe, Y. Park, J. T. Vogelstein, K. Levin, V. Lyzinski, and Y. Qin. Statistical inference on random dot product graphs: a survey. *The Journal of Machine Learning Research*, 18(1):8393–8484, 2017.

[14] K. Athreya and P. Ney. *Branching Processes*. Dover Books on Mathematics. Dover Publications, 2004.

[15] L. Backstrom, P. Boldi, M. Rosa, J. Ugander, and S. Vigna. Four degrees of separation. *Proceedings of the 4th Annual ACM Web Science Conference*, pages 33–42. ACM, 2012.

[16] P. Bajardi, A. Barrat, F. Natale, L. Savini, and V. Colizza. Dynamical patterns of cattle trade movements. *PLoS One*, 6(5):e19869, 2011.

[17] F. Ball and D. Sirl. An sir epidemic model on a population with random network and household structure, and several types of individuals. *Advances in Applied Probability*, 44(1):63–86, 2012.

[18] F. Ball, D. Sirl, and P. Trapman. Analysis of a stochastic sir epidemic on a random network incorporating household structure. *Mathematical Biosciences*, 224(2):53–73, 2010.

[19] F. Ball, D. Sirl, and P. Trapman. Epidemics on random intersection graphs. *The Annals of Applied Probability*, 24(3):1081 –1128, 2014.

[20] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[21] E. Baroni, R. v. d. Hofstad, and J. Komjáthy. Tight fluctuations of weight-distances in random graphs with infinite-variance degrees. *Journal of Statistical Physics*, 174:906–934, 2019.

[22] E. Baroni, R. van der Hofstad, and J. Komjáthy. Nonuniversality of weighted random graphs with infinite variance degree. *Journal of Applied Probability*, 54(1):146–164, 2017.

[23] N. Berger, C. Borgs, J. T. Chayes, and A. Saberi. On the spread of viruses on the internet. *Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 301–310. Society for Industrial and Applied Mathematics, 2005.

[24] N. Berger, C. Borgs, J. T. Chayes, and A. Saberi. Asymptotic behavior and distributional limits of preferential attachment graphs. *The Annals of Probability*, 42(1):1–40, 2014.

[25] S. Bhamidi, S. Dhara, and R. van der Hofstad. Multiscale genesis of a tiny giant for percolation on scale-free random graphs. *Preprint arXiv:2107.04103*, 2021.

[26] S. Bhamidi, R. v. d. Hofstad, and G. Hooghiemstra. First passage percolation on random graphs with finite mean degrees. *The Annals of Applied Probability*, 20(5):1907–1965, 2010.

[27] S. Bhamidi, R. v. d. Hofstad, and G. Hooghiemstra. First passage percolation on the Erdős-Rényi random graph. *Combinatorics, Probability & Computing*, 20:683–707, 2011.

[28] S. Bhamidi, R. v. d. Hofstad, and G. Hooghiemstra. Universality for first passage percolation on sparse random graphs. *The Annals of Probability*, 45(4):2568–2630, 2017.

[29] S. Bhamidi, R. v. d. Hofstad, and J. Komjáthy. The front of the epidemic spread and first passage percolation. *Journal of Applied Probability*, 51(A):101–121, 2014.

[30] M. Biskup. On the scaling of the chemical distance in long-range percolation models. *The Annals of Probability*, 32(4):2938–2977, 2004.

[31] T. Bläsius, T. Friedrich, M. Katzmann, U. Meyer, M. Penschuck, and C. Weyand. Efficiently generating geometric inhomogeneous and hyperbolic random graphs. *Proceedings of the 27th Annual European Symposium on Algorithms (ESA 2019)*, volume 144 of *Leibniz International Proceedings in Informatics (LIPIcs)*, 21:1–21:14, Dagstuhl, Germany, 2019.

[32] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang. Complex networks: structure and dynamics. *Physics Reports*, 424(4-5):175–308, 2006.

[33] B. Bollobás, O. Riordan, J. Spencer, and G. Tusnády. The degree sequence of a scale-free random graph process, *The Structure and Dynamics of Networks*, pages 385–395. Princeton University Press, 2011.

[34] B. Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combin.*, 1(4):311–316, 1980.

[35] B. Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combin.*, 1(4):311–316, 1980.

[36] B. Bollobás. *Random graphs*, volume 73 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2001, pages xviii+498.

[37] B. Bollobás, S. Janson, and O. Riordan. The phase transition in inhomogeneous random graphs. *Random Structures & Algorithms*, 31(1):3–122, 2007.

[38] B. Bollobás and O. Riordan. The diameter of a scale-free random graph. *Combinatorica*, 24(1):5–34, 2004.

[39] K. Bringmann, R. Keusch, and J. Lengler. Average distance in a general class of scale-free networks with underlying geometry. arXiv:1602.05712, 2016.

[40] K. Bringmann, R. Keusch, and J. Lengler. Geometric inhomogeneous random graphs. *Theoretical Computer Science*, 760:35–54, 2019.

[41] T. Britton, S. Janson, and A. Martin-Löf. Graphs with specified degree distributions, simple epidemics, and local vaccination strategies. *Advances in Applied Probability*, 39(4):922–948, 2007.

[42] S. R. Broadbent and J. M. Hammersley. Percolation processes: I. Crystals and mazes. *Mathematical proceedings of the Cambridge philosophical society*, 53(3):629–641, 1957.

[43] V. H. Can. Metastability for the contact process on the preferential attachment graph. *Preprint arXiv:1502.05633*, 2015.

[44] E. Candellero and N. Fountoulakis. Bootstrap percolation and the geometry of complex networks. *Stochastic Processes and their Applications*, 126(1):234–264, 2016.

