PERCOLATION PHASE TRANSITION IN WEIGHT-DEPENDENT RANDOM CONNECTION MODELS

PETER GRACAR [ID],* **
LUKAS LÜCHTRATH [ID],* *** AND
PETER MÖRTERS [ID],* **** *University of Cologne*

Abstract

We investigate spatial random graphs defined on the points of a Poisson process in *d*-dimensional space, which combine scale-free degree distributions and long-range effects. Every Poisson point is assigned an independent weight. Given the weight and position of the points, we form an edge between any pair of points independently with a probability depending on the two weights of the points and their distance. Preference is given to short edges and connections to vertices with large weights. We characterize the parameter regime where there is a non-trivial percolation phase transition and show that it depends not only on the power-law exponent of the degree distribution but also on a geometric model parameter. We apply this result to characterize robustness of age-based spatial preferential attachment networks.

Keywords: Subcritical regime; random geometric graph; Boolean model; scale-free percolation; long-range percolation; spatial network; robustness; age-based spatial preferential attachment

2010 Mathematics Subject Classification: Primary 60K35
Secondary 05C80

1. Introduction and statement of results

1.1. Motivation

In classical continuum percolation theory, a graph is built with a Poisson point process in $\mathbb{R}^d$ as the vertex set. Two points are connected by an edge if their Euclidean distance is below a fixed or variable threshold. Assuming the resulting graph has an infinite component, one asks whether there exists an infinite component in the percolated graph where every edge is independently removed with probability $1 - p$ (and respectively retained with probability p). We say that the graph has a *percolation phase transition* if there is a critical probability $p_c > 0$ such that, almost surely, if $p < p_c$ there is no infinite component, and if $p > p_c$ there exists an infinite component in the percolated graph. It is known that there exists a percolation phase transition for the fixed threshold model in $\mathbb{R}^d$, often called the Boolean model, and for variable threshold models where the threshold is the sum of independent radii with finite *d*th moment

associated with the points [7, 8]. The result also extends to long-range percolation models, where the probability that two points are connected is a decreasing function of their distance; see [16, 18].

By contrast, the continuous version of the scale-free percolation model of van der Hofstad, Hooghiemstra and Deijfen [4] does not have a percolation phase transition if the power-law exponent satisfies $\tau < 3$; see for example [6, 12]. In fact, for many graphs combining scale-free degree distributions and long-range effects the problem of existence of a percolation phase transition is open. This includes, for example, models where the connection probability of two points is a decreasing function of the ratio of their distance and the sum or maximum of their radii. In this paper we look at a broad class of such graphs, the *weight-dependent random connection models*, and characterize the parameter regimes where there is a percolation phase transition. Other than in the scale-free percolation model, in this class a subcritical phase can only fail to exist if there is a sufficiently small power-law exponent combined with a strong long-range effect. The weight-dependent random connection models include the weak local limits of the age-based preferential attachment model introduced in [9]. We use this result to characterize the regimes when these network models are robust under random removal of edges, offering new insight into the notoriously difficult topic of spatial preferential attachment networks; see [14].

1.2. Framework

We introduce the weight-dependent random connection model as in [10]. The vertex set of the graph $\mathscr{G}$ is a Poisson point process of unit intensity on $\mathbb{R}^d \times (0, 1]$. We think of a Poisson point $\mathbf{x} = (x, t)$ as a *vertex* at *position* x with *weight* t^{-1}. Two vertices $\mathbf{x}$ and $\mathbf{y}$ are connected by an edge in $\mathscr{G}$ independently of any other (possible) edge with probability $\varphi(\mathbf{x}, \mathbf{y})$. Here, φ is a connectivity function

$$\varphi : \left(\mathbb{R}^d \times (0, 1]\right) \times \left(\mathbb{R}^d \times (0, 1]\right) \to [0, 1],$$

of the form

$$\varphi(\mathbf{x}, \mathbf{y}) = \varphi((x, t), (y, s)) = \rho\left(g(t, s)|x - y|^d\right)$$

for a non-increasing, integrable *profile function* $\rho : \mathbb{R}_+ \to [0, 1]$ and a function $g : (0, 1) \times (0, 1) \to \mathbb{R}_+$, which is symmetric and non-decreasing in both arguments. Hence, we give preference to short edges or edges that are connected to vertices with large weights. We also assume (without loss of generality) that

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

Then the degree distribution of a vertex depends only on the function g. However, the profile function controls the intensity of long edges in the graph.

We next give explicit examples for the function g we will focus on throughout the paper. We define the functions in terms of two parameters $\gamma \in (0, 1)$ and $\beta \in (0, \infty)$. The parameter γ describes the strength of the influence of the vertices' weights on the connection probability; the larger γ, the stronger the preference for connecting to vertices with large weight. All kernel functions we consider lead to models that are *scale-free* with power law exponent

$$\tau = 1 + \frac{1}{\gamma};$$

see [9, 10]. In particular, all graphs are locally finite; i.e. every vertex has finite degree. The parameter β is used to control the edge density; i.e. increasing β increases the expected number

of edges connected to a typical vertex [9]. Our focus is on the following three functions (for further examples, see [10]):

- The *sum kernel*, defined as

$$g^{\mathrm{sum}}(s, t) = \beta^{-1}(s^{-\gamma/d} + t^{-\gamma/d})^{-d}.$$

  The interpretation of $(\beta as^{-\gamma})^{1/d}, (\beta at^{-\gamma})^{1/d}$ as random radii together with $\rho(r) = \mathbb{1}_{[0,a]}(r)$ leads to the Boolean model in which two vertices are connected by an edge when their associated balls intersect.

- The *min kernel*, defined as

$$g^{\mathrm{min}}(s, t) = \beta^{-1}(s \wedge t)^{\gamma}.$$

  Here, in the case of an indicator profile function as above, two vertices are connected by an edge when one of them lies inside the ball associated with the other one. As $2^{-d}g^{\mathrm{min}} \leq g^{\mathrm{sum}} \leq g^{\mathrm{min}}$, the min kernel and the sum kernel show qualitatively similar behaviour.

- The *preferential attachment kernel*, defined as

$$g^{\mathrm{pa}}(s, t) = \beta^{-1}(s \vee t)^{1-\gamma}(s \wedge t)^{\gamma}. \tag{2}$$

  It gives rise to the *age-dependent random connection model* introduced by Gracar *et al.* [9]. This model is the weak local limit of the age-based spatial preferential attachment model which is an approximation of the spatial preferential attachment model introduced by Jacob and Mörters [13].

As we want to study the influence of long-range effects on the percolation problem, we focus primarily on profile functions that are *regularly varying* with index $-\delta$ for some $\delta > 1$, that is,

$$\lim_{r \uparrow \infty} \frac{\rho(cr)}{\rho(r)} = c^{-\delta} \quad \text{for all } c \geq 1. \tag{3}$$

A comparison argument can be used to derive the behaviour of profile functions with lighter tails (including those with bounded support) from a limit $\delta \uparrow \infty$.

We fix one of the kernels above, as well as γ, β, and δ. Let $p \in [0, 1]$ and perform Bernoulli bond percolation with retention parameter p on the graph $\mathscr{G}$; i.e., every edge of $\mathscr{G}$ remains intact independently with probability p, or is removed with probability $1 - p$. We denote the graph we obtain by $\mathscr{G}^p$ and ask whether there exists an infinite cluster, or equivalently an infinite self-avoiding path, in $\mathscr{G}^p$. If so, we say that the graph *percolates*. We define the *critical percolation parameter* p_c as the infimum of all parameters $p \in [0, 1]$ such that the percolation probability is positive. By the Kolmogorov 0–1 law, for all $1 \geq p > p_c$ the graph percolates and for all $0 \leq p < p_c$ the graph does not percolate, almost surely. We call the parameter range $(p_c, 1]$ the *supercritical phase* and $[0, p_c)$ the *subcritical phase*.

1.3. Main result: percolation phase transition

Our main result characterizes the parameter regime where there is a percolation phase transition in the weight-dependent random connection model.

Theorem 1.1. (Percolation phase transition.) *Suppose ρ satisfies (3) for some $\delta > 1$. Then, for the weight-dependent random connection model with preferential attachment kernel, sum kernel, or min kernel and parameters $\beta > 0$, $0 < \gamma < 1$, we have that*

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

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

Remarks:

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

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

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

$$p_c > \frac{1}{A}\Big(\frac{d(\delta(1-\gamma)-\gamma)(\delta-1)}{2^{d\delta+4}J(d)\beta\delta}\Big)^{\delta},$$

where $J(d) = \prod_{j=0}^{d-2}\int_0^\pi \sin^j(\alpha_j)d\alpha_j$ is the Jacobian of the d-dimensional sphere coordinates.

(ii) If $\gamma < \frac{\delta}{\delta+1}$, one can follow the argument for long-range percolation (see [17]) and check that if $d \geq 2$ or if $d = 1$ and $\delta < 2$, then there exists $\beta_c < \infty$ such that the graph percolates for all $\beta > \beta_c$, and fixing such a β we then get $p_c < 1$.

(iii) If $\gamma = \frac{\delta}{\delta+1}$, we do not expect a universal result; i.e. it depends on the exact form of the kernel g and the profile ρ whether $p_c = 0$ or not.

(iv) A variant of our arguments shows that if $\gamma < \frac{\delta}{\delta+1}$ and either $d \geq 2$ or $d = 1$ and $\delta < 2$, there exists $0 < \beta_c < \infty$ such that there does not exist an infinite component in $\mathcal{G}$ if $\beta < \beta_c$, but an infinite component does exist if $\beta > \beta_c$. By scaling the Poisson process we see that, if γ and δ are as above, $\beta > 0$ is fixed, and the intensity of the Poisson process is variable, say $\lambda > 0$, then there exists $0 < \lambda_c < \infty$ such that there does not exist an infinite component in $\mathcal{G}$ if $\lambda < \lambda_c$, but an infinite component does exist if $\lambda > \lambda_c$. If however $\gamma > \frac{\delta}{\delta+1}$, then there exists an infinite component in $\mathcal{G}$ regardless of the values of λ, $\beta > 0$.

