Scale-free geometric random graphs

Peter Mörters

University of Cologne

Contents

Chapter 1. Scale-free random graphs 3
1. A simple preferential attachment model 3
2. Weak local limits: Random trees 4
3. Inhomogeneous random graphs: Other kernels 5

Chapter 2. Spatial preferential attachment 6
1. The simple spatial preferential attachment model 6
2. Weak local limit: The age-dependent random-connection model 7
3. Clustering in the spatial preferential attachment model 8

Chapter 3. Scale-free geometric random graphs 9
1. Weight-dependent random connection model 9
2. Some interesting kernels 10
3. Exercise 1: The weak kernel: Non-spatial and spatial 11

Chapter 4. Phase transitions in scale-free geometric random graphs 13
1. Existence and non-existence of a subcritical phase 13
2. Existence and non-existence of a supercritical phase 15
3. Exercise 2: Calculate δ_{eff} for the most important kernels. 16

Chapter 5. Shortest paths 17
1. Ultrasmallness with or without connectors 17
2. Non-ultrasmallness and a limit theorem 18
3. Refining the truncated moment method 18

Bibliography 21

CHAPTER 1

Scale-free random graphs

In this course we are concerned with sequences $(\mathscr{G}_N)$ of finite random graphs (also called a *network*) such that the number of vertices goes to infinity as $N \to \infty$ but the empirical degree distribution

$$\frac{1}{|\mathscr{G}_N|} \sum_{v \in \mathscr{G}_N} \delta_{\text{degree}(v)}$$

converges to a limiting probability distribution μ. Such a network is called *scale-free* if

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

for some $\tau > 2$, which we call the *power-law exponent*. The first lecture will look at abstract graphs, but for most of the course the vertices are points in geometric space and the structure of the graph depends on its embedding into space, whence the graphs are called *geometric*.

1. A simple preferential attachment model

There is an abundance of models for scale-free networks, but a particularly interesting concept which is naturally leading to scale-free networks is *preferential attachment*. The idea, popularised 20 years ago by Barabasi and Albert, is that a graph is built by adding new vertices, which connect themselves at random but preferably to powerful vertices. In the classical models vertices arriving early are the most powerful. I will now discuss the probably simplest incarnation of preferential attachment.*

The model. Vertices arrive one-by-one and vertex n attaches to each vertex $m \in \{1, \dots, n-1\}$ independently with a probability proportional to $m^{-\gamma}$ for some parameter $0 < \gamma < 1$ parametrising the strength of the preference of early vertices. The proportionality factor is chosen so that the expected number of connections of a vertex is asymptotically constant. As

$$\sum_{m=1}^{n-1} m^{-\gamma} \sim c n^{1-\gamma}$$

the proportionality factor has to be of order $n^{\gamma-1}$. Altogether, the connection probabilities of two distinct vertices with number (or rank) i and j is

$$p_{ij} = \beta (i \vee j)^{\gamma-1} (i \wedge j)^{-\gamma}$$

and all connections are independent.

Degree distribution. To check that this network is indeed scale-free we let X_{ij} be the indicator that there is an edge between vertices ranked i and j. We look at a vertex with rank $k = aN$. Its expected degree is then

$$\sum_{m=1}^{N} p_{m,aN} = \sum_{m=1}^{aN} \beta (aN)^{\gamma-1} m^{-\gamma} + \sum_{m=aN}^{N} \beta m^{\gamma-1} (aN)^{-\gamma} \sim d(a) := \beta\big(\tfrac{1}{1-\gamma} + \tfrac{1}{\gamma}(a^{-\gamma} - 1)\big).$$

*This is a significant simplification of the model studied in [**10**], and as such not published.

By the Chernoff inequality

$$\mathbb{P}\Big(\Big|\sum_{m=1}^{N} X_{m,aN} - d(a)\Big| > \delta d(a)\Big) \le 2e^{-\delta^2 d(a)/3}.$$

As for different k these events are all independent the law of large numbers gives that the symmetric difference of the set of vertices with degree $\ge d(x)$ and the set of vertices with rank $\le xN$ contains a proportion $\mathrm{o}(x)$ of vertices, as x gets small. As $d^{-1}(x) = x^{-1/\gamma+o(1)}$ our preferential attachment model is therefore scale-free with power-law exponent

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

2. Weak local limits: Random trees

The main tool in analysing scale-free random graphs as above is local approximation by a (nice) random tree. Picking a vertex uniformly at random from $\mathscr{G}_N$ we look at a graph neighbourhood of depth a_N of this vertex, where a_N is slowly increasing. The distributional limit of these rooted random graphs, if it exists, is called the weak local limit. In our case it is the genealogical tree of a killed branching random walk, as we now explain.

To keep a record of the rank of the vertices we see when exploring a neighbourhood of a vertex define

$$t_n = \sum_{k=1}^{n-1} \frac{1}{k} \sim \log n$$

and the mapping

$$\phi_N \colon \{1, \dots, N\} \to (-\infty, 0], \quad n \mapsto t_n - t_N.$$

A uniformly chosen vertex is then mapped into a position which for large N converges to the law of $\log U$ for U uniform on $(0,1)$, which equals $-X$ for X standard exponentially distributed.

We now fix a point $uN \in \mathscr{G}_N$ and look at the point process consisting of all points $\phi_N(j)$ with $j \sim uN$. The claim is that this process converges to $\log u + \pi$ restricted to $(-\infty, 0]$, where π is a Poisson point process with intensity measure

$$\nu(dz) = \beta(e^{z(1-\gamma)} 1_{z<0} + e^{z\gamma} 1_{z>0})\, dz.$$

Proof. Take $a < b < u$, then

$$\mathbb{P}\big(uN \text{ does not connect to any } j \in (aN, bN)\big) = \prod_{i=aN}^{bN} \big(1 - \beta(uN)^{\gamma-1} i^{-\gamma}\big)$$

$$\sim \exp\Big(-\beta(uN)^{\gamma-1} \sum_{i=aN}^{bN} i^{-\gamma}\Big) \to \exp\Big(-\beta u^{\gamma-1} \int_a^b x^{-\gamma}\, dx\Big) = \exp\Big(-\beta \int_{\log a - \log u}^{\log b - \log u} e^{z(1-\gamma)}\, dz\Big).$$

Hence the probability that there are no points $\phi_N(j) \in [\log a, \log b]$ with $j \sim un$ converges to the probability that $\pi([\log a - \log u, \log b - \log u]) = 0$ and, by the same calculation, the expected number of points $\phi_N(j) \in [\log a, \log b]$ with $j \sim un$ converges to the expectation of $\pi([\log a - \log u, \log b - \log u])$ as well. By Kallenberg's theorem, see e.g. [**23**, Proposition 3.22], this implies convergence to the Poisson process $\log u + \pi$ on $(-\infty, \log u]$ and a similar calculation gives convergence on $(\log u, 0]$. □

Suppose we have explored the offspring of uN and move to the next generation. If vN is an offspring vertex, then we can do the same calculation omitting the vertices that we have already seen in the exploration so far. If these are $\mathrm{o}(N)$ many we also get offspring distributed like an independent point

process with distribution $\log v + \pi$ restricted to $(-\infty, 0]$. As a result we can couple the exploration process up to some generation a_N to a killed branching random walk on $(-\infty, 0]$ started with one particle in position $-X$, where particles in position v have Poisson$(\nu(-\infty, -v))$ many offspring in position shifted by a step distribution distributed according to $\nu|_{(-\infty,-v)}/\nu(-\infty, -v)$.