[45] F. Caravenna, A. Garavaglia, and R. v. d. Hofstad. Diameter in ultra-small scale-free random graphs: extended version. *Preprint arXiv:1605.02714*, 2016.

[46] F. Caravenna, A. Garavaglia, and R. v. d. Hofstad. Diameter in ultra-small scale-free random graphs. *Random Structures & Algorithms*, 54(3):444–498, 2019.

[47] R. Cerf. *Large deviations for three dimensional supercritical percolation*. Société mathématique de France, 2000.

[48] S. Chatterjee and R. Durrett. Contact processes on random graphs with power law degree distributions have critical value 0. *The Annals of Probability*, 37(6):2332–2356, 2009.

[49] F. Chung and L. Lu. Connected components in random graphs with given expected degree sequences. *Annals of Combinatorics*, 6(2):125–145, 2002.

[50] F. Chung and L. Lu. The average distances in random graphs with given expected degrees. *Proceedings of the National Academy of Sciences*, 99(25):15879–15882, 2002.

[51] A. Cipriani and A. Fontanari. Dynamical fitness models: evidence of universality classes for preferential attachment graphs. *Journal of Applied Probability*, 59(3):609–630, 2022.

[52] A. Cipriani and M. Salvi. Scale-free percolation mixing time. *Preprint arXiv:2111.05201*, 2021.

[53] A. Clark, M. Jit, C. Warren-Gash, B. Guthrie, H. H. Wang, S. W. Mercer, C. Sanderson, M. McKee, C. Troeger, K. L. Ong, et al. Global, regional, and national estimates of the population at increased risk of severe covid-19 due to underlying health conditions in 2020: a modelling study. *The Lancet Global Health*, 8(8):e1003–e1017, 2020.

[54] R. Cohen, S. Havlin, and D. Ben-Avraham. Efficient immunization strategies for computer networks and populations. *Physical Review Letters*, 91(24):247901, 2003.

[55] D. Collaris, P. Gajane, J. Jorritsma, J. J. van Wijk, and M. Pechenizkiy. LEMON: Alternative Sampling for More Faithful Explanation through Local Surrogate Models. *Proceedings of the International Symposium on Data Analysis*, 2023.

[56] D. Contreras, S. Martineau, and V. Tassion. Supercritical percolation on graphs of polynomial growth. *Preprint arXiv:2107.06326*, 2021.

[57] C. Cooper, A. Frieze, and J. Vera. Random deletion in a scale-free random graph process. *Internet Mathematics*, 1(4):463–483, 2004.

[58] N. Crawford and A. Sly. Simple random walk on long range percolation clusters I: heat kernel bounds. *Probability Theory and Related Fields*, 154(3-4):753–786, 2012.

[59] N. Crawford and A. Sly. Simple random walk on long-range percolation clusters II: scaling limits. *The Annals of Probability*, 41(2):445–502, 2013.

[60] J. Dalmau and M. Salvi. Scale-free percolation in continuous space: quenched degree and clustering coefficient. *Journal of Applied Probability*, 58(1):106–127, 2021.

[61] S. Davis, B. Abbasi, S. Shah, S. Telfer, and M. Begon. Spatial analyses of wildlife contact networks. *Journal of The Royal Society Interface*, 12(102):20141004, 2015.

[62] D. R. de Souza and T. Tomé. Stochastic lattice gas model describing the dynamics of the SIRS epidemic process. *Physica A: Statistical Mechanics and its Applications*, 389(5):1142–1150, 2010.

[63] M. Deijfen, R. van der Hofstad, and G. Hooghiemstra. Scale-free percolation. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 49(3):817–838, 2013.

[64] M. Deijfen, R. v. d. Hofstad, and G. Hooghiemstra. Scale-free percolation. *Annales de l'institut Henri Poincare (B) Probability and Statistics*, 49(3):817–838, 2013.

[65] M. Deijfen and M. Lindholm. Growing networks with preferential deletion and addition of edges. *Physica A: Statistical Mechanics and its Applications*, 388(19):4297–4303, 2009.

[66] P. Deprez, R. S. Hazra, and M. V. Wüthrich. Inhomogeneous long-range percolation for real-life network modeling. *Risks*, 3(1):1–23, 2015.

[67] P. Deprez and M. V. Wüthrich. Scale-free percolation in continuum space. *Communications in Mathematics and Statistics*, 2018.

[68] S. Dereich, C. Mailler, and P. Mörters. Nonextensive condensation in reinforced branching processes. *The Annals of Applied Probability*, 27(4):2539–2568, 2017.

[69] S. Dereich, C. Mönch, and P. Mörters. Typical distances in ultrasmall random networks. *Advances in Applied Probability*, 44(2):583–601, 2012.

[70] S. Dereich and P. Mörters. Random networks with sublinear preferential attachment: degree evolutions. *Electronic Journal of Probability*, 14:1222–1267, 2009.

[71] S. Dereich and P. Mörters. Random networks with sublinear preferential attachment: the giant component. *The Annals of Probability*, 41(1):329–384, 2013.

[72] S. Dereich, C. Mönch, and P. Mörters. Distances in scale free networks at criticality. *Electronic Journal of Probability*, 22:38 pp. 2017.

[73] S. Dereich and M. Ortgiese. Robust analysis of preferential attachment models with fitness. *Combininatorics, Probability and Computing*, 23(3):386–411, 2014.

[74] S. Dereich and M. Ortgiese. Local neighbourhoods for first-passage percolation on the configuration model. *Journal of Statistical Physics*:1–17, 2018.

