UNIVERSITY OF LEEDS

This is a repository copy of *The age-dependent random connection model*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/200619/

Version: Published Version

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

The age-dependent random connection model

Peter Gracar[1] · Arne Grauer[1] · Lukas Lüchtrath[1] · Peter Mörters[1]

Abstract

We investigate a class of growing graphs embedded into the d-dimensional torus where new vertices arrive according to a Poisson process in time, are randomly placed in space and connect to existing vertices with a probability depending on time, their spatial distance and their relative birth times. This simple model for a scale-free network is called the *age-based spatial preferential attachment network* and is based on the idea of preferential attachment with spatially induced clustering. We show that the graphs converge weakly locally to a variant of the random connection model, which we call the *age-dependent random connection model*. This is a natural infinite graph on a Poisson point process where points are marked by a uniformly distributed age and connected with a probability depending on their spatial distance and both ages. We use the limiting structure to investigate asymptotic degree distribution, clustering coefficients and typical edge lengths in the age-based spatial preferential attachment network.

Keywords Scale-free networks · Benjamini–Schramm limit · Random connection model · Preferential attachment · Geometric random graphs · Spatially embedded graphs · Clustering coefficient · Power-law degree distribution · Edge lengths

Mathematics Subject Classification Primary 05C80; Secondary 60K35

1 Motivation: scale-free networks and clustering

Networks arising in different contexts, be it social, communication, technological or biological networks, often have strikingly similar features. The question of interest to us is why this is the case. Can these features be explained by a few basic principles underpinning the construction of these networks?

✉ Peter Mörters
  moerters@math.uni-koeln.de

1 Department Mathematik/Informatik, Universität zu Köln, 50931 Cologne, Germany

The probabilistic methodology to approach this question is to build a network model as a growing sequence of random graphs defined from very simple interaction principles and to prove the emerging features in the form of limit theorems when the number of vertices is going to infinity. The purpose of this paper is to present a tractable network model which is based on simple and natural construction principles and for which such limit theorems can be proved rigorously. We will call this model the *age-based spatial preferential attachment model*.

Potential features of networks we are interested in include:

- Networks are *scale-free*: For very large network size and large k, the proportion $\mu(k)$ of nodes with exactly k neighbours is of order $k^{-\tau+o(1)}$ for some *power-law exponent* τ.
- Networks are *ultrasmall*: The shortest path between two randomly chosen nodes in the graph is doubly logarithmic in the number of vertices.
- Networks are *robust* under random attack: If an arbitrarily large proportion of links is randomly removed from the network, the qualitative topological features of the network remain unchanged.
- Networks are *vulnerable* under targeted attack: Even if only a small number of the most influential nodes are removed, the topological features of the network change dramatically.
- Networks show strong *clustering*: Nodes picked from the neighbourhood of a typical node have a much higher chance of being connected by a link than randomly picked nodes.

These features should emerge solely from the principles on which our model rests. In this paper, the focus is on the scale-free and clustering properties of our model. We also believe that the other properties hold in certain parameter ranges; these (harder) properties are left for future research we will undertake.

The simple building principles for our network are:

- They are built dynamically by adding nodes successively.
- When a new node is introduced, it prefers to establish links to existing nodes that are either
    - powerful or old;
    - or similar to the new node.

The idea of building a network by connecting incoming nodes to existing nodes with a probability depending increasingly on their power was introduced into network theory by Barabási and Albert [2]. They use the degree of the existing node as the indicator of its power; we speak of *degree-based preferential attachment*. There is now a substantial body of work showing rigorously that the resulting networks are scale-free [4] and, for power-law exponent $\tau < 3$, they are ultrasmall and robust [8–10]. The key technical tool in the proofs of the latter properties is the coupling of neighbourhoods of typical vertices to well-studied random tree models often coming from the genealogy of branching processes. This technique rests therefore crucially on the absence of clustering, as clustering leads to presence of short cycles that destroy the local tree structure of the graphs.

To include clustering the idea of preferential attachment has to be developed further. An attractive approach is to include the idea that nodes have individual features and similarity of the features of nodes is a further incentive to form links between two nodes. This is realised by embedding the graphs into space and giving preference to short edges. We speak of *spatially induced clustering*. Spatial preferential attachment models were studied by Manna and Sen [22], Flaxman et al. [11,12], Aiello et al. [1], Jordan [17,18], Janssen et al. [16], Jordan and Wade [19], and Jacob and Mörters [14, 15].

The spatial preferential attachment models studied in these papers appear to be too complicated to fully characterise features like robustness or ultrasmallness. We therefore propose a simpler spatial model where preferential attachment is not to vertices with high degree but to vertices with old age; we speak of *age-based preferential attachment*. Age-based models are easier to study because while the actual degree of a vertex depends in a complex way on the graph geometry, the age of a vertex is a given quantity. At the same time, there is a strong link between degree and age, as the expected degree is a simple function of the age of a vertex. Our simplification therefore removes complicated but (on a large scale) inessential correlations between edges and allows us to focus on the important correlations, namely those coming from the spatial embedding.

2 The age-based spatial preferential attachment network

The *age-based spatial preferential attachment model* is a growing sequence of graphs $(G_t)_{t>0}$ in continuous time. The vertices of the graphs are embedded in the d-dimensional torus $\mathbb{T}_1^d = (-1/2, 1/2]^d$ of side-length one, endowed with the torus metric d defined by

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

where, here and throughout the paper, $|\cdot|$ denotes the Euclidean norm. Vertices are denoted by $\mathbf{y} = (y, s)$ and they are characterised by their birth time $s > 0$ and by their position $y \in \mathbb{T}_1^d$.

At time $t = 0$ the graph 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 $\mathbb{T}_1^d$.
- Given the graph 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

$$\varphi\left(\frac{t \cdot d(x, y)^d}{\beta \cdot \left(\frac{t}{s}\right)^\gamma}\right), \tag{2.1}$$

where

(a) $\varphi \colon [0, \infty) \to [0, 1]$ is the *profile function*. It is nonincreasing, integrable and normalised in the sense that

$$\int_{\mathbb{R}^d} \varphi(|x|^d)\, \mathrm{d}x = 1. \tag{2.2}$$

The profile function can be used to control the occurrence of long edges.

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

(c) $\beta \in (0, \infty)$ is a parameter to control the edge density, which is asymptotically equal to $\frac{\beta}{1-\gamma}$, hence the smaller β, the sparser the graph.

Some comments on our choices in (2.1) are in order.

(i) For any $r > 0$, the profile function φ and parameter β define the same model as the profile function $x \mapsto \varphi(rx)$ and parameter $r\beta$. Hence, the normalisation convention (2.2) represents no loss of generality. Similarly, if the intensity of the arrival process is taken as $\lambda > 0$, the process $(G_{t/\lambda})_{t>0}$ is the original process with the same profile function φ and parameter $\beta\lambda$.

(ii) The form of the connection probability (2.1) is natural for the following reasons: To ensure that the probability of a new vertex connecting to its nearest neighbour does not degenerate as $t \to \infty$, it is necessary to scale $d(x, y)$ by $t^{-1/d}$, which is the order of the distance of a point to its nearest neighbour at time t. Further, the integrability condition on φ ensures that the expected number of edges connecting a new vertex to the already existing ones remains bounded from zero and infinity, as $t \to \infty$.

(iii) In the degree-based spatial preferential attachment model of Jacob and Mörters [14], the term $(t/s)^\gamma$ that creates the age dependence in our model is replaced by a function of the indegree, the number of younger vertices **y** is connected to at time t. If this function is asymptotically linear with slope γ, the network is scale-free with power-law exponent $\tau = 1 + \frac{1}{\gamma}$. In this case, the expected indegree is of order $(t/s)^\gamma$ so that the models remain comparable and this is the natural choice to $(t/s)^\gamma$ that our network model will be scale-free.

(iv) For the profile function φ, one has different choices. We normally assume that φ is either regularly varying at infinity with index $-\delta$, for some $\delta > 1$, or φ decays quicker than any regularly varying function, in which case we set $\delta = \infty$. In the latter case, a natural choice is to consider $\varphi(x) = \frac{1}{2a}\mathbb{1}_{[0,a]}(x)$ 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 and only if their positions are within distance

$$\left(\frac{1}{t}\beta a \, (t/s)^\gamma\right)^{1/d}.$$

In the case $a = 1/2$, the profile function φ only takes the values zero and one; thus, the decision is not random and we connect two vertices whenever they are

close enough. The degree-based preferential attachment model in discrete time for this choice of φ was introduced in [1] and further studied in [5,16]. This particular choice for the profile function helps to get a better understanding of the problems and properties of this model; see, for example, Sect. 5. However, for our ultimate purpose this choice is too restrictive as it does not allow the networks to be robust or ultrasmall.