One of the principal questions of interest in the theory of (sparse) random graphs $(\mathscr{G}_N)$ is whether there *exists a giant component*, i.e. whether the proportion of vertices in the largest component of $\mathscr{G}_N$ converges to a strictly positive limit. We can use our coupling to show that this limit agrees with the survival probability of our killed branching random walk. By general branching process theory [**4**] a killed branching random walk dies out almost surely iff there exists α such that

$$A_\alpha := \int_{-\infty}^{\infty} e^{-\alpha t} \nu(dt) < 1.$$

This can be calculated as

$$A_\alpha = \beta\Big(\int_{-\infty}^{0} e^{-\alpha z + (1-\gamma)z}\, dz + \int_{0}^{\infty} e^{-\alpha z + \gamma z}\, dz \Big),$$

which is finite only if $\gamma < \alpha < 1 - \gamma$. Such a choice is only possible if $\gamma < \frac{1}{2}$ and in this case

$$= \frac{\beta}{1 - \alpha - \gamma} + \frac{\beta}{\alpha - \gamma}.$$

This expression is minimal if $\alpha = \frac{1}{2}$ and then equals $\frac{4\beta}{1-2\gamma}$. Hence no giant component exists iff

$$\gamma < \frac{1}{2} \text{ and } \beta \le \frac{1}{4} - \frac{\gamma}{2}.$$

In particular if $\gamma \ge \frac{1}{2}$ then a giant component exists even if the edge density (i.e. the total number of edges per vertex, which is proportional to β in our model) is arbitrarily small. A feature common to most models that are locally tree-like is that there is a giant component irrespective of the particle density if $\tau < 3$ but not if $\tau > 3$. This turns out different in the spatial models we consider below.

3. Inhomogeneous random graphs: Other kernels

There are other ways of defining scale-free networks, and our focus is on the class of inhomogeneous random graphs. These graphs are defined by a symmetric kernel $p\colon (0,1] \times (0,1] \to [0, \infty)$ which is non-decreasing in every coordinate. The vertex set of $\mathscr{G}_N$ is again the set $\{1, \dots, N\}$ and distinct vertices i, j are connected by an edge independently with probability

$$p_{ij} = \frac{1}{N}\, p\Big(\frac{i}{N}, \frac{j}{N}\Big).$$

The following kernels all give scale-free networks with power-law exponent $\tau = 1 + \frac{1}{\gamma}$:

- Preferential attachment kernel: $p(x, y) = \beta(x \vee y)^{\gamma-1}(x \wedge y)^{-\gamma}$,
- Product kernel: $p(x, y) = \beta x^{-\gamma} y^{-\gamma}$,
- Strong kernel: $p(x, y) = \beta(x \wedge y)^{-\gamma}$,
- Weak kernel: $p(x, y) = \beta(x \vee y)^{-\gamma-1}$.

Most other models are variants of inhomogeneous random graphs with suitable kernels (and often with minor dependencies that do not influence the qualitative behaviour of the model). For example, the configuration model is intimately related to the product kernel. Note that our model (corresponding to the preferential attachment kernel) is the only one which is dynamical in the sense that the connection probabilities do not depend on N.

CHAPTER 2

Spatial preferential attachment

The fact that the scale-free networks we looked at so far are locally tree-like was very useful for the analysis. But it is not desirable for modelling purpose, as most networks have clustering, i.e. the probability that two vertices in the neighbourhood of a given vertex are connected is much larger than the probability that two arbitrary vertices are connected. Spatial models address this problem: Points are now embedded in space and the probability of edges connecting two points depends on their geometric distance, in that far away points are much less likely to be connected. This creates the desired clustering effect in a natural way, but makes the model harder to analyse.

Geometric random graphs like this have been analysed for many years, but the intensive study of graphs that are *both scale-free and geometric* is relatively recent. We will start by having a closer look at a simple version of a geometric preferential attachment model, the material is taken from [**14**].

1. The simple spatial preferential attachment model

As our vertices are placed in continuous space it is natural to also let time be continuous, so that we are dealing with a growing sequence of graphs $(\mathscr{G}_t)_{t>0}$. The vertices of the graphs are embedded in the d-dimensional torus $\mathbb{T}_1 = (-1/2, 1/2]^d$ of side-length one, endowed with the torus metric d.

At time $t = 0$ the graph $\mathscr{G}_0$ has no vertices or edges. Then

- Vertices arrive according to a standard Poisson process in time and are placed independently uniformly on the d-dimensional torus $\mathbb{T}_1$.
- Given the graph $\mathscr{G}_{t-}$ a vertex $\mathbf{x} = (x, t)$ born at time t and placed in position x is connected by an edge to each existing node $\mathbf{y} = (y, s)$ independently with probability
$$\rho\big(\beta^{-1} s^{\gamma} t^{1-\gamma}\, d(x,y)^d\big),$$

where $\rho\colon [0,\infty) \to [0,1]$ is a nonincreasing *profile function* standardised to satisfy

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

With these choice of parametrisation the degree distribution of $(\mathscr{G}_t)$ does not depend on the choice of ρ and is scale-free with power-law exponent $\tau = 1 + \frac{1}{\gamma}$ as before. Indeed, the expected degree of a vertex $\mathbf{x} = (x, s)$ at time t is

$$\int_0^s dr \int_{\mathbb{T}} dy\, \rho\big(\beta^{-1} r^{\gamma} s^{1-\gamma}\, d(x,y)^d\big) + \int_s^t dr \int_{\mathbb{T}} dy\, \rho\big(\beta^{-1} s^{\gamma} r^{1-\gamma}\, d(x,y)^d\big),$$

and when $t \to \infty$ and $\frac{s}{t} \to a$ we obtain

$$\begin{aligned} &= s \int_0^1 dr \int_{\mathbb{T}} dy\, \rho\big(\beta^{-1} r^{\gamma} s\, d(0,y)^d\big) + s \int_1^{t/s} dr \int_{\mathbb{T}} dy\, \rho\big(\beta^{-1} s r^{1-\gamma}\, d(0,y)^d\big) \\ \longrightarrow &\int_0^1 dr \beta r^{-\gamma} \int_{\mathbb{R}^d} \rho(|y|^d)\, dy + \int_1^{1/a} dr \beta r^{\gamma-1} \int_{\mathbb{R}^d} \rho(|y|^d)\, dy = \frac{\beta}{1-\gamma} + \frac{\beta}{\gamma}(a^{-\gamma} - 1). \end{aligned}$$

We can therefore use the profile function ρ to tune the influence of the geometry independently from the degree distribution, the lighter its tails the stronger the geometric restrictions. Most variants of spatial preferential attachment in the literature are using $\rho(x) = \frac{1}{2a}\mathbb{1}_{[0,a]}(x)$ for $a \geq 1/2$. In this case, a vertex born at time s is linked to a new vertex at time t with probability $1/(2a)$ if their positions are within distance $(\frac{1}{t}\beta a\,(t/s)^\gamma)^{1/d}$. However, for our ultimate purpose this choice is too restrictive. We normally assume that ρ is regularly varying at infinity with index $-\delta$, for some parameter $\delta > 1$ as in this case we will see interesting phase transitions.

2. Weak local limit: The age-dependent random-connection model

We will find the weak local limit of our spatial preferential attachment network using a very elegant trick. Despite the fact that this limit is no longer a tree, it is still a useful and fascinating object, the *age-dependent random connection model*.