(v) To understand the occurrence of the critical value $\gamma = \frac{\delta}{\delta+1}$, the calculation in Lemma 2.2 is key. There it is shown that for $\gamma < \frac{\delta}{\delta+1}$ and small p the probability that two sufficiently distant vertices are connected using an intermediate vertex of smaller weight is smaller than the probability of existence of a direct edge. If $\gamma > \frac{\delta}{\delta+1}$ a converse statement holds, and it is more likely that two vertices of large weight are connected by an intermediate vertex of small weight. The corresponding strategy enters into the construction of long paths in Lemma 3.1.

(vi) A continuum version of the *scale-free percolation* model introduced by Deijfen *et al.* [4, 12] is given by the *product kernel*

$$g^{\mathrm{prod}}(s, t) = \beta^{-1}s^\gamma t^\gamma;$$

see [5, 6] for more details. For this model it is known that there is no percolation phase transition if $\gamma > \frac{1}{2}$, but there is one if $\gamma < \frac{1}{2}$. As the product kernel and the preferential attachment kernel coincide for $\gamma = \frac{1}{2}$, it follows that the scale-free percolation model has $p_c > 0$ at the critical parameter $\gamma = \frac{1}{2}$ for a general class of profile functions ρ. For more information how to translate the parameters of that model to our setting, see [10, Table 2].

(vii) Our result also shows that for profile functions ρ that decay faster than any polynomial, there always exists a subcritical phase. This applies in particular to the Boolean model mentioned above where ρ is the indicator function; see also [7].

1.4. Robustness of age-based preferential attachment networks

Let $\mathscr{G}_0$ be the age-dependent random connection model with a vertex at the origin. That is, $\mathscr{G}_0$ is the graph with

- vertex set obtained from a standard Poisson point process in $\mathbb{R}^d \times (0, 1]$ with an additional point $\mathbf{0} = (0, U)$ placed at the origin with inverse weight, resp. birth time U, sampled independently from everything else from the uniform distribution on (0,1], and

- edges laid down independently with connection probabilities given by the preferential attachment kernel (2), i.e.

$$\varphi((x, t), (y, s)) = \rho(g^{\mathrm{pa}}(t, s)|x - y|^d).$$

Theorem 1.1 applies to the graph $\mathscr{G}_0$, which plays a special role as weak local limit in the sense of Benjamini and Schramm [1] of the age-based spatial preferential attachment model, which we now describe.

Let $\mathbb{T}_a^d = (-a^{1/d}/2, a^{1/d}/2]^d$ be the d-dimensional torus of volume a, endowed with the torus metric d defined by

$$d(x, y) = \min \left\{ |x - y + u| : u \in \left\{ -a^{1/d}, 0, a^{1/d} \right\}^d \right\}, \quad \text{for } x, y \in \mathbb{T}_a^d.$$

The *age-based (spatial) preferential attachment model* is a growing sequence of graphs $(\mathscr{G}_t)_{t \geq 0}$ on $\mathbb{T}_1^d$ defined as follows:

- The graph $\mathscr{G}_t$ at time $t = 0$ has neither vertices nor edges.

- Vertices arrive successively after exponential waiting times with parameter one and are placed uniformly on $\mathbb{T}_1^d$. We denote a vertex created at time s and placed in $y \in \mathbb{T}_1^d$ by $\mathbf{y} = (y, s)$.

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

$$\rho \left(\frac{1}{\beta} \frac{t \, d(x,y)^d}{(t/s)^\gamma} \right). \tag{4}$$

Note that the connection probability has the same form as the previously defined connection function φ, where the Euclidean distance is replaced by the torus distance.

We say that such a network $(\mathscr{G}_t)_{t \geq 0}$ has a *giant component* if its largest connected component is asymptotically of linear size. More precisely, let $|\mathscr{C}_t|$ be the size of the largest component in $\mathscr{G}_t$. Then $(\mathscr{G}_t)_{t \geq 0}$ has a giant component if

$$\lim_{\varepsilon \downarrow 0} \limsup_{t \to \infty} \mathbb{P}\left\{ \frac{1}{t} |\mathscr{C}_t| < \varepsilon \right\} = 0.$$

We say $(\mathscr{G}_t)_{t \geq 0}$ is *robust* if the percolated sequence $(\mathscr{G}_t^p)_{t \geq 0}$ has a giant component for every retention parameter $p > 0$. Otherwise we say the network is *non-robust*. The idea of this

definition is that a random attack cannot significantly affect the connectivity of a robust network.

Theorem 1.2. *Suppose ρ satisfies (3) for some $\delta > 1$ and $(\mathscr{G}_t)_{t\geq 0}$ is the age-based preferential attachment network with parameters $\beta > 0$ and $0 < \gamma < 1$. Then the network $(\mathscr{G}_t)_{t\geq 0}$ is robust if $\gamma > \frac{\delta}{\delta+1}$, but non-robust if $\gamma < \frac{\delta}{\delta+1}$.*

Remarks:

- As $\tau = 1 + \frac{1}{\gamma}$ the condition $\gamma < \frac{\delta}{\delta+1}$ is equivalent to $\tau > 2 + \frac{1}{\delta}$. Hence the qualitative change in the behaviour does not occur when τ passes the critical value 3 as in the classical scale-free network models without spatial correlations, but when it passes a strictly smaller value. This shows the significant effect of clustering on the network topology.

- Replacing $(t/s)^\gamma$ in (4) by $f(\text{indegree of } (y, s) \text{ in } \mathscr{G}_{t-})$, for some increasing function f, we obtain the spatial preferential attachment model of [13]. If f is a function of asymptotic linear slope γ, then $(t/s)^\gamma$ is the asymptotic expected degree at time t of a vertex born at time s. The age-based preferential attachment model is therefore a simplification and approximation of the spatial preferential attachment model showing very similar behaviour. In [14] Jacob and Mörters show that the spatial preferential attachment model is robust for $\gamma > \frac{\delta}{\delta+1}$, but it remains an open problem to show non-robustness for $\gamma < \frac{\delta}{\delta+1}$ for this model. Theorem 1.2 is a strong indication that this is the case.

The remainder of the paper is organized as follows. In Section 2 we prove existence of a percolation phase transition as claimed in Theorem 1.1(a). This proof is based on a novel path decomposition argument and constitutes the main new contribution of this paper. The remaining proofs are similar to the corresponding arguments for spatial preferential attachment in [13, 14], namely the absence of a phase transition in Theorem 1.1(b) in Section 3 and the proof of Theorem 1.2, in Section 4, and will only be sketched. Some technical calculations are deferred to the appendix.

2. Existence of a subcritical phase

In this section, we prove Theorem 1.1(a). This proof works for all kernels g which are bounded from below by a constant multiple of the preferential attachment kernel g^{pa}; similarly, the proof of Theorem 1.1(b) given in Section 3 works for all kernels bounded from above by a multiple of the min kernel $g^{\min}$.

2.1. Graphical construction of the model

We explicitly construct the weight-dependent random connection model on a given countable set $\mathcal{Y} \subset \mathbb{R}^d \times (0, 1]$. Let $E(\mathcal{Y}) = \{\{\mathbf{x}, \mathbf{y}\} : \mathbf{x}, \mathbf{y} \in \mathcal{Y}\}$ be the set of potential edges and $\mathcal{V} = (\mathcal{V}(e))_{e \in E(\mathcal{Y})}$ a sequence in $[0,1]$ indexed by the potential edges. We then construct the graph $\mathcal{G}_\varphi(\mathcal{Y}, \mathcal{V})$ through its vertex set $\mathcal{Y}$ and edge set

$$\big\{\{\mathbf{x}, \mathbf{y}\} : \mathcal{V}(\{\mathbf{x}, \mathbf{y}\}) \leq \varphi(\mathbf{x}, \mathbf{y})\big\}.$$

Let $\mathcal{X}$ be a Poisson point process on $\mathbb{R}^d \times (0, 1]$ and $\mathcal{U} = (\mathcal{U}(e))_{e \in E(\mathcal{X})}$ an independent sequence of random variables uniformly distributed in $(0,1)$; then $\mathscr{G} = \mathcal{G}_\varphi(\mathcal{X}, \mathcal{U})$ is the weight-dependent random connection model with connectivity function φ. If $p \in (0, 1]$ then $\mathscr{G}^p =$

$\mathcal{G}_{p\varphi}(\mathcal{X}, \mathcal{U})$ is the percolated model with retention parameter p. Add to $\mathcal{X}$ a vertex $\mathbf{0} = (0, U)$, placed at the origin with inverse weight U distributed uniformly on $(0,1)$, independent of everything else, and denote the resulting point process by $\mathcal{X}_0$. Insert further independent uniformly distributed random variables $(U_{\{\mathbf{0},\mathbf{x}\}})_{\mathbf{x}\in\mathcal{X}}$ into the family $\mathcal{U}$, and denote the result by $\mathcal{U}_0$ and the underlying probability measure by $\mathbb{P}_0$. The graph $\mathcal{G}_0^p = \mathcal{G}_{p\varphi}(\mathcal{X}_0, \mathcal{U}_0)$ is the Palm version of $\mathcal{G}^p$; we denote its law by $\mathbb{P}_0^p$ and expectation by $\mathbb{E}_0^p$. Writing $\mathbb{P}_{(x,t)}^p$ for the law of $\mathcal{G}^p$ conditioned on the event that (x, t) is a vertex of $\mathcal{G}^p$, we have $\mathbb{P}_0^p = \mathbb{P}_{(0,u)}^p \mathrm{d}u$. Roughly speaking, this construction ensures that $\mathbf{0}$ is a typical vertex in $\mathcal{G}_0^p$.