In the following sections, we use $g = o(h)$ to indicate that g/h converges to zero, $g \asymp h$ if g/h is bounded from zero and infinity and $g \sim h$ if g/h converges to one.

3 Weak local limit: the age-dependent random connection model

In this section, we introduce a graphical representation of the network G_t. This representation allows a simple rescaling, and the rescaled graphs turn out to converge to a limiting graph, which is denoted as the *age-dependent random connection model*. This also turns out to be the weak local limit of the graph sequence $(G_t)_{t>0}$, which enables us to achieve results for the network $(G_t)_{t\geq 0}$ by studying the age-dependent random connection model.

Let $\mathcal{X}$ denote a Poisson point process of unit intensity on $\mathbb{R}^d \times (0, \infty)$. We say a point $\mathbf{x} = (x, s) \in \mathcal{X}$ is born at time s and placed at position x. Observe that, almost surely, two points of $\mathcal{X}$ neither have the same birth time nor the same position. We say that (x, s) is *older* than (y, u) if $s < u$. For $t > 0$ write $\mathcal{X}_t$ for $\mathcal{X} \cap (\mathbb{T}_1^d \times (0, t])$, the set of vertices on the torus already born at time t. We denote by

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

the set of *potential edges* in $\mathcal{X}$. Given $\mathcal{X}$ we introduce a family $\mathcal{V}$ of independent random variables, uniformly distributed on $(0, 1)$, indexed by the set of potential edges. We denote these variables by $\mathcal{V}_{\mathbf{x},\mathbf{y}}$ or $\mathcal{V}(\mathbf{x}, \mathbf{y})$. A realisation of $\mathcal{X}_t$ and $\mathcal{V}_t$, defined as the restriction of $\mathcal{V}$ to indices in $\mathcal{X}_t \times \mathcal{X}_t$, defines a network $G(\mathcal{X}_t, \mathcal{V}_t)$ 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

$$\mathcal{V}(\mathbf{x}, \mathbf{y}) \leq \varphi \left(\frac{u \cdot d(x, y)^d}{\beta \left(\frac{u}{s} \right)^\gamma} \right). \tag{3.1}$$

Observe that the graph sequence $(G(\mathcal{X}_t, \mathcal{V}_t))_{t>0}$ has the law of our age-based spatial preferential attachment model and is therefore constructed on the probability space carrying the Poisson process $\mathcal{X}$ and the sequence $\mathcal{V}$. Moreover, G extends to a deterministic mapping associating a graph structure to any locally finite set of points in $\mathcal{Y} \subseteq \mathbb{T}_a^d \times (0, \infty)$ and sequence $\mathcal{V}$ in $(0, 1)$ indexed by $E(\mathcal{Y}) = \{(\mathbf{x}, \mathbf{y}) \in \mathcal{Y} \times \mathcal{Y} : \mathbf{x} \text{ younger than } \mathbf{y}\}$, where $\mathbb{T}_a^d = (-\frac{1}{2}a^{1/d}, \frac{1}{2}a^{1/d}]^d$ is the torus of volume a equipped with its canonical metric $d(\cdot, \cdot)$ and $\mathbf{x}, \mathbf{y}$ are connected if and only if (3.1) holds. We permit the case $a = \infty$, with $\mathbb{T}_\infty^d = \mathbb{R}^d$ equipped with the Euclidean metric.

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

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

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

$$\varphi\left(\frac{u/t \cdot d(t^{1/d}x, t^{1/d}y)^d}{\beta\left(\frac{u/t}{s/t}\right)^\gamma}\right) = \varphi\left(\frac{u \cdot d(x, y)^d}{\beta\left(\frac{u}{s}\right)^\gamma}\right),$$

the operation h_t preserves the rule (3.1) and therefore

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

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; see Fig. 1.
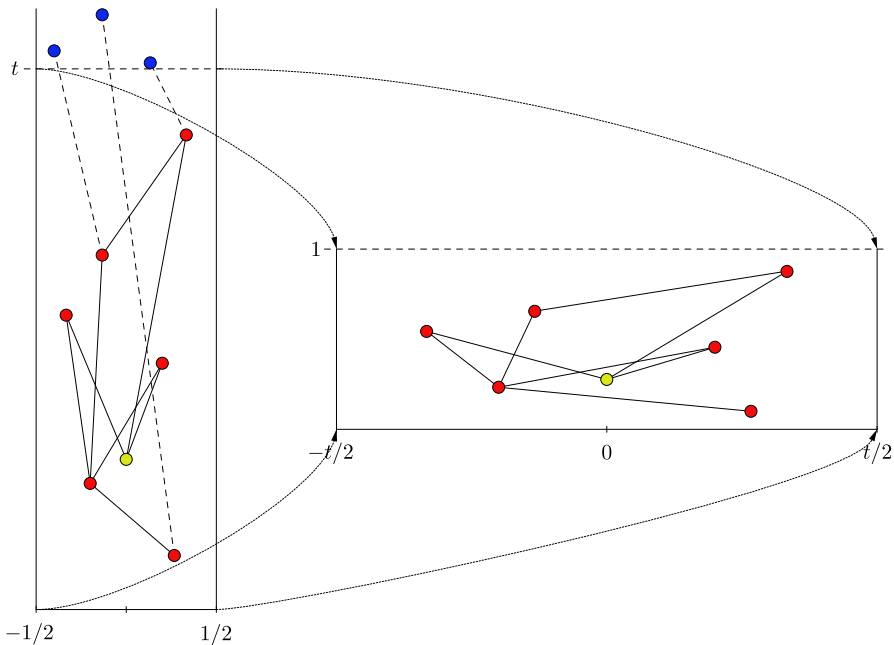

Fig. 1 The graph G_t on the left and its rescaling $h_t(G_t)$ on the right. The blue vertices are born after time t and, therefore, the corresponding edges do not exist yet and the vertices are not part of the rescaled graph. The yellow vertex is placed at position 0 and remains in the centre after the rescaling (Color figure online)

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

In order to formulate also a local version of this convergence result, we add a point at the origin to our Poisson process, denoting $\mathcal{X}_0 := \mathcal{X} \cup \{(0, U)\}$, where U is an independent birth time, uniformly distributed on $(0,]$. As before, let $\mathcal{V}_0$ be a family of independent uniformly distributed random variables indexed by the potential edges in $\mathcal{X}_0$, and, for $0 < t \le \infty$, let $\mathcal{X}_0^t = \mathcal{X}_0 \cap (\mathbb{T}_t^d \times (0, 1])$ and denote by $\mathcal{V}_0^t$ the restriction of $\mathcal{V}_0$ to indices in $\mathcal{X}_0^t \times \mathcal{X}_0^t$. We define rooted graphs $G_0^t := G(\mathcal{X}_0^t, \mathcal{V}_0^t)$ with the root being the vertex placed at the origin. For $p > 0$, define the class $\mathcal{H}_p$ of nonnegative functions H acting on locally finite rooted graphs and depending only on a bounded graph neighbourhood of the root with the property that

$$\sup_{0<t<\infty} \mathbb{E}[H(G_0^t)^p] < \infty.$$

Theorem 3.1 *(i)* G^∞ *is almost surely locally finite, i.e., almost surely all its vertices have finite degree.*
(ii) Almost surely, the graph sequence (G^t) converges to G^∞ in the sense that for each $\mathbf{x} \in \mathcal{X}^\infty$ the neighbours of $\mathbf{x}$ in G^t and in G^∞ coincide for large t.
(iii) In probability, the graph sequence (G_t) converges weakly locally to G_0^∞ in the sense that, for any $H \in \mathcal{H}_p$, $p > 1$, we have

$$\lim_{t\to\infty} \frac{1}{t} \sum_{\mathbf{x}\in G_t} H(\theta_\mathbf{x} G_t) = \mathbb{E}[H(G_0^\infty)] \qquad \text{in probability}, \qquad (3.2)$$

where $\theta_\mathbf{x}$ acts on points $\mathbf{y} = (y, s)$ as $\theta_\mathbf{x}(\mathbf{y}) = (y - x, s)$ and on graphs accordingly.

Theorem 3.1 will be proved in Section 4. The limiting graph G^∞ in *(ii)* is what we call the *age-dependent random connection model*. This model is of independent interest as a natural generalisation of the random connection model; see Meester and Roy [23] or Last et al. [20] for a recent paper. 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

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

The rooted graph G_0^∞ occurring as the local limit is the *Palm version* of the age-dependent random connection model G^∞; loosely speaking the graph G^∞ with a typical point shifted to the origin.

