*J. Appl. Prob.* **56,** 672–700 (2019)
*doi:*10.1017/jpr.2019.45
© *Applied Probability Trust* 2019

# DEGREE CORRELATIONS IN SCALE-FREE RANDOM GRAPH MODELS

CLARA STEGEHUIS,* *Eindhoven University of Technology*

**Abstract**

We study the average nearest-neighbour degree $a(k)$ of vertices with degree $k$. In many real-world networks with power-law degree distribution, $a(k)$ falls off with $k$, a property ascribed to the constraint that any two vertices are connected by at most one edge. We show that $a(k)$ indeed decays with $k$ in three simple random graph models with power-law degrees: the erased configuration model, the rank-1 inhomogeneous random graph, and the hyperbolic random graph. We find that in the large-network limit for all three null models, $a(k)$ starts to decay beyond $n^{(\tau-2)/(\tau-1)}$ and then settles on a power law $a(k) \sim k^{\tau-3}$, with $\tau$ the degree exponent.

*Keywords:* Degree correlations; random graph

2010 Mathematics Subject Classification: Primary 05C80
Secondary 60F05

## 1. Introduction

Complex networks are often studied via mathematical analysis of null models that can match the network degree distribution. For scale-free networks, this degree distribution follows a power law. In many real-world networks, such as the Internet, social networks, and biological networks, the power-law exponent $\tau$ was found to be between 2 and 3 [1, 19, 30, 43]. In such scale-free networks, high-degree vertices called hubs are likely present, and give rise to scale-free properties such as small distances and fast information spreading. The hubs also crucially influence local properties such as clustering [26, 41] and the occurrence of subgraphs [36]. Clustering can be measured in terms of the probability $c(k)$ that a degree-$k$ vertex creates triangles. Both empirically [33, 38] and theoretically [17, 41], it has been shown that $c(k)$ falls off with $k$, and hence that hubs are less likely to take part in triangles.

Whereas triangles and even larger subgraphs require us to study the correlation between at least three vertices, in this paper we study the degree correlation between pairs of two vertices in terms of $a(k)$, the average degree of a neighbour of a vertex of degree $k$. According to several studies [2, 15], this degree–degree correlation is an essential local network property, because it also falls off with $k$ and can largely explain the fall-off of $c(k)$ [9, 15, 40]. We provide support for this statement by identifying an explicit relation between $a(k)$ and $c(k)$ for large $k$. But the main goal of this paper is to explain the full spectrum $k \mapsto a(k)$ for all $k$, and to provide theoretical underpinning for the widely observed $a(k)$ fall-off.

There exists a vast array of papers, empirical, non-rigorous, and rigorous, on $a(k)$ [2, 3, 9, 10, 15, 34, 37, 38, 42, 45]. The function $k \mapsto a(k)$ describes the correlation between the degrees on the two sides of an edge, and classifies the network into one of the following

Received 29 May 2018; revision received 24 April 2019.
* Current address: Twente University, The Netherlands. Email address: c.stegehuis@utwente.nl


Downloaded from https://www.cambridge.org/core. University of Bath, on 09 Feb 2020 at 11:24:30, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2019.45

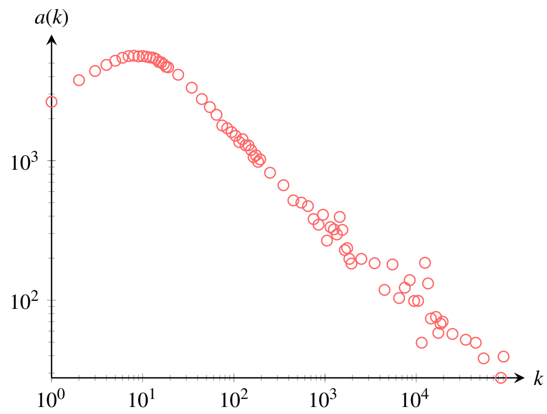



FIGURE 1: $a(k)$ for the YouTube friendship network [32].

three categories [35]. When $a(k)$ increases with $k$, the network is said to be *assortative*: vertices with high degrees mostly connect to other vertices with high degrees. When $a(k)$ decreases in $k$, the network is said to be *disassortative*. Then high-degree vertices typically connect to low-degree vertices. When $a(k)$ forms a flat curve in $k$, the network is said to be *uncorrelated*. In this case, the degrees on the two different sides of an edge can be viewed as independent of each other, a desirable property when studying the mathematical properties of networks. But the fact is that the majority of real-world networks with power-law degrees and unbounded degree fluctuations ($\tau \in (2, 3)$) show a clear decay of $a(k)$ as $k$ grows large [33, 38]. Figure 1 illustrates this for the YouTube friendship network [32]. Hence, scale-free networks are inherently disassortative, and hubs are predominantly connected to small-degree vertices. In complex network theory, such a well-established empirical fact then asks for a theoretical explanation. Typically, this explanation comes in the form of a null model that matches the degree distribution and has the empirical observation as a property, in this case disassortivity, or more specifically, the essential features of the curve $k \mapsto a(k)$.

The popular configuration model [11] generates random networks with any prescribed degree distribution, but only results in uncorrelated networks when including self-loops and multi-edges. Hence, the configuration model can never explain the $a(k)$ fall-off. We therefore resort to different null models that, contrary to the configuration model, generate random networks without self-loops and multi-edges. The resulting *simple* random networks are therefore prone to the structural correlations that come with the presence of hubs. We study $a(k)$ for three widely used random graph models: the erased configuration model, the rank-1 inhomogeneous random graph (also called hidden variable model), and the hyperbolic random graph. We show that these models display universal $a(k)$-behaviour: For $k$ sufficiently small, $a(k)$ is independent of $k$. Thus, in simple scale-free networks, neighbours of small-degree vertices are similar. We then identify the value of $k$ for which $a(k)$ starts decaying. An intuitive explanation for the $a(k)$ fall-off is that in simple networks, high-degree vertices have so many neighbours that they must reach out to lower-degree vertices, because networks typically only contain a small amount of high-degree vertices. This causes the average degree of a neighbour of a high-degree vertex to be smaller. Thus, single-edge constraints may cause the decay of $a(k)$.

Downloaded from https://www.cambridge.org/core. University of Bath, on 09 Feb 2020 at 11:24:30, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2019.45

## 2. Main results

We first define the average nearest-neighbour degree $a(k, G)$ of a graph $G$ in more detail. Let $(D_i)_{i\in[n]}$ be the degree sequence of the graph, where $[n] = 1, \ldots, n$. Furthermore, let $N_k$ denote the total number of degree $k$ vertices in the graph, and let $\mathcal{N}_i$ denote the neighbourhood of vertex $i$. The average nearest-neighbour degree of graph $G$ is then defined as

$$a(k, G) = \frac{1}{kN_k} \sum_{i:D_i=k} \sum_{j\in\mathcal{N}_i} D_j. \tag{1}$$

Note that it is possible that no vertex of degree $k$ exists in the graph, and in this situation we set $a(k, G) = 0$. We therefore analyse

$$a_\varepsilon(k, G) = \frac{1}{k|M_\varepsilon(k)|} \sum_{i\in M_\varepsilon(k)} \sum_{j\in\mathcal{N}_i} D_j, \tag{2}$$

where

$$M_\varepsilon(k) = \{i \in [n] \colon D_i \in [k(1-\varepsilon), k(1+\varepsilon)]\}.$$

When $M_\varepsilon(k) = \varnothing$, we set $a_\varepsilon(k, G) = 0$. We will show that in the models we analyse, $M_\varepsilon(k)$ is non-empty with high probability, so that $a_\varepsilon(k, G)$ is well-defined with high probability. Note that $a(k, G) = a_0(k, G)$. We now analyse $a_\varepsilon(k, G)$, first for the erased configuration model in Section 2.1 and then for the rank-1 inhomogeneous random graph and the hyperbolic random graph in Sections 2.3 and 2.4.