2.2. Percolation

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

$$\theta(p) = \mathbb{P}_0^p\{\mathbf{0} \leftrightarrow \infty\} = \int_0^1 \mathrm{d}u\, \mathbb{P}_{(0,u)}^p\{(0, u) \leftrightarrow \infty\}, \tag{5}$$

which can be interpreted as the probability that a typical vertex belongs to the infinite cluster. We define the critical percolation parameter as

$$p_c := \inf\{p \in (0, 1]: \theta(p) > 0\}. \tag{6}$$

2.3. Existence of a subcritical phase: case $\gamma < \frac{1}{2}$

We fix $\delta > 1$, $\beta > 0$, and $\gamma < \frac{\delta}{\delta+1}$. Since $g^{\mathrm{pa}} \leq g^{\min} \leq 2^d g^{\mathrm{sum}}$, we have

$$\mathbb{P}_0\{\mathbf{0} \leftrightarrow \infty \text{ in } \mathcal{G}_0^p(\rho \circ g^{\mathrm{pa}})\} \geq \mathbb{P}_0\{\mathbf{0} \leftrightarrow \infty \text{ in } \mathcal{G}_0^p(\rho \circ g^{\min})\}$$

$$\geq \mathbb{P}_0\{\mathbf{0} \leftrightarrow \infty \text{ in } \mathcal{G}_0^{2^d p}(\tilde{\rho} \circ g^{\mathrm{sum}})\}$$

for $\tilde{\rho}(x) = \frac{1}{2^d}\rho(2^d x)$ by a simple coupling argument. Thus, we focus on the preferential attachment kernel and show that we can choose a $p > 0$ such that $\theta(p) = 0$. Consequently, in the following we work exclusively in the age-dependent random connection model, and we therefore use the corresponding terminology. For a vertex $\mathbf{x} = (x, t)$ we refer to t as the *birth time* of $\mathbf{x}$ and, for another vertex $\mathbf{y} = (y, s)$ with $s < t$, we say $\mathbf{y}$ is *older* than $\mathbf{x}$. We also say $\mathbf{y}$ is born before $\mathbf{x}$, or before t.

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 $\frac{1}{2} \leq \gamma < \frac{\delta}{\delta+1}$. We denote by $\mathbf{E}$ the expectation of a Poisson point process on $\mathbb{R}^d \times (0, 1]$ of unit intensity, by $\mathbb{P}_{\mathcal{X}}^p$ the law of $\mathcal{G}^p$ conditioned on the whole vertex set $\mathcal{X}$, and by $\mathbb{P}_{\mathbf{x}_1,\dots,\mathbf{x}_n}^p$ the law of $\mathcal{G}^p$ conditioned on the event that $\mathbf{x}_1, \dots, \mathbf{x}_n$ are points of the vertex set.

Lemma 2.1. *If* $0 < \gamma < \frac{1}{2}$, *then* $\theta(p) = 0$ *for all* $p < \frac{1-2\gamma}{4\beta}$ *or, equivalently,* $p_c \geq \frac{1-2\gamma}{4\beta}$.

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

$$\theta(p) = \lim_{n \to \infty} \mathbb{P}_0^p\{\exists \text{ a path of length } n \text{ starting in } \mathbf{x}_0\}$$

$$\leq \lim_{n \to \infty} \int_0^1 \mathrm{d}t_0 \, \mathbf{E}\left[\sum_{\substack{\mathbf{x}_1,\ldots,\mathbf{x}_n \in \mathcal{X} \\ x_i \neq x_j \forall i \neq j}} \mathbb{P}_{\mathcal{X} \cup \{(0,t_0)\}}^p \left(\bigcap_{j=1}^n \{\mathbf{x}_j \sim \mathbf{x}_{j-1}\} \right) \right].$$

The inner probability is a measurable function of the Poisson process and the points $\mathbf{x}_1, \ldots, \mathbf{x}_n$, and by Mecke's equation [15, Theorem 4.4] we get, with η denoting an independent copy of $\mathcal{X}$,

$$\int_0^1 \mathrm{d}t_0 \int_{(\mathbb{R}^d \times (0,1])^n} \bigotimes_{j=1}^n \mathrm{d}\mathbf{x}_j \, \mathbf{E}\left[\mathbb{P}_{\eta \cup \{(0,t_0),\mathbf{x}_1,\ldots,\mathbf{x}_n\}}^p \left(\bigcap_{j=1}^n \{\mathbf{x}_{j-1} \sim \mathbf{x}_j\} \right) \right]$$

$$= \int_0^1 \mathrm{d}t_0 \int_{(\mathbb{R}^d \times (0,1])^n} \bigotimes_{j=1}^n \mathrm{d}\mathbf{x}_j \, \mathbb{P}_{\mathbf{x}_0,\ldots,\mathbf{x}_n}^p \left(\bigcap_{j=1}^n \{\mathbf{x}_{j-1} \sim \mathbf{x}_j\} \right).$$

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

$$\int_0^1 \mathrm{d}t_0 \int_{(\mathbb{R}^d \times (0,1])^n} \bigotimes_{j=1}^n \mathrm{d}(x_j, t_j) \left(\prod_{j=1}^n p\rho\big(g^{\mathrm{pa}}(t_{j-1}, t_j)|x_j - x_{j-1}|^d\big) \right)$$

$$= p^n \beta^n \int_0^1 \mathrm{d}t_0 \int_0^1 \mathrm{d}t_1 \ldots \int_0^1 \mathrm{d}t_n \left(\prod_{j=1}^n (t_j \wedge t_{j-1})^{-\gamma}(t_j \vee t_{j-1})^{\gamma-1} \right),$$

where we used the normalization condition (1). Since $\gamma < \frac{1}{2}$, Lemma 17 of [14] states that

$$\int_0^1 \mathrm{d}t_0 \int_0^1 \mathrm{d}t_1 \ldots \int_0^1 \mathrm{d}t_n \left(\prod_{j=1}^n (t_j \wedge t_{j-1})^{-\gamma}(t_j \vee t_{j-1})^{\gamma-1} \right) \leq \left(\frac{1}{1+\alpha-\gamma} - \frac{1}{\alpha+\gamma} \right)^n,$$

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 $p < \frac{1-2\gamma}{4\beta}$, we achieve

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

$\square$

Existence of a subcritical phase: case $\gamma \geq \frac{1}{2}$

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 to prove our claim. Intuitively, as 'powerful' vertices are typically far apart from each other, 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, $\mathscr{G}$ 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 $\gamma < \frac{1}{2}$, the degrees

of typical vertices are small enough so that the number of possible paths does not increase too much. This is not true anymore 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 birth-time structure. In the end, we will use these paths to complete the proof of Theorem 1.1(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 analogously define $i_k = \max\{i > i_{k-1} : v_i \sim v_{i_{k-1}}\}$ 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 $P = ((v_0, t_0), (v_1, t_1), \dots, (v_n, t_n))$ be a path of length n in some graph G where every vertex v_i carries a distinct birth time t_i. Then precisely one of the vertices in P is the oldest; let $k_{\min} = \{k \in \{0, \dots, n\} : t_k < t_j, \ \forall j \neq k\}$ be its index. Starting from (v_0, t_0), we now choose the first vertex of the path that has birth time smaller than t_0 and call it (v_{i_1}, t_{i_1}). Continuing from this vertex, we choose the next vertex of the path that is older still, call it (v_{i_2}, t_{i_2}), and continue analogously until we reach the oldest vertex $(v_{k_{\min}}, t_{k_{\min}})$. We then repeat the same procedure starting from the end vertex (v_n, t_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 $P = ((v_0, t_0), \dots, (v_n, t_n))$, there exist unique $0 \leq k \leq n$ and $k \leq m \leq n$ as well as a set of indices $\{i_0, i_1, \dots, i_{k-1}, i_k, i_{k+1}, \dots, i_m\}$ such that

$$i_0 = 0, \ i_k = k_{\min}, \ \text{and} \ i_m = n,$$

$$t_{i_{\ell-1}} > t_{i_\ell} \ \text{and} \ t_i > t_{i_{\ell-1}} \ \forall i_{\ell-1} < i < i_\ell, \quad \text{for} \ \ell = 1, \dots, k, \ \text{and}$$

$$t_{i_{\ell-1}} < t_{i_\ell} \ \text{and} \ t_i > t_{i_\ell} \ \forall i_{\ell-1} < i < i_\ell, \quad \text{for} \ \ell = k+1, \dots m.$$

The *skeleton of P* is then given by $((v_{i_j}, t_{i_j}))_{j=0,\dots,m}$. 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 $(v_i, t_i) \in P \setminus \{(v_0, t_0), (v_n, t_n)\}$ is called a *local maximum* if $t_i > t_{i-1}$ and $t_i > t_{i+1}$. We successively remove all local maxima from P as follows. First, we take the local maximum in P with the greatest birth time, remove it from P, and connect its former neighbours by a direct edge. In the resulting path, we take the local maximum of greatest birth time and remove it, repeating until there is no local maximum left; see Figure 1. Therefore, the final path is decreasing in birth times of its vertices until the oldest vertex is reached, and only increasing in birth times afterwards. Hence, it is the uniquely determined skeleton of the path. Note that the skeleton is not necessarily an actual path of the graph. In fact, the skeleton of a shortcut-free path is not itself a path unless the path is its own skeleton.
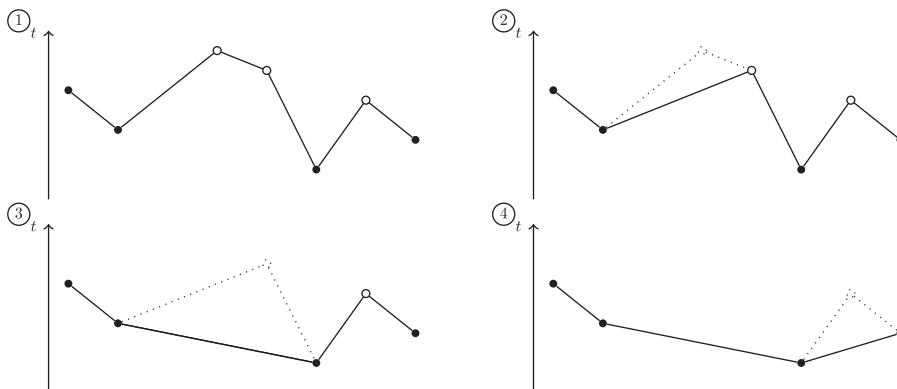
FIGURE 1: A path where a vertex's birth time is denoted on the *t*-axis. The vertices of the skeleton are in black. We successively remove all local maxima, starting with the youngest, and replace them by direct edges until the path containing only the skeleton vertices is left.

Graph surgery.