Remarks:

- Weak local limits were introduced by Benjamini and Schramm [3] as distributional limits for deterministic sequences of finite graphs randomised by a uniform choice of root. The result in (iii) allows that H additionally depends continuously on the ages of the vertices and the length of the edges if taken in the scaled graphs $h_t(\theta_{\mathbf{x}} G_t)$. Further generalisations of the results hold; see Yukich and Penrose [26] for seminal work on random geometric graphs and Jacob and Mörters [14] for a similar proof in the case of the degree-based model which can be adapted to our situation. We will not need these more general results here.
- The age-dependent random connection model is in a different universality class than other established models of infinite spatial scale-free graphs. For example, the scale-free percolation model of Deijfen et al. [6] and its continuous version studied by Deprez and Wüthrich [7] correspond (when rewritten in our framework) to a connection probability of the form

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

Models of this type do not arise naturally from sequences of growing finite random graphs on a fixed space as our model does.
- There is a similar convergence result for the degree-based spatial preferential attachment model, but the limiting graph is not as natural as the age-dependent random connection model as the existence of edges between vertices with given position and age depends in this graph on the existence of edges between the older vertex and other vertices that may lie arbitrarily far away; see Jacob and Mörters [14].

4 Convergence of neighbourhoods and degree distributions

In this section, we will study the asymptotic degree distribution and show that the age-based spatial preferential attachment model is scale-free. To this end, we study the neighbourhood of a fixed vertex $\mathbf{x} = (x, u)$ in the graphs G^t. We think of edges as oriented from the younger to the older endvertex, so that the *indegree* of $\mathbf{x}$ is the number of younger vertices that connect to it and the *outdegree* is the number of older vertices it connects to. As our construction is based on Poisson processes and conditionally independent edges, the indegree and outdegree of a fixed vertex are independent and Poisson distributed.

If G is a graph with vertices in $\mathbb{T}_t^d \times (0, \infty)$, we write $\mathbf{x} \leftrightarrow \mathbf{y}$ to indicate that there is an edge between $\mathbf{x}$ and $\mathbf{y}$ in G. Now, let $\mathbf{x} = (x, u)$ be a vertex in G and define its older neighbours,

$$\mathcal{Y}_{\mathbf{x}}(G) := \big\{ \mathbf{y} = (y, s) \in G : \mathbf{x} \leftrightarrow \mathbf{y}, s \le u \big\},$$

and its younger neighbours born before time s,

$$\mathcal{Z}_{\mathbf{x}}(s, G) := \left\{ \mathbf{y} = (y, r) \in G \colon \mathbf{y} \leftrightarrow \mathbf{x}, u < r \leq s \right\}.$$

For $t \in (0, \infty]$ and $0 < u < s \leq 1$, we write $\mathcal{Y}_{\mathbf{x}}^t := \mathcal{Y}_{\mathbf{x}}(G^t)$ and $\mathcal{Z}_{\mathbf{x}}^t(s) := \mathcal{Z}_{\mathbf{x}}(s, G^t)$, adding the point $\mathbf{x} = (x, u)$ to the underlying Poisson process $\mathcal{X}$ if it is not already there.

Proposition 4.1 *(a) For every $t \in (0, \infty]$, the older neighbours $\mathcal{Y}_{\mathbf{x}}^t$ of $\mathbf{x} = (x, u)$ form a Poisson point process on $\mathbb{T}_t^d \times [0, u)$ with intensity measure*

$$\lambda_{\mathcal{Y}_{\mathbf{x}}^t} := \varphi \left(\beta^{-1} u \left(\frac{s}{u} \right)^{\gamma} d(x, y)^d \right) \, \mathrm{d}y \, \mathrm{d}s.$$

(b) For every $t \in (0, \infty]$, the younger neighbours $\mathcal{Z}_{\mathbf{x}}^t(s_0)$ of $\mathbf{x} = (x, u)$ at time $s_0 \in (u, 1]$ form a Poisson point process on $\mathbb{T}_t^d \times (u, s_0]$ with intensity measure

$$\lambda_{\mathcal{Z}_{\mathbf{x}}^t(s_0)} := \varphi \left(\beta^{-1} s \left(\frac{u}{s} \right)^{\gamma} d(x, y)^d \right) \, \mathrm{d}y \, \mathrm{d}s.$$

(c) The outdegree of the origin in G_0^{∞} is Poisson distributed with parameter $\frac{\beta}{1-\gamma}$ and independent of the age U of the origin.

(d) The indegree of the origin in G_0^{∞} is mixed Poisson distributed, where the mixing distribution has the density

$$f(\lambda) = \beta^{1/\gamma} (\gamma \lambda + \beta)^{-(1+1/\gamma)} \qquad \text{for } \lambda > 0. \tag{4.1}$$

Proof The older neighbours of $\mathbf{x} = (x, u)$ are all neighbours with birth time smaller than u; therefore, $\mathcal{X} \cap (\mathbb{T}_t^d \times [0, u))$ is the set of all potential vertices connected to $\mathbf{x}$ by an outgoing edge. Now, given $\mathcal{X}$, a vertex $\mathbf{y} = (y, s) \in \mathcal{X} \cap (\mathbb{T}_t^d \times [0, u))$ is connected to $\mathbf{x}$ independently with probability $\varphi(\beta^{-1} u^{1-\gamma} s^{\gamma} d(x, y)^d)$. Thus, $\mathcal{Y}_{\mathbf{x}}^t$ defines a thinning of $\mathcal{X} \cap (\mathbb{T}_t^d \times [0, u))$ and (a) follows. The analogous argument for the vertices in $\mathcal{X} \cap (\mathbb{T}_t^d \times (u, s_0])$ proves (b).

Applying (a) to $\mathbf{x} = (0, u)$ and $t = \infty$ gives that the number of older neighbours is Poisson distributed with parameter

$$\lambda_{\mathcal{Y}_{\mathbf{x}}^{\infty}} \left(\mathbb{R}^d \times [0, u] \right) = \int_0^u \int_{\mathbb{R}^d} \varphi \left(\beta^{-1} u^{1-\gamma} s^{\gamma} |y|^d \right) \mathrm{d}y \, \mathrm{d}s$$

$$= \int_0^u \beta u^{\gamma-1} s^{-\gamma} \, \mathrm{d}s \int_{\mathbb{R}^d} \varphi \left(|y|^d \right) \mathrm{d}y = \frac{\beta}{1 - \gamma},$$

using the normalisation of φ. The claimed independence follows as the distribution does not depend on u, completing the proof of (c).

Applying (b) to $\mathbf{x} = (0, u)$ and $t = \infty$ gives that the number of younger neighbours up to time s is Poisson distributed with parameter

$$
\lambda_{\mathcal{Z}^{\infty}_{\mathbf{x}}(s)}\left(\mathbb{R}^d \times (u, s]\right) = \int_u^s \int_{\mathbb{R}^d} \varphi\left(\beta^{-1} v^{1-\gamma} u^{\gamma} |y|^d\right) \mathrm{d}y \, \mathrm{d}v
$$

$$
= \beta \int_u^s v^{\gamma-1} u^{-\gamma} \int_{\mathbb{R}^d} \varphi(|y|^d) \, \mathrm{d}y \, \mathrm{d}v
$$

$$
= \beta \int_u^s v^{\gamma-1} u^{-\gamma} \, \mathrm{d}v = \beta \frac{s^{\gamma} u^{-\gamma} - 1}{\gamma}.
$$

As U is independent of $\mathcal{X}$ and $\mathcal{V}$, the probability that the indegree equals k is therefore

$$
\int_0^1 \exp\left(-\beta \frac{u^{-\gamma} - 1}{\gamma}\right) \cdot \frac{\left(\beta \frac{u^{-\gamma} - 1}{\gamma}\right)^k}{k!} \, \mathrm{d}u
$$

$$
= \int_0^{\infty} \exp(-\lambda) \cdot \frac{\lambda^k}{k!} \cdot \left(\beta^{1/\gamma} (\gamma \lambda + \beta)^{-(1+1/\gamma)}\right) \mathrm{d}\lambda,
$$

as claimed (Fig. 2).   □

Remark Since, by construction, $\mathcal{Y}^t_{\mathbf{x}}$ and $\mathcal{Z}^t_{\mathbf{x}}(1)$ are independent Poisson point processes, the neighbourhood of a point $\mathbf{x} = (0, u)$ added (if necessary) to G^t is a Poisson point process with intensity $\lambda_{\mathcal{Z}^t_{\mathbf{x}}(1)} + \lambda_{\mathcal{Y}^t_{\mathbf{x}}}$. Now let t be finite and pick a vertex X uniformly at random from the finite graph $\tilde{G}_t$. We easily see that $h_t(\theta_{\mathbf{x}} G_t) = G^t_0$ in distribution. Hence, Proposition 4.1, parts (a) and (b), give a precise description of the neighbourhood of a randomly chosen vertex in G_t.