Let $\mathcal{X}$ denote a Poisson point process of unit intensity on $\mathbb{R}^d \times (0,\infty)$ and $\mathbb{T}_a \subset \mathbb{R}^d$ the centred torus of volume a endowed with the torus metric. We say a point $\mathbf{x} = (x,s) \in \mathcal{X}$ is born at time s and placed at position x. For $t > 0$ write $\mathcal{X}_t$ for $\mathcal{X} \cap (\mathbb{T}_1 \times (0,t])$, and note that this coincides with the set of vertices on the torus born by time t. Given $\mathcal{X}$ we introduce a family $\mathcal{V} = (\mathcal{V}_{\mathbf{x},\mathbf{y}})$ of independent random variables, uniformly distributed on $(0,1)$, indexed by the set of potential edges and we denote by $\mathcal{V}_t$ its restriction to indices in $\mathcal{X}_t \times \mathcal{X}_t$. We can define $\mathscr{G}(\mathcal{X}_t, \mathcal{V}_t)$ with vertex set $\mathcal{X}_t$ placing an edge between $\mathbf{x} = (x,u)$ and $\mathbf{y} = (y,s)$ with $s < u$, if and only if

$$\mathcal{V}_{\mathbf{x},\mathbf{y}} \leq \rho\big(\beta^{-1} s^\gamma u^{1-\gamma}\, d(x,y)^d\big). \tag{2.1}$$

Observe that the graph sequence $(\mathscr{G}(\mathcal{X}_t, \mathcal{V}_t))_{t>0}$ has the law of our spatial preferential attachment model and is therefore constructed on the probability space carrying the Poisson process $\mathcal{X}$ and the sequence $\mathcal{V}$. Moreover, it extends to a deterministic mapping associating a graph structure to any locally finite set of points in $\mathcal{Y} \subseteq \mathbb{T}_a \times (0,\infty)$ and sequence $\mathcal{V}$ in $(0,1)$, where $\mathbf{x}$, $\mathbf{y}$ are connected if and only if (2.1) holds. We permit the case $a = \infty$, with $\mathbb{T}_\infty = \mathbb{R}^d$.

For finite $t > 0$, we define the *rescaling mapping*

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

which expands space by a factor of $t^{1/d}$ and time by a factor of $1/t$. The mapping h_t operates canonically on the set $\mathcal{X}_t$ as well as on $\mathcal{V}_t$ by $h_t(\mathcal{V}_t)(h_t(\mathbf{x}), h_t(\mathbf{y})) := \mathcal{V}_t(\mathbf{x},\mathbf{y})$, and also on graphs with vertex set in $\mathcal{X}_t$ 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 $\mathbf{y}$. As

$$\rho\big(\beta^{-1}(s/t)^\gamma (u/t)^{1-\gamma}\, d(t^{1/d}x, t^{1/d}y)^d\big) = \rho\big(\beta^{-1} s^\gamma u^{1-\gamma}\, d(x,y)^d\big)$$

the operation h_t preserves the rule (2.1) and therefore

$$\mathscr{G}(h_t(\mathcal{X}_t), h_t(\mathcal{V}_t)) = h_t(\mathscr{G}(\mathcal{X}_t, \mathcal{V}_t)).$$

In plain words, it is the same to construct the graph and then rescale the picture, or to first rescale the picture and then construct the graph on the rescaled picture.

We now denote $\mathcal{X}^t = \mathcal{X} \cap (\mathbb{T}_t \times (0,1])$ and by $\mathcal{V}^t$ the restriction of $\mathcal{V}$ to indices in $\mathcal{X}^t \times \mathcal{X}^t$. This gives rise to a graph $\mathscr{G}^t := \mathscr{G}(\mathcal{X}^t, \mathcal{V}^t)$. As $h_t(\mathcal{X}_t)$ is a Poisson point process of unit intensity on $\mathbb{T}_t \times (0,1]$ and $h_t(\mathcal{V}_t)$ are independent uniform marks attached to the potential edges, the graph $\mathscr{G}^t$ has the same law as $h_t(\mathscr{G}_t)$. However, the *process* $(\mathscr{G}^t)_{t>0}$ behaves differently from the original process $(\mathscr{G}_t)_{t>0}$. Indeed, while the degree of any fixed vertex in $(\mathscr{G}_t)_{t>0}$ goes to infinity, the degree of any fixed vertex in $(\mathscr{G}^t)_{t>0}$ stabilizes and the graph sequence converges to the graph $\mathscr{G}^\infty := \mathscr{G}(\mathcal{X}^\infty, \mathcal{V}^\infty)$.

The graph $\mathscr{G}^\infty$ is the *age-dependent random connection model*. Like in the classical geometric random graph models points are placed according to a Poisson point process $\mathcal{Y} \subseteq \mathbb{R}^d$, but now every point additionally carries a mark drawn independently from the uniform distribution on $(0,1)$. Given points and marks, we independently connect two points in position x with mark u, resp. position y with mark s, with probability

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

We define the rooted graph $\mathscr{G}_0^\infty$ as the *Palm version* of the age-dependent random connection model $\mathscr{G}^\infty$; which differs from $\mathscr{G}^\infty$ by an additional point with uniform mark placed at the origin. By a law of large numbers as in Penrose and Yukich [**22**], for any bounded, "local" nonnegative function H acting on locally finite rooted graphs,

$$\lim_{t\to\infty} \frac{1}{t} \sum_{\mathbf{x}\in\mathscr{G}_t} H(\theta_{\mathbf{x}}\mathscr{G}_t) = \mathbb{E}[H(\mathscr{G}_0^\infty)] \qquad \text{in probability,} \tag{2.2}$$

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 means precisely that $\mathscr{G}_0^\infty$ is the weak local limit of $(\mathscr{G}_t)_{t>0}$.

3. Clustering in the spatial preferential attachment model

If G is a finite graph, we call a pair of edges in G a *wedge* if they share an endpoint (called its *tip*). To show that the age-based spatial preferential attachment model is clustering we compare the number of triangles and wedges with tip at a fixed vertex $\mathbf{x}$. For a vertex $\mathbf{x}$ with at least two neighbours, define the *local clustering coefficient* of a graph G as

$$c_{\mathbf{x}}^{\mathrm{loc}}(G) := \frac{\text{Number of triangles in } G \text{ containing vertex } \mathbf{x}}{\text{Number of wedges with tip } \mathbf{x} \text{ in } G},$$

which is also an element of $[0,1]$. Let $V_2(G) \subseteq G$ be the set of vertices in G with degree at least two, and define the *average clustering coefficient* by

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

For the average clustering coefficient we then get from the weak local limit theorem (2.2),

$$c^{\mathrm{av}}(\mathscr{G}_t) \longrightarrow \int_0^1 \mathbb{P}\big\{(X_u^{(1)}, S_u^{(1)}) \sim (X_u^{(2)}, S_u^{(2)})\big\}\, \pi(du),$$

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

$$\frac{1}{\lambda_u}\Big(\rho(\beta^{-1}s^{1-\gamma}u^{\gamma}|x|^d)\mathbb{1}_{\,(u,1]}(s) + \rho(\beta^{-1}u^{1-\gamma}s^{\gamma}|x|^d)\mathbb{1}_{\,[0,u]}(s)\Big)\, dx\, ds, \tag{2.1}$$

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 $1 - e^{-\lambda_u} - \lambda_u e^{-\lambda_u}$. We see that the average clustering coefficient of $(\mathscr{G}_t)$ converges to a positive constant indicating strong clustering of the model.

CHAPTER 3

Scale-free geometric random graphs

We now look at generalisation of the age-dependent random-connection model. The age parameter is replaced by a mark, which can play the role of a weight or radius. Although the behaviour of the general models can differ vastly, there is some common ground and synergy in studying them together.

1. Weight-dependent random connection model

We introduce the *weight-dependent random connection model* alternatively known as *kernel-based random geometric graphs*. The vertex set of the graph $\mathscr{G} = \mathscr{G}_\beta$ 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 *mark* t. 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

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

of the form

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

for a non-increasing, integrable *profile function*

$$\rho : \mathbb{R}_+ \to [0,1]$$

and a *kernel*

$$g\colon (0,1] \times (0,1] \to \mathbb{R}_+,$$

which is symmetric and non-decreasing in both arguments, so that we give preference to short edges or edges that are connected to vertices with large weight. We assume (without loss of generality) that

$$\int_{\mathbb{R}^d} \rho(|x|^d)\,dx = 1. \tag{3.1}$$

Then, the degree distribution of a vertex only depends on the function g. Indeed, the expected degree of a vertex $\mathbf{x} = (x,s)$ is

$$\begin{aligned}\int_0^1 dr \int_{\mathbb{R}^d} dy\, \rho\big(g(r,s)|x-y|^d\big) &= \int_0^1 dr \int_{\mathbb{R}^d} dy\, \rho\big(|g(r,s)^{1/d}y|^d\big)\\ &= \int_0^1 dr\, g(r,s)^{-1} \int_{\mathbb{R}^d} \rho(|x|^d)\,dx = \int_0^1 g(r,s)^{-1} dr.\end{aligned}$$

While the profile function ρ has no influence on the degree distribution, it does influence the intensity of long edges in the graph. Roughly speaking, the heavier the tails of ρ the weaker the influence of the geometry. If ρ is an indicator function the model is *purely geometric*, i.e. whether $\mathbf{x} \sim \mathbf{y}$ depends only on the vertex positions in $\mathbb{R}^d \times (0,1]$ and not on extra randomness. But when the tails of ρ are decaying polynomially at infinity long edges can occur spontaneously and geometric restrictions are softened. These are known as *weak links* in the networks literature.