To bound the probability of existence of an infinite self-avoiding path in $\mathscr{G}_0^p$ starting at the origin, we increase the number of short edges in $\mathscr{G}_0^p$, 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

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

by the Potter bound [3, Theorem 1.5.6]. We define

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

Note that p enters the definition of $\tilde{\rho}$ at two places. Namely, it determines the range where edges are put deterministically and also scales the profile function. We now choose $\tilde{\rho}$ as a profile function together with the preferential attachment kernel (2) and construct $\mathcal{G}_{\tilde{\varphi}}(\mathcal{X}_0, \mathcal{U}_0)$ where

$$\tilde{\varphi}((x, t), (y, s)) = \tilde{\rho}\big(g^{\mathrm{pa}}(t, s)|x - y|^d\big).$$

In other words, we connect two given vertices (x, t) and (y, s) with probability

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

Note that in general $\tilde{\rho}$ does not satisfy the normalization condition (1). However, $\tilde{\rho}$ is still integrable, and therefore the resulting graph $\mathcal{G}_{\tilde{\varphi}}(\mathcal{X}_0, \mathcal{U}_0)$ is still locally finite with unchanged power law and shows the same qualitative behaviour. Since $p\rho \leq \tilde{\rho}$, it follows by a simple coupling argument that

$$\theta(p) \leq \mathbb{P}_0\{\mathbf{0} \leftrightarrow \infty \text{ in } \mathcal{G}_{\tilde{\varphi}}(\mathcal{X}_0, \mathcal{U}_0)\}.$$

By the above, there is no loss of generality in considering the unpercolated graph $\mathscr{G}$, resp. $\mathscr{G}_0$, where the profile function ρ is of the form

$$\rho(x) = 1 \wedge (pAx^{-\delta}), \tag{7}$$

which is what we do from now on. Note that we can no longer assume that (1) holds; instead we have

$$I_\rho := \int_{\mathbb{R}^d} \rho(|x|^d)\,\mathrm{d}x = (pA)^{1/\delta}\big(J(d)\tfrac{\delta}{d(\delta-1)}\big), \tag{8}$$

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 $P = ((x_1, t_1), (x_2, t_2), \dots)$ exists in $\mathscr{G}$. By choice of ρ, such a path satisfies

$$|x_i - x_j|^d > (pA)^{\frac{1}{\delta}} g^{\mathrm{pa}}(t_i, t_j)^{-1} \quad \text{for all } |i - j| \geq 2.$$

Strategy of the proof.

To build a long path, one needs to use old vertices. Every path is divided into a skeleton, which encodes how it moves to increasingly old vertices, and subpaths connecting consecutive points of the skeleton by any number of younger vertices, 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 $(KI_\rho)^k$ times the probability that the two extremal vertices are connected by an edge, for some constant $K > 1$. Combining this estimate with the van den Berg–Kesten (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 truncated first moment method with the truncation applied to the birth time of the oldest point on the skeleton. We therefore obtain that the probability of existence of a shortcut-free path of length n starting at $\mathbf{0}$ is bounded from above by $(KI_\rho)^n$, and hence

$$\theta(p) \leq \lim_{n\to\infty} (KI_\rho)^n = 0$$

for $p > 0$ small enough to ensure $I_\rho < 1/K$.

Connecting two old 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 age from oldest to youngest. We assume without loss of generality that $\mathbf{x}$ is younger 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 with fixed vertex set $\{\mathbf{y}_2, \dots, \mathbf{y}_k\}$ such that every child has birth time greater than its parent. Here, a binary tree is a rooted tree in which each vertex can have either no child, a left child, a right child, or both a left and a right child. With the path P we associate a tree in $\mathscr{T}_{k-1}$ as follows (see Figure 2):

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

Step two: Suppose the tree with vertices $\mathbf{y}_2, \dots, \mathbf{y}_{i-1}$ has been 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 at each vertex, branch to the left if the path P visits $\mathbf{y}_i$ before the vertex and to the right otherwise. If this
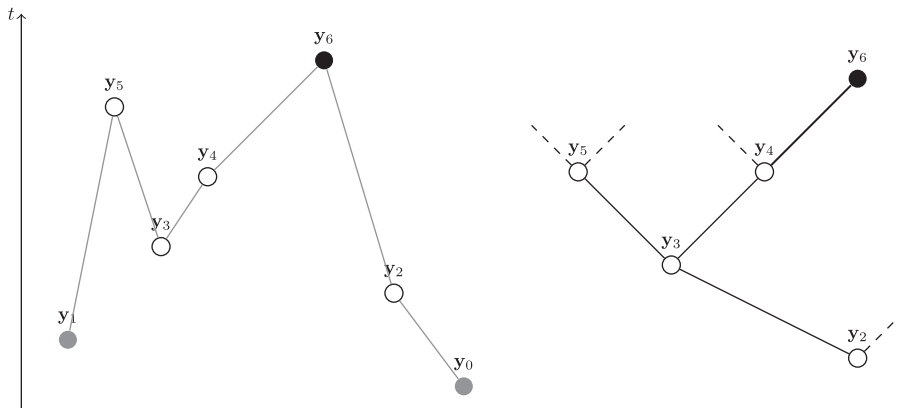
FIGURE 2: The left panel shows the path P, with the t-axis giving the vertices' birth times. The vertices $\mathbf{y}_1$ and $\mathbf{y}_0$, which will not appear in the tree, are in grey. We insert the vertex $\mathbf{y}_6$ at the end of the branch that goes left at $\mathbf{y}_2$, right at $\mathbf{y}_3$, and right at $\mathbf{y}_4$.
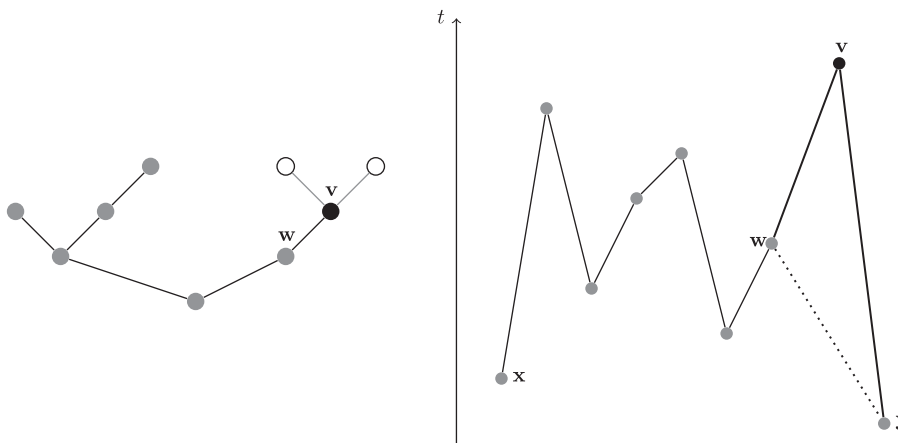

FIGURE 3: The left panel shows the binary tree T. The grey vertices have already been explored by a depth-first search. The black vertex $\mathbf{v}$ is the vertex currently being explored. The white vertices have not been discovered yet. The right panel shows the path P corresponding to the already explored tree. The t-axis gives the vertices' birth times. The start and end vertices, $\mathbf{x}$ 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.

means going to a place where there is no vertex, we attach $\mathbf{y}_i$ there. We continue in this way until all $\mathbf{y}_2, \ldots, \mathbf{y}_k$ have been attached.

Next, we explain how to construct a path P connecting $\mathbf{x}$ and $\mathbf{y}$ when $T \in \mathcal{T}_{k-1}$ is given; see Figure 3. 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 in which they occur in L). We proceed as follows:

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 vertex. 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.

It is clear that for given $\mathbf{y}_0, \ldots, \mathbf{y}_k$ the two procedures establish a bijection between the paths with vertices $\mathbf{y}_0, \ldots, \mathbf{y}_k$ 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 $\left\{ \mathbf{x} \overset{2}{\underset{\mathbf{x},\mathbf{y}}{\longleftrightarrow}} \mathbf{y} \right\}$ the event that $\mathbf{x}$ and $\mathbf{y}$ are connected by a path of length two where the connector is younger than both of them.

Lemma 2.2. *Let* $\gamma \in \left(0, \frac{\delta}{\delta+1} \right)$. *Let* $\mathbf{x} = (x, t)$ *and* $\mathbf{y} = (y, s)$ *be two given vertices satisfying* $|x - y|^d \geq (pA)^{1/\delta} g^{\mathrm{pa}}(t, s)^{-1}$. *Then*

$$\mathbb{P}^p_{\mathbf{x},\mathbf{y}} \left\{ \mathbf{x} \overset{2}{\underset{\mathbf{x},\mathbf{y}}{\longleftrightarrow}} \mathbf{y} \right\} \leq \int\limits_{\mathbb{R}^d \times ((t \vee s), 1]} d\mathbf{z} \, \mathbb{P}^p_{\mathbf{x},\mathbf{z}} \{ \mathbf{x} \sim \mathbf{z} \} \mathbb{P}^p_{\mathbf{y},\mathbf{z}} \{ \mathbf{z} \sim \mathbf{y} \} \leq I_\rho \, C_1 \mathbb{P}^p_{\mathbf{x},\mathbf{y}} \{ \mathbf{x} \sim \mathbf{y} \},$$

where

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

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

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

(see [9]), from which the first inequality follows. For the second inequality, we have

$$\int_t^1 du \int_{\mathbb{R}^d} dz \, \rho(\beta^{-1} t^\gamma u^{1-\gamma} |x - z|^d) \rho(\beta^{-1} s^\gamma u^{1-\gamma} |z - y|^d)$$

$$\leq \int_t^1 du \Big[\int_{\mathbb{R}^d} dz \, \rho(\beta^{-1} t^\gamma u^{1-\gamma} |x - z|^d) \rho \left((2^d \beta)^{-1} s^\gamma u^{1-\gamma} |x - y|^d \right)$$

$$+ \int_{\mathbb{R}^d} dz \, \rho \left((2^d \beta)^{-1} t^\gamma u^{1-\gamma} |x - y|^d \right) \rho(\beta^{-1} s^\gamma u^{1-\gamma} |z - y|^d) \Big].$$

Here, the inequality holds because, 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

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

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

$$pA2^{d\delta}\beta^{1+\delta}I_{\rho} \int_t^1 du\ s^{-\gamma\delta} t^{-\gamma}|x - y|^{-d\delta} u^{-\delta(1-\gamma)+\gamma-1}$$

$$\leq pA2^{d\delta}I_{\rho}\frac{\beta^{\delta+1}}{\delta(1 - \gamma) - \gamma}(s^{\gamma}t^{1-\gamma}|x - y|^d)^{-\delta}$$

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

$$\mathbb{P}^p_{\mathbf{x},\mathbf{y}}\{\mathbf{x} \sim \mathbf{y}\} = pA\left(\beta^{-1}s^{\gamma}t^{1-\gamma}|x - y|^d\right)^{-\delta},$$

which proves the claim. $\qquad\square$

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 oldest vertices within the path. We denote this event by $\{\mathbf{x} \overset{k}{\underset{\mathbf{x},\mathbf{y}}{\longleftrightarrow}} \mathbf{y}\}$.