Proof of Theorem 3.1(i) By Proposition 4.1, parts (c) and (d), almost surely, the origin has finite degree in G^{∞}_0. Hence, by the refined Campbell theorem (see Theorem 9.1 in [21]), almost surely, every vertex in G^{∞} has finite degree.   □

Proof of Theorem 3.1(ii) We work conditionally on $\mathbf{x} = (x, s) \in \mathcal{X}^{\infty}$. Our aim is to show that there exists an almost surely finite random variable M such that, for all $t \in (0, \infty]$ and $\mathbf{y} \in \mathcal{X}^{\infty}$ with distance at least M from $\mathbf{x}$, the vertices $\mathbf{x}$ and $\mathbf{y}$ are not connected in G^t. To this end, observe that the distance between $\mathbf{x}$ and any $\mathbf{y} \in \mathbb{T}^d_t$ can be up to $2\sqrt{d} |x|$ smaller than it would be in $\mathbb{R}^d$. Consider the model where the vertices within distance $2\sqrt{d} |x|$ of $\mathbf{x}$ are deleted from $\mathcal{X}^{\infty}$ and all the other vertices are moved towards $\mathbf{x}$ by a distance of $2\sqrt{d} |x|$. It is easy to see that all vertices $\mathbf{y} \in \mathcal{X}^{\infty}$, which are at least $2\sqrt{d} |x|$ away from $\mathbf{x}$ and connected to $\mathbf{x}$ in the finite graph G^t for some $t > 0$, are also linked to $\mathbf{x}$ in this new model. Furthermore, the degree of $\mathbf{x}$ is still almost surely finite. Hence, we define the random variable M as the distance of $\mathbf{x}$ to the furthest vertex it is linked to in this new model, plus $2\sqrt{d} |x|$. Then M is almost surely finite and, as for $t > |x| + M$ the vertices in $\mathcal{X}^{\infty}$ and in $\mathcal{X}^t$ within distance M from $\mathbf{x}$ coincide, the edges of $\mathbf{x}$ linking it to another vertex $\mathbf{y}$ that is at most M away coincide in G^t and G^{∞} for sufficiently large t.   □
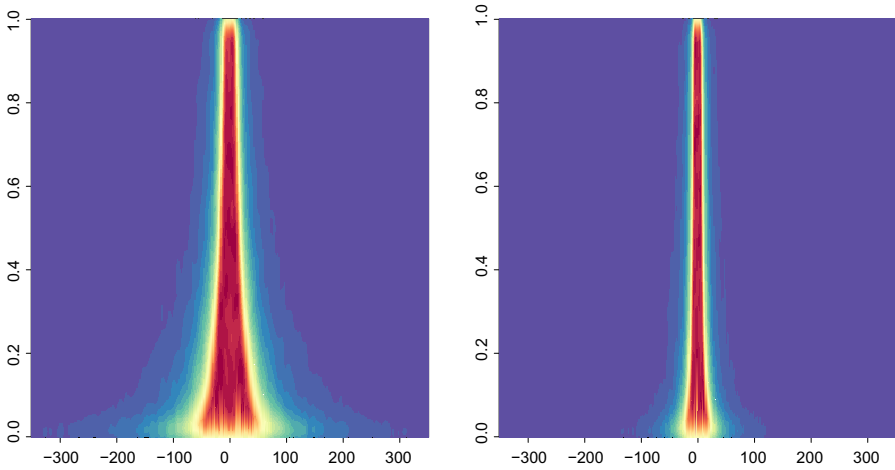
Fig. 2 Heatmaps of the neighbourhood of a relatively old root (left, birth time 0.2) and of a relatively young root (right, birth time 0.8) in G_0^∞ with $\beta = 5$, $\gamma = 1/3$ and $\varphi(x) = 1 \wedge x^{-2}$

Proof of Theorem 3.1(iii) We can replace the left-hand side in (3.2) by the limit of $\frac{1}{t} \sum_{\mathbf{x} \in G^t} H(\theta_{\mathbf{x}} G^t)$, which has the same distribution and, due to Campbell's formula, has expectation $\mathbb{E}[H(G_0^t)]$. Furthermore, the neighbourhoods of the origin in G_0^t and in G_0^∞ agree for sufficiently large t. As the family $(H(G_0^t))_{t>0}$ is bounded in L^p and therefore uniformly integrable, we infer that $\mathbb{E}[H(G_0^t)]$ converges to $\mathbb{E}[H(G_0^\infty)]$. Hence, the first moments in (3.2) converge, and we now argue that for bounded H the second moments converge, too.

Spelling out the second moment of $\frac{1}{t} \sum_{\mathbf{x} \in G^t} H(\theta_{\mathbf{x}} G^t)$, we get a term corresponding to choosing the same $\mathbf{x} \in G^t$ twice, which by the first moment calculation applied to H^2 converges to zero, and the term

$$\mathbb{E}\left[\frac{1}{t^2} \sum_{\substack{\mathbf{x},\mathbf{x}' \in G^t \\ \mathbf{x} \neq \mathbf{x}'}} H(\theta_{\mathbf{x}} G^t) H(\theta_{\mathbf{x}'} G^t)\right].$$

Using the boundedness of H, we can chose $\varepsilon > 0$ so that the contribution from pairs $\mathbf{x}$, $\mathbf{x}'$ for which one is born before time ε is arbitrarily small. We can then find a large radius R so that the graph neighbourhood of the origin on which H depends is contained in $\{\mathbf{y} : d(0, y) \leq R\}$ for $\theta_{\mathbf{x}} G^t$ for a proportion of vertices $\mathbf{x} \in G^t$ born after time ε arbitrarily close to one, for all sufficiently large t. We can neglect the small proportion of exceptional vertices as well as pairs $\mathbf{x}$, $\mathbf{x}'$ with distance smaller than R using again the boundedness of H. On the remaining part, the expectation factorises and we see that second moment converges to $\mathbb{E}[H(G_0^\infty)]^2$. Hence, we get convergence in L^2.

It remains to remove the condition of boundedness of H. Let $k \in \mathbb{N}$ and observe that our result applies to the bounded functional $H \wedge k$. Note that

$$\mathbb{E}\left[\frac{1}{t}\sum_{\mathbf{x}\in G^t} H(\theta_{\mathbf{x}}G^t) - H\wedge k(\theta_{\mathbf{x}}G^t)\right] = \mathbb{E}\left[H(G_0^t) - H\wedge k(G_0^t)\right]$$

and the right-hand side goes to zero uniformly in t as $k\to\infty$ by the uniform integrability implied in our L^p bound. This implies the required convergence. $\qquad\square$

We define the *empirical outdegree distribution* ν_t of the graph G_t by

$$\nu_t(k) = \frac{1}{t}\sum_{\mathbf{x}\in G_t} \mathbb{1}_{\{|\mathcal{Y}_{\mathbf{x}}(G_t)|=k\}} \qquad \text{for } k\in\mathbb{N},$$

and note that (for convenience) we have normalised ν_t so that its mass converges to one without necessarily being equal to one for small t. We now show that the empirical outdegree distribution ν_t converges to a deterministic limit.

Theorem 4.2 *For any function $g:\mathbb{N}_0\to[0,\infty)$ growing no faster than exponentially, we have*

$$\frac{1}{t}\sum_{\mathbf{x}\in G_t} g\big(|\mathcal{Y}_{\mathbf{x}}(G_t)|\big) = \int g\,\mathrm{d}\nu_t \longrightarrow \int g\,\mathrm{d}\nu$$

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

Proof For a finite graph G with vertices marked by birth times and a root vertex $\mathbf{x}$, we can define $H(G) = g(|\mathcal{Y}_{\mathbf{x}}(G)|)$, where $\mathcal{Y}_{\mathbf{x}}(G)$ is the set of edges from the root to older vertices in G. Note that the function H depends only on the neighbourhood of the root within graph distance one and the relative birth times of these vertices. Moreover, $H(G_0^t) = g(|\mathcal{Y}_{\mathbf{x}}(G_0^t)|)$, where $\mathbf{x}\in G_0^t$ is the vertex placed at the origin, for arbitrary t, and as $|\mathcal{Y}_{\mathbf{x}}(G_0^t)|$ is Poisson distributed with a bounded parameter, the integrability condition $H\in\mathcal{H}_p$ is satisfied as long as g is not growing faster than exponentially. As $H(\theta_{\mathbf{x}}G_t) = g(|\mathcal{Y}_{\mathbf{x}}(G_t)|)$ for all $\mathbf{x}\in\mathcal{X}_t$ and finite t, we infer the result from Theorem 3.1(iii). $\qquad\square$

Define the *empirical indegree distribution* μ_t of the graph G_t by

$$\mu_t(k) = \frac{1}{t}\sum_{\mathbf{x}\in G_t} \mathbb{1}_{\{|\mathcal{Z}_{\mathbf{x}}(t,G_t)|=k\}}.$$

Similar to above, the empirical indegree distribution μ_t also converges to a deterministic limit.

