Percolation in weight-dependent random connection models

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Lukas Lüchtrath

geboren in Freiburg im Breisgau

Köln, 2022

Berichterstatter:

Prof. Dr. Peter Mörters, Universität zu Köln
Prof. Dr. Júlia Komjáthy, Delft University of Technology

Tag der mündlichen Prüfung: 31.10.2022

Abstract

We study a general class of inhomogeneous spatial random graphs, the weight-dependent random connection model. Vertices are given through a standard Poisson point process in Euclidean space and each vertex carries additionally an i.i.d weight. Edges are drawn in such a way that short edges and edges to large weight vertices are preferred. This allows in particular the study of models that combine long-range interactions and heavy-tailed degree distributions. The occurrence of long edges together with the hierarchy of the vertices coming from the weights typically leads to very well connected graphs. We identify a sharp phase transition where the existence of a subcritical percolation phase becomes possible. This transition depends on both, the power-law of the degree distribution and on the geometric model parameter, showing the significant effect of clustering on the graph's topology. We further study the specifics of dimension one in parameter regimes where a subcritical phase exists. Natural examples that are contained in our framework are for instance the random connection model, the Poisson Boolean model, scale-free percolation and the age-dependent random connection model. We use our results to characterize robustness of age-based spatial preferential attachment networks.

Acknowledgement

My greatest thanks go to the best research group one can imagine here in Cologne. It was a great pleasure to work with you. First to mention here of course is Prof. Peter Mörters who is our supervisor and brought together this wonderful group. Dear Peter, many many thanks for all your guidance and advices over the last years, the many discussions from which I learned a lot and all the fun and interesting lectures we organised together. Also many thanks for all the trust you had in me and letting me try my own things. Your door was always open for me and all the work happened in a fruitful and encouraging environment. I couldn't imagine a better place to write this thesis.

The next people from our group I want to thank are Arne Grauer and Peter Gracar. Dear Arne and Pete, thank you for discovering the path of doing math in spatial random graphs together with me, all the discussions, tips and ways to solutions you had for me, the nice papers we wrote and all the great times we spent together. Special thanks to Pete for all the memes and Cappuccini I received from him. I hope that at least the memes keep coming.

The group is completed by Céline Kerriou, Marilyn Düsterbeck and Vera Sausen. Dear Céline, Marilyn and Vera, thank you for helping me with organising our great summer school and that you had always an open door for me when I needed someone to chat and distract from work.

I would also like to thank the whole probability group in Cologne and colleagues I had there, Alexis Prevóst, Lars Schmitz, Leonie Brinker, Katharina Bata, Gioele Gallo, Heidi Anderka and especially Prof. Alexander Drewitz for being my personal mentor. It was always a pleasure to come to the institute and meeting you.

Further, I like to thank Júlia Komjáthy who was the second examiner and referee of my thesis. I couldn't imagine a referee more interested in my results and every math detail. Dear Júlia, many thank you for all the effort and work you put into this.

Last but not least there are many other people who supported and encouraged me since I have started studying math. First of all my family; but also my many friends, Richard, Hans, Chris, Philipp, Laura, Adriana, Matthias, Julia, Regina, Robin, Lydia, Timo, Rainer, Jaina and all the other people who made Köln my real home town. Thank you.

Contents

Abstract **1**

Acknowledgement **3**

1. Introduction and summary of results **7**

1.1. Preferential attachment 11

1.1.1. A preferential attachment model and results 13

1.2. Inhomogeneous percolation 19

1.2.1. Transition to the age-dependent random connection model 20

1.2.2. A general inhomogeneous percolation model 23

1.2.3. Main percolation results 31

1.3. Structure of the thesis 40

2. The weight-dependent random connection model **43**

2.1. A formal construction 44

2.1.1. Percolation notation 46

2.2. The interpolation-kernel 47

2.2.1. Neighbourhoods and degree distribution 48

2.2.2. The percolation picture 53

2.3. FKG- and BK-inequality 55

2.4. Existence of a subcritical phase 57

2.4.1. Existence of a subcritical phase: Case $\gamma \geq 1/2$. . . . 59

2.5. Absence of a subcritical phase 75

2.6. Existence of a supercritical phase in dimension one 78

2.6.1. Proof of Proposition 2.13 81

2.7. Non existence of a supercritical phase in dimension one . . 94

2.7.1. Proof of Proposition 2.19 95

2.7.2. Proof of Corollary 1.13 99

3. The age-based spatial preferential attachment **101**
3.1. Construction and weak local limit 101
3.2. Degree distribution, clustering and edge length distribution 106
3.2.1. Degree distribution 106
3.2.2. Local and global clustering 108
3.2.3. Asymptotics for typical edge lengths 113
3.3. Robustness vs. non-robustness 116

Bibliography **I**

A. Related work and open problems **VII**

B. Integration results **IX**

C. Frequently used notation **XIII**

Declaration for the doctoral thesis **XV**

CHAPTER 1

Introduction and summary of results

Complex real-world systems can be seen as a collection of numerous objects interacting with each other in specific ways. This is applicable in many different contexts and fields such as biology (e.g. communicating neurons), physics (e.g. interacting (quantum-) particles), telecommunication (e.g. cell phone users and cell towers), social sciences (e.g. social networks), information technology (e.g. web pages that are linked by hyperlinks) and many more. Put differently, many complex systems can be seen as a network where the objects are described by the network's nodes and a link between two nodes in the network indicates the interaction of the corresponding objects. In many of the previous examples further dynamics can be considered on top of the system, such as the spread of (fake) information through a social network or the spread of a virus through society as present day examples. These events can then be seen as a dynamic process happening on a complex network. For these reasons, complex networks have become a key tool used to describe real-world systems and related problems over the last 20 years. Despite the large amount of uncertainty and complexity arising from their dynamical nature, it is of great importance to understand the structure of the underlying network when analysing a complex system. Do global phenomena arise in the system and if so, can they be explained by the way the network is built? These are typical questions in the science community but which are also of public interest as their answers may affect decisions made by political or economic leaders.

With growing computing power more and more real-world networks were able to be analysed empirically in the recent years. Most interestingly, many of them, despite different contexts often ended up having similar structural properties, see e.g. [15, 23]. That implies that there exists some universal structure for many of the above described systems. The question of interest here is, why is this the case? Can it be explained in a comparable way by basic building mechanisms which are shared by the different networks.

The study of complex networks has attracted a lot of attention from the mathematical community in recent years as this area contains important and interesting mathematical problems starting with providing suitable models replicating the observed behaviour. The probabilistic approach is to build a network as a growing sequence of random graphs and to prove the emerging features as limit theorems when the number of vertices tends to infinity. An important aim of this thesis is to present a suitable network model, the *age-based spatial preferential attachment model*. For this model, a limit structure can not only be proven to exist but also constructed explicitly. We shall see that the limit model is contained in a large class of inhomogeneous spatial random graphs, the *weight-dependent random connection model*. Here, a random graph is constructed on a countably infinite vertex set and edges are drawn dependent on the spatial distances of vertices and additional introduced vertex weights. Such models provide a different approach to model large interacting systems and their study is of independent interest. From a mathematical perspective the study of the weight-dependent random connection model contributes the most new results and proof techniques in this thesis. These results can then be used to derive results for the original network model using the limit structure.

We call the limit model *the age-dependent random connection model* and it will play a key role in this thesis as an interesting representative showing newly observed behaviour in the class of weight-dependent random connection models as well as the limit model that connects the two approaches of modelling interacting systems.

We start however with the above described network model and formulate its key features. We then explain how the limit model can be derived with the help of a proper rescaling and how to generalise the picture from there to a universal class of graphs. The features of networks we will focus on in this thesis, as formulated in [32], are

- Networks are *scale-free*: When the number of vertices tends to infinity, the asymptotic proportion of vertices with exactly k neighbours is of order $k^{-\tau}$ for some *power-law exponent* τ as $k \to \infty$. As a result, on each scale one can find vertices with a large degree compared to the majority of the other vertices; the so called *hubs* or *stars*.

- Networks show strong *clustering*: Nodes sharing a common neighbour are much more likely to be connected by an edge themselves than nodes that are picked randomly.

- Networks are *robust* under random attack: If an arbitrarily large proportion of links is randomly removed from the network, the qualitatively topological features of the network remains unchanged.

An example for the first property are celebrities in a social network which have considerably more followers than a typical user. Clustering essentially says that people with a common friend are much more likely being friends themselves then two randomly picked people. The robustness is the hardest to prove and will contribute one of the main results of this thesis. An example for it is the following: Consider a network of cell towers that are linked when their signals reach each other. A cell phone user is automatically connected to the nearest cell tower and we assume that there is always a cell tower close enough for this to happen. Then, it only depends on the connectivity of the cell tower network whether the cell phone user can reach people in every part of the world. Now imagine that each link (or signal) between two towers has a failure probability q which essentially says that a proportion of $1-q$ of the links is not working properly. Robustness then means that if q increases due to some global event such as global warming, the connectivity of the network remains qualitatively the same. In other words, the cell phone user can still reach people all over the world. We shall see that the age-based spatial preferential attachment model is a model providing all three features. We shall further see that robustness can only happen under the right interplay of the two former properties and that the question of robustness is directly linked to the existence of infinite clusters in the age-dependent random connection model.

In the following section we introduce the idea of *preferential attachment* which is an established network building mechanism in the literature. We give a brief overview over some interesting models and their evolution which will lead to our model. Before doing so, we state two more interesting features of networks for completeness, cf. [32]

- Networks are *ultra-small*: The shortest path between two randomly picked vertices in the graph is of doubly logarithmic order in the number of nodes.

- Networks are *vulnerable* under target attack: The topological features of the network changes dramatically, if a small number of highly influential nodes is strategically removed.

1.1. Preferential attachment

The idea of preferential attachment was introduced into network theory by Barabási and Albert in 1999 [4], establishing a mechanism that replicates the *rich get richer* concept. Here, at each time step a new node joins the graph and connects to already existing nodes with a probability proportional to the degree of current nodes. In this network, nodes are hence ranked by their degree and a node with a large degree can be seen as highly influantial or powerful. Having a large degree also enables a node to collect further links over time and thus to increase its degree even further. We speak therefore of *degree-based preferential attachment*. Due to this mechanism, it is reasonable to epect that there exist nodes with an exceptional large degree at later times. Indeed, this model creates a scale-free network with power-law exponent $\tau = 3$ [4]. With small adjustments on the model one can also attain all power-law exponents $\tau > 2$ without changing the construction principle [10, 22, 48]. This allows more flexibility in the study of the other properties and one can ask how these properties are affected when the power-law exponent changes. In fact, it is known that this model is robust and ultra-small if $\tau \leq 3$ and not if $\tau > 3$ [19, 20, 21, 26]. In other words, there occurs a *phase transition* when the empirical degree distribution loses its second moment. To prove these results the authors crucially rely on the fact that the exploration of the network, starting from a typical vertex, can be coupled with rather complicated, yet well-studied branching processes [20]. The coupling then yields that the network is robust and ultra-small if and only if the branching process survives. The offspring distribution of the branching process however is determined by the degree distribution and power-law exponents $\tau \leq 3$ lead to supercritical branching processes whereas $\tau > 3$ lead to subcritical ones. However, this representation of neighbourhoods immediately implies that these models are locally tree-like and cannot have clustering.

An natural idea to tackle this issue is to embed the graphs into space and give preference to short edges. We then speak of *spatial preferential* *attachment*. The idea is that the spatial location of a node represents its individual features and nodes at a short distance can be seen as similar. In 2002 Manna and Sen studied such a model [56]. Here, at every time step the new node is placed uniformly at random into space and is connected to exactly one of the older vertices where each node is chosen for connection with a probability proportional to the product of its degree and its distance to the new vertex. Another model was introduced by Flaxman et al. in 2006 [27, 28] which combines the ideas of preferential attachment and *random geometric graphs*, cf. [63]. In this model, each newly incoming node only connects to vertices within a certain distance. Amongst all possible nodes a fixed number is chosen for connection with a probability proportional to current degrees. This model was further studied by Jordan [49, 50] and Jordan and Wade [51]. An extension of this model was introduced by Aiello et al. in 2008 [1] and further studied by Cooper et al. [16] and Janssen et al. [47]. In their extension each node has a sphere of influence that grows with the node's degree. When a new node joins the graph, it connects to each node in whose sphere of influence it has been placed independently with a fixed probability p. We finally present the spatial preferential attachment model of Jacob and Mörters, introduced in 2015 [44, 45]. In their model, a growing sequence of random graphs in continuous time is built as follows: the vertices arrive according to a standard Poisson process and are placed onto the unit torus. Given the graph up to time t, a vertex x arriving at time t connects to each already present vertex y independently with probability

$$\rho\Big(\frac{t\,\mathrm{d}_1(x,y)}{f(\mathrm{indeg}_t(y))}\Big). \tag{1.1}$$

Here, ρ is a non increasing and integrable *profile function* with image in $[0,1]$ and d_1 denotes the torus metric. In the denominator, $\mathrm{indeg}_t(y)$ denotes the number of connections of y to vertices that have been added to the graph after it and up up time t and f is a function of asymptotic linear slope $\gamma \in (0,1)$, i.e. $f(x)/x \to \gamma$ as $x \to \infty$ modelling the influence of the vertices degree. Further comments and details on the above mentioned objects are given in the section below. The choice of $\rho = p\mathbb{1}_{[0,1/p]}$ reduces to a continuous time version of the model of Aiello et al.. However, the introduction of a general profile function allows for more flexibility in the model. In particular, one can chose functions of unbounded support, softening geometric restrictions.

The previously described spatial preferential attachment models appear to be too complicated to fully characterise features like robustness and ultra-smallness. This is because the actual degree of each vertex depends in a complex way on the graph geometry. At the same time there is a strong link between the age of a vertex and its degree due to the building principles. The age-based preferential attachment model we study in this thesis can be seen as a simplification and approximation of Jacob and Mörter's model where the denominator in the connection function is replaced by the expected degree of the vertex y at time t. As the latter is just a function of the age of a vertex which is a given quantity, this removes complicated but on large scales inessential correlations between edges. This allows more explicit calculations and we can focus on the more complex question of robustness and its dependence on the spatial embedding. In the next section, we introduce the model properly and formulate the main results about it.

1.1.1. A preferential attachment model and results

In this section we introduce our model, state and discuss its features of interest. In Section 1.2 a limit structure of the model which is necessary to prove the features of our preferential attachment model is introduced. This limit furthermore leads to a new class of graphs, *the weight-dependent random connection model*. Both sections combined can be seen as a summary and explanation of the results of [32, 35, 36]. It is further explained how the results of the papers connect. All three papers are published or submitted for publication and build the core of my thesis. A more detailed statement about which part of the thesis is built on which result and paper and how I was involved in the research and writing process is provided in Section 1.3.

We denote by $\mathbb{T}_1^d = (-1/2, 1/2]^d$ the d-dimensional torus of unit volume, endowed with the torus metric d_1 defined by

$$\mathrm{d}_1(x,y) = \min\Big\{|x-y+u| : u \in \{-1,0,1\}^d\Big\} \text{ for } x,y, \in \mathbb{T}_1^d.$$

We introduce the *age-based spatial preferential attachment model* as a growing sequence of undirected graphs $(\mathscr{G}_t : t \geq 0)$ in continuous time where the vertex set is embedded into $\mathbb{T}_1^d$. We denote a vertex by $\mathbf{y} = (y,s)$ and identify its *spatial position* by $y \in \mathbb{T}_1^d$ and by $s > 0$ its *birth-time*, the time the vertex has been added to the graph. We start with the empty graph $\mathscr{G}_0$. Then

- Vertices arrive according to a standard Poisson process in time and are placed independently uniformly on the d-dimensional torus $\mathbb{T}_1^d$.

- Given the graph $\mathscr{G}_{t-}$, a vertex $\mathbf{x} = (x,t)$ born at time t and placed in position x connects by an edge to each existing node $\mathbf{y} = (y,s)$ independently with probability

$$\rho\left(\frac{t \cdot \mathrm{d}_1(x,y)^d}{\beta\left(\frac{t}{s}\right)^\gamma}\right), \tag{1.2}$$

where

(a) $\rho : (0,\infty) \to [0,1]$ is the *profile function*. It is non increasing, integrable and normalized in the sense that

$$\int_{\mathbb{R}^d} \rho(|x|^d)\, dx = 1. \tag{1.3}$$

The profile function can be used to control the occurrence of long edges and models the strength of the geometrical restrictions.

(b) $\gamma \in (0, 1)$ is a parameter that quantifies the strength of the preferential attachment mechanism. We shall see that it alone determines the power-law exponent of the network.

(c) $\beta \in (0, \infty)$ is an edge density parameter such that larger values of β lead to more edges on average.

Some comments on the choices of the model parameter are in order.

Remark 1.1.

(i) For any $r > 0$, the profile function ρ and the parameter β define the same model as the profile function $x \mapsto \rho(rx)$ and the parameter $r\beta$. Hence the normalisation convention (1.3) represents no loss of generality. Similarly, if the intensity of the arrival process is taken as $\lambda > 0$ the process $(\mathscr{G}_{t/\lambda})_{t>0}$ is the original process with the same profile function ρ and parameter $\beta\lambda$.

(ii) The form of the connection probability (1.2) is natural for the following reasons: To ensure that the probability of a new vertex connecting to its nearest neighbour does not degenerate, as $t \to \infty$, it is necessary to scale $\mathrm{d}_1(x, y)$ by $t^{-1/d}$, which is the order of the distance of a point to its nearest neighbour at time t. Further, the integrability condition of ρ ensures that the expected number of edges connecting a new vertex to the already existing ones, remains bounded from zero and infinity, as $t \to \infty$.

(iii) In the degree-based spatial preferential attachment model of Jacob and Mörters [44], constructed via (1.1), a vertex $\mathbf{y}$ born at time s has expected degree $({}^{t}\!/\!{}_{s})^{\gamma}$ at time t. The age-based spatial preferential model therefore is indeed the aforementioned simplification and approximation of their model. We believe our approximation to be accurate for universal features such as robustness and ultra-smallness as these questions depend only on the global structure of the network which remains largely unchanged by our simplification. However, there are certain questions such as the concrete value of critical parameters or the behaviour at criticality, where small changes in the model description lead to significantly different answers. Answering these questions for the simpler age-based models yields at best insights for the same question for the degree-based version. Indeed, in a non spatial setting certain critical parameters can be calculated explicitly [26] and a difference between degree-based and age-based preferential attachment can be observed.

Before we state the main theorem about preferential attachment, we define the interesting key features of networks from the beginning of this chapter in a rigorous mathematical way. In the definitions, we will use the Landau notation $f = o(g)$ to indicate that $f(x)/g(x) \to 0$ as $x \to \infty$ and $f \asymp g$ if $f(x)/g(x)$ is bounded from 0 and ∞. We denote by $V(G)$ the vertex set of a graph G and by $\sharp A$ the number of elements in a finite set A. We denote by $\mathscr{N}_{\mathbf{x}}(G)$ the neighbours of $\mathbf{x}$ in G.

Definition 1.2 (Scale-free networks)**.** We say a network $(G_t : t \geq 0)$ is *scale-free* with *power-law exponent* τ if there exists a sequence $(\mu(k) : k \in \mathbb{N}_0)$ such that it holds for all $k \in \mathbb{N}_0$

$$\frac{\sharp\{\mathbf{x} \in V(G_t) : \sharp\mathscr{N}_x(G_t) = k\}}{\sharp V(G_t)} \longrightarrow \mu(k)$$

in probability as $t \to \infty$ and

$$\mu(k) = k^{-\tau + o(1)}, \quad \text{as } k \to \infty.$$

Let $V_2(G) \subset V(G)$ be the set of vertices of the finite graph G that have at least two neighbours. For $\mathbf{x} \in V_2(G)$ we call any distinct pair of vertices $\mathbf{y}, \mathbf{z} \in \mathscr{N}_{\mathbf{x}}(G)$ a *wedge* with tip in x. Note that there are $\binom{\sharp\mathscr{N}_{\mathbf{x}}(G)}{2}$ different wedges with tip in $\mathbf{x}$. If further $\mathbf{y}$ and $\mathbf{z}$ are connected in G, i.e. $\mathbf{y} \in N_{\mathbf{z}}(G)$, then we call $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ a *triangle* in G. We now define the *local clustering coefficient* of $\mathbf{x} \in V_2(G)$ by

$$c_{\mathbf{x}}^{\mathrm{loc}}(G) := \frac{\sharp\{\text{triangles in } G \text{ containing } \mathbf{x}\}}{\binom{\sharp \mathscr{N}_{\mathbf{x}}(G)}{2}}$$

and the *average clustering coefficient* of G by

$$c^{\mathrm{av}}(G) := \frac{1}{\sharp V_2(G)} \sum_{\mathbf{x} \in V_2(G)} c_{\mathbf{x}}^{\mathrm{loc}}$$

if $V_2(G) \neq \emptyset$ and as zero otherwise.

Definition 1.3 (Clustering)**.** We say that a network $(G_t : t \geq 0)$ shows *clustering* if there exists a constant $c > 0$ only depending on model parameters such that

$$c^{\mathrm{av}}(G_t) \longrightarrow c,$$

in probability as $t \to \infty$.

To define robustness we introduce *Bernoulli bond percolation* with retention parameter $p \in (0, 1]$. In a graph G any edge is independently removed with probability $1 - p$ or kept with probability p. We denote the resulting graph by $G(p)$. We denote by $\mathscr{C}(G)$ the largest connected component of the graph G. We define robustness as in [36, 45].

Definition 1.4 (Robustness)**.** Let $(G_t : t \geq 0)$ be a network. We say it has a *giant component* if its largest connected component is asymptotically of linear size, that is

$$\lim_{\varepsilon \downarrow 0} \limsup_{t \to \infty} \mathbb{P}\{\sharp \mathscr{C}(G_t) < \varepsilon t\} = 0.$$

We say the network is *robust* if the bond percolated sequence $(G_t(p) : t \geq 0)$ has a giant component for every retention parameter p. Otherwise, we say the network is *non-robust*.

Theorem 1.5 (Age-based spatial preferential attachment)**.**
Let ρ be a profile function, $\gamma \in (0,1)$ and $\beta > 0$. Let $(\mathscr{G}_t : t \geq 0)$ be the age-based spatial preferential attachment model constructed with ρ, γ, β according to (1.2). Then $(\mathscr{G}_t : t \geq 0)$ is a scale-free *network with power-law exponent $\tau = 1 + {}^1\!/\!{}_\gamma$* that exhibits clustering*. If ρ is further of the form*

$$\rho(x) \asymp 1 \wedge \ell(x) x^{-\delta}$$

for some $\delta > 1$ and a slowly varying function ℓ, cf. [7], it holds

(a) if $\gamma < \frac{\delta}{\delta+1}$, the network is non-robust *and if*

(b) if $\gamma > \frac{\delta}{\delta+1}$, the network is robust*.*

Remark 1.6.

(i) The condition $\gamma < {}^\delta\!/\!{}_{(\delta+1)}$ is equivalent to $\tau > 2 + {}^1\!/\!{}_\delta$. Hence, the qualitative change in the robustness behaviour does not occur when τ passes the critical value 3 as observed for the non spatial models but when it passes a strictly smaller value. This shows the significant effect of clustering on the network topology.

(ii) In [44] Jacob and Mörters show that their degree-based spatial preferential attachment model is scale-free with the same power-law exponent $\tau = 1 + {}^1\!/\!{}_\gamma$. They show robustness for their model if $\gamma > {}^\delta\!/\!{}_{(\delta+1)}$ but it remains an open problem to show non-robustness for $\gamma < {}^\delta\!/\!{}_{(\delta+1)}$ [45]. Theorem 1.5 is a strong indication that this is the case.

(iii) Theorem 1.5 shows that the power-law exponent is independent of ρ and only determined by γ and that the model shows clustering for any choice of γ and ρ. However, it also shows that robustness can never occur if ρ has bounded support or decays faster than any polynomial. This applies in particular to an age-based version of the model of Aiello et al [1]. Hence the theorem also indicates that their model is never robust.

(iv) In [25] Eckhoff and Mörters show that robust non-spatial preferential attachment models are vulnerable under targeted attack. More precisely, they show in the robust regime $\gamma > 1/2$ that the removal of an arbitrarily small proportion of the oldest vertices leads to a network that is no longer scale-free and robust. Theorem 1.5 shows that the spatial embedding makes it harder for the network to be robust and one can show for the spatial model that the vulnerability statement remains true, cf. Chapter A.

(v) Combining the limit structure outlined in the next section and frequently used to prove Theorem 1.5 with results of Gracar et al. [33] shows that the age-based spatial preferential attachment model is ultra-small in the robust regime $\gamma > \delta/(\delta+1)$. All together, the age-based spatial preferential attachment model is a tractable model only relying on simple building mechanisms which provides all our features of interest in a fully characterised parameter regime.

(vi) The proof of Theorem 1.5 is done in Chapter 3 and in particular in the Corollaries 3.7, 3.10 and 3.19.

1.2. Inhomogeneous percolation

Percolation theory was introduced by Broadbent and Hammersley in 1957 [13] to model how random properties of a porous medium effects the percolation of a fluid through it. Typically, the medium is modelled as a random graph and the question can be interpreted as the study of this graph's component structure. One of the first models that were studied rigorously are graphs on a lattice and in particular the nearest-neighbour graph $\mathbb{Z}^d$ where edges are present independently with a given probability p and absent with probability $1 - p$. In this case we also speak of *bond percolation*. Another approach is to remove vertices instead in which case we speak of *site percolation*. These lattice models are well studied and more general types of connections have been established; we refer to the monograph of Gilbert [38] for an overview. If the vertex set is embedded into continuum space, typically through a Poisson point process, we speak of *continuum percolation*. One may think here of the spread of bark beetles in a forest. The vertices of the graph then represents the trees and an edge indicates that the bark beetles were transmitted from one tree towards the other. A transmission is more likely to happen when the trees are spatially close. Even though the trees are still a countable number of objects the position of each tree is in continuum space; cf. [59] by Meester and Roy for more examples.

We next consider a rescaling of the age-based spatial preferential attachment model which leads in the limit to an inhomogeneous continuum percolation model, the *age-dependent random connection model*. From this model we derive a general framework that covers many of the models established in the literature which we will then discuss.

1.2.1. Transition to the age-dependent random connection model

Recall the age-based preferential attachment network $(\mathscr{G}_t : t \geq 0)$ constructed according to (1.2). We denote by $\mathbb{T}_t^d = (-\sqrt[d]{t}/2, \sqrt[d]{t}/2]^d$ the torus of volume $t \in (0, \infty]$ and its associated torus metric by d_t where

$$\mathrm{d}_t(x, y) = \min\Big\{|x - y + u| : u \in \{-t^{1/d}, 0, t^{1/d}\}^d\Big\} \text{ for } x, y, \in \mathbb{T}_t^d. \quad (1.4)$$

We identify the case $t = \infty$ with $\mathbb{T}_\infty^d = \mathbb{R}^d$ equipped with the standard Euclidean metric.

For $t > 0$, we define the *rescaling map*

$$\begin{array}{rccc} h_t: & \mathbb{T}_1^d \times (0,t] & \longrightarrow & \mathbb{T}_t^d \times (0,1], \\ & (x,s) & \longmapsto & \left(t^{1/d}x, \frac{s}{t}\right), \end{array}$$

which stretches the space by the factor $\sqrt[d]{t}$ and time by $1/t$. The map also operates on the graph $\mathscr{G}_t$ in the following way: The vertex set of $h_t(\mathscr{G}_t)$ is given by $h_t(V(\mathscr{G}_t))$ and each pair of vertices $h_t(\mathbf{x}), h_t(\mathbf{y})$ is connected by an edge if and only if $\mathbf{x}$ and $\mathbf{y}$ are connected in $\mathscr{G}_t$. Since $V(\mathscr{G}_t)$ is a unit intensity Poisson process on $\mathbb{T}_1^d \times (0,t]$, the set $h_t(V(\mathscr{G}_t))$ is a unit intensity Poisson process on $\mathbb{T}_t^d \times (0,1)$, that is a Poisson process on $\mathbb{T}_t^d$ where each point is marked with an independent random variable distributed uniformly on $(0,1]$ [55, Chapter 5]. Moreover,

$$\rho\left(\frac{u/t \cdot \mathrm{d}_t(\sqrt[d]{t}x, \sqrt[d]{t}y)^d}{\beta\left(\frac{u/t}{s/t}\right)^{\gamma}}\right) = \rho\left(\frac{u \cdot \mathrm{d}_1(x,y)^d}{\beta\left(\frac{u}{s}\right)^{\gamma}}\right). \tag{1.5}$$

Hence, h_t preserves the connection rule (1.2) and it is the same to construct the graph $\mathscr{G}_t$ and then rescale the marked space or to first rescale the marked space and then construct the graph, see Figure 1.1.

Let $\mathscr{G}^t$ be the graph, constructed on the points of a unit intensity Poisson process on $\mathbb{T}_t^d \times (0,1]$ where, given the vertex set, each pair of vertices $\mathbf{x} = (x,u)$ and $\mathbf{y} = (y,s)$ with $s < u$ is connected by an edge independently with probability

$$\rho\left(\frac{u \cdot d_t(x,y)^d}{\beta\left(\frac{u}{s}\right)^{\gamma}}\right).$$

Hence, the connection probability coincides with (1.5) and $\mathscr{G}^t$ has the same law as $h_t(\mathscr{G}_t)$. Hence, to study degree-distribution, clustering and robustness for $\mathscr{G}_t$ is the same as studying these questions for $\mathscr{G}^t$ which we will do in Chapter 3. However, as a process $(\mathscr{G}^t : t \geq 0)$ behaves differently as the original process $(\mathscr{G}_t : t \geq 0)$; while the degree of any given vertex in the original process $(\mathscr{G}_t : t \geq 0)$ goes to infinity, the degree of any fixed vertex in $(\mathscr{G}^t : t \geq 0)$ stabilises. Due to the latter, $(\mathscr{G}^t : t \geq 0)$ converges locally to a limit graph $\mathscr{G}^\infty$ in the sense that for large t bounded graph neighbourhoods in $\mathscr{G}^t$ and $\mathscr{G}^\infty$ coincide. This will be proven rigorously in

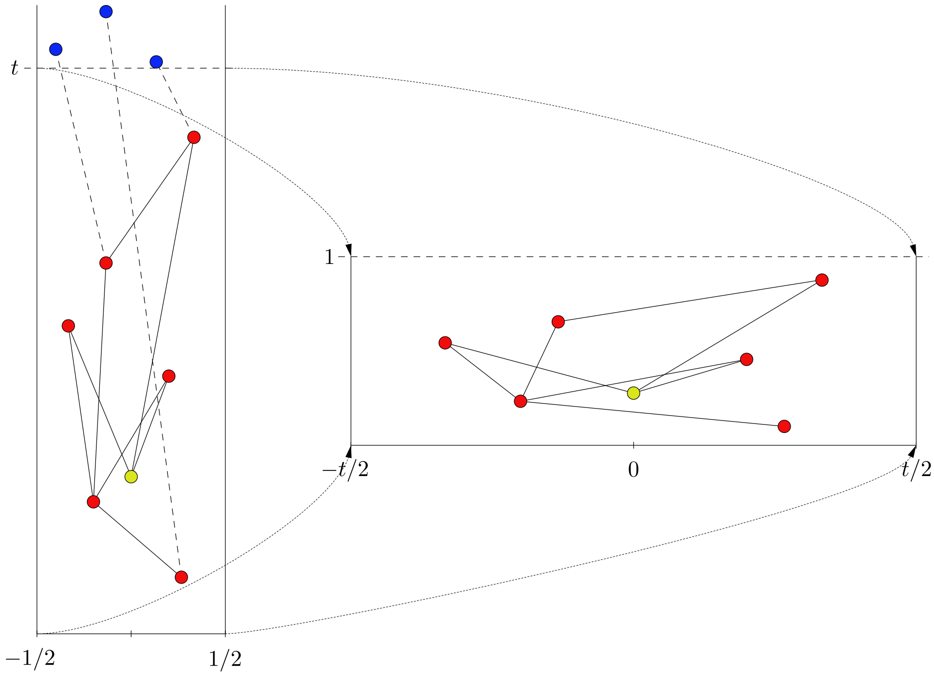

Figure 1.1.: The graph $\mathscr{G}_t$ on the left and its rescaling $h_t(\mathscr{G}_t)$ on the right. The blue vertices are born after time t and, therefore, the corresponding edges do not exist yet and the vertices are not part of the rescaled graph. The yellow vertex is placed at position 0 and remains in the centre after the rescaling; cf. [32, Figure 1].

Theorem 3.1. We will further show in Theorem 3.3 that $\mathscr{G}^\infty$ also plays the role of the weak local limit for $\mathscr{G}_t$ in the sense of Benjamini and Schramm [5]. Intuitively, this concept states that asymptotic local properties of a graph sequence can be studied as they are seen by a typical vertex in the limit graph. This concept has also been independently introduced by Aldous and Steele [3]. The local limit structure will be shown to be crucial for the proof of Theorem 1.5. We conclude this section with a closer look on the limit graph $\mathscr{G}^\infty$. By construction its vertex set is given by a Poisson process on $\mathbb{R}^d$ where each vertex carries an independent and uniformly on $(0,1)$ distributed mark and, given their locations and marks, any pair of vertices $\mathbf{x} = (x, u)$ and $\mathbf{y} = (y, s)$ is connected by an edge independently with probability

$$\rho\Big(\tfrac{1}{\beta}(u \wedge s)^\gamma (u \vee s)^{1-\gamma} |x - y|^d\Big).$$

The connection probability therefore depends on the Euclidean distance of the vertices and their marks in a way where edges between spatially close vertices and connections to vertices with small marks are more probable. Here, small marks coincide with early birth times of vertices. We call this model the *age-dependent random connection model*. It is a model of independent interest as an inhomogeneous version of the random connection model [59] which arises naturally. However, by making the way the marks influence the connection mechanism more flexible this approach gives rise to a general class of inhomogeneous spatial random graphs which we call the *weight-dependent random connection model*. It was first introduced by Gracar et al. [34] in 2019. We shall see that many well-established models in the literature fall within this class.

1.2.2. A general inhomogeneous percolation model

We introduce the weight-dependent random connection model as in [34]. The model's building mechanism is similar as above and has two principal components:

- the *kernel*, a symmetric function $g : (0,1)^2 \to (0,\infty)$ which is non decreasing in both arguments and satisfies

$$\int_0^1 \mathrm{d}s \int_0^1 \mathrm{d}t \, \frac{1}{g(s,t)} < \infty; \tag{1.6}$$

- the *profile*, a non increasing function $\rho : (0,\infty) \to [0,1]$ satisfying

$$\int_{\mathbb{R}^d} \rho(|x|^d)\mathrm{d}x = 1 \tag{1.7}$$

  as above in (1.3).

We generate a graph in two steps: Firstly, we sample vertices. Let η be a unit intensity Poisson point process on $\mathbb{R}^d$. We refer to the points of η as *vertex locations*. Given η, each location $x \in \eta$ is assigned an independent *vertex mark* t_x, distributed uniformly on $(0,1)$. The pairs $\mathbf{x} = (x, t_x)$ form the vertex set and we denote by $\mathcal{X}$ the collection of all vertices. Secondly, we fix a $\beta > 0$, a profile ρ and a kernel g and, given $\mathcal{X}$, we connect every pair of vertices $\mathbf{x}, \mathbf{y} \in \mathcal{X}$ independently by an edge with probability

$$\rho\Big(\tfrac{1}{\beta} g(t_x, t_y) |x - y|^d\Big) \tag{1.8}$$

and we denote this event by $\mathbf{x} \sim \mathbf{y}$. We denote the resulting graph by $\mathscr{G}(\beta)$ and by $(\mathscr{G}(\beta) : \beta > 0)$ the family of graphs of various edge densities.

Remark 1.7.

(i) As before, the normalisation condition of ρ and the fixed unit intensity of the underlying Poisson process provide no loss of generality.

(ii) By the monotonicity assumptions on ρ and g, short edges and connections to vertices with small marks are more probable. One can think of the marks as inverse vertex weights, which give the model its name.

(iii) By monotonicity of ρ, the parameter $\beta > 0$ affects the edge probability in the way where larger values of β result in more edges on average. We hence call β the *edge density*.

(iv) As we shall see in Chapter 2, the integrability conditions of g and ρ guarantee that $\mathscr{G}(\beta)$ has finite mean degree which in particular means that the considered graphs are sparse. Moreover, due to the normalisation of ρ, the degree distribution only depends on g and the edge density β whereas the profile ρ controls the intensity of long edges in the graph. This allows to independently tune the degree distribution and the occurrence of long edges.

(v) By construction, edges are conditionally independent given $\mathcal{X}$. However, due to the introduction of marks, edges are positively correlated whenever they share a common end vertex.

(vi) Throughout the manuscript, we always consider a Poisson point process as the underlying vertex locations. However, the Poisson point process can easily be replaced by a site percolated lattice, that is the lattice $\mathbb{Z}^d$ where each site is independently removed with a fixed probability $p \in [0, 1)$.

We now turn our attention to the second focus of this thesis: Does a critical density parameter $\beta_c \in (0, \infty)$ exist, such that the graph $\mathscr{G}(\beta)$ contains an infinite connected component, or *infinite cluster,* for $\beta > \beta_c$ and no infinite cluster for $\beta < \beta_c$? The ergodicity of the vertex set together with the conditional independence of edges yields that the existence of an infinite cluster is a 0-1-event. Moreover, one can adapt [45, Proposition 4] to our setting to deduce that an infinite cluster, if it exists, is almost surely unique. This proof follows an established technique of Burton and Keane [14]. We will call the parameter regime $(0, \beta_c)$ the *subcritical phase* and the regime (β_c, ∞) the *supercritical phase* of the model. In light of our discussion about robust networks a natural equivalent to this question in the context of percolation is whether there exist sparse models without a subcritical phase, i.e. $\beta_c = 0$.

Many models from the literature fit into the framework of the weight-dependent random connection model by tuning ρ and g. In Chapter 2 we will discuss a general kernel, the *interpolation-kernel*, which contains all the examples below and allows us to draw a rather complete picture of percolation results from our perspective. However, in the remainder of this chapter, we restrict ourselves to the following kernels corresponding to concrete models from the literature, which we now discuss. An overview about these kernels and the models they represent can be found in [34, Table 1].

Homogeneous random connection model In this class of graphs, vertices are sampled according to a standard Poisson process and given two points the probability of drawing an edge between them is a function of their spatial distance. In the framework of the weight-dependent random connection model the vertex marks have no influence on the connection probability and these graphs are defined according to the *plain-kernel* $g^{\text{plain}}(s,t) = 1$. The connection probability (1.8) then reads $\rho({}^1\!/\!_\beta|x-y|^d)$. This class was first introduced by Penrose in 1991 [64] under the name *random connection model* and various models can be derived by varying ρ. The first model belonging to this class was introduced by Gilbert in 1961 [30]. Here, any pair of vertices is connected by an edge if their distance is beneath a fixed threshold β corresponding to the choice of $\rho = \mathbb{1}_{[0,1/2]}$. It is well-known that this model has a non-trivial phase transition at $\beta_c \in (0,\infty)$ in dimension $d \geq 2$ and no supercritical phase, i.e. $\beta_c = \infty$, in dimension $d = 1$. This behaviour extends to all models with bounded edge length and is essentially a consequence of the existence of a supercritical phase in two dimensional nearest-neighbour percolation on the lattice. In dimension one however one finds gaps in the Poisson process without any vertex of arbitrary length and an infinite component can never be formed. The same holds true if ρ decays faster than polynomial at infinity. Such a ρ allows arbitrary long edges but with such small probabilities that there are no qualitative changes to β_c, cf. [60] in particular for profile-functions with exponential tails. Hence, one may also speak of *short-range* percolation. The behaviour in dimension $d \geq 2$ extends to all following models and the existence of a supercritical phase is the regular case. The remaining question in this case is then whether $\beta_c > 0$ or not.