2. Some interesting kernels

We next give explicit examples for the function g. 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' marks on the connection probability; the larger γ, the stronger the preference of connecting to vertices with small mark. All kernel functions we consider lead to models that are *scale-free* with power law exponent

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

see [**16, 14**]. 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 [**14**]. Our focus is on the following examples, see [**16**].

- The *sum kernel*, defined as
$$g(s,t) = \beta^{-1}(s^{-\gamma/d} + t^{-\gamma/d})^{-d}.$$
The interpretation of $(\beta a s^{-\gamma})^{1/d}, (\beta a t^{-\gamma})^{1/d}$ 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. Indeed, with $r(\mathbf{x}) = (\beta a)^{1/d} s^{-\gamma/d}$ we have
$$g(s,t)|x-y|^d \le a \Leftrightarrow |x-y| \le (\beta a)^{1/d} s^{-\gamma/d} + (\beta a)^{1/d} t^{-\gamma/d}$$
$$\Leftrightarrow \frac{|x-y|}{r(\mathbf{x}) + r(\mathbf{y})} \le 1.$$
A natural generalisation replaces the hard constraint by a soft constraint
$$\frac{|x-y|}{r(\mathbf{x}) + r(\mathbf{y})} \le X,$$
where X is an independent random variable for every pair $\mathbf{x}, \mathbf{y}$ of vertices. If ρ is the (decreasing) tail probability function of aX^d then $U = \rho(aX^d)$ is uniformly distributed and hence the probability of a connection between $\mathbf{x}$ and $\mathbf{y}$ is
$$\rho(g(t,s)|x-y|^d).$$
We call this model the *soft Boolean model*. It allows for *weak links* between far away vertices even if they are not particularly powerful.

- The *strong kernel*, defined as
$$g(s,t) = \beta^{-1}(s \wedge t)^{\gamma}.$$
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 the ratio of strong kernel and sum kernel is bounded from zero and infinity they show qualitatively similar behaviour.

- The *product kernel*, defined as
$$g(s,t) = \beta^{-1} s^{\gamma} t^{\gamma}.$$
This is a continuum space version of *scale-free percolation*, a model introduced by Deijfen, van der Hofstad, Hooghiemstra [**8**]. Results specifically for the continuum can be found, for example, in [**9, 7**]. Note that the much studied *random hyperbolic graphs* can also be mapped on this model, see [**7, 21**].

- The *preferential attachment kernel*, defined as
$$g(s,t) = \beta^{-1}(s \vee t)^{1-\gamma}(s \wedge t)^{\gamma} \tag{3.1}$$
gives rise to the *age-dependent random connection model* introduced by Gracar et al. [**14**]. As seen above this model is the weak local limit of the age-based spatial preferential attachment model. In this model, s and t play the role of the vertices' birth times in the underlying dynamic network. We therefore also refer to vertices with small s as old vertices.
- The *weak kernel*, defined as
$$g(s,t) = \beta^{-1}(s \vee t)^{1+\gamma}$$
gives rise to a model where connections require both kernels to be strong to make a connection likely. It generalises the scale-free random graph model of Yukich [**24**] by allowing a general profile function and places points randomly rather than at lattice points. If $\rho = 1_{[0,a]}$ is an indicator function we connect two vertices (x,s) and (y,t) if both
$$x \in B(y, (a\beta)^{1/d} t^{-\frac{1+\gamma}{d}}) \text{ and } y \in B(x, (a\beta)^{1/d} s^{-\frac{1+\gamma}{d}}).$$
- One obtains the classical *random connection model* or *long-range percolation* by choosing the kernel g constant and allowing for a general profile function ρ. As this graph is (interesting but) not scale-free it is not in our focus in this course. However, in some cases we will be able to extend methods designed for long-range percolation in our proofs.

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
$$\lim_{r\uparrow\infty} \frac{\rho(cr)}{\rho(r)} = c^{-\delta} \quad \text{for all } c \geq 1. \tag{3.2}$$

3. Exercise 1: The weak kernel: Non-spatial and spatial

(a) Identify the weak local limit of the inhomogeneous random graph with weak kernel and find the parameters $0 < \beta, \gamma < 1$ for which these random graphs have a giant component.
(b) Find the parameters for which the weight-dependent random connection model with weak kernel has an infinite component.

Sketch of solution (a): We still use
$$t_n = \sum_{k=1}^{n-1} \frac{1}{k}$$
and the mapping
$$\phi_N \colon \{1, \dots, N\} \to (-\infty, 0], \quad n \mapsto t_n - t_N.$$
We fix a point $uN \in \mathscr{G}_N$ and look at the point process consisting of all points $\phi_N(j)$ with $j \sim uN$. The claim is that this process converges to the Poisson point process $\log u + \pi_u$ with intensity measure
$$\nu_u(dz) = \beta u^{-\gamma}(e^z 1_{z<0} + e^{-\gamma z} 1_{z>0})\, dz.$$
Take $a < b < u$, then
$$\mathbb{P}\big(uN \text{ does not connect to any } j \in (aN, bN)\big) = \prod_{i=aN}^{bN} \big(1 - \beta \frac{1}{N} u^{-\gamma-1}\big)$$
$$\to \exp\big(-\beta(b-a)u^{-\gamma-1}\big) = \exp\big(-\beta u^{-\gamma} \int_{\log a - \log u}^{\log b - \log u} e^z\, dz\big).$$

Take $u < a < b < 1$, then

$$\mathbb{P}\big(uN \text{ does not connect to any } j \in (aN, bN)\big) = \prod_{i=aN}^{bN} \big(1 - \beta \frac{1}{N}(i/N)^{-\gamma-1}\big)$$

$$\to \exp\big(-\beta \frac{1}{N} \sum_{i=aN}^{bN} (i/N)^{-\gamma-1}\big) = \exp\big(-\beta \int_a^b z^{-\gamma-1}\, dz\big) = \exp\big(-\beta u^{-\gamma} \int_{\log a - \log u}^{\log b - \log u} e^{-\gamma y}\, dy\big).$$

Hence the probability that there are no points $\phi_N(j) \in [\log a, \log b]$ with $j \sim un$ converges to the probability that $\pi_u([\log a - \log u, \log b - \log u]) = 0$ and, by essentially the same calculation, the expected number of points $\phi_N(j) \in [\log a, \log b]$ with $j \sim un$ converges to the expectation of $\pi_u([\log a - \log u, \log b - \log u])$. By Kallenberg's theorem this implies convergence to π_u.

The branching process with the given (state-dependent) offspring distribution is the weak local limit of the inhomogeneous random graph with weak kernel. Note that this is not a branching random walk anymore. But just restricting attention to the offspring left of the parent one can see that starting from a sufficiently powerful vertex this process is dominating a supercritical Galton-Watson process and hence a giant component exists for any $\gamma, \beta > 0$.

Sketch of solution (b): Again such a component exists for any $\gamma, \beta > 0$. The non-spatial model has taught us that in this case we can build a large connected component by going to more and more powerful vertices without detour.