Theorem 4.3 *For any function $g:\mathbb{N}_0\to[0,\infty)$ growing no faster than linearly, we have*

$$\frac{1}{t}\sum_{\mathbf{x}\in G_t} g\big(|\mathcal{Z}_{\mathbf{x}}(t,G_t)|\big) = \int g\,\mathrm{d}\mu_t \longrightarrow \int g\,\mathrm{d}\mu$$

in probability, as $t \to \infty$, where μ is the mixed Poisson distribution with density f as in (4.1).

Proof For a finite graph G with vertices marked by birth times and a root vertex $\mathbf{x}$, we can define $H(G) = g(|\mathcal{Z}_\mathbf{x}(G)|)$, where $\mathcal{Z}_\mathbf{x}(G)$ is the set of edges from younger vertices in G to the root. Note that the function H depends only on the neighbourhood of the root within graph distance one and the relative birth times of these vertices. Moreover, $H(G_0^t) = g(|\mathcal{Z}_\mathbf{x}(G_0^t)|)$, where $\mathbf{x} \in G_0^t$ is the vertex placed at the origin, for arbitrary t. Now, $|\mathcal{Z}_\mathbf{x}(G_0^t)|$ is dominated by $|\mathcal{Z}_\mathbf{x}(G_0^\infty)|$ whose distribution μ has tails (calculated in Lemma 4.4) that vanish fast enough to ensure that $H \in \mathcal{H}_p$ for some $p > 1$. As $H(\theta_\mathbf{x} G_t) = g(|\mathcal{Z}_\mathbf{x}(G_t)|)$ for all $\mathbf{x} \in \mathcal{X}_t$ and finite t, we infer the result from Theorem 3.1(iii). □

To complete the proof that the age-based preferential attachment model is scale-free with power-law exponent $\tau = 1 + 1/\gamma$, we observe that, by a similar argument as in Theorems 4.2 and 4.3, the empirical degree distribution in G_t converges in probability to the convolution of ν and μ. As ν has superexponentially light tails, the tail behaviour of the convolution is inherited from that of μ, which we now calculate.

Lemma 4.4 $\mu(k) = k^{-(1+\frac{1}{\gamma})+o(1)}$ *as $k \uparrow \infty$.*

Proof Observe that

$$\mu(k) = \beta^{1/\gamma} \int_0^\infty \frac{\lambda^k}{k!} e^{-\lambda} (\gamma\lambda + \beta)^{-(1+\frac{1}{\gamma})} \, d\lambda \leq \frac{\beta^{1/\gamma} \gamma^{-1-\frac{1}{\gamma}}}{\Gamma(k+1)} \int_0^\infty \lambda^{(k-\frac{1}{\gamma})-1} e^{-\lambda} \, d\lambda$$

$$= \frac{\beta^{1/\gamma}}{\gamma^{1+1/\gamma}} \frac{\Gamma(k-\frac{1}{\gamma})}{\Gamma(k+1)} = k^{-1-\frac{1}{\gamma}+o(1)},$$

as $k \uparrow \infty$, by Stirling's formula. On the other hand, note that for some fixed bound $A > 0$, there exists a constant $c > 0$ such that $\gamma x + \beta \leq c\gamma x$ for all $x \geq A$. Hence,

$$\mu(k) \geq \frac{c^{-1-\frac{1}{\gamma}} \beta^{1/\gamma}}{\Gamma(k+1)} \int_A^\infty \lambda^k e^{-\lambda} (\gamma\lambda)^{-1-\frac{1}{\gamma}} \, d\lambda$$

$$= \tilde{c} \frac{\Gamma(k-\frac{1}{\gamma})}{\Gamma(k+1)} - \frac{\tilde{c}}{\Gamma(k+1)} \int_0^A \lambda^{(k-\frac{1}{\gamma})-1} e^{-\lambda} \, d\lambda,$$

for some positive constant $\tilde{c}$. As the subtracted term, for fixed A, is of smaller order as $k \to \infty$, we obtain the lower bound. □

5 Global and local clustering coefficients

To show that the age-based spatial preferential attachment model has clustering features, we use two metrics well established in the applied networks literature; see, for example, [25,27] for some early papers. If G is a finite graph, we call a pair of edges

in G a *wedge* if they share an endpoint (called its *tip*). Define the *global clustering coefficient* or *transitivity* of G as

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

if there is at least one wedge in G, and $c^{\text{glob}}(G) := 0$ otherwise. By definition, $c^{\text{glob}}(G) \in [0, 1]$.

Another way of thinking about clusters is locally, i.e., to count only the triangles and wedges containing a fixed vertex $\mathbf{x}$. For a vertex $\mathbf{x}$ with at least two neighbours, define the *local clustering coefficient* by

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

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

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

if $V_2(G)$ is not empty, and as $c^{\text{av}}(G) := 0$ otherwise. Note that this metric places more weight on the low-degree nodes, while the transitivity places more weight on the high-degree nodes.

Theorem 5.1 (Clustering coefficients)

(a) *For the average clustering coefficient, we have*

$$c^{\text{av}}(G_t) \longrightarrow \int_0^1 \mathbb{P}\big\{(X_u^{(1)}, S_u^{(1)}) \leftrightarrow (X_u^{(2)}, S_u^{(2)})\big\} \, \pi(\mathrm{d}u)$$

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

$$\frac{1}{\lambda_u}\Big(\varphi(\beta^{-1}s^{1-\gamma}u^\gamma |x|^d)\mathbb{1}_{(u,1]}(s) + \varphi(\beta^{-1}u^{1-\gamma}s^\gamma |x|^d)\mathbb{1}_{[0,u]}(s)\Big) \, \mathrm{d}x \, \mathrm{d}s, \quad (5.1)$$

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 $1 - e^{-\lambda_u} - \lambda_u e^{-\lambda_u}$.

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

$$c^{\text{glob}}(G_t) \longrightarrow c_\infty^{\text{glob}}$$

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

Remark The limiting average clustering coefficient can be interpreted as the probability that in G_0^∞ two neighbours of the vertex at the origin are connected by an edge. The density of the birth time of the vertex at the origin here is not uniform but given by the measure π, which is the conditional distribution of the birth time of a vertex given that it has degree at least two. Observe that this coefficient is always positive. By contrast, the global clustering coefficient vanishes asymptotically when preferential attachment to old nodes is strong (i.e., when γ is large). In this case, the collection of wedges is dominated by those with an untypically old tip. These vertices have small local clustering as they are endvertices to a significant amount of long edges.

Proof Let G be a finite rooted graph and define the function $H(G) = c_{\mathbf{x}}^{\mathrm{loc}}(G)$ if the root $\mathbf{x}$ has degree at least two, and $H(G) = 0$ otherwise. As H is bounded, we have $H \in \mathcal{H}_p$ for any $p > 1$ and, by Theorem 3.1 (iii), we get

$$\frac{1}{t} \sum_{\mathbf{x} \in G_t} H(\theta_{\mathbf{x}} G_t) \longrightarrow \mathbb{E}\left[H(G_0^\infty)\right]$$

in probability, as $t \to \infty$. To calculate the limit, observe that, for a vertex $\mathbf{x}$ with degree k, the number of wedges with tip $\mathbf{x}$ is $k(k-1)/2$. It follows that

$$\mathbb{E}\left[H(G_0^\infty)\right]$$
$$= \int_0^1 \sum_{k \geq 2} \mathbb{E}\left[\frac{2}{k(k-1)} \sum_{(x,s) \leftrightarrow (0,u)} \sum_{\substack{(y,v) \leftrightarrow (0,u) \\ v < s}} \mathbb{1}_{\{(x,s) \leftrightarrow (y,v)\}} \mathbb{1}_{\{|\mathcal{Y}_{(0,u)}^\infty| + |\mathcal{Z}_{(0,u)}^\infty(1)| = k\}}\right] du.$$