A different approach to building the graph is to connect vertices over a long distance where the connection probability decays polynomially in the distance. This model was first introduced on the one-dimensional lattice as *long-range percolation* [67] and extended to and studied on the Poisson process by Penrose [64] and Meester et al. [57]. In the language of the weight-dependent random connection model, this model coincides with a profile function

$$\rho(x) \asymp 1 \wedge \ell(x)x^{-\delta}, \tag{1.9}$$

where $\delta > 1$ and $\ell(x)$ is a slowly varying function. In contrast to Gilbert's model above, extra randomness on the edges is introduced. In this setting, it again holds $\beta_c \in (0, \infty)$ in any dimension $d \geq 2$. The picture for $d = 1$ however changes and it is known that $\beta_c < \infty$ for $\delta \in (1, 2)$ and $\beta_c = \infty$ for $\delta > 2$ [61, 67]. Moreover, even the boundary regime $\delta = 2$ is known in this model when the slowly varying correction term ℓ is constant. This is also referred to as *scale-invariant long-range percolation* in the literature. It holds $\beta_c < \infty$ in this case and it is even known that there exists an infinite cluster at the critical point $\beta = \beta_c$ [2, 24] which is rather atypical. These proofs are given for the original lattice model but can be transfered to the Poisson process as well. The behaviour at criticality of the homogeneous random connection model in higher dimensions are studied by Heydenreich et al. [40]. The results of the preceding discussion can be found summarised in the monograph of Meester and Roy [59]. These models also have a finite volume analogue where the vertices are embedded into a bounded domain. For the short-range regime, we refer to the monograph of Penrose [63] about random geometric graphs. More general profile functions are studied by Penrose in [65] under the name *soft random geometric graphs*. In the following, we refer to all models constructed with the plain-kernel as *homogeneous random connection models*.

Poisson Boolean model and associated graphs In the classical Poisson Boolean model each Poisson point is assigned an i.i.d. radius and the model is given by the union of the balls centred around the Poisson points with the assigned radii. One is interested in the space covered by the balls and whether the covered space contains an unbounded connected component. A similar representation is given by a graph where every pair of vertices is connected by an edge if their associated balls intersect which then can be seen as a generalisation of the Gilbert graph above. The question of existence of an unbounded connected component in the union of balls is equivalent to the existence of an infinite component in the graph. However, a graph contains more structural informations and allows for further variants and generalisations of the model. To ensure finite expected degrees the radius distribution must have finite d-th moment. The Boolean model has been studied by various authors, e.g. Hall in 1985 [39] and Meester and Roy in 1994 [58]. Again, the results are summarised in detail in the monograph [59]. It is known, that in $d \geq 2$ there exists a non-trivial $\beta_c \in (0, \infty)$ and in $d = 1$ it holds $\beta_c = \infty$. This holds true for all valid radius distributions with finite d-th moment and applies in particular to heavy-tailed radii which allow vertices with an exceptionally large radius and degree. These are the radius distributions we will focus on in the following. In our framework, the profile function is given by the indicator $\mathbb{1}_{[0,1/2]}$ and the radius distribution is derived from the vertices' marks. We then define the graph derived from the Boolean model via the *sum-kernel*

$$g^{\mathrm{sum}}(s,t) = \left(s^{-\gamma/d} + t^{-\gamma/d}\right)^{-d}, \text{ for } \gamma \in (0,1). \tag{1.10}$$

The radius of a vertex $\mathbf{x} = (x, t_x)$ is then given by $((\beta/2)t_x^{-\gamma})^{1/d}$ and hence heavy-tailed. We shall see that these radii lead to a scale-free degree distribution with power-law exponent $\tau = 1 + 1/\gamma$. Note that this is the power-law exponent of the age-based spatial preferential attachment model. To allow a direct comparison of our models, we will parametrise all models in terms of $\gamma \in (0, 1)$ in the following leading to the same power-law exponent. A closely related variant of the Boolean model graph is given by the *strong-kernel*

$$g^{\mathrm{str}}(s,t) = (s \wedge t)^{\gamma}, \text{ for } \gamma \in (0,1). \tag{1.11}$$

Here, two vertices are connected by an edge if the vertex with the smaller radius is contained in the ball associated with the vertex with the larger radius. Hence, whether the vertices are connected or not only depends on the *stronger* vertex. However, the radii are heavy-tailed and the order of the sum of two radii is the same as the order of the larger radius. Indeed,

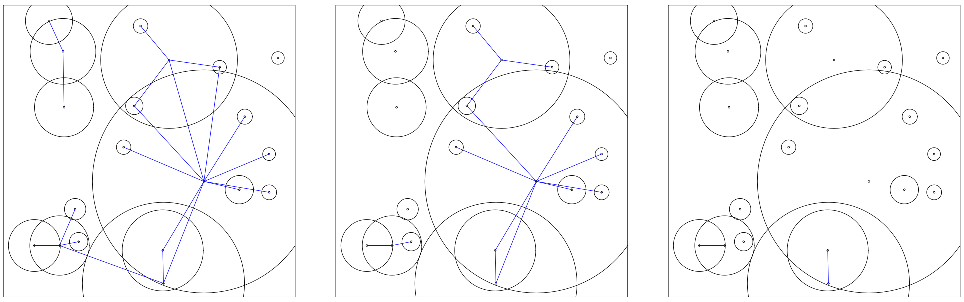

Figure 1.2.: Examples for the model with the sum-kernel (left), strong-kernel (middle) and weak-kernel (right) where ρ is an indicator function.

as

$$2^{-d} g^{\mathrm{str}} \le g^{\mathrm{sum}} \le g^{\mathrm{str}}, \tag{1.12}$$

both kernels show qualitatively similar behaviour. The strong-kernel model has also been studied also under the name *scale-free Gilbert graph* by Hirsch in 2017 [41]. If one allows $\gamma = 0$, one obtains the Gilbert graph of the previous paragraph. Due to (1.12), in all our calculations we will work with the strong kernel only.

A similar yet much more restrictive idea is to make the connection probability dependent on the *weaker* vertex. That is, two vertices are connected whenever they are contained in each others' associated ball. This corresponds to the *weak-kernel*

$$g^{\mathrm{weak}}(s,t) = (s \vee t)^{\alpha}, \ \text{for } \alpha > 0. \tag{1.13}$$

Typically, the weaker radius is of much smaller order than the stronger one leading to a much less connected graph. If the radius distribution remains integrable as before, i.e. $\alpha \in (0,1)$, the degree distribution no longer follows a power-law. To achieve the same power-law $\tau = 1 + {}^{1}\!/\!{}_{\gamma}$ as above, one has to choose $\alpha = 1 + \gamma$. This model then coincides to a continuum version of Yukich's *ultra-small scale-free geometric network* [70]. In this situation, the radii are big enough that we always have $\beta_c = 0$.

All of the three discussed models have in common that each vertex has assigned a sphere of influence and edges are drawn according to a mechanism depending on the size of the involved spheres. Due to (1.12) we refer to a model that is constructed with either the sum-kernel or the strong kernel as *Boolean model*. Since the weak kernel behaves significantly different it is not contained in the Boolean model. Instead we simply refer to it as weak-kernel model. Examples for the three models are given in Figure 1.2.

Scale-free percolation This model was introduced by Deijfen et al. on the lattice in 2013 [17] and translated to continuum space by Deprez and Wüthrich in 2019 [18]. It can be seen as a generalisation of *hyperbolic random graphs* [9, 54, 68]. Here, each vertex carries an independent and heavy-tailed weight and the weights enter the connection probability as a product. In [18] the weights are Pareto distributed. In [17] a further slowly varying correction term is allowed in the weights' tail distribution function. However, for their percolation results they restrict themselves to tail distribution functions that can be suitably bounded by a Pareto tail with the same tail index. Hence, for their results Pareto distributed weights are paradigmatic. In the notion of the weight-dependent random connection model this coincides with the *product-kernel*

$$g^{\text{prod}}(s,t) = s^{\gamma}t^{\gamma}, \text{ for } \gamma \in (0,1) \tag{1.14}$$

which again lead to a power-law degree distribution with $\tau = 1 + {}^{1}\!/\!_{\gamma}$. For a translation from our parametrisation to the original one, we refer to [34, Table 2]. The typical choices for profile functions in this model are the ones of long-range percolation given by (1.9). In this model it is known that $\beta_c = 0$ if $\gamma > {}^{1}\!/\!_{2}$ and $\beta_c > 0$ if $\gamma < {}^{1}\!/\!_{2}$. This question hence only depends on the degree distribution and not on the geometry modelled by ρ and the phase transition is the same as for the non spatial models discussed in Section 1.1. If ρ fulfils (1.9) for some $\delta > 1$ and $\gamma < {}^{1}\!/\!_{2}$, we observe in dimension $d = 1$ that $\beta_c < \infty$ if $\delta < 2$ and $\beta_c = \infty$ if $\delta > 2$ [18, Theorem 3.2]. Here, the picture equals the one of long-range percolation and the weights have no qualitative effect. Shortly, we discuss some models where both questions depend on the combination of both the power-law and the long-range effects. This model has also been studied under the name *geometric inhomogeneous random graphs* by Bringmann et al. [11, 12] and Komjáthy et al. [52]. They use a parametrisation closely related to the weight-dependent random connection model that allows generalisations in a very similar way.

The age-dependent random connection model For completeness, we recall the previously discussed age-dependent random connection model that is given by the *preferential attachment-kernel*

$$g^{\mathrm{pa}}(s,t) = (s \wedge t)^{\gamma}(s \vee t)^{1-\gamma}, \text{ for } \gamma \in (0,1). \tag{1.15}$$

1.2.3. Main percolation results

In the examples above we have seen models that have parameter regimes providing no subcritical phase. However, in both situations this only depends on the degree distribution: the influence of the vertex weights on the connection probability is so strong that geometric restrictions have no influence. Beside the question of this behaviour in the age-dependent random connection model, we are therefore interested in whether there exist models in general where the existence of a subcritical phase depends on both the degree distribution and the geometry. Both of these questions will be answered in the upcoming Theorem 1.8. Before stating it, let us comment on the geometry of the graph and how the choice of ρ effects it. The two canonical choices for profile-functions in the examples above are the indicator function, i.e. the short-range regime, and a function of polynomial decay, i.e. the long-range regime. In the homogeneous short-range models whether two vertices are connected or not only depends on their distance and hence the underlying Poisson process. In the homogeneous long-range models, (very) long edges can occur spontaneously often enough so that geometric restrictions are softened. These edges are also referred to as *weak links* in the networks literature. This extends naturally to the general setting. If $\rho = \mathbb{1}_{[0,1/2]}$, two vertices are connected if their "mark and kernel scaled distance" fulfils

$$g(t_x, t_y)|x - y|^d \leq \frac{\beta}{2}.$$

There is a hard bound given by β for the product of kernel and distance. We therefore refer to these models as *hard models*. For a profile-function that fulfils (1.9) we instead have the following. Given the Vertex set $\mathcal{X}$, each potential edge $\{\mathbf{x}, \mathbf{y}\}$ is assigned an independent heavy tailed random variable $Z(\mathbf{x}, \mathbf{y})$ with tail distribution function $\mathbb{P}\{Z(\mathbf{x}, \mathbf{y}) > z\} = \rho(z)$ and the edge is drawn if

$$g(t_x, t_y)|x - y|^d \leq \beta Z(\mathbf{x}, \mathbf{y}). \tag{1.16}$$

The bound on the right-hand side is therefore randomised and we speak of *soft models*. An advantage of this representation is the justification of considering $(\mathscr{G}(\beta) : \beta > 0)$ as a growing family of graphs (in β) allowing a direct coupling between graphs with various values of β. For each given realisation of the vertex set $\mathcal{X}$ and the collection $Z = (Z(\mathbf{x}, \mathbf{y}) : \mathbf{x} \neq \mathbf{y} \in \mathcal{X})$, the "mark and kernel scaled distance" and the edge weights Z are given quantities and whether the above inequality is fulfilled or not depends only on β. Obviously, for the fixed realisation and $\beta_1 < \beta_2$, the graph $\mathscr{G}(\beta_2)$ contains all the edges of $\mathscr{G}(\beta_1)$ and possibly more. The picture is particularly nice for the *soft Boolean model*. Plugging in the sum kernel (1.10), inequality (1.16) becomes

$$|x - y| \leq \left(\beta Z(\mathbf{x}, \mathbf{y}) t_x^{-\gamma}\right)^{1/d} + \left(\beta Z(\mathbf{x}, \mathbf{y}) t_y^{-\gamma}\right)^{1/d}.$$

Hence, for each potential edge, the radii of the two end vertices are stretched by the same independent heavy tailed random variable and the edge is drawn if the stretched balls intersect, cf. [33]. In the following our main focus lies on soft models and we therefore consider profile functions ρ that are *regularly varying* at infinity with index $-\delta$ for some $\delta > 1$. More precisely, ρ fulfils

$$\lim_{r\to\infty} \frac{\rho(\lambda r)}{\rho(r)} = \lambda^{-\delta} \quad \text{for all } \lambda \geq 1; \tag{1.17}$$

see [7] for a monograph on regularly varying functions. Note that in particular functions of the form (1.9) are regularly varying. We can use a comparison argument to derive the behaviour of profile-functions with lighter tails (including those with bounded support) from a limit $\delta \uparrow \infty$. Therefore, to include the hard models also, we identify the case $\rho = \mathbb{1}_{[0,1/2]}$ with $\delta = \infty$. We set $1/\infty := 0$. We now state the promised theorem giving a positive answer to the question of percolation being affected by both marks and spatial embedding and contextualising the behaviour of the age-dependent random connection model.

Theorem 1.8 (Existence vs. non-existence of the subcritical phase)**.** *Let $(\mathscr{G}(\beta) : \beta > 0)$ be the weight-dependent random connection model constructed with either the preferential attachment-kernel* (1.15)*, the sum-kernel* (1.10) *or the strong-kernel* (1.11) *for some $\gamma \in [0, 1)$ and a profile-function ρ satisfying* (1.17) *for some $\delta \in (1, \infty]$. It holds*

(a) if $\gamma < \frac{\delta}{\delta+1}$, then $\beta_c > 0$;

(b) if $\gamma > \frac{\delta}{\delta+1}$, then $\beta_c = 0$.

Remark 1.9.

(i) We obtain the following estimates for β_c from our proof:

- If $\gamma < 1/2$, then $\beta_c \geq \frac{1-2\gamma}{4}$.

- If $\rho(x) \leq Ax^{-\delta}$ for $A > 1$ and $1/2 \leq \gamma < \delta/(\delta+1)$, then

$$\beta_c \geq \frac{d(\delta-1)(\delta(1-\gamma)-\gamma)}{2^{d\delta+4}A^{1/\delta}\delta}.$$

(ii) We can use a simple coupling with the homogeneous random connection model to obtain that if $\gamma < \delta/(\delta+1)$ in $d \geq 2$ and for $d = 1$ if additionally $\delta < 2$ we have $\beta_c < \infty$. We deal with the case $\delta > 2$ in $d = 1$ below in Theorem 1.11.

(iii) From the perspective of drawing edges, the edge density parameter β scales the distance between points. Hence varying β is equivalent to varying the intensity of the Poisson process which can be seen easily by performing a linear coordinate transform on the underlying space and using that a homogeneous Poisson point process is uniquely characterised by its intensity and the condition that point counts in disjoint Borel sets be independent [55]. Hence, if we build the graph with a fixed $\beta > 0$ and vary the intensity λ of the Poisson process, we obtain $\beta_c = 0 \Leftrightarrow \lambda_c = 0$ and $\lambda_c < \infty \Leftrightarrow \beta_c < \infty$. Similarly, performing Bernoulli bond percolation with retention parameter p on the graph $\mathscr{G}(\beta)$ coincides with constructing the graph with profile-function $p\rho$ by the conditional independence of the edges given the vertices. Since $p\rho$ has the same tail as ρ and in particular fulfils (1.17) with the same δ, we get $p_c = 0 \Leftrightarrow \beta_c = 0$. In a parameter regime where $\beta_c < \infty$, we can build the graph with a fixed $\beta > \beta_c$ and have $p_c < 1$ in this case.

(iv) If we choose $\gamma = 0$ and the strong or the sum-kernel, we obtain the known results for homogeneous random connection model. If we choose the strong or the sum-kernel for some $\gamma > 0$ and a profile-function of bounded support, i.e. $\delta = \infty$, we reproduce the known results for the hard Boolean model. Choosing $\gamma = 0$ and the preferential attachment-kernel, we obtain the weak-kernel at critical $\gamma = 0$ and find $\beta_c > 0$ in that case.

(v) If $\gamma = \delta/(\delta+1)$, we do not expect a universal result and it may depend on the exact form of the kernel g and the profile ρ whether $\beta_c = 0$ or not. However, for $\gamma = 1/2$ the product-kernel (1.14) and the preferential attachment-kernel coincide. Hence, the scale-free percolation model in continuum space has $\beta_c > 0$ at $\gamma = 1/2$ for a universal class of profile-functions. The findings of the last two comments are summarized in the following corollary.

Corollary 1.10. *Consider the weight-dependent random connection model* $(\mathscr{G}(\beta) : \beta > 0)$ *constructed with a profile-function* ρ *satisfying* (1.17) *for some* $\delta \in (1, \infty]$. *Then we have* $\beta_c > 0$ *for*

(a) the weak-kernel at critical $\gamma = 0$, *that is* $g^{weak}(s,t) = s \vee t$, *and*

(b) the product-kernel at critical $\gamma = 1/2$, *that is* $g^{prod}(s,t) = \sqrt{st}$.

The final important result of this thesis addresses the existence of a supercritical phase in dimension $d = 1$, i.e. when the underlying Poisson process is embedded into the real line. We have already seen in the model description above that in the homogeneous setting the supercritical phase can only exist when there are sufficiently many long-range edges. In other words, the critical density parameter β_c is finite if ρ fulfils (1.17) for $\delta \in (1, 2)$. We however consider inhomogeneous models where additional edges are drawn through the influence of the vertex weights. Of particular interest are the models constructed with some $\delta > 2$ together with the kernels introduced above and the question at hand is whether they can contain an infinite cluster or not. The following theorem characterises the overall long-range connectivity of the models and relates it to the finiteness of β_c.

Theorem 1.11 (Existence vs. non-existence of a supercritical phase). *Consider the one-dimensional weight-dependent random connection model* $(\mathscr{G}(\beta) : \beta > 0)$ *with kernel-function* g *and profile-function* ρ.

(a) The percolation threshold β_c *is finite, whenever*

$$-\liminf_{n\to\infty} \frac{\log\left(\int_{1/n}^{1}\int_{1/n}^{1}\rho(g(s,t)n)\,\mathrm{d}s\,\mathrm{d}t\right)}{\log n} < 2.$$

(b) The percolation threshold β_c *is infinite, whenever*

$$-\limsup_{n\to\infty} \frac{\log\left(\int_{1/n}^{1}\int_{1/n}^{1}\rho(g(s,t)n)\,\mathrm{d}s\,\mathrm{d}t\right)}{\log n} > 2.$$

Remark 1.12.

(i) For the concrete kernels above and our choices for profile-functions, fulfilling (1.17), the limits appearing in the theorem coincide. If this is the case, we define the *effective decay exponent* associated with ρ and g as

$$\delta_{\text{eff}} := -\lim_{n\to\infty} \frac{\log\left(\int_{1/n}^{1}\mathrm{d}s\int_{1/n}^{1}\mathrm{d}t\ \rho(g(s,t)n)\right)}{\log n}. \tag{1.18}$$

In this situation Theorem 1.11 states that, $\beta_c < \infty$ if $\delta_{\text{eff}} < 2$ and $\beta_c = \infty$ if $\delta_{\text{eff}} > 2$. Note that in case of a regularly varying profile-function it always holds that $\delta \geq \delta_{\text{eff}}$ and one obtains a strict inequality if g vanishes sufficiently fast at $(0,0)$.

(ii) Note, that the order in n of the integrals appearing in the theorem are independent of β for our choices of ρ except for the case when $g(s,s) \sim s$ as $s \downarrow 0$ and $\rho = \mathbb{1}_{[0,1/2]}$. Then the integral is always zero. However, we are interested in the behaviour for large β. Hence, we fix a $\beta > 1$ and have

$$\int_{1/n}^{1}\mathrm{d}s\int_{1/n}^{1}\mathrm{d}t\ \mathbb{1}_{[0,1/2]}\left(\beta^{-1}g(s,t)n\right) \asymp \int_{1/n}^{\beta/n}\mathrm{d}s\int_{1/n}^{\beta/n}\mathrm{d}t\ 1 = (\beta-1)^2 n^{-2}$$

and hence $\delta_{\text{eff}} = 2$, the case which is not covered by the theorem.

To give an intuition for δ_{eff}, consider two disjoint sets of n vertices at distance roughly n. As n grows large the smallest mark in each of the two sets is of order $1/n$ so that the integral appearing in (1.18) is essentially the probability that two vertices picked randomly from each of the two sets are connected by an edge. Ignoring the correlations between edges arising from the vertex marks, the number of edges between the two vertex sets of size n is roughly given by a binomial experiment with n^2 trials and success probability $n^{-\delta_{\text{eff}}}$. If $\delta_{\text{eff}} < 2$ the probability of having an edge connecting the two sets increases with n whereas it decreases for $\delta_{\text{eff}} > 2$. The effective decay exponent δ_{eff} hence measures the occurrence of long edges in a way comparable to homogeneous long-range percolation models, seen from a coarse grained perspective. The truncation of the integral bounds in (1.18) is crucial to control the correlations arising from the vertex marks which is a necessity to identify the phase transition for the existence of a supercritical phase correctly. Indeed, at first glance, one might only want to calculate the decay exponent of the marginal distributions of single edges which is the rate at which the annealed probability of two typical vertices at distance n being connected decays. In analogy with homogeneous long-range percolation one might now assume that the graph contains an infinite cluster if this decay exponent is smaller than two and does not contain an infinite cluster if this exponent is larger than two. However, this does not take the aforementioned correlations into account and does not capture the behaviour accurately. In Remark 1.14 (i) below we give an explicit example showing that this heuristic is indeed not sufficient.

Before applying the theorem to our examples, let us comment on the special case $\delta_{\text{eff}} = 2$. This can be seen as a generalisation of the '$1/|x-y|^2$-model' of long-range percolation [2]. This explicit homogeneous model coincides with $\rho(x) \asymp 1 \wedge x^{-2}$ and hence $\delta_{\text{eff}} = \delta = 2$ and it is known that in this $\beta_c < \infty$. If we replace however the profile-function with $\rho(x) \asymp 1 \wedge (x \log(1+x))^{-2}$ we still have $\delta_{\text{eff}} = \delta = 2$ but since now

$$\int_0^\infty \mathrm{d}x \int_0^\infty \mathrm{d}y \ \rho(|x-y|) < \infty,$$

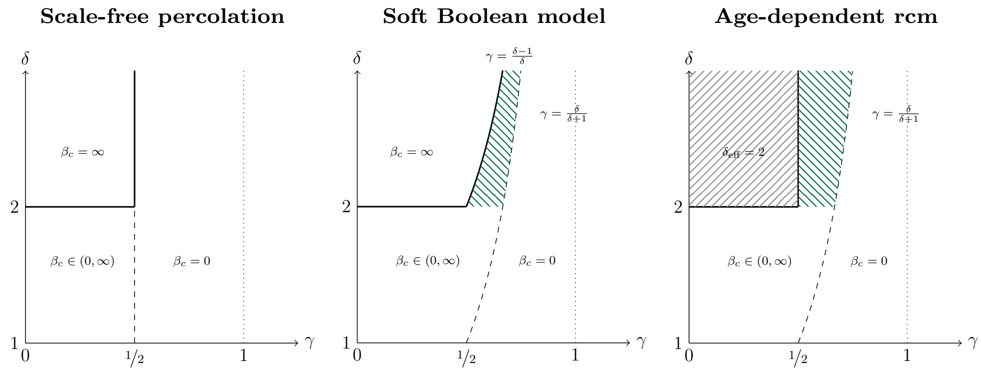

Figure 1.3.: The diagrams show the parameter regimes for γ and δ and the corresponding new results for β_c for the soft Boolean and the age-dependent random connection model and also the known results for the scale-free percolation model for comparison. The green shaded areas are the newly observed regimes where $\beta_c \in (0,\infty)$ even though $\delta > 2$. The gray shaded $\delta_{\mathrm{eff}} = 2$ phase is the part of the scale-invariant regime of the age-dependent random connection model where the finiteness or infiniteness of β_c is unknown.

a standard first moment argument yields $\beta_c = \infty$, cf. [67]. Hence, this regime does not depend on the exponent δ_{eff} alone and we cannot adapt our proof techniques to this boundary case. In analogy with long-range-percolation, we refer to the $\delta_{\mathrm{eff}} = 2$ case as *(weakly) scale invariant*. We now apply Theorem 1.11 to the instances from our examples where the (non-)existence of a supercritical phase in $d = 1$ has not been proven before. The findings are summarised in Figure 1.3. Corollary 1.13 is proven in Section 2.7.2.

Corollary 1.13. *Let $(\mathscr{G}(\beta) : \beta > 0)$ be the weight-dependent random connection in dimension $d = 1$ with profile-function ρ satisfying (1.17) for some $\delta \in (2,\infty]$ and a kernel g.*

(a) For the soft Boolean model, i.e. $g = g^{str}$ (1.11) or $g = g^{sum}$ (1.10), we have

- *if $\gamma < 1 - 1/\delta$, then $\delta_{\mathit{eff}} > 2$ and hence $\beta_c = \infty$ and*
- *if $\gamma > 1 - 1/\delta$, then $\delta_{\mathit{eff}} < 2$ and hence $\beta_c < \infty$.*

(b) For the age-dependent random connection model, i.e. $g = g^{pa}$ (1.15),

we have

- *if* $\gamma \leq 1 - {}^1\!/\delta$*, then* $\delta_{\mathit{eff}} = 2$ *and*
- *if* $\gamma > 1 - {}^1\!/\delta$*, then* $\delta_{\mathit{eff}} < 2$ *and hence* $\beta_c < \infty$*.*

(c) In the soft non scale-free weak-kernel model, i.e. $g = g^{weak}$ (1.13) *with* $\alpha \in (0,1)$*, we have* $\delta_{\mathit{eff}} > 2$ *and hence* $\beta_c = \infty$*.*

Remark 1.14.

(i) The corollary shows together with Theorem 1.8 that the one dimensional soft Boolean model provides three open parameter regimes: one where the graph always contains an infinite cluster, an intermediate regime where an infinite cluster exists for large values of β and one where no infinite cluster can exists. Let us further calculate the decay exponent of the marginal distribution of single edges as mentioned in the explanation above the corollary. If we choose $\rho(x) \asymp 1 \wedge x^{-\delta}$ for $\delta > 2$ and the strong kernel $g^{\mathrm{str}}(s,t) = (s \wedge t)^{\gamma}$ for some ${}^1\!/_2 < \gamma < 1 - {}^1\!/\delta$, we calculate

$$\int_0^1 \mathrm{d}s \int_0^1 \mathrm{d}t \, \Big(1 \wedge (s\wedge t)^{-\gamma\delta} n^{-\delta}\Big) \asymp n^{-\frac{1}{\gamma}} + n^{-\delta} \int_{n^{-1/\gamma}}^1 \mathrm{d}s \; s^{-\gamma\delta} \asymp n^{-\frac{1}{\gamma}}.$$

Since $\gamma > {}^1\!/_2$, the decay of the single edge distribution decays with exponent ${}^1\!/\gamma < 2$ and one might guess $\beta_c < \infty$. However, since also $\gamma < 1 - {}^1\!/\delta$, it holds $\delta_{\mathrm{eff}} > 2$ and we actually have $\beta_c = \infty$ by Corollary 1.13.

(ii) Usually one would expect that the $\delta_{\mathrm{eff}} = 2$ cases correspond to boundary regimes. Corollary 1.13 however shows that the age-dependent random connection model has $\delta_{\mathrm{eff}} = 2$ for the whole parameter regime $0 \leq \gamma \leq 1 - {}^1\!/\delta$ for all profile-functions with $\delta > 2$. This illustrates a scale invariance property that is built into the model which might be a result of the dynamics within the model coming from the role of vertex marks beeing birth times. For $\gamma = 1/2$ the preferential attachment-kernel and the product-kernel coincide and in [12, 43, 53] it is shown that the 'KPKVB-model', a hyperbolic random graph model, has the one-dimensional product-kernel model as weak local limit after a change of coordinates. Results from Bode et al. [8] for the KPKVB-model then show that there exists and infinite cluster in the one-dimensional product-kernel model for all our choices of ρ whenever $\gamma = 1/2$ and β is sufficiently large. By monotonicity, it follows that the same holds true for the preferential attachment-kernel for any $\gamma \geq 1/2$. It remains an interesting open problem to show whether there can be percolation for $\gamma < 1/2$ for natural choices of profile-functions. The findings of this remark are summarised in the following corollary.

Corollary 1.15. *Let $(\mathscr{G}(\beta) : \beta > 0)$ be the weight-dependent random connection model in dimension $d = 1$ with profile-function ρ satisfying (1.17) for some $\delta \in (2, \infty]$ and a kernel g that is either*

(a) the product-kernel at criticality, i.e. $g(s,t) = \sqrt{st}$, or

(b) the preferential attachment-kernel $g = g^{pa}$ (1.15) with $\gamma \in [1/2, \delta/(\delta+1))$,

then it holds $\beta_c \in (0, \infty)$.

1.3. Structure of the thesis

In this section I lay out how the thesis is organised. I further explain, which papers the models and results originate from and what my contribution was. This thesis is based on the results of three papers: [32] with Arne Grauer, Peter Gracar and Peter Mörters, [36] with Peter Gracar and

Peter Mörters and [35] with Peter Gracar and Christian Mönch. The organisation of the thesis is as follows:

The current chapter gives a motivation and introduction to the topics of the thesis, introduces the models under consideration and presents the main results. The preferential attachment model of Section 1.1 and its rescaling in Section 1.2.1 is introduced as origianlly done in [32]. Likewise, the scale-free and clustering property of Theorem 1.5 are results of [32]. Robustness is a result of [36]. The framework of the weight-dependent random connection model in Section 1.2.2 is introduced and discussed as in [35, 36]. Theorem 1.8, its remarks and its corollaries are results of [36]. Theorem 1.11, its remarks and its corollaries are results of [35].

In Chapter 2, the weight-dependent random connection model is constructed more formally, as originally appeared in [35]. In Section 2.4, the existence of a subcritical phase for a certain parameter regime and in Section 2.5 the non existence of the subcritical phase in the other parameter regime is proved; combined proving Theorem 1.8. Chapter B of the appendix contains technical integration results that are needed for the proofs. Both sections and the appendix are from [36]. In the Sections 2.6 and 2.7 the existence and non existence of a supercritical phase in dimension one, based on the value of δ_{eff} is proved as in [35]; concluding the proof of Theorem 1.11.

In Chapter 3, the age-based spatial preferential attachment model is studied. In Section 3.1, its weak local limit is constructed which is used in Section 3.2 to prove results about the degree distribution, clustering and edge lengths as in [32]. In Section 3.3, robustness is proved as in [36], but in greater detail.

In all chapters, the notation might have slightly changed from the relevant papers to guarantee a better readability and connection of the topics. I had major contributions to all three papers as I explain now:

- I contributed significantly to [32], as I participated and contributed in all group discussions leading to the new techniques and ideas and co-authored together with Arne Grauer. I had major involvement in the writing process and its revision.

- My contribution to [36] was essential, as under joint discussions with my colleagues I was responsible for the development of the key ideas and formulations and carried out the technical part of proving the main results in this work. I also contributed significantly to the writing process of this work and its revision. As the proofs of this paper identify the precise parameter regimes for $\beta_c > 0$ and $\beta_c = 0$ in the models of interest as well as robustness and non-robustness in the age-based spatial preferential attachment model, these proofs can be seen as the core of this thesis.

- I contributed significantly to [35], as all ideas and techniques were developed in group discussions and each one of the authors (including myself) contributed crucial ideas. I also did a large part of the writing.

CHAPTER 2

The weight-dependent random connection model

In this chapter we study intensively the weight-dependent connection model introduced in Section 1.2.2. We start with a formal construction of the model as a functional of Point processes. Afterwards we introduce the *interpolation-kernel* which can be seen as a general kernel that describes all of the previously discussed ones. We will use that kernel to describe the neighbourhood and the degree distribution of the graph and draw a complete picture of percolation for our setting. As outlined in Section 1.3, from Section 2.4 on forward we prove the main results. The proofs of Section 2.2.1 follow the arguments of [32]. The proofs of the Sections 2.4 and 2.5 can be found in [36]. The proofs of the Sections 2.6 and 2.7 can be found in [35].

2.1. A formal construction

We construct the weight-dependent random connection model as a functional $\mathcal{G}^{\beta,\rho,g}$ of marked point processes and edge marks similarly as in [36] and [40]. Let η be a stationary ergodic simple point process on $\mathbb{R}^d$ with finite intensity $\lambda > 0$. We use the notation of η_0 for its *Palm version* containing a point at the origin, cf. [55, Chapter 9]. We may and shall frequently use the representation

$$\eta_0 = \{X_j : j \in \mathbb{Z} \text{ such that } X_k <_{\text{lex}} X_\ell \text{ for } k < \ell\}$$

with $X_0 = 0$ where $<_{\text{lex}}$ denotes the strict lexicographic order on $\mathbb{R}^d$. We call the elements of η_0 the *vertex locations* and denote its law by P_0^η. Let further $\mathcal{T}_0 = \{T_j : j \in \mathbb{Z}\}$ be a sequence of independent Uniform$(0,1)$ random variables, independent of η_0, which we call the *vertex marks*. The vertex set is then given by

$$\mathcal{X}_0 := \{\mathbf{X}_j = (X_j, T_j) : X_j \in \eta_0, T_j \in \mathcal{T}_0, j \in \mathbb{Z}\}.$$

We denote its law and expectation by $\mathbf{P}_0$ and $\mathbf{E}_0$. By [55, Proposition 5.5], this is the Palm version of a point process on $\mathbb{R}^d \times (0,1)$ with intensity measure $\mathrm{P}^\eta \otimes \text{Uniform}(0,1)$. We denote the vertex at the origin by $\mathbf{0} := \mathbf{X}_0 = (0, T_0)$ and call it the *root vertex*. The Palm version construction hence ensures that the root is a distinguished typical vertex and we will study the local properties of the graph from its perspective. We finally introduce another family of independent Uniform$(0,1)$ random variables

$$\mathcal{U}_0 := \{U_{i,j} : i < j \in \mathbb{Z}\}$$

which is independent of $\mathcal{X}_0$ and we call its elements *edge marks*.

Now, fix $\beta > 0$, a profile-function ρ, cf. (1.7), and a kernel-function g, cf. (1.6). Then $\mathcal{G}^{\beta,\rho,g}(\mathcal{X}_0, \mathcal{U}_0)$ is the undirected graph with vertex set $\mathcal{X}_0$

and edge set

$$\Big\{\{\mathbf{X}_i,\mathbf{X}_j\} : U_{i,j} \leq \rho\Big(\tfrac{1}{\beta} g(T_i,T_j)|X_i - X_j|^d\Big),\ i < j\Big\}.$$

We denote its law by $\mathbb{P}^{\beta,\rho,g}$ and its expectation by $\mathbb{E}^{\beta,\rho,g}$. To keep notation concise, we write $\mathcal{G}^\beta = \mathcal{G}^{\beta,\rho,g}(\mathcal{X}_0,\mathcal{U}_0)$ as well as $\mathbb{P}^\beta = \mathbb{P}^{\beta,\rho,g}$ and $\mathbb{E}^\beta = \mathbb{E}^{\beta,\rho,g}$ whenever the profile and the kernel are fixed and clear from the context. We denote by $(\Omega,\mathscr{A})$ the underlying measurable space. Note that if η is a unit intensity Poisson process then $\mathcal{X}_0$ is the Palm version of a unit intensity Poisson process on $\mathbb{R}^d \times (0,1)$ and $\mathcal{G}^\beta$ is the Palm version of the graph $\mathscr{G}(\beta)$ from Section 1.2.2. That is the graph $\mathscr{G}(\beta)$ where an additional vertex is added at the origin, marked with a Uniform$(0,1)$ random variable, independent of everything else, and connected to the graph by the same mechanism (1.8) used to build the graph. We will usually assume that η is a Poisson process however this formal construction allows more flexibility in the vertex locations. Indeed, in the construction η can for instance be chosen to be the site percolated lattice $\mathbb{Z}^d$ with retention parameter $p \in (0,1]$. In this case, the stationarity is with respect to shifts induced by $\mathbb{Z}^d$ and η_0 is the percolated lattice conditioned on the event that the root survived the percolation. In fact all our proves hold true in either case. In Theorem 1.11 about one-dimensional percolation even more general vertex locations are covered by our proof. We will comment on this in Section 2.6. Also note that we may have used the Palm version from the beginning as it allows the a priori ordering of the vertices and is the object we work with in our proofs. However, the functional $\mathcal{G}^\beta$ can be used to build graphs from any sets of vertex locations, vertex marks and edge marks. This applies in particular for the non Palm version of the point process above.

Distinguishable vertices To formulate the events we are interested in, we often rely on the existence of distinguishable vertices. The first one of this kind is the root vertex $\mathbf{0} = \mathbf{X}_0 = (0,T_0)$ given by the Palm version. Sometimes it is useful to consider a root with a fixed mark. In order to do so, we use the probability kernel

$$\mathscr{A} \times (0,1) \longrightarrow [0,1], \; (A,u) \longmapsto \mathbb{P}^{\beta}_{(0,u)}(A)$$

for the conditional probability of $\mathbb{P}^{\beta}$ given $T_0 = u$ and we get $\mathbb{P}^{\beta} = \mathbb{P}^{\beta}_{(0,u)}\mathrm{d}u$. Similarly, we can add further given vertices to the graph, say $\mathbf{v}_1 = (v_1, s_1), \dots, \mathbf{v}_n = (v_n, s_n)$, by doing the following: We add the vertices to the graph's vertex set and write $\mathcal{X}_0^{\mathbf{v}_1,\dots,\mathbf{v}_n} = \mathcal{X}_0 \cup \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$. Afterwards, for each vertex $\mathbf{v}_i$, $i = 1, \dots, n$, we sample a sequence $(U_{j,\mathbf{v}_i})_{j\in\mathbb{Z}}$ of independent edge marks and connect it to each vertex $\mathbf{X}_j \in \mathcal{X}_0$ as before if

$$U_{j,\mathbf{v}_i} \leq \rho\Big(\tfrac{1}{\beta} g(T_j, s_i)|X_j - v_i|^d\Big).$$

The resulting graph is given by

$$\mathscr{G}^{\beta}_{\mathbf{v}_1,\dots,\mathbf{v}_n} := \mathscr{G}^{\beta,\rho,g}\Big(\mathcal{X}_0 \cup \{\mathbf{v}_1, \dots, \mathbf{v}_n\}, \mathcal{U}_0 \cup \bigcup_{i=1}^{n} (U_{j,\mathbf{v}_i})_{j\in\mathbb{Z}}\Big)$$

and we denote its law by $\mathbb{P}^{\beta}_{\mathbf{v}_1,\dots,\mathbf{v}_n}$. If η is given by a Poisson process, the Palm version η_0 is given by η where an additional typical vertex is added at the origin. Hence the notation is consistent with the probability kernel given above. In the Poisson process case, the Mecke equation [55, Theorem 4.4] tells us how to use $\mathscr{G}^{\beta}_{\mathbf{v}_1,\dots,\mathbf{v}_n}$ as the graph $\mathscr{G}^{\beta}$ conditioned on the event that $\mathbf{v}_1, \dots, \mathbf{v}_n$ are vertices of the graph; a terminology we will stick to in the following. In the percolated lattice case, no further vertices have to be added. Instead, $\mathbb{P}^{\beta}_{\mathbf{v}_1,\dots,\mathbf{v}_n}$ becomes the probability kernel for the conditional probability given that $\mathbf{v}_1, \dots, \mathbf{v}_n$ survived the site percolation and have marks $s_1, \dots, s_n$.