[75] A. Desvars-Larrive, E. Dervic, N. Haug, T. Niederkrotenthaler, J. Chen, A. Di Natale, J. Lasser, D. S. Gliga, A. Roux, J. Sorger, et al. A structured open dataset of government interventions in response to covid-19. *Scientific data*, 7(1):285, 2020.

[76] J.-D. Deuschel and A. Pisztora. Surface order large deviations for high-density percolation. *Probability Theory and Related Fields*, 104(4):467–482, 1996.

[77] S. Dhara, R. Van Der Hofstad, and J. S. Van Leeuwaarden. Critical percolation on scale-free random graphs: new universality class for the configuration model. *Communications in Mathematical Physics*, 382(1):123–171, 2021.

[78] S. Dhara and R. van der Hofstad. Barely supercritical percolation on poissonian scale-free networks. *Preprint arXiv:2107.04106*, 2021.

[79] M. Dickson and M. Heydenreich. The triangle condition for the marked random connection model. *Preprint arXiv:2210.07727*, 2022.

[80] O. Diekmann, J. A. P. Heesterbeek, and J. A. J. Metz. On the definition and the computation of the basic reproduction ratio Ro in models for infectious diseases in heterogeneous populations. *Journal of Mathematical Biology*, 28:365–382, 1990.

[81] S. Dommers, R. Hofstad, and G. Hooghiemstra. Diameters in preferential attachment models. *Journal of Statistical Physics*, 139(1):72–107, 2010.

[82] S. N. Dorogovtsev and J. F. F. Mendes. Evolution of networks. *Advances in Physics*, 51(4):1079–1187, 2002.

[83] S. Dorogovtsev and J. Mendes. *Evolution of networks: From biological nets to the Internet and WWW*. Oxford University Press, 2003.

[84] S. N. Dorogovtsev, A. V. Goltsev, and J. F. Mendes. Critical phenomena in complex networks. *Reviews of Modern Physics*, 80(4):1275, 2008.

[85] C. Drent. *Structured Learning and Decision Making for Maintenance*. PhD thesis, Eindhoven University of Technology, 2022.

[86] H. Duminil-Copin, C. Garban, and V. Tassion. long-range models in 1d revisited.

[87] K. T. Eames and M. J. Keeling. Contact tracing and disease control. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270(1533):2565–2571, 2003.

[88] M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the internet topology. *SIGCOMM Comput. Commun. Rev.*, 29(4):251–262, 1999.

[89] L. Federico and R. Van der Hofstad. Critical window for connectivity in the configuration model. *Combinatorics, Probability and Computing*, 26(5):660–680, 2017.

[90] S. L. Feld. Why your friends have more friends than you do. *American Journal of Sociology*, 96(6):1464–1477, 1991.

[91] N. Fountoulakis, P. Van Der Hoorn, T. Müller, and M. Schepers. Clustering in a hyperbolic model of complex networks. *Electronic Journal of Probability*, 26:1–132, 2021.

[92] N. Fountoulakis and J. Yukich. Limit theory for isolated and extreme points in hyperbolic random geometric graphs. *Electronic Journal of Probability*, 25(none):1 –51, 2020.

[93] C. Fransson and P. Trapman. SIR epidemics and vaccination on random graphs with clustering. *Journal of Mathematical Biology*, 78:2369–2398, 2019.

[94] N. Freeman, J. Jordan, et al. Extensive condensation in a model of preferential attachment with fitness. *Electronic Journal of Probability*, 25, 2020.

[95] M. Galanti and J. Shaman. Direct observation of repeated infections with endemic coronaviruses. *The Journal of Infectious Diseases*, 223(3):409–415, 2020.

[96] A. Gandolfi, M. S. Keane, and C. M. Newman. Uniqueness of the infinite component in a random graph with applications to percolation and spin glasses. *Probability Theory and Related Fields*, 92(4):511–527, 1992.

[97] A Garavaglia. *Preferential attachment models for dynamic networks*. PhD thesis, Eindhoven University of Technology, 2019.

[98] A. Garavaglia, R. S. Hazra, R. van der Hofstad, and R. Ray. Universality of the local limit of preferential attachment models. *Preprint arXiv:2212.05551*, 2022.

[99] A. Garavaglia and C. Stegehuis. Subgraphs in preferential attachment models. *Advances in applied probability*, 51(3):898–926, 2019.

[100] L. A. Goldberg, J. Jorritsma, J. Komjáthy, and J. Lapinskas. Increasing efficacy of contact-tracing applications by user referrals and stricter quarantining. *PLoS ONE*, 16(5):1–36, May 2021.

[101] S. Gómez, A. Arenas, J. Borge-Holthoefer, S. Meloni, and Y. Moreno. Discrete-time markov chain approach to contact-based disease spreading in complex networks. *EPL (Europhysics Letters)*, 89(3):38009, 2010.

[102] M. C. González, C. A. Hidalgo, and A.-L. Barabási. Understanding individual human mobility patterns. *Nature*, 453:779 –782, 2008.

[103] J.-B. Gouéré et al. Subcritical regimes in the Poisson Boolean model of continuum percolation. *The Annals of Probability*, 36(4):1209–1220, 2008.

[104] P. Gracar and A. Grauer. The contact process on scale-free geometric random graphs. *Preprint arXiv:2208.08346*, 2022.

[105] P. Gracar, A. Grauer, L. Lüchtrath, and P. Mörters. The age-dependent random connection model. *Queueing Systems*, 93(3):309–331, 2019.

[106] P. Gracar, A. Grauer, and P. Mörters. Chemical distance in geometric random graphs with long edges and scale-free degree distribution. *Communications in Mathematical Physics*, 395(2):859–906, 2022.