By Proposition 4.1, the neighbourhood of the root $(0, u)$ is given by a Poisson point process with intensity measure

$$\lambda_{\mathcal{Z}_{(0,u)}^\infty}(1) + \lambda_{\mathcal{Y}_{(0,u)}^\infty}.$$

Conditioned on the number of neighbours, the neighbours of the root $(0, u)$ are independent and identically distributed as the normalised intensity measure of the neighbourhood given in (5.1); see [21, Proposition 3.8]. Therefore,

$$\mathbb{E}\left[H(G_0^\infty)\right] = \int_0^1 \mathbb{P}\left\{(X_u^{(1)}, S_u^{(1)}) \leftrightarrow (X_u^{(2)}, S_u^{(2)})\right\} \mathbb{P}\left\{|\mathcal{Y}_{(0,u)}^\infty| + |\mathcal{Z}_{(0,u)}^\infty(1)| \geq 2\right\} du,$$

where $(X_u^{(1)}, S_u^{(1)})$ and $(X_u^{(2)}, S_u^{(2)})$ are independent and identically distributed as claimed. Choosing $H(G)$ as the indicator of the event that the root has degree at least two, Theorem 3.1 (iii) gives

$$\frac{|V_2(G_t)|}{t} \longrightarrow \int_0^1 \mathbb{P}\left\{|\mathcal{Y}_{(0,u)}^\infty| + |\mathcal{Z}_{(0,u)}^\infty(1)| \geq 2\right\} du,$$

in probability. As $|\mathcal{Y}_{(0,u)}^{\infty}| + |\mathcal{Z}_{(0,u)}^{\infty}(1)|$ is Poisson distributed with intensity λ_u, we conclude that

$$c^{\mathrm{av}}(G_t) \longrightarrow \frac{\int_0^1 \mathbb{P}\big\{(X_u^{(1)}, S_u^{(1)}) \leftrightarrow (X_u^{(2)}, S_u^{(2)})\big\} \left(1 - e^{-\lambda_u} - \lambda_u e^{-\lambda_u}\right) \mathrm{d}u}{\int_0^1 1 - e^{-\lambda_u} - \lambda_u e^{-\lambda_u} \, \mathrm{d}u},$$

as claimed in part (a).

For the global clustering coefficient, we count the number of triangles and wedges separately. To this end, define $H(G)$ to be the number of triangles which have their youngest vertex in the root of G, and $\hat{H}(G)$ to be the number of wedges with tip in the root $\mathbf{x}$ of G. Note that $H(G_0^t) \le |\mathcal{Y}_{\mathbf{x}}(G_0^{\infty})|^2$ and thus $H \in \mathcal{H}_p$ for any $p > 1$. Moreover,

$$\hat{H}(G_0^t) = \frac{1}{2}|\mathcal{Y}_{\mathbf{x}}^t|(|\mathcal{Y}_{\mathbf{x}}^t| - 1) + \frac{1}{2}|\mathcal{Z}_{\mathbf{x}}^t(1)|(|\mathcal{Z}_{\mathbf{x}}^t(1)| - 1)$$
$$+ |\mathcal{Y}_{\mathbf{x}}^t||\mathcal{Z}_{\mathbf{x}}^t(1)| \le 2\big(|\mathcal{Y}_{\mathbf{x}}^{\infty}|^2 + |\mathcal{Z}_{\mathbf{x}}^{\infty}(1)|^2\big).$$

If $\gamma < 1/2$ and $1 < p < 1/(2\gamma)$, we hence have $\hat{H} \in \mathcal{H}_p$ and Theorem 3.1(iii) gives that

$$c^{\mathrm{glob}}(G_t) = \frac{\sum_{\mathbf{x} \in G_t} H(\theta_{\mathbf{x}} G_t)}{t} \cdot \frac{t}{\sum_{\mathbf{x} \in G_t} \hat{H}(\theta_{\mathbf{x}} G_t)} \longrightarrow \frac{\mathbb{E}[H(G_0^{\infty})]}{\mathbb{E}[\hat{H}(G_0^{\infty})]} > 0$$

in probability. If $\gamma > 1/2$, applying the theorem to the bounded functions $\hat{H}(G_t) \wedge k$ and then sending k to ∞, we get $\frac{1}{t} \sum_{\mathbf{x} \in G_t} \hat{H}(\theta_{\mathbf{x}} G_t) \to \infty$ and hence $c^{\mathrm{glob}}(G_t) \to 0$ in probability, as $t \to \infty$. $\square$

The local and average clustering coefficients cannot be calculated explicitly, but can be simulated; see Appendix A of this paper for a discussion of the simulation techniques used here. We focus on the profile functions $\varphi = \frac{1}{2a}\mathbb{1}_{[0,a]}$, for $a \ge 1/2$, dimension $d = 1$ and fixed edge density $\beta/(1-\gamma)$. Figure 3 shows the local clustering coefficient of a vertex of age u in G^{∞} showing monotone dependence on the age, i.e., the empirical probability that two neighbours of a given vertex are connected to each other is larger for younger vertices. This coincides with our intuitive understanding of the local structure of the networks, in which a young vertex, typically, is connected to either very close or very old vertices such that two randomly chosen neighbours have a decent chance of being connected to each other as well. By contrast, an old vertex typically has more long edges to younger vertices. Thus, two of its neighbours are typically further apart, which reduces the chance of them being each other's neighbour. This monotonicity occurs independently of the choice of β, γ and a.

In Fig. 4, we see that the dependence of the average clustering coefficient with respect to the width a of the profile function is of order $\frac{1}{a}$, a scaling that we also see in the analysis of the global clustering coefficient in the case $\gamma < \frac{1}{2}$. Hence, the average clustering coefficient and the global clustering coefficient (if $\gamma < \frac{1}{2}$) can be varied by the choice of φ and can be made arbitrarily small by choosing a
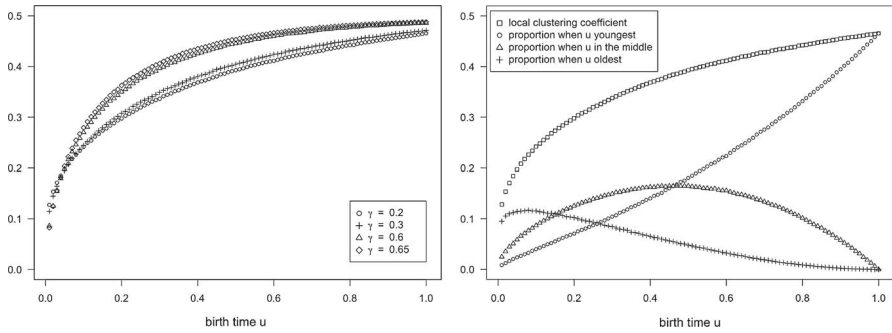
Fig. 3 Local clustering coefficient of a vertex $(0, u)$ for parameters $a = 1$ and $\beta = c_{ed}(1 - \gamma)$ chosen such that the asymptotic edge density is fixed at c_{ed}. The plot on the left displays the behaviour of the model for high edge density ($c_{ed} = 10$) for various values of γ. We remark that the shown behaviour is qualitatively independent of the edge density. In the plot on the right, the clustering coefficient for $\gamma = 0.2$ is shown, along with the probabilities of the event that u is younger (resp. in the middle or older) than two randomly picked neighbours, which are connected
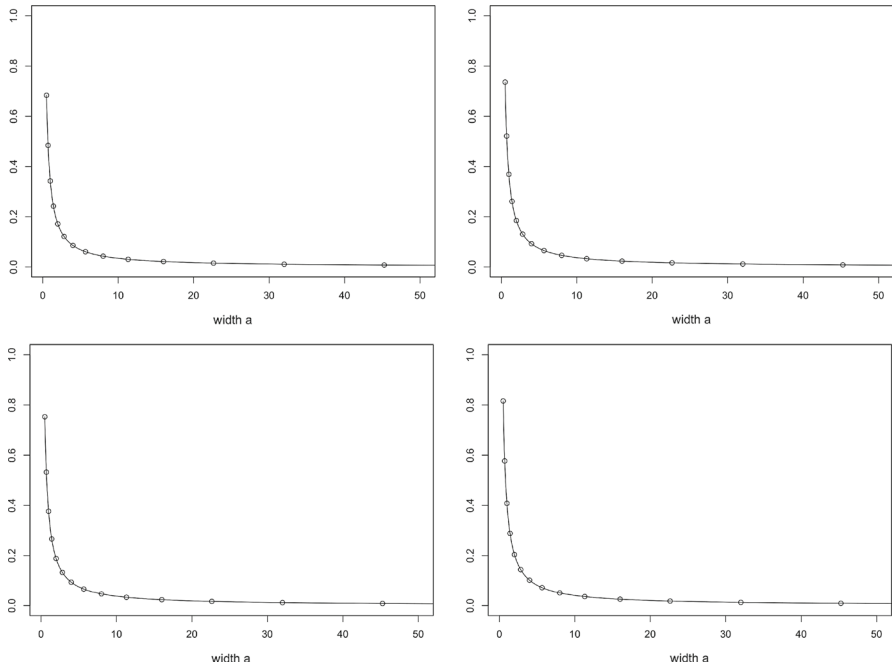

Fig. 4 Average clustering coefficient for the network with profile function $\varphi = \frac{1}{2a} \mathbb{1}_{[0,a]}$ plotted against the width a, for $\gamma = 0.3$ in the left resp. $\gamma = 0.6$ in the right graphs. The graphs in the top row correspond to fixed edge density 1, while the bottom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

large. Unlike with the global clustering coefficient, there is a mild dependence on β. Again, roughly speaking, a large width of φ encourages long edges and reduces clustering.