Suppose we start with a vertex (x, s), for s small. This vertex is at the center of a ball with volume of order $s^{-\gamma-1}$, it will connect to a vertex (y, t) in this ball with $t < s$ with probability bounded from zero. For every $1 < \eta < \gamma + 1$ with high probability there exists a large number of vertices (y, t) in this ball with $0 < t < s^\eta$. They all connect independently with a fixed positive probability to (x, s). Hence with high probability we find an edge from (x, s) to (y, t) with $t < s^\eta$. As the error probability can be made exponentially decreasing we can continue this chain ad infinitum and get an infinite component with a probability that goes to one as the initial mark $s \downarrow 0$.

CHAPTER 4

Phase transitions in scale-free geometric random graphs

1. Existence and non-existence of a subcritical phase

We have seen in Section 1 that in the non-spatial preferential attachment model a subcritical phase, i.e. a phase where for small $\beta > 0$ there is no giant component almost surely, exists if $\gamma < \frac{1}{2}$ but not if $\gamma \geq \frac{1}{2}$. We now ask the analogous question for the weight-dependent random connection model, i.e. whether

$$\beta_c := \inf\{\beta > 0 \colon \mathscr{G}_\beta \text{ has an infinite component}\}$$

is positive. It is shown in [**8, 9**] that for the product kernel we have the same behaviour as in the non-spatial models, i.e. if $\gamma < \frac{1}{2}$, then $\beta_c > 0$, if $\gamma \geq \frac{1}{2}$, then $\beta_c = 0$. For the weak kernel we always have $\beta_c = 0$. The interesting problem are the other kernels. The following result is from [**17**].

Theorem. For the weight-dependent random connection model with preferential attachment, sum or strong kernel and parameters $\delta > 1$ and $0 < \gamma < 1$, we have

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

Remark 1. This means that the weak links are necessary to make a scale-free geometric random graph have $\beta_c = 0$. On the other hand graphs with weak links that are not scale-free (like long-range percolation) do not have $\beta_c = 0$ either, so that only the combination of weak links and powerful vertices is successful in qualitatively changing the graph topology.

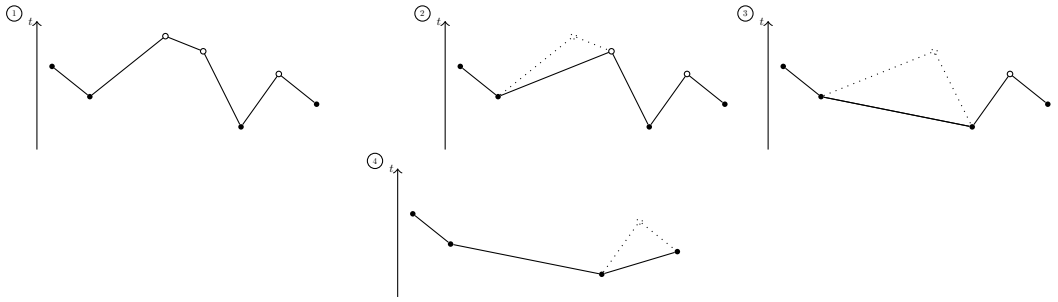
Remark 2. With some additional effort the result can be transferred to the spatial preferential attachment model, i.e. the dynamic network that motivates our study.

Sketch of upper bound. We can work with the strong kernel, which dominates the preferential attachment and sum kernels. If $\gamma > \frac{\delta}{\delta+1}$ we can find $\alpha \in (1, \frac{\gamma}{\delta(1-\gamma)})$. Starting from a vertex $\mathbf{x} = (x, s)$ we are looking for a vertex $\mathbf{y} = (y, t)$ with $t < s^\alpha$ that is connected to $\mathbf{x}$. Typically, the nearest such vertex is in a ball of volume $s^{-\alpha}$, but the probability of $\mathbf{y}$ connecting to $\mathbf{x}$ is at most $\rho(\frac{1}{\beta} s^{(\gamma-1)\alpha})$, which is very small. However, if we look at all vertices in a ball of volume $s^{-\gamma}$ around $\mathbf{y}$, for each one the probability that it is connected to both $\mathbf{x}$ and $\mathbf{y}$ is at least $\rho(\frac{1}{\beta} s^\gamma s^{-\alpha}) \rho(\frac{1}{\beta} s^{\alpha\gamma} s^{-\gamma}) \gg s^\gamma$, so that $\mathbf{x}$ and $\mathbf{y}$ are with high probability connected by a path of length two.

Hence, as $s \downarrow 0$ the probability that for a fixed vertex (x_0, s_0) with $s_0 < s$ there exists an infinite sequence of vertices $(x_0, s_0), (x_1, s_1), (x_2, s_2), \ldots$ such that

- $s_{k+1} < s_k^\alpha$ and $|x_{k+1} - x_k|^d < \frac{\beta}{2} s_k^{-\alpha}$, and
- (x_k, s_k) is connected to (x_{k+1}, s_{k+1}) by *a path of length two*;

converges to one.



Sketch of lower bound. As the strong kernel and 2^d times the sum kernel dominate the preferential attachment kernel we can work with the latter. A first idea is to bound the expected number of shortcut-free paths by

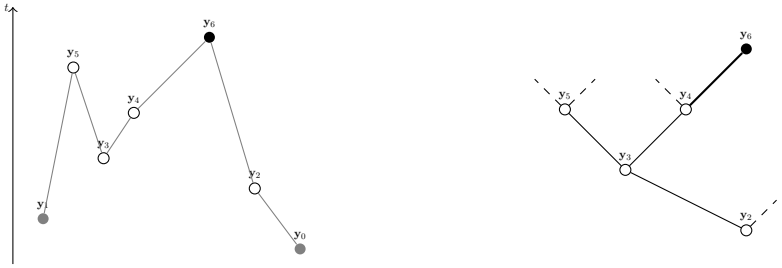
$$\mathbb{E}[\sharp\text{shortcut-free paths of length } n \text{ starting in } 0] \leq (C\beta)^n.$$

This would gives that the origin is not in an infinite component almost surely and hence there is none. This method works only if $\gamma < \frac{1}{2}$. If $\frac{1}{2} \leq \gamma < \frac{\delta}{\delta+1}$ the expectation is dominated by highly unlikely events and a more detailed analysis is needed.

We look at a path of length n and identify its *skeleton*, see figure above. Let $E(\mathbf{x} \overset{k}{\leftrightarrow} \mathbf{y})$ be the set of all shortcut-free paths connecting $\mathbf{x}$ and $\mathbf{y}$ by $k-1$ vertices with larger marks, then

$$\mathbb{E}_{\mathbf{x},\mathbf{y}}\left[\sharp E(\mathbf{x} \overset{k}{\leftrightarrow} \mathbf{y})\right] \leq (C\beta)^{k-1}\mathbb{P}_{\mathbf{x},\mathbf{y}}\{\mathbf{x} \sim \mathbf{y}\}.$$

The ways in which the $k-1$ vertices are inserted is encoded by a binary tree.



Add local maxima successively according to the tree using

$$\int_{t_0 \vee t_1}^1 dt_2 \int_{\mathbb{R}^d} dy_2 \; \varphi((y_0,t_0),(y_2,t_2))\,\varphi((y_2,t_2),(y_1,t_1)) \leq (C\beta)\,\mathbb{P}_{\mathbf{y}_0,\mathbf{y}_1}\{\mathbf{y}_0 \sim \mathbf{y}_1\}.$$

Now use a *truncated first moment method* on the *skeletons*. We only count paths that are free of shortcuts. A skeleton $S = (\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_m)$ of length m is *regular* if its lowest mark is larger than 2^{-m}. A path is *regular* if its skeleton is regular. An infinite path $(\mathbf{x}_0, \mathbf{x}_1, \dots)$ is *regular* if there exists N such that all paths $(\mathbf{x}_0, \dots, \mathbf{x}_n)$, $n \geq N$ are regular. By stopping a path when it goes below the threshold and using our tool, we show that if β is below some positive constant depending only on ρ, γ and d, almost surely every infinite path is regular. The probability that there exists a regular path of length n can be bounded by $(C\beta)^n$ combining our tool with the first moment method. Hence for $0 < \beta < \frac{1}{C}$ there are no infinite regular paths.