Lemma 2.3. *Let $\gamma \in \left(0, \frac{\delta}{\delta+1}\right)$, and let $\mathbf{x} = (x, t)$, $\mathbf{y} = (y, s)$ be two Poisson points satisfying $|x - y|^d > (pA)^{1/\delta}g^{\mathrm{pa}}(t, s)^{-1}$. Then, for all $k \in \mathbb{N}$, we have*

$$\mathbb{P}^p_{\mathbf{x},\mathbf{y}}\left\{\mathbf{x} \overset{k}{\underset{\mathbf{x},\mathbf{y}}{\longleftrightarrow}} \mathbf{y}\right\} \leq (I_{\rho}C_2)^{k-1}\mathbb{P}^p_{\mathbf{x},\mathbf{y}}\{\mathbf{x} \sim \mathbf{y}\}, \tag{9}$$

where

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

Proof. For $k = 1$ there is nothing to show, so we assume $k \geq 2$. 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

$$\mathbb{P}^p_{\mathbf{x},\mathbf{y}}\{\mathbf{x} \overset{k}{\underset{\mathbf{x},\mathbf{y}}{\longleftrightarrow}} \mathbf{y}\} \leq \sum_{\substack{T \text{ binary tree} \\ \text{on } k - 1 \text{ vertices}}} \mathbb{E}^p_{\mathbf{x},\mathbf{y}}[X(T)],$$

and as the number of binary trees on $k - 1$ vertices is bounded from above by the Catalan numbers $(2(k - 1))!/(k - 1)!k!) \leq 4^{k-1}$, it suffices to show

$$\mathbb{E}^p_{\mathbf{x},\mathbf{y}}[X(T)] \leq (I_{\rho}C_1)^{k-1}\mathbb{P}^p_{\mathbf{x},\mathbf{y}}\{\mathbf{x} \sim \mathbf{y}\}$$

for all binary trees T with $k - 1$ vertices. 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 2.2. 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 **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 **y**.

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

(i) The new leaf is younger 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 as 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 $\mathbf{x}_\ell = (x_\ell, t_\ell)$ and $\mathbf{x}_r = (x_r, t_r)$, respectively, we get using (i) that

$$
\mathbb{E}^p_{\mathbf{x},\mathbf{y}}\big[X(T')\big]
$$
$$
= \mathbb{E}^p_{\mathbf{x},\mathbf{y}}\big[\sharp\{\text{paths } \mathbf{x} \leftrightarrow \mathbf{x}_\ell \sim \mathbf{x}_{\text{new}} \sim \mathbf{x}_r \leftrightarrow \mathbf{y} \text{ associated with } T'\}\big]
$$
$$
\leq \mathbb{E}^p_{\mathbf{x},\mathbf{y}}\left[\sum_{\substack{\text{almost complete}\\ \text{labellings of } T}} \int_{t_\ell \vee t_r}^1 du \int_{\mathbb{R}^d} dz \varphi((x_\ell, t_\ell), (z, u)) \varphi((z, u), (x_r, t_r)) \right].
$$

As the paths associated to T' are shortcut-free, we have $|x_\ell - x_r|^d > (pA)^{1/\delta} \beta g^{\text{pa}}(t_\ell, t_r)^{-1}$, and hence Lemma 2.2 ensures that this is bounded by

$$
I_\rho\, C_1\, \mathbb{E}^p_{\mathbf{x},\mathbf{y}}\left[\sum_{\substack{\text{almost complete}\\ \text{labelling of } T}} \mathbb{P}^p_{\mathbf{x}_\ell, \mathbf{x}_r}\{\mathbf{x}_\ell \sim \mathbf{x}_r\} \right]
$$
$$
\leq I_\rho\, C_1\, \mathbb{E}^p_{\mathbf{x},\mathbf{y}}\big[X(T)\big] \leq (I_\rho C_1)^k \mathbb{P}^p_{\mathbf{x},\mathbf{y}}\{\mathbf{x} \sim \mathbf{y}\},
$$

using (ii) and the induction hypothesis. $\qquad\square$

BK inequality.

We use a version of the famous van den Berg–Kesten (BK) inequality [2], where the application to our setting is described in detail in [11, pp. 10–13]. For a path with given skeleton,

the BK inequality allows us to focus on the individual subpaths between any two consecutive skeleton vertices instead of considering the whole path at once.

For given Poisson points $\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_m$, we write

$$\left\{ \mathbf{x}_0 \xleftrightarrow[\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_m]{k} \mathbf{x}_m \right\}$$

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

Conditioned on the event that the three distinct points $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ are vertices of $\mathscr{G}$, define E to be the event that $\mathbf{x}_1$ is connected by a path of length n_1 to $\mathbf{x}_2$ and $\mathbf{x}_2$ is connected by a path of length n_2 to $\mathbf{x}_3$, where the two paths share only $\mathbf{x}_2$ as a common vertex; we say that the two paths *occur disjointly*. We denote this disjoint occurrence by $\circ$ and write $E = \{\mathbf{x}_1 \xleftrightarrow{n_1} \mathbf{x}_2\} \circ \{\mathbf{x}_2 \xleftrightarrow{n_2} \mathbf{x}_3\}$. Furthermore, both events are *increasing* in the following sense. Given any realization of the Poisson point process such that there is such a path between, say, $\mathbf{x}_1$ and $\mathbf{x}_2$, there also exists such a path in any realization with additional vertices. Recall that $\mathbb{P}^p_{\mathbf{x}_1, \ldots, \mathbf{x}_n}$ denotes the law of $\mathscr{G}$ conditioned on $\mathbf{x}_1, \ldots, \mathbf{x}_n$ being vertices in $\mathcal{X}$. Then the BK inequality from [11, Theorem 2.1] yields

$$\mathbb{P}^p_{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3} \left(\{\mathbf{x}_1 \xleftrightarrow{n_1} \mathbf{x}_2\} \circ \{\mathbf{x}_2 \xleftrightarrow{n_2} \mathbf{x}_3\} \right) \le \mathbb{P}^p_{\mathbf{x}_1, \mathbf{x}_2}\{\mathbf{x}_1 \xleftrightarrow{n_1} \mathbf{x}_2\} \mathbb{P}^p_{\mathbf{x}_2, \mathbf{x}_3}\{\mathbf{x}_2 \xleftrightarrow{n_2} \mathbf{x}_3\}. \tag{10}$$

Next, let $S = (\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_m)$ be a given skeleton and recall that all paths we consider are self-avoiding. Then the event that the root $\mathbf{0} = \mathbf{x}_0$ starts a path of length n that has skeleton S can be written as

$$\left\{ \mathbf{x}_0 \xleftrightarrow[\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_m]{n} \mathbf{x}_m \right\} = \bigcup_{\substack{(n_1, \ldots, n_m) \in \mathbb{N}^m: \\ n_1 + \cdots + n_m = n}} \left\{ \mathbf{x}_0 \xleftrightarrow[\mathbf{x}_0, \mathbf{x}_1]{n_1} \mathbf{x}_1 \right\} \circ \cdots \circ \left\{ \mathbf{x}_{m-1} \xleftrightarrow[\mathbf{x}_{m-1}, \mathbf{x}_m]{n_m} \mathbf{x}_m \right\}.$$

Inductively, we derive as in (10) that

$$\mathbb{P}^p_{\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_m} \left\{ \mathbf{x}_0 \xleftrightarrow[\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_m]{n} \mathbf{x}_m \right\} \le \sum_{\substack{(n_1, \ldots, n_m) \in \mathbb{N}^m: \\ n_1 + \cdots + n_m = n}} \prod_{j=1}^m \mathbb{P}^p_{\mathbf{x}_{j-1}, \mathbf{x}_j} \left\{ \mathbf{x}_{j-1} \xleftrightarrow[\mathbf{x}_{j-1}, \mathbf{x}_j]{n_j} \mathbf{x}_j \right\}. \tag{11}$$

Proof of the subcritical phase.

We now use the results of the previous paragraphs to bound the probability that a shortcut-free path of length n exists by some exponential, thus proving Theorem 1.1(a). To this end, we have to distinguish between regular and irregular paths. Let $S = (\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_m)$ be a skeleton of length m. We say S is *regular* if its oldest vertex is born after time 2^{-m}. We say S is *irregular* if its oldest vertex is born before time 2^{-m}. 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 $n \ge k$ such that the path (of length n) $(\mathbf{v}_0, \ldots, \mathbf{v}_n)$ 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.

Proof of Theorem 1.1(a). Observe that if an irregular path of length n exists, then an irregular path of length $k \leq n$ whose end vertex is the oldest vertex of the path also exists. Let $A_{\mathrm{irreg}}(k)$ be the event that $\mathbf{0}$ starts an irregular path of length k where the end vertex is the oldest one. We will prove in the following lemma that $\mathbb{P}_0^p(A_{\mathrm{irreg}}(k)) \leq (C_3 I_\rho)^k$ for some constant C_3. We then choose p such that $I_\rho < C_3^{-1}$ and achieve

$$\sum_{k=1}^{\infty} \mathbb{P}_0^p(A_{\mathrm{irreg}}(k)) < \infty.$$

Hence, Borel–Cantelli yields that almost surely any long enough path is regular.