6 Asymptotics for typical edge lengths

In this section, we study the distribution of the length of typical edges in G_t. We denote by $E(G)$ the set of edges of the graph G and define λ_t, the (rescaled) *empirical edge length distribution* in G_t, by

$$\lambda_t = \frac{1}{|E(G_t)|} \sum_{(\mathbf{x},\mathbf{y}) \in E(G_t)} \delta_{t^{1/d} d(x,y)}.$$

Theorem 6.1 *For every continuous and bounded* $g \colon [0, \infty) \to \mathbb{R}$, *we have*

$$\frac{1}{|E(G_t)|} \sum_{(\mathbf{x},\mathbf{y}) \in E(G_t)} g\big(t^{1/d} d(x, y)\big) = \int g \, d\lambda_t \to \int g \, d\lambda$$

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

$$\lambda([a, b)) = \frac{1 - \gamma}{\beta} \int_0^1 \int_0^u \int_{a \le |y| < b} \varphi\big(\beta^{-1} u^{1-\gamma} s^\gamma |y|^d\big) \, dy \, ds \, du. \qquad (6.1)$$

Proof For a finite graph G with vertices positioned in $\mathbb{R}^d$ and marked by birth times and with a root vertex $\mathbf{x}$ placed at the origin, define, for $a < b \in [0, \infty]$, the function

$$H_{a,b}(G) = \sum_{\mathbf{y} \in \mathcal{Y}_{\mathbf{x}}(G)} \mathbb{1}_{[a,b)}(|y|). \qquad (6.2)$$

Observe that the law of $\lambda_t([a, b))$ in G_t equals the law of

$$\frac{1}{|E(G^t)|} \sum_{\mathbf{x} \in \mathcal{X}^t} H_{a,b}(\theta_{\mathbf{x}} G^t).$$

As mentioned in the remark following the theorem, Theorem 3.1 is applicable to functions $H_{a,b}$ depending on the length of edges in the rescaled graphs $(G^t)_{t>0}$. Since the sum in (6.2) is dominated by the outdegree, $H_{a,b} \in \mathcal{H}_p$ for some $p > 1$. We thus get

$$\frac{1}{t} \sum_{\mathbf{x} \in \mathcal{X}^t} H_{a,b}(\theta_{\mathbf{x}} G^t) \longrightarrow \mathbb{E}[H_{a,b}(G_0^\infty)],$$

and since Theorem 3.1 (iii) also gives $|E(G^t)|/t \to \frac{\beta}{1-\gamma}$ and $\lambda([a, b)) = \frac{1-\gamma}{\beta} \mathbb{E}[H_{a,b}(G_0^\infty)]$ we infer that $\lambda_t([a, \infty)) \longrightarrow \lambda([a, \infty))$ in probability, as $t \to \infty$. Therefore, convergence in probability of λ_t to λ in the space of probability measures on $\mathbb{R}_+$, equipped with the Lévy–Prokhorov metric follows. $\qquad \square$

Remark Suppose there exists $\delta > 1$ such that the profile function satisfies $\varphi(x^d) \asymp 1 \wedge x^{-d\delta}$. Then, the explicit formula for λ in (6.1) can be used to calculate the tail behaviour of λ. More precisely, separating the integration into several domains, depending on whether we are integrating over the tail domain of φ or not, results in the terms of order d, $d(\frac{1}{\gamma} - 1)$ and $d(\delta - 1)$. This gives that $\lambda([K, \infty)) \asymp 1 \wedge (\beta^{-1/d} K)^{-\eta}$, where

$$\eta := \min \left\{ d, d(\frac{1}{\gamma} - 1), d(\delta - 1) \right\}. \tag{6.3}$$

In particular, λ has finite expectation if $\eta > 1$ and infinite expectation if $\eta < 1$.

We denote by M_0^∞ the length of the longest outgoing edge of the origin in G_0^∞. By the construction of λ above, $\lambda([K, \infty))$ is the expected number of outgoing edges of length bigger than K divided by the total number of outgoing edges from the origin. If K is large, this should be of similar order to the probability that $M_0^\infty \geq K$. This is confirmed in the following lemma.

Lemma 6.2 *Suppose there exists $\delta > 1$ such that the profile function satisfies $\varphi(x^d) \asymp 1 \wedge x^{-d\delta}$. Then, $\mathbb{E}\left[(M_0^\infty)^a\right]$ is finite if $a < \eta$ and infinite if $a > \eta$, where η is as defined in (6.3).*

Proof We show that the tail probability $\mathbb{P}\left\{(M_0^\infty)^a \geq K\right\}$ is of order $K^{-\eta/a}$ as $K \to \infty$. The number of outgoing edges with length at least $K^{1/a}$ in G_0^∞ from the vertex $(0, u)$ at the origin is Poisson distributed with parameter

$$\lambda_{K^{1/a}, u} := \lambda \mathcal{Y}^\infty_{(0,u)}\left(\mathbb{R}^d \setminus (\{|x| < K^{1/a}\}) \times (0, u]\right),$$

and hence

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

recalling the asymptotic edge length distribution λ defined in (6.1). The established tail behaviour of the measure λ yields $\mathbb{P}\left\{(M_0^\infty)^a \geq K\right\} \asymp 1 \wedge K^{-\eta/a}$. □

Using this, we can establish a result about the average rescaled length in the network G_t.

Theorem 6.3 *Suppose that there exists $\delta > 1$ such that the profile function satisfies $\varphi(x^d) \asymp 1 \wedge x^{-d\delta}$. Then, for all $a > 0$ and $b \in [0, \frac{\eta}{a})$, there exists a positive constant C, depending on $a, b, \gamma, \beta, \varphi$, such that*

$$\frac{1}{|E(G_t)|} \sum_{\mathbf{x} \in G_t} \left(\sum_{\substack{\mathbf{y} \in G_t \\ \mathbf{x} \leftrightarrow \mathbf{y}}} \left(t^{1/d} d(\mathbf{x}, \mathbf{y})\right)^a \right)^b \to C \tag{6.4}$$

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

Remark If $\eta > 1$, one can choose $a = b = 1$, and this yields that the mean edge length in G_t is of order $t^{-1/d}$. If $\eta < 1$ (and in particular always if $d = 1$), the mean edge length is of larger order.

Proof Consider again a finite graph G with vertices positioned in $\mathbb{R}^d$ and marked by birth times and with a root vertex $\mathbf{x}$ placed at the origin. Define

$$H(G) := \Big(\sum_{\mathbf{y} \in \mathcal{Y}_\mathbf{x}(G)} |y|^a \Big)^b$$

and observe that the law of the left-hand side in (6.4) equals the law of

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

It suffices to show that $H \in \mathcal{H}_p$ for some $p > 1$, since Theorem 3.1 (iii) then ensures the convergence in probability to $\frac{1-\gamma}{\beta} \mathbb{E}[H(G_0^\infty)]$, which is a positive constant. To this end, recall M_0^∞, the length of the longest outgoing edge of the root $\mathbf{x}$ in G_0^∞, and observe that, almost surely, $H(G_0^t) \leq (M_0^\infty)^{ab} |\mathcal{Y}_\mathbf{x}^\infty|^b$. Since, by choice, $ab < \eta$, there exist some $p, q > 1$ such that $\alpha := pqab < \eta$. Lemma 6.2 then ensures $\mathbb{E}[(M_0^\infty)^\alpha] < \infty$ and, by applying Hölder's inequality to the observed bound for $H(G_0^t)$, we get

$$\sup_{t>0} \mathbb{E}\big[H(G_0^t)^p\big] \leq \big(\mathbb{E}\big[(M_0^\infty)^\alpha\big]\big)^{1/q} \Big(\mathbb{E}\Big[|\mathcal{Y}_\mathbf{x}^\infty|^{\frac{\alpha}{a(q-1)}}\Big]\Big)^{\frac{q-1}{q}} < \infty.$$

$\square$

7 Conclusion and outlook

We have seen that properties of real networks, like scale-free degree distributions and clustering, emerge from the simple building principle of preferential attachment to old and near nodes. Our model simplifies the spatial preferential attachment model in the literature, and this allows for more explicit calculations of asymptotic network metrics. Moreover, we are therefore confident that we can also cover more complicated features that have proved elusive in the full, degree-based, model. In particular, for the small-world property, robustness and vulnerability of the age-based spatial preferential attachment model only partial results have been possible [13,15]. A full study of these problems has been initiated in our group, and we hope to be able to report on new results soon.