[107] P. Gracar, M. Heydenreich, C. Mönch, and P. Mörters. Recurrence versus transience for weight-dependent random connection models. *Electronic Journal of Probability*, 27:1–31, 2022.

[108] P. Gracar, L. Lüchtrath, and C. Mönch. Finiteness of the percolation threshold for inhomogeneous long-range models in one dimension. *Preprint arXiv:2203.11966*, 2022.

[109] P. Gracar, L. Lüchtrath, and P. Mörters. Percolation phase transition in weight-dependent random connection models. *Advances in Applied Probability*, 53(4):1090–1114, 2021.

[110] G. R. Grimmett and J. M. Marstrand. The supercritical phase of percolation is well behaved. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, 430(1879):439–457, 1990.

[111] A. Grover and J. Leskovec. Node2vec: scalable feature learning for networks. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, 855–864. Association for Computing Machinery, 2016.

[112] D. Gruhl, R. Guha, D. Liben-Nowell, and A. Tomkins. Information diffusion through blogspace. *Proceedings of the 13th international conference on World Wide Web*, pages 491–501. ACM, 2004.

[113] J. M. Hammersley and D. J. A. Welsh. First-passage percolation, subadditive processes, stochastic networks, and generalized renewal theory, *Proc. Internat. Res. Semin., Statist. Lab., Univ. California, Berkeley, Calif*, pages 61–110. Springer-Verlag, New York, 1965.

[114] J. M. Hammersley and D. J. Welsh. First-passage percolation, subadditive processes, stochastic networks, and generalized renewal theory, *Bernoulli 1713, Bayes 1763, Laplace 1813*, pages 61–110. Springer, 1965.

[115] H. Han. Estimate the incubation period of coronavirus 2019 (covid-19). *medRxiv*, 2020.

[116] N. Hao and M. Heydenreich. Graph distances in scale-free percolation: the logarithmic case. *Journal of Applied Probability*, 60(1):295–313, 2023.

[117] T. E. Harris. *The theory of branching processes*. Courier Corporation, 2002.

[118] J. He, Y. Guo, R. Mao, and J. Zhang. Proportion of asymptomatic coronavirus disease 2019: a systematic review and meta-analysis. *Journal of medical virology*, 2020.

[119] A. M. Henao-Restrepo, I. M. Longini, M. Egger, N. E. Dean, W. J. Edmunds, A. Camacho, M. W. Carroll, M. Doumbia, B. Draguez, S. Duraffour, et al. Efficacy and effectiveness of an rvsv-vectored vaccine expressing ebola surface glycoprotein: interim results from the guinea ring vaccination cluster-randomised trial. *The Lancet*, 386(9996):857–866, 2015.

[120] H. W. Hethcote. The mathematics of infectious diseases. *SIAM Review*, 42(4):599–653, 2000.

[121] M. Heydenreich, T. Hulshof, and J. Jorritsma. Structures in supercritical scale-free percolation. *The Annals of Applied Probability*, 27(4):2569–2604, 2017.

[122] C. Hirsch. From heavy-tailed Boolean models to scale-free Gilbert graphs. *Brazilian Journal of Probability and Statistics*, 31(1):111 –143, 2017.

[123] C. Hirsch and C. Mönch. Distances and large deviations in the spatial preferential attachment model. *Bernoulli*, 26(2):927–947, 2020.

[124] R. v. d. Hofstad. *Random graphs and complex networks. Vol. 1*. Cambridge University Press, Cambridge, 2017.

[125] R. v. d. Hofstad. Stochastic processes on random graphs, 2017. `https://www.win.tue.nl/~rhofstad/SaintFlour_SPoRG.pdf`.

[126] R. v. d. Hofstad. Random graphs and complex networks, Volume 2, 2020+. `https://www.win.tue.nl/~rhofstad/NotesRGCNII.pdf`.

[127] R. v. d. Hofstad. *Random graphs and complex networks. Vol. 2*. 2021+. Available at `http://www.win.tue.nl/~rhofstad/NotesRGCNII.pdf`.

[128] R. v. d. Hofstad. The giant in random graphs is almost local. *arXiv:2103.11733*, 2021.

[129] R. v. d. Hofstad, G. Hooghiemstra, and D. Znamenski. Distances in random graphs with finite mean and infinite variance degrees. *Electronic Journal Of Probability*, 12:703–766, 2007.

[130] R. v. d. Hofstad and G. Hooghiemstra. Universality for distances in power-law random graphs. *Journal of Mathematical Physics*, 49(12), 2008.

[131] R. v. d. Hofstad, G. Hooghiemstra, and P. Van Mieghem. Distances in random graphs with finite variance degrees. *Random Structures & Algorithms*, 27(1):76–123, 2005.

[132] R. v. d. Hofstad, G. Hooghiemstra, and D. Znamenski. Distances in random graphs with finite mean and infinite variance degrees. *Electronic Journal of Probability*, 12(25):703–766, 2007.

[133] R. v. d. Hofstad, P. v. d. Hoorn, C. Kerriou, N. Maitra, and P. Mörters. Condensation in scale-free geometric graphs with excess edges. In preparation, 2023+.

[134] R. v. d. Hofstad, P. v. d. Hoorn, and N. Maitra. Local limits of spatial inhomogeneous random graphs. *Preprint arXiv:2107.08733*, 2021.

[135] R. v. d. Hofstad, P. v. d. Hoorn, and N. Maitra. Scaling of the clustering function in spatial inhomogeneous random graphs. *Preprint arXiv:2212.12885*, 2022.

[136] R. v. d. Hofstad and J. Komjáthy. Explosion and distances in scale-free percolation. *Preprint arXiv:1706.02597*, 2017.