Lemma 2.4. *Let* $\gamma \in \left[0, \frac{\delta}{\delta+1}\right)$. *Then, for all* $k \in \mathbb{N}$,

$$\mathbb{P}_0^p(A_{\mathrm{irreg}}(k)) \leq (C_3 I_\rho)^k,$$

where

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

Proof. A path of length k whose oldest vertex is also the end vertex has a skeleton whose vertices' birth times are decreasing. Thus, we again write $\mathbf{0} = \mathbf{x}_0 = (x_0, t_0)$ and have by the Mecke equation as in the proof of Lemma 2.1 that

$$\mathbb{P}_0^p(A_{\mathrm{irreg}}(k)) \leq \sum_{m=1}^{k} \mathbf{E}\left[\sum_{\substack{(x_1,t_1),\ldots,(x_m,t_m)\in\mathcal{X} \\ t_0>t_1>\cdots>t_m \\ t_m<2^{-m}}} \mathbb{P}_{\mathcal{X}_0}^p\left\{ (x_0, t_0) \xleftarrow[\;\;(x_0,t_0),\ldots,(x_m,t_m)\;\;]{k} (x_m, t_m) \right\} \right]$$

$$= \sum_{m=1}^{k} \int_0^1 \mathrm{d}t_0 \int_{\substack{(\mathbb{R}^d\times(0,1])^m \\ t_0>t_1>\cdots>t_m \\ t_m<2^{-m}}} \bigotimes_{j=1}^{m} \mathrm{d}(x_j, t_j) \mathbb{P}_{\mathbf{x}_0,\ldots,\mathbf{x}_m}^p\left\{ (x_0, t_0) \xleftarrow[\;\;(x_0,t_0),\ldots,(x_m,t_m)\;\;]{k} (x_m, t_m) \right\},$$

where we have written $\mathbf{x}_j = (x_j, t_j)$ for $j = 1, \ldots, m$ as usual. Using the BK inequality (11) and Lemma 2.3, for the last probability we get

$$\mathbb{P}_{\mathbf{x}_0,\ldots,\mathbf{x}_m}^p\left\{ (x_0, t_0) \xleftarrow[\;\;(x_0,t_0),\ldots,(x_m,t_m)\;\;]{k} (x_m, t_m) \right\}$$

$$\leq \sum_{\substack{(n_1,\ldots,n_m)\in\mathbb{N}^m: \\ n_1+\cdots+n_m=n}} \prod_{j=1}^{m} \mathbb{P}_{\mathbf{x}_{j-1},\mathbf{x}_j}^p\left\{ (x_{j-1}, t_{j-1}) \xleftarrow[\;\;\mathbf{x}_{j-1},\mathbf{x}_j\;\;]{n_j} (x_j, t_j) \right\}$$

$$\leq \sum_{\substack{(n_1,\ldots,n_m)\in\mathbb{N}^m: \\ n_1+\cdots+n_m=n}} (C_2 I_\rho)^{k-m} \prod_{j=1}^{m} \mathbb{P}_{\mathbf{x}_{j-1},\mathbf{x}_j}^p\{ (x_{j-1}, t_{j-1}) \sim (x_j, t_j) \}$$

$$= \binom{k-1}{m-1} (C_2 I_\rho)^{k-m} \prod_{j=1}^{m} \mathbb{P}_{\mathbf{x}_{j-1},\mathbf{x}_j}^p\{ (x_{j-1}, t_{j-1}) \sim (x_j, t_j) \}.$$

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

$$\mathbb{P}_0^p(A_{\text{irreg}}(k)) \tag{12}$$

$$\leq \sum_{m=1}^{k} \binom{k-1}{m-1} (C_2 I_\rho)^{k-m}$$

$$\times \int_0^1 dt_0 \int_0^{t_0} dt_1 \int_{\mathbb{R}^d} dx_1 \ldots \int_0^{2^{-m} \wedge t_{m-1}} dt_m \int_{\mathbb{R}^d} dx_m \left(\prod_{i=1}^{m} \rho\left(\beta^{-1} t_{i-1}^{1-\gamma} t_i^{\gamma} |x_{i-1} - x_i|^d \right) \right)$$

$$\leq \sum_{m=1}^{k} \binom{k-1}{m-1} C_2^{k-m} I_\rho^k \beta^m \int_0^1 dt_0 \int_0^{t_0} dt_1 \ldots \int_0^{2^{-m} \wedge t_{m-1}} dt_m \, t_0^{\gamma-1} t_m^{-\gamma} \prod_{i=1}^{m-1} t_i^{-1}$$

$$\leq I_\rho^k \sum_{m=1}^{k} \binom{k-1}{m-1} \beta^m C_2^{k-m} (1-\gamma)^{-m} \leq (C_2 I_\rho)^k \sum_{m=1}^{k} \binom{k-1}{m-1} \leq (C_3 I_\rho)^k,$$

where the third inequality follows from Lemma A.5. $\qquad \square$

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

Lemma 2.5. *Let $\gamma \in \left[\frac{1}{2}, \frac{\delta}{\delta+1} \right)$. Then, for all $n \in \mathbb{N}$, we have*

$$\mathbb{P}_0^p(A_{\text{reg}}(n)) \leq K(C_3 I_\rho)^n,$$

where

$$C_3 = 2C_2 = \frac{\beta 2^{d\delta+4}}{\delta(1-\gamma) - \gamma}$$

and K is some constant.

Proof. Writing $\mathbf{0} = \mathbf{x}_0 = (x_0, t_0)$ and following the same arguments based on the Mecke equation, the BK inequality, and Lemma 2.3 as in the proof of Lemma 2.4 above, we get for large enough n that

$$\mathbb{P}_0^p(A_{\text{reg}}(n)) \leq \sum_{m=1}^{n} \sum_{k=0}^{m} \int_{2^{-m}}^{1} dt_0 \binom{n-1}{m-1} (C_2 I_\rho)^{n-m} \tag{13}$$

$$\times \int_{\substack{(x_1,t_1),\ldots,(x_m,t_m) \in \mathbb{R}^d \times (0,1] \\ t_0 > t_1 > \ldots > t_k > 2^{-m} \\ t_k < t_{k+1} < \cdots < t_m}} \bigotimes_{j=1}^{m} d(x_j, t_j) \prod_{j=1}^{m} \varphi((x_{j-1}, t_{j-1}), (x_j, t_j)).$$

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

$$\varphi((x_{j-1}, t_{j-1}), (x_j, t_j)) = \rho\left(g^{\text{pa}}(t_{j-1}, t_j) |x_{j-1} - x_j|^d \right).$$

Therefore, the right-hand side of (13) reads

$$
\sum_{m=1}^{n} \binom{n-1}{m-1} (C_2 I_\rho)^{n-m}
$$

$$
\times \sum_{k=0}^{m} I_\rho^m \int\limits_{\substack{1 > t_0 > t_1 > \ldots > t_k > 2^{-m} \\ t_k < t_{k+1} < \cdots < t_m}} \bigotimes_{j=0}^{m} dt_j \prod_{j=1}^{m} g^{\mathrm{pa}}(t_{j-1}, t-j)^{-1}. \tag{14}
$$

For $k = 0$ the integral from (14) can be written as

$$
\beta^m \int_{2^{-m}}^{1} dt_0 \int_{t_0}^{1} dt_1 \ldots \int_{t_{m-1}}^{1} dt_m t_0^{-\gamma} t_m^{\gamma-1} \prod_{j=1}^{m-1} t_j^{-1} \leq \left(\frac{\beta}{1-\gamma} \right)^m,
$$

by Lemma A.1. For $k = m$, we obtain for the integral from (14)

$$
\beta^m \int_{2^{-m}}^{1} dt_0 \int_{2^{-m}}^{t_0} dt_1 \ldots \int_{2^{-m}}^{t_{m-1}} dt_m t_0^{\gamma-1} t_m^{-\gamma} \prod_{j=1}^{m-1} t_j^{-1} \leq \left(\frac{\beta}{1-\gamma} \right)^m,
$$

by Lemma A.5. For $1 \leq k \leq m-1$, we infer for the integral from (14), using Lemma A.4,

$$
\beta^m \sum_{k=1}^{m-1} \int_{2^{-m}}^{1} dt_0 \int_{2^{-m}}^{t_0} dt_1 \ldots \int_{2^{-m}}^{t_{k-1}} dt_k \left[t_0^{\gamma-1} \left(\prod_{j=1}^{k-1} t_j^{-1} \right) t_k^{-\gamma} \right.
$$

$$
\times \int_{t_k}^{1} dt_{k+1} \ldots \int_{t_{m-1}}^{1} dt_m \left[t_k^{-\gamma} \left(\prod_{j=k+1}^{m-1} t_j^{-1} \right) t_m^{\gamma-1} \right] \Bigg]
$$

$$
\leq \beta^m \frac{2^{-m(1-2\gamma)}(m \log (2))^{m-2}}{\gamma^2 (2\gamma - 1)(m-2)!} \sum_{k=1}^{m-1} \binom{m-2}{k-1}.
$$

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 (13) and (14) that

$$
\mathbb{P}_0^p(A_{\mathrm{reg}}(n)) \leq I_\rho^n K \sum_{m=1}^{n} \binom{n-1}{m-1} \beta^m C_2^{n-m} \left((1-\gamma)^{-m} + (2^{2\gamma + \log_2 (e)} \log (2))^m \right),
$$

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

$$
\mathbb{P}_0^p(A_{\mathrm{reg}}(n)) \leq K(I_\rho C_3)^n.
$$

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

Setting p small enough that $C_3 I_\rho < 1$ concludes the proof of Theorem 1.1(a). $\qquad \square$

3. Absence of a subcritical phase

In this section, we prove Theorem 1.1(b) using a strategy of Jacob and Mörters [14]. Starting from a sufficiently old vertex, we use a young connector to connect the old vertex with a much older one; we repeat this indefinitely, moving to older and older 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 old vertices sum to a probability strictly less than one.

To this end, we show that an old vertex is *with extreme probability* connected to a much older 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 $g^{\mathrm{pa}}, g^{\mathrm{sum}} \leq g^{\mathrm{min}}$, we can fix the kernel g to be the min kernel g^{min} throughout this section. We retain the terminology of *old* and *young* vertices motivated by preferential attachment for convenience and better comparison with the previous section. Hence, for two given vertices $\mathbf{x} = (x, t)$ and $\mathbf{y} = (y, s)$, the connection probability is given by

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

Recall that ρ is regularly varying with index $-\delta$ for $\delta > 1$. Furthermore, $\gamma > \delta/(\delta + 1)$. 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 1.1(b).