Mathematically, our research is a step in the important direction of developing methods for networks that, due to clustering, cannot be locally approximated by trees. We have seen that in our case there is still a valuable description of a local limit given in terms of a tractable graph, the age-dependent random connection model. This model is interesting in its own right and methods from the theory of random geometric graphs

as well as techniques to investigate long-range percolation models can presumably be developed and enhanced to provide a powerful toolbox for its investigation.

The results we have achieved (and hope to achieve soon) are to some extent universal, i.e., other network models that follow our building principles should show qualitatively very similar behaviour. But these results do not necessarily explain the full picture, other building principles could lead to similar behaviour and add to the explanation of the abundance of networks with the mentioned features that describe complex systems. A particularly interesting way to generate networks based on universal principles is to define random graphs that try to optimise certain functionals, for example in the form of Gibbs measures on graphs with Hamiltonians that reward connectivity and punish long edges; see, for example, recent work of Mourrat and Valesin [24]. There remain a lot of interesting challenges for probabilists in the area of random networks.

Appendix A: Simulation of the model

In this section, we give an overview of the code used to generate the pictures shown throughout the paper. It is also used for estimating the limiting average clustering coefficient in Sect. 5. The code can be freely accessed at: http://www.mi.uni-koeln.de/~moerters/LoadPapers/adrc-model.R.

The main objective of the code is to sample neighbours of a given vertex (x, u) in the age-dependent random connection model in dimension 1 for given parameters β and γ and the profile function φ. Due to Proposition 4.1, which gives an explicit description of the neighbourhood of a given vertex, we can use rejection sampling to achieve this. The distribution in (5.1), defined on $\mathbb{R} \times (0, 1]$, that we use to sample the neighbours of (x, u) may be unbounded and heavy tailed in the first parameter. To deal with this, we restrict the sampling to a region with mass $q = 0.99$ with respect to this distribution. This sampling works for arbitrary but reasonable choices of the profile function φ and parameters β, γ; we provide and use an optimised sampling algorithm for $\varphi = \frac{1}{2a} \mathbb{1}_{[0,a]}$ with $a \geq \frac{1}{2}$. The advantage of studying this class of φ is that expressions can be analytically simplified, which allows us to improve the algorithm by dividing the region from which the points are sampled into sub-regions with equal mass with respect to φ, thus increasing the acceptance rate for points sampled far away from (x, u). That is, the code first selects one of these equally likely sub-regions uniformly at random and then points are sampled therein until one is accepted. The numerical optimisation method `nlminb` is used to calculate the boundaries of the ranges, i.e., quantiles of the distribution from (5.1).

A first application of the sampling is the estimation of the expected local clustering coefficient of a vertex $(0, u)$ in the age-dependent random connection model (see Fig. 3) and by Theorem 5.1 also the average clustering coefficient for the age-based preferential attachment network (see Fig. 4). To this end, the code samples pairs of

neighbours of $(0, u)$ and averages the probability that the pair is connected. A second application of the sampling is generating heatmaps of the neighbourhoods of a given vertex (see Fig. 2). The heatmaps are generated using the R library `MASS` and function `kde2d` by estimating the heat kernel for the sampled neighbouring vertices. Further properties thereof can be studied with additional heatmap generating functions that we provide.

References

1. Aiello, W., Bonato, A., Cooper, C., Janssen, J., Prałat, P.: A spatial web graph model with local influence regions. Int. Math. **5**(1–2), 175–196 (2008)
2. Barabási, A.-L., Albert, R.: Emergence of scaling in random networks. Science **286**(5439), 509–512 (1999)
3. Benjamini, I., Schramm, O.: Recurrence of distributional limits of finite planar graphs. Electron. J. Probab. 6:13 pp. (2001)
4. Bollobás, B., Riordan, O., Spencer, J., Tusnády, G.: The degree sequence of a scale-free random graph process. Random Struct. Algorithms **18**(3), 279–290 (2001)
5. Cooper, C., Frieze, A., Prałat, P.: Some typical properties of the spatial preferred attachment model. In: Algorithms and Models for the Web Graph, Volume 7323 of Lecture Notes in Computer Science, pp. 29–40. Springer, Heidelberg (2012)
6. Deijfen, M., van der Hofstad, R., Hooghiemstra, G.: Scale-free percolation. Ann. Inst. Henri Poincaré Probab. Stat. **49**(3), 817–838 (2013)
7. Deprez, P., Wüthrich, M.V.: Scale-free percolation in continuum space. Commun. Math. Stat. (2018). https://doi.org/10.1007/s40304-018-0142-0
8. Dereich, S., Mönch, C., Mörters, P.: Typical distances in ultrasmall random networks. Adv. Appl. Probab. **44**(2), 583–601 (2012)
9. Dereich, S., Mörters, P.: Random networks with sublinear preferential attachment: the giant component. Ann. Probab. **41**(1), 329–384 (2013)
10. Dommers, S., van der Hofstad, R., Hooghiemstra, G.: Diameters in preferential attachment models. J. Stat. Phys. **139**(1), 72–107 (2010)
11. Flaxman, A.D., Frieze, A.M., Vera, J.: A geometric preferential attachment model of networks. Int. Math. **3**(2), 187–205 (2006)
12. Flaxman, A.D., Frieze, A.M., Vera, J.: A geometric preferential attachment model of networks. II. In: Algorithms and Models for the Web-Graph, Volume 4863 of Lecture Notes in Computer Science, pp. 41–55. Springer, Berlin (2007)
13. Hirsch, C., Mönch, C.: Distances and large deviations in the spatial preferential attachment model. ArXiv e-prints (2018)
14. Jacob, E., Mörters, P.: Spatial preferential attachment networks: power laws and clustering coefficients. Ann. Appl. Probab. **25**(2), 632–662 (2015)
15. Jacob, E., Mörters, P.: Robustness of scale-free spatial networks. Ann. Probab. **45**(3), 1680–1722 (2017)
16. Janssen, J., Prałat, P., Wilson, R.: Geometric graph properties of the spatial preferred attachment model. Adv. Appl. Math. **50**(2), 243–267 (2013)
17. Jordan, J.: Degree sequences of geometric preferential attachment graphs. Adv. Appl. Probab. **42**(2), 319–330 (2010)
18. Jordan, J.: Geometric preferential attachment in non-uniform metric spaces. Electron. J. Probab. **18**(8), 15 (2013)
19. Jordan, J., Wade, A.R.: Phase transitions for random geometric preferential attachment graphs. Adv. Appl. Probab. **47**(2), 565–588 (2015)
20. Last, G., Nestmann, F., Schulte, M.: The random connection model and functions of edge-marked Poisson processes: second order properties and normal approximation. ArXiv e-prints (2018)
21. Last, G., Penrose, M.: Lectures on the Poisson Process. Cambridge University Press, Cambridge (2017)
22. Manna, S.S., Sen, P.: Modulated scale-free network in Euclidean space. Phys. Rev. E **66**, 066114 (2002)
23. Meester, R., Roy, R.: Continuum Percolation. Cambridge University Press, Cambridge (1996)
24. Mourrat, J.-C., Valesin, D.: Spatial Gibbs random graphs. Ann. Appl. Probab. **28**(2), 751–789 (2018)

25. Newman, M.E.J., Strogatz, S.H., Watts, D.J.: Random graphs with arbitrary degree distributions and their applications. Phys. Rev. E **64**, 026118 (2001)
26. Penrose, M.D., Yukich, J.E.: Weak laws of large numbers in geometric probability. Ann. Appl. Probab. **13**(1), 277–303 (2003)
27. Watts, D.J., Strogatz, S.H.: Collective dynamics of 'small-world' networks. Nature **393**, 440–442 (1998)