[137] P. Holme and N. Litvak. Cost-efficient vaccination protocols for network epidemiology. *PLoS Computational Biology*, 2017.

[138] T. Hutchcroft. Power-law bounds for critical long-range percolation below the upper-critical dimension. *Probability Theory and Related Fields*, 181(1):533–570, 2021.

[139] T. Hutchcroft. Transience and anchored isoperimetric dimension of supercritical percolation clusters. *Preprint arXiv:2207.05226*, 2022.

[140] J. Illenberger, K. Nagel, and G. Flötteröd. The role of spatial interaction in social networks. *Networks and Spatial Economics*:255–282, 2013.

[141] V. Isham, J. Kaczmarska, and M. Nekovee. Spread of information and infection on finite random networks. *Physical Review E*, 83(4):046128, 2011.

[142] E. Jacob and P. Mörters. Spatial preferential attachment networks: power laws and clustering coefficients. *The Annals of Applied Probability*, 25(2):632–662, 2015.

[143] E. Jacob and P. Mörters. Robustness of scale-free spatial networks. *The Annals of Probability*, 45(3):1680–1722, 2017.

[144] B. Jahnel and L. Lüchtrath. Existence of s2302.05396ubcritical percolation phases for generalised weight-dependent random connection models. *Preprint arXiv:2302.05396*, 2023.

[145] S. Janson, T. Łuczak, and A. Ruciński. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000, pages xii+333.

[146] S. Janson. On percolation in random graphs with given vertex degrees. *Electronic Journal of Probability*, 14:86–118, 2009.

[147] S. Janson and L. Warnke. Preferential attachment without vertex growth: emergence of the giant component. *The Annals of Applied Probability*, 31(4):1523–1547, 2021.

[148] J. Janssen and A. Mehrabian. Rumors spread slowly in a small-world spatial network. *SIAM Journal on Discrete Mathematics*, 31(4):2414–2428, 2017.

[149] J. Jorritsma, J. Komjáthy, and D. Mitsche. Cluster-size decay in supercritical kernel-based spatial random graphs. *Preprint arXiv:2303.00724*, 2023.

[150] J. Jorritsma, J. Komjáthy, and D. Mitsche. Cluster-size decay in supercritical long-range percolation. *Preprint arXiv:2303.00712*, 2023.

[151] J. Jorritsma, J. Komjáthy, and D. Mitsche. Speed of large deviations of the giant component in supercritical kernel-based spatial random graphs. In preparation, 2023+.

[152] J. Jorritsma, T. Hulshof, and J. Komjáthy. Not all interventions are equal for the height of the second peak. *Chaos, Solitons & Fractals*, 139:109965, 2020.

[153] J. Jorritsma and J. Komjáthy. Distance evolutions in growing preferential attachment graphs. *The Annals of Applied Probability*, 32(6):4356 –4397, 2022.

[154] J. Jorritsma and J. Komjáthy. Weighted distances in scale-free preferential attachment models. *Random Structures & Algorithms*, 57(3):823–859, 2020.

[155] J. Jorritsma and J. Lapinskas. Software for "increasing efficacy of contact-tracing applications by user referrals and stricter quarantining", version v1.0.0, Apr. 2021.

[156] J. Jorritsma, J. Lengler, and D. Sudholt. Comma selection outperforms plus selection on onemax with randomly planted optima. *Submitted*, 2023+.

[157] M. Karsai, M. Kivelä, R. K. Pan, K. Kaski, J. Kertész, A.-L. Barabási, and J. Saramäki. Small but slow world: how network topology and burstiness slow down spreading. *Physical Review E*, 83(2):025102, 2011.

[158] H. Kesten and Y. Zhang. The probability of a large finite cluster in supercritical Bernoulli percolation. *The Annals of Probability*:537–555, 1990.

[159] R. D. Kirkcaldy, B. A. King, and J. T. Brooks. COVID-19 and Postinfection Immunity: Limited Evidence, Many Remaining Questions. *JAMA*, 2020.

[160] I. Z. Kiss, D. M. Green, and R. R. Kao. Infectious disease control using contact tracing in random and scale-free networks. *Journal of The Royal Society Interface*, 3(6):55–62, 2006.

[161] M. Kiwi and D. Mitsche. Spectral gap of random hyperbolic graphs and related parameters. *The Annals of Applied Probability*, 28(2):941–989, 2018.

[162] M. Kiwi and D. Mitsche. On the second largest component of random hyperbolic graphs. *SIAM Journal on Discrete Mathematics*, 33(4):2200–2217, 2019.

[163] M. Kiwi, M. Schepers, and J. Sylvester. Cover and hitting times of hyperbolic random graphs. *Preprint arXiv:2207.06956*, 2022.

[164] C. Koch and J. Lengler. Bootstrap Percolation on Geometric Inhomogeneous Random Graphs. *43rd International Colloquium on Automata, Languages, and Programming (ICALP 2016)*, volume 55 of *Leibniz International Proceedings in Informatics (LIPIcs)*, 147:1–147:15, Dagstuhl, Germany, 2016.

[165] J. Komjáthy, J. Lapinskas, and J. Lengler. Penalising transmission to hubs in scale-free spatial random graphs. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 57(4):1968 –2016, 2021.

[166] J. Komjáthy and B. Lodewijks. Explosion in weighted hyperbolic random graphs and geometric inhomogeneous random graphs. *Stochastic Processes and their Applications*:1309–1367, 3, 2020.

[167] M. E. Kretzschmar, G. Rozhnova, M. C. Bootsma, M. van Boven, J. H. van de Wijgert, and M. J. Bonten. Impact of delays on effectiveness of contact tracing strategies for covid-19: a modelling study. *The Lancet Public Health*, 5(8):e452–e459, 2020.