2.1.1. Percolation notation

For two given vertices $\mathbf{x}$ and $\mathbf{y}$ we denote by $\{\mathbf{x} \sim \mathbf{y}\}$ the event that $\mathbf{x}$ and $\mathbf{y}$ are connected by an edge. We define $\{\mathbf{0} \leftrightarrow \infty\}$ as the event that the root $\mathbf{0}$ is starting point of an *infinite self-avoiding path* $(\mathbf{0} := \mathbf{z}_0, \mathbf{z}_1, \mathbf{z}_2, \dots)$. That is, $\mathbf{z}_i \in \mathcal{X}_0$ for all i, $\mathbf{z}_i \neq \mathbf{z}_j$ for all $i \neq j$ and $\mathbf{z}_i \sim \mathbf{z}_{i-1}$ for all i. Throughout the manuscript, all considered paths are assumed being self-avoiding. If $\{\mathbf{0} \leftrightarrow \infty\}$ occurs for $\mathcal{G}^\beta$, we say that the graph percolates. We denote the percolation probability by

$$\theta(\beta) = \mathbb{P}^\beta\{\mathbf{0} \leftrightarrow \infty\} = \int_0^1 \mathrm{d}u\ \mathbb{P}^\beta_{(0,u)}\{(0,u) \leftrightarrow \infty\}$$

which can be seen as the probability that a typical vertex belongs to the infinite cluster. If $\theta(\beta) > 0$, the graph almost surely contains an infinite cluster. Vice versa, if the graph contains an infinite cluster, the root $\mathbf{0}$ is connected to it with a positive probability and we have $\theta(\beta) > 0$. We can therefore write the critical edge density as

$$\beta_c := \inf\{\beta > 0 : \theta(\beta) > 0\}. \tag{2.1}$$

2.2. The interpolation-kernel

In Section 1.2.2 we have seen various kernels which lead to instances of models from the literature. In all of them we transfer the vertex marks to vertex weights by taking the mark to a negative power. The only exception here is the plain-kernel where the exponent is zero and the marks do not play any role. The idea now is to represent all of these kernels as a single one. This allows a better and direct comparison of the various models. We define the *interpolation-kernel*

$$g_{\gamma,\alpha}(s,t) := (s \wedge t)^\gamma (s \vee t)^\alpha, \quad \text{for } \gamma, \alpha \geq 0. \tag{2.2}$$

Note, that the integrability condition (1.6) for kernels is fulfilled precisely if $\gamma < 1$ and $\alpha < 2 - \gamma$. From the proof of Proposition 2.1 below we can immediately derive that this condition is indeed a necessity for finite mean degrees. Note that all previously discussed kernels can be written as instances of the interpolation-kernel or can be bound from below and above by a constant factor of it. We study the neighbourhood and degree distribution in the following section. In Section 2.2.2 we summarise the percolation results from Chapter 1 in terms of the new kernel. From now on, we work explicitly on a standard Poisson process η. However, all results remain valid if η is chosen to be a site percolated lattice. We comment at the end where small changes in the proofs are to be made.

2.2.1. Neighbourhoods and degree distribution

We describe the neighbourhoods in terms of the root vertex $\mathbf{0}$. We write $\mathscr{N}_{\mathbf{0}} := \mathscr{N}_{\mathbf{0}}(\mathcal{G}^\beta)$ for the neighbourhood of the root in $\mathcal{G}^\beta$. For a given vertex $\mathbf{y} = (y, s)$ we also abbreviate $\mathscr{N}_{\mathbf{y}} := \mathscr{N}_{\mathbf{y}}(\mathcal{G}^\beta_{\mathbf{y}})$. If the mark $T_0 = u$ of the root is given, we also write $\mathscr{N}_{(0,u)}$. It will show helpful to distinguish between connections to vertices with smaller marks, the more influential vertices, and connections to vertices with larger marks, the less influential ones. In order to do so precisely, we think of edges as oriented from the vertex with the larger mark to the vertex with smaller mark. We denote by

$$\mathscr{N}^<_{(y,s)} := \{\mathbf{x} = (x,t) \in \mathcal{X}_0 : (x,t) \sim (y,s), t < s\}$$

the neighbours connected to $\mathbf{y} = (y, s)$ in $\mathcal{G}^\beta_{\mathbf{y}}$ by an outgoing edges and we call $\sharp\mathscr{N}^<_{\mathbf{y}}$ the outdegree of $\mathbf{y}$. Similarly, we denote by

$$\mathscr{N}^>_{(y,s)} := \{\mathbf{x} = (x,t) \in \mathcal{X}_0 : (x,t) \sim (y,s), t > s)\}$$

the neighbours connected by an ingoing edges and call $\sharp\mathscr{N}^>_{\mathbf{y}}$ the indegree of $\mathbf{y}$. The following proposition describes the neighbourhood of the root $\mathbf{0}$ in $\mathcal{G}^\beta$, see Figure 2.1.

Proposition 2.1 (Neighbourhood and degree distribution).
Let $\mathcal{G}^\beta$ be the graph constructed with the interpolation-kernel $g_{\gamma,\alpha}$ for $\gamma \in [0,1)$ and $\alpha \in [0, 2-\gamma)$ and a profile-function ρ.

(a) *For every $u \in (0,1)$, under $\mathbb{P}^{\beta}_{(0,u)}$ the outgoing edges $\mathscr{N}^{<}_{(0,u)}$ form a Poisson process on $\mathbb{R}^d \times (0,u)$ with intensity measure*

$$\lambda^{<}_{(0,u)} := \rho\Big(\beta^{-1} s^{\gamma} u^{\alpha} |y|^d\Big) \,\mathrm{d}s \,\mathrm{d}y.$$

(b) *For every $u \in (0,1)$, under $\mathbb{P}^{\beta}_{(0,u)}$ the ingoing edges $\mathscr{N}^{>}_{(0,u)}$ form a Poisson process on $\mathbb{R}^d \times (u,1)$ with intensity measure*

$$\lambda^{>}_{(0,u)} := \rho\Big(\beta^{-1} s^{\alpha} u^{\gamma} |y|^d\Big) \,\mathrm{d}s \,\mathrm{d}y.$$

(c) *If $\alpha \le 1-\gamma$, the outdegree of $\mathbf{0}$ in $\mathscr{G}^{\beta}$ is stochastically dominated by a Poisson distributed random variable with parameter $\beta/(1-\gamma)$.*
If $\alpha > 1-\gamma$, the outdegree of $\mathbf{0}$ in $\mathscr{G}^{\beta}$ is mixed Poisson distributed with mixing density

$$f^{<}(\lambda) = \tfrac{1}{\alpha+\gamma-1}\Big(\tfrac{\beta\lambda^{-(\alpha+\gamma)}}{1-\gamma}\Big)^{1/(\alpha+\gamma-1)}, \quad \textit{for } \lambda > \tfrac{\beta}{1-\gamma}. \tag{2.3}$$

(d) *If $\alpha < 1$, then the indegree of $\mathbf{0}$ in $\mathscr{G}^{\beta}$ is mixed Poisson distributed where the mixing density fulfils*

$$f^{>}(\lambda) \asymp \tfrac{1}{\gamma}\Big(\tfrac{\beta}{1-\alpha}\Big)^{1/\gamma}\Big(\lambda + \tfrac{\beta}{1-\alpha}\Big)^{-1-1/\gamma}, \quad \textit{for } \lambda > 0. \tag{2.4}$$

If $\alpha > 1$, then the mixing density fulfils

$$f^{>}(\lambda) \asymp \tfrac{1}{\alpha+\gamma-1}\Big(\tfrac{\beta}{\alpha-1}\Big)^{1/(\alpha+\gamma-1)}\Big(\lambda + \tfrac{\beta}{\alpha-1}\Big)^{-(\gamma+\alpha)/(\gamma+\alpha-1)}, \quad \textit{for } \lambda > 0. \tag{2.5}$$

Proof. All outgoing edges of $(0,u)$ connect to end vertices with marks smaller than u. Hence, $\mathscr{N}^{<}_{(0,u)} \subset \mathcal{X}_0 \cap (\mathbb{R}^d \times (0,u))$. Now, given $\mathcal{X}_0$ and $T_0 = u$ each vertex $\mathbf{y} = (y,s) \in \mathcal{X}_0 \cap (\mathbb{R}^d \times (0,u))$ is connected to $(0,u)$ independently with probability $\rho(\beta^{-1} s^{\gamma} u^{\alpha} |y|^d)$. Thus, $\mathscr{N}^{<}_{(0,u)}$ defines a thinning, cf. [55], of $\mathcal{X}_0 \cap (\mathbb{R}^d \times (0,1))$ and (a) follows. Part (b) is proven by analogous argumentation.

Applying (a), we get for fixed $u \in (0,1)$.

$$\begin{aligned}\mathbb{E}^{\beta}_{(0,u)}[\sharp\mathscr{N}_{\mathbf{0}}^{<}] = \lambda^{<}_{(0,u)}(\mathbb{R}^d \times (0,u)) &= \int_0^u \mathrm{d}s \int_{\mathbb{R}^d} \mathrm{d}y\ \rho\Big(\beta^{-1} s^{\gamma} u^{\alpha} |y|^d\Big) \\ &= \tfrac{\beta}{1-\gamma} u^{1-\gamma-\alpha},\end{aligned}$$

using $\gamma < 1$ and the normalisation (1.3). If $1-\gamma-\alpha \geq 0$ this is bound from above by $\beta/(1-\gamma)$ proving the first part of (c). If instead $1-\gamma-\alpha < 0$, we have for each $k \in \mathbb{N}_0$ by independence of T_0 and $\mathscr{X}_0 \backslash \{\mathbf{0}\}$

$$\begin{aligned}\mathbb{P}^{\beta}\{\sharp\mathscr{N}_{\mathbf{0}}^{<} = k\} &= \int_0^1 \mathrm{d}u\ \mathbb{P}^{\beta}_{(0,u)}\{\sharp\mathscr{N}^{<}_{(0,u)} = k\} \\ &= \int_0^1 \mathrm{d}u\ \ \exp\Big(- \tfrac{\beta}{1-\gamma} u^{1-\gamma-\alpha}\Big) \frac{\big(\tfrac{\beta}{1-\gamma} u^{1-\gamma-\alpha}\big)^k}{k!} \\ &= \int_{\frac{\beta}{1-\gamma}}^{\infty} \mathrm{d}\lambda\ \Big(e^{-\lambda} \tfrac{\lambda^k}{k!}\Big)\Big(\tfrac{1}{\alpha+\gamma-1}\Big(\tfrac{\beta\lambda^{-(\alpha+\gamma)}}{1-\gamma}\Big)^{1/(\alpha+\gamma-1)}\Big)\end{aligned}$$

by a change of variables, proving the second part of (c).

We start the proof of (d) with the case $\alpha < 1$. Similar as above using (b), we get

$$\lambda^{>}_{(0,u)}(\mathbb{R}^d \times (u,1)) = \int_u^1 \mathrm{d}s\ \beta u^{-\gamma} s^{-\alpha} \asymp \tfrac{\beta}{1-\alpha}(u^{-\gamma} - 1).$$

Hence, for $k \in \mathbb{N}$, we have

$$\begin{aligned}\mathbb{P}^{\beta}\{\mathscr{N}_0^{>} = k\} &\asymp \int_0^1 \mathrm{d}u\ \ \exp\Big(- \tfrac{\beta}{1-\alpha}(u^{-\gamma} - 1)\Big) \frac{\Big(\tfrac{\beta}{1-\alpha}(u^{-\gamma}-1)\Big)^k}{k!} \\ &= \int_0^{\infty} \mathrm{d}\lambda\ \Big(e^{-\lambda}\tfrac{\lambda^k}{k!}\Big)\Big(\tfrac{1}{\gamma}\Big(\tfrac{\beta}{1-\alpha}\Big)^{1/\gamma}\Big(\lambda + \tfrac{\beta}{1-\alpha}\Big)^{-1-1/\gamma}\Big)\end{aligned}$$

as claimed. The proof of the second statement of (d) works completely analogous. $\square$

The previous theorem shows that the in- and the outdegree are independent and hence the degree of the root is the convolution of both. In the following

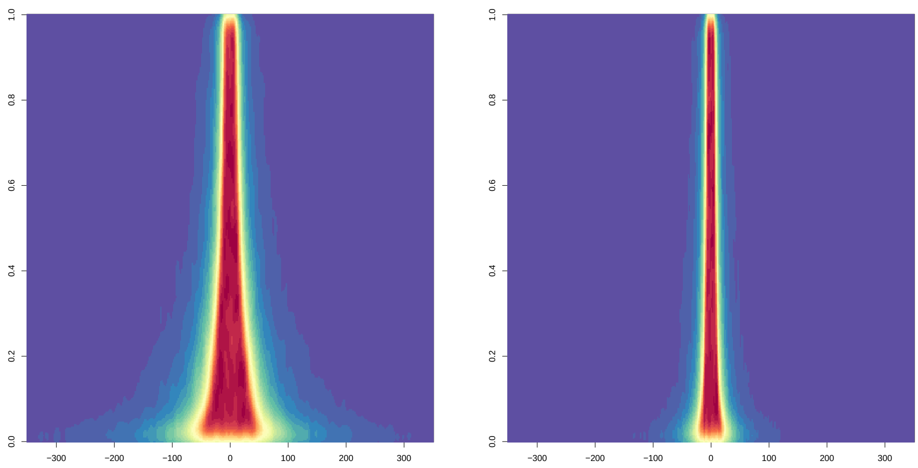

Figure 2.1.: Heatmaps of the neighbourhood of a relatively small mark root (left, mark 0.2) and of a relatively large mark root (right, mark 0.8) in $\mathcal{G}^\beta$ constructed with the preferential attachment kernel and $\beta = 5$, $\gamma = 1/3$ (hence $\alpha = 2/3$) and $\varphi(x) = 1 \wedge x^{-2}$; cf. [32, Fig. 2].

lemma, we show that the indegree is heavy-tailed distributed with index

$$\tau^{>} := 1 + \left(\tfrac{1}{\gamma} \wedge \tfrac{1}{(\alpha+\gamma-1)^+}\right)$$

where $(\alpha+\gamma-1)^+ := (\alpha+\gamma-1 \vee 0)$ and $1/0 := \infty$. We further show that the outdegree is heavy-tailed with index

$$\tau^{<} := 1 + \tfrac{1}{\alpha+\gamma-1}$$

in the mixed Poisson case $\alpha > 1-\gamma$. In the other case, the outdegree follows no power-law which we identify with the exponent $\tau^{<} = \infty$. Hence, the root's degree is heavy-tailed distributed with index $\tau = \tau^{<} \wedge \tau^{>}$ proving the claimed power-law exponents of Section 1.2.2. There are hence three phases depending on the value of α. If $\alpha \leq 1-\gamma$, the indegree is heavy-tailed and the outdegree is light-tailed. If $1-\gamma < \alpha < 1$ both are heavy-tailed but the indegree is of higher order and if $\alpha > 1$ both are of the same order.

Lemma 2.2. *Let $\mathcal{G}^\beta$ be the graph constructed with the interpolation-kernel $g_{\gamma,\alpha}$ with $\gamma \in (0,1)$ and $\alpha \in [0, 2-\gamma)$.*

(a) It holds

$$\mathbb{P}^\beta\{\sharp\mathscr{N}_0^> = k\} = k^{-\tau^> + o(1)}, \quad \text{as } k \uparrow \infty.$$

(b) If $\alpha > 1-\gamma$, it holds

$$\mathbb{P}^\beta\{\sharp\mathscr{N}_0^< = k\} = k^{-\tau^< + o(1)}, \quad \text{as } k \uparrow \infty.$$

Proof. Let $\alpha < 1$ and hence $\tau^> = 1 + {}^1\!/\!_\gamma$. By Proposition 2.1 (d), we have

$$\begin{aligned}\mathbb{P}^\beta\{\mathscr{N}_0^> = k\} &\asymp \int_0^\infty \mathrm{d}\lambda\, \Big(e^{-\lambda}\tfrac{\lambda^k}{k!}\Big)\Big(\tfrac{1}{\gamma}\big(\tfrac{\beta}{1-\alpha}\big)^{1/\gamma}\big(\lambda + \tfrac{\beta}{1-\alpha}\big)^{-1-1/\gamma}\Big)\\ &\asymp \frac{1}{\Gamma(k+1)}\int_0^\infty \mathrm{d}\lambda\, \lambda^{(k-1/\gamma)-1}e^{-\lambda}\\ &= \frac{\Gamma(k-1/\gamma)}{\Gamma(k+1)}\\ &= k^{-1-1/\gamma+o(1)},\end{aligned}$$

as $k \uparrow \infty$ by Stirling's formula. The remaining statements are proven analogously. □

Lemma 2.2 shows that the root vertex has finite expected degree. By the refined Campbell theorem [55, Theorem 9.1], almost surely, this holds true for every vertex in $\mathcal{G}^\beta$.

Corollary 2.3. *Let $\mathcal{G}^\beta$ be the graph constructed with the interpolation-kernel $g_{\gamma,\alpha}$ with $\gamma \in (0,1)$ and $\alpha \in [0, 2-\gamma)$. Then $\mathcal{G}^\beta$ is almost surely locally finite.*

In the lattice case, we do not have the Poisson structure of the neighbourhoods any longer. However, the expected number of incoming or outgoing

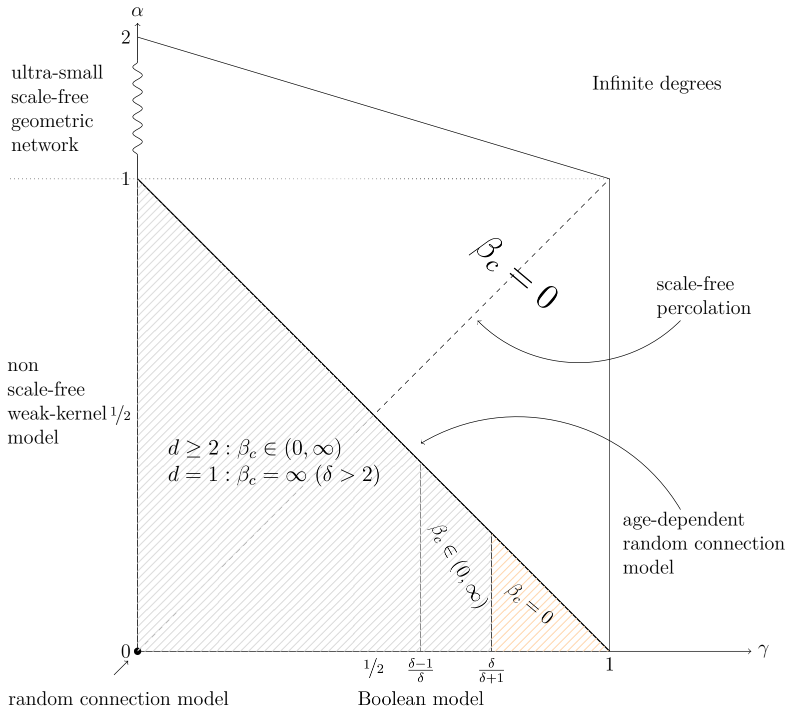

Figure 2.2.: Phase diagram for the interpolation-kernel depending on the values of γ and α. Dotted or dashed lines represent no change of behaviour. Shaded the main results of Proposition 2.4 (a) in orange (i) and grey (ii). The $d = 1$ specifics only occur for $\delta > 2$. In hard models ($\delta = \infty$) the two δ-dependent phases at the bottom right do not exist.

edges of the fixed root remains of the same order. Hence, one can apply the proof method of [70, Theorem 1.1], also outlined in [17] to obtain the same power-law exponents.

2.2.2. The percolation picture

The introduction of the interpolation-kernel allows us to study all models of Section 1.2.2 at once and to compare their behaviour. Together with the previous section we can further analyse how the observed behaviour is influenced by the degree distribution. In the next proposition, we summarise known results discussed in Section 1.2.2 and the main results of Section 1.2.3 to draw a complete picture of percolation in the weight-dependent random connection model with the interpolation kernel. Recall the definition of regularly varying profile-functions (1.17) and the identification of $\rho = \mathbb{1}_{[0,1/2]}$ with $\delta = \infty$. The Proposition is summarised in Figure 2.2.

Proposition 2.4. *Let $(\mathcal{G}^\beta : \beta > 0)$ be the weight-dependent random connection model constructed with a profile-function ρ fulfilling (1.17) for some $\delta \in (1, \infty]$ and the interpolation kernel $g_{\gamma,\alpha}$ (2.2) for $\gamma \in [0, 1)$ and $\alpha \in [0, 2 - \gamma)$.*

(a) For $\alpha \leq 1 - \gamma$, we have:

(i) If $\gamma > \frac{\delta}{\delta+1}$, then $\beta_c = 0$.

(ii) If $\gamma < \frac{\delta}{\delta+1}$ and

- *either $d \geq 2$ or $d = 1$ and $\delta < 2$ or $d = 1$ and $\gamma > 1 - 1/\delta$ or $d = 1$ and $\alpha = 1 - \gamma$ and $\gamma \geq 1/2$, then $\beta_c \in (0, \infty)$;*

- *$d = 1$ and $\alpha < 1 - \gamma$ and $\delta > 2$ and $\gamma < 1 - 1/\delta$, then $\beta_c = \infty$.*

(b) For $\alpha > 1 - \gamma$, we have $\beta_c = 0$.

The statements of (a) are direct consequences of Theorem 1.8 and Corollary 1.13 together with the monotonicity of $g_{\gamma,\alpha}$ in γ and α. To prove (b) it is important to note that in this case the outdegree is also heavy-tailed. Hence, the smaller a vertex's mark is the more neighbours with even smaller marks this vertex has. This suffices to adapt the proof [18, Theorem 3.2 (a1)] of Deprez and Wüthrich showing that one can always build an infinite self-avoiding path going directly from heavy to heavier vertices if starting with a sufficiently heavy vertex. In contrast we shall see that for light-tailed outdegree distributions it is a better strategy to connect two heavy vertices through a connector, any vertex with potentially large mark but connected to both of the heavy ones. However, this strategy only works for strong enough indegree distribution (determining the number of potential connectors a heavy vertex has) combined with sufficiently many long edges to overcome spatial restrictions. This idea is the core of the proof of Theorem 1.8. In the remaining Chapter this theorem as well as Theorem 1.11 and Corollary 1.13 are proven as outlined in more detail in Section 1.3.

2.3. FKG- and BK-inequality

In this section, we present two standard tools in percolation theory that are helpful dealing with correlated events; the so called *FKG-inequality* named after Fortuin, Kasteleyn, and Ginibre [29] and the *BK-inequality* named after van den Berg and Kesten [6]. We use variants of both from [40]. Let f be a function, defined on a point process on $\mathbb{R}^d \times (0,1)$ and an edge mark collection $\mathcal{U}_0$. We say such a function is *increasing* if it is non decreasing in the underlying point process η with respect to set inclusion, coordinate-wise non increasing with respect to vertex marks as well as coordinate-wise non increasing with respect to edge marks. We say an event E is increasing if $\mathbb{1}_E$ is increasing. Put differently, the probability of E increases if either additional vertices are added to the graph or if the vertex weights are increased (corresponding to decreasing the vertex marks) or additional edges are drawn (corresponding to decreasing the edge marks). If f_1 and f_2 are two such increasing functions, we have by repeating the arguments of [40, Equation (2.21)]

$$\begin{aligned}\mathbb{E}^\beta[f_1(\mathcal{X}_0,\mathcal{U}_0)f_2(\mathcal{X}_0,\mathcal{U})] &= \mathbb{E}^\beta\Big[\mathbb{E}^\beta[f_1(\mathcal{X}_0,\mathcal{U})f_2(\mathcal{X}_0,\mathcal{U}_0) \mid \eta]\Big] \\ &\geq \mathbb{E}^\beta\Big[\mathbb{E}^\beta[f_1(\mathcal{X}_0,\mathcal{U}_0) \mid \eta]\mathbb{E}^\beta[f_2(\mathcal{X}_0,\mathcal{U}_0) \mid \eta]\Big] \qquad (2.6) \\ &\geq \mathbb{E}^\beta[f_1(\mathcal{X}_0,\mathcal{U}_0)]\mathbb{E}^\beta[f_2(\mathcal{X}_0,\mathcal{U}_0)].\end{aligned}$$

This inequality is the FKG-inequality and it essentially says that the occurrence of an increasing event gives the positive information on the existence of sufficiently many vertices, large vertex weights and edges which favours the occurrence of another increasing event.

For the other inequality, consider again two increasing events on the vertices and edges of a graph but which are now meant to occur disjointly. Here, *disjoint occurrence* means that both events have to occur on disjoint subsets of the vertices and edges. Again, both events profit from the existence of many large weight vertices and edges on the one hand but they are also obstructing themselves as each vertex and edge can only be used by one of the events. Hence, it is naturally to assume that the probability of independent occurrence of the events in two independent copies of the graph is larger than the probability of disjoint occurrence in the same graph which is the statement of the BK-inequality. In full generality it is outlined in [40, Section 2.4]. For the sakes of this thesis it suffices to deal with the disjoint occurrence of certain paths. Let $\mathbf{z}_1, \mathbf{z}_2$ and $\mathbf{z}_3$ be given points and denote by $\{\mathbf{z}_1 \leftrightarrow \mathbf{z}_2\}$ the event that $\mathbf{z}_1$ is connected to $\mathbf{z}_2$ via a self-avoiding path. We denote by $\{\mathbf{z}_1 \leftrightarrow \mathbf{z}_2\} \circ \{\mathbf{z}_2 \leftrightarrow \mathbf{z}_3\}$ the disjoint occurrence of the two paths; that is $\mathbf{z}_1$ is connected by a path to $\mathbf{z}_2$ and $\mathbf{z}_2$ is connected by a path to $\mathbf{z}_3$ where both paths only share $\mathbf{z}_2$ as a common vertex. Then [40, Equation (2.20)] yields

$$\mathbb{P}^{\beta}_{\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3}(\{\mathbf{z}_1 \leftrightarrow \mathbf{z}_2\} \circ \{\mathbf{z}_2 \leftrightarrow \mathbf{z}_3\}) \leq \mathbb{P}^{\beta}_{\mathbf{z}_1,\mathbf{z}_2}\{\mathbf{z}_1 \leftrightarrow \mathbf{z}_2\}\mathbb{P}^{\beta}_{\mathbf{z}_2,\mathbf{z}_3}\{\mathbf{z}_2 \leftrightarrow \mathbf{z}_3\}.$$

This extends easily to the existence of paths with further restrictions such as required paths length or vertices from a certain range of locations or vertex marks. Denote by $\{\mathbf{z}_1 \overset{k}{\leftrightarrow} \mathbf{z}_2\}$ the event that $\mathbf{z}_1$ is connected to $\mathbf{z}_2$ by a path of length k. Here, the length of a path is defined as the number of the edges on the path. We immediately get again by [40]

$$\mathbb{P}^{\beta}_{\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3}(\{\mathbf{z}_1 \overset{k_1}{\leftrightarrow} \mathbf{z}_2\} \circ \{\mathbf{z}_2 \overset{k_2}{\leftrightarrow} \mathbf{z}_3\}) \leq \mathbb{P}^{\beta}_{\mathbf{z}_1,\mathbf{z}_2}\{\mathbf{z}_1 \overset{k_1}{\leftrightarrow} \mathbf{z}_2\}\mathbb{P}^{\beta}_{\mathbf{z}_2,\mathbf{z}_3}\{\mathbf{z}_2 \overset{k_2}{\leftrightarrow} \mathbf{z}_3\}.$$

Let us further restrict the paths to vertices with certain vertex marks. Let $E^k_{a_1,a_2}(\mathbf{z}_1, \mathbf{z}_2)$ be the event that $\mathbf{z}_1$ is connected by a path of length k to $\mathbf{z}_2$ where all vertices of the paths have marks in the interval (a_1, a_2) which can be written as

$$
\begin{aligned}
E^k_{a_1,a_2}(\mathbf{z}_1,\mathbf{z}_2) := \Big\{\exists \mathbf{y}_1,\dots,&\mathbf{y}_{k-1} \in \mathcal{X}_0 \cap \mathbb{R}^d \times (a_1,a_2) : \\
&\mathbf{z}_1 \sim \mathbf{y}_1 \sim \dots \sim \mathbf{y}_{k-1} \sim \mathbf{z}_2\Big\}
\end{aligned}
$$

Here, it is to note that this is still an increasing event with respect to decreasing the vertex marks. It can be seen as being defined on the point process $\mathcal{X}_0$ where all vertices with marks outside (a_1,a_2) have been removed. Now, if a vertex mark is decreased such that it leaves the interval this simply coincides with removing this vertex from the graph and the monotonicity goes in the right direction still. We hence get

$$
\begin{aligned}
&\mathbb{P}^\beta_{\mathbf{z}_1,\mathbf{z}_2,\mathbf{z}_3}\Big(E^{k_1}_{a_1,a_2}(\mathbf{z}_1,\mathbf{z}_2) \circ E^{k_2}_{b_1,b_2}(\mathbf{z}_2,\mathbf{z}_3)\Big) \\
&\qquad \leq \mathbb{P}^\beta_{\mathbf{z}_1,\mathbf{z}_2}\Big(E^{k_1}_{a,1,a_2}(\mathbf{z}_1,\mathbf{z}_2)\Big)\mathbb{P}^\beta_{\mathbf{z}_2,\mathbf{z}_3}\Big(E^{k_2}_{b_1,b_2}(\mathbf{z}_2,\mathbf{z}_3)\Big).
\end{aligned}
\tag{2.7}
$$

2.4. Existence of a subcritical phase

We fix $\delta > 1$ and $\gamma < \frac{\delta}{\delta+1}$. Since $g^{\mathrm{pa}} \leq g_{\gamma,\alpha}$ for all $\alpha \leq 1-\gamma$, we have

$$
\mathbb{P}^{\beta,\rho,g^{\mathrm{pa}}}\{\mathbf{0} \leftrightarrow \infty\} \geq \mathbb{P}^{\beta,\rho,g_{\gamma,\alpha}}\{\mathbf{0} \leftrightarrow \infty\}
$$

by a simple coupling argument. This includes in particular the strong-kernel $g^{\mathrm{str}} = g_{\gamma,0}$. Since, it further holds $g^{\mathrm{pa}} \leq 2^d g^{\mathrm{sum}}$, we have by the same argument

$$
\mathbb{P}^{\beta,\rho,g^{\mathrm{pa}}}\{\mathbf{0} \leftrightarrow \infty\} \geq \mathbb{P}^{\beta,\tilde{\rho},g^{\mathrm{sum}}}\{\mathbf{0} \leftrightarrow \infty\},
$$

where $\tilde{\rho}(x) = \frac{1}{2^d}\rho(2^d x)$. Thus, we focus on the preferential attachment kernel and show that we can choose a $\beta > 0$ such that $\theta(\beta) = 0$. Consequently, we work in the following exclusively in the age-dependent random connection model.

We use a *first moment method* approach for the number of paths of length n. We start with $\gamma < \frac{1}{2}$ and explicitly calculate the expected number of such paths. This turns out to be independent of the spatial geometry of the model and therefore cannot be used to prove the statement for $\frac{1}{2} \leq \gamma < \frac{\delta}{\delta+1}$. Recall the notation of $\mathbf{E}_0$ as the expectation of the Palm version of a unit intensity Poisson point process on $\mathbb{R}^d \times (0,1)$ and $\mathbb{P}^\beta_{\mathbf{v}_1,\dots,\mathbf{v}_n}$ as the law of $\mathcal{G}^\beta$ given $\mathbf{v}_1,\dots,\mathbf{v}_n$ are points of the vertex set. If the whole vertex set is given, we write $\mathbb{P}^\beta_{\mathcal{X}_0}$.

Lemma 2.5. *If* $0 < \gamma < \frac{1}{2}$, *then* $\theta(\beta) = 0$ *for all* $\beta < \frac{1-2\gamma}{4}$ *or, equivalently,* $\beta_c \geq \frac{1-2\gamma}{4}$.

Proof. We set $\mathbf{0} = \mathbf{z}_0 = (0, t_0)$ and get

$$\begin{aligned}\theta(\beta) &= \lim_{n\to\infty} \mathbb{P}^\beta\{\exists \text{ a path of length } n \text{ starting in } \mathbf{z}_0\} \\ &\leq \lim_{n\to\infty} \int_0^1 \mathrm{d}t_0 \; \mathbf{E}_{(0,t_0)}\Bigg[\sum_{\substack{\mathbf{z}_1,\dots,\mathbf{z}_n\in\mathcal{X} \\ \mathbf{z}_i\neq\mathbf{z}_j \forall i\neq j}} \mathbb{P}^\beta_{\mathcal{X}\cup\{(0,t_0)\}}\Big(\bigcap_{j=1}^n \{\mathbf{z}_j \sim \mathbf{z}_{j-1}\}\Big)\Bigg].\end{aligned}$$

The inner probability is a measurable function of the Poisson process and the points $\mathbf{z}_1,\dots,\mathbf{z}_n$ and by Mecke's equation [55, Theorem 4.4] we get, with $\mathcal{X}^*$ denoting an independent copy of $\mathcal{X}$,

$$\begin{aligned}&\int_0^1 \mathrm{d}t_0 \int_{(\mathbb{R}^d\times(0,1])^n} \bigotimes_{j=1}^n \mathrm{d}\mathbf{z}_j \; \mathbf{E}_{(0,t_0)}\Bigg[\mathbb{P}^\beta_{\mathcal{X}^*\cup\{(0,t_0),\mathbf{z}_1,\dots,\mathbf{z}_n\}}\Bigg(\bigcap_{j=1}^n \{\mathbf{z}_{j-1} \sim \mathbf{z}_j\}\Bigg)\Bigg] \\ &= \int_0^1 \mathrm{d}t_0 \int_{(\mathbb{R}^d\times(0,1])^n} \bigotimes_{j=1}^n \mathrm{d}\mathbf{z}_j \; \mathbb{P}^\beta_{\mathbf{z}_0,\dots,\mathbf{z}_n}\Bigg(\bigcap_{j=1}^n \{\mathbf{z}_{j-1} \sim \mathbf{z}_j\}\Bigg).\end{aligned}$$

Given the vertices, edges are drawn independently so we get by writing $\mathbf{z}_j = (z_j, s_j)$ for all $j \in \{1,\dots,n\}$ that the previous expression equals

$$\begin{aligned}&\int_0^1 \mathrm{d}s_0 \int_{(\mathbb{R}^d\times(0,1])^n} \bigotimes_{j=1}^n \mathrm{d}(z_j,s_j)\Bigg(\prod_{j=1}^n \rho\Big(g^{\mathrm{pa}}(s_{j-1},s_j)|z_j - z_{j-1}|^d\Big)\Bigg) \\ &= \beta^n \int_0^1 \mathrm{d}s_0 \int_0^1 \mathrm{d}s_1 \cdots \int_0^1 \mathrm{d}s_n \Bigg(\prod_{j=1}^n (s_j \wedge s_{j-1})^{-\gamma}(s_j \vee s_{j-1})^{\gamma-1}\Bigg),\end{aligned}$$

where we used the normalization condition (1.7). Since $\gamma < \frac{1}{2}$, Lemma 17 of [45] states that

$$\int_0^1 \mathrm{d}s_0 \int_0^1 \mathrm{d}s_1 \cdots \int_0^1 \mathrm{d}s_n \Bigg(\prod_{j=1}^n (s_j \wedge s_{j-1})^{-\gamma} (s_j \vee s_{j-1})^{\gamma-1} \Bigg)$$
$$\leq \left(\frac{1}{1+\alpha-\gamma} - \frac{1}{\alpha+\gamma} \right)^n,$$

for $\alpha \in (\gamma - 1, -\gamma)$. The minimum of the right-hand side over this non-empty interval equals $\frac{4}{1-2\gamma}$ and thus, setting $\beta < \frac{1-2\gamma}{4}$ we achieve

$$\theta(\beta) \leq \lim_{n\to\infty} \left(\tfrac{4\beta}{1-2\gamma}\right)^n = 0.$$

□

2.4.1. Existence of a subcritical phase: Case $\gamma \geq 1/2$.

We now turn to the more interesting case when $\gamma \in [\frac{1}{2}, \frac{\delta}{\delta+1})$ where we have to use the spatial properties of our model in order to prove our claim. Intuitively, as "powerful" vertices are typically far apart from each other, in order to create an infinite path in this spatial network one has to use long edges often enough to reach them. Therefore, where the long edges are used is the crucial and most interesting part of a path. On the other hand $\mathcal{G}^\beta$ is locally dense. Therefore, considering paths that stay for a long time in a neighbourhood of a vertex before using long edges greatly increases the number of possible paths we can construct. For $\gamma < 1/2$, the degrees of typical vertices are small enough so that the number of possible paths does not increase too much. This is not true any longer for $\gamma > \frac{1}{2}$ where the degree distribution has an infinite second moment. Thus, it becomes difficult to bound the probability of the existence of an arbitrary path of length n. In order to prove the existence of a subcritical phase, we start by explaining how to limit our counting to paths that are not stuck in local clusters. Then, we define what we call the *skeleton* of a path, which will help with counting the valid paths. As we will see, the skeleton is a collection of key vertices from a path ordered in a specific vertex mark structure. In the end, we will use these paths to complete the proof of Theorem 1.8(a).

Shortcut-free paths Let $P = (v_0, v_1, v_2, \dots)$ be a path in some graph G. We say (v_i, v_j) is a *shortcut* in P if $j > i + 1$ and v_i and v_j are connected by an edge in G. If P does not contain any shortcut, we say P is *shortcut-free*. If G is locally finite, i.e. all vertices of G are of finite degree, then there exists an infinite path if and only if there exists one that is also shortcut-free. To see how an infinite path $P = (v_0, v_1, v_2 \dots)$ in G can be made shortcut-free define $i_0 = \max\{i \geq 1 : v_i \sim v_0\}$. If $i_0 = 1$, then v_1 is the only neighbour v_0 has in P. If $i_0 \geq 2$, then (v_0, v_{i_0}) is a shortcut in P so we remove the vertices $v_1, \dots, v_{i_0 - 1}$ from P. We have thus removed all shortcuts starting from v_0 and since $v_0 \sim v_{i_0}$ the new P is still a path. We define analogously $i_k = \max\{i > i_{k-1} : v_i \sim v_{i_{k-1}}\}$ for every $k \geq 1$ and remove the intermediate vertices as needed. The resulting path $(v_0, v_{i_0}, v_{i_1}, \dots)$ is then still infinite but also shortcut-free.

Skeleton of a path Let $P = ((v_0, s_0), (v_1, s_1), \dots, (v_n, s_n))$ be a path of length n in some graph G where every vertex v_i carries a distinct vertex mark s_i. Then, precisely one of the vertices in P has the smallest mark; let $k_{\min} = \{k \in \{0, \dots, n\} : s_k < s_j, \ \forall j \neq k\}$ be its index. Starting from (v_0, s_0), we now choose the first vertex of the path that has a mark smaller than s_0 and call it (v_{i_1}, s_{i_1}). Continuing from this vertex, we choose the next vertex of the path that carries a smaller mark still, call it (v_{i_2}, s_{i_2}) and continue analogously until we reach the vertex with the smallest mark $(v_{k_{\min}}, s_{k_{\min}})$. We then repeat the same procedure starting from the end vertex (v_n, s_n) and going backwards across the indices. The union of the two subset of vertices is what we call the *skeleton* of the path P. More precisely, for every path $P = ((v_0, s_0), \dots, (v_n, s_n))$, there exists unique $0 \leq k \leq n$ and $k \leq m \leq n$ as well as a set of indices $\{i_0, i_1, \dots, i_{k-1}, i_k, i_{k+1}, \dots, i_m\}$

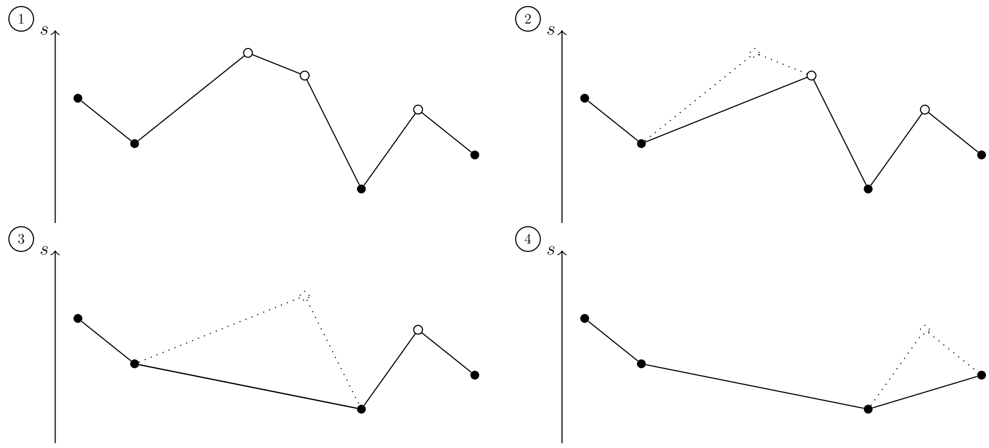

Figure 2.3.: A path where a vertex's mark is denoted on the s-axis. The vertices of the skeleton are in black. We successively remove all local maxima, starting with the largest mark vertex, and replace them by direct edges until the path, only containing the skeleton vertices, is left.

such that

$$\begin{aligned}
&i_0 = 0, i_k = k_{\min}, \text{ and } i_m = n \text{ as well as}\\
&s_{i_{\ell-1}} > s_{i_\ell} \text{ and } s_i > s_{i_{\ell-1}}, \ \forall i_{\ell-1} < i < i_\ell, \text{ for } \ell = 1, \dots, k \text{ and}\\
&s_{i_{\ell-1}} < s_{i_\ell} \text{ and } s_i > s_{i_\ell}, \ \forall i_{\ell-1} < i < i_\ell, \text{ for } \ell = k+1, \dots m.
\end{aligned}$$

The *skeleton of* P is then given by $((v_{i_j}, s_{i_j}))_{j=0,\dots,m}$. We say it is of length m and has its minimum at k.