Lemma 3.1. *Let* $\gamma > \tfrac{\delta}{\delta+1}$ *and* ρ *be regularly varying with index* $-\delta$ *for* $\delta > 1$. *Let* α_1, α_2 *be as defined as above. Let* $\mathbf{x}_0 = (x_0, s_0)$ *be a given Poisson point with* $s_0 < \tfrac{1}{2}$. *Then, for any retention parameter* $p > 0$, *wep($1/s_0$), there exists a sequence* $(\mathbf{x}_k)_{k \in \mathbb{N}}$ *of vertices* $\mathbf{x}_k = (x_k, s_k) \in \mathcal{X}$ *such that*

(i) $s_k < s_{k-1}^{\alpha_1}$ *and* $|x_k - x_{k-1}|^d < \tfrac{\beta}{2}s_{k-1}^{-\alpha_2}$, *and*

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

for all $k \in \mathbb{N}$.

Proof. It suffices to show that, *wep($1/s_0$)*, there exists a vertex $\mathbf{x}_1 = (x_1, s_1)$ satisfying (i) and (ii). The result then follows by induction. The number of vertices born before time $s_0^{\alpha_1}$ and within distance $((\beta/2)s_0^{-\alpha_2})^{1/d}$ from x_0 is Poisson distributed with parameter

$$\mathrm{Vol}\left(\{|x_1 - x_0|^d < \tfrac{\beta}{2}s_0^{-\alpha_2}\} \times (0, s_0^{\alpha_1})\right) = O\left(s_0^{\alpha_1 - \alpha_2}\right),$$

where $O(\cdot)$ again is the standard Landau symbol. Since $\alpha_2 > \alpha_1$, there exists such a vertex $\mathbf{x}_1$, *wep($1/s_0$)*. To connect $\mathbf{x}_0$ to $\mathbf{x}_1$ via a young vertex, we focus on connectors (y, t) born after

time $1/2$ and within distance $((\beta/2)s_0^{-\gamma})^{1/d}$ from $\mathbf{x}_0$. Since, for such choices of (y, t), we have

$$
|x_1 - y|^d \leq \left(\left(\frac{\beta s_0^{-\alpha_2}}{2} \right)^{1/d} + \left(\frac{\beta s_0^{-\gamma}}{2} \right)^{1/d} \right)^d \leq \beta s_0^{-\alpha_2},
$$

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

$$
p^2 \int_{1/2}^1 \mathrm{d}t \int\limits_{\{|y-x_0|^d \leq \frac{\beta}{2} s_0^{-\gamma}\}} \mathrm{d}y \, \rho\big(\beta^{-1} s_0^{\gamma} |y - x_0|^{-d}\big) \rho\big(s_0^{\alpha_1 \gamma - \alpha_2}\big)
$$

$$
= p^2 \frac{1}{2} \beta s_0^{-\gamma} \rho(s_0^{\alpha_1 \gamma - \alpha_2}) \int\limits_{\{|y-x_0|^d \leq 1/2\}} \mathrm{d}y \, \rho(|y - x_0|^d). \tag{15}
$$

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 [3, Theorem 1.5.6] that

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

for some $A < 1$ and s_0 small enough. Additionally, $\rho(|x|^d) \geq \rho(1/2) > 0$ for all $|x|^d < 1/2$. Hence, (15) is bounded from below by

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

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

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

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

4. Proof of Theorem 1.2

We first introduce for finite $t > 0$ the rescaling map

$$
\begin{aligned}
h_t \colon \mathbb{T}_1^d \times (0, t] &\longrightarrow \mathbb{T}_t^d \times (0, 1], \\
(x, s) &\longmapsto \big(t^{1/d} x, s/t\big).
\end{aligned}
$$

It gives rise to a new graph $h_t(\mathscr{G}_t^p)$ whose vertices live on $\mathbb{T}_t^d \times (0, 1]$ and where two rescaled vertices are connected in $h_t(\mathscr{G}_t^p)$ if they were originally connected in $\mathscr{G}_t^p$. It is easy to see that $h_t(\mathscr{G}_t^p)$ is the graph with vertex set given by a standard Poisson process on $\mathbb{T}_t^d \times (0, 1]$ and independent edges with the same connection probability as in (4); see [9]. The process $t \mapsto h_t(\mathscr{G}_t^p)$ converges almost surely to the graph $\mathscr{G}^p$, in the sense that if a randomly selected point in $h_t(\mathscr{G}_t^p)$ is shifted to the origin, the embedded graph in any ball around the origin converges in distribution as $t \to \infty$, to the same ball centred at the origin of $\mathscr{G}_0^p$; see [9, Theorem 3.1]. To obtain the weak local limit structure for the age-based preferential attachment network, let $h_t^0(\mathscr{G}_t^p)$ be the graph $h_t(\mathscr{G}_t^p)$ with a root vertex $\mathbf{0}$ added at the origin. If G is a locally finite graph equipped with a root $\mathbf{x} \in G$ and $\xi_t(\mathbf{x}, G)$ is a non-negative functional acting on rooted graphs that satisfy

(A) $\xi_t(\mathbf{0}, h_t^0(\mathscr{G}_t^p)) \to \xi_\infty(\mathbf{0}, \mathscr{G}_0^p)$ in probability as $t \to \infty$ and

(B) $\sup_{t>0} \mathbb{E}[(\xi_t(\mathbf{0}, h_t^0(\mathscr{G}_t^p)))^q] < \infty$ for some $q > 1$,

then from Theorem 7 of [13] we get

$$\lim_{t\to\infty} \frac{1}{t} \sum_{\mathbf{x}\in\mathscr{G}_t^p} \xi_t(\theta_\mathbf{x}(\mathbf{x}), \theta_\mathbf{x}(\mathscr{G}_t^p)) = \mathbb{E}[\xi_\infty(\mathbf{0}, \mathscr{G}_0^p)] \tag{16}$$

in probability, where $\theta_\mathbf{x}$ acts on points $\mathbf{y} = (y, s)$ as $\theta_\mathbf{x}(\mathbf{y}) = (y - x, s)$ and on graphs accordingly. This weak law of large numbers is an adaptation of a general weak law of large numbers for point processes due to Penrose and Yukich [19]. For the proof of non-robustness in Theorem 1.2, define $\xi^k(\mathbf{x}, G)$ as the indicator that the component of the root vertex $\mathbf{x}$ is of size at most k. By the weak law of large numbers,

$$\lim_{t\to\infty} \frac{1}{t} \sum_{\mathbf{x}\in\mathscr{G}_t^p} \xi^k(\theta_\mathbf{x}(\mathbf{x}), \theta_\mathbf{x}(\mathscr{G}_t^p)) = \mathbb{E}[\xi^k(\mathbf{0}, \mathscr{G}^p)].$$

The left-hand side is asymptotically the proportion of vertices that are in components no bigger than k. As $k \to \infty$, the right-hand side converges to $1 - \theta(p)$ and if we choose a $p > 0$ such that $\theta(p) = 0$, there is no giant component in $(\mathscr{G}_t^p)_{t>0}$.

For the proof of robustness in Theorem 1.2, define $\xi_t(\mathbf{x}, G)$ as the indicator that the root $\mathbf{x}$ of G belongs to the connected component of the oldest vertex in the finite graph G, and $\xi_\infty(\mathbf{0}, G)$ as the indicator that the root $\mathbf{0}$ of G belongs to an infinite component in the infinite graph G. Then one has to show that

$$\xi_t(\mathbf{0}, h_t^0(\mathscr{G}_t^p)) \to \xi_\infty(\mathbf{0}, \mathscr{G}_0^p) \quad \text{in probability as } t \to \infty.$$

This is done in detail in [14] for the spatial preferential attachment model and can easily be adapted to the simpler age-based preferential attachment model. The weak law of large numbers then yields

$$\lim_{t\to\infty} \frac{1}{t} \sum_{\mathbf{x}\in\mathscr{G}_t^p} \xi_t(\theta_\mathbf{x}(\mathbf{x}), \theta_\mathbf{x}(\mathscr{G}_t^p)) = \mathbb{E}[\xi_\infty(\mathbf{0}, \mathscr{G}^p)] = \theta(p)$$

in probability. Again, we see from this that if $\theta(p) > 0$ there is a giant component, and the result follows from Theorem 1.1; further details are exactly as in [14].

A. Integration results

Lemma A.1. *Let* $\gamma \in (0, 1)$ *and* $t_0 \in (0, 1)$. *Then the following hold:*

(a) *For all* $k \in \mathbb{N}$, *we have*

$$\int_{t_0}^1 dt_1 \int_{t_1}^1 dt_2 \ldots \int_{t_{k-1}}^1 dt_k \left[t_0^{-\gamma} \left(\prod_{j=1}^{k-1} t_j^{-1} \right) t_k^{\gamma-1} \right] \le \frac{t_0^{-\gamma} \log^{k-1}(1/t_0)}{\gamma(k-1)!}.$$

(b) *For all* $k \in \mathbb{N}$, *we have*

$$\int_0^1 dt \frac{t^{-\gamma} \log^k(1/t)}{k!} = \left(\frac{1}{1-\gamma} \right)^{k+1}.$$

Proof. We prove (a) by induction. For $k = 1$, we have

$$t_0^{-\gamma} \int_{t_0}^1 dt_1 t_1^{\gamma-1} \leq \frac{t_0^{-\gamma}}{\gamma}.$$

For $k + 1$, using the induction hypothesis we get

$$\int_{t_0}^1 dt_1 \int_{t_1}^1 dt_2 \ldots \int_{t_k}^1 dt_{k+1} \left[t_0^{-\gamma} \left(\prod_{j=1}^k t_j^{-1} \right) t_{k+1}^{\gamma-1} \right] \leq t_0^{-\gamma} \int_{t_0}^1 dt_1 \frac{t_1^{-1} \log^{k-1}(1/t_1)}{\gamma(k-1)!}$$

$$= \frac{t_0^{-\gamma}(-1)^{k-1}}{\gamma(k-1)!} \int_{t_0}^1 dt_1 \log(t_1)' \log^{k-1}(t_1) = \frac{t_0^{-\gamma} \log^k(1/t_0)}{\gamma k!}.$$

We prove (b) by induction as well. As $\gamma < 1$, for $k = 1$, using integration by parts, we get

$$\int_0^1 dt \frac{t^{-\gamma} \log(1/t)}{1!} = \int_0^1 dt \frac{t^{-\gamma}}{1-\gamma} = \frac{1}{(1-\gamma)^2}.$$