[168] M. Kretzschmar, G. Rozhnova, and M. van Boven. Isolation and contact tracing can tip the scale to containment of covid-19 in populations with social distancing. *Available at SSRN 3562458*, 2020.

[169] D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, and M. Boguná. Hyperbolic geometry of complex networks. *Physical Review E*, 82(3):036106, 2010.

[170] T. Kuchler, D. Russel, and J. Stroebel. The geographic spread of COVID-19 correlates with the structure of social networks as measured by facebook. 2020.

[171] H. Kunz and B. Souillard. Essential singularity in percolation problems and asymptotic behavior of cluster size distribution. *Journal of Statistical Physics*, 19(1):77–106, 1978.

[172] A. Lambert. A mathematical assessment of the efficiency of quarantining and contact tracing in curbing the COVID-19 epidemic. *Mathematical Modelling of Natural Phenomena*, 16:53, 2021. E. Augeraud, M. Banerjee, J.-S. Dhersin, A. d'Onofrio, T. Lipniacki, S. Petrovskii, C. Tran, A. Veber-Delattre, E. Vergu, and V. Volpert, editors.

[173] G. Last and M. Penrose. *Lectures on the Poisson process*, volume 7. Cambridge University Press, 2017.

[174] S. A. Lauer, K. H. Grantz, Q. Bi, F. K. Jones, Q. Zheng, H. R. Meredith, A. S. Azman, N. G. Reich, and J. Lessler. The incubation period of coronavirus disease 2019 (covid-19) from publicly reported confirmed cases: estimation and application. *Annals of internal medicine*, 172(9):577–582, 2020.

[175] L. Lichev, B. Lodewijks, D. Mitsche, and B. Schapira. Bernoulli percolation on the random geometric graph. *Preprint arXiv:2205.10923*, 2022.

[176] T. M. Liggett, R. H. Schonmann, and A. M. Stacey. Domination by product measures. *The Annals of Probability*, 25(1):71–95, 1997.

[177] A. Linker, D. Mitsche, B. Schapira, and D. Valesin. The contact process on random hyperbolic graphs: metastability and critical exponents. *The Annals of Probability*, 49(3):1480–1514, 2021.

[178] L. H. Loomis and H. Whitney. An inequality related to the isoperimetric inequality. *Bulletin of the American Mathematical Society*, 55(10):961 –962, 1949.

[179] L. Lüchtrath. *Percolation in weight-dependent random connection models*. PhD thesis, Universität zu Köln, 2022. Ph. D. thesis.

[180] Y. Malyshkin and E. Paquette. The power of choice combined with preferential attachment. *Electronic Communications in Probability*, 19(44):1–13, 2014.

[181] S. L. McKenna and I. R. Dohoo. Using and interpreting diagnostic tests. *Veterinary Clinics of North America: Food Animal Practice*, 22(1):195–205, Mar. 2006.

[182] R. Meester and R. Roy. *Continuum percolation*, volume 119 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1996, pages x+238.

[183] C. Menni, A. M. Valdes, M. B. Freidin, C. H. Sudre, L. H. Nguyen, D. A. Drew, S. Ganesh, T. Varsavsky, M. J. Cardoso, J. S. E.-S. Moustafa, et al. Real-time tracking of self-reported symptoms to predict potential covid-19. *Nature medicine*:1–4, 2020.

[184] S. Merler, M. Ajelli, L. Fumanelli, M. F. Gomes, A. P. y Piontti, L. Rossi, D. L. Chao, I. M. Longini Jr, M. E. Halloran, and A. Vespignani. Spatiotemporal spread of the 2014 outbreak of ebola virus disease in liberia and the effectiveness of non-pharmaceutical interventions: a computational modelling analysis. *The Lancet Infectious Diseases*, 15(2):204–211, 2015.

[185] D. Mitsche and M. D. Penrose. Limit theory of combinatorial optimization for random geometric graphs. *The Annals of Applied Probability*, 31(6):2721 –2771, 2021.

[186] M. Mitzenmacher and E. Upfal. *Probability and computing: Randomized algorithms and probabilistic analysis*. Cambridge University Press, 2005.

[187] K. Mizumoto, K. Kagaya, A. Zarebski, and G. Chowell. Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the diamond princess cruise ship, yokohama, japan, 2020. *Eurosurveillance*, 25(10), 2020.

[188] M. Molloy and B. Reed. A critical point for random graphs with a given degree sequence, *The Structure and Dynamics of Networks*, pages 240–258. Princeton University Press, 2011.

[189] M. Molloy and B. Reed. The size of the giant component of a random graph with a given degree sequence. *Combinatorics, Probability and Computing*, 7(3):295–305, 1998.

[190] C. Mönch. *Distances in preferential attachment networks*. PhD thesis, University of Bath, 2013.

[191] J. M. Montoya and R. V. Solé. Small world patterns in food webs. *Journal of theoretical biology*, 214(3):405–412, 2002.

[192] Y. Moreno, R. Pastor-Satorras, and A. Vespignani. Epidemic outbreaks in complex heterogeneous networks. *The European Physical Journal B - Condensed Matter and Complex Systems*, 26:521–529, 2002.

[193] T. F. Móri. On random trees. *Studia Sci. Math. Hungar.*, 39(1-2):143–155, 2002.

[194] J. Mossong, N. Hens, M. Jit, P. Beutels, K. Auranen, R. Mikolajczyk, M. Massari, S. Salmaso, G. S. Tomba, J. Wallinga, et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med*, 5(3):e74, 2008.