### 2.1. The erased configuration model

Given a positive integer $n$ and a degree sequence $(D_1, D_2, \ldots, D_n)$ such that the sum of the degrees is even, the configuration model is a (multi)graph where vertex $i$ has degree $D_i$ [11]. We start with $D_j$ free half-edges adjacent to vertex $j$, for $j = 1, \ldots, n$. The configuration model is then constructed by pairing free half-edges uniformly at random into edges, until no free half-edges remain. Conditionally on obtaining a simple graph, the resulting graph is a uniform graph with the prescribed degrees. This is why the configuration model is often used as a null model for real-world networks with given degrees. When the degree distribution has an infinite second moment, however, the probability of obtaining a simple graph tends to zero as $n$ grows large (see e.g. [23, Chapter 7]). In this setting the configuration model can no longer be used as a null model for simple real-world networks. The erased configuration model is the model where all multiple edges are merged and all self-loops are removed [13]. We take the original degree sequence to be an i.i.d. sample from the distribution

$$\mathbb{P}(D = k) = ck^{-\tau}, \tag{3}$$

where $\tau \in (2, 3)$ so that $\mathbb{E}[D^2] = \infty$. We denote $\mathbb{E}[D] = \mu$. When this sample constructs a degree sequence such that the sum of the degrees is odd, we add an extra half-edge to the last vertex. This does not affect our computations. We denote the actual degree sequence of the graph after merging the multiple edges and removing self-loops by $(D^{(\mathrm{er})})_{i\in[n]}$, and we call these the resulting degrees.

2.1.1. *Stable random variables.* The limit theorem of $a(k, G_n)$ for the erased configuration model contains stable random variables. A random variable $X$ follows a stable distribution if,

Downloaded from https://www.cambridge.org/core. University of Bath, on 09 Feb 2020 at 11:24:30, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2019.45

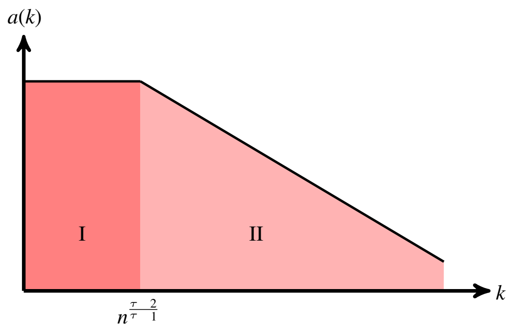



FIGURE 2: Illustration of the behaviour of $a(k, G_n)$ in the erased configuration model.

for any positive numbers $a_1$ and $a_2$, there exists a real number $b_1 = b_1(a_1, a_2)$ and a positive number $b_2 = b_2(a_1, a_2)$ such that

$$a_1X_1 + a_2X_2 \overset{d}{=} b_1 + b_2X, \tag{4}$$

where $X_1$ and $X_2$ are independent copies of $X$. Stable random variables can be parametrized by four parameters, and are usually denoted by $\mathcal{S}_\alpha(\sigma, \beta, \mu)$ (see e.g. [44, Chapter 4]). Throughout this paper, we will only use stable distributions with $\sigma = 1$, $\beta = 1$, $\mu = 0$ and we denote $\mathcal{S}_\alpha = \mathcal{S}_\alpha(1, 1, 0)$ to ease notation.

We now state the main result for the erased configuration model.

**Theorem 1.** ($a(k, G)$ in the erased configuration model.) *Let $(G_n)_{n\geq 1}$ be a sequence of erased configuration models on n vertices, where the degrees are an i.i.d. sample from (3). Take $\varepsilon_n = 1/(\log(\log(n)))$ and let $\Gamma$ denote the gamma function.*

(i) *For $k \ll n^{(\tau-2)/(\tau-1)}/\log(n)$,*

$$\frac{a_{\varepsilon_n}(k, G_n)}{n^{(3-\tau)/(\tau-1)}} \xrightarrow{d} \frac{1}{\mu}\left(\frac{2c\Gamma(\frac{5}{2}-\frac{1}{2}\tau)}{(\tau-1)(3-\tau)}\cos\left(\frac{\pi(\tau-1)}{4}\right)\right)^{2/(\tau-1)}\mathcal{S}_{(\tau-1)/2}, \tag{5}$$

*where $\mathcal{S}_{(\tau-1)/2}$ is a stable random variable.*

(ii) *For $n^{(\tau-2)/(\tau-1)} \ll k \ll n^{1/(\tau-1)}/\log(n)$,*

$$\frac{a_{\varepsilon_n}(k, G_n)}{n^{3-\tau}k^{\tau-3}} \xrightarrow{\mathbb{P}} -c\mu^{2-\tau}\Gamma(2-\tau). \tag{6}$$

**Remark 1.** The convergence in (5) also holds jointly in $k$ and $n$, so that for fixed $m \geq 1$ and $1 \leq k_1 < k_2 < \cdots < k_m \ll n^{(\tau-2)/(\tau-1)}/\log(n)$,

$$\frac{(a_{\varepsilon_n}(k_i, G_n))_{i\in[m]}}{n^{(3-\tau)/(\tau-1)}} \xrightarrow{d} \frac{1}{\mu}\left(\frac{2c\Gamma(\frac{5}{2}-\frac{1}{2}\tau)}{(\tau-1)(3-\tau)}\cos\left(\frac{\pi(\tau-1)}{4}\right)\right)^{2/(\tau-1)}\mathcal{S}_{(\tau-1)/2}\mathbf{1},$$

where $\mathbf{1} \in \mathbb{R}^m$ is a vector with $m$ entries equal to 1. Thus, $a_{\varepsilon_n}(k_i, G_n)$ and $a_{\varepsilon_n}(k_j, G_n)$ converge to the same realization of the random variable $\mathcal{S}_{(\tau-1)/2}$. We will prove this remark at the end of the proof of Theorem 1.

Downloaded from https://www.cambridge.org/core. University of Bath, on 09 Feb 2020 at 11:24:30, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2019.45

Figure 2 illustrates the behaviour of $a_{\varepsilon_n}(k, G_n)$. First, it stays flat and does not depend on $k$. After that, $a(k, G_n)$ starts decreasing in $k$, which shows that the erased configuration model indeed is a disassortative random graph. Theorem 1 shows that $n^{(\tau-2)/(\tau-1)}$ serves as a threshold. Thus, the negative degree–degree correlations due to the single-edge constraint only affect vertices of degrees at least $n^{(\tau-2)/(\tau-1)}$. This can be understood as follows. In the erased configuration model the maximum contribution to $a_{\varepsilon_n}(k, G)$ (see Propositions 1 and 2) comes from vertices with degrees proportional to $n/k$. The maximal degree in an observation of $n$ i.i.d. power-law distributed samples is proportional to $n^{1/(\tau-1)}$ w.h.p. Therefore, if $k \ll n^{(\tau-2)/(\tau-1)}$, such vertices with degree proportional to $n/k$ do not exist w.h.p. This explains the two regimes.

For $k$ small, $a_{\varepsilon_n}(k, G_n)$ converges to a stable random variable, as was also shown in [45] for $k$ fixed. Thus, for $k$ small, different instances of the erased configuration model show wild fluctuations. The joint convergence in $k$ of Remark 1 shows that $a(k, G_n)$ still forms a flat curve in $k$ for one realization of an erased configuration model when $k$ is small. In contrast, $a_{\varepsilon_n}(k, G_n)$ converges to a constant for large $k$-values, so that different realizations of erased configuration models result in similar $a_{\varepsilon_n}(k, G_n)$-values.

### 2.2. Sketch of the proof

We now give a heuristic proof of Theorem 1. Conditionally on the degrees, the probability that vertices with degrees $D_i$ and $D_j$ are connected in the erased configuration model can be approximated by [24] $1-\mathrm{e}^{-D_iD_j/\mu n}$. Let $v \in M_{\varepsilon_n}(k)$, and let $X_{iv}$ denote the indicator that vertex $i$ is connected to $v$. The expected degree of a neighbour of $v$ can then be approximated by

$$a_{\varepsilon_n}(k, G_n) \approx k^{-1} \sum_{i\in[n]} D_i \mathbb{P}(X_{iv}=1) \approx k^{-1} \sum_{i\in[n]} D_i(1-\mathrm{e}^{-D_ik/(\mu n)}). \tag{7}$$

The maximum degree in an i.i.d. sample from (3) scales as $n^{1/(\tau-1)}$ w.h.p. Thus, as long as $k \ll n^{(\tau-2)/(\tau-1)}$, we can Taylor-expand the exponential so that

$$a_{\varepsilon_n}(k, G_n) \approx \frac{1}{\mu n} \sum_{i\in[n]} D_i^2.$$

Because $(D_i)_{i\in[n]}$ are samples from a power-law distribution with infinite second moment, the Stable Law Central Limit Theorem gives Theorem 1(i).

When $k \gg n^{(\tau-2)/(\tau-1)}$, we approximate the sum in (7) by the integral

$$\begin{aligned} a_{\varepsilon_n}(k, G_n) &\approx cnk^{-1} \int_1^\infty x^{1-\tau}(1-\mathrm{e}^{-xk/(\mu n)})\,\mathrm{d}x \\ &= c\mu^{2-\tau}\left(\frac{n}{k}\right)^{3-\tau} \int_{k/(\mu n)}^\infty y^{1-\tau}(1-\mathrm{e}^{-y})\,\mathrm{d}y, \end{aligned}$$

using the degree distribution (3) and the change of variables $y = xk/(\mu n)$. When $k \ll n$, we can approximate this by

$$a_{\varepsilon_n}(k, G_n) \approx c\mu^{2-\tau}\left(\frac{n}{k}\right)^{3-\tau} \int_0^\infty y^{1-\tau}(1-\mathrm{e}^{-y})\,\mathrm{d}y = -c\mu^{2-\tau}\left(\frac{n}{k}\right)^{3-\tau} \Gamma(2-\tau).$$

The proof of Theorem 1(ii) then consists of showing that the above approximations are indeed valid. We prove Theorem 1 in detail in Sections 3.2 and 3.3.

Downloaded from https://www.cambridge.org/core. University of Bath, on 09 Feb 2020 at 11:24:30, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2019.45

### 2.3. Rank-1 inhomogeneous random graphs

We now turn to the rank-1 inhomogeneous random graph (or hidden variable model). This model constructs simple graphs with soft constraints on the degree sequence [9, 16]. The graph consists of $n$ vertices with weights $(h_i)_{i\in[n]}$. These weights are an i.i.d. sample from the power-law distribution (3). We denote the average value of the weights by $\mu$. Then, every pair of vertices with weights $(h, h')$ is connected with probability $p(h, h')$. In this paper, we take

$$p(h, h') = \min(hh'/(\mu n), 1),$$

which is the Chung–Lu version of the rank-1 inhomogeneous random graph [16]. This connection probability ensures that the degree of a vertex with weight $h$ will be close to $h$ [9]. We show the following result.

**Theorem 2.** ($a(k, G_n)$ in the rank-1 inhomogeneous random graph.) *Let $(G_n)_{n\geq 1}$ be a sequence of rank-1 inhomogeneous random graphs on n vertices, where the weights are an i.i.d. sample from (3). Take $\varepsilon_n = 1/(\log(\log(n)))$ and let $\Gamma$ denote the gamma function.*

(i) *For $1 \ll k \ll n^{(\tau-2)/(\tau-1)}/\log(n)$,*

$$\frac{a_{\varepsilon_n}(k, G_n)}{n^{(3-\tau)/(\tau-1)}} \xrightarrow{d} \frac{1}{\mu}\left(\frac{2c\Gamma(\frac{5}{2}-\frac{1}{2}\tau)}{(\tau-1)(3-\tau)}\cos\left(\frac{\pi(\tau-1)}{4}\right)\right)^{2/(\tau-1)} \mathcal{S}_{(\tau-1)/2},$$

*where $\mathcal{S}_{(\tau-1)/2}$ is a stable random variable.*

(ii) *For $n^{(\tau-2)/(\tau-1)} \ll k \ll n^{1/(\tau-1)}/\log(n)$,*

$$\frac{a_{\varepsilon_n}(k, G_n)}{n^{3-\tau}k^{\tau-3}} \xrightarrow{\mathbb{P}} \frac{c\mu^{2-\tau}}{(3-\tau)(\tau-2)}. \tag{8}$$

Theorem 2 is almost identical to Theorem 1. The proof of Theorem 2 exploits the deep connection between the two models, and essentially carries over the results for the erased configuration model to the rank-1 inhomogeneous random graph. The similarity can be understood by noticing that in the erased configuration model the probability that vertices $i$ and $j$ with degrees $D_i$ and $D_j$ are connected can be approximated by $1 - \exp(-D_iD_j/L_n)$ which is close to

$$\min\left(1, \frac{D_iD_j}{\mu n}\right),$$

the connection probability in the rank-1 inhomogeneous random graph. Arguments similar to those that led to (7) show that $a_{\varepsilon_n}(k, G_n)$ can be approximated by

$$a_{\varepsilon_n}(k, G_n) \approx k^{-1} \sum_{i\in[n]} h_i \min(h_ik/\mu n, 1).$$

This sum behaves very similarly to the sum in (7), so that the only difference between Theorem 1 and 2 is the limiting constants in (6) and (8). The main difference between the two models is that in the rank-1 inhomogeneous random graph the presence of all edges is independent as soon as the weights are sampled. This is not true in the erased configuration model, because we know that a vertex with sampled degree $D_i$ cannot have more than $D_i$ neighbours, creating dependence between the presence of edges incident to vertex $i$. We show that these correlations between the presence of different edges in the erased configuration model are small enough for $a_{\varepsilon_n}(k, G_n)$ to behave similarly in the erased configuration model and the rank-1 inhomogeneous random graph.

Downloaded from https://www.cambridge.org/core. University of Bath, on 09 Feb 2020 at 11:24:30, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2019.45

### 2.4. Hyperbolic random graphs

The third random graph model we consider is the hyperbolic random graph. This model was introduced in [31] and samples $n$ vertices on a disk of radius $R = 2\log(n/\nu)$, where the density of the radial coordinate $r$ a vertex $p = (r, \phi)$ is

$$\rho(r) = \alpha \frac{\sinh(\alpha r)}{\cosh(\alpha R) - 1}$$

with $\alpha = (\tau - 1)/2$. The angle of $p$ is sampled uniformly from $[0, 2\pi]$. Then, two vertices are connected if their hyperbolic distance is at most $R$. The hyperbolic distance of points $u = (r_u, \phi_u)$ and $v = (r_v, \phi_v)$ is defined by

$$\cosh(\mathrm{d}(u, v)) = \cosh(r_u)\cosh(r_v) - \sinh(r_u)\sinh(r_v)\cos(\theta_{uv}), \tag{9}$$

where $\theta_{uv}$ denotes the relative angle between $\phi_u$ and $\phi_v$. This creates a simple random graph with power-law degrees with exponent $\tau$ [31]. The parameter $\nu$ fixes the average degree of the graph.

The hyperbolic random graph creates simple sparse random graphs with power-law degrees, but in contrast to the erased configuration model and the rank-1 inhomogeneous random graph, can at the same time create many triangles due to its geometric nature [14, 31]. In both the rank-1 inhomogeneous random graph and the erased configuration model, the connection probabilities of different pairs of vertices are (almost) independent. In the hyperbolic random graph, this is not true. When $u$ is connected to $v$ and $u$ is connected to $w$, then $v$ and $w$ should also be close to one another by the triangle inequality. However, if we define the type of a vertex as

$$t(u) = \mathrm{e}^{(R - r_u)/2},$$

then we show that we can approximate the probability that vertices $u$ and $v$ are connected by

$$\mathbb{P}(X_{uv} = 1 \mid t(u), t(v)) = \begin{cases} \dfrac{2}{\pi}\sin^{-1}(\nu t(u)t(v)/n) & \nu t(u)t(v)/n < 1, \\ 1 & \nu t(u)t(v)/n \geq 1, \end{cases}$$

which behaves similarly to the connection probability in the rank-1 inhomogeneous random graph. Furthermore, by [7, Lemma 1.3], the density of $2\ln(t(u))$ can be written as

$$f_{2\ln(t(u))}(x) = \left(\frac{\tau - 1}{2}\right)\mathrm{e}^{-(\tau-1)x/2}(1 + \mathrm{o}(1)),$$

where the $\mathrm{o}(1)$ term is with respect to the network size $n$. Therefore,

$$\mathbb{P}(t(u) > x) = \mathbb{P}(2\ln(t(u)) > 2\ln(x)) = x^{-\tau+1}(1 + \mathrm{o}(1)), \tag{10}$$

so that on a high level the hyperbolic random graph can be interpreted as a rank-1 inhomogeneous random graph with $(t(u))_{u\in[n]}$ as weights (see [7, Section 1.1.1] for a more elaborate discussion).

The next theorem shows that indeed the behaviour of $a_{\varepsilon_n}(k, G_n)$ in the hyperbolic random graph is similar to the rank-1 inhomogeneous random graph.

Downloaded from https://www.cambridge.org/core. University of Bath, on 09 Feb 2020 at 11:24:30, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2019.45

**Theorem 3.** ($a(k, G_n)$ in the hyperbolic random graph.) *Let $(G_n)_{n\geq 1}$ be a sequence of hyperbolic random graphs on n vertices with power-law degrees with exponent $\tau$ and parameter $\nu$. Take $\varepsilon_n = 1/\log(\log(n))$ and let $\Gamma$ denote the gamma function.*

(i) *For $1 \ll k \ll n^{(\tau-2)/(\tau-1)}/\log(n)$,*

$$\frac{a_{\varepsilon_n}(k, G_n)}{n^{(3-\tau)/(\tau-1)}} \xrightarrow{d} \frac{2\nu}{\pi}\left(\frac{2}{3-\tau}\Gamma\left(\frac{5}{2}-\frac{1}{2}\tau\right)\cos\left(\frac{\pi(\tau-1)}{4}\right)\right)^{2/(\tau-1)} \mathcal{S}_{(\tau-1)/2},$$

*where $\mathcal{S}_{(\tau-1)/2}$ is a stable random variable.*

(ii) *For $n^{(\tau-2)/(\tau-1)} \ll k \ll n^{1/(\tau-1)}/\log(n)$,*

$$\frac{a_{\varepsilon_n}(k, G_n)}{n^{3-\tau}k^{\tau-3}} \xrightarrow{\mathbb{P}} \frac{(\tau-1)\nu\sqrt{\pi}\,\Gamma(\frac{3}{2}-\frac{\tau}{2})}{2(\tau-2)\Gamma(2-\frac{\tau}{2})}\left(\frac{2(\tau-1)}{\pi(\tau-2)}\right)^{3-\tau}.$$

## 2.5. Discussion

2.5.1. *Universality.* The behaviour of $a_{\varepsilon_n}(k, G_n)$ is universal across the three null models we consider. The erased configuration model and the rank-1 inhomogeneous random graph are closely related. They are known to behave similarly, for example, under critical percolation [4, 6], in terms of distances [18] when $\tau > 3$, and in terms of clustering when $\tau \in (2, 3)$ [41]. The hyperbolic random graph typically shows different behaviour, for example in terms of clustering [14, 22], or connectivity [7, 8]. Still, the behaviour of $a_{\varepsilon_n}(k, G_n)$ is similar in the hyperbolic random graph and the other two null models. In all three null models, the main contribution for $k \gg n^{(\tau-2)/(\tau-1)}$ comes from vertices with degrees proportional to $n/k$ (see Propositions 1 and 2). In the hyperbolic random graph, we can relate this maximum contribution to the geometry of the hyperbolic sphere. A vertex $i$ of degree $k$ has radius $r_i \approx R - 2\log(k)$. Similarly, a vertex $j$ of degree $n/(\nu k)$ has radius $r_j \approx R - 2\log(n/(k\nu)) = 2\log(k)$. Then, $r_j \approx R - r_i$, so that the major contributing vertices have radial coordinate proportional to $R - r_i$.

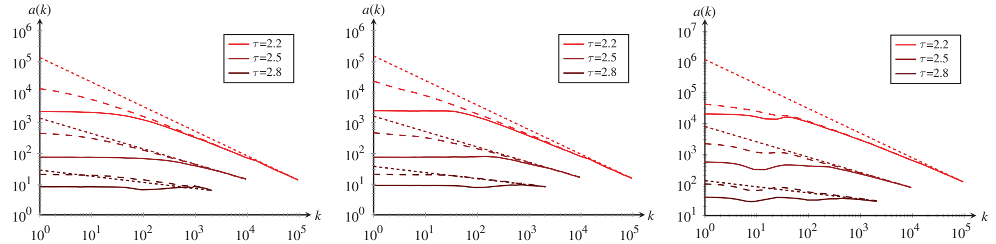
2.5.2. *Expected average nearest-neighbour degree.* In Theorems 1–3 we show that $a_{\varepsilon_n}(k, G_n)$ converges in probability to a stable random variable when $k$ is small. Thus, when we generate many samples of random graphs, for fixed $k$, the distribution of the values of $a_{\varepsilon_n}(k, G_n)$ across the different samples will look like a stable random variable. We can also study the expected value of $a(k, G_n)$ across the different samples. For the erased configuration model for example, we can show that (see Section 3.4)

$$\lim_{n\to\infty} \frac{\mathbb{E}[a(k, G_n)]}{(n/k)^{3-\tau}} = -c\mu^{2-\tau}\Gamma(2-\tau). \tag{11}$$

The difference between the scaling of the expected value of $a(k, G_n)$ and the typical behaviour of $a(k, G_n)$ in Theorem 1(i) is caused by high-degree vertices. In typical degree sequences, the maximum degree is proportional to $n^{1/(\tau-1)}$. It is unlikely that vertices with higher degrees are present, but if they are, they have a high impact on the average nearest-neighbour degree of low-degree vertices, causing the difference between the expected average nearest-neighbour degree and the typical average nearest-neighbour degree. Thus, the expected value of $a(k, G_n)$ is not very informative when $k$ is small, since Theorem 1 shows that $a(k, G_n)$ will almost always be smaller than its expected value when $k$ is small.

Figure 3 illustrates this difference in terms of the mean and median value of $a(k, G_n)$ over many realizations of the erased configuration model, the rank-1 inhomogeneous random graph,

Downloaded from https://www.cambridge.org/core. University of Bath, on 09 Feb 2020 at 11:24:30, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2019.45




(a) Erased configuration model (b) Inhomogeneous random graph (c) Hyperbolic random graph

FIGURE 3: $a(k, G_n)$ for different random graph models with $n = 10^6$. The solid line is the median of $a(k, G_n)$ over $10^4$ realizations of the random graph, and the dashed line is the average over these realizations. The dotted line is the asymptotic slope $k^{\tau-3}$.

and the hyperbolic random graph. Here indeed we see that the expected average neighbour degree scales as a power of $k$ over the entire range of $k$, where the median shows the straight part of the curve from Theorem 1. Thus, it is important to distinguish between mean and median of $a(k, G_n)$ when simulating random graphs.

2.5.3. *Vertices of degree k.* Definition (1) assumes that a vertex of degree $k$ is present. For large values of $k$, this is a rare event, by (3). Indeed, vertices of degree at most $n^{1/\tau}$ are present with high probability in the erased configuration model, whereas the probability that a vertex of degree $k \gg n^{1/\tau}$ is present tends to zero in the large network limit [45]. We avoid this problem by averaging $a(k, G_n)$ over a small range of degrees. Another option is to condition on the event that a vertex of degree $k$ is present. Our proofs for $k \ll n^{(\tau-2)/(\tau-1)}$ for the erased configuration model can easily be adjusted to condition on this event. For $k$ larger, we leave the behaviour of $a(k, G_n)$ conditionally on a vertex of degree $k$ being present open for further research.

2.5.4. *Fixed degrees.* In the proof of Theorem 1 we show that the fluctuations that come with the stable laws for small $k$ are not present when we condition on the degree sequence. Thus, the large fluctuations in $a_{\varepsilon_n}(k, G_n)$ for small $k$ are caused by fluctuations of the i.i.d. degrees, weights, or radii. For a given real-world network, its degrees are often preserved, and many erased configuration models or inhomogeneous random graphs with the same observed degree sequence are created. In this fixed-degree setting, the sample-to-sample fluctuations of $a_{\varepsilon_n}(k, G_n)$ are relatively small.

2.5.5. *Relation with local clustering.* The local clustering coefficient $c(k)$ of vertices of degree $k$ measures the probability that two randomly chosen neighbours af a randomly chosen vertex of degree $k$ are connected. In many real-world networks as well as simple null models, $c(k)$ decreases as a function of $k$ [9, 26, 39, 41, 43]. The relation between the decay rate of $c(k)$ and the decay rate of $a(k)$ has been investigated for the rank-1 inhomogeneous random graph, where it was shown that $c(k) < a(k)/k$ [40]. Using recent results for $c(k)$ on the erased configuration model and the rank-1 inhomogeneous random graph, we can make the relation between $c(k)$ and $a(k)$ more precise. When $k \gg \sqrt{n}$, $c(k)$ in the erased configuration model satisfies [26]

$$c(k) = c^2\Gamma(2-\tau)^2\mu^{3-2\tau}n^{5-2\tau}k^{2\tau-6}(1 + o_{\mathbb{P}}(1)).$$

Then, by Theorem 1, when $k \gg \sqrt{n}$,

$$c(k) = \frac{a(k)^2}{\mu n}(1 + o_{\mathbb{P}}(1)). \tag{12}$$

Downloaded from https://www.cambridge.org/core. University of Bath, on 09 Feb 2020 at 11:24:30, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2019.45

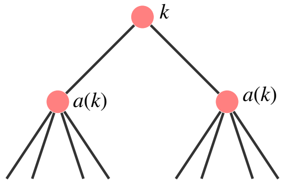



FIGURE 4: The neighbours of a vertex of degree $k$ have average degree $a(k)$.

Intuitively, we can see this relationship in the following way. Pick two neighbours of a vertex of degree $k$. By definition, these vertices have degree $a(k)$ on average. See Figure 4. Since $k \gg \sqrt{n}$, by Theorem 2, $a(k) \ll \sqrt{n}$. Therefore, the probability of two vertices with weight $a(k)$ being connected is approximately $1 - \mathrm{e}^{-a(k)^2/\mu n} \approx a(k)^2/\mu n$. Since the clustering coefficient can be interpreted as the probability that two randomly chosen neighbours are connected, the clustering coefficient should satisfy $c(k) \sim a(k)^2/\mu n$ when $k \gg \sqrt{n}$. In particular, the decay of the clustering coefficient should be twice as fast as the decay of the average neighbour degree. Analytical results on $c(k)$ on the rank-1 inhomogeneous random graph show that (12) is also the correct relation between clustering and degree correlations in the rank-1 inhomogeneous random graph [41]. Future research might explore the relation between $c(k)$ and $a(k)$ in other null models, such as the hyperbolic random graph or the preferential attachment model. It would also be interesting to see if the difference between expectation and typical behaviour that is present in $a(k)$ also occurs for the local clustering coefficient $c(k)$.

2.5.6. *Correlations in the hyperbolic random graph.* The relation in (12) is based on the fact that in the erased configuration model and the rank-1 inhomogeneous random graph the connection probabilities of pairs of vertices $(i, j)$, $(i, k)$, and $(j, k)$ are (almost) independent. In the hyperbolic random graph, the geometry causes a strong dependence between these connection probabilities. If vertices $j$ and $k$ are neighbours of $i$, they are likely to be geometrically close to one another due to the triangle inequality. This makes the probability that $j$ and $k$ are connected larger than in the rank-1 inhomogeneous random graph or the erased configuration model. These correlations do not play a role when computing $a(k, G_n)$, since it only involves the connection probability of two different vertices. When computing statistics of the hyperbolic random graph that include three-point correlations, the equivalence between the hyperbolic random graph and the rank-1 inhomogeneous random graph may fail to hold, as in the example of $c(k)$.

Interestingly, the number of cliques was also shown to be similar in the hyperbolic random graph, the rank-1 inhomogeneous random graph and the erased configuration model [20], even though cliques clearly involve three-point correlations. Cliques in the hyperbolic random graph are typically formed between vertices at radius proportional to $R/2$ [20], so that their degrees are proportional to $\sqrt{n}$ [7]. These vertices form a dense core, which is very similar to what happens in the erased configuration model and the rank-1 inhomogeneous random graph [29]. In the erased configuration model, many other small subgraphs typically occur between vertices of degrees proportional to $\sqrt{n}$ [27]. It would be interesting to see if the number of these small subgraphs behaves similarly in the hyperbolic random graph.

## 3. Average nearest-neighbour degree in the ECM

In this section, we prove Theorem 1. For $k \ll n^{(\tau-2)/(\tau-1)}/\log(n)$, we couple the degrees of neighbours of a uniformly chosen vertex of degree $k$ to i.i.d. samples of the size-biased degree

Downloaded from https://www.cambridge.org/core. University of Bath, on 09 Feb 2020 at 11:24:30, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2019.45

distribution in Section 3.2. When $k \gg n^{(\tau-2)/(\tau-1)}$, this coupling is no longer valid. We then show in Section 3.3 that a specific range of degrees contributes most to $a_{\varepsilon_n}(k, G_n)$.

### 3.1. Preliminaries

We say that $X_n = \mathrm{O}_{\mathbb{P}}(b_n)$ for a sequence of random variables $(X_n)_{n\geq 1}$ if $|X_n|/b_n$ is a tight sequence of random variables, and $X_n = \mathrm{o}_{\mathbb{P}}(b_n)$ if $X_n/b_n \xrightarrow{\mathbb{P}} 0$. Let $L_n$ denote the total number of half-edges, so that $L_n = \sum_i D_i$. We define the events

$$
\begin{aligned}
J_n &= \{|L_n - \mu n| \leq n^{2/\tau}\},\\
\mathcal{A}_n &= \{|M_{\varepsilon_n}(k)| \geq nk^{1-\tau}/\log(n)\},\\
\mathcal{B}_n &= \{\max_i D_i \leq n^{1/(\tau-1)}\log(n)\}.
\end{aligned}
\tag{13}
$$

By [25, Lemma 2.3], $\mathbb{P}(J_n) \to 1$ as $n \to \infty$. Furthermore, since the degree sequence is an i.i.d. sample from (3), also $\mathbb{P}(\mathcal{B}_n) \to 1$. Note that $\mathcal{A}_n$ implies that $|M_{\varepsilon_n}(k)| \geq 1$ as long as $k \ll n^{1/(\tau-1)}/\log(n)$. Below, we will show that $\mathbb{P}(\mathcal{A}_n) \to 1$ for $k \ll n^{1/(\tau-1)}/\log(n)$. Throughout the rest of the paper, we will often condition on the event

$$
\Lambda_n = J_n \cap \mathcal{A}_n \cap \mathcal{B}_n. \tag{14}
$$

Note that $\mathbb{P}(\Lambda_n) \to 1$.

In the rest of this paper, we will often condition on the degree sequence. For some event $\mathcal{E}$, we use the notation $\mathbb{P}_n(\mathcal{E}) = \mathbb{P}(\mathcal{E} \mid (D_i)_{i\in[n]})$, and we define $\mathbb{E}_n$ and $\mathrm{Var}_n$ similarly, where we often assume that the event $\Lambda_n$ holds. In this notation $\mathbb{P}_n(\mathcal{E}_1 \mid \mathcal{E}_2) = \mathbb{P}(\mathcal{E}_1 \mid \mathcal{E}_2, (D_i)_{i\in[n]})$. In this notation, we often work on the event $\Lambda_n$. That is, we assume that the degree sequence satisfies the event $\Lambda_n$, and we compute $\mathbb{P}(\mathcal{E} \mid (D_i)_{i\in[n]} \in \Lambda_n)$.

We often want to interchange the sampled degree of a vertex $i$, $D_i$ and its resulting degree $D^{(\mathrm{er})}{}_i$. The next lemma shows that $D_i$ and $D^{(\mathrm{er})}{}_i$ are close.

**Leema 1.** *Let $G$ be an erased configuration model where the degrees are i.i.d. samples from a power-law distribution with $\tau \in (2, 3)$. Assume the event $J_n \cap \mathcal{B}_n \cap \mathcal{D}_n(D_i)$, where*

$$
\mathcal{D}_n(D_i) = \left\{\sum_{j\in[n]} \psi\left(\frac{D_iD_j}{\mu n}\right) \leq D_i^{\tau-1} n^{1-\tau}\log(n)\right\},
$$

*and $\psi(x) = x - 1 + \mathrm{e}^{-x}$. Then,*

$$
\mathbb{P}_n(D_i - D^{(\mathrm{er})}{}_i > \varepsilon D_i) \leq \mathrm{O}(\varepsilon^{-1}(D_i/n)^{\tau-2}\log(n)) + \mathrm{O}(\varepsilon^{-1}D_i/n). \tag{15}
$$

Note that by [28, Theorem 3(ii)], for $D$ distributed as in (3), $\mathbb{E}[\psi(D/t)] = \mathrm{O}(t^{-(\tau-1)})$ so that

$$
\mathbb{E}[\psi(D_iD_j/(\mu n)) \mid D_i] = \mathrm{O}(D_i^{\tau-1}n^{1-\tau}).
$$

This shows that for all $D_i$, $\mathbb{P}(\mathcal{D}_n(D_i)) \to 1$.

*Proof.* We use [24, (4.9)], which calculates $p(n, m, L)$, the probability that a set of $n$ half-edges does not connect to another set of $m$ half-edges, when the total number of half-edges equals $L$. Then, choosing $n = D_i$, $m = D_j$ yields $p(D_i, D_j, L_n) = \mathbb{P}_n(X_{ij} = 0)$, so that by [24, (4.9)]

$$
\mathbb{P}_n(X_{ij} = 0) \leq \prod_{s=0}^{D_i-1}\left(1 - \frac{D_i}{L_n - 2D_i - 1}\right) + \frac{D_i^2 D_j}{(L_n - 2D_i)^2} \leq \mathrm{e}^{-D_iD_j/L_n} + \frac{D_i^2 D_j}{(L_n - 2D_i)^2}.
$$

Downloaded from https://www.cambridge.org/core. University of Bath, on 09 Feb 2020 at 11:24:30, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2019.45

Let $\psi(x) = x - 1 + \mathrm{e}^{-x}$. The expected number of erased edges at vertex $i$ satisfies

$$\begin{aligned}
\mathbb{E}_n[D_i - D^{(\mathrm{er})}{}_i] &= D_i - \sum_{j\in[n]} (1 - \mathbb{P}_n(X_{ij} = 0)) \\
&\leq \sum_{j\in[n]} \left( \frac{D_i D_j}{L_n} - 1 + \mathrm{e}^{-D_i D_j / L_n} + \frac{D_i^2 D_j}{(L_n - 2D_i)^2} \right) \\
&= \sum_{j\in[n]} \psi\left(\frac{D_i D_j}{L_n}\right) + \mathrm{O}\left(\frac{D_i^2}{L_n}\right) \\
&= \sum_{j\in[n]} \psi\left(\frac{D_i D_j}{\mu n}\right)(1 + \mathrm{o}(1)) + \mathrm{O}\left(\frac{D_i^2}{n}\right).
\end{aligned}$$

Thus, on the event $\mathcal{D}_n$,

$$\mathbb{E}_n[D_i - D^{(\mathrm{er})}{}_i] = \mathrm{O}(n^{-\tau+2} D_i^{\tau-1} \log(n)) + \mathrm{O}(D_i^2/n). \tag{16}$$

Using the Markov inequality, we obtain

$$\mathbb{P}_n(D_i - D^{(\mathrm{er})}{}_i > \varepsilon D_i) \leq \mathrm{O}(\varepsilon^{-1}(D_i/n)^{\tau-2} \log(n)) + \mathrm{O}(\varepsilon^{-1} D_i/n).$$

□

We now use Lemma 1 to show that $\mathbb{P}(\mathcal{A}_n) \to 1$, with $\mathcal{A}_n$ as in (13). Let $v$ be a uniformly chosen vertex and denote $I(k, \varepsilon_n) = [k(1-\varepsilon_n), k(1+\varepsilon_n)]$. Then,

$$\begin{aligned}
\mathbb{E}[|M_{\varepsilon_n}(k)|] &= n\mathbb{P}(D^{(\mathrm{er})}{}_v \in I(k, \varepsilon_n)) \\
&\geq n\mathbb{P}(D_v \in I(k, \varepsilon_n/2))\mathbb{P}(D^{(\mathrm{er})}{}_v \in I(k, \varepsilon_n) \mid D_v \in I(k, \varepsilon_n/2)) \\
&\geq n\mathbb{P}(D_v \in I(k, \varepsilon_n/2))\mathbb{P}(D_v - D^{(\mathrm{er})}{}_v \leq k\varepsilon_n/2 \mid D_v = k(1-\varepsilon_n/2), J_n, \mathcal{D}_n(k)) \\
&\quad \times \mathbb{P}(J_n \cap \mathcal{D}_n(k)).
\end{aligned} \tag{17}$$

Furthermore, for some $\tilde{C}_1 > 0$,

$$\mathbb{P}(D_v \in I(k, \varepsilon_n/2)) = \sum_{i=k(1-\varepsilon_n)}^{k(1+\varepsilon_n)} i^{-\tau} \geq \tilde{C}_1 \int_{k(1-\varepsilon_n)}^{k(1+\varepsilon_n)} x^{-\tau}\, \mathrm{d}x.$$

Therefore, using (17), for some $\tilde{C}_2 > 0$,

$$\begin{aligned}
\mathbb{E}[|M_{\varepsilon_n}(k)|] &\geq \tilde{C}_1 n \int_{k(1-\varepsilon_n/2)}^{k(1+\varepsilon_n/2)} x^{-\tau}\, \mathrm{d}x (1 - \mathrm{O}(\varepsilon_n^{-1}(k/n)^{\tau-2} \log(n))) \\
&= \tilde{C}_2 n k^{1-\tau} \varepsilon_n (1 + \mathrm{o}(1)),
\end{aligned}$$

where we used (15) combined with the choice of $\varepsilon_n$ in Theorem 1. Thus, $\mathbb{P}(\mathcal{A}_n) \to 1$ for $k \ll n^{1/(\tau-1)}/\log(n)$ by the choice of $\varepsilon_n$ in Theorem 1.

Downloaded from https://www.cambridge.org/core. University of Bath, on 09 Feb 2020 at 11:24:30, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2019.45

### 3.2. Small $k$: coupling to i.i.d. random variables

In this section we investigate the behaviour of $a_{\varepsilon_n}(k, G_n)$ when $k \ll n^{(\tau-2)/(\tau-1)}/\log(n)$. We first pick a random vertex $v \in M_{\varepsilon_n}(k)$. We couple the degrees of the neighbours of $v$ to i.i.d. copies of the size-biased degree distribution $D_n^*$, where

$$\mathbb{P}_n(D_n^* = j) = \frac{j}{L_n} \sum_{i \in [n]} \mathbf{1}_{\{D_i = j\}}.$$

We then use this coupling to compute $a_{\varepsilon_n}(k, G_n)$.

*Proof of Theorem 1(i).* We first condition on the degree sequence $(D_i)_{i\in[n]} \in \Lambda_n$ of (14), since $\mathbb{P}(\Lambda_n) \to 1$. Let $v$ be a vertex of degree $d$. In the configuration model, neighbours of $v$ are constructed by pairing the half-edges of $v$ uniformly to other half-edges. The distribution of the degree of a vertex attached to a uniformly chosen half-edge is given by $D_n^*$. However, the degrees of the neighbours of $v$ in the erased configuration model are not an i.i.d. sample of $D_n^*$ due to the fact that the half-edges should attach to distinct vertices that are different from $v$. We now couple the degrees of the neighbours of $v$ by an i.i.d. sample of $D_n^*$. Denote degrees of neighbours of $v$ by $B_1, \ldots, B_d$. Let $Y_1, \ldots, Y_d$ be i.i.d. samples of $D_n^*$. We use a similar coupling as in [5, Construction 4.2] to couple $B_i$ to $Y_i$. Set $V_0 = v$. Then, for $i \in [d]$ the coupling is defined in the following way.

- Take a uniformly chosen half-edge (with replacement), let $v'_i$ be the vertex attached to it, and set $Y_i = D_{v'_i}$.
- Sample a uniformly chosen half-edge from the set of half-edges not incident to $V_{i-1}$. Let $w_i$ denote the vertex incident to the chosen half-edge.
- If $v'_i \notin V_{i-1}$, then set $B_i = Y_i$. and $V_i = V_{i-1} \cup v'_i$.
- If $v'_i \in V_{i-1}$, set $B_i = D_{w_i}$ and $V_i = V_{i-1} \cup w_i$.

Summarizing,

$$(Y_i, B_i) = \begin{cases} (D_{v'_i}, D_{v'_i}) & \text{when } v'_i \notin V_{i-1}, \\ (D_{v'_i}, D_{w_i}) & \text{when } v'_i \in V_{i-1}, \end{cases}$$

where $(v'_i)_{i\in[d]}$ are vertices attached to uniformly chosen half-edges (with replacement), and $w_i$ are vertices attached to uniformly chosen half-edges from the set of half-edges not incident to $V_{i-1}$.

Informally, at every step $i$ we sample a uniformly chosen half-edge from all half-edges, and select the vertex $v'_i$ incident to it. If this vertex is different from all previously selected vertices and unequal to $v$, we declare the vertex to be a neighbour of $v$, and $B_i = Y_i = D_{v'_i}$. If not, we redraw the selected half-edge to ensure that all neighbours of $v$ are distinct and unequal to $v$, and declare the vertex attached to this half-edge, $w_i$ to be a neighbour of $v$ and set $B_i = D_{w_i}$.

Downloaded from https://www.cambridge.org/core. University of Bath, on 09 Feb 2020 at 11:24:30, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2019.45

When $B_s \neq Y_s$, $B_s$ is drawn in a size-biased manner from the vertices that are not chosen yet. Then, because $D_i \geq 0$ for all $i$,

$$\begin{aligned}\mathbb{E}_n[B_s \mid V_s, B_s \neq Y_s] &= \frac{\sum_{i\notin V_s} D_i^2}{\sum_{i\notin V_s} D_i} \\ &\leq \frac{\sum_{i\in[n]} D_i^2}{\sum_{i\in[n]} D_i - \sum_{i\in V_s} D_i} \\ &= \frac{\sum_{i\in[n]} D_i^2}{\sum_{i\in[n]} D_i}\left(1 + \frac{\sum_{i\in V_s} D_i}{\sum_{i\in[n]} D_i - \sum_{i\in V_s} D_i}\right).\end{aligned}$$

On the event $\mathcal{B}_n$, $D_{\max} \leq n^{1/(\tau-1)} \log(n)$, so that $\sum_{i\in V_s} D_i \leq sD_{\max} = \mathrm{O}(sn^{1/(\tau-1)} \log(n))$ for all possible $V_s$, so that for $s = \mathrm{o}(n^{(\tau-2)/(\tau-1)}/\log(n))$

$$\mathbb{E}_n[B_s \mid B_s \neq Y_s] = \frac{\sum_{i\in[n]} D_i^2}{L_n}(1 + \mathrm{o}(1)) = \mathbb{E}[D_n^*](1 + \mathrm{o}(1)). \tag{18}$$

Let $\mathcal{N}_v(U)$ denote a uniformly chosen neighbour of vertex $v$ of degree $\mathrm{o}(n^{(\tau-2)/(\tau-1)}/\log(n))$. Then, by (18) and the fact that $\mathbb{E}_n[Y_s] = \mathbb{E}_n[D_n^*]$,

$$\mathbb{E}_n[D_{\mathcal{N}_v(U)}] = \mathbb{E}[D_n^*](1 + \mathrm{o}(1)).$$

When $i \in M_{\varepsilon_n}(k)$, $D^{(\mathrm{er})}{}_i = k(1 + \mathrm{o}(1))$. Let $E_n$ denote the total number of erased edges in the erased configuration model. By [28, Theorem 1],

$$E_n = \mathrm{O}_{\mathbb{P}}(\min(n^{1/(\tau-1)}, n^{4/(\tau-1)-2})). \tag{19}$$

Then, conditionally on the degree sequence

$$\begin{aligned}a_{\varepsilon_n}(k, G_n) &= \frac{1}{k|M_{\varepsilon_n}(k)|} \sum_{i\in M_{\varepsilon_n}(k)} \sum_{j\in\mathcal{N}_i} D^{(\mathrm{er})}{}_j \\ &= \frac{1}{k|M_{\varepsilon_n}(k)|} \sum_{i\in M_{\varepsilon_n}(k)} \sum_{j\in\mathcal{N}_i} D_j - \frac{1}{k|M_{\varepsilon_n}(k)|} \sum_{i\in M_{\varepsilon_n}(k)} \sum_{j\in\mathcal{N}_i} (D_j - D^{(\mathrm{er})}{}_j) \\ &= (1 + \mathrm{o}(1))\mathbb{E}_n[D_{\mathcal{N}_{V_k}(U)}] + \mathrm{O}(E_n k^{-1} |M_{\varepsilon_n}(k)^{-1}|) \\ &= (1 + \mathrm{o}(1))\mathbb{E}_n[D_{\mathcal{N}_{V_k}(U)}] + \mathrm{O}_{\mathbb{P}}(k^{\tau-2} n^{-1} \log(n) \min(n^{1/(\tau-1)}, n^{4/(\tau-1)-2})),\end{aligned}$$

where $V_k$ denotes a uniformly chosen vertex in $M_{\varepsilon_n}(k)$, and $\mathcal{N}_{V_k}(U)$ is a uniformly chosen neighbour of vertex $V_k$. Here the third equality holds because the average nearest-neighbour degree averages over all neighbours of vertex $j$ and over all vertices in $M_{\varepsilon_n}(k)$, together with the fact that $D_{V_k} = k(1 + \mathrm{o}(1))$. The fourth equality follows from (19) and on the event $\mathcal{A}_n$. Note that for $k \ll n^{(\tau-2)/(\tau-1)}/\log(n)$ and $\tau \in (2, 3)$, the last term is $\mathrm{o}_{\mathbb{P}}(1)$. Then, conditionally on the degree sequence, on the event $\Lambda_n$,

$$a_{\varepsilon_n}(k, G_n) = (1 + \mathrm{o}(1))\mathbb{E}_n[D_{\mathcal{N}_{V_k}(U)}] + \mathrm{o}_{\mathbb{P}}(1) = (1 + \mathrm{o}(1))(\mu n)^{-1} \sum_{i\in[n]} D_i^2 + \mathrm{o}_{\mathbb{P}}(1). \tag{20}$$

For $t$ large, we obtain from (3) that

$$\mathbb{P}(D^2 > t) = \mathbb{P}(D > \sqrt{t}) = \frac{c}{\tau - 1} t^{(1-\tau)/2}(1 + \mathrm{o}(1)).$$

Downloaded from https://www.cambridge.org/core. University of Bath, on 09 Feb 2020 at 11:24:30, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2019.45

We then use the Stable Law Central Limit Theorem (see e.g. [44, Theorem 4.5.2]) to conclude that

$$\sum_{i\in[n]} D_i^2 \Big/ \left( n^{2/(\tau-1)} \left( \frac{2c}{(\tau-1)(3-\tau)} \Gamma\left(\frac{5}{2} - \frac{1}{2}\tau\right) \cos\left(\frac{\pi(\tau-1)}{4}\right)\right)^{2/(\tau-1)} \right) \xrightarrow{d} \mathcal{S}_{(\tau-1)/2},$$

where $\mathcal{S}_{(\tau-1)/2}$ is a stable random variable. Combining this with (20) results in

$$\frac{a_{\varepsilon_n}(k, G_n)}{n^{(3-\tau)/(\tau-1)}} \xrightarrow{d} \frac{1}{\mu} \left( \frac{2c\Gamma(\frac{5}{2} - \frac{1}{2}\tau)}{(\tau-1)(3-\tau)} \cos\left(\frac{\pi(\tau-1)}{4}\right)\right)^{2/(\tau-1)} \mathcal{S}_{(\tau-1)/2}.$$

To prove the joint convergence of Remark 1, note that, for fixed $m$, the probability that $M_{\varepsilon_n}(k_i)$ satisfies the condition in (13) for all $i \in [m]$ tends to 1, by a proof similar to the proof that $\mathcal{A}_n$ defined in (13) satisfies $\mathbb{P}(\mathcal{A}_n) \to 1$. Thus, we may condition on the event that $M_{\varepsilon_n}(k_i) \geq nk^{1-\tau}/\log(n)$ for all $i \in [m]$. Then, the fact that (20) is the same for all $k_i \ll n^{(\tau-2)/(\tau-1)}/\log(n)$ and it only depends on the degree sequence proves the joint convergence. □

### 3.3. Large $k$

Now we study the value of $a_{\varepsilon_n}(k, G_n)$ when $k \gg n^{(\tau-2)/(\tau-1)}$. We show that there exists a range of degrees $W_n^k(\delta)$ which gives the largest contribution to $a_{\varepsilon_n}(k, G_n)$. For ease of notation, we write $a_{\varepsilon_n}(k)$ for $a_{\varepsilon_n}(k, G_n)$ in this section. We define

$$W_n^k(\delta) = \{u \colon D_u \in [\delta\mu n/k, \mu n/(\delta k)]\}, \tag{21}$$

and we write

$$\begin{aligned} a_{\varepsilon_n}(k) &= \frac{1}{k|M_{\varepsilon_n}(k)|} \sum_{i\in M_{\varepsilon_n}(k)} \sum_{j\in W_n^k(\delta)} D^{(\mathrm{er})}{}_j + \frac{1}{k|M_{\varepsilon_n}(k)|} \sum_{i\in M_{\varepsilon_n}(k)} \sum_{j\notin W_n^k(\delta)} D^{(\mathrm{er})}{}_j \\ &=: a_{\varepsilon_n}(k, W_n^k(\delta)) + a_{\varepsilon_n}(k, \bar{W}_n^k(\delta)), \end{aligned} \tag{22}$$

where $a_{\varepsilon_n}(k, W_n^k(\delta))$ gives the contribution to $a_{\varepsilon_n}(k)$ from vertices in $W_n^k(\delta)$ and $a_{\varepsilon_n}(k, \bar{W}_n(\varepsilon))$ denotes the contribution from vertices not in $W_n^k(\delta)$. In the rest of this section, we prove the following two propositions, which together show that the largest contribution to $a_{\varepsilon_n}(k)$ indeed comes from vertices in $W_n^k(\delta)$.

**Proposition 1.** (Minor contributions.) *There exists $\kappa > 0$ such that, for $k \ll n^{1/(\tau-1)}$,*

$$\limsup_{n\to\infty} \frac{\mathbb{E}[a_{\varepsilon_n}(k, \bar{W}_n^k(\delta))]}{(n/k)^{3-\tau}} = \mathrm{O}(\delta^{\kappa}).$$

**Proposition 2.** (Major contributions.) *For $n^{(\tau-2)/(\tau-1)} \ll k \ll n^{1/(\tau-1)}/\log(n)$,*

$$\frac{a_{\varepsilon_n}(k, W_n^k(\delta))}{(n/k)^{3-\tau}} \xrightarrow{\mathbb{P}} c\mu^{2-\tau} \int_{\delta}^{1/\delta} x^{1-\tau}(1-\mathrm{e}^{-x})\,\mathrm{d}x$$

We now show how these propositions prove part (ii) of Theorem 1.

Downloaded from https://www.cambridge.org/core. University of Bath, on 09 Feb 2020 at 11:24:30, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2019.45

*Proof of Theorem 1(ii).* By the Markov inequality and Proposition 1,

$$\frac{a_{\varepsilon_n}(k, \bar{W}_n^k(\delta))}{(n/k)^{3-\tau}} = \mathrm{O}_{\mathbb{P}}(\delta^{\kappa}).$$

Combining this with Proposition 2 results in

$$\frac{a_{\varepsilon_n}(k)}{(n/k)^{3-\tau}} \xrightarrow{\mathbb{P}} c\mu^{2-\tau} \int_{\delta}^{1/\delta} x^{1-\tau}(1-\mathrm{e}^{-x})\,\mathrm{d}x + \mathrm{O}_{\mathbb{P}}(\delta^{\kappa}). \tag{23}$$

Taking the limit of $\delta \to 0$ then proves the theorem. □

The rest of this section is devoted to proving Propositions 1 and 2.

3.3.1. *Conditional expectation.* We first compute the expectation of $a_{\varepsilon_n}(k, W_n^k(\delta))$ conditionally on the degree sequence and on the event $(D_i)_{i\in[n]} \in \Lambda_n \cap \mathcal{D}_n(\mu n/k)$. Define the event

$$\mathcal{W}_n = \Big\{ \sum_{u \in W_n^k(\delta)} D_u - D^{(\mathrm{er})}{}_u \le |W_n^k(\delta)| n k^{1-\tau} \log(n) \Big\}.$$

Using (16), we obtain

$$\mathbb{E}_n\Big[ \sum_{u\in W_n^k(\delta)} D^{(\mathrm{er})}{}_u - D_u \Big] = \sum_{u\in W_n^k(\delta)} \mathrm{O}(D_u^{\tau-1} n^{2-\tau}) = |W_n^k(\delta)| \mathrm{O}(n k^{1-\tau}),$$

so that $\mathbb{P}(\mathcal{W}_n) \to 1$ by the Markov inequality.

**Leema 2.** *When $k \gg n^{(\tau-2)/(\tau-1)}$, on the event $(D_i)_{i\in[n]} \in \Lambda_n \cap \mathcal{W}_n \cap \mathcal{D}_n(\mu n/k)$,*

$$\mathbb{E}_n[a_{\varepsilon_n}(k, W_n^k(\delta))] = (1+\mathrm{o}(1)) \frac{1}{k} \sum_{u\in W_n^k(\delta)} D_u(1-\mathrm{e}^{-D_u k/L_n}).$$

*Proof.* Let $X_{ij}$ denote the indicator that $i$ and $j$ are connected. By (22),

$$\mathbb{E}_n[a_{\varepsilon_n}(k, W_n^k(\delta))] = \frac{1}{k|M_{\varepsilon_n}(k)|} \sum_{v\in M_{\varepsilon_n}(k)} \sum_{u\in W_n^k(\delta)} D^{(\mathrm{er})}{}_u \mathbb{P}_n(X_{uv}=1).$$

By [24, (4.6) and (4.9)]

$$\begin{aligned}\mathbb{P}_n(X_{uv}=1) &= \prod_{s=0}^{D_i-1}\Big(1 - \frac{D_i}{L_n - 2D_i - 1}\Big) + \mathrm{O}\Big(\frac{D_v^2 D_u + D_u^2 D_v}{L_n^2}\Big) \\ &= 1 - \mathrm{e}^{-D_u D_v/L_n} + \mathrm{O}\Big(\frac{D_v^2 D_u + D_u^2 D_v}{L_n^2}\Big),\end{aligned} \tag{24}$$

since $D_i, D_j = \mathrm{o}(n)$ and $L_n = \mu n(1+\mathrm{o}(1))$ on $\Lambda_n$. Thus, when $D_u \in \mu n/k[\delta, 1/\delta]$ and $v \in M_{\varepsilon_n}(k)$,

$$\mathbb{P}_n(X_{uv}=1) = (1-\mathrm{e}^{-D_u k/L_n})(1+\mathrm{o}(1)).$$

Downloaded from https://www.cambridge.org/core. University of Bath, on 09 Feb 2020 at 11:24:30, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2019.45

Since $\sum_{u\in W_n^k(\delta)} D_u = |W_n^k(\delta)|\Theta(n/k)$ and $n/k \gg nk^{1-\tau}\log(n)$ for $k \gg n^{(\tau-2)/(\tau-1)}$, on the event $\mathcal{W}_n$,

$$\sum_{u\in W_n^k(\delta)} D^{(\mathrm{er})}{}_u = \sum_{u\in W_n^k(\delta)} (D^{(\mathrm{er})}{}_u - D_u) + \sum_{u\in W_n^k(\delta)} D_u = (1+\mathrm{o}(1)) \sum_{u\in W_n^k(\delta)} D_u$$

Because $1-\mathrm{e}^{-D_uk/L_n} \in [1-\mathrm{e}^{-\delta}, 1-\mathrm{e}^{\delta}]$ when $u\in W_n^k(\delta)$, this also shows that

$$\sum_{u\in W_n^k(\delta)} D^{(\mathrm{er})}{}_u(1-\mathrm{e}^{-D_uk/L_n}) = (1+\mathrm{o}(1)) \sum_{u\in W_n^k(\delta)} D_u(1-\mathrm{e}^{-D_uk/L_n}). \tag{25}$$

Thus, we obtain

$$\begin{aligned}\mathbb{E}_n[a_{\varepsilon_n}(k, W_n^k(\delta))] &= (1+\mathrm{o}(1))\frac{1}{k}\sum_{u\in W_n^k(\delta)} D^{(\mathrm{er})}{}_u(1-\mathrm{e}^{-D_uk/L_n})\\ &= (1+\mathrm{o}(1))\frac{1}{k}\sum_{u\in W_n^k(\delta)} D_u(1-\mathrm{e}^{-D_uk/L_n}).\end{aligned}$$

□

3.3.2. *Convergence of conditional expectation.* We shall now show that $\mathbb{E}_n[a_{\varepsilon_n}(k, W_n^k(\delta))]$ of Lemma 2 converges to a constant when we take the i.i.d. degrees into account.

**Leema 3.** *When $k \gg n^{(\tau-2)/(\tau-1)}$, on the event $(D_i)_{i\in[n]} \in \Lambda_n \cap \mathcal{W}_n \cap \mathcal{D}_n(\mu n/k)$,*

$$\frac{\mathbb{E}_n[a_{\varepsilon_n}(k, W_n^k(\delta))]}{n^{3-\tau}k^{\tau-3}} \xrightarrow{\mathbb{P}} c\mu^{2-\tau}\int_{\delta}^{1/\delta} x^{1-\tau}(1-\mathrm{e}^{-x})\,\mathrm{d}x.$$

*Proof.* First of all, notice that for $k \gg n^{(\tau-2)/(\tau-1)}$

$$\begin{aligned}\mathbb{P}(D\in[a,b]\mu n/k) &= (1+\mathrm{o}(1))\sum_{t=\lceil a\mu n/k\rceil}^{\lfloor b\mu n/k\rfloor} ct^{-\tau} \le (1+\mathrm{o}(1))\int_{a\mu n/k-1}^{b\mu n/k} cx^{-\tau}\,\mathrm{d}x\\ &= (1+\mathrm{o}(1))\frac{c}{\tau-1}\big((b\mu n/k)^{1-\tau} - (a\mu n/k-1)^{1-\tau}\big)\\ &= (1+\mathrm{o}(1))\frac{c}{\tau-1}\big((b\mu n/k)^{1-\tau} - (a\mu n/k)^{1-\tau}\big)\\ &= (1+\mathrm{o}(1))\int_{a\mu n/k}^{b\mu n/k} cx^{-\tau}\,\mathrm{d}x.\end{aligned}$$