Analogously, for $k + 1$,

$$\int_0^1 dt \frac{t^{-\gamma} \log^{k+1}(1/t)}{(k+1)!} = \int_0^1 dt \frac{t^{-\gamma} \log^k(1/t)}{(1-\gamma)k!} = \frac{1}{(1-\gamma)^{k+2}}$$

by the induction hypothesis. □

Lemma A.2. *Let $\gamma \in (1/2, 1)$ and $x \in (0, 1)$. Then, for all $k \in \mathbb{N}$, it holds that*

$$\int_x^1 dt \, \frac{t^{-2\gamma} \log^k(1/t)}{k!} \leq \frac{x^{1-2\gamma} \log^k(1/x)}{(2\gamma-1)k!}.$$

Proof. Integration by parts yields

$$\int_x^1 dt \, \frac{t^{-2\gamma} \log^k(1/t)}{k!} = \frac{x^{1-2\gamma} \log^k(1/x)}{(2\gamma-1)k!} - \int_x^1 dt \frac{t^{-2\gamma} \log^{k-1}(1/t)}{(2\gamma-1)(k-1)!} \leq \frac{x^{1-2\gamma} \log^k(1/x)}{(2\gamma-1)k!},$$

as the second integral is bounded from below by 0. □

Lemma A.3. *Let $\gamma \in (0, 1)$, $x \in (0, 1)$, and $t_0 \in (x, 1)$. Then, for all $k \in \mathbb{N}$, it holds that*

$$\int_x^{t_0} dt_1 \int_x^{t_1} dt_2 \ldots \int_x^{t_{k-1}} dt_k \left(t_0^{\gamma-1} \prod_{j=1}^k t_j^{-1} \right) = \frac{t_0^{\gamma-1} \log^k(t_0/x)}{k!}.$$

Proof. For $k = 1$, we get

$$\int_x^{t_0} dt_1 \, t_0^{\gamma-1} t_1^{-1} = t_0^{\gamma-1} \log(t_0/x).$$

For $k + 1$, using the induction hypothesis, we get

$$\int_x^{t_0} dt_1 \int_x^{t_1} dt_2 \ldots \int_x^{t_k} dt_{k+1} \left(t_0^{\gamma-1} \prod_{j=1}^{k+1} t_j^{-1} \right) = t_0^{\gamma-1} \int_x^{t_0} dt_1 \frac{t_1^{-1} \log^k(t_1/x)}{k!}$$

$$= t_0^{\gamma-1} \int_0^{\log(t_0/x)} dy \frac{y^k}{k!} = \frac{t_0^{\gamma-1} \log^{k+1}(t_0/x)}{(k+1)!}. \qquad \square$$

Lemma A.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*

$$
\int_x^1 dt_0 \int_x^{t_0} dt_1 \ldots \int_x^{t_{k-1}} dt_k \left[t_0^{\gamma-1} \left(\prod_{j=1}^{k-1} t_j^{-1} \right) t_k^{-\gamma} \int_{t_k}^1 dt_{k+1} \ldots \int_{t_{m-1}}^1 dt_m \left[t_k^{-\gamma} \left(\prod_{j=k+1}^{m-1} t_j^{-1} \right) t_m^{\gamma-1} \right] \right]
$$

$$
\leq \binom{m-2}{k-1} \frac{x^{1-2\gamma} \log^{m-2}(1/x)}{\gamma^2 (2\gamma - 1)(m-2)!}. \tag{17}
$$

Proof. We apply the previous lemmas. By Lemma A.1, we get

$$
\int_{t_k}^1 dt_{k+1} \ldots \int_{t_{m-1}}^1 dt_m \left[t_k^{-\gamma} \left(\prod_{j=k+1}^{m-1} t_j^{-1} \right) t_m^{\gamma-1} \right] \leq \frac{t_k^{-\gamma} \log^{m-k-1}(1/t_k)}{\gamma(m-k-1)!}.
$$

Therefore, the integral in (17) can be bounded by

$$
\int_x^1 dt_0 \int_x^{t_0} dt_1 \ldots \int_x^{t_{k-2}} dt_{k-1} \left[t_0^{\gamma-1} \left(\prod_{j=1}^{k-1} t_j^{-1} \right) \int_x^{t_{k-1}} dt_k \frac{t_k^{-2\gamma} \log^{m-k-1}(1/t_k)}{\gamma(m-k-1)!} \right].
$$

By Lemma A.2,

$$
\int_x^{t_{k-1}} dt_k \frac{t_k^{-2\gamma} \log^{m-k-1}(1/t_k)}{\gamma(m-k-1)!} \leq \frac{x^{1-2\gamma} \log^{m-k-1}(1/x)}{\gamma(2\gamma - 1)(m-k-1)!},
$$

and by Lemma A.3,

$$
\int_x^{t_0} dt_1 \ldots \int_x^{t_{k-2}} dt_{k-1} \, t_0^{\gamma-1} \left(\prod_{j=1}^{k-1} t_j^{-1} \right) = \frac{t_0^{\gamma-1} \log^{k-1}(t_0/x)}{(k-1)!}.
$$

Therefore, the integral in (17) can be further bounded by

$$
\int_x^1 dt_0 \frac{t_0^{\gamma-1} \log^{k-1}(t_0/x)}{(k-1)!} \frac{x^{1-2\gamma} \log^{m-k-1}(1/x)}{\gamma(2\gamma - 1)(m-k-1)!}
$$

$$
\leq \binom{m-2}{k-1} \frac{x^{1-2\gamma} \log^{m-2}(1/x)}{\gamma(2\gamma - 1)(m-2)!} \int_x^1 dt_0 \, t_0^{\gamma-1}.
$$

The result follows by integrating with respect to t_0. □

Lemma A.5. *Let $\gamma \in (0, 1)$ and $k \in \mathbb{N}$. Then*

$$
\int_0^1 dt_0 \int_0^{t_0} dt_1 \ldots \int_0^{t_{k-1}} dt_k \, t_0^{\gamma-1} \left(\prod_{j=1}^{k-1} t_j^{-1} \right) t_k^{-\gamma} \leq \left(\frac{1}{1-\gamma} \right)^k.
$$

Proof. We have

$$
\int_0^1 dt_0 \int_0^{t_0} dt_1 \ldots \int_0^{t_{k-2}} dt_{k-1} \left[t_0^{\gamma-1} \left(\prod_{j=1}^{k-1} t_j^{-1} \right) \int_0^{t_{k-1}} dt_k \, t_k^{-\gamma} \right]
$$

$$
= \frac{1}{1-\gamma} \int_0^1 dt_0 \int_0^{t_0} dt_1 \ldots \int_0^{t_{k-2}} dt_{k-1} \, t_0^{\gamma-1} \left(\prod_{j=1}^{k-2} t_j^{-1} \right) t_{k-1}^{-\gamma},
$$

and the result follows by repeating this across all integrals. □

Acknowledgements

References

[1] BENJAMINI, I. AND SCHRAMM, O. (2001). Recurrence of distributional limits of finite planar graphs. *Electron. J. Prob.* **6**, paper no. 23, 13 pp.

[2] VAN DEN BERG, J. (1996). A note on disjoint-occurrence inequalities for marked Poisson point processes. *J. Appl. Prob.* **33**, 420–426.

[3] BINGHAM, N. H., GOLDIE, C. M. AND TEUGELS, J. L. (1987). Regular variation. *Encyclopaedia Math. Appl.* Cambridge University Press.

[4] DEIJFEN, M., VAN DER HOFSTAD, R. AND HOOGHIEMSTRA, G. (2013). Scale-free percolation. *Ann. Inst. H. Poincaré Prob. Statist.* **49**, 817–838.

[5] DEPREZ, P., SUBHRA HAZRA, R. AND WÜTHRICH, M. V. (2015). Inhomogeneous long-range percolation for real-life network modeling. *Risks* **3**, 1–23.

[6] DEPREZ, P. AND WÜTHRICH, M. V. (2019). Scale-free percolation in continuum space. *Commun. Math. Statist.* 7, 269–308.

[7] GOUÉRÉ, J.-B. (2008). Subcritical regimes in the Poisson Boolean model of continuum percolation. *Ann. Prob.* **36**, 1209–1220.

[8] GOUÉRÉ, J.-B. (2009). Subcritical regimes in some models of continuum percolation. *Ann. Appl. Prob.* **19**, 1292–1318.

[9] GRACAR, P., GRAUER, A., LÜCHTRATH, L. AND MÖRTERS, P. (2019). The age-dependent random connection model. *Queueing Systems* **93**, 309–331.

[10] GRACAR, P., HEYDENREICH, M., MÖNCH, C. AND MÖRTERS, P. (2019). *Recurrence versus transience for weight-dependent random connection models*. Preprint. Available at https://arxiv.org/abs/1911.04350.

[11] HEYDENREICH, M., VAN DER HOFSTAD, R., LAST, G. AND MATZKE, K. (2019). *Lace expansion and mean-field behavior for the random connection model*. Preprint. Available at https://arxiv.org/abs/1908.11356.

[12] HEYDENREICH, M., HULSHOF, T. AND JORRITSMA, J. (2017). Structures in supercritical scalefree percolation. *Ann. Appl. Prob.* **27**, 2569–2604.

[13] JACOB, E. AND MÖRTERS, P. (2015). Spatial preferential attachment networks: power laws and clustering coefficients. *Ann. Appl. Prob.* **25**, 632–662.

[14] JACOB, E. AND MÖRTERS, P. (2017). Robustness of scale-free spatial networks. *Ann. Prob.* **45**, 1680–1722.

[15] LAST, G. AND PENROSE, M. (2017). *Lectures on the Poisson Process*. Cambridge University Press.

[16] MEESTER, R. AND ROY, R. (1996). *Continuum Percolation*. Cambridge University Press.

[17] NEWMAN, C. M. AND SCHULMAN, L. S. (1986). One-dimensional $1/|j - i|^s$ percolation models: the existence of a transition for $s \leq 2$. *Commun. Math. Phys.* **104**, 547–571.

[18] PENROSE, M. D. (1991). On a continuum percolation model. *Adv. Appl. Prob.* **23**, 536–556.

[19] PENROSE, M. D. AND YUKICH, J. E. (2003). Weak laws of large numbers in geometric probability. *Ann. Appl. Prob.* **13**, 277–303.