[195] L. Muchnik, S. Pei, L. C. Parra, S. D. S. Reis, J. S. J. Andrade, S. Havlin, and H. A. Makse. Origins of power-law degree distribution in the heterogeneity of human activity in social networks. *Scientific Reports*, 3(1783), 2013.

[196] J. Müller and M. Kretzschmar. Contact tracing–old models and new challenges. *Infectious Disease Modelling*, 6:222–231, 2021.

[197] M. E. J. Newman. Spread of epidemic disease on networks. *Physical Review E*, 66:016128, 1, 2002.

[198] M. E. J. Newman. The structure and function of complex networks. *SIAM Review*, 45(2):167–256, 2003.

[199] M. E. J. Newman and D. J. Watts. Scaling and percolation in the small-world network model. *Physical Review E*, 60:7332–7342, 6, 1999.

[200] M. Newman, A.-L. Barabasi, and D. J. Watts. *The structure and dynamics of networks*, volume 19. Princeton University Press, 2011.

[201] M. E. Newman. The structure and function of complex networks. *SIAM review*, 45(2):167–256, 2003.

[202] M. E. Newman. *Networks*. Oxford University Press, 2018.

[203] H. Nishiura, T. Kobayashi, T. Miyama, A. Suzuki, S.-m. Jung, K. Hayashi, R. Kinoshita, Y. Yang, B. Yuan, A. R. Akhmetzhanov, et al. Estimation of the asymptomatic ratio of novel coronavirus infections (covid-19). *International journal of infectious diseases*, 94:154, 2020.

[204] I. Norros and H. Reittu. On a conditionally Poissonian graph process. *Advances in Applied Probability*, 38(1):59–75, 2006.

[205] L. Pan, M. Mu, P. Yang, Y. Sun, R. Wang, J. Yan, P. Li, B. Hu, J. Wang, C. Hu, et al. Clinical characteristics of covid-19 patients with digestive symptoms in hubei, china: a descriptive, cross-sectional, multicenter study. *The American Journal of Gastroenterology*, 115, 2020.

[206] P. K. Pandey and B. Adhikari. Why lockdown : on the spread of sars-cov-2 in india, a network approach. arXiv:2004.11973, 2020.

[207] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani. Epidemic processes in complex networks. *Reviews of modern physics*, 87(3):925, 2015.

[208] R. Pastor-Satorras and A. Vespignani. Epidemic spreading in scale-free networks. *Physical Review Letters*, 86:3200–3203, 14, 2001.

[209] R. Pastor-Satorras and A. Vespignani. *Evolution and structure of the Internet: A statistical physics approach*. Cambridge University Press, 2007.

[210] M. Penrose. *Random geometric graphs*, volume 5. Oxford University Press, 2003.

[211] M. D. Penrose. Giant component of the soft random geometric graph. *Preprint arXiv:2204.10219*, 2022.

[212] M. Penrose. Connectivity of soft random geometric graphs. *The Annals of Applied Probability*, 26(2):986–1028, 2016.

[213] B. Pittel. On a random graph evolving by degrees. *Advances in Mathematics*, 223(2):619–671, 2010.

[214] P. Poletti et al. Association of age with likelihood of developing symptoms and critical disease among close contacts exposed to patients with confirmed SARS-CoV-2 infection in italy. *JAMA Network Open*, 4(3):e211085, 2021.

[215] Y. Raaijmakers. *Job-Replication Trade-Offs: Performance Analysis of Redundancy Systems*. PhD thesis, Eindhoven University of Technology, 2021.

[216] L. S. Schulman. Long range percolation in one dimension. *Journal of Physics A: Mathematical and General*, 16(17):L639–L641, 1983.

[217] C. Stegehuis, R. van Der Hofstad, and J. S. van Leeuwaarden. Scale-free network clustering in hyperbolic and other random graphs. *Journal of Physics A: Mathematical and Theoretical*, 52(29):295101, 2019.

[218] S. D. Stovitz, H. R. Banack, and J. S. Kaufman. 'Depletion of the susceptibles' taught through a story, a table and basic arithmetic. *BMJ Evidence-Based Medicine*, 23(5):199–199, 2018.

[219] M. A. Strassburg. The global eradication of smallpox. *American Journal of Infection Control*, 10(2):53–59, 1982.

[220] O. Tange. GNU Parallel - The Command-Line Power Tool. *The USENIX Magazine*, 36(1):42–47, 2011.

[221] S. Thurner, P. Klimek, and R. Hanel. A network-based explanation of why most COVID-19 infection curves are linear. 2020.

[222] J. Travers and S. Milgram. The small world problem. *Phychology Today*, 1(1):61–67, 1967.

[223] V. Vadon, J. Komjáthy, and R. van der Hofstad. A new model for overlapping communities with arbitrary internal structure. *Applied Network Science*, 4(42), 2019.

[224] P. Van den Driessche and J. Watmough. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical Biosciences*, 180(1):29–48, 2002.

[225] C. Viboud, O. N. Bjørnstad, D. L. Smith, L. Simonsen, M. A. Miller, and B. T. Grenfell. Synchrony, waves, and spatial hierarchies in the spread of influenza. *Science*, 312(5772):447–451, 2006.

[226] H.-Y. Wang, X.-L. Li, Z.-R. Yan, X.-P. Sun, J. Han, and B.-W. Zhang. Potential neurological symptoms of covid-19. *Therapeutic Advances in Neurological Disorders*, 13:1756286420917830, 2020.

[227] S. J. Young and E. R. Scheinerman. Random dot product graph models for social networks. *International Workshop on Algorithms and Models for the Web-Graph*, pages 138–149. Springer, 2007.