2. Existence and non-existence of a supercritical phase

As the weight-dependent random connection models are, thanks to the presence of long edges, more connected than a standard nearest neighbour or Gilbert graph model, the existence of a supercritical phase is obvious if $d \geq 2$. However this is an interesting issue in dimension one, when these models fail to have an infinite component for any choice of β. This problem is studied by Gracar, Lüchtrath and Mönch [**18**].

The idea is that in the long-range percolation model, i.e. when $g = \frac{1}{\beta}$ a supercritical phase exists (which means $\beta_c < \infty$) when $\delta < 2$ and it does not exist when $\delta > 2$. A heuristic argument for this is the following. We divide space in boxes of length n, or with n particles. The number of edges connecting two boxes at distance n is then approximately binomially distributed with parameters n^2 and $p \approx \frac{1}{\beta} n^{-\delta}$. Hence if $\delta > 2$ the probability that no edge of length $\ell \in (2n, 100n)$ across the origin exists is approximately

$$(1-p)^{n^2} \approx e^{-\frac{1}{\beta} n^{2-\delta}} \approx 1 - \tfrac{1}{\beta} n^{2-\delta},$$

whence the complementary event is summable for $n_k = 50^k n_0$, for $k = 1, 2, \ldots$. By a union bound the probability that there is no edge crossing the origin is positive. We conclude with the ergodic theorem that almost surely there is no infinite connected component.

Showing that for $\delta < 2$, when two boxes with n particles at distance of order n are typically connected by an abundance of edges, there is an infinite component for sufficiently large β, requires an elegant renormalisation argument of Duminil-Copin, Garban, and Tassion [**12**]. The rough idea is, for each scale, to divide the line into overlapping blocks so that each point is covered exactly twice. A block is called good if it contains a connected component of density $\theta > \frac{3}{4}$. In the smallest scale blocks can be made good with high probability by increasing β. If all subblocks of a block are good, then so is the big block containing them. The estimates on the connection probability of subblocks are used to show that if just one subblock is bad, then a block is still likely to be good, and then iteratively use the bound on the probability that two or more subblocks are bad.

This argument can be extended to our weight-dependent random connection models. Except for significant technical difficulties, the main thing that changes is the probability p that two vertices in boxes of length n and at distance n are connected. As the smallest mark in a box of length n is typically of order $1/n$, we assume that such a vertex has uniformly distributed mark on the interval $(1/n, 1)$ and hence

$$p = \int_{1/n}^{1} ds \int_{1/n}^{1} dt \rho(g(s,t)n),$$

which leads to the definition of an effective delta

$$\delta_{\text{eff}} := -\lim_{n\to\infty} \frac{\log \int_{1/n}^{1} ds \int_{1/n}^{1} dt \rho(g(s,t)n)}{\log n}.$$

Gracar, Lüchtrath and Mönch [**18**] prove that

$$\delta_{\text{eff}} < 2 \;\Rightarrow\; \beta_c < \infty$$

and

$$\delta_{\text{eff}} > 2 \;\Rightarrow\; \beta_c = \infty.$$

3. Exercise 2: Calculate δ_{eff} for the most important kernels.

Check that for $\delta > 2$,

- for the product kernel if $\gamma < \frac{1}{2}$, then $\delta_{\text{eff}} > 2$, but if $\gamma > \frac{1}{2}$, then $\delta_{\text{eff}} < 2$.
- for the sum and strong kernel if $\gamma < \frac{\delta-1}{\delta}$, then $\delta_{\text{eff}} > 2$, but if $\gamma > \frac{\delta-1}{\delta}$, then $\delta_{\text{eff}} < 2$.
- for the preferential attachment kernel if $\gamma < \frac{\delta-1}{\delta}$, then $\delta_{\text{eff}} = 2$, but if $\gamma > \frac{\delta-1}{\delta}$, then $\delta_{\text{eff}} < 2$.

Infer that for the sum kernel and strong kernel we have

$$\beta_c = \infty \text{ if } \gamma < \frac{\delta-1}{\delta},$$
$$0 < \beta_c < \infty \text{ if } \frac{\delta-1}{\delta} < \gamma < \frac{\delta}{\delta+1},$$
$$\beta_c = 0 \text{ if } \gamma > \frac{\delta}{\delta+1}.$$

For the preferential attachment kernel the behaviour for $\gamma < \frac{\delta-1}{\delta}$ is an open problem.

CHAPTER 5

Shortest paths

A hard problem, already for long-range percolation but particularly for scale-free geometric graphs, is to find the asymptotic length of the shortest path connecting two vertices on the infinite component that are far away. In the case of long-range percolation, for example, one expects

- shortest paths to be linear in the Euclidean distance if $\delta > 2$,
- and behave like a power of the logarithm of the distance if $\delta < 2$,

see Biskup [**5, 6**] and Berger [**2**] for partial results. For weight-dependent random connection models one expects an ultrasmall phase of shortest paths of loglog length in the Euclidean distance if γ is above some threshold. In this chapter we focus on this situation. We present results from a recent paper of Gracar, Grauer and Mörters [**15**].

1. Ultrasmallness with or without connectors

Let us first see how we can construct short paths between two vertices $\mathbf{x}$ and $\mathbf{y}$.

- for the *weak kernel* and *product kernel*.

  As $\mathbf{x}$ is in the infinite component it is connected in finitely many steps with a powerful vertex (x_0, s_0), which we use to start a path of increasingly powerful vertices (x_n, s_n) with $s_n = s_{n-1}^{\alpha}$ for some $\alpha > 1$. We do this starting from both $\mathbf{x}$ and $\mathbf{y}$ until we reach a scale s_n at which points (x_n, s_n) and (y_n, t_n) are connected with high (or positive) probability. Then the path constructed has length of order $2n$.

  We quickly calculate this for the weak kernel, which is the easiest case, see [**8, 9**] for calculations in the product kernel case if $\gamma > \frac{1}{2}$. In the weak kernel case we do not benefit from weak links, so we may assume that ρ has bounded support. Suppose $\mathbf{x} = (x, s)$ is given, then to be connected to $\mathbf{x}$ points need to be contained in a ball with volume of order $s^{-\gamma-1}$. The most powerful vertices $\mathbf{y} = (y, t)$ in this ball have $t \approx s^{\gamma+1}$ and, with high probability, we can find such a $\mathbf{y}$ connecting to $\mathbf{x}$. Hence the connecting strategy works for $1 < \alpha < \gamma + 1$.

- for the *strong*, *sum* and *preferential attachment kernel*.

  As indicated in Chapter 3 one cannot connect $\mathbf{x}_n = (x_n, s_n)$ directly to a vertex $\mathbf{x}_{n+1} = (x_{n+1}, s_{n+1})$ with $s_{n+1} < s_n^{\alpha}$ in these cases. Instead one looks for a *connector*, a vertex in a ball of radius $\approx s_n^{-\gamma}$ around $\mathbf{x}_{n+1}$ that connects to both $\mathbf{x}_n$ and $\mathbf{x}_{n+1}$. If $\gamma > \frac{\delta}{\delta+1}$ this is possible for $\alpha \in (1, \frac{\gamma}{\delta(1-\gamma)})$.

We obtain that the weight-dependent random connection model is ultrasmall

- for the weak kernel with any $\gamma > 0$,
- for the product kernel if $\gamma > \frac{1}{2}$,
- for the strong, sum and preferential attachment kernel if $\gamma > \frac{\delta}{1+\delta}$.

2. Non-ultrasmallness and a limit theorem

The criteria above are sharp, i.e. the weight-dependent random connection model is not ultrasmall for the product kernel if $\gamma < \frac{1}{2}$, and for the strong, sum and preferential attachment kernel if $\gamma < \frac{\delta}{1+\delta}$. To prove this and sharp limit theorems in the ultrasmall phase we need lower bounds. While these can be proved by similar means as in the non-geometric case for the weak and product kernel [**10, 7**], they are much harder to come by when we have a genuine dependence on the parameter δ. We get the following limit theorems, where the principal result is specialising results from [**15**].