We now give an alternative construction of the skeleton of P, which we call the *local maxima construction*. A vertex $(v_i, s_i) \in P\backslash\{(v_0, s_0), (v_n, s_n)\}$ is called a *local maximum* if $s_i > s_{i-1}$ and $s_i > s_{i+1}$. We successively remove all local maxima from P as follows: First, take the local maximum in P with the greatest vertex mark, remove it from P and connect its former neighbours by a direct edge. In the resulting path, we take the local maximum of greatest vertex mark and remove it, repeating until there is no local maximum left, see Figure 2.3. Therefore, the final path is decreasing in the marks of its vertices until the minimum mark vertex is reached, and only increasing in vertex marks afterwards. Hence, it is the uniquely determined skeleton of the path. Note that the skeleton is not necessarily an actual path of the graph. Actually, the skeleton of a shortcut-free path is not itself a path unless the path is its own skeleton.

Graph surgery In order to bound the probability of existence of an infinite self-avoiding and shortcut-free path in $\mathcal{G}^\beta$ starting in the origin we increase the number of short edges in $\mathcal{G}^\beta$, which then allows us to make better use of the shortcut-free condition. We choose $\varepsilon > 0$ such that

$$\tilde{\delta} := \delta - \varepsilon > \frac{\gamma}{1-\gamma}.$$

This is equivalent to $\gamma < \frac{\tilde{\delta}}{\tilde{\delta}+1}$. As ρ is regularly varying and bounded there exists $A > 1$ such that

$$\rho(x) \leq A x^{-\tilde{\delta}} \quad \text{for all } x > 0,$$

by the Potter bound [7, Theorem 1.5.6]. We define

$$\tilde{\rho}(x) = \mathbb{1}_{[0, A^{1/\tilde{\delta}}]}(x) + A x^{-\tilde{\delta}} \mathbb{1}_{(A^{1/\tilde{\delta}}, \infty)}(x).$$

We now choose $\tilde{\rho}$ as a profile function together with the preferential attachment kernel (1.15) and construct $\mathcal{G}^{\beta,\tilde{\rho},g^{\text{pa}}}(\mathcal{X}_0, \mathcal{U}_0)$. Put differently, we connect two given vertices (x,t) and (y,s) with probability

$$\begin{cases} 1, & \text{if } |x-y|^d \leq A^{\frac{1}{\tilde{\delta}}} g^{\text{pa}}(t,s)^{-1} \\ A\left(g^{\text{pa}}(t,s)|x-y|^d\right)^{-\tilde{\delta}}, & \text{otherwise.} \end{cases}$$

Note that in general $\tilde{\rho}$ does not satisfy the normalization condition (1.7). However, $\tilde{\rho}$ is still integrable and therefore the resulting graph $\mathcal{G}^{\beta,\tilde{\rho},g^{\text{pa}}}$ is still locally finite with unchanged power law and shows qualitatively the same behaviour. Since $\rho \leq \tilde{\rho}$, it follows by a simple coupling argument that

$$\theta(\beta) \leq \mathbb{P}^{\beta,\tilde{\rho},g^{\text{pa}}}\{\mathbf{0} \leftrightarrow \infty\}.$$

Due to the above it is no loss of generality to consider the graph $\mathscr{G}^\beta = \mathscr{G}^{\beta,\rho,g^{\mathrm{pa}}}$ where the profile function ρ is of the form

$$\rho(x) = 1 \wedge (Ax^{-\delta}),$$

which is what we do from now on. Note that we can no longer assume that (1.7) holds, instead we have

$$I_\rho := \int_{\mathbb{R}^d} \rho(|x|^d)\,\mathrm{d}x = A^{1/\delta}\Big(J(d)\tfrac{\delta}{d(\delta-1)}\Big)$$

where $J(d) = \prod_{j=0}^{d-2}\int_0^\pi \sin^j(\alpha_j)\mathrm{d}\alpha_j$ is the Jacobian of the d-dimensional sphere coordinates. We look at the probability that a shortcut-free path $P = ((z_1,s_1),(z_2,s_2),\dots)$ exists in $\mathscr{G}^\beta$. By choice of ρ, such a path satisfies

$$|z_i - z_j|^d > A^{\frac{1}{\delta}} g^{\mathrm{pa}}(s_i,s_j)^{-1}, \qquad \text{for all } |i-j| \geq 2.$$

Strategy of the proof To build a long path, one needs to use vertices with a small mark as those are the vertices with large degrees in our model. Every path is divided into a skeleton, which encodes how it moves to vertices with smaller and smaller marks, and subpaths connecting consecutive points of the skeleton by any number of vertices with larger marks, which we call *connectors*. We encode a characteristic feature of such a subpath by an unlabelled binary tree using the local maxima construction. We show that whenever $\gamma < \delta/(\delta+1)$ the expected number of shortcut-free subpaths with a given tree of size k is bounded by $(K\beta I_\rho)^k$ times the probability that the two extremal vertices are connected by an edge, for some constant $K > 1$. Combining this estimate with the BK-inequality allows us to bound the probability of existence of a path with a given skeleton in terms of the probability that this skeleton is a path. The probability of existence of paths of the latter type can be estimated by a first moment bound. We therefore obtain that the probability of existence of a shortcut-free path of length n starting in $\mathbf{0}$ is bounded from above by $(K\beta I_\rho)^n$ and hence

$$\theta(\beta) \leq \lim_{n\to\infty} (K\beta I_\rho)^n = 0$$

for $0 < \beta < 1/(KI_\rho)$.

Connecting two powerful vertices Let P be a path of length k that can be reduced to a skeleton with two vertices $\mathbf{x}$ and $\mathbf{y}$. Let $\mathbf{y}_0, \dots, \mathbf{y}_k$ be the vertices of P, ordered by their marks from smallest to largest. We assume without loss of generality that $\mathbf{x}$ carries a larger mark than $\mathbf{y}$ and therefore $\mathbf{x} = \mathbf{y}_1$ and $\mathbf{y} = \mathbf{y}_0$. We denote by $\mathscr{T}_{k-1}$ the set of all binary trees[1] with fixed vertex set $\{\mathbf{y}_2, \dots, \mathbf{y}_k\}$ such that every child carries a larger mark than its parent. With the path P we associate a tree in $\mathscr{T}_{k-1}$ as follows, see Figure 2.4.

Step one: $\mathbf{y}_2$ is the root of the tree.

Step two: Suppose the tree with vertices $\mathbf{y}_2, \dots, \mathbf{y}_{i-1}$ is constructed. Attach $\mathbf{y}_i$ as a new leaf of the tree. To find the place to attach the leaf start at the root and branch at every vertex to the left if the path P visits $\mathbf{y}_i$ before the vertex and to the right otherwise. If this means going to a place where there is no vertex, we attach $\mathbf{y}_i$ there. We continue like this until all $\mathbf{y}_2, \dots, \mathbf{y}_k$ are attached.

Next, we explain how to construct a path P connecting $\mathbf{x}$ and $\mathbf{y}$ when $T \in \mathscr{T}_{k-1}$ is given, see Figure 2.5. Here, given a path $(v_i)_{i=1}^n$ and any subpath (v_{j-1}, v_j, v_{j+1}), we call v_{j-1} the *preceding vertex* of v_j and v_{j+1} the *subsequent vertex* of v_j. We explore T using depth-first search and add the vertex currently being explored to the path. Let $P = (\mathbf{x}, \mathbf{y})$ and let $\mathbf{u}$ be the root of T. We define $L = (\mathbf{u})$ to be the list of vertices to be explored next (in the order as they are in L). We proceed as follows.

[1]Here, a binary tree is a rooted tree in which every vertex can have either (i) no child, (ii) a left child (iii) a right child, or (iv) a left and a right child.

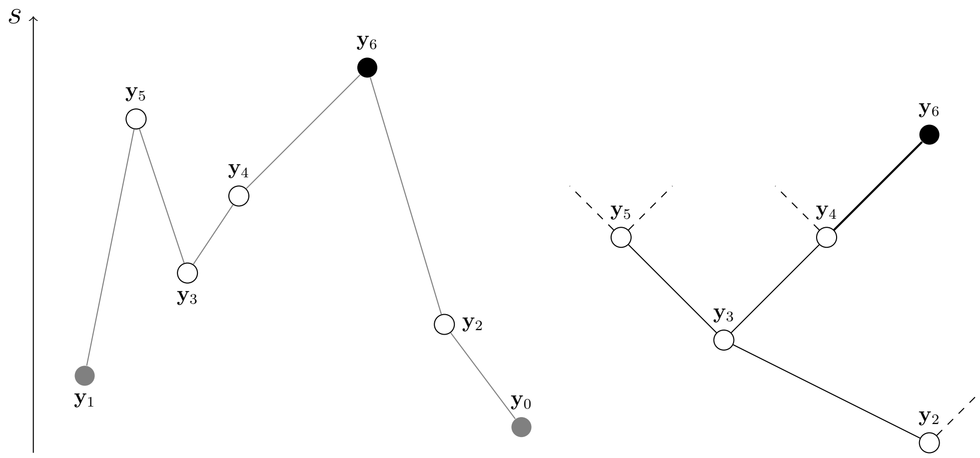

Figure 2.4.: On the left the path P where the s-axis denotes the vertices' marks. The vertices $\mathbf{y}_1$ and $\mathbf{y}_0$, which will not appear in the tree, are in grey. We insert the vertex $\mathbf{y}_6$ at the end of the branch that goes left at $\mathbf{y}_2$, right at $\mathbf{y}_3$, and right at $\mathbf{y}_4$.

Step one: We insert $\mathbf{u}$ into P as a local maximum between $\mathbf{x}, \mathbf{y}$. As a result $P = (\mathbf{x}, \mathbf{u}, \mathbf{y})$. We remove $\mathbf{u}$ from L and if $\mathbf{u}$ has children in T, we add them to L, ordered from left to right.

Step two: While L is not empty, we do the following:

1. We take the first vertex in L, denote it by $\mathbf{v}$ and remove it from L.

2. If $\mathbf{v}$ has children in T, we insert them at the beginning of L, ordered from left to right. Having done that, we consider $\mathbf{v}$ explored.

3. Let $\mathbf{w}$ be the parent of $\mathbf{v}$ in T and $\{\mathbf{z}_1, \mathbf{w}\}$, $\{\mathbf{w}, \mathbf{z}_2\}$ its incident edges in P, where $\mathbf{z}_1$ is the preceding vertex of $\mathbf{w}$ in P and $\mathbf{z}_2$ the subsequent one. If $\mathbf{v}$ is the left child of $\mathbf{w}$, we insert $\mathbf{v}$ as a local maximum between $\mathbf{z}_1$ and $\mathbf{w}$ in P by adding it to the path and replacing the edge $\{\mathbf{z}_1, \mathbf{w}\}$ in P by the two edges $\{\mathbf{z}_1, \mathbf{v}\}$ and $\{\mathbf{v}, \mathbf{w}\}$. If $\mathbf{v}$ is a right child, we insert $\mathbf{v}$ as a local maximum between $\mathbf{w}$ and $\mathbf{z}_2$ in an analogous way.

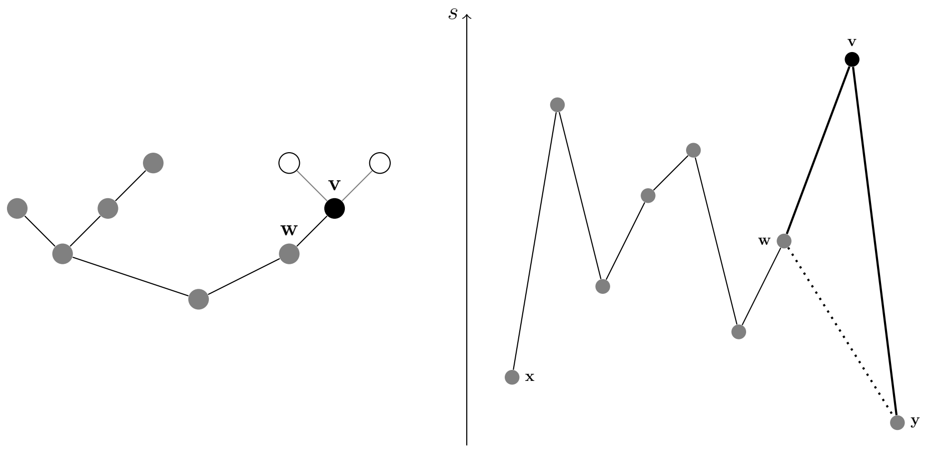

Figure 2.5.: On the left the binary tree T. The grey vertices are already explored by depth-first search. The black vertex $\mathbf{v}$ is the vertex currently being explored. The white vertices have not been discovered yet. On the right, the path P corresponding to the already explored tree. The s-axis denotes the vertices' birth times. Start and end vertex, $\mathbf{x}$ and $\mathbf{y}$, do not appear in the tree. Since $\mathbf{v}$ is the right child of $\mathbf{w}$, we insert $\mathbf{v}$ as a local maximum between $\mathbf{w}$ and $\mathbf{y}$ in the path P.

It is clear that for given $\mathbf{y}_0, \dots, \mathbf{y}_k$ the two procedures establish a bijection between the paths with vertices $\mathbf{y}_0, \dots, \mathbf{y}_k$ that can be reduced to a skeleton with two vertices $\mathbf{y}_0$ and $\mathbf{y}_1$ on the one hand, and the trees $T \in \mathscr{T}_{k-1}$ on the other hand. Removing the labels from a tree in $\mathscr{T}_k$ yields a binary tree, which encodes important structural information about the path.

The following lemma shows that, if $\gamma < \delta/(\delta+1)$, the probability of two vertices being connected through a single connector is bounded by a small multiple of the probability that there exists a direct edge between them.

For two given vertices $\mathbf{x}$ and $\mathbf{y}$, we denote by $\{\mathbf{x} \overset{2}{\underset{\mathbf{x},\mathbf{y}}{\longleftrightarrow}} \mathbf{y}\}$ the event that $\mathbf{x}$ and $\mathbf{y}$ are connected by a path of length two where the connector carries a larger mark than both of them.

Lemma 2.6. *Let $\gamma \in (0, \frac{\delta}{\delta+1})$. Let $\mathbf{x} = (x,t)$ and $\mathbf{y} = (y,s)$ be two given vertices satisfying $|x-y|^d \geq A^{1/\delta} g^{pa}(t,s)^{-1}$. Then*

$$\mathbb{P}^\beta_{\mathbf{x},\mathbf{y}}\{\mathbf{x} \overset{2}{\underset{\mathbf{x},\mathbf{y}}{\longleftrightarrow}} \mathbf{y}\} \leq \int_{\mathbb{R}^d \times ((t\vee s),1]} \mathrm{d}\mathbf{z}\, \mathbb{P}^\beta_{\mathbf{x},\mathbf{z}}\{\mathbf{x} \sim \mathbf{z}\}\mathbb{P}^\beta_{\mathbf{y},\mathbf{z}}\{\mathbf{z} \sim \mathbf{y}\} \leq \beta C_1 \mathbb{P}^\beta_{\mathbf{x},\mathbf{y}}\{\mathbf{x} \sim \mathbf{y}\},$$

where $C_1 = \frac{2^{d\delta+1} I_\rho}{\delta(1-\gamma)-\gamma}$.

Proof. Without loss of generality let $t > s$ in which case $g^{\mathrm{pa}}(t,s) = s^\gamma t^{1-\gamma}$. Recall that $\{\mathbf{x} \xleftrightarrow[\mathbf{x},\mathbf{y}]{2} \mathbf{y}\}$ is the event that $\mathbf{x}$ and $\mathbf{y}$ share a common neighbour that carries a larger mark than both of them. Such neighbours form a Poisson point process on $\mathbb{R}^d \times (t,1]$ with intensity measure

$$\rho(\beta^{-1} t^\gamma u^{1-\gamma} |x-z|^d) \rho(\beta^{-1} s^\gamma u^{1-\gamma} |z-y|^d) \, \mathrm{d}z \, \mathrm{d}u,$$

cf. the arguments of Section 2.2.1, from which the first inequality follows. For the second inequality, we have

$$\begin{aligned}
&\int_t^1 \mathrm{d}u \int_{\mathbb{R}^d} \mathrm{d}z \; \rho(\beta^{-1} t^\gamma u^{1-\gamma} |x-z|^d) \rho(\beta^{-1} s^\gamma u^{1-\gamma} |z-y|^d) \\
&\leq \int_t^1 \mathrm{d}u \Big[\int_{\mathbb{R}^d} \mathrm{d}z \; \rho(\beta^{-1} t^\gamma u^{1-\gamma} |x-z|^d) \rho\left((2^d\beta)^{-1} s^\gamma u^{1-\gamma} |x-y|^d \right) \\
&\qquad + \int_{\mathbb{R}^d} \mathrm{d}z \; \rho\left((2^d\beta)^{-1} t^\gamma u^{1-\gamma} |x-y|^d \right) \rho(\beta^{-1} s^\gamma u^{1-\gamma} |z-y|^d) \Big].
\end{aligned}$$

Here, the inequality holds as for all $z \in \mathbb{R}^d$ either $|x-z| \geq \frac{1}{2}|x-y|$ or $|y-z| \geq \frac{1}{2}|x-y|$, and ρ is non-increasing. For the first integral, a change of variables leads to

$$\int_t^1 \mathrm{d}u \; \beta t^{-\gamma} u^{\gamma-1} \rho\left((2^d\beta)^{-1} s^\gamma u^{1-\gamma} |x-y|^d \right) I_\rho.$$

As $\rho(x) = 1 \wedge (Ax^{-\delta})$ this can be further bound by

$$\begin{aligned}
A 2^{d\delta} \beta^{1+\delta} I_\rho \int_t^1 \mathrm{d}u \; & s^{-\gamma\delta} t^{-\gamma} |x-y|^{-d\delta} u^{-\delta(1-\gamma)+\gamma-1} \\
&\leq A\beta^{\delta+1} \frac{2^{d\delta} I_\rho}{\delta(1-\gamma)-\gamma} (s^\gamma t^{1-\gamma} |x-y|^d)^{-\delta}
\end{aligned}$$

using that $\gamma < \delta/(\delta+1)$. A similar calculation for the second integral yields the same bound and as $|x-y|^d > A^{1/\delta} \beta s^{-\gamma} t^{\gamma-1}$ implies

$$A(\beta^{-1} s^\gamma t^{1-\gamma} |x-y|^d)^{-\delta} \leq 1,$$

and therefore

$$\mathbb{P}^{\beta}_{\mathbf{x},\mathbf{y}}\{\mathbf{x} \sim \mathbf{y}\} = A\Big(\beta^{-1}s^{\gamma}t^{1-\gamma}|x-y|^d\Big)^{-\delta},$$

which proves the claim. $\square$

We now extend this result to bound the probability that the two given vertices $\mathbf{x}$ and $\mathbf{y}$ are connected through $k-1$ connectors. That is, $\mathbf{x}$ and $\mathbf{y}$ are connected by a path of length k and $\mathbf{x}$ and $\mathbf{y}$ are the two vertices with smallest marks within the path. We denote this event by $\{\mathbf{x} \xleftrightarrow[\mathbf{x},\mathbf{y}]{k} \mathbf{y}\}$.

Lemma 2.7. *Let $\gamma \in (0, \frac{\delta}{\delta+1})$ and $\mathbf{x} = (x,t), \mathbf{y} = (y,s)$ be two Poisson points satisfying $|x-y|^d > A^{1/\delta} g^{pa}(t,s)^{-1}$. Then, for all $k \in \mathbb{N}$, we have*

$$\mathbb{P}^{\beta}_{\mathbf{x},\mathbf{y}}\{\mathbf{x} \xleftrightarrow[\mathbf{x},\mathbf{y}]{k} \mathbf{y}\} \le (\beta C_2)^{k-1}\mathbb{P}^{\beta}_{\mathbf{x},\mathbf{y}}\{\mathbf{x} \sim \mathbf{y}\}, \tag{2.8}$$

where $C_2 = \frac{2^{d\delta+3}I_\rho}{\delta(1-\gamma)-\gamma}$.

Proof. For $k=1$ there is nothing to show, so we assume $k \ge 2$. If T is an unlabelled binary tree with $k-1$ vertices we denote by $X(T)$ the number of paths connecting $\mathbf{x}$ and $\mathbf{y}$ through $k-1$ connectors, which are associated with a labelling of T. Taking the union over all (unlabelled) binary trees on $k-1$ vertices we get

$$\mathbb{P}^{\beta}_{\mathbf{x},\mathbf{y}}\{\mathbf{x} \xleftrightarrow[\mathbf{x},\mathbf{y}]{k} \mathbf{y}\} \le \sum_{\substack{T \text{ binary tree} \\ \text{on } k-1 \text{ vertices}}} \mathbb{E}^{\beta}_{\mathbf{x},\mathbf{y}}\Big[X(T)\Big],$$

and as the number of binary trees on $k-1$ vertices is bounded from above by[2] 4^{k-1} it suffices to show

$$\mathbb{E}^{\beta}_{\mathbf{x},\mathbf{y}}\Big[X(T)\Big] \le C_1^{k-1}\mathbb{P}^{\beta}_{\mathbf{x},\mathbf{y}}\{\mathbf{x} \sim \mathbf{y}\},$$

[2] The number of binary rooted trees of size n is given by the Catalan numbers $(2n)!/(n!(n+1)!)$.

for all binary trees T with $k-1$ vertices where C_1 is the constant of Lemma 2.6. We show this by induction on k starting with the case $k=2$, when T consist of just the root, which is shown in Lemma 2.6. For the induction step we fix an unlabelled binary tree T with $k-1$ vertices and insert a new leaf. Denote the new tree with k vertices by T'. We identify the two vertices in the tree, which correspond to the preceding and subsequent vertex of the new leaf in any path associated with T' as follows:

- If the new leaf is a left child, its subsequent vertex in the path is its parent, and its preceding vertex is determined by following its ancestral line backwards along the tree until we find a vertex which has a right child on the ancestral line. If there is no such vertex its preceding vertex is $\mathbf{x}$.
- If the new leaf is a right child, its preceding vertex in the path is its parent, and its subsequent vertex is determined by following its ancestral line backwards along the tree until we find a vertex which has a left child on the ancestral line. If there is no such vertex its subsequent vertex is $\mathbf{y}$.

From the construction of the tree we make the following two observations if a path is associated with T',

(i) the new leaf carries a larger mark than its parent and the path contains two sequential edges, one connecting the preceding vertex to the new leaf, and one connecting the new leaf to its subsequent vertex,

(ii) if the preceding and subsequent vertex of the new leaf are connected by an edge, then the path using that edge instead of the the two edges adjacent to the new leaf is associated with T.

We call a labelling of T by points of the Poisson process *almost complete* if it becomes the labelling associated with a path when the preceding and subsequent vertex of the new leaf are connected by an edge. Hence (ii) can be restated saying that the labelling of T obtained by association of a path with T' is almost complete.

Denoting the labels of the preceding and subsequent vertices of the new leaf by $\mathbf{z}_\ell = (z_\ell, s_\ell)$ resp. $\mathbf{z}_r = (z_r, s_r)$ we get using (i) that

$$\begin{aligned}
&\mathbb{E}^\beta_{\mathbf{x},\mathbf{y}}\Big[X(T')\Big] \\
&= \mathbb{E}^\beta_{\mathbf{x},\mathbf{y}}\Big[\sharp\Big\{\text{paths } \mathbf{x} \leftrightarrow \mathbf{z}_\ell \sim \mathbf{z_{new}} \sim \mathbf{z}_r \leftrightarrow \mathbf{y} \text{ associated with } T'\Big\}\Big] \\
&\leq \mathbb{E}^\beta_{\mathbf{x},\mathbf{y}} \sum_{\substack{\text{almost complete} \\ \text{labellings of } T}} \int_{s_\ell \vee s_r}^1 \mathrm{d}u \int_{\mathbb{R}^d} \mathrm{d}z\, \mathbb{P}^\beta_{\mathbf{z}_\ell,(z,u)}\{\mathbf{z}_\ell \sim (z,u)\} \mathbb{P}^\beta_{(z,u),\mathbf{z}_r}\{(z,u) \sim \mathbf{z}_r\}.
\end{aligned}$$

As the paths associated to T' are shortcut-free we have

$$|z_\ell - z_r|^d > A^{1/\delta} \beta g^{\mathrm{pa}}(s_\ell, s_r)^{-1}$$

and hence Lemma 2.6 ensures that this is bounded by

$$\begin{aligned}
\beta C_1 \mathbb{E}^\beta_{\mathbf{x},\mathbf{y}}\Bigg[\sum_{\substack{\text{almost complete} \\ \text{labelling of } T}} \mathbb{P}^\beta_{\mathbf{z}_\ell,\mathbf{z}_r}\{\mathbf{z}_\ell \sim \mathbf{z}_r\}\Bigg] &\leq \beta C_1\, \mathbb{E}^\beta_{\mathbf{x},\mathbf{y}} X(T) \\
&\leq (\beta C_1)^k \mathbb{P}^\beta_{\mathbf{x},\mathbf{y}}\{\mathbf{x} \sim \mathbf{y}\},
\end{aligned}$$

using (ii) and the induction hypothesis. □

Proof of the subcritical phase We now use the results of the previous paragraphs to bound the probability of a shortcut-free path of length n existing by some exponential, thus showing Theorem 1.8(a). To this end, we have to distinguish between regular and irregular paths. Let $S = (\mathbf{z}_0, \mathbf{z}_1, \ldots, \mathbf{z}_m)$ be a skeleton of length m. We say S is *regular* if its vertex with smallest mark has a mark larger than 2^{-m}. We say S is *irregular* if it is not regular. Similarly, we say a path P of finite length is regular if its underlying skeleton is regular and conversely, P is irregular if its skeleton is irregular. Finally, let $P = (\mathbf{v}_0, \mathbf{v}_1, \ldots)$ be an infinite path. We say P is *irregular* if for all $k \in \mathbb{N}$ there exists $n \geq k$ such that the path $(\mathbf{v}_0, \dots, \mathbf{v}_n)$ (of length n) is irregular. An infinite path P is *regular* if it is not irregular. In other words, an infinite path is irregular if it has irregular subpaths of arbitrarily large lengths. We first show that almost surely any path is regular on a large enough scale, that is any irregular path becomes regular if it is extended by enough additional vertices. Therefore, $\{\mathbf{0} \leftrightarrow \infty\}$ equals the event that the root $\mathbf{0}$ starts an infinite path that is regular and we then show that no such path exists.

For given Poisson points $\mathbf{z}_0, \mathbf{z}_1, \dots, \mathbf{z}_m$, we write

$$\Big\{\mathbf{z}_0 \xleftrightarrow[\mathbf{z}_0,\mathbf{z}_1,\dots,\mathbf{z}_m]{k} \mathbf{z}_m\Big\}$$

for the event that $\mathbf{z}_0$ and $\mathbf{z}_m$ are connected by a path of length k, that has skeleton $\mathbf{z}_0, \mathbf{z}_1, \dots, \mathbf{z}_m$. Recall that the length of a path is the number of edges on the path and note that this definition is consistent with the previously introduced notation $\{\mathbf{x} \xleftrightarrow[\mathbf{x},\mathbf{y}]{2} \mathbf{y}\}$ and $\{\mathbf{x} \xleftrightarrow[\mathbf{x},\mathbf{y}]{k} \mathbf{y}\}$.

Proof of Theorem 1.8(a). Observe that if an irregular path of length n exists, then an irregular path of length $k \leq n$, whose end vertex carries the smallest mark of the path also exists. Let $A_{\text{irreg}}(k)$ be the event that $\mathbf{0}$ starts an irregular path of length k where the end vertex has the smallest mark. We will prove in the following lemma that

$$\mathbb{P}^\beta(A_{\text{irreg}}(k)) \leq (C_3\beta)^k$$

for some constant C_3. We then choose $\beta < C_3^{-1}$ and achieve

$$\sum_{k=1}^{\infty} \mathbb{P}^\beta(A_{\text{irreg}}(k)) < \infty.$$

Hence, the Borel-Cantelli-Lemma yields that almost surely any long enough path is regular.

Lemma 2.8. *Let $\gamma \in [0, \frac{\delta}{\delta+1})$. Then, for all $k \in \mathbb{N}$,*

$$\mathbb{P}^{\beta}(A_{irreg}(k)) \leq (C_3\beta)^k,$$

where $C_3 = 2C_2 = \frac{2^{d\delta+4} I_\rho}{\delta(1-\gamma)-\gamma}$.

Proof. A path of length k whose minimum mark vertex is also the end vertex has a skeleton whose vertices' marks are decreasing. Thus, we again write $\mathbf{0} = \mathbf{z}_0 = (z_0, s_0)$ and have by the Mecke equation as in the proof of Lemma 2.5 that

$$\begin{aligned}
&\mathbb{P}^{\beta}(A_{\mathrm{irreg}}(k)) \\
&\leq \sum_{m=1}^{k} \mathbf{E}_0\Bigg[\sum_{\substack{(z_1,s_1),\dots,(z_m,s_m)\in\mathcal{X}_0 \\ s_0>s_1>\dots>s_m \\ s_m<2^{-m}}} \mathbb{P}^{\beta}_{\mathcal{X}_0}\Big\{(z_0, s_0) \xleftrightarrow[(z_0,s_0),\dots,(z_m,s_m)]{k} (z_m, s_m)\Big\}\Bigg] \\
&= \sum_{m=1}^{k} \int_0^1 \mathrm{d}s_0 \int\limits_{\substack{(\mathbb{R}^d\times(0,1])^m \\ s_0>s_1>\dots>s_m \\ s_m<2^{-m}}} \bigotimes_{j=1}^{m} \mathrm{d}(z_j, s_j) \mathbb{P}^{\beta}_{\mathbf{z}_0,\dots,\mathbf{z}_m}\Big\{(z_0, s_0) \xleftrightarrow[\mathbf{z}_0,\dots,\mathbf{z}_m]{k} (z_m, s_m)\Big\},
\end{aligned}$$

where we have written $\mathbf{z}_j = (z_j, s_j)$ for $j = 1, \dots, m$ as usual. Using the BK-Inequality (2.7) and Lemma 2.7, we get for the last probability,

$$\begin{aligned}
&\mathbb{P}^{\beta}_{\mathbf{z}_0,\dots,\mathbf{z}_m}\Big\{(z_0, s_0) \xleftrightarrow[(z_0,s_0),\dots,(z_m,s_m)]{k} (z_m, s_m)\Big\} \\
&\leq \sum_{\substack{(n_1,\dots,n_m)\in\mathbb{N}^m: \\ n_1+\dots+n_m=k}} \prod_{j=1}^{m} \mathbb{P}^{\beta}_{\mathbf{z}_{j-1},\mathbf{z}_j}\{(z_{j-1}, s_{j-1}) \xleftrightarrow[\mathbf{z}_{j-1},\mathbf{z}_j]{n_j} (z_j, s_j)\} \\
&\leq \sum_{\substack{(n_1,\dots,n_m)\in\mathbb{N}^m: \\ n_1+\dots+n_m=k}} (\beta C_2)^{k-m} \prod_{j=1}^{m} \mathbb{P}^{\beta}_{\mathbf{z}_{j-1},\mathbf{z}_j}\{(z_{j-1}, s_{j-1}) \sim (z_j, s_j)\} \\
&= \binom{k-1}{m-1} (\beta C_2)^{k-m} \prod_{j=1}^{m} \mathbb{P}^{\beta}_{\mathbf{z}_{j-1},\mathbf{z}_j}\{(z_{j-1}, s_{j-1}) \sim (z_j, s_j)\}.
\end{aligned}$$

Here, we used that either the consecutive skeleton vertices $\mathbf{z}_{i-1}$ and $\mathbf{z}_i$ fulfil the minimum distance for shortcut-free paths or $n_i = 1$. Therefore,

$$\begin{aligned}
&\mathbb{P}^\beta(A_{\text{irreg}}(k)) \\
&\le \sum_{m=1}^{k} \binom{k-1}{m-1} (\beta C_2)^{k-m} \int_0^1 \mathrm{d}s_0 \int_0^{s_0} \mathrm{d}s_1 \int_{\mathbb{R}^d} \mathrm{d}z_1 \dots \\
&\qquad\qquad \dots \int_0^{2^{-m}\wedge s_{m-1}} \mathrm{d}s_m \int_{\mathbb{R}^d} \mathrm{d}z_m \Big(\prod_{i=1}^{m} \rho(\beta^{-1} s_{i-1}^{1-\gamma} s_i^{\gamma} |z_{i-1} - z_i|^d) \Big) \\
&\le \sum_{m=1}^{k} \binom{k-1}{m-1} I_\rho^m C_2^{k-m} \beta^m \int_0^1 \mathrm{d}s_0 \int_0^{s_0} \mathrm{d}s_1 \cdots \int_0^{2^{-m}\wedge s_{m-1}} \mathrm{d}s_m \; s_0^{\gamma-1} s_m^{-\gamma} \prod_{i=1}^{m-1} s_i^{-1} \\
&\le \sum_{m=1}^{k} \binom{k-1}{m-1} \beta^m I_\rho^m C_2^{k-m} (1-\gamma)^{-m} \\
&\le (\beta C_2)^k \sum_{m=1}^{k} \binom{k-1}{m-1} \le (\beta C_3)^k,
\end{aligned}$$

where the third inequality follows from Lemma B.5. □

The previous lemma shows that for $\beta < C_3^{-1}$, it suffices to show that $\mathbf{0}$ does not start an infinite path that is regular in order to obtain $\theta(\beta) = 0$. Let $A_{\text{reg}}(n)$ be the event that $\mathbf{0}$ starts a regular path of length n.

Lemma 2.9. *Let $\gamma \in [\frac{1}{2}, \frac{\delta}{\delta+1})$. Then, for all $n \in \mathbb{N}$, we have*

$$\mathbb{P}^\beta(A_{reg}(n)) \le K(\beta C_3)^n,$$

where $C_3 = 2C_2 = \frac{I_\rho 2^{d\delta+4}}{\delta(1-\gamma)-\gamma}$ and K is some constant.

Proof. Writing $\mathbf{0} = \mathbf{z}_0 = (z_0, s_0)$ and following the same arguments of Mecke equation, BK-Inequality and Lemma 2.7 as done in the previous proof of Lemma 2.8, we get for large enough n that

$$
\begin{aligned}
&\mathbb{P}^{\beta}(A_{\mathrm{reg}}(n)) \\
&\leq \sum_{m=1}^{n} \sum_{k=0}^{m} \int_{2^{-m}}^{1} \mathrm{d}s_0 \binom{n-1}{m-1} (C_2\beta)^{n-m} \qquad (2.9)\\
&\quad \times \int\limits_{\substack{(z_1,s_1),\dots(z_m,s_m)\in\mathbb{R}^d\times(0,1] \\ s_0>s_1>\dots>s_k>2^{-m} \\ s_k<s_{k+1}<\dots<s_m}} \bigotimes_{j=1}^{m} \mathrm{d}(z_j,s_j) \prod_{j=1}^{m} \mathbb{P}^{\beta}_{\mathbf{z}_{j-1},\mathbf{z}_j}\{(z_{j-1},s_{j-1}) \sim (z_j,s_j)\}.
\end{aligned}
$$

Here, the two sums and integrals describe all regular skeletons a regular path of length n can have. For the calculation, we focus on $\gamma > 1/2$. For $\gamma = 1/2$ minor changes are needed; we comment on this below. Recall that

$$
\mathbb{P}^{\beta}_{\mathbf{z}_{j-1},\mathbf{z}_j}\{(z_{j-1},s_{j-1}) \sim (z_j,s_j)\} = \rho(\beta^{-1} g^{\mathrm{pa}}(s_{j-1},s_j)|z_{j-1}-z_j|^d).
$$

Therefore, the right-hand side of (2.9) reads

$$
\begin{aligned}
&\sum_{m=1}^{n} \binom{n-1}{m-1} (C_2\beta)^{n-m} \\
&\quad \times \sum_{k=0}^{m} (\beta I_\rho)^m \int\limits_{\substack{1>s_0>s_1>\dots>s_k>2^{-m} \\ s_k<s_{k+1}<\dots<s_m}} \bigotimes_{j=0}^{m} \mathrm{d}s_j \prod_{j=1}^{m} g^{\mathrm{pa}}(s_{j-1},s_j)^{-1}. \qquad (2.10)
\end{aligned}
$$

For $k=0$ the integral from (2.10) can be written as

$$
\int_{2^{-m}}^{1} \mathrm{d}s_0 \int_{s_0}^{1} \mathrm{d}s_1 \cdots \int_{s_{m-1}}^{1} \mathrm{d}s_m \; s_0^{-\gamma} s_m^{\gamma-1} \prod_{j=1}^{m-1} s_j^{-1} \leq \left(\frac{1}{1-\gamma}\right)^m,
$$

by Lemma B.1. For $k=m$, we obtain for the integral from (2.10)

$$
\int_{2^{-m}}^{1} \mathrm{d}s_0 \int_{2^{-m}}^{s_0} \mathrm{d}s_1 \cdots \int_{2^{-m}}^{s_{m-1}} ds_m \; s_0^{\gamma-1} s_m^{-\gamma} \prod_{j=1}^{m-1} s_j^{-1} \leq \left(\frac{1}{1-\gamma}\right)^m,
$$

by Lemma B.5. For $1 \leq k \leq m-1$, we infer for the integral from (2.10), using Lemma B.4,

$$\sum_{k=1}^{m-1} \int_{2^{-m}}^{1} \mathrm{d}s_0 \int_{2^{-m}}^{s_0} \mathrm{d}s_1 \cdots \int_{2^{-m}}^{s_{k-1}} \mathrm{d}s_k \left[t_0^{\gamma-1} \left(\prod_{j=1}^{k-1} s_j^{-1} \right) s_k^{-\gamma} \right.$$
$$\left. \times \int_{s_k}^{1} \mathrm{d}s_{k+1} \cdots \int_{s_{m-1}}^{1} \mathrm{d}s_m \left[s_k^{-\gamma} \left(\prod_{j=k+1}^{m-1} s_j^{-1} \right) s_m^{\gamma-1} \right] \right]$$
$$\leq \frac{2^{-m(1-2\gamma)}(m\log(2))^{m-2}}{\gamma^2(2\gamma-1)(m-2)!} \sum_{k=1}^{m-1} \binom{m-2}{k-1}.$$

Since $m^{m-2}/(m-2)!$ asymptotically equals $2^{\log_2(e)(m-2)}/\sqrt{2\pi(m-2)}$ by Stirling's formula, and $\sum_{k=1}^{m-1}\binom{m-2}{k-1} \leq 2^{m-2}$, we infer from (2.9) and (2.10)

$$\mathbb{P}^\beta(A_{\mathrm{reg}}(n))$$
$$\leq \beta^n K \sum_{m=1}^{n} \binom{n-1}{m-1} I_\rho^m C_2^{n-m} \left((1-\gamma)^{-m} + (2^{2\gamma+\log_2(e)}\log(2))^m \right),$$

for some constant $K \geq 2$. As $C_2 > (1-\gamma)^{-1}$ and $C_2 \geq 2^{2\gamma+\log_2(e)}\log(2)$ we infer that

$$\mathbb{P}^\beta(A_{\mathrm{reg}}(n)) \leq K(\beta C_3)^n.$$

For $\gamma = \frac{1}{2}$, Lemma B.2 and Lemma B.4 have to be modified slightly. The changes in the calculations only influence the value of K and not the constant C_3. □

Setting $\beta < C_3^{-1}$ concludes the proof of Theorem 1.8(a). □

2.5. Absence of a subcritical phase

In this section, we prove Theorem 1.8(b) using a strategy of Jacob and Mörters [45]. Starting from a sufficiently powerful vertex, we use a connector to connect the powerful vertex with a much more powerful one; we repeat this indefinitely, moving to more and more powerful vertices as we go along. To ensure that this procedure generates an infinite path with positive probability, we have to show that the failure probabilities of connecting the pairs of increasingly powerful vertices sum to a probability strictly less than one.

To this end, we show that an powerful vertex is *with extreme probability* connected to a much more powerful one by a single connector. Here, if $(A(t))_{t>0}$ is a family of events, we say an event $A(t)$ holds with extreme probability, or $wep(t)$, if it holds with probability at least $1 - \exp(-\Omega(\log^2(t)))$, as $t \to \infty$, where $\Omega(t)$ is the standard Landau symbol. Observe, if $(A(t)_n)_{n\in\mathbb{N}}$ is a sequence of events, holding simultaneously $wep(t)$ in the sense that

$$\inf_n \mathbb{P}(A(t)_n) \geq 1 - \exp(-\Omega(\log^2(t))),$$

as $t \to \infty$, then $\bigcap_{k\leq\lfloor t\rfloor} A(t)_k$ holds $wep(t)$.

Because $g^{\mathrm{pa}}, g^{\mathrm{sum}} \leq g^{\mathrm{str}}$ we can fix the kernel g to be the strong-kernel g^{str} (1.11) throughout this section. Hence, for two given vertices $\mathbf{x} = (x, t)$ and $\mathbf{y} = (y, s)$, the connection probability is given by

$$\rho(\beta^{-1}(s \wedge t)^{-\gamma}|x-y|^d).$$

Recall that ρ is regularly varying with index $-\delta$ for $\delta > 1$, cf. (1.17). Further, $\gamma > \delta/(\delta+1)$. Thus, we can choose

$$\alpha_1 \in \left(1, \tfrac{\gamma}{\delta(1-\gamma)}\right) \text{ and then fix } \alpha_2 \in \left(\alpha_1, \tfrac{\gamma}{\delta}(1+\alpha_1\delta)\right).$$

The following lemma shows that the outlined strategy for an infinite path works and thus proves Theorem 1.8(b).

Lemma 2.10. *Let $\gamma > \frac{\delta}{\delta+1}$ and ρ be regularly varying with index $-\delta$ for $\delta > 1$. Let α_1, α_2 be as defined as above. Let $\mathbf{z}_0 = (z_0, s_0)$ be a given Poisson point with $s_0 < 1/2$. Then, for any $\beta > 0$, there exists, $wep(1/s_0)$, a sequence $(\mathbf{z}_k)_{k\in\mathbb{N}}$ of vertices $\mathbf{z}_k = (z_k, s_k) \in \mathcal{X}_0$ such that for all $k \in \mathbb{N}$*

(i) $s_k < s_{k-1}^{\alpha_1}$ *and* $|z_k - z_{k-1}|^d < \frac{\beta}{2}s_{k-1}^{-\alpha_2}$ *and*

(ii) $\mathbf{z}_{k-1} \xleftrightarrow[\mathbf{z}_{k-1},\mathbf{z}_k]{2} \mathbf{z}_k$.

Proof. It suffices to show that, $wep(1/s_0)$, there exists a vertex $\mathbf{z}_1 = (z_1, s_1)$ satisfying (i) and (ii). The result then follows by induction. The number of vertices, with mark smaller than $s_0^{\alpha_1}$ and within distance $((\beta/2)s_0^{-\alpha_2})^{1/d}$ from z_0 is Poisson distributed with parameter

$$\mathrm{Volume}\Big(\{|z_1 - z_0|^d < \tfrac{\beta}{2} s_0^{-\alpha_2}\} \times (0, s_0^{\alpha_1})\Big) = O(s_0^{\alpha_1 - \alpha_2}),$$

where $O(\cdot)$ again is the standard Landau symbol. Since $\alpha_2 > \alpha_1$, such a vertex $\mathbf{z}_1$ exists $wep(1/s_0)$. To connect $\mathbf{z}_0$ to $\mathbf{z}_1$ via a connector, we focus on connectors (y, t), with mark larger than $1/2$ and within distance $((\beta/2)s_0^{-\gamma})^{1/d}$ from $\mathbf{z}_0$. Since, for such choices of (y, t), we have

$$|z_1 - y|^d \le \Big((\tfrac{\beta s_0^{-\alpha_2}}{2})^{1/d} + (\tfrac{\beta s_0^{-\gamma}}{2})^{1/d} \Big)^d \le \beta s_0^{-\alpha_2},$$

the number of such connectors is Poisson distributed with its parameter bounded from below by

$$\begin{aligned} &\int_{1/2}^{1} \mathrm{d}t \int\limits_{\{|y - z_0|^d \le \frac{\beta}{2} s_0^{-\gamma}\}} \mathrm{d}y\ \rho(\beta^{-1} s_0^{\gamma} |y - z_0|^{-d}) \rho(s_0^{\alpha_1 \gamma - \alpha_2}) \\ &\quad = \tfrac{1}{2} \beta s_0^{-\gamma} \rho(s_0^{\alpha_1 \gamma - \alpha_2}) \int\limits_{\{|y - z_0|^d \le 1/2\}} \mathrm{d}y\ \rho(|y - z_0|^d). \end{aligned} \tag{2.11}$$

Now, we choose $\varepsilon > 0$ such that $\tilde{\delta} := \delta + \varepsilon < \frac{\gamma}{1-\gamma}$, or equivalently $\gamma > \tilde{\delta}/(\tilde{\delta} + 1)$, and infer by the Potter bound [7, Theorem 1.5.6],

$$\rho(s_0^{\alpha_1 \gamma - \alpha_2}) \ge A s_0^{-\tilde{\delta}(\alpha_1 \gamma - \alpha_2)},$$

for some $A < 1$ and s_0 small enough. Additionally, $\rho(|x|^d) \ge \rho(1/2) > 0$ for all $|x|^d < 1/2$. Hence, (2.11) is bounded from below by

$$\Omega\Big(s_0^{-\tilde{\delta}(\alpha_1 \gamma - \alpha_2) - \gamma}\Big).$$

Therefore, $wep(1/s_0)$, the vertex $\mathbf{z}_1$ satisfies (ii) as

$$\mathbb{P}^{\beta}_{\mathbf{z}_0,\mathbf{z}_1}\{\mathbf{z}_0 \xleftrightarrow[\mathbf{z}_0,\mathbf{z}_1]{2} \mathbf{z}_1\} \geq 1 - \exp(-\Omega(s_0^{-\tilde{\delta}(\alpha_1\gamma-\alpha_2)-\gamma}))$$

and $-\tilde{\delta}(\alpha_1\gamma - \alpha_2) - \gamma < 0$. □

The proofs of this section and Section 2.4 remain valid if the underlying Poisson process is replaced by a site percolated lattice. The only calculations in which we explicitly used the Poisson process are the applications of Mecke's equation and where we calculated the expected number of connectors to get bounds for the probability of certain paths existing in the graph. In the lattice case Mecke's equation can be easily replaced by the sum over all lattice point conditioned on the event that the considered point survived the percolation, which is of the same order as the integral coming from Mecke's equation with an additional factor p for the survival probability. Further, the expected number of points remains of the same order and similar as for the degree distribution at the end of Section 2.2.1, we can use the arguments of [17, 70].

2.6. Existence of a supercritical phase in dimension one

In this section, we prove Theorem 1.11(a). Therefore, we work now explicitly in dimension $d = 1$. Recall that the vertex locations η_0 are constructed in a way such that $X_i < X_j$ for all $X_i, X_j \in \eta_0$ if $i < j$. For the proof the only requirement on the vertex locations is that they behave 'lattice-like' on large scales. Hence the proof of Theorem 1.11 works whenever η_0 fulfils the following regularity condition.

Definition 2.11 (Evenly spaced point process)**.** Let η be a stationary ergodic simple point process on $\mathbb{R}$ and denote by $\mathbf{P}_0$ the law of its Palm version η_0. We say that η (and also η_0) is *evenly spaced*, if

(a) there exists a constant $a_1 > 0$ such that

$$\sum_{n\in\mathbb{N}} \mathbf{P}_0\{|X_{-K_n} - X_{K_n-1}| > a_1 K_n\} = o(1), \text{ as } K \to \infty,$$

where $K_n = (n!)^3 K^n, n \geq 1$;

(b) there exists a constant $a_2 > 0$ such that we have

$$\mathbf{P}_0\{|X_{-2^{n+1}} - X_{2^n}| < a_2 2^n\} = o(1), \text{ as } n \to \infty.$$

Remark 2.12.

(i) There is some leeway in the choice of sequences in Definition 2.11 – for canonical examples of evenly spaced processes other sequences can be used to derive the same results. For example, choose $a_1 = 2(1+\varepsilon)\lambda$ in (a), where $\varepsilon > 0$ and $\lambda > 0$ is the intensity of η. Then, by stationarity, $\mathbf{E}_0\eta(-K_n, K_n - 1) = 2\lambda K_n$ and the speed at which K_n needs to grow can be derived from existing large deviation estimates for η.

(ii) Property (a) is a stronger requirement than property (b) because to prove the existence of a supercritical phase we need that vertex locations are sufficiently dense uniformly over all scales whereas it suffices for the non-existence proof that vertices are spaced out on each scale.

(iii) Examples of evenly spaced point processes are:

- A Poisson process of intensity $\lambda > 0$.

- The point process induced by performing i.i.d. Bernoulli(p) site percolation on $\mathbb{Z}$ with retention probability $p \in (0,1]$. Note that in this case, stationarity is with respect to shifts induced by $\mathbb{Z}$.

- Depending on the underlying random intensity measure, Cox processes may be evenly spaced, see [46, Section 3] for examples.

We now state a proposition which proves a sufficient condition for the existence of an infinite cluster.

Proposition 2.13 (Existence of an infinite component)**.** *Let η be a simple and ergodic point process in dimension $d = 1$ satisfying assumption (a) of Definition 2.11 and let $\mathcal{X}_0$ and $\mathcal{U}_0$ be defined as in Section 2.1. Let ρ be a profile-function and g be a kernel-function and $(\mathcal{G}^\beta = \mathcal{G}^{\beta,\rho,g}(\mathcal{X}_0,\mathcal{U}_0) : \beta > 0)$ the associated weight-dependent random connection model. Let $K_n := (n!)^3 K^n, n \in \mathbb{N}$ for some fixed $K \in \mathbb{N}$. Assume that there exist $\mu \in (0, 1/2)$ such that*

$$\lim_{K\to\infty} \sup_{n\geq 2} n^3 K \exp\Big(- K_{n-1}^2 \int\limits_{[K_{n-1}^{\mu-1}, 1-K_{n-1}^{\mu-1}]^2} \mathrm{d}s\mathrm{d}t \; \rho\big(g(s,t)K_n\big)\Big) = 0. \quad \text{(A1)}$$

Then $\beta_c < \infty$.

The proof of Proposition 2.13 is in part based on [24, Theorem 1(i)]. Condition (A1) can be seen as a version of [24, Eq. (8)]. Note that (A1) is a slightly more technical assumption then the limit condition required in Theorem 1.11(a). It precisely quantifies the order of the probability that certain long edges are not present in the graph we need for our proof to work. However, it is straightforward to deduce that (A1) is satisfied whenever

$$-\liminf_{n\to\infty} \frac{\log\Big(\int_{1/n}^1 \int_{1/n}^1 \rho(g(s,t)n)\,\mathrm{d}s\,\mathrm{d}t\Big)}{\log n} < 2.$$

Another application of our results is the existence of a component of linear size in finite intervals. The paper [43] elaborates how weight-dependent random connection-type models arise as weak local limits of models on finite domains. There, growing sequences of graphs $(\mathscr{G}_n(\beta))$ are constructed where $\mathscr{G}_n(\beta)$ consists of n vertices which are independently placed into the unit interval $(-1/2, 1/2)$. Translated to our parametrisation each vertex carries an independent uniform mark and, given locations and marks, each pair $(x,t),(y,s)$ of vertices is connected independently with probability $\rho(\beta^{-1}g(s,t)n|x-y|)$. The scaling factor n ensures that the graph remains sparse and the n vertices can hence be considered as being embedded into $(-n/2, n/2)$. Note that the graph sequence $(\mathscr{G}_t : t \geq 0)$ of Section 1.2.1 arising from the rescaling of the age-based spatial preferential attachment network is a special instance of this type of models which arises naturally from this context. Our proofs yield the following corollary regarding finite versions of the weight-dependent random connection model.

Corollary 2.14. *Let $(\mathscr{G}_n(\beta) : n \in \mathbb{N})$ be the above sequence of finite graphs on intervals with weak local limit $\mathscr{G}(\beta)$ given by an instance of the weight-dependent random connection model with kernel g and profile ρ. If ρ and g satisfy the assumptions of Theorem 1.11(a), then $(\mathscr{G}_n(\beta) : n \in \mathbb{N})$ contains a giant component for large enough β. Conversely, if ρ and g satisfy the assumptions of Theorem 1.11(b), then there is not giant component for any $\beta > 0$.*

It is well-known (cf. Lemma 3.16) that a graph sequence cannot have a giant component when its weak local limit does not percolate. Hence it remains to prove the first statement of Corollary 2.14 which is done at the end of this section.

2.6.1. Proof of Proposition 2.13

We now proceed to prove the existence of an infinite component under the assumptions of Proposition 2.13. Our proof is based on an argument given in [24, Theorem 1(i)], where the existence of an infinite path in classical long range percolation on $\mathbb{Z}$ is proven in the scale-invariant regime $\delta = 2$. In their proof, Duminil-Copin et al. define a renormalisation scheme that works roughly as follows: At stage n, the lattice is covered by half-way overlapping blocks of K_n lattice points. The overlap has the effect that if two adjacent blocks contain a rather dense connected component each, the two components must intersect by necessity. They then argue that since enough blocks at stage n have such components, a larger block of size K_{n+1} containing several K_n-blocks must also contain a component of positive (but ever so slightly smaller) density. The result then follows by iterating this construction and taking the limit $n \to \infty$. In our model we have to control the additional randomness of the marks, so we cannot quite tackle the scale-invariant situation as in [24], but we demonstrate below that a modified version of the strategy works under the assumptions of Proposition 2.13, i.e. in particular if $\delta_{\text{eff}} < 2$.

To apply ideas developed for the lattice model to the more general underlying point process η, we make use of the evenly spaced property. This property guarantees that in all ways that matter for the proof, η behaves like the lattice on large scales. Let $K_n = (n!)^3 K^n$ be as in Definition 2.11(a). By the same assumption, we can choose a sufficiently large K such that the probability measure

$$\widetilde{\mathbb{P}}^\beta(\,\cdot\,) = \mathbb{P}^\beta\Big(\,\cdot\,\Big|\bigcap_{n\in\mathbb{N}}\{|X_{-K_n} - X_{K_n-1}| \leq aK_n\}\Big)$$

is well-defined. Define also $\tilde{\theta}(\beta) = \widetilde{\mathbb{P}}^\beta\{\mathbf{0} \leftrightarrow \infty\}$ and note that if $\tilde{\theta}(\beta) > 0$, then $\theta(\beta) > 0$ as well. Throughout this section, we make the standing assumption that K is chosen large enough for $\widetilde{\mathbb{P}}$ to be well-defined.

For $N \in \mathbb{N}$ and $i \in \mathbb{Z}$ let

$$B_N^i := \{\mathbf{X}_{N(i-1)}, \ldots, \mathbf{X}_{Ni}, \ldots, \mathbf{X}_{N(i+1)-1}\}$$

and $B_N := B_N^0 = \{\mathbf{X}_{-N}, \ldots, X_{N-1}\}$. Each set B_N^i consists of precisely $2N$

consecutive vertices. If η is the lattice, then $X_j = j$ for each $j \in \mathbb{Z}$ and B_N^i is simply the lattice interval $[N(i-1), N(i+1)) \cap \mathbb{Z}$, matching the notation of [24]. In the general setting, the sets B_N^i are blocks of vertices that all contain the same number of vertices but with random distances between consecutive vertices. Note that two consecutive blocks B_N^i and B_N^{i-1} overlap on half of their vertices. The blocks at stage n are then given by the blocks $B_{K_n}^i$ for $i \in \mathbb{Z}$.

Connecting vertex sets that are far apart. To make sure that the strategy outlined at the beginning of this section works and that a K_{n+1}-block at stage $n+1$ contains a "large" connected component (we will specify this shortly), it is necessary that two stage n blocks at a given distance are connected with a sufficiently high probability to overcome potentially bad regions.

Recall that $K_n = (n!)^3 K^n$ for some $K \in \mathbb{N}$. For $\vartheta^* \in (0,1)$ and $n \geq 2$, we define the 'leftmost' and 'rightmost' parts of B_{K_n}

$$V_\ell^n(\vartheta^*) := \{\mathbf{X}_{-K_n}, \dots, \mathbf{X}_{-K_n + \lfloor \vartheta^* K_{n-1} \rfloor - 1}\} \quad \text{and}$$
$$V_r^n(\vartheta^*) := \{\mathbf{X}_{K_n - \lfloor \vartheta^* K_{n-1} \rfloor}, \dots, \mathbf{X}_{K_n - 1}\}.$$

Note that $K_{n-1} \ll K_n$ and so $V_\ell^n(\vartheta^*)$ (resp. $V_r^n(\vartheta^*)$) is only a relatively small number of vertices at the very left (resp. right) end of the block B_{K_n}. Before calculating the probability of the two sets $V_\ell^n(\vartheta^*)$ and $V_r^n(\vartheta^*)$ being connected, we need to understand the behaviour of the vertex marks inside each set.

For $\mu \in (0, 1/2)$, we denote for all $i = 1, \dots, \lfloor (\vartheta^* K_{n-1})^{1-\mu} \rfloor$ the empirical mark counts in $V_\ell^n(\vartheta^*)$ by

$$N_\ell^n(i) := \sum_{S : (X,S) \in V_\ell^n(\vartheta^*)} \mathbb{1}_{\left\{S \leq \frac{i}{\lfloor (\vartheta^* K_{n-1})^{1-\mu} \rfloor}\right\}}$$

and we say that $V_\ell^n(\vartheta^*)$ has *μ-regular* vertex marks if

$$N_\ell^n(i) \geq \frac{i\vartheta^* K_{n-1}}{2\lfloor(\vartheta^* K_{n-1})^{1-\mu}\rfloor}$$

for all $i = 1, \ldots, \lfloor(\vartheta^* K_{n-1})^{1-\mu}\rfloor$. A simple calculation yields that

$$\mathbb{E}^\beta N_\ell^n(i) = \frac{\vartheta^* K_{n-1} i}{\lfloor(\vartheta^* K_{n-1})^{1-\mu}\rfloor}.$$

Hence, by using a Chernoff bound for Uniform(0,1) random variables, we have that

$$\mathbb{P}^\beta\Big\{N_\ell^n(i) < \mathbb{E}^\beta N_\ell^n(i)/2\Big\} \leq \exp\Big(-c\cdot i(K_{n-1})^\mu\Big)$$

for some constant $c > 0$, depending only on the value of ϑ^*. Therefore,

$$\mathbb{P}^\beta\{V_\ell^n(\vartheta^*) \text{ is } \mu\text{-regular}\} \geq 1 - K_{n-1}^{1-\mu}\exp(-cK_{n-1}^\mu). \tag{2.12}$$

The same holds verbatim for $V_r^n(\vartheta^*)$. Hence, both sets are μ-regular with a stretched exponential error bound already in the first stage for a sufficient large K. We therefore focus on the case when both sets have μ-regular vertex marks when calculating the probability of both sets being connected, which we do now.

Denote by $\{V_\ell^n(\vartheta^*) \sim V_r^n(\vartheta^*)\}$ the event that the two sets are connected by a direct edge, i.e. there exist $\mathbf{X} \in V_\ell^n(\vartheta^*)$ and $\mathbf{Y} \in V_r^n(\vartheta^*)$ such that $\mathbf{X} \sim \mathbf{Y}$.

Lemma 2.15. *Let $\vartheta^* \in (0,1)$ and write $v := v_n := \vartheta^* K_{n-1}$. There exists a constant $C = C(g,\rho) > 0$ such that for all $\mu \in (0, 1/2)$ and $n \geq 2$ we have*

$$\begin{aligned}&\widetilde{\mathbb{P}}^\beta\Big(V_\ell^n(\vartheta^*) \sim V_r^n(\vartheta^*) \mid V_\ell^n(\vartheta^*) \text{ and } V_r^n(\vartheta^*) \text{ are } \mu\text{-regular}\Big)\\ &\quad\geq 1 - \exp\Big(-Cv^2\int_{v^{\mu-1}}^{1-v^{\mu-1}} \mathrm{d}s \int_{v^{\mu-1}}^{1-v^{\mu-1}} \mathrm{d}t\ \rho\Big(\beta^{-1}g(s,t)aK_n\Big)\Big).\end{aligned}$$

Proof. To lighten notation we write $V_\ell = V_\ell^n(\vartheta^*)$ and $V_r = V_r^n(\vartheta^*)$ and denote

$$E = \{V_\ell\ \mu\text{-regular}\} \cap \{V_r\ \mu\text{-regular}\}.$$

Denote by F_i the empirical distribution function of the vertex marks corresponding to V_i, where $i \in \{\ell, r\}$. Writing $h := \lfloor(\vartheta^* K_{n-1})^{1-\mu}\rfloor$, we have on the event E for $t \in [0,1]$ by the definition of μ-regularity

$$\begin{aligned} vF_\ell(t) &= \sum_{i=-K_n}^{-K_n+\lfloor\vartheta^* K_{n-1}\rfloor-1} \mathbb{1}_{\{T_i \le t\}} \ge \sum_{j=1}^{h} N_\ell^n(j-1)\mathbb{1}_{\{(j-1)/h<t\le j/h\}} \\ &= N_\ell^n(\lfloor th\rfloor) \ge \frac{v\lfloor th\rfloor}{2h} \ge \frac{v}{2}(t-1/h) \ge \frac{v}{3}(t-v^{\mu-1}). \end{aligned} \tag{2.13}$$

The same holds for F_r. Under $\widetilde{\mathbb{P}}^\beta$, the point process η is concentrated on point configurations that satisfy for any $(X_i, T_i) \in V_1$ and $(X_j, T_j) \in V_2$ that $|X_i - X_j| \le aK_n$. We call such a point configuration $\omega = (x_i, i \in \mathbb{Z})$ *properly spaced*. Further, for any fixed properly spaced ω, we denote by $\widetilde{\mathbb{P}}^\beta_{\eta=\omega}$ the measure on edge and vertex mark configurations given the vertex locations ω. By construction $\widetilde{\mathbb{P}}^\beta_{\eta=\omega}$ is a product measure with Uniform$(0,1)$-marginals. Under $\widetilde{\mathbb{P}}^\beta_{\eta=\omega}$, two vertices (x_i, T_i) and (x_j, T_j) are connected, whenever their corresponding edge mark satisfies

$$U_{i,j} \le 1 - \exp(-\rho(\beta^{-1} g(T_i, T_j) aK_n)), \tag{2.14}$$

since ω is properly spaced. In particular, there always exists a direct edge connecting V_ℓ and V_r if

$$\Sigma := \sum_{\substack{(x_i,T_i)\in V_\ell,\\ (x_j,T_j)\in V_r}} \mathbb{1}_{\left\{U_{i,j} \le 1-\exp(-\rho(\beta^{-1} g(T_i,T_j)aK_n)\right\}} > 0.$$

Since the edge marks are independent of the vertex marks and locations, we have

$$
\begin{aligned}
&\widetilde{\mathbb{E}}^\beta_{\eta=\omega}(\mathbb{1}_{\{\Sigma=0\}}\mathbb{1}_E)\\
&\quad\le \widetilde{\mathbb{E}}^\beta_{\eta=\omega}\Bigg[\mathbb{1}_E \prod_{\substack{(x,T)\in V_\ell\\(y,S)\in V_r}} \exp\Big(-\rho(\beta^{-1}g(T,S)aK_n)\Big)\Bigg]\\
&\quad= \widetilde{\mathbb{E}}^\beta_{\eta=\omega}\Bigg[\exp\Big(-\sum_{\substack{(x,T)\in V_\ell\\(y,S)\in V_r}}\rho\Big(\beta^{-1}g(T,S)aK_n\Big)\Big)\mathbb{1}_E\Bigg]\\
&\quad\le \widetilde{\mathbb{E}}^\beta_{\eta=\omega}\Bigg[\exp\Big(-\frac{v^2}{9}\int_{v^{\mu-1}}^{1-v^{\mu-1}}\mathrm{d}t\int_{v^{\mu-1}}^{1-v^{\mu-1}}\mathrm{d}s\ \rho\left(\beta^{-1}g(t,s)aK_n\right)\Big)\mathbb{1}_E\Bigg]\\
&\quad= \widetilde{\mathbb{P}}^\beta_{\eta=\omega}(E)\exp\Big(-\frac{v^2}{9}\int_{v^{\mu-1}}^{1-v^{\mu-1}}\mathrm{d}t\int_{v^{\mu-1}}^{1-v^{\mu-1}}\mathrm{d}s\ \rho\left(\beta^{-1}g(t,s)aK_n\right)\Big),
\end{aligned}
$$

where we have used in the second inequality that on E and with (2.13)

$$
\begin{aligned}
\sum_{\substack{(x,T)\in V_\ell\\(y,S)\in V_r}}\rho\left(\beta^{-1}g(T,S)aK_n\right) &= v^2\int_0^1 F_\ell(\mathrm{d}t)\int_0^1 F_r(\mathrm{d}s)\rho\left(\beta^{-1}g(t,s)aK_n\right)\\
&\ge \frac{v^2}{9}\int_{v^{\mu-1}}^{1-v^{\mu-1}}\mathrm{d}t\int_{v^{\mu-1}}^{1-v^{\mu-1}}\mathrm{d}s\ \rho\left(\beta^{-1}g(t,s)aK_n\right).
\end{aligned}
$$

The proof now concludes with the observation that the established bound is uniform in all properly spaced configurations ω. □

Renormalisation scheme. For $N\in\mathbb{N}, i\in\mathbb{Z}$, we denote by $\mathscr{C}(B_N^i)$ the largest connected component of the subgraph of $\mathcal{G}^\beta$ on the vertices of B_N^i.

For $\vartheta\in(0,1)$, we say a block B_N^i is ϑ-*good*, if it contains a connected component of size at least $2\vartheta N$; otherwise we call it ϑ-*bad*. We denote by

$$p_\beta(N,\vartheta) := \widetilde{\mathbb{P}}^\beta\Big\{\sharp\mathscr{C}(B_N) < 2N\vartheta\Big\}$$

the probability that the block B_N is ϑ-bad. Here, $\sharp A$ denotes again the number of points in a finite set A. We will show that the probability of the scale n Block B_{K_n} being ϑ-bad can be bounded by the probability that the smaller block $B_{K_{n-1}}$ is bad with a slightly larger value of ϑ.

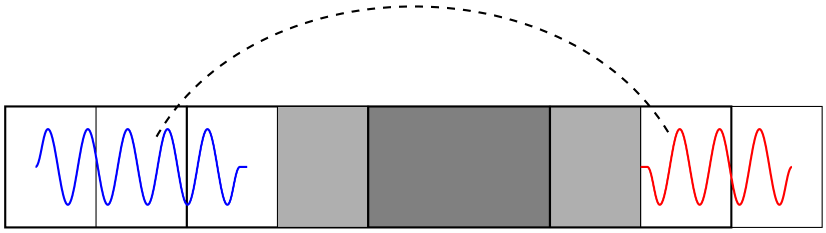

Figure 2.6.: The overlapping blocks of scale $n-1$ that together form the scale n block. The cluster $\mathscr{C}^-$ (resp. $\mathscr{C}^+$) on the left in blue (resp. on the right in red). The dark block is the bad block and in light gray are the non overlapping halves of the two neighbouring blocks. The dotted line indicates the existence of a direct edge connecting $\mathscr{C}^-$ and $\mathscr{C}^+$ avoiding the bad region.

As a first step, we show that it is probable that the subgraph on $B_{K_{n+1}}$ contains a connected component of size proportional to $\vartheta-\varepsilon$, provided that B_{K_n} is ϑ-good with a sufficiently large probability. This is an adaptation of [24, Lemma 2] to our setting. Afterwards, we show that for a large enough K determining the initial scale, the subgraph on B_{K_1} contains a sufficiently large cluster. Combining both results yields Proposition 2.13.

Recall that we have $K_n = (n!)^3 K^n$. Define a sequence $(C_n)_{n\in\mathbb{N}}$ as $C_n = n^3 K$. Then $K = K_1 = C_1$ and $K_n = C_n K_{n-1}$ for $n \geq 2$.

Lemma 2.16. *Let $\vartheta^* \in (3/4, 1)$ and $\vartheta \in (\vartheta^*, 1)$. Under the assumptions of Proposition 2.13 there exists $M > 0$ such that for all $K \geq M$ and $n \geq 2$ it holds $\vartheta - {}^2/{}_{C_n} \geq \vartheta^*$ and*

$$p_\beta(K_n, \vartheta - 2/C_n) \leq \tfrac{1}{100} p_\beta(K_{n-1}, \vartheta) + 2C_n^2 p_\beta(K_{n-1}, \vartheta)^2.$$

Proof. Let $\vartheta' := \vartheta - 2/C_n$. We begin by considering the blocks $B^i_{K_{n-1}}$ for $|i| \in \{0, \dots, C_n - 1\}$, which together form B_{K_n}, and their largest connected components $\mathscr{C}(B^i_{K_{n-1}})$. Since $\vartheta > 3/4$, $\mathscr{C}(B^i_{K_{n-1}})$ is unique if $B^i_{K_{n-1}}$ is ϑ-good. Furthermore, due to the overlapping property of neighbouring blocks, the largest components of two adjacent ϑ-good blocks have to intersect in at least one vertex. Hence, if all the blocks $B^i_{K_{n-1}}$ are ϑ-good, then B_{K_n} is ϑ-good as well.

Define now for every $|i| \in \{0, \dots, C_n - 1\}$ the event

$$E_i = \Big\{\sharp\mathscr{C}(B^i_{K_{n-1}}) < 2\vartheta K_{n-1}\Big\} \cap \Bigg[\bigcap_{\substack{|j|=0 \\ j\notin\{i-1,i,i+1\}}}^{C_n-1} \Big\{\sharp\mathscr{C}(B^j_{K_{n-1}}) \geq 2\vartheta K_{n-1}\Big\}\Bigg].$$

That is, the set $B^i_{K_{n-1}}$ is ϑ-bad but all blocks $B^j_{K_{n-1}}$ with which it does not intersect are ϑ-good. If we write

$$F_i = E_i \cap \{B_{K_n} \text{ is } \vartheta'\text{-bad}\},$$

then B_{K_n} being ϑ'-bad implies that either F_i occurs for some i or at least two disjoint blocks of scale $n-1$ are ϑ-bad, since otherwise B_{K_n} would be ϑ- and therefore also ϑ'-good. Consequently,

$$\begin{aligned} &p_\beta(K_n, \vartheta') \\ &\quad = p_\beta(C_n K_{n-1}, \vartheta') \leq \sum_{|i|=0}^{C_n-1} \widetilde{\mathbb{P}}^\beta(F_i) + \binom{C_n}{2} p_\beta(K_{n-1}, \vartheta)^2 \\ &\quad \leq p_\beta(K_{n-1}, \vartheta) \sum_{|i|=0}^{C_n-1} \widetilde{\mathbb{P}}^\beta(B_{C_n K_{n-1}} \text{ is } \vartheta'\text{-bad} \mid E_i) + 2C_n^2 p_\beta(K_{n-1}, \vartheta)^2. \end{aligned}$$

To finish the proof it therefore remains to bound the sum of the conditional probabilities by $1/100$. To this end, define

$$\mathscr{C}_i^- := \bigcup_{j=1-C_n}^{i-2} \mathscr{C}(B^j_{K_{n-1}}) \quad \text{and} \quad \mathscr{C}_i^+ := \bigcup_{j=i+2}^{C_n-1} \mathscr{C}(B^j_{K_{n-1}}).$$

Conditioned on E_i both sets $\mathscr{C}_i^-$ and $\mathscr{C}_i^+$ are connected sets. Further, if $i \in \{C_n - 2, C_n - 1\}$, then

$$\sharp\mathscr{C}_i^- \geq 2(C_n - 2)\vartheta K_{n-1} \geq 2\vartheta' C_n K_{n-1}$$

and hence B_{K_n} is ϑ'-good. The same holds for $\mathscr{C}_i^+$ if $i \in \{1 - C_n, 2 - C_n\}$. Therefore, the bad block and any neighbouring block cannot be the left-or the right-most ones in B_{K_n}. This then guarantees that $\mathscr{C}_i^-, \mathscr{C}_i^+ \neq \emptyset$. Further, if $\mathscr{C}_i^-$ and $\mathscr{C}_i^+$ are connected by a direct edge, it is the case that

$$\begin{aligned}\sharp\mathscr{C}_i^- + \sharp\mathscr{C}_i^+ &\geq \vartheta K_{n-1}(C_n + i - 2) + \vartheta K_{n-1}(C_n - i - 2) \\ &= 2\vartheta K_{n-1}(C_n - 2) \geq 2\vartheta' C_n K_{n-1},\end{aligned}$$

and B_{K_n} is again ϑ'-good, see Figure 2.6. Writing

$$A_i := \left[\bigcap_{\substack{|j|=0 \\ j \notin \{i-1,i,i+1\}}}^{C_n-1} \left\{ \sharp\mathscr{C}(B^j_{K_{n-1}}) \geq 2\vartheta K_{n-1} \right\} \right]$$

so that $E_i = \left\{ \sharp\mathscr{C}(B^i_{K_{n-1}}) < 2\vartheta K_{n-1} \right\} \cap A_i$ we have

$$\widetilde{\mathbb{P}}^\beta(B_{K_n} \text{ is } \vartheta'\text{-bad} \mid E_i) \leq \widetilde{\mathbb{P}}^\beta(\mathscr{C}_i^- \not\sim \mathscr{C}_i^+ \mid A_i)$$

since $\mathscr{C}_i^-$ and $\mathscr{C}_i^+$ do not share any vertex or edge with the bad block $B^i_{K_{n-1}}$ or its adjacent blocks. Now, A_i is an increasing event in the sense of Section 2.3 and $\{\mathscr{C}_i^- \not\sim \mathscr{C}_i^+\}$ is a decreasing event in the sense that $-\mathbb{1}_{\{\mathscr{C}_i^- \not\sim \mathscr{C}_i^+\}}$ is an increasing function. Hence, by the FKG-inequality (2.6) we have

$$\widetilde{\mathbb{E}}^\beta \mathbb{1}_{A_i} \mathbb{1}_{\{\mathscr{C}_i^- \not\sim \mathscr{C}_i^+\}} \leq \widetilde{\mathbb{P}}^\beta(A_i) \widetilde{\mathbb{P}}^\beta\{V_\ell^n(\vartheta^*) \not\sim V_r^n(\vartheta^*)\}$$

where we have used $\sharp\mathscr{C}_i^- \geq \sharp V_\ell^n(\vartheta^*)$ and $\sharp\mathscr{C}_i^+ \geq \sharp V_r^n(\vartheta^*)$, since $\vartheta' > \vartheta^*$, together with

$$\min\{|x - y| : \mathbf{x} \in \mathscr{C}_i^-, \mathbf{y} \in \mathscr{C}_i^+\} \leq \min\{|x - y| : \mathbf{x} \in V_\ell^n(\vartheta^*), \mathbf{y} \in V_r^n(\vartheta^*)\}.$$

Combining Lemma 2.15, Assumption (A1) of Proposition 2.13 and (2.12), we can choose $\mu \in (0, 1/2)$ such that for sufficiently large $C_1 = K_1 = K$ it holds

$$\begin{aligned}&\sum_{|i|=0}^{C_n-1} \widetilde{\mathbb{P}}^\beta(B_{C_n K_{n-1}} \text{ is } \vartheta'\text{-bad} \mid E_i) \\ &\qquad \leq 2n^3 C_1 \widetilde{\mathbb{P}}^\beta\{V_\ell^n(\vartheta^*) \not\sim V_r^n(\vartheta^*)\}\end{aligned}$$

$$\leq 2n^3 C_1 \Big[\exp\Big(- Cv^2 \int_{[v^{\mu-1}, 1-v^{\mu-1}]^2} \mathrm{d}(t,s)\ \rho\big(\beta^{-1} g(s,t) a K_n\big)\Big) + K_{n-1}^{1-\mu} \exp(-cK_{n-1}^{\mu})\Big]$$
$$\leq \frac{1}{100}.$$

□

Lemma 2.17. *Let η be a point process such that assumption (a) of Definition 2.11 is fulfilled. Then for every kernel g, every profile-function ρ and every $\vartheta \in (0,1)$, there exist constants $M > 0$, $\alpha > 0$, and $\kappa > 0$ such that for all $K = C_1 > M$ and $\beta > \alpha C_1$, it holds*

$$p_\beta(\vartheta, C_1) \leq \exp(-\kappa C_1).$$

Proof. Denote by $ER_n(q)$ an Erdös-Renyi-graph on n vertices with edge probability q and denote its law by P_n^q. If $\lambda > 1$, then $ER_n(\lambda/n)$ is supercritical, i.e. for all $\varepsilon_1 > 0$, there exists $c > 0$ and $N(\varepsilon_1, \lambda) > 0$ such that

$$\mathrm{P}_n^{\lambda/n}\Big\{\sharp\mathscr{C}(ER_n(\lambda/n)) > cn\Big\} \geq 1 - \varepsilon_1, \quad n \geq N(\varepsilon_1, \lambda), \tag{2.15}$$

where $\mathscr{C}(ER_n(\lambda/n))$ denotes the largest connected component of the graph $ER_n(\lambda/n)$ [42]. The idea is to compare this behaviour with the behaviour of the finite graph in the finite block B_{C_1} by making use of the evenly spaced property. We assume without loss of generality that $\rho(1) > 0$. We further assume that g is bounded and remark on the unbounded case below. Let a_1 be the constant from the evenly spaced condition and fix $\beta > a_1 ||g||_\infty C_1$. Then, for all $\mathbf{X}_i, \mathbf{X}_j \in B_{C_1}$, we have

$$\widetilde{\mathbb{P}}^\beta\{\mathbf{X}_i \sim \mathbf{X}_j\} \geq \rho(1)$$

and we focus on the subgraph on B_{C_1} where only the edges with marks smaller than $\rho(1)$ are present which is now independent of vertex marks and locations. For a fixed $\lambda > 1$ set c accordingly as above, fix $\varepsilon_2 < \vartheta/c$ and choose C_1 large enough such that

$$2\varepsilon_2 C_1 \rho(\alpha) \geq \lambda \text{ and } \lfloor 2\varepsilon_2 C_1 \rfloor \geq N(\varepsilon_1, \lambda).$$

Denote by H the subgraph on the vertices $\{\mathbf{X}_0, \dots, \mathbf{X}_{\lfloor 2\varepsilon_2 C_1 \rfloor}\} \subset B_{C_1}$. By (2.15), we have

$$\begin{aligned}\widetilde{\mathbb{P}}^\beta\Big\{\sharp\mathscr{C}(H) > c \cdot 2\varepsilon_2 C_1\Big\} &\geq \mathrm{P}^\lambda_{2\varepsilon_2 C_1}\left\{\sharp\mathscr{C}\Big(ER_{\lfloor 2\varepsilon_2 C_1\rfloor}(\tfrac{\lambda}{2\varepsilon_2 C_1})\Big) > c \cdot 2\varepsilon_2 C_1\right\} \\ &\geq 1 - \varepsilon_1.\end{aligned}$$

On $\{\sharp\mathscr{C}(H) > c\cdot 2\varepsilon_2 C_1\}$, the block B_{C_1} is ϑ-good if enough of the remaining vertices in $B_{C_1} \backslash H$ are connected to $\mathscr{C}(H)$. Each such remaining vertex is connected to $\mathscr{C}(H)$ with a probability of at least

$$q := q(C_1) := 1 - (1 - \rho(1))^{2\varepsilon_2 C_1}.$$

For $\xi > (\vartheta - c\varepsilon_2)/(1-\varepsilon_2)$ and C_1 large enough such that $q > \xi$, we have by writing $F_{\mathrm{Bin}(n,p)}$ for the distribution function of a binomial random variable with parameters n and p that

$$\begin{aligned}&\widetilde{\mathbb{P}}^\beta\Big\{\sharp\mathscr{C}(B_{C_1}) > 2\vartheta C_1\Big\} \\ &\qquad\geq (1-\varepsilon_1)\widetilde{\mathbb{P}}^\beta\Big(\sharp\mathscr{C}(B_{C_1}) > 2\vartheta C_1 \;\Big|\; \sharp\mathscr{C}(H) > c(2\varepsilon_2 C_1)\Big) \\ &\qquad\geq (1-\varepsilon_2)\Big(1 - F_{\mathrm{Bin}(2(1-\varepsilon_2)C_1, q)}\Big(2\xi(1-\varepsilon_2)C_1\Big)\Big) \\ &\qquad\geq 1 - \exp(-\kappa C_1),\end{aligned}$$

for some $\kappa > 0$ by a standard Chernoff bound.

If g is not bounded we can instead do the following. We fix a small $\varepsilon > 0$ and only consider vertices with marks smaller than $1-\varepsilon$ and therefore each vertex is removed independently with probability ε due to independence of marks and locations. However, the new block B_{C_1} still consists of order $(1-\varepsilon)C_1$ vertices with an error term exponentially small in C_1 by Chernoff's bound. Furthermore, the thinned process η is still evenly spaced and we can repeat the proof above since it holds that $g(s,t) \le g(1-\varepsilon, 1-\varepsilon) < \infty$ for all remaining marks s and t. $\square$

Finalising the proof of Proposition 2.13. We are now ready to prove Proposition 2.13 which we do following the arguments of the proof of Theorem 1(i) of [24] in the following lemma.

Lemma 2.18. *Let the assumptions of Proposition 2.13 be fulfilled. Then there exist* $\beta_c \in (0,\infty)$ *such that*

$$\widetilde{\mathbb{P}}^\beta\{\mathbf{0} \leftrightarrow \infty\} \ge \tfrac{3}{8}$$

for all $\beta > \beta_c$.

Proof. Fix $\vartheta^* \in (3/4, 1)$ and $\vartheta_1 \in (\vartheta^*, 1)$. Choose $K = C_1$ and afterwards β both large enough, such that the Lemmas 2.16 and 2.17 hold. Recall also that $C_n = n^3 K$, $K_1 = C_1$ and $K_n = C_n K_{n-1}$. Define $\vartheta_n := \vartheta_1 - 2/C_{n+1}$ for $n \ge 2$. Since the assumptions of Lemma 2.16 are satisfied, we have that $\vartheta_n > \vartheta^*$ for all n. We have by Lemma 2.17 that $p_\beta(C_1, \vartheta_1) \le (400 C_1^2)^{-1}$, and by Lemma 2.16 that

$$p_\beta(K_n, \vartheta_n) \le \frac{1}{100} p_\beta(K_{n-1}, \vartheta_{n-1}) + 2C_n^2 p_\beta(K_{n-1}, \vartheta_{n-1})^2, \quad \forall n \ge 2.$$

Inductively, this yields that $p_\beta(K_n, \vartheta_n) \le (400 C_n^2)^{-1}$. Hence, we have

$$\widetilde{\mathbb{P}}^\beta\{B_{K_n} \text{ is } \vartheta_n\text{-good}\} \ge 1 - \tfrac{1}{400 C_n^2} \ge \tfrac{1}{2}.$$

We derive from this that

$$\begin{aligned} \tfrac{3}{4} K_n &\le \tfrac{3}{4}(2K_n) \widetilde{\mathbb{P}}^\beta\{B_{K_n} \text{ is } {}^3\!/\!_4\text{-good}\} \\ &\le \widetilde{\mathbb{E}}^\beta\Big[\sharp \mathscr{C}(B_{K_n}) \mathbb{1}_{\{B_{K_n} \text{ is } {}^3\!/\!_4\text{-good}\}}\Big] \\ &\le 2K_n \widetilde{\mathbb{P}}^\beta\{\exists \text{ a cluster of size at least } \tfrac{3}{2} K_n\}. \end{aligned}$$

Dividing both sides by $2K_n$ and then sending $n \to \infty$ together with the translation invariance yields the desired result. $\square$

It remains in this section to prove Corollary 2.14 and specifically the existence of a component of linear size in the constructed graph sequence for large enough β if the assumption of Theorem 1.11(a) is fulfilled.

Proof of Corollary 2.14. Consider the sequence $(\mathscr{G}_{2m}(\beta) : m \geq m_0)$, where m_0 will be determined below. Let $(R_i)_{i\in\mathbb{N}}$ be a sequence of natural numbers sufficiently large for Lemmas 2.16 and 2.17 to hold for $K = R_i$, such that

$$\bigcup_i \{2K_n : n \in \mathbb{N},\ K = R_i\}$$

contains all sufficiently large natural numbers, and let m_0 be an arbitrary natural number contained in this set. Note that the sequence $(R_i)_{i\in\mathbb{N}}$ can be chosen to be simply all sufficiently large natural numbers.

Consider now the sequence $(\mathscr{G}_{2K_n}(\beta) : n \in \mathbb{N})$ for fixed $K \in (R_i)_{i\in\mathbb{N}}$. Since the assumption of Theorem 1.11(a) on the kernel g and the profile ρ is fulfilled, Assumption (A1) in particular is also satisfied. Now, the $2K_n$ points can be seen as embedded into the interval $(-K_n, K_n)$ and therefore the construction above (for the block B_{K_n}) guarantees that the proportion of the largest connected component of the graphs $\mathscr{G}_{2K_n}(\beta)$ is bounded from below by $3/4$ with a probability larger than $3/8$ and this holds uniformly in n. This holds furthermore uniformly along the sequence R_i and therefore uniformly for all intervals $(-m, m)$ with $m > m_0$. Our claim then follows by ergodicity. $\square$

2.7. Non existence of a supercritical phase in dimension one

We start this section with stating a Proposition comparably to the previous section which then implies Theorem 1.11(b). We denote by $f(a+)$ the limit $\lim_{x\downarrow a} f(a)$.

Proposition 2.19 (Finite Components)**.** *Let η be a simple and ergodic Point process in dimension $d = 1$ fulfilling condition (b) of Definition 2.11 and let $\mathcal{X}_0$ and $\mathcal{U}_0$ be defined as above in Section 2.1. Let g be a kernel-function and ρ be a profile-function for which $\rho(0+) < 1$. Let $(\mathcal{G}^\beta = \mathcal{G}^{\beta,\rho,g}(\mathcal{X}_0,\mathcal{U}_0) : \beta > 0)$ be the associated weight-dependent random connection model. Assume that there exists $\mu \in (0, {}^1\!/_2)$ such that*

$$\sum_{n\in\mathbb{N}} 2^{2n} \int_{2^{-(1+\mu)n}}^{1} \mathrm{d}s \int_{2^{-(1+\mu)n}}^{1} \mathrm{d}t \; \rho(g(s,t)2^n) < \infty. \tag{A2}$$

Then $\beta_c = \infty$.

The assumption $\rho(0+) < 1$ in Proposition 2.19 is a technical requirement needed in the proof below and can essentially be viewed as a continuum version of the analogous requirement in long-range percolation on $\mathbb{Z}$ that not all nearest-neighbour-edges be present. If however η is a homogeneous Poisson process, the additional condition $\rho(0+) < 1$ can be dropped which is proved at the end of Section 2.7.1..

Corollary 2.20. *Let η be a standard Poisson point process on $\mathbb{R}$ and let $\mathcal{X}_0$ and $\mathcal{U}_0$ be as above. Let g be a kernel and ρ be a profile-function such that assumption (A2) of Proposition 2.19 is fulfilled and consider the weight-dependent random connection model $(\mathcal{G}^\beta = \mathcal{G}^{\beta,\rho,g}(\mathcal{X}_0,\mathcal{U}_0) : \beta > 0)$. Then $\beta_c = \infty$.*

2.7.1. Proof of Proposition 2.19

Throughout the section, the vertex locations are assumed to belong to a simple and ergodic point process η that fulfills assumption (b) of Definition 2.11. We consider edges that connect a vertex left of the origin to a vertex right of the origin and say that such an edge *crosses the origin*. We will show that the probability that there are no edges crossing the origin is bounded away from zero, and since by stationarity the same bound holds for crossings of any arbitrary point, we get by ergodicity that there exist infinitely many points that are not crossed by an edge and all components must be finite. Technically, the root $\mathbf{X}_0$ is neither right or left but at the origin as we work on the palm version of η which allows the canonical ordering of the vertices. However, we declare without loss of generality $\mathbf{X}_0$ to be on the right of the origin.

We define the disjoint sets of vertices

$$\Gamma_k^\ell := \{\mathbf{X}_{-2^k}, \dots, \mathbf{X}_{-1}\}, \qquad \Gamma_k^{\ell\ell} := \{\mathbf{X}_{-2^{k+1}}, \dots, \mathbf{X}_{-2^k-1}\}$$
$$\Gamma_k^r := \{\mathbf{X}_0, \dots, \mathbf{X}_{2^k-1}\}, \qquad \Gamma_k^{rr} := \{\mathbf{X}_{2^k}, \dots, \mathbf{X}_{2^{k+1}-1}\}$$

for each $k \in \mathbb{N}$. We say that a *crossing of the origin* occurs at stage

$k = 1$, if any edge connects the sets $\Gamma_1^\ell \cup \Gamma_1^{\ell\ell}$ and $\Gamma_1^r \cup \Gamma_1^{rr}$ or at

$k \geq 2$, if any edge connects either $\Gamma_k^{\ell\ell}$ to Γ_k^{rr}, $\Gamma_k^{\ell\ell}$ to Γ_k^r or Γ_k^ℓ to Γ_k^{rr}. Note that any edges between Γ_k^ℓ and Γ_k^r have by necessity already been considered at a smaller scale.

We denote by $\chi(k) \in \{0,1\}$ the indicator that a crossing of the origin occurs at stage $k \in \mathbb{N}$. The event that there is no edge crossing the origin is then given by $\bigcap_k \{\chi(k) = 0\}$. Note that these events are all decreasing and therefore positively correlated, implying that

$$\mathbb{P}_\beta\Big(\bigcap_{k\in\mathbb{N}} \{\chi(k) = 0\}\Big) \geq \prod_{k\in\mathbb{N}} \mathbb{P}_\beta\{\chi(k) = 0\}.$$

To show that the product on the right-hand side is bounded away from zero, it suffices to show the equivalent statement that

$$\sum_{k\in\mathbb{N}} \mathbb{P}_\beta\{\chi(k)=1\}<\infty.$$

For $k\geq 2$ we have by symmetry,

$$\mathbb{P}_\beta\{\chi(k)=1\}=\mathbb{P}_\beta\Big\{\Gamma_k^{rr}\sim\Gamma_k^{\ell\ell}\Big\}+2\mathbb{P}_\beta\Big\{\Gamma_k^{\ell\ell}\sim\Gamma_k^{r}\Big\}\leq 3\mathbb{P}_\beta\Big\{\Gamma_k^{\ell\ell}\sim\Gamma_k^{r}\Big\}.$$

The following lemma shows that for profile-functions satisfying $\rho(0+)<1$ the probability of the right hand side is bounded by the summand occurring in Assumption (A2), which immediately implies Proposition 2.19.

Lemma 2.21. *Assume that ρ satisfies $\rho(0+)<1$. Then for all $\beta>0$, there exist constants $c>0$ and $K\in\mathbb{N}$ such that for all $k\geq K$, it holds*

$$\mathbb{P}^\beta\{\Gamma_k^{\ell\ell}\sim\Gamma_k^{r}\}\leq c2^{2k}\int_{2^{-(1+\mu)k}}^1 \mathrm{d}s\int_{2^{-(1+\mu)k}}^1 \mathrm{d}t\ \rho(\beta^{-1}g(s,t)2^k).$$

Proof. We begin by modifying the definition of μ-regularity, since we are now interested in upper bounds on connection probabilities. Throughout this proof, we say a set Γ_k^o, $o\in\{\ell,\ell\ell,r,rr\}$ is *μ-regular* if, for all $i\in\{1,\dots,\lceil 2^{k(1-\mu)}\rceil\}$,

(i) $\sum_{T:(X,T)\in\Gamma_k^o}\mathbb{1}_{\{T\leq\lceil 2^{-(1+\mu)k}\rceil\}}=0,$

(ii) $\sum_{T:(X,T)\in\Gamma_k^o}\mathbb{1}_{\{T\leq i/\lceil 2^{(1-\mu)k}\rceil\}}\leq\frac{i2^{k+1}}{\lceil 2^{(1-\mu)k}\rceil}.$

Note, that assumption (i) is fulfilled with a probability of order $1-2^{-k\mu}$ for large k. For assumption (ii), we can again use Chernoff's bound to deduce

$$\mathbb{P}^\beta\Big\{\sum_{T:(X,T)\in\Gamma_k^o}\mathbb{1}_{\{T\leq i/\lceil 2^{(1-\mu)k}\rceil\}}\leq\frac{i2^{k+1}}{\lceil 2^{(1-\mu)k}\rceil}\Big\}\geq 1-2^{(1-\mu)k}\exp(-c2^{\mu k}).$$

Hence, the event

$$\bigcap_{o\in\{\ell,\ell\ell,r,rr\}} \{\Gamma_k^o \text{ is } \mu\text{-regular}\}.$$

holds with a probability at least $1-\varepsilon$ for any $\varepsilon\in(0,1)$, if k is large enough. Together with the evenly spaced condition (b) from Definition 2.11, we have that the event

$$E_k := \bigcap_{o\in\{\ell,\ell\ell,r,rr\}} \{\Gamma_k^o \text{ is } \mu\text{-regular}\} \cap \{|X_{-2^{k+1}} - X_{2^k}| > a_2 2^k\}$$

holds with probability greater than $1-\varepsilon$ for large enough k. We now argue as in the proof of Lemma 2.15. A point configuration $\omega = (x_i, i\in\mathbb{Z})$ is *properly spaced* if it satisfies the spacing condition of Definition 2.11(b), namely $|X_{-2^{n+1}} - X_{2^n}| < a_2 2^n$. The μ-regularity is measurable with respect to vertex marks only and thus independent of the vertex locations and edge marks. Thus, denoting by $\mathbb{P}^\beta_{\eta=\omega}$ the law induced on vertex and edge mark configurations given a fixed properly spaced vertex locations ω, we obtain

$$\begin{aligned}
&\mathbb{E}^\beta_{\eta=\omega}\Big[\mathbb{1}_{\{\Gamma_k^{\ell\ell}\nsim\Gamma_k^r\}}\mathbb{1}_{E_k}\Big] \\
&\quad\geq \mathbb{E}^\beta_{\eta=\omega}\Bigg[\mathbb{1}_{E_k}\prod_{\substack{(x,T)\in\Gamma_k^{\ell\ell}\\(y,S)\in\Gamma_k^r}}(1-\rho\Big(\beta^{-1}g(S,T)a2^k\Big)\Bigg] \\
&\quad\geq \mathbb{E}^\beta_{\eta=\omega}\Bigg[\mathbb{1}_{E_k}\exp\Big(-c\sum_{\substack{(x,T)\in\Gamma_k^{\ell\ell}\\(y,S)\in\Gamma_k^r}}\rho(\beta^{-1}g(S,T)a2^k)\Big)\Bigg] \\
&\quad= \mathbb{E}^\beta_{\eta=\omega}\Bigg[\mathbb{1}_{E_k}\exp\Big(-c\int_0^1 2^k F_{\Gamma_k^{\ell\ell}}(\mathrm{d}s)\int_0^1 2^k F_{\Gamma_k^r}(\mathrm{d}t)\ \rho(\beta^{-1}g(s,t)a2^k)\Big)\Bigg]
\end{aligned} \tag{2.16}$$

for some constant $c>0$, where the second to last inequality follows from the fact that $\rho(x)\leq\rho(0+)<1$ for all $x>0$ by assumption. As before, $F_{\Gamma_k^{\ell\ell}}$ denotes the empirical distribution function of the vertex marks in $\Gamma_k^{\ell\ell}$. By the μ-regularity of the marks, we have similarly as done in (2.13) that

$$2^k F_{\Gamma_k^{\ell\ell}}(t) \leq 2^k \sum_{j=1}^{\lceil 2^{k(1+\mu)}\rceil} \frac{2j}{\lceil 2^{k(1+\mu)}\rceil}\mathbb{1}_{\{j-1<t\lceil 2^{k(1+\mu)}\rceil\leq j\}}$$

$$\leq \frac{\left\lceil t\lceil 2^{k(1+\mu)}\rceil\right\rceil}{\lceil 2^{k(1+\mu)}\rceil}2^k \leq c' 2^k(t+2^{-k(1+\mu)})$$

for some $c' \geq 2$ uniformly for all μ-regular vertex mark configurations. Plugging this into (2.16), we get

$$\begin{aligned}\mathbb{P}_\beta^{\eta=\omega}(\{\Gamma_k^{\ell\ell} \not\sim \Gamma_k^r\} \cap E_k)& \\ \geq \exp\Big(-c2^{2k}\int_{2^{-k(1+\mu)}}^1 \mathrm{d}s \int_{2^{-k(1+\mu)}}^1 \mathrm{d}t\ \rho\Big(\beta^{-1}g(s,t)a2^k\Big)\Big)\mathbb{P}_\beta^{\eta=\omega}(E_k).\end{aligned}$$

Since this estimate is uniform in the properly spaced configuration ω for large enough k, it follows that we have

$$\begin{aligned}\mathbb{P}_\beta\{\Gamma_k^{\ell\ell} \sim \Gamma_k^r\} &\leq 1-\exp\Big(-c2^{2k}\int_{2^{-k(1+\mu)}}^1 \mathrm{d}s \int_{2^{-k(1+\mu)}}^1 \mathrm{d}t\ \rho\Big(\beta^{-1}g(s,t)a2^k\Big)\Big) \\ &\asymp 2^{2k}\int_{2^{-k(1+\mu)}}^1 \mathrm{d}s \int_{2^{-k(1+\mu)}}^1 \mathrm{d}t\ \rho\Big(\beta^{-1}g(s,t)a2^k\Big),\end{aligned}$$

which concludes the proof. □

It remains to show that Corollary 2.20 holds when the vertex set is given by a Poisson process without requiring $\rho(0+) < 1$, an assumption which is crucial when doing the calculation leading up to (2.16).

Proof of Corollary 2.20. Let η be a Poisson point process of intensity $\lambda > 0$. In this case, β can be seen as a scaling parameter of the Euclidean distance between the vertices and therefore varying β is equivalent to varying the intensity of the Poisson process, cf. Remark 1.9 (iii).

We now fix an arbitrarily $\beta > 0$ and show that no infinite component exists in $\mathcal{G}^{\beta,\rho,g}$ constructed on the Poisson process η, or rather on its Palm version η_0. By Poisson thinning, we can interpret $\mathcal{G}^{\beta,\rho,g}$ as the graph resulting from i.i.d. Bernoulli site percolation of the graph $\mathcal{G}^{\beta/p,\rho,g}$ for some arbitrary $p < 1$. We perform Bernoulli bond percolation on the graph $\mathcal{G}^{\beta/p,\rho,g}$ with retention parameter $p' \in (p,1)$. That is, each edge is independently removed with probability $1-p'$. By construction, this coincides with constructing the graph where the profile-function ρ is replaced by $p'\rho$. Hence, we are working with the graph $\mathcal{G}^{\beta/p,\rho,g}$, its bond percolated version $\mathcal{G}^{\beta/p,p'\rho,g}$ and its site percolated version $\mathcal{G}^{\beta,\rho,g}$. Since vertex percolation removes at least as many edges from the graph as bond percolation and $p' > p$, we have

$$\mathbb{P}^{\beta,\rho,g}\{\mathbf{0} \leftrightarrow \infty\} \leq \mathbb{P}^{\beta/p,p'\rho,g}\{\mathbf{0} \leftrightarrow \infty\},$$

see e.g. [37]. Note now that $p'\rho(0+) < 1$ by construction and assumption (A2) is still fulfilled, so the right hand side equals zero by Proposition 2.19, proving the claim. □

2.7.2. Proof of Corollary 1.13

In this short section, we calculate the effective decay exponent δ_{eff} for our examples and thus proving Corollary 1.13. For all our calculations we assume without loss of generality that

$$\rho(x) \asymp x^{-\delta} \text{ for } x > 1.$$

for $\delta \in (2, \infty)$. The results for the hard models follow either by a limit $\delta \to \infty$ or by Remark 1.12(ii). We start with the soft Boolean model and the strong-kernel and get for $\gamma > {}^1\!/\!_\delta$

$$\int_{1/n}^1 \mathrm{d}s \int_{1/n}^1 \mathrm{d}t \rho(g^{\text{str}}(s,t)n) \asymp n^{-\delta} \int_{1/n}^1 \mathrm{d}s\ s^{-\delta\gamma} = n^{-\delta(1-\gamma)-1}$$

and

$$1 + \delta(1-\gamma) \begin{cases} > 2, & \gamma < 1 - \frac{1}{\delta}, \\ = 2, & \gamma = 1 - \frac{1}{\delta}, \\ < 2, & \gamma > 1 - \frac{1}{\delta}. \end{cases}$$

For the preferential attachment-kernel, we get for $\gamma < 1 - {}^1\!/\!_\delta$

$$\begin{aligned}\int_{1/n}^{1}\mathrm{d}s\int_{1/n}^{1}\mathrm{d}t\rho(g^{\mathrm{pa}}(s,t)n) &\asymp \int_{1/n}^{1}\mathrm{d}s\int_{s}^{1}\mathrm{d}t\ s^{-\gamma\delta}t^{-\delta(1-\gamma)}n^{-\delta} \\ &\asymp n^{-\delta}\int_{1/n}^{1}\mathrm{d}s\ s^{1-\delta}\asymp n^{-2}.\end{aligned}$$

Finally, for the non scale-free weak kernel model, we get for $1>\alpha>{}^2\!/\!_\delta$

$$\begin{aligned}\int_{1/n}^{1}\mathrm{d}s\int_{1/n}^{1}\mathrm{d}t\rho(g^{\mathrm{weak}}(s,t)n) &\asymp n^{-\delta}\int_{1/n}^{1}\mathrm{d}s\int_{1/n}^{s}\mathrm{d}ts^{-\alpha\delta}\asymp\int_{1/n}^{1}\mathrm{d}ss^{1-\alpha\delta} \\ &\asymp n^{-2-\delta(1-\alpha)}\end{aligned}$$

and $2+\delta(1-\alpha)>2$.

CHAPTER 3

The age-based spatial preferential attachment

This chapter is devoted to the age-based spatial preferential attachment model, discussed in Section 1.1, and the proof of Theorem 1.5. To this end, we extend the formal construction of Section 2.1 and recap the rescaling arguments of Section 1.2.1. We then prove a law of large numbers connecting the age-based spatial preferential attachment model with the age-dependent random connection model. We use this to prove Theorem 1.5. The proofs of the Sections 3.1 and 3.2 can be found in [32]. The proofs of Section 3.3 can be found in [36] and are in particular based on arguments of [45] as outlined in Section 1.3.

3.1. Construction and weak local limit

Recall the construction of the weight-dependent random connection model as a functional of the points of a marked Poisson point process and edge marks. We extend the notion of $\mathcal{X}$ to be a unit intensity Poisson point process on $\mathbb{R}^d \times (0,\infty)$. The points of $\mathcal{X}$ play the role of the vertices of the age-based spatial preferential attachment model and we say a vertex $\mathbf{x} = (x,s) \in \mathcal{X}$ is born at time s and placed at position x. We say that (x,s) is *older* than (y,u) if $s < u$. We denote by

$$E(\mathcal{X}) := \{(\mathbf{x},\mathbf{y}) \in \mathcal{X} \times \mathcal{X} : \mathbf{y} \text{ older than } \mathbf{x}\}$$

the set of *potential edges*. Given $\mathcal{X}$ we denote by $\mathcal{U}$ a family of independent random variables, uniformly distributed on $(0,1)$, indexed by the set of potential edges. We denote these variables by $U_{\mathbf{x},\mathbf{y}}$ and again call them *edge marks*. For $t > 0$ write $\mathcal{X}_t$ for $\mathcal{X} \cap \mathbb{T}_1^d \times (0,t]$, the set of vertices on the unit torus already born at time t. Also, we write $\mathcal{U}_t$ for the restriction of $\mathcal{U}$ to indices in $\mathcal{X}_t \times \mathcal{X}_t$. Fix $\beta > 0$, a profile function ρ and $\gamma \in (0,1)$ and define the graph $\mathcal{G}_1^{\beta,\rho,\gamma}(\mathcal{X}_t,\mathcal{U}_t)$ with vertex set $\mathcal{X}_t$ and two vertices $\mathbf{x} = (x,s)$, $\mathbf{y} = (y,u)$ with $s > u$ are connected by an edge if and only if

$$U_{\mathbf{x},\mathbf{y}} \leq \rho\Big(\tfrac{s \cdot d_1(x,y)^d}{\beta(s/u)^\gamma}\Big).$$

Observe that the graph sequence $(\mathcal{G}_1^{\beta,\rho,\gamma}(\mathcal{X}_t,\mathcal{U}_t) : t \geq 0)$ has the law of the age-based spatial preferential attachment model $(\mathscr{G}_t : t \geq 0)$ constructed according to (1.2). Similarly, we define $\mathcal{X}^t$ as the restriction of $\mathcal{X}$ to $\mathbb{T}_t^d \times (0,1]$ and $\mathcal{U}^t$ the restriction of $\mathcal{U}$ to indices in $\mathcal{X}^t \times \mathcal{X}^t$. Then the graph $\mathcal{G}_t^{\beta,\rho,\gamma}(\mathcal{X}^t,\mathcal{U}^t)$ is defined through the vertex set $\mathcal{X}^t$ and $\mathbf{x} = (x,s)$, $\mathbf{y} = (y,u)$ with $s > u$ are connected by an edge if and only if

$$U_{\mathbf{x},\mathbf{y}} \leq \rho\Big(\tfrac{s \cdot d_t(x,y)^d}{\beta(s/u)^\gamma}\Big).$$

Then the law of $(\mathcal{G}_t^{\beta,\rho,\gamma}(\mathcal{X}^t,\mathcal{U}^t) : t \geq 0)$ is the same as of the family $(\mathscr{G}^t : t \geq 0)$ of Section 1.2.1. Hence, from now on we identify $\mathscr{G}_t = \mathcal{G}_1^{\beta,\rho,\gamma}(\mathcal{X}_t,\mathcal{U}_t)$ and $\mathscr{G}^t = \mathcal{G}_t^{\beta,\rho,\gamma}(\mathcal{X}^t,\mathcal{U}^t)$. All graphs are constructed on the probability space carrying the Poisson process $\mathcal{X}$ and the edge marks $\mathcal{U}$, whose joint probability measure and expectation we denote by $\mathbb{P}$ and $\mathbb{E}$. Recall further the rescaling map h_t of Section 1.2.1. It works canonically on the set $\mathcal{X}_t$ as well as on $\mathcal{U}_t$ by $h_t(U)_{h_t(\mathbf{x}),h_t(\mathbf{y})} := U_{\mathbf{x},\mathbf{y}}$, and also on graphs with vertex set $\mathcal{X}_t$ by mapping points $\mathbf{x}$ to $h_t(\mathbf{x})$ and introducing an edge between $h_t(\mathbf{x})$ and $h_t(\mathbf{y})$ if and only if there is one between $\mathbf{x}$ and $\mathbf{y}$. By the arguments of Section 1.2.1 and (1.5) it follows

$$\mathcal{G}_t^{\beta,\rho,\gamma}(h_t(\mathcal{X}_t), h_t(\mathcal{U}_t)) = h_t(\mathcal{G}_1^{\beta,\rho,\gamma}(\mathcal{X}_t,\mathcal{U}_t)) = h_t(\mathscr{G}_t).$$

Since $h_t(\mathcal{X}_t)$ has the law of $\mathcal{X}^t$, it follows that the law of $\mathscr{G}^t$ is the same as the law of $h_t(\mathscr{G}_t)$, making the arguments for Figure 1.1 rigorous. Therefore, to study degree distribution, clustering and robustness for the network $(\mathscr{G}_t : t \geq 0)$ is the same as studying $(\mathscr{G}^t : t \geq 0)$ instead.

We start by showing that, almost surely, $t \mapsto \mathscr{G}^t$ converges locally to the graph $\mathscr{G}^\infty := \mathcal{G}_\infty^{\beta,\rho,\gamma}(\mathcal{X}^\infty,\mathcal{U}^\infty)$ which is the age-dependent random connection model. Recall that $\mathscr{G}^\infty$ is locally finite by Corollary 2.3.

Theorem 3.1. *Almost surely, the graph sequence $(\mathscr{G}^t : t \geq 0)$ converges to $\mathscr{G}^\infty$ in the sense that for each $\mathbf{x} \in \mathcal{X}^\infty$ the neighbours of $\mathbf{x}$ in $\mathscr{G}^t$ and $\mathscr{G}^\infty$ coincide for large t.*

Proof. We work conditionally on $\mathbf{x} = (x, s) \in \mathcal{X}^\infty$. Our aim is to show that there exists an almost surely finite random variable M such that, for all $t \in (0, \infty]$ and $\mathbf{y} \in \mathcal{X}^\infty$ with distance at least M from $\mathbf{x}$, the vertices $\mathbf{x}$ and $\mathbf{y}$ are not connected in $\mathscr{G}^t$. To this end, observe that the distance between $\mathbf{x}$ and any $\mathbf{y} \in \mathbb{T}_t^d$ can be up to $2\sqrt{d}|x|$ smaller than it would be in $\mathbb{R}^d$. Consider the model where the vertices within distance $2\sqrt{d}|x|$ of $\mathbf{x}$ are deleted from $\mathcal{X}^\infty$ and all the other vertices are moved towards $\mathbf{x}$ by a distance of $2\sqrt{d}|x|$. It is easy to see that all vertices $\mathbf{y} \in \mathcal{X}^\infty$, that are at least $2\sqrt{d}|x|$ away from $\mathbf{x}$ and connected to $\mathbf{x}$ in the finite graph $\mathcal{G}^t$ for some $t > 0$, are also linked to $\mathbf{x}$ in this new model. Furthermore, the degree of $\mathbf{x}$ is still almost surely finite. Hence, we define the random variable M as the distance of $\mathbf{x}$ to the furthest vertex it is linked to in this new model, plus $2\sqrt{d}|x|$. Then M is almost surely finite and, as for $t > |x| + M$ the vertices in $\mathcal{X}^\infty$ and in $\mathcal{X}^t$ within distance M from $\mathbf{x}$ coincide, the edges of $\mathbf{x}$ linking it to another vertex $\mathbf{y}$ that is at most M away coincide in $\mathscr{G}^t$ and $\mathscr{G}^\infty$ for sufficiently large t. $\square$

The above theorem only states the local convergence of the neighbourhood of vertices. Global results require a specific law of large numbers for the graphs rooted in a randomly chosen point. In order to do so, we add a point at the origin to our Poisson process denoting $\mathcal{X}_0 := \mathcal{X}^\infty \cup \{(0, U)\}$ where U is an independent, uniformly on $(0,1]$ distributed birth time and we denote $\mathbf{0} = (0, U)$. As before let $\mathcal{U}_0$ be a family of independent uniformly distributed random variables indexed by the potential edges in $\mathcal{X}_0$, and, for $0 < t \leq \infty$, let $\mathcal{X}_0^t = \mathcal{X}_0 \cap (\mathbb{T}_t^d \times (0,1])$ and denote by $\mathcal{U}_0^t$ the restriction of $\mathcal{U}_0$ to indices in $\mathcal{X}_0^t \times \mathcal{X}_0^t$. We define rooted graphs $\mathscr{G}_0^t := \mathcal{G}_t^{\beta,\rho,\gamma}(\mathcal{X}_0^t, \mathcal{U}_0^t)$ with the root being the vertex placed at the origin. Note that this is consistent with the construction in Section 2.1 and in particular $\mathscr{G}_0^\infty = \mathcal{G}^{\beta,\rho,g^{\mathrm{pa}}}(\mathcal{X}_0, \mathcal{U}_0)$. The graphs depend now on the Poisson process, the additional root and the extended edge marks. We denote their joint law and expectation by $\mathbb{P}_0$ and $\mathbb{E}_0$.

Definition 3.2. Consider a family $(H_t : t \geq 0)$ of non negative functionals H_t acting on the root of finite rooted graphs and the whole graph. We say such a family *belongs to the class* $\mathcal{H}_p$ for $p > 0$ if there exists a functional H_∞ acting on the root and the graph of an infinite rooted graph such that

(A) the random variable $H_t(\mathbf{0}, \mathscr{G}_0^t)$ converges in probability to the random variable $H_\infty(\mathbf{0}, \mathscr{G}_0^\infty)$ as $t \to \infty$,

(B) the uniform moment condition

$$\sup_{t \geq 0} \mathbb{E}_0 H_t(\mathbf{0}, \mathscr{G}_0^t)^p < \infty$$

holds and

(C) for every $t \in (0, \infty]$, the functional H_t is invariant under shifts

$$\theta_x^t : \mathbb{T}_t^d \times (0,1) \to \mathbb{T}_t^d \times (0,1), (y, s) \mapsto (y - x, s),$$

where θ_x^t acts on graphs canonically by shifting all vertices and placing an edge between shifted vertices if and only if there has been an edge between the original vertices in the original graph.

Theorem 3.3 (Weak law of large numbers)**.** *Suppose* $(H_t : t \geq 0) \in \mathcal{H}_p$ *for some* $p > 1$*. Then*

$$\lim_{t\to\infty} \frac{1}{t} \sum_{\mathbf{x}\in\mathcal{X}^t} H_t(\mathbf{x}, \mathscr{G}^t) = \mathbb{E}_0 H_\infty(\mathbf{0}, \mathscr{G}_0^\infty) \tag{3.1}$$

in probability as $t \to \infty$*.*

Remark 3.4. The proof can be found in [44, Theorem 7] where Jacob and Mörters prove the the weak law of large numbers for the degree-based model. Their arguments work mutatis mutandis in our setting. The theorem particularly allows that H_t additionally depend continuously on the ages of vertices and the length of edges as it is formulated for the graphs $\mathscr{G}^t$ on the rescaled marked space. The theorem is an adaption of [66, Theorem 2.1] by Penrose and Yukich. It shows that the age-dependent random connection model is indeed the weak local limit of the age-based spatial preferential attachment in the sense of Benjamini and Schramm [5] as outlined in Section 1.2.1.

With the law of large numbers at hand we are ready to prove Theorem 1.5. We prove the scale-freeness and clustering in Section 3.2 and robustness in Section 3.3.

3.2. Degree distribution, clustering and edge length distribution

In this section, we consider the degree distribution and clustering of the network. We further state some results about the empirical rescaled edge length. To this end, we frequently define functionals of the form $H_t(\mathbf{x}, \mathscr{G}^t)$ depending on a given vertex $\mathbf{x}$ of the graph $\mathscr{G}^t$ and the graph $\mathscr{G}^t$ to make use of the weak law of large numbers.

3.2.1. Degree distribution

In the original network $(\mathscr{G}_t : t \geq 0)$ a new vertex joins the graph and connects to already existing vertices. Afterwards, one can think of the vertex as waiting for new vertices connecting to it. We hence think of edges as oriented from young to old. In the graph family $(\mathscr{G}^t : t \geq 0)$ the birth times are represented by the vertices' marks. Consequently, we use the notation of Section 2.2.1 and write $\mathscr{N}_{\mathbf{x}}(t) := \mathscr{N}_{\mathbf{x}}(\mathscr{G}^t)$ for the neighbours of $\mathbf{x}$ in $\mathscr{G}^t$. For the older neighbours of $\mathbf{x}$ in $\mathscr{G}^t$, we write $\mathscr{N}_{\mathbf{x}}^{<}(t)$ and denote the outdegree by $\sharp\mathscr{N}_{\mathbf{x}}^{<}(t)$. The younger neighbours are denoted by $\mathscr{N}_{\mathbf{x}}^{>}(t)$ and the indegree by $\sharp\mathscr{N}_{\mathbf{x}}^{>}(t)$. If we work in the limiting graph, we simply write $\mathscr{N}_{\mathbf{x}} = \mathscr{N}_{\mathbf{x}}(\infty)$ for the neighbourhood of $\mathbf{x}$ in $\mathscr{G}^\infty$ and do so like wise for the younger and older neighbours.

We define the *empirical outdegree distribution* ν_t of the graph $\mathscr{G}^t$ by

$$\nu_t(k) := \frac{1}{t} \sum_{\mathbf{x} \in \mathcal{X}^t} \mathbb{1}_{\{\sharp\mathscr{N}_{\mathbf{x}}^{<}(t) = k\}} \qquad \text{for } k \in \mathbb{N}_0,$$

and note that (for convenience) we have normalised ν_t so that is mass converges to one without necessarily equal to one for small t.

Theorem 3.5. *For any function $g : \mathbb{N}_0 \to [0, \infty)$ growing no faster than exponentially, we have*

$$\frac{1}{t} \sum_{\mathbf{x} \in \mathcal{X}^t} g(\sharp \mathscr{N}_{\mathbf{x}}^{<}(t)) = \int g \, \mathrm{d}\nu_t \longrightarrow \int g \, \mathrm{d}\nu$$

in probability, as $t \to \infty$, where ν is the Poisson distribution with parameter $\beta/(1-\gamma)$.

Proof. We define the functionals $H_t(\mathbf{x}, \mathscr{G}^t) = g(\sharp \mathscr{N}_{\mathbf{x}}^{<}(t))$ for all $t \in (0, \infty]$. Since H_t only depends on the neighbourhood of $\mathbf{x}$, assumption (A) of Definition 3.2 is fulfilled by Theorem 3.1. Moreover, $\mathscr{N}_{\mathbf{x}}^{<}(t)$ is Poisson distributed with a bounded parameter by the same arguments as in Proposition 2.1, and the uniform moment condition (B) is fulfilled as long as g is not growing faster than exponentially. As (C) is trivially fulfilled, we infer the result from the weak law of large numbers, Theorem 3.3, and Proposition 2.1. $\square$

Define the *empirical indegree distribution* μ_t of the graph $\mathscr{G}^t$ by

$$\mu_t(k) = \frac{1}{t} \sum_{\mathbf{x} \in \mathcal{X}^t} \mathbb{1}_{\{\sharp \mathscr{N}_{\mathbf{x}}^{>}(t) = k\}}.$$

Theorem 3.6. *For any function $g : \mathbb{N}_0 \to [0, \infty)$ growing no faster than linearly, we have*

$$\frac{1}{t} \sum_{\mathbf{x} \in \mathcal{X}^t} g(\sharp \mathscr{N}_{\mathbf{x}}^{>}(t)) = \int g \, \mathrm{d}\mu_t \longrightarrow \int g \, \mathrm{d}\mu$$

in probability, as $t \to \infty$, where μ is mixed Poisson distributed with mixing density f as in (2.4).

Proof. We define $H_t(\mathbf{x}, \mathscr{G}^t) = g(\sharp \mathscr{N}_{\mathbf{x}}^{>}(t))$ for all $t \in (0, \infty]$. Assumption (A) of Definition 3.2 holds with the same arguments as above. As

$\sharp\mathscr{N}_{\mathbf{x}}^{>}(t)$ is dominated by $\sharp\mathscr{N}_{\mathbf{x}}^{>}$ whose distribution μ has tails (calculated in Lemma 2.2) that vanishes fast enough to ensure the uniform moment condition (B), the result follows again by Theorem 3.3 and Proposition 2.1. □

Corollary 3.7. *The age-based spatial preferential attachment network* $(\mathscr{G}_t : t \geq 0)$ *is scale-free with power-law exponent* $\tau = 1 + {}^1\!/\!_\gamma$.

Proof. This is immediate by the above and Lemma 2.2. □

3.2.2. Local and global clustering

Recall the definition of the average clustering coefficient, c^{av} above Definition 1.3. Let us further introduce another metric, the *global clustering coefficient* or *transitivity* of a graph G, defined as

$$c^{\mathrm{glob}}(G) := 3\frac{\sharp\{\text{triangles in } G\}}{\sharp\{\text{wedges in } G\}},$$

if there is at least one wedge in G, and $c^{\mathrm{glob}}(G) = 0$ otherwise. Both coefficients, c^{av} and c^{glob}, are well-established in the applied networks literature, see e.g. [62, 69] for some early papers.

Theorem 3.8 (Clustering coefficents)**.**

(a) For the average clustering coefficient we have

$$c^{av}(\mathscr{G}^t) \longrightarrow \int_0^1 \mathbb{P}\Big\{(X_u^{(1)}, S_u^{(1)}) \sim (X_u^{(2)}, S_u^{(2)})\Big\}\,\pi(\mathrm{d}u),$$

in probability as $t \to \infty$*, where* $(X_u^{(1)}, S_u^{(1)})$ *resp.* $(X_u^{(2)}, S_u^{(2)})$ *are two independent random variables on* $\mathbb{R}^d \times (0,1)$ *with distribution*

$$\frac{1}{\lambda_u}\Big(\rho\big(\tfrac{s^{1-\gamma}u^{\gamma}}{\beta}|x|^d\big)\mathbb{1}_{(u,1)}(s) + \rho\big(\tfrac{u^{1-\gamma}s^{\gamma}}{\beta}|x|^d\big)\mathbb{1}_{(0,u]}(s)\Big)\,\mathrm{d}x\mathrm{d}s, \qquad (3.2)$$

where $\lambda_u = \frac{\beta}{\gamma}(\frac{2\gamma-1}{1-\gamma} + u^{-\gamma})$ *is the normalising factor, and* π *is the probability measure on* $(0,1)$ *with density proportional to* $1 - e^{-\lambda_u} - \lambda_u e^{-\lambda_u}$*.*

(b) For the global clustering coefficient, there exists a number $c_\infty^{glob} \geq 0$ *such that*

$$c^{glob}(\mathscr{G}^t) \longrightarrow c_\infty^{glob}$$

in probability, as $t \to \infty$*. The limiting global clustering coefficient* c_∞^{glob} *is positive if and only if* $\gamma < 1/2$*.*

Remark 3.9. The limiting average clustering coefficient can be interpreted as the probability that in $\mathscr{G}_0^\infty$ two neighbours of the vertex at the origin are connected by an edge. The density of the birth time of the vertex at the origin here is not uniform but given by the measure π, which is the conditional distribution of the birth time of a vertex given that it has degree at least two. Observe that this coefficient is always positive. By contrast the global clustering coefficient vanishes asymptotically when preferential attachment to old nodes is strong (i.e. when γ is large). In this case the collection of wedges is dominated by those with an untypically old tip. These vertices have small local clustering as they are endvertices to a significant amount of long edges.

Proof. Define $H_t(\mathbf{x}, \mathscr{G}^t) = H(\mathbf{x}, \mathscr{G}^t) = c_{\mathbf{x}}^{\text{loc}}(\mathscr{G}^t)$ if $\mathbf{x}$ has at least two neighbours in $\mathscr{G}^t$ and $H_t(\mathbf{x}, \mathscr{G}^t) = H(\mathbf{x}, \mathscr{G}^t) = 0$ otherwise where $c_{\mathbf{x}}^{\text{loc}}$ is the local clustering coefficient used to define c^{av} above Definition 1.3. As H only depends on a bounded graph neighbourhood of $\mathbf{x}$, is bounded and hence uniformly integrable and translation invariant, we get by Theorem 3.3

$$\frac{1}{t} \sum_{\mathbf{x} \in \mathcal{X}^t} H(\mathbf{x}, \mathscr{G}^t) \longrightarrow \mathbb{E}_0 H(\mathbf{0}, \mathscr{G}_0^\infty) = \mathbb{E}_0 c_{\mathbf{0}}^{\text{loc}}(\mathscr{G}^\infty)$$

in probability, as $t \to \infty$. To calculate the limit, observe that, for a vertex

$\mathbf{x}$ with degree k, the number of wedges with tip $\mathbf{x}$ is $k(k-1)/2$. Therefore,

$$\mathbb{E}_0 c_{\mathbf{0}}^{\mathrm{loc}}(\mathscr{G}^\infty) = \int_0^1 \mathrm{d}u \sum_{k\geq 2} \mathbb{E}\Big[\tfrac{2}{k(k-1)} \sum_{(x,s)\sim(0,u)} \sum_{\substack{(y,t)\sim(0,u)\\ t<s}} \mathbb{1}_{\{(x,s)\sim(y,t)\}} \mathbb{1}_{\{\sharp\mathscr{N}_{(0,u)}=k\}}\Big].$$

By Proposition 2.1, the neighbourhood of the root $(0,u)$ is given by a Poisson point process with intensity measure $\lambda^<_{(0,u)} + \lambda^>_{(0,u)}$. Conditioned on the number of neighbours, the neighbours of $(0,u)$ are independent and identically distributed as the normalised intensity measure of the neighbourhood given in (3.2); see [55, Proposition 3.8]. It follows,

$$\mathbb{E}_0 c_{\mathbf{0}}^{\mathrm{loc}}(\mathscr{G}_0^\infty) = \int_0^1 \mathrm{d}u \; \mathbb{P}\Big\{(X_u^{(1)}, S_u^{(1)}) \sim (X_u^{(2)}, S_u^{(2)})\Big\} \mathbb{P}_{(0,u)}\Big\{\sharp\mathscr{N}_{(0,u)} \geq 2\Big\},$$

where $(X_u^{(1)}, S_u^{(1)})$ and $(X_u^{(2)}, S_u^{(2)})$ are independent and identically distributed as claimed. Choosing now $H(\mathbf{x},\mathscr{G}^t) = \mathbb{1}_{\{\sharp\mathscr{N}_{\mathbf{x}}(t)\geq 2\}}$ we get by the law of large numbers for the number of vertices of degree at least two

$$\frac{\sharp V_2(\mathscr{G}^t)}{t} \longrightarrow \int_0^1 \mathrm{d}u \; \mathbb{P}_{(0,u)}\Big\{\sharp\mathscr{N}_{(0,u)} \geq 2\Big\},$$

in probability. As $\sharp\mathscr{N}_{(0,u)}$ is Poisson distributed with intensity λ_u, we conclude

$$c^{\mathrm{av}}(\mathscr{G}^t) \longrightarrow \frac{\int_0^1 \mathrm{d}u (1 - e^{-\lambda_u} - \lambda_u e^{-\lambda_u}) \; \mathbb{P}\Big\{(X_u^{(1)}, S_u^{(1)}) \sim (X_u^{(2)}, S_u^{(2)})\Big\}}{\int_0^1 \mathrm{d}u \; (1 - e^{-\lambda_u} - \lambda_u e^{-\lambda_u})},$$

in probability, as claimed in (a).

For the global clustering coefficient, we count the number of triangles and wedges separately. To this end, define $H(\mathbf{x},\mathscr{G}^t)$ to be the number of triangles in $\mathscr{G}^t$ which have their youngest vertex in $\mathbf{x}$, and $\hat{H}(\mathbf{x},\mathscr{G}^t)$ to be the number of wedges with tip in $\mathbf{x}$. Again assumption (A) and (C) of Definition 3.2 are fulfilled. The moment condition is fulfilled for any $p>1$ as $H(\mathbf{x},\mathscr{G}^t) \leq (\sharp\mathscr{N}_{\mathbf{x}}^<)^2$. Moreover,

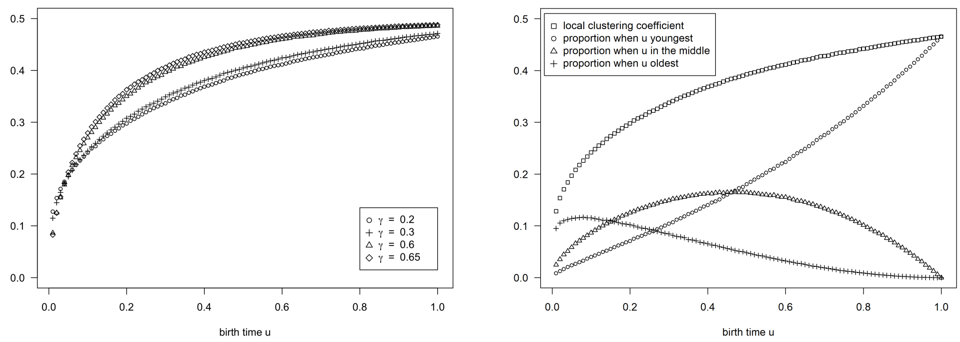

Figure 3.1.: Local clustering coefficient of a vertex $(0, u)$ for parameters $a = 1$ and $\beta = c_{\text{ed}}(1-\gamma)$ chosen such that the asymptotic edge density is fixed at c_{ed}. The plot on the left displays the behaviour of the model for high edge density ($c_{\text{ed}} = 10$) for various values of γ. We remark that the shown behaviour is qualitatively independent of the edge density. In the plot on the right, the clustering coefficient for $\gamma = 0.2$ is shown, along with the probabilities of the event that u is younger (resp. in the middle or older) than two randomly picked neighbours, which are connected, see [32, Fig 3].

$$\begin{aligned}\hat{H}(\mathbf{x}, \mathscr{G}^t) &= \tfrac{1}{2}(\sharp\mathscr{N}_{\mathbf{x}}^{<}(t))(\sharp\mathscr{N}_{\mathbf{x}}^{<}(t) - 1) + \tfrac{1}{2}(\sharp\mathscr{N}_{\mathbf{x}}^{>}(t))(\sharp\mathscr{N}_{\mathbf{x}}^{>}(t) - 1)\\ &\quad + (\sharp\mathscr{N}_{\mathbf{x}}^{<}(t))(\sharp\mathscr{N}_{\mathbf{x}}^{>}(t))\\ &\le 2\Big((\sharp\mathscr{N}_{\mathbf{x}}^{<})^2 + (\sharp\mathscr{N}_{\mathbf{x}}^{>})^2\Big).\end{aligned}$$

If $\gamma < 1/2$ and $1 < p < 1/2\gamma$, we have $\hat{H} \in \mathcal{H}_p$ and Theorem 3.3 gives that

$$c^{\text{glob}}(\mathscr{G}^t) = \frac{\sum_{\mathbf{x} \in \mathcal{X}^t} H(\mathbf{x}, \mathscr{G}^t)}{t} \cdot \frac{t}{\hat{H}(\mathbf{x}, \mathscr{G}^t)} \longrightarrow \frac{\mathbb{E}_0 H(\mathbf{0}, \mathscr{G}_0^\infty)}{\mathbb{E}_0 \hat{H}(\mathbf{0}, \mathscr{G}_0^\infty)} > 0,$$

in probability. If $\gamma > 1/2$, we apply the theorem to the bounded functions $\hat{H} \wedge k$ and send then $k \to \infty$ and we get $t^{-1} \sum_{\mathbf{x}} \hat{H}(\mathbf{x}, \mathscr{G}^t) \to \infty$ and hence $c^{\text{glob}}(\mathscr{G}^t) \to 0$ in probability as $t \to \infty$. $\square$

Corollary 3.10. *The age-based spatial preferential attachment network* $(\mathscr{G}_t : t \geq 0)$ *shows clustering in the sense of Definition 1.3.*

The local and average clustering coefficients cannot be calculated explicitly, but can be simulated. We focus on the profile functions $\rho = \frac{1}{2a}\mathbb{1}_{[0,a]}$, for

$a \geq 1/2$, dimension $d = 1$, and fixed edge density $\beta/(1-\gamma)$. Figure 3.1 shows the local clustering coefficient of a vertex of age u in $\mathscr{G}^\infty$ showing monotone dependence on the age, i.e. the empirical probability that two neighbours of a given vertex are connected to each other is larger for younger vertices. This coincides with our intuitive understanding of the local structure of the networks, in which a young vertex, typically, is connected to either very close or very old vertices such that two randomly chosen neighbours have a decent chance of being connected to each other as well. By contrast, an old vertex typically has more long edges to younger vertices. Thus, two of its neighbours are typically further apart, which reduces the chance of them being each others neighbour. This monotonicity occurs independently of the choice of β, γ and a.

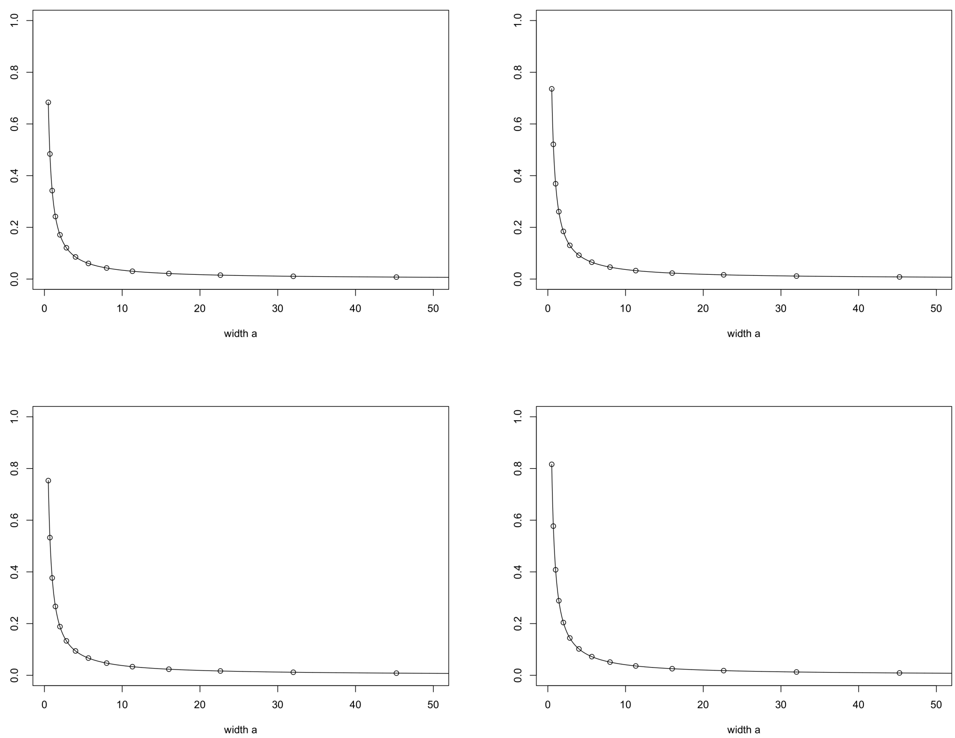

Figure 3.2.: Average clustering coefficient for the network with profile function $\rho = \frac{1}{2a}\mathbb{1}_{[0,a]}$ plotted against the width a, for $\gamma = 0.3$ in the left resp. $\gamma = 0.6$ in the right graphs. The graphs in the top row correspond to fixed edge density 1 while the bottom row corresponds to edge density 10, see [32, Fig 4].

In Figure 3.2 we see that the dependence of the average clustering coefficient with respect to the width a of the profile function is of order $1/a$, a scaling that we also see in the analysis of the global clustering coefficient in the case $\gamma < \frac{1}{2}$. Hence, the average clustering coefficient and the global clustering coefficient (if $\gamma < \frac{1}{2}$) can be varied by the choice of ρ and can be made arbitrarily small by choosing a large. Unlike with the global clustering coefficient, there is a mild dependence on β. Again, roughly speaking, large width of ρ encourages long edges and reduces clustering.

3.2.3. Asymptotics for typical edge lengths

We define the *empirical edge length distribution* in $\mathscr{G}^t$ by

$$\lambda_t = \frac{1}{\sharp E(\mathscr{G}^t)} \sum_{\{\mathbf{x},\mathbf{y}\} \in E(\mathscr{G}^t)} \delta_{\mathrm{d}_t(x,y)}.$$

Theorem 3.11. *For every continuous and bounded $g : [0,\infty) \to \mathbb{R}$, we have*

$$\frac{1}{\sharp E(\mathscr{G}^t)} \sum_{\{\mathbf{x},\mathbf{y}\} \in E(\mathscr{G}^t)} g(d_t(x,y)) = \int g \; \mathrm{d}\lambda_t \longrightarrow \int g \mathrm{d}\lambda$$

in probability, as $t \to \infty$, where the limiting probability measure λ on $(0,\infty)$ is given by

$$\lambda([a,b)) = \tfrac{1-\gamma}{\beta} \int_0^1 \mathrm{d}u \int_0^u \mathrm{d}s \int_{\{a \le |y| < b\}} \mathrm{d}y \; \rho\Big(\beta^{-1} u^{1-\gamma} s^{\gamma} |y|^d\Big). \tag{3.3}$$

Proof. For $a < b \in [0,\infty]$ and $t \in (0,\infty]$, define the functional

$$H_t^{a,b}(\mathbf{x}, \mathscr{G}^t) = \sum_{\mathbf{y} \in \mathscr{N}_{\mathbf{x}}^{<}(t)} \mathbb{1}_{[a,b)}(\mathrm{d}_t(x,y)). \tag{3.4}$$

Observe that the law of $\lambda_t[a,b))$ in $\mathscr{G}^t$ equals the law of

$$\frac{1}{\sharp E(\mathscr{G}^t)} \sum_{\mathbf{x} \in \mathcal{X}^t} H_t^{a,b}(\mathbf{x}, \mathscr{G}^t).$$

Since the sum in (3.4) is dominated by the outdegree, $(H_t^{a,b} : t \geq 0) \in \mathcal{H}_p$ for some $p > 1$. We thus get by Theorem 3.3

$$\frac{1}{t} \sum_{\mathbf{x} \in \mathcal{X}^t} H_t^{a,b}(\mathbf{x}, \mathscr{G}^t) \longrightarrow \mathbb{E}_0 H_\infty^{a,b}(\mathbf{0}, \mathscr{G}_0^\infty)$$

as well as $\sharp E(\mathscr{G}^t)/t \to \beta/(1-\gamma)$ in probability. By definition $\lambda([a,b)) = \frac{1-\gamma}{\beta} \mathbb{E}_0 H_\infty^{a,b}(\mathbf{0}, \mathscr{G}_0^\infty)$, we infer that $\lambda_t([a,\infty)) \to \lambda([a,\infty))$ in probability, as $t \to \infty$. Therefore, convergence in probability of λ_t to λ in the space of probability measures on $(0,\infty)$, equipped with the Lévy-Prokhorov metric follows. $\square$

Remark 3.12. Suppose there exists $\delta > 1$ such that the profile function satisfies $\rho(x) \asymp 1 \wedge x^{-\delta}$. Then, the explicit formula for λ in (3.3) can be used to calculate the tail behaviour of λ. More precisely, separating the integration into several domains, depending on whether one is integrating the tail domain of ρ or not, results in terms decaying polynomial of order $d, d(1/\gamma - 1)$ and $d(\delta - 1)$. This gives that $\lambda([K,\infty)) \asymp 1 \wedge (\beta^{-1/d} K)^{-\eta}$, where

$$\eta := \min\Big\{d, d\Big(\tfrac{1}{\gamma} - 1\Big), d(\delta - 1)\Big\}. \tag{3.5}$$

If ρ is a regularly varying profile function for $\delta > 1$, cf. (1.17), one observes the same tail behaviour for $K \to \infty$ by the Potter bounds [7]. In particular, λ has finite expectation if $\eta > 1$ and infinite expectation if $\eta < 1$.

We denote by M_0^∞ the length of the longest outgoing edge of the origin in $\mathscr{G}_0^\infty$. By the construction of λ above, $\lambda([K,\infty))$ is the expected number of outgoing edges of length bigger than K divided by the total number of outgoing edges from the origin. If K is large, this should be of similar order to the probability that $M_0^\infty \geq K$. This is confirmed in the following lemma.

Lemma 3.13. *Suppose ρ fulfils (1.17) for some $\delta > 1$. Then $\mathbb{E}[(M_0^\infty)^a]$ is finite if $a < \eta$ and infinite if $a > \eta$ where η is defined as in (3.5).*

Proof. We show that the tail probability $\mathbb{P}\{M_0^\infty \geq K^{1/a}\}$ is of order $K^{-\eta/a}$ as $K \to \infty$. The number of outgoing edges with length at least $K^{1/a}$ in $\mathscr{G}_0^\infty$ from the vertex $(0, u)$ at the origin is Poisson distributed with parameter

$$\lambda_{K^{1/a},u} := \lambda^{<}_{(0,u)}\Big(\{|x| \geq K^{1/a}\} \times (0, u]\Big)$$

by Proposition 2.1. Hence,

$$\begin{aligned}\mathbb{P}\Big\{(M_0^\infty)^a \geq K\Big\} &= \int_0^1 \mathrm{d}u\, \Big(1 - \exp(-\lambda_{K^{1/a},u})\Big) \asymp \int_0^1 \mathrm{d}u\, \lambda_{K^{1/a},u} \\ &\asymp \lambda([K^{1/a}, \infty)),\end{aligned}$$

as $\lambda_{K^{1/a},u}$ is bounded. The proof concludes by Remark 3.12. $\square$

Using this, we can establish a result about the average edge length in $\mathscr{G}^t$.

Theorem 3.14. *Suppose that ρ fulfils (1.17) for some $\delta > 1$. Then, for all $a > 0$ and $b \in [0, \eta/a)$, there exists a positive constant C, depending on $a, b, \gamma, \beta, \rho$ such that*

$$\frac{1}{\sharp E(\mathscr{G}^t)} \sum_{\mathbf{x} \in \mathcal{X}^t} \Bigg(\sum_{\mathbf{y} \in \mathscr{N}^{<}_{\mathbf{x}}(t)} \mathrm{d}_t(x, y)^a \Bigg)^b \longrightarrow C \tag{3.6}$$

in probability as $t \to \infty$.

Remark 3.15. This theorem can also be applied on the preferential attachment network $(\mathscr{G}_t : t \geq 0)$ by replacing $\mathscr{G}^t$ and $\mathcal{X}^t$ by $\mathscr{G}_t$ and $\mathcal{X}_t$ and d_t by $t^{-1/d}\,\mathrm{d}_1$. It hence gives a result about the average rescaled edge length in the network. If $\eta > 1$, one can choose $a = b = 1$, and this yields that the mean edge length in $\mathscr{G}_t$ is of order $t^{-1/d}$. If $\eta \leq 1$ (and in particular always if $d = 1$), the mean edge length is of larger order.

Proof. Define

$$H(\mathbf{x}, \mathscr{G}^t) := H_t(\mathbf{x}, \mathscr{G}^t) := \Big(\sum_{\mathbf{y} \in \mathscr{N}_{\mathbf{x}}^<(t)} \mathrm{d}_t(x,y)^a \Big)^b$$

so that the left-hand side in (3.6) can be written as

$$\frac{1}{\sharp E(\mathscr{G}^t)} \sum_{\mathbf{x} \in \mathcal{X}^t} H(\mathbf{x}, \mathscr{G}^t).$$

If suffices to show that H fulfils the uniform moment condition of Definition 3.2 to prove the theorem since the weak law of large numbers then ensures the convergence in probability to $\frac{1-\gamma}{\beta}\mathbb{E}_0 H(\mathbf{0}, \mathscr{G}_0^\infty)$. To this end, recall M_0^∞, the length of the longest outgoing edge of the root $\mathbf{0}$ in $\mathscr{G}_0^t$ and observer that, almost surely, $H(\mathbf{0}, \mathscr{G}_0^t) \leq (M_0^\infty)^{ab}(\sharp\mathscr{N}_{\mathbf{0}}^<)^b$. Since by choice, $ab < \eta$, there exists some $p, q > 1$ such that $\alpha := pqab \leq \eta$. Lemma 3.13 then ensures $\mathbb{E}_0[(M_0^\infty)^\alpha] < \infty$ and by applying Hölder's inequality to the observed bound for $H(\mathbf{0}, \mathscr{G}_0^t)$, we get

$$\sup_{t>0} \mathbb{E}_0 H(\mathbf{0}, \mathscr{G}_0^t)^p \leq \Big(\mathbb{E}_0[(m_0^\infty)^\alpha] \Big)^{1/q} \Big(\mathbb{E}_0 \Big[(\sharp \mathscr{N}_{\mathbf{0}}^<)^{\alpha/(a(q-1))} \Big] \Big)^{(q-1)/q} < \infty.$$

$\square$

3.3. Robustness vs. non-robustness

The proof of robustness and non-robustness follows the proof done by Jacob and Mörters for the more complicated degree-based spatial preferential attachment model [45] which can be adapted easily for the simpler age-based model. Recall Bernoulli bond percolation with retention parameter $p \in (0, 1]$ above Definition 1.4 and the definition itself to refamiliarise yourself with the notion of robustness.

We fix now $\beta > 0$, $\gamma \in (0, 1)$ and a profile function ρ satisfying (1.17) for some $\delta \in (1, \infty)$ and perform Bernoulli bond percolation on the graphs $\mathscr{G}^t$.

We denote the resulting graphs by $\mathscr{G}^t(p)$.

Recall that an event $A(t)$ holds with extreme probability, *wep*(t) if it holds with probability at least $1 - \exp(-\Omega(\log^2(t))$, cf. Section 2.5. We say, that $A(t)$ holds with high probability, *whp*(t), if the probability of $A(t)$ converges to one, as $t \to \infty$. If the parameter is clear, we simply write *whp* and *wep*. Also, we use the standard Landau notation $f = O(g)$ if $\limsup_{x\to\infty} f(x)/g(x) < \infty$ throughout this section. Further, since p is now our main percolation parameter instead of β, we write

$$\theta(p) := \mathbb{P}_0\{\mathbf{0} \leftrightarrow \infty \text{ in } \mathscr{G}_0^\infty(p)\},$$

cf. Section 2.1.1.

We start with the proof of non-robustness for $\gamma < \delta/(\delta+1)$. Also recall that for the age-dependent random connection model we have $p_c > 0$ if and only if $\beta_c > 0$ by Remark 1.9(iii).

Lemma 3.16. *Let $\gamma < \frac{\delta}{\delta+1}$, then there exist $p_c > 0$ such that for any $p < p_c$ the graph family $(\mathscr{G}^t(p) : t \geq 0)$ contains no giant component.*

Proof. Define for $k \in \mathbb{N}$ the functional $H^k(\mathbf{x}, \mathscr{G}^t(p))$ as the indicator that the component of the vertex $\mathbf{x}$ in $\mathscr{G}^t(p)$ is of size at most k. Since H^k is bounded and depends only on a bounded graph neighbourhood of $\mathbf{x}$ we can apply Theorem 3.3 and get

$$\lim_{t\to\infty} \frac{1}{t} \sum_{\mathbf{x}\in\mathcal{X}^t} H^k(\mathbf{x}, \mathscr{G}^t(p)) = \mathbb{E}_0 H^k(\mathbf{0}, \mathscr{G}^t(p))$$

in probability. The left-hand side is asymptotically in t the proportion of vertices that are in components no bigger than k. As $k \to \infty$, the right-hand side converges to

$$1 - \mathbb{P}_0\{\mathbf{0} \leftrightarrow \infty \text{ in } \mathscr{G}_0^\infty(p)\} = 1,$$

for a small enough $p > 0$ by Theorem 1.8 and Remark 1.9(iii) since $\gamma < \delta/(\delta+1)$. Hence, with high probability, all vertices belong to finite sized components and there is no giant component in $(\mathscr{G}^t(p) : t \geq 0)$. $\square$

We turn to the more complicated proof of robustness if $\gamma > \delta/(\delta+1)$.

Proposition 3.17. *Let $\gamma > \frac{\delta}{\delta+1}$ and $p > 0$. With high probability, the largest connected component $\mathscr{C}(\mathscr{G}^t(p))$ of the graph $\mathscr{G}^t(p)$ is of size $(\theta(p) + o(1))t$, while the second largest component is of size $o(t)$. Hence, there is a unique giant component of asymptotic density $\theta(p) > 0$.*

Proof. If we prove that $\sharp\mathscr{C}(\mathscr{G}^t(p))$ is asymptotically of linear density $\theta(p)$, the uniqueness follows by the arguments of the proof of Lemma 3.16. The proof is done in two steps: First, we study the component of the oldest vertex in $\mathscr{G}^t(p)$. Second, we use the weak law of large numbers to show that the asymptotic proportion of vertices belonging to the cluster of the oldest vertex equals the probability that the root $\mathbf{0}$ in the limiting graph belongs to the infinite cluster. Since our arguments do not depend on the value of p, we carry out the proof for $p = 1$ and the graphs $\mathscr{G}^t = \mathscr{G}^t(1)$ to simplify notation. Recall the exponents

$$\alpha_1 \in \left(1, \tfrac{\gamma}{\delta(1-\gamma)}\right) \text{ and } \alpha_2 \in \left(\alpha_1, \tfrac{\gamma}{\delta}(1+\alpha_1\delta)\right)$$

of Section 2.5. Since, with high probability, the oldest vertex in $\mathscr{G}^t$ is born before time $\log t/t$, we work conditioned on this event in the following. The first step of the proof is a version of Lemma 2.10 of connecting old vertices through connectors in the finite graphs $\mathscr{G}^t$.

Lemma 3.18. *Fix $k > 1$. Given that the oldest vertex in $\mathscr{G}^t$ is born before time $\log t/t$, every vertex $\mathbf{x} = (x, u)$ with birth time $u < t^{-1/\alpha_1^k}$ is connected through a connector born after time $1/2$ to a vertex $\mathbf{y} = (y, s)$ with $s < u^{\alpha_1}$*

and $|x-y|^d < u^{-\alpha_2}$ *or is connected through a connector to the oldest vertex or is the oldest vertex, wep*(t).

Proof. Let $\mathbf{x} = (x,u)$ be a given vertex with $u < t^{-1/\alpha_1^k}$. First consider $u \in \left(t^{-1/\alpha_2}, t^{-1/\alpha_1^k}\right)$. Then u is large enough such that the ball with radius $u^{-\alpha_2}$ is completely contained in the torus $\mathbb{T}_t^d$ and we find the vertex $\mathbf{y}$, *wep*, by the same arguments as in Lemma 2.10. If $u < t^{-1/\alpha_2}$, either $\mathbf{x}$ is the oldest vertex or, due to the finite volume of $\mathbb{T}_t^d$, the oldest vertex is within distance $t^{1/d} < u^{-\alpha_2/d}$. As the oldest vertex is born before time $(\log t)/t$, we can apply the arguments of Lemma 2.10 once more and, *wep*, the vertex $\mathbf{x}$ is connected to the oldest one through a connector. $\square$

An important consequence of Lemma 3.18 is that for a fixed $k > 1$, all vertices born before time t^{-1/α_1^k} in $\mathscr{G}^t$ belong with extreme probability to the same cluster which is particularly the cluster of the oldest vertex in $\mathscr{G}^t$.

For the second step define $H_t(\mathbf{x}, \mathscr{G}^t)$ as the indicator that the vertex $\mathbf{x}$ belongs to the component of the oldest vertex in $\mathscr{G}^t$ and $H_\infty(\mathbf{x}, \mathscr{G}^\infty)$ as the indicator that $\mathbf{x}$ belongs to the infinite cluster of $\mathscr{G}^\infty$. Since the functionals are bounded, the uniform moment condition (B) of Definition 3.2 is fulfilled. Furthermore, the functionals are invariant under shifts. Assume that $H_t(\mathbf{0}, \mathscr{G}_0^t) \to H_\infty(\mathbf{0}, \mathscr{G}_0^\infty)$ in probability, then Theorem 3.3 ensures

$$\frac{1}{t}\sum_{\mathbf{x}\in\mathcal{X}^t} H_t(\mathbf{x}, \mathscr{G}^t) \to \mathbb{E}_0 H_\infty(\mathbf{0}, \mathscr{G}_0^\infty) = \theta(1).$$

Hence, the proof concludes if we show the convergence assumption (A). To this end, define by A_t the event that the root $\mathbf{0} = (0, U)$ is connected by a path to the oldest vertex in $\mathscr{G}_0^t$ and $A_\infty := \{\mathbf{0} \leftrightarrow \infty \text{ in } \mathscr{G}_0^\infty\}$. Thus, we have to show that

$$\textbf{(a)}\ \mathbb{P}_0(A_t \cap A_\infty^c) \longrightarrow 0 \quad\text{and}\quad \textbf{(b)}\ \mathbb{P}_0(A_\infty \cap A_t^c) \longrightarrow 0,$$

as $t \to \infty$. To prove **(a)**, observe that on A_∞^c, the component of the root in $\mathscr{G}_0^\infty$ is finite and depends only on a finite graph distance. Hence, for large enough t the components of the root in $\mathscr{G}_0^t$ and $\mathscr{G}_0^\infty$ coincide by Theorem 3.1. On A_t the oldest vertex of $\mathscr{G}^t$ belongs to that component. Therefore, for a sufficiently large t, the oldest vertex in $\mathscr{G}_0^t$ must remain the oldest vertex forever which only happens with a probability converging to zero as $t \to \infty$.

To prove **(b)**, we introduce some notation:

- Let $a, b > 0$ be small parameters and $m > 1$ be a large parameter which we specify later.

- Let B_t be the event that neither in $\mathscr{G}_0^{t^m}$ nor in $\mathscr{G}_0^\infty$ a vertex located in $\mathcal{B}(0, t^{1/d})$ is incident to an edge longer than $\frac{t^{m/d}}{2} - t^{1/d}$ where $\mathcal{B}(0, t^{1/d})$ denotes the ball around the origin with radius $t^{1/d}$.

- On the event $\{U > t^{-a}\}$, that holds with high probability, we introduce further objects:

  - Denote by $\mathscr{C}_0$ the component of the root $\mathbf{0} = (0, U)$ in $\mathscr{G}_0^\infty$, restricted to vertices located in $\mathcal{B}(0, t^{1/d})$ and born after time t^{-a}.

  - Let $\widehat{A}_t \subset \{U > t^{-a}\}$ be the event that $\mathscr{C}_0$ is connected in $\mathscr{G}_0^\infty$ by a direct edge to a vertex located in $\mathcal{B}(0, t^{1/d})$ that is born before time t^{-a}.

The proof of **(b)** is carried out in the following three steps:

(I) $\mathbb{P}_0(B_t^c) \to 0$, as $t \to \infty$,

(II) $\mathbb{P}_0\Big((A_\infty \cap B_t) \cap \widehat{A}_t^c\Big) \to 0$, as $t \to \infty$ and

(III) $\mathbb{P}_0\Big((\widehat{A}_t \cap B_t) \cap A_{t^m}^c\Big) \to 0$, as $t \to \infty$.

Proof of (I). The number of vertices located inside $\mathcal{B}(0, t^{1/d})$ is of order t, $wep(t)$. Each of the $O(t)$ vertices carries an independent birth time distributed uniformly on $(0,1)$. By Theorem 3.11 and Remark 3.12, the probability that a vertex with uniform birth time is incident to an edge longer than $t^{m/d}/2 - t^{1/d}$ in $\mathscr{G}_0^\infty$ is of order

$$\Big(\tfrac{t^{m/d}}{2} - t^{1/d}\Big)^{-\eta}, \quad \text{where } \eta = d \cdot \min\Big\{1, \tfrac{1}{\gamma} - 1, \delta - 1\Big\}. \tag{3.7}$$

For large enough t the same holds true in the finite graph $\mathscr{G}_0^{tm}$. Hence, the number of vertices inside $\mathcal{B}(0, t^{1/d})$, incident to an edge longer than ${}^{t^{m/d}}\!/_2 - t^{1/d}$ is bounded by a Binomial with $O(t)$ trials and success probability (3.7). We choose ${}^m\!/_d > \eta^{-1}$, say ${}^m\!/_d > \eta^{-1} + \varepsilon$, and infer

$$t\Big(\tfrac{t^{m/d}}{2} - t^{1/d}\Big)^{-\eta} \le \Big(t^{(m/d)-(1/\eta)}\Big)^{-\eta} \le t^{-\eta\varepsilon} = o(1),$$

from which we conclude (I).

Proof of (II). We work on the event $\{U > t^{-a}\}$ and split the event depending on the size of $\mathscr{C}_0$. First, we show that

$$\lim_{t\to\infty} \mathbb{P}\Big(\{\sharp\mathscr{C}_0 < t^{b/d}\} \cap A_\infty \cap B_t \cap \widehat{A}_t^c\Big) = 0.$$

On $A_\infty \cap \widehat{A}_t^c$, a vertex of $\mathscr{C}_0$ is connected by an edge to a vertex outside of $\mathcal{B}(0, t^{1/d})$. On $B_t \cap \widehat{A}_t^c$, this vertex has to be located in

$$\Big[-\tfrac{t^{m/d}}{2}, \tfrac{t^{m/d}}{2}\Big]^d \cap \mathcal{B}(0, t^{1/d})^c.$$

Hence, on $\{\sharp\mathscr{C}_0 < t^{b/d}\}$, one of the vertices of $\mathscr{C}_0$ is incident to an edge longer than $t^{(1-b)/d}$, called a *long edge* in the following. This edge either connects two vertices inside $\mathscr{C}_0$ or it connects one vertex of $\mathscr{C}_0$ to some vertex in $\Big[-\tfrac{t^{m/d}}{2}, \tfrac{t^{m/d}}{2}\Big]^d \cap \mathcal{B}(0, t^{1/d})^c$. We define the three subsets

$$\mathcal{Y}_1 := \mathcal{X}_0^\infty \cap \Big(\mathcal{B}(0, t^{1/d}) \times (t^{-a}, 1)\Big),$$
$$\mathcal{Y}_2 := \mathcal{X}_0^\infty \cap \Big(\Big(\Big[-\tfrac{t^{m/d}}{2}, \tfrac{t^{m/d}}{2}\Big]^d \cap \mathcal{B}(0, t^{1/d})^c\Big) \times (t^{-a}, 1)\Big) \text{ and}$$

$$\mathcal{Y}_3 := \mathcal{X}_0^\infty \cap \Big(\Big(\Big[-\tfrac{t^{m/d}}{2}, \tfrac{t^{m/d}}{2} \Big]^d \cap \mathcal{B}(0, t^{1/d})^c \Big) \times (0, t^{-a}] \Big).$$

Note that $\mathscr{C}_0 \subset \mathcal{Y}_1$ and $\mathcal{Y}_2, \mathcal{Y}_3$ are sets of potential end vertices of a long edge. For two vertices $\mathbf{x} = (x, u), \mathbf{y} = (u, s_y) \in \mathcal{Y}_1 \cup \mathcal{Y}_2$ it holds

$$\mathbb{P}_0\{\mathbf{x} \sim \mathbf{y}\} \leq \rho\Big(\tfrac{t^{-a}|x-y|^d}{t^{a\gamma}}\Big) = O\Big(t^{a(1+\gamma)(\delta-\varepsilon)}|x-y|^{-d(\delta-\varepsilon)}\Big), \tag{3.8}$$

for some $\varepsilon > 0$ using the Potter bound [7]. Similarly, for $\mathbf{x} = (x, u) \in \mathcal{Y}_1$ and $\mathbf{y} = (y, s_y) \in \mathcal{Y}_3$, we have

$$\begin{aligned}\mathbb{P}_0\{\mathbf{x} \sim \mathbf{y}\} \leq \rho\Big(\tfrac{t^{-a}|x-y|^d}{s_y^{a\gamma}}\Big) &= O\Big(t^{a(\delta-\varepsilon)} s_y^{-\gamma(\delta-\varepsilon)}|x-y|^{-d(\delta-\varepsilon)}\Big) \wedge 1, \\ &\leq O\Big(t^{a(\delta-\varepsilon)}\Big(s_y^{-\gamma(\delta-\varepsilon)}|x-y|^{-d(\delta-\varepsilon)}\Big) \wedge 1\Big).\end{aligned} \tag{3.9}$$

Given $\mathscr{C}_0$, we fix $\mathbf{x} = (x, u) \in \mathscr{C}_0$. Our aim is to bound the probability that $\mathbf{x}$ is connected a vertex $\mathbf{y} = (y, s_y)$ with $|x - y| > t^{(1-b)/d}$ separately for $\mathbf{y} \in \mathcal{Y}_i$, $i = 1, 2, 3$. Given $\mathcal{X}^\infty$, the probability that $\mathbf{x}$ is connected by a long edge to some $\mathbf{y} \in \mathcal{Y}_1$ is bounded by

$$E_1 := ct^{a(1+\gamma)(\delta-\varepsilon)} t^{(\delta-\varepsilon)(1-b)}(\sharp \mathcal{Y}_1),$$

for some constant $c > 1$ using (3.8). Hence, on $\{\sharp\mathscr{C}_0 < t^{b/d}\}$, the expected number of vertices in $\mathscr{C}_0$ connected by a long edge to another vertex in $\mathscr{C}_0$, is bounded by

$$t^{b/d}\mathbb{E}_0 E_1 \asymp t^{a(1+\gamma)(\delta-\varepsilon)+b(1/d+\delta-\varepsilon)+(1+\varepsilon-\delta)} = o(1),$$

as $t \to \infty$ for a, b, ε chosen small enough. Next for $\mathbf{y} \in \mathcal{Y}_2 \cup \mathcal{Y}_3$, it holds

$$|x - y| \geq t^{(1-b)/d} \vee (|y| - t^{1/d}) \geq \tfrac{1}{2}(|y| + t^{(1-b)/d} - t^{1/d}).$$

Hence, the conditional probability of x being connected by a long edge to a vertex in $\mathcal{Y}_2$, is bounded by

$$E_2 := ct^{a(1+\gamma)(\delta-\varepsilon)} \sum_{\mathbf{y} \in \mathcal{Y}_2} \Big(|y| + t^{(1-b)/d} - t^{1/d}\Big)^{-d(\delta-\varepsilon)},$$

by (3.8). Similarly, for the conditional probability of $\mathbf{x}$ being connected to $\mathcal{Y}_3$ by a long edge, we observe the bound

$$E_3 := ct^{a(\delta-\varepsilon)} \sum_{\mathbf{y}\in\mathcal{Y}_3} \Big(s_y^{-\gamma(\delta-\varepsilon)}(|y| - t^{1/d} + t^{(1-b)/d})^{-d(\delta-\varepsilon)} \wedge 1\Big),$$

by (3.9). Observe that both, E_2 and E_3, are independent of the location of $\mathbf{x}$. Using Campbell's formula [55], we get

$$\begin{aligned}\mathbb{E}_0 E_2 &\asymp t^{a(1+\gamma)(\delta-\varepsilon)}(1-t^{-a}) \int_{t^{1/d}}^{t^{m/d}} \mathrm{d}r\ r^{d-1}(|r| + t^{(1-b)/d} - t^{1/d}))^{-d(\delta-\varepsilon)} \\ &\asymp t^{a(1+\gamma)(\delta-\varepsilon)} t^{-(1-b)(\delta-\varepsilon)+1}\end{aligned}$$

and hence

$$t^{b/d}\mathbb{E}_0 E_2 \asymp t^{1+\varepsilon-\delta+a(1+\gamma)(\delta-\varepsilon)+b(1/d+\delta-\varepsilon)} = o(1),$$

as $t\to\infty$ for sufficiently small a, b, ε. Similarly,

$$\begin{aligned}\mathbb{E}E_3 &\asymp t^{a(\delta-\varepsilon)} \int_{t^{1/d}}^{t^{m/d}} \mathrm{d}r\ r^{d-1} \int_0^{t^{-a}} \mathrm{d}s \Big(\Big(s^{\gamma/d}(r - t^{1/d} + t^{(1-b)/d})\Big)^{-d(\delta-\varepsilon)} \wedge 1\Big) \\ &\asymp t^{a(\delta-\varepsilon)} \int_{t^{(1-b)/d}}^{t^{m/d}-t^{1/d}+t^{(1-b)/d}} \mathrm{d}y\ \Big(y + t^{\frac{1}{d}} - t^{\frac{1-b}{d}}\Big)^{d-1} \\ &\qquad\qquad \times \Big(y^{-\frac{d}{\gamma}} + \int_{y^{-d/\gamma}}^{1} \mathrm{d}s\ s^{-\gamma(\delta-\varepsilon)} y^{-d(\delta-\varepsilon)}\Big) \\ &\le t^{a(\delta-\varepsilon)+1-(1-b)((1/\gamma)\wedge(\delta-\varepsilon))}\end{aligned}$$

and hence, writing $\eta' := ({}^1/_\gamma \wedge (\delta-\varepsilon))$,

$$t^{b/d}\mathbb{E}_0 E_3 \le t^{a(\delta-\varepsilon)+b(1/d+\eta')+1-\eta'} = o(1)$$

as $t\to\infty$, for small enough a, b, ε. The claim follows as

$$\lim_{t\to\infty} \mathbb{P}_0\Big(\{\sharp\mathscr{C}_0 < t^{b/d}\} \cap A_\infty \cap B_t \cap \widehat{A}_t^c\Big) \le \lim_{t\to\infty} ct^{b/d}\mathbb{E}_0[E_1 + E_2 + E_3] = 0.$$

The second part of the proof of part (II) is to show that

$$\lim_{t\to\infty} \mathbb{P}_0\Big(A_\infty \cap \{\sharp\mathscr{C}_0 \geq t^{b/d}\} \cap \widehat{A}_t^c\Big) = 0.$$

On the event $A_\infty \cap \{\sharp\mathscr{C}_0 \geq t^{b/d}\}$, we try to connect a vertex of $\mathscr{C}_0$ to some vertex in $\mathcal{B}(0, t^{1/d}) \times (0, t^{-a})$ which coincides with the event $\widehat{A}_t$ happening. Given $\mathscr{C}_0$ and a vertex $\mathbf{x} \in \mathscr{C}_0$, pick some $\varepsilon > 0$, then *wep* there exists a vertex $\mathbf{y}$ with birth time in $(0, t^{-a}])$ and within distance $t^{a/d+\varepsilon}$ of $\mathbf{x}$. The two vertices are connected by a direct edge with probability at least

$$\rho\Big(\tfrac{t^{a+\varepsilon}}{t^{a\gamma}}\Big) = \Omega\Big(t^{-(a(1-\gamma)+\varepsilon)(\delta+\varepsilon')}\Big). \tag{3.10}$$

Hence, the expected number of edges between $\mathscr{C}_0$ and $\mathcal{B}(0, t^{1/d}) \times (0, t^{-a}]$ is bounded from below by the expectation of a Binomial with parameters $\lfloor t^{b/d} \rfloor$ and (3.10). Therefore, $\mathscr{C}_0$ is linked by a direct edge to $\mathcal{B}(0, t^{1/d}) \times (0, t^{-a}]$, *whp*, if $b > d(a(1-\gamma)+\varepsilon)(\delta+\varepsilon')$. The proof of (II) concludes by reducing a, ε and ε' if necessary to ensure that this inequality is satisfied.

Proof of (III). Recall the claim

$$\lim_{t\to\infty} \mathbb{P}_0\{\widehat{A}_t \cap B_t \cap A^c_{t^m}\} = 0.$$

On $\widehat{A}_t$, the root $\mathbf{0} = (0, U)$ is connected in $\mathscr{G}_0^\infty$ to a vertex $\mathbf{x} = (x, s)$ located inside $\mathcal{B}(0, t^{1/d})$ with birth time in $(0, t^{-a}]$. On B_t, all these connections remain in the finite graph $\mathscr{G}_0^{t^m}$. Thus, it suffices to show that $\mathbf{x}$ is, *whp* connected by a path to the oldest vertex of $\mathscr{G}_0^{t^m}$. To this end, observe that, for $k > \log({}^m\!/_a)/\log\alpha_1$, we have $s \leq t^{-a} < t^{-m/\alpha_1^k}$ and the proof concludes with Lemma 3.18. $\square$

Corollary 3.19. *Assume that ρ fulfils (1.17) for some $\delta \in (1, \infty)$. Then the age-based spatial preferential attachment network $(\mathscr{G}_t : t \geq 0)$, with parameters $\beta > 0$, $\gamma \in (0,1)$ and ρ, is robust if $\gamma > \delta/(\delta+1)$ and non-robust if $\gamma < \delta/(\delta+1)$.*

Bibliography

[1] W. Aiello, A. Bonato, C. Cooper, J. Janssen, and P. Prałat. "A spatial web graph model with local influence regions". *Internet Math.* 5.1-2 (2008), pages 175–196. ISSN: 1542-7951.

[2] M. Aizenman and C. M. Newman. "Discontinuity of the percolation density in one-dimensional $1/|x-y|^2$ percolation models". *Comm. Math. Phys.* 107.4 (1986), pages 611–647. ISSN: 0010-3616.

[3] D. Aldous and J. M. Steele. "The Objective Method: Probabilistic Combinatorial Optimization and Local Weak Convergence". In: *Probability on Discrete Structures*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pages 1–72. ISBN: 978-3-662-09444-0.

[4] A.-L. Barabási and R. Albert. "Emergence of Scaling in Random Networks". *Science* 286.5439 (1999), pages 509–512. ISSN: 0036-8075. eprint: http://science.sciencemag.org/content/286/5439/509.full.pdf.

[5] I. Benjamini and O. Schramm. "Recurrence of Distributional Limits of Finite Planar Graphs". *Electronic Journal of Probability* 6.none (2001), pages 1–13.

[6] J. van den Berg and H. Kesten. "Inequalities with applications to percolation and reliability". *J. Appl. Probab.* 22.3 (1985), pages 556–569. ISSN: 0021-9002.

[7] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular Variation*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1987.

[8] M. Bode, N. Fountoulakis, and T. Müller. "On the largest component of a hyperbolic model of complex networks". *Electron. J. Combin.* 22.3 (2015), Paper 3.24, 46.

[9] M. Boguñá, F. Papadopoulos, and D. Krioukov. “Sustaining the Internet with hyperbolic mapping”. *Nature Communications* 1.1 (2010), page 62.

[10] B. Bollobás, O. Riordan, J. Spencer, and G. Tusnády. “The degree sequence of a scale-free random graph process”. *Random Struct Algorithms* 18.3 (2001), pages 279–290. ISSN: 1042-9832.

[11] K. Bringmann, R. Keusch, and J. Lengler. “Sampling geometric inhomogeneous random graphs in linear time”. In: *25th European Symposium on Algorithms*. Volume 87. LIPIcs. Leibniz Int. Proc. Inform. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2017, Art. No. 20, 15.

[12] K. Bringmann, R. Keusch, and J. Lengler. “Geometric inhomogeneous random graphs”. *Theoret. Comput. Sci.* 760 (2019), pages 35–54. ISSN: 0304-3975.

[13] S. R. Broadbent and J. M. Hammersley. “Percolation processes: I. Crystals and mazes”. *Mathematical Proceedings of the Cambridge Philosophical Society* 53.3 (1957), pages 629–641.

[14] R. M. Burton and M. Keane. “Density and uniqueness in percolation”. *Comm. Math. Phys.* 121.3 (1989), pages 501–505. ISSN: 0010-3616.

[15] F. Chung and L. Lu. *Complex graphs and networks*. Volume 107. CBMS Regional Conference Series in Mathematics. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 2006, pages viii+264. ISBN: 978-0-8218-3657-6; 0-8218-3657-9.

[16] C. Cooper, A. Frieze, and P. Prałat. “Some typical properties of the spatial preferred attachment model”. In: *Algorithms and models for the web graph*. Volume 7323. Lecture Notes in Comput. Sci. Springer, Heidelberg, 2012, pages 29–40.

[17] M. Deijfen, R. van der Hofstad, and G. Hooghiemstra. “Scale-free percolation”. *Ann. Inst. Henri Poincaré Probab. Stat.* 49.3 (2013), pages 817–838. ISSN: 0246-0203.

[18] P. Deprez and M. V. Wüthrich. “Scale-free percolation in continuum space”. *Commun. Math. Stat.* 7.3 (2019), pages 269–308. ISSN: 2194-6701.

[19] S. Dereich, C. Mönch, and P. Mörters. “Typical distances in ultrasmall random networks”. *Adv. in Appl. Probab.* 44.2 (June 2012), pages 583–601.

[20] S. Dereich and P. Mörters. "Random networks with sublinear preferential attachment: The giant component". *Ann. Probab.* 41.1 (Jan. 2013), pages 329–384.

[21] S. Dommers, R. van der Hofstad, and G. Hooghiemstra. "Diameters in preferential attachment models". *J. Stat. Phys.* 139.1 (2010), pages 72–107. ISSN: 0022-4715.

[22] S. Dorogovtsev, J. F. Mendes, and A. Samukhin. "Structure of Growing Networks with Preferential Linking". *Phys. Rev. Lett.* 85 (Nov. 2000), pages 4633–6.

[23] S. N. Dorogovtsev. *Lectures on complex networks*. Volume 20. Oxford Master Series in Physics. Oxford Master Series in Statistical Computational, and Theoretical Physics. Oxford University Press, Oxford, 2010, pages x+134. ISBN: 978-0-19-954893-4.

[24] H. Duminil-Copin, C. Garban, and V. Tassion. *Long-range models in 1D revisited*. 2020. arXiv: `2011.04642 [math.PR]`.

[25] M. Eckhoff and P. Mörters. "Vulnerability of robust preferential attachment networks". *Electron. J. Probab.* 19 (2014), no. 57, 47.

[26] M. Eckhoff, P. Mörters, and M. Ortgiese. "Near critical preferential attachment networks have small giant components". *J. Stat. Phys.* 173.3-4 (2018), pages 663–703. ISSN: 0022-4715.

[27] A. D. Flaxman, A. M. Frieze, and J. Vera. "A geometric preferential attachment model of networks". *Internet Math.* 3.2 (2006), pages 187–205. ISSN: 1542-7951.

[28] A. D. Flaxman, A. M. Frieze, and J. Vera. "A geometric preferential attachment model of networks. II". In: *Algorithms and models for the web-graph*. Volume 4863. Lecture Notes in Comput. Sci. Springer, Berlin, 2007, pages 41–55.

[29] C. M. Fortuin, P. W. Kasteleyn, and J. Ginibre. "Correlation inequalities on some partially ordered sets". *Comm. Math. Phys.* 22 (1971), pages 89–103. ISSN: 0010-3616.

[30] E. N. Gilbert. "Random plane networks". *J. Soc. Indust. Appl. Math.* 9 (1961), pages 533–543. ISSN: 0368-4245.

[31] J.-B. Gouéré. "Subcritical regimes in the Poisson Boolean model of continuum percolation". *Ann. Probab.* 36.4 (2008), pages 1209–1220. ISSN: 0091-1798.

[32] P. Gracar, A. Grauer, L. Lüchtrath, and P. Mörters. "The age-dependent random connection model". *Queueing Syst.* 93.3-4 (2019), pages 309–331. ISSN: 0257-0130.

[33] P. Gracar, A. Grauer, and P. Mörters. *Chemical distance in geometric random graphs with long edges and scale-free degree distribution.* 2021. arXiv: 2108.11252 [math.PR].

[34] P. Gracar, M. Heydenreich, C. Mönch, and P. Mörters. *Recurrence versus Transience for Weight-Dependent Random Connection Models.* 2019. arXiv: 1911.04350 [math.PR].

[35] P. Gracar, L. Lüchtrath, and C. Mönch. *Finiteness of the percolation threshold for inhomogeneous long-range models in one dimension.* 2022. arXiv: 2203.11966 [math.PR].

[36] P. Gracar, L. Lüchtrath, and P. Mörters. "Percolation phase transition in weight-dependent random connection models". *Adv. in Appl. Probab.* 53.4 (2021), pages 1090–1114. ISSN: 0001-8678.

[37] G. R. Grimmett and A. M. Stacey. "Critical probabilities for site and bond percolation models". *Ann. Probab.* 26.4 (1998), pages 1788–1812. ISSN: 0091-1798.

[38] G. Grimmett. *Percolation*. Second. Volume 321. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, 1999, pages xiv+444. ISBN: 3-540-64902-6.

[39] P. Hall. "On continuum percolation". *Ann. Probab.* 13.4 (1985), pages 1250–1266. ISSN: 0091-1798.

[40] M. Heydenreich, R. van der Hofstad, G. Last, and K. Matzke. *Lace Expansion and Mean-Field Behavior for the Random Connection Model.* 2020. arXiv: 1908.11356 [math.PR].

[41] C. Hirsch. "From heavy-tailed Boolean models to scale-free Gilbert graphs". *Brazilian Journal of Probability and Statistics* 31.1 (2017), pages 111–143.

[42] R. van der Hofstad. *Random graphs and complex networks. Vol. 1.* Cambridge Series in Statistical and Probabilistic Mathematics, [43]. Cambridge University Press, Cambridge, 2017, pages xvi+321. ISBN: 978-1-107-17287-6.

[43] R. van der Hofstad, P. van der Hoorn, and N. Maitra. *Local limits of spatial inhomogeneous random graphs.* 2021. arXiv: 2107.08733 [math.PR].

[44] E. Jacob and P. Mörters. "Spatial preferential attachment networks: power laws and clustering coefficients". *Ann. Appl. Probab.* 25.2 (2015), pages 632–662. ISSN: 1050-5164.

[45] E. Jacob and P. Mörters. "Robustness of scale-free spatial networks". *Ann. Probab.* 45.3 (2017), pages 1680–1722. ISSN: 0091-1798.

[46] B. Jahnel, A. Tóbiás, and E. Cali. *Phase transitions for the Boolean model of continuum percolation for Cox point processes*. 2020. arXiv: 2003.06206 [math.PR].

[47] J. Janssen and R. Prałat Pawełand Wilson. "Geometric graph properties of the spatial preferred attachment model". *Adv. in Appl. Math.* 50.2 (2013), pages 243–267. ISSN: 0196-8858.

[48] J. Jordan. "The degree sequences and spectra of scale-free random graphs". *Random Struct Algorithms* 29.2 (2006), pages 226–242.

[49] J. Jordan. "Degree sequences of geometric preferential attachment graphs". *Adv. in Appl. Probab.* 42.2 (2010), pages 319–330. ISSN: 0001-8678.

[50] J. Jordan. "Geometric preferential attachment in non-uniform metric spaces". *Electron. J. Probab.* 18 (2013), no. 8, 15. ISSN: 1083-6489.

[51] J. Jordan and A. R. Wade. "Phase transitions for random geometric preferential attachment graphs". *Adv. in Appl. Probab.* 47.2 (2015), pages 565–588. ISSN: 0001-8678.

[52] 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 (2021), pages 1968–2016.

[53] J. Komjáthy and B. Lodewijks. "Explosion in weighted hyperbolic random graphs and geometric inhomogeneous random graphs". *Stochastic Processes and their Applications* 130.3 (2020), pages 1309–1367. ISSN: 0304-4149.

[54] D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, and M. Boguñá. "Hyperbolic geometry of complex networks". *Phys. Rev. E (3)* 82.3 (2010), pages 036106, 18. ISSN: 1539-3755.

[55] G. Last and M. Penrose. *Lectures on the Poisson Process*. Cambridge University Press, Oct. 2017.

[56] S. S. Manna and P. Sen. "Modulated scale-free network in Euclidean space". *Phys. Rev. E* 66 (6 Dec. 2002), page 066114.

[57] R. Meester, M. D. Penrose, and A. Sarkar. "The random connection model in high dimensions". *Statistics & Probability Letters* 35.2 (1997), pages 145–153. ISSN: 0167-7152.

[58] R. Meester and R. Roy. "Uniqueness of unbounded occupied and vacant components in Boolean models". *Ann. Appl. Probab.* 4.3 (1994), pages 933–951. ISSN: 1050-5164.

[59] R. Meester and R. Roy. *Continuum percolation.* Volume 119. Cambridge Tracts in Mathematics. Cambridge University Press, Cambridge, 1996, pages x + 238. ISBN: 0-521-47504-X.

[60] R. Meester and J. E. Steif. "On the continuity of the critical value for long range percolation in the exponential case". *Comm. Math. Phys.* 180.2 (1996), pages 483–504. ISSN: 0010-3616.

[61] C. M. Newman and L. S. Schulman. "One-dimensional $1/|j-i|^s$ percolation models: the existence of a transition for $s \leq 2$". *Comm. Math. Phys.* 104.4 (1986), pages 547–571. ISSN: 0010-3616.

[62] M. E. J. Newman, S. H. Strogatz, and D. J. Watts. "Random graphs with arbitrary degree distributions and their applications". *Phys. Rev. E* 64 (2 July 2001), page 026118.

[63] M. Penrose. *Random geometric graphs.* Volume 5. Oxford Studies in Probability. Oxford University Press, Oxford, 2003. ISBN: 0-19-850626-0.

[64] M. D. Penrose. "On a continuum percolation model". *Advances in Applied Probability* 23.3 (1991), pages 536–556.

[65] M. D. Penrose. "Connectivity of soft random geometric graphs". *The Annals of Applied Probability* 26.2 (2016), pages 986–1028.

[66] M. D. Penrose and J. E. Yukich. "Weak laws of large numbers in geometric probability". *Ann. Appl. Probab.* 13.1 (2003), pages 277–303. ISSN: 1050-5164.

[67] L. S. Schulman. "Long range percolation in one dimension". *J. Phys. A* 16.17 (1983), pages L639–L641. ISSN: 0305-4470.

[68] M. Á. Serrano, D. Krioukov, and M. Boguñá. "Self-Similarity of Complex Networks and Hidden Metric Spaces". *Phys. Rev. Lett.* 100 (7 Feb. 2008), page 078701.

[69] D. J. Watts and S. H. Strogatz. "Collective dynamics of 'small-world' networks". *Nature* 393 (June 1998), pages 440–442.

[70] J. E. Yukich. "Ultra-small scale-free geometric networks". *J. Appl. Probab.* 43.3 (2006), pages 665–677. ISSN: 0021-9002.

APPENDIX A

Related work and open problems

We list some related work on the weight-dependent random connection model and state some open problems. This list is far from comprehensive as a lot of interesting work in the context of inhomogeneous percolation models is currently happening.

- In [34], Gracar et al. study the question of recurrence vs. transience of the weight-dependent random connection model.
- In [33], Gracar et al. study chemical distances for general graphs which also includes instances of the weight-dependent random connection model. They precisely identify the ultra small regime for the graphs studied in this thesis.
- Recall the open problem stated in Remark 1.14 (ii): Does the age-dependent random connection model have a supercritical phase in dimension one for natural examples of profile-functions in the regime $\gamma < 1/2$. This includes the weak-kernel model at critical $\gamma = 0$ also.
- In [31], Gouéré studies the tail of the Euclidean diameter of the root's component in a subcritical regime of the hard Poisson Boolean model. Denote by
$$D := \sup\{|x|^d : \mathbf{x} \leftrightarrow \mathbf{0}\}.$$

Then, translated to our parametrisation, Gouéré proves that there exists a $\beta_c > 0$ such that for all $\beta < \beta_c$

$$\mathbb{P}^\beta\{D \geq K\} \asymp K^{1-1/\gamma}, \qquad \text{as } K \to \infty.$$

What is the tail of D in the soft version of the model? The event $\{D \geq K\}$ is the event that there exists a vertex $\mathbf{x}$ outside $\mathcal{B}(0, K^{1/d})$ that is connected to $\mathbf{0}$ by a path where all the other vertices on the path lie within the ball of radius $K^{1/d}$. Using the strong kernel (1.11), a lower bound is given by the probability of the event that there exists a vertex with marker in $(K^{-\delta/\gamma(\delta+1)}, t_x \vee t_0$ within the ball $\mathcal{B}(0, K^{1/d})$ that is connected to x and $\mathbf{0}$. The probability of this event happening is of order $K^{1-\frac{\delta}{\gamma(\delta+1)}}$. For $\gamma < 1/2$ an upper bound of the same order can be calculated for significantly small β using a first moment bound on the number of paths of length n connecting $\mathbf{0}$ and $\mathbf{x}$ where the smallest mark within the path is larger than $K^{-\delta/\gamma(\delta+1)}$. The probability that there is vertex with even smaller mark within the ball is of the same order and hence

$$\mathbb{P}^\beta\{D \geq K\} \asymp K^{1-\frac{\delta}{\gamma(\delta+1)}}, \qquad \text{as } K \to \infty$$

if $\gamma < 1/2$. As this exponent matches the one calculated by Gouéré for the hard model when sending $\delta \to \infty$, one might expect that this exponent also holds in the remaining regime $\gamma \in [1/2, \delta/(\delta+1))$. An interesting open problem is to prove this. A project to work on this problem with Marcel Ortgiese has been initiated.

- In [25], Eckhoff and Mörters prove vulnerability for the non spatial preferential attachment model. They show that in the robust regime the network is no longer scale-free and robust if an ε-proportion of the oldest vertices is removed, cf. Remark 1.6(iv). As the geometric restrictions of our model make it harder to be robust, the same holds true in our model. Indeed, if working in the rescaled graph $\mathscr{G}^t$, the removal of the εt oldest vertices coincides with restricting the vertex marks to the interval $(\varepsilon, 1)$. Therefore, we can couple this graph with a homogeneous random geometric graph and immediately get the absence of scale-freeness, and non-robustness. The question of interest is hence: what is the correct amount of old vertices that have to be removed? Is it still necessary to remove a linear proportion or is a strictly smaller order, e.g. $\log t$, sufficient.

APPENDIX B

Integration results

Lemma B.1. *Let* $\gamma \in (0,1)$ *and* $t_0 \in (0,1)$*. Then,*

(a) for all $k \in \mathbb{N}$*, we have*

$$\int_{t_0}^{1} \mathrm{d}s_1 \int_{s_1}^{1} \mathrm{d}s_2 \cdots \int_{s_{k-1}}^{1} \mathrm{d}s_k \left[s_0^{-\gamma} \left(\prod_{j=1}^{k-1} s_j^{-1} \right) s_k^{\gamma-1} \right] \leq \frac{s_0^{-\gamma} \log^{k-1}(1/s_0)}{\gamma(k-1)!}.$$

(b) for all $k \in \mathbb{N}$*, we have*

$$\int_0^1 \mathrm{d}s \frac{s^{-\gamma} \log^k(1/s)}{k!} = \left(\frac{1}{1-\gamma} \right)^{k+1}.$$

Proof. We prove (a) by induction. For $k = 1$, we have

$$s_0^{-\gamma} \int_{s_0}^{1} \mathrm{d}s_1 \, s_1^{\gamma-1} \leq \frac{s_0^{-\gamma}}{\gamma}.$$

For $k+1$ we get using the induction hypothesis

$$\begin{aligned}\int_{s_0}^1 \mathrm{d}s_1 \int_{s_1}^1 \mathrm{d}s_2 \cdots \int_{s_k}^1 \mathrm{d}s_{k+1}\left[s_0^{-\gamma}\left(\prod_{j=1}^k s_j^{-1}\right)s_{k+1}^{\gamma-1}\right]&\\ \leq s_0^{-\gamma}\int_{s_0}^1 \mathrm{d}s_1 \frac{s_1^{-1}\log^{k-1}(1/s_1)}{\gamma(k-1)!}&\\ = \frac{s_0^{-\gamma}(-1)^{k-1}}{\gamma(k-1)!}\int_{s_0}^1 \mathrm{d}s_1 \log(s_1)'\log^{k-1}(s_1)&\\ = \frac{s_0^{-\gamma}\log^k(1/s_0)}{\gamma k!}.&\end{aligned}$$

We prove (b) by induction as well. As $\gamma < 1$, we get, for $k=1$ using integration by parts

$$\int_0^1 \mathrm{d}s \frac{s^{-\gamma}\log(1/s)}{1!} = \int_0^1 \mathrm{d}s \frac{t^{-\gamma}}{1-\gamma} = \frac{1}{(1-\gamma)^2}.$$

Analogously for $k+1$,

$$\int_0^1 \mathrm{d}s \frac{s^{-\gamma}\log^{k+1}(1/s)}{(k+1)!} = \int_0^1 \mathrm{d}s \frac{s^{-\gamma}\log^k(1/s)}{(1-\gamma)k!} = \frac{1}{(1-\gamma)^{k+2}}$$

by the induction hypothesis. □

Lemma B.2. *Let $\gamma \in (1/2,1)$ and $x \in (0,1)$. Then, for all $k \in \mathbb{N}$, it holds*

$$\int_x^1 \mathrm{d}s\ \frac{s^{-2\gamma}\log^k(1/s)}{k!} \leq \frac{x^{1-2\gamma}\log^k(1/x)}{(2\gamma-1)k!}.$$

Proof. Integration by parts yields

$$\begin{aligned}\int_x^1 \mathrm{d}s\ \frac{s^{-2\gamma}\log^k(1/s)}{k!} &= \frac{x^{1-2\gamma}\log^k(1/x)}{(2\gamma-1)k!} - \int_x^1 \mathrm{d}s \frac{s^{-2\gamma}\log^{k-1}(1/s)}{(2\gamma-1)(k-1)!}\\ &\leq \frac{x^{1-2\gamma}\log^k(1/x)}{(2\gamma-1)k!},\end{aligned}$$

as the second integral is bounded from below by 0. □

Lemma B.3. *Let $\gamma \in (0,1)$, $x \in (0,1)$ and $s_0 \in (x,1)$. Then, for all $k \in \mathbb{N}$, it holds*

$$\int_x^{s_0} \mathrm{d}s_1 \int_x^{s_1} \mathrm{d}s_2 \cdots \int_x^{s_{k-1}} \mathrm{d}s_k \Big(s_0^{\gamma-1} \prod_{j=1}^{k} s_j^{-1} \Big) = \frac{s_0^{\gamma-1} \log^k(s_0/x)}{k!}.$$

Proof. For $k=1$, we get

$$\int_x^{s_0} \mathrm{d}s_1 \; s_0^{\gamma-1} s_1^{-1} = s_0^{\gamma-1} \log(s_0/x).$$

For $k+1$, using induction hypothesis, we get

$$\begin{aligned} \int_x^{s_0} \mathrm{d}s_1 \int_x^{s_1} \mathrm{d}s_2 \cdots \int_x^{s_k} \mathrm{d}s_{k+1} \Big(s_0^{\gamma-1} \prod_{j=1}^{k+1} s_j^{-1} \Big) &= s_0^{\gamma-1} \int_x^{s_0} \mathrm{d}s_1 \frac{s_1^{-1} \log^k(s_1/x)}{k!} \\ &= s_0^{\gamma-1} \int_0^{\log(s_0/x)} \mathrm{d}y \frac{y^k}{k!} = \frac{s_0^{\gamma-1} \log^{k+1}(s_0/x)}{(k+1)!}. \end{aligned}$$

□

Lemma B.4. *Let $\gamma \in (1/2,1)$ and $m,k \in \mathbb{N}$, such that $m \geq 2$ and $1 \leq k \leq m-1$. Further, let $x \in (0,1)$. Then,*

$$\begin{aligned} &\int_x^1 \mathrm{d}s_0 \int_x^{s_0} \mathrm{d}s_1 \cdots \int_x^{s_{k-1}} \mathrm{d}s_k \Bigg[s_0^{\gamma-1} \Big(\prod_{j=1}^{k-1} s_j^{-1} \Big) s_k^{-\gamma} \\ &\qquad\qquad \times \int_{s_k}^1 \mathrm{d}s_{k+1} \cdots \int_{s_{m-1}}^1 \mathrm{d}s_m \Big[s_k^{-\gamma} \Big(\prod_{j=k+1}^{m-1} s_j^{-1} \Big) s_m^{\gamma-1} \Big] \Bigg] \qquad \text{(B.1)} \\ &\leq \binom{m-2}{k-1} \frac{x^{1-2\gamma} \log^{m-2}(1/x)}{\gamma^2 (2\gamma-1)(m-2)!}. \end{aligned}$$

Proof. We apply the previous lemmas. By Lemma B.1, we get

$$\int_{s_k}^1 \mathrm{d}s_{k+1} \cdots \int_{s_{m-1}}^1 \mathrm{d}s_m \Big[s_k^{-\gamma} \Big(\prod_{j=k+1}^{m-1} s_j^{-1} \Big) s_m^{\gamma-1} \Big] \leq \frac{s_k^{-\gamma} \log^{m-k-1}(1/s_k)}{\gamma(m-k-1)!}.$$

Therefore, the integral in (B.1) can be bound by

$$\int_x^1 \mathrm{d}s_0 \int_x^{s_0} \mathrm{d}s_1 \cdots \int_x^{s_{k-2}} \mathrm{d}s_{k-1} \left[s_0^{\gamma-1} \left(\prod_{j=1}^{k-1} s_j^{-1} \right) \int_x^{s_{k-1}} \mathrm{d}s_k \frac{s_k^{-2\gamma} \log^{m-k-1}(1/s_k)}{\gamma(m-k-1)!} \right].$$

By Lemma B.2

$$\int_x^{s_{k-1}} \mathrm{d}s_k \frac{s_k^{-2\gamma} \log^{m-k-1}(1/s_k)}{\gamma(m-k-1)!} \leq \frac{x^{1-2\gamma} \log^{m-k-1}(1/x)}{\gamma(2\gamma-1)(m-k-1)!}$$

and by Lemma B.3

$$\int_x^{s_0} \mathrm{d}s_1 \cdots \int_x^{s_{k-2}} \mathrm{d}s_{k-1}\ s_0^{\gamma-1} \left(\prod_{j=1}^{k-1} s_j^{-1} \right) = \frac{s_0^{\gamma-1} \log^{k-1}(s_0/x)}{(k-1)!}.$$

Therefore, the integral in (B.1) can be further bound by

$$\begin{aligned} \int_x^1 \mathrm{d}s_0 \frac{t_0^{\gamma-1} \log^{k-1}(s_0/x)}{(k-1)!} & \frac{x^{1-2\gamma} \log^{m-k-1}(1/x)}{\gamma(2\gamma-1)(m-k-1)!} \\ & \leq \binom{m-2}{k-1} \frac{x^{1-2\gamma} \log^{m-2}(1/x)}{\gamma(2\gamma-1)(m-2)!} \int_x^1 \mathrm{d}s_0\ s_0^{\gamma-1}. \end{aligned}$$

The result follows by integrating with respect to s_0. □

Lemma B.5. *Let $\gamma \in (0,1)$ and $k \in \mathbb{N}$. Then*

$$\int_0^1 \mathrm{d}s_0 \int_0^{s_0} \mathrm{d}s_1 \cdots \int_0^{s_{k-1}} \mathrm{d}s_k\ s_0^{\gamma-1} \left(\prod_{j=1}^{k-1} s_j^{-1} \right) s_k^{-\gamma} \leq \left(\frac{1}{1-\gamma} \right)^k.$$

Proof. We have

$$\begin{aligned} \int_0^1 \mathrm{d}s_0 \int_0^{s_0} \mathrm{d}s_1 \cdots \int_0^{s_{k-2}} \mathrm{d}s_{k-1} & \left[s_0^{\gamma-1} \left(\prod_{j=1}^{k-1} s_j^{-1} \right) \int_0^{s_{k-1}} \mathrm{d}s_k\ s_k^{-\gamma} \right] \\ &= \frac{1}{1-\gamma} \int_0^1 \mathrm{d}s_0 \int_0^{s_0} \mathrm{d}s_1 \cdots \int_0^{s_{k-2}} \mathrm{d}s_{k-1}\ s_0^{\gamma-1} \left(\prod_{j=1}^{k-2} s_j^{-1} \right) s_{k-1}^{-\gamma} \end{aligned}$$

and the result follows by repeating this across all integrals. □

APPENDIX C

Frequently used notation

Euclidean space and sets

$\mathbb{R}$	the real numbers
$\mathbb{R}^d$	d-dimensional column vectors with real entries, i.e $\mathbb{R}^d = \{(x_1, \dots, x_d)^T : x_j \in \mathbb{R} \text{ for } j = 1, \dots, d\}$
$\lvert x\rvert$	The Euclidean norm $\lvert x\rvert = \left(\sum_1^d x_j^2\right)^{1/2}$ of $x \in \mathbb{R}^d$
$\mathbb{T}_t^d$	the d-dimensional torus of volume t, cf. above (1.4)
d_t	the torus metric, cf. (1.4)
$<_{\text{lex}}$	strict lexicographic order in $\mathbb{R}^d$
$\mathcal{B}(x, r)$	the ball around x with radius r
$\sharp A$	number of elements of the set A
A^c	the complement of the set A

Graph Theory

$V(G)$	the vertex set of the graph G
$\mathscr{C}(G)$	the largest connected component of the graph G
$\mathscr{N}_{\mathbf{x}}(G)$	the set of neighbours in G of the vertex $\mathbf{x}$
$\mathscr{N}_{\mathbf{x}}^{<}$	set of neighbours of $\mathbf{x}$ with smaller mark
$\mathscr{N}_{\mathbf{x}}^{>}$	set of neighbours of $\mathbf{x}$ with larger mark
$\mathbf{x} \sim \mathbf{y}$	the vertices $\mathbf{x}$ and $\mathbf{y}$ are neighbours.
$\mathbf{x} \leftrightarrow \mathbf{y}$	$\mathbf{x}$ and $\mathbf{y}$ are connected by a path

Measures and random object

$(\mathscr{G}_t : t > 0)$	The age-based preferential attachment network
$(\mathscr{G}(\beta) : \beta > 0)$	The weight-dependent random connection model
η, η_0	A stationary, ergodic, simple point process on $\mathbb{R}^d$ and its Palm version. Mostly a Poisson process
P_0^{η}	The law of η_0
Uniform(A)	The uniform distribution on the set A
$\mathcal{T}_0$	A sequence of i.i.d. Uniform$(0,1)$ random variables, the vertex marks
$\mathcal{X}_0$	A point process on $\mathbb{R}^d \times (0,1)$, the marked vertex set
$\mathbf{P}_0, \mathbf{E}_0$	Law and expectation w.r.t. $\mathcal{X}$
$\mathcal{U}_0$	A sequence of i.i.d. Uniform$(0,1)$, the edge marks
$\mathcal{G}^{\beta}$	Graph functional that builds $\mathscr{G}(\beta)$
$\mathbb{P}^{\beta}, \mathbb{E}^{\beta}$	Law and expectation w.r.t $\mathcal{G}^{\beta}$

Asymptotics of non negative functions

$f = o(g)$	$\lim_{x\to\infty} f(x)/g(x) = 0$
$f = O(g)$	$\limsup_{x\to\infty} f(x)/g(x) < \infty$
$f = \Omega(g)$	$g = O(f)$
$f \asymp g$	f/g is bounded from zero and infinity

Erklärung zur Dissertation

gemäß der Promotionsordnung vom 12. März 2020

Diese Erklärung muss in der Dissertation enthalten sein.
(*This version must be included in the doctoral thesis*)

„Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation selbstständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel und Literatur angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten Werken dem Wortlaut oder dem Sinn nach entnommen wurden, sind als solche kenntlich gemacht. Ich versichere an Eides statt, dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie - abgesehen von unten angegebenen Teilpublikationen und eingebundenen Artikeln und Manuskripten - noch nicht veröffentlicht worden ist sowie, dass ich eine Veröffentlichung der Dissertation vor Abschluss der Promotion nicht ohne Genehmigung des Promotionsausschusses vornehmen werde. Die Bestimmungen dieser Ordnung sind mir bekannt. Darüber hinaus erkläre ich hiermit, dass ich die Ordnung zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit wissenschaftlichem Fehlverhalten der Universität zu Köln gelesen und sie bei der Durchführung der Dissertation zugrundeliegenden Arbeiten und der schriftlich verfassten Dissertation beachtet habe und verpflichte mich hiermit, die dort genannten Vorgaben bei allen wissenschaftlichen Tätigkeiten zu beachten und umzusetzen. Ich versichere, dass die eingereichte elektronische Fassung der eingereichten Druckfassung vollständig entspricht."

Teilpublikationen:

1) ‚The age-dependent random connection model‘. Queueing Syst. (2019), 93.3. Mit Peter Gracer, Arne Grauer, Peter Mörters

2) ‚Percolation phase transition in weight-dependent random connection models’. Adv. Appl. Probab (2021), 53.4. Mit Peter Gracar, Peter Mörters

3) ‚Finiteness of the percolation threshold for inhomogeneous long-range models in one dimension’. arXiv- Preprint: arXiv:2203.11966. Mit Peter Gracar, Christian Mönch

Datum, Name und Unterschrift

04.04.2022, Lukas Lüchtrath