[228] J. E. Yukich. Ultra-small scale-free geometric networks. *Journal of Applied Probability*, 43(3):665–677, 2016.

[229] G. U. Yule. II.—A mathematical theory of evolution, based on the conclusions of Dr. JC Willis, FR S. *Philosophical transactions of the Royal Society of London. Series B*, 213(402-410):21–87, 1925.

SUMMARY

This thesis is centered around the analysis of supercritical random graphs in which the parameters are tuned such that the degree distribution has infinite variance, similar to many real-world networks. The theoretical understanding of the models that is obtained in the first two parts of the thesis, is used in the third part to give qualitative answers for questions that arised during the COVID-19 pandemic, based on simulations.

Part I: Distances in preferential attachment models. In Chapter 2 we study first-passage percolation, a theoretical model for information or disease spreading, on three preferential attachment models: a dynamic random graph model in which vertices arrive over time, and connect to vertices that are already present in the graph, favouring connecting to already high-degree vertices. Each edge is equipped with a non-negative i.i.d. weight, representing the transmission time of a message to cross an edge. The weighted distance between two vertices is the sum of the edge weights along the least-weighted path in the graph between the two vertices. We study the weighted distance between two typical vertices, i.e., vertices chosen uniformly at random from the largest component when the graph has size t, and the number of edges on this path, the typical hopcount.

We prove that there are precisely two universality classes of weight distributions, called the explosive and conservative class, and describe the limiting behavior precisely for any weight distribution. If the edge-weight distribution is in the *explosive* class, the typical weighted distance converges in distribution to a finite random variable that we describe in terms of the explosion time of the local weak limit of the graph. If the edge-weight distribution is in the *conservative* class, we prove that the typical weighted distance tends to infinity, and we give explicit expression for the main growth term around which the fluctuations are tight under mild conditions on the weight distribution L.

In Chapter 3 we study how the weighted distance between two *fixed* typical vertices u_t, v_t (sampled uniformly from the graph at time t) *evolves* as the surrounding preferential attachment graph grows for time $t' > t$. We identify a function $Q_{t,t'}$ such that the weighted distance between u_t and v_t measured at time t', stays within a tight strip around the function $Q_{t,t'}$ for all $t' \geqslant t$ and any t sufficiently large.

Part II: Cluster-size decay in kernel-based spatial random graphs. In Part ii we study the distribution of the sizes of *finite connected components* in a

framework called *kernel-based spatial random graph models* (KSRGs). In these models, the vertices are embedded in a d-dimensional Euclidean space (nearby vertices are more likely to be connected), and equipped with a positive mark (high-mark vertices are more likely to form connections). The parameters are tuned such that the model is supercritical, that is, the graph contains an infinite connected component almost surely. We show for a wide range of KSRGs that the probability that a typical vertex is in a finite component of size at least k, decays stretched exponentially with k, i.e., $\exp(-\Theta(k^{\zeta}))$ for some $\zeta \in (0,1)$. We explicitly identify a formula for ζ which undergoes several phase transitions in terms of the model parameters: it depends amongst others on the amount of degree-inhomogeneity and long-range connections in the graph.

In Chapter 5, we are specifically concerned with KSRGs that are similar to continuum scale-free percolation and/or hyperbolic random graphs, but also give intuition for the cluster-size decay for other models, such as the age-dependent random connection model, that we study in more detail in an accompanying paper that is not included in the thesis [151].

Our proofs always work if the structural inhomogeneity is "sufficiently large" compared to the spatial dimension d. When there is less inhomogeneity in the graph, we expect that the cluster-size decay is similar to the graph obtained by nearest-neighbor bond percolation on $\mathbb{Z}^d$ (NNP), and that the formula ζ depends only on the dimension d: in Chapter 6, we study supercritical long-range percolation on $\mathbb{Z}^d$, where two vertices x, y are connected by an edge with probability proportional to $\|x-y\|^{-\alpha d}$, independently of other edges. A small value of α corresponds to more long-range edges. We show that when $\alpha > 1 + 1/d$, under the (technical) condition that the edge density is sufficiently high, the cluster-size decay is stretched exponential with exponent $\zeta = (d-1)/d$, similar to NNP.

Part III: Agent-based modeling: infection spreading. Motivated by COVID-19, we conduct two simulation-based studies of spreading processes on networks: we demonstrate the influence of space and the presence of structural inhomogeneity on various characteristics of spreading processes (using a special instance of a KSRG), using spreading processes that are more advanced and realistic than first-passage percolation. We compare these characteristics to more traditional models for disease spreading (such as compartmental models). We present how the KSRGs can be used to model intervention strategies. This gives intuition for qualitative implications of the intervention strategies in Chapter 7, and the usage of contact-tracing applications to mitigate infection spread in Chapter 8.

ABOUT THE AUTHOR

Joost Jorritsma was born on January 26, 1992 in Terneuzen, The Netherlands. He completed his secondary education at Reynaertcollege in Hulst in 2010, and then started his studies in industrial and applied mathematics at Eindhoven University of Technology. After obtaining a bachelor's degree in 2014, he continued his studies at the same university by pursuing a master's degree in industrial and applied mathematics, with a specialization in statistics, probability, and operations research. He spend three months at Ludwig-Maximilians-Universität München for a research visit after which he obtained his master's degree in 2016.

Joost started his PhD project at Eindhoven university of Technology in September 2018 under the supervision of Júlia Komjáthy. His research focuses on structures in large networks. The results of this research are presented in this dissertation and provided the basis for several scientific publications. During his PhD he spent three months at ETH Zürich, and one month at Pontificia Universidad Católica de Chile.