THEOREM 5.1. *Let $\mathscr{G}$ be the weight-dependent random connection model. For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d \times (0,1)$ and*

- *the strong, sum or preferential attachment kernel with $\gamma > \frac{\delta}{\delta+1}$ we have*
$$d(\mathbf{x}, \mathbf{y}) = (4 + o(1)) \frac{\log\log|x-y|}{\log\left(\frac{\gamma}{\delta(1-\gamma)}\right)},$$
- *the product kernel with $\gamma > \frac{1}{2}$ we have*
$$d(\mathbf{x}, \mathbf{y}) = (2 + o(1)) \frac{\log\log|x-y|}{\log\left(\frac{\gamma}{1-\gamma}\right)},$$
- *the weak kernel with $0 < \gamma < 1$ we have*
$$d(\mathbf{x}, \mathbf{y}) = (2 + o(1)) \frac{\log\log|x-y|}{\log(1+\gamma)},$$

under $\mathbb{P}_{\mathbf{x},\mathbf{y}}(\ \cdot\ \mid \mathbf{x} \leftrightarrow \mathbf{y})$ with high probability as $|x-y| \to \infty$.

In the last section we illustrate the argument used for all lower bounds, the refined truncated moment method, in the case of the preferential attachment kernel with $\gamma > \frac{\delta}{\delta+1}$.

3. Refining the truncated moment method

To prove lower bounds we find an upper bound for $\mathbb{P}_{\mathbf{x},\mathbf{y}}\{d(\mathbf{x}, \mathbf{y}) \le 2\Delta\}$ and choose Δ as large as possible while keeping the probability sufficiently small. This event is equivalent to the existence of at least one path between $\mathbf{x}$ and $\mathbf{y}$ of length smaller than 2Δ. As seen before, estimating its probability by the expected number of short-cut free paths is only good enough if $\gamma < \frac{1}{2}$, as otherwise the expectation is dominated by very unlikely events. We therefore truncate the admissible mark of the vertices of a possible path between $\mathbf{x}$ and $\mathbf{y}$. We define a decreasing sequence $(\ell_k)_{k\in\mathbb{N}_0}$ of thresholds and call a tuple of vertices $(\mathbf{x}_0, \dots, \mathbf{x}_n)$ *good* if their marks satisfy $t_k \wedge t_{n-k} \ge \ell_k$ for all $k \in \{0, \dots, n\}$. A path consisting of a good tuple of vertices is called a good path. We denote by $A_k^{(\mathbf{x})}$ the event that there exists a path starting in $\mathbf{x}$ which fails this condition after exactly k steps, i.e. a path $((x,t),(x_1,t_1),\dots(x_k,t_k))$ with $t \ge \ell_0, t_1 \ge \ell_1, \dots, t_{k-1} \ge \ell_{k-1}$, but $t_k < \ell_k$. Furthermore we denote by $B_n^{(\mathbf{x},\mathbf{y})}$ the event that there exists a good path of length n between $\mathbf{x}$ and $\mathbf{y}$. Then, for given vertices $\mathbf{x}$ and $\mathbf{y}$

$$\mathbb{P}_{\mathbf{x},\mathbf{y}}\{d(\mathbf{x}, \mathbf{y}) \le 2\Delta\} \le \sum_{n=1}^{\Delta} \mathbb{P}_{\mathbf{x}}(A_n^{(\mathbf{x})}) + \sum_{n=1}^{\Delta} \mathbb{P}_{\mathbf{y}}(A_n^{(\mathbf{y})}) + \sum_{n=1}^{2\Delta} \mathbb{P}_{\mathbf{x},\mathbf{y}}(B_n^{(\mathbf{x},\mathbf{y})}). \tag{5.1}$$

This decomposition is the same as for the non-spatial models in [**11**]. The main feature of our proof is to show that the geometric restrictions and resulting correlations in our spatial random graphs make it much more difficult for a path to connect to a vertex with small mark. Hence a larger sequence (ℓ_k) of thresholds can be chosen that still makes the two first sums on the right of (5.1) small, allowing the third sum to be small for a larger choice of Δ.

The proof of the upper bound suggests that the best way to connect (x,s) to a more powerful vertex (y,t) is to use a path of length two. In the main technical part of the proof this is verified. This is difficult as the abundance of other possibilities introduces an entropy term in the estimates that needs to be controlled by 'path surgery' and clever coarse graining. Supposing this can be done, we now show how the truncation sequences $(\ell_k)_{k\in\mathbb{N}_0}$ can be constructed using the geometry of the optimal paths.

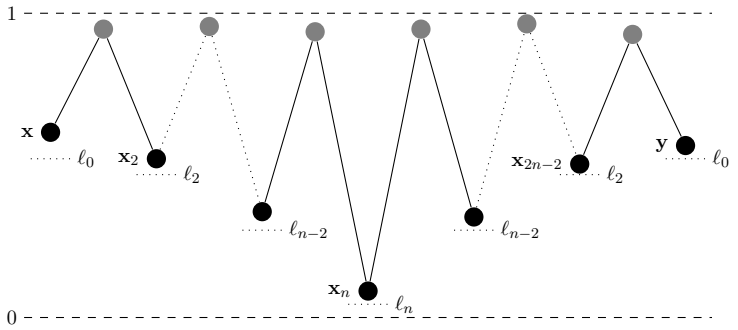


FIGURE 1. An example of a path with optimal connection type for $\gamma > \frac{\delta}{\delta+1}$. The horizontal axis corresponds to the sequential numbering of vertices on the path, the vertical axis represents the mark space. Powerful vertices (indicated by black dots) alternate with connectors (indicated by grey dots).

We now establish an upper bound for the probability of the event $A_n^{(\mathbf{x})}$ that there exists a path starting in $\mathbf{x}$ whose n-th vertex is the first vertex which has a mark smaller than the corresponding ℓ_n. We denote by $N(\mathbf{x},\mathbf{y},n)$ the number of paths of length n from $\mathbf{x}=(x,t)$ to a vertex $\mathbf{y}=(y,s)$ whose vertices $(x_1,t_1),\dots(x_{n-1},t_{n-1})$ fulfill $t_{2(k+1)} < t_{2k} < t_{2k+1}$ for all $k=0,\dots,\lceil n/2\rceil-1$ and which is one half of a good path, i.e. $t\geq\ell_0, t_1\geq\ell_1,\dots,t_{n-1}\geq\ell_{n-1}$. The mark of $\mathbf{y}$ is not restricted in this definition and is therefore allowed to be smaller than ℓ_n. Hence, in this case the event $A_n^{(\mathbf{x})}$ can only occur for n even and we have, with $\mathbf{z}_0=\mathbf{x}$,

$$\mathbb{P}_{\mathbf{x}}(A_n^{(\mathbf{x})}) \leq \int\limits_{\mathbb{R}^d\times(\ell_2,t_0]} d\mathbf{z}_1\cdots \int\limits_{\mathbb{R}^d\times(\ell_{n-2},t_0]} d\mathbf{z}_{n/2-1} \int\limits_{\mathbb{R}^d\times(0,\ell_n]} d\mathbf{z}_{n/2}\, \mathbb{E}_{\mathbf{z}_0,\dots,\mathbf{z}_{n/2}} \prod_{i=1}^{n/2} e_K(\mathbf{z}_i,\mathbf{z}_{i-1},2),$$

where, for $\mathbf{x}=(x,t)$, $\mathbf{y}=(y,s)$, $\mathbf{z}=(z,u)$, we set

$$e_K(\mathbf{x},\mathbf{y},2) = \int\limits_{\mathbb{R}^d\times(t\vee s,1)} d\mathbf{z}\, \rho\big(\tfrac{1}{\beta}t^\gamma u^{1-\gamma}|x-z|^d\big)\rho\big(\tfrac{1}{\beta}s^\gamma u^{1-\gamma}|y-z|^d\big).$$

As either $|x-z|\geq\frac{|x-y|}{2}$ or $|y-z|\geq\frac{|x-y|}{2}$, we have for two given vertices $\mathbf{x},\mathbf{y}$ far enough from each other, after a calculation

$$\begin{aligned} e_K(\mathbf{x},\mathbf{y},2) \leq{}& \int\limits_{t\vee s}^1 du\rho\big(2^{-d}\tfrac{1}{\beta}t^\gamma u^{1-\gamma}|x-y|^d\big)\int\limits_{\mathbb{R}^d} dz\rho\big(\tfrac{1}{\beta}s^\gamma u^{1-\gamma}|y-z|^d\big) \\ &+ \int\limits_{t\vee s}^1 du\rho\big(2^{-d}\tfrac{1}{\beta}s^\gamma u^{1-\gamma}|x-y|^d\big)\int\limits_{\mathbb{R}^d} dz\rho\big(\tfrac{1}{\beta}t^\gamma u^{1-\gamma}|x-z|^d\big) \\ \leq{}& C_\beta\,\rho\big((t\wedge s)^\gamma(t\vee s)^{\gamma/\delta}|x-y|^d\big). \end{aligned}$$

Note the influence of the spatial embedding via the parameter δ. For a sufficiently large constant $c > 0$ the probability of $A_n^{(\mathbf{x})}$ can be bounded by

$$c^{n/2}\ell_n^{1-\gamma}\ell_0^{-\gamma/\delta}\prod_{i=1}^{n/2-1}\ell_{2i}^{1-\gamma-\gamma/\delta}.$$

With ℓ_0 smaller than the mark of $\mathbf{x}$ we choose the truncation sequence so that the product above equals $\frac{\varepsilon}{\pi^2 n^2}$. Then we have

$$\sum_{n=1}^{\Delta}\mathbb{P}_{\mathbf{x}}(A_n^{(\mathbf{x})}) = \sum_{\substack{n=1\\ n\ even}}^{\Delta}\mathbb{P}_{\mathbf{x}}(A_n^{(\mathbf{x})}) \leq \sum_{\substack{n=1\\ n\ even}}^{\Delta} c^{n/2}\ell_n^{1-\gamma}\ell_0^{-\gamma/\delta}\prod_{i=1}^{n/2-1}\ell_{2i}^{1-\gamma-\gamma/\delta} \leq \sum_{n=1}^{\infty}\frac{\varepsilon}{\pi^2 n^2} = \frac{\varepsilon}{6}.$$

Writing $\eta_n := \ell_n^{-1}$ we can deduce a recursive description of $(\ell_n)_{n\in\mathbb{N}_0}$ such that

$$\eta_{n+2}^{1-\gamma} = \frac{(n+2)^2}{n^2}c\eta_n^{\gamma/\delta}.$$

Consequently there exist $b > 0$ and $B > 0$ such that $\eta_n \leq b\exp(B(\gamma/(\delta(1-\gamma)))^{n/2})$. We close the argument with heuristics that leads from this truncation sequence to a lower bound for the chemical distance. Let $\mathbf{x}$ and $\mathbf{y}$ be two given vertices. If there exists a path of length $n < \log|x-y|$ between them, there must exist at least one edge in this path which is longer than $\frac{|x-y|}{\log|x-y|}$. For $|x-y|$ large, this edge typically must have an endvertex whose mark is, up to a multiplicative constant, smaller than $|x-y|^{-d}$. Hence, if we choose

$$\Delta < (2+o(1))\frac{\log\log|x-y|}{\log\left(\frac{\gamma}{\delta(1-\gamma)}\right)}$$

we ensure ℓ_Δ is of larger order than $|x-y|^{-d}$. Therefore there is no good path whose vertices are old enough to be an endvertex of an edge longer than $\frac{|x-y|}{\log|x-y|}$ and consequently no good path of length shorter than 2Δ can exist between $\mathbf{x}$ and $\mathbf{y}$.

Bibliography

[1] Itaj Benjamini and Oded Schramm. "Recurrence of distributional limits of finite planar graphs." Electron. J. Probab. 6 13 pp. (2001).

[2] Noam Berger. "A lower bound for the chemical distance in sparse long-range percolation models." Preprint arxiv:math.PR/0409021.

[3] Shankar Bhamidi. "Universal techniques to analyze preferential attachment tree and networks: Global and Local analysis." Published on author's homepage.

[4] J.D. Biggins, B.D. Lubachevsky, A. Shwartz and A. Weiss. "A branching random walk with a barrier." Ann. Appl. Probab. 1, 573-581 (1991).

[5] Marek Biskup. "On the scaling of the chemical distance in long-range percolation models." Ann. Probab. 32, 2938-2977 (2004).

[6] Marek Biskup and Jeffrey Lin. "Sharp asymptotic for the chemical distance in long-range percolation." Random Struct. Alg. 55 (2019) 560–583.

[7] Karl Bringmann, Ralph Keusch and Johannes Lengler. "Average distance in a general class of scale-free networks with underlying geometry." Preprint arXiv:1602.05712.

[8] Maria Deijfen, Remco van der Hofstad, and Gerard Hooghiemstra. "Scale-free percolation." Annales de l'Institut Henri Poincaré, 49, 817-838 (2013).

[9] Philippe Deprez and Mario Wüthrich. "Scale-free percolation in continuum space." Communications in Mathematics and Statistics, 7, 69-308 (2019).

[10] Steffen Dereich, Peter Mörters. "Random networks with sublinear preferential attachment: The giant component." The Annals of Probability, 41, 329-384 (2013).

[11] Steffen Dereich, Peter Mörters and Christian Mönch. "Typical distances in ultrasmall random networks." Advances in Applied Probability. 44, 583-601 (2012).

[12] Hugo Duminil-Copin, Christophe Garban, Vincent Tassion. "Long-range models in 1D revisited." Preprint arXiv:2011.04642.

[13] Maren Eckhoff, Peter Mörters and Marcel Ortgiese. "Near critical preferential attachment networks have small giant components." Journal of Statistical Physics 173, 663-703 (2018).

[14] Peter Gracar, Arne Grauer, Lukas Lüchtrath and Peter Mörters. "The age-dependent random connection model." Queueing Systems 93, 309-331 (2019).

[15] Peter Gracar, Arne Grauer and Peter Mörters. "Chemical distance in geometric random graphs with long edges and scale-free degree distribution" Preprint arXiv:2108.11252.

[16] Peter Gracar, Markus Heydenreich, Christian Mönch and Peter Mörters. "Recurrence versus transience for weight-dependent random connection models" Preprint arXiv:1911.04350.

[17] Peter Gracar, Lukas Lüchtrath and Peter Mörters. "Percolation phase transition in weight-dependent random connection models." Advances in Applied Probability 53, 1090-1114 (2021).

[18] Peter Gracar, Lukas Lüchtrath and Christian Mönch. "Finiteness of the percolation threshold for inhomogeneous long-range models in one dimension." Preprint arXiv:2203.11966.

[19] Remco van der Hofstad. "Random graphs and complex networks". Cambridge University Press, 2011.

[20] Emmanuel Jacob, Peter Mörters. "Robustness of scale-free spatial networks." Ann. Probab. 45, 1680-1722 (2017).

[21] Dmitri Krioukov et al. "Hyperbolic geometry of complex networks." Phys. Rev. E 82, 036106.

[22] Mathew D. Penrose, J. E. Yukich. "Weak laws of large numbers in geometric probability." Ann. Appl. Probab. 13, 277-303 (2003).

[23] Sidney Resnick. "Extreme values, regular variation, and point processes." Springer-Verlag, 1987.

[24] J. E. Yukich "Ultra-small scale-free geometric networks." Journal of Applied Probability, 43, 665-677 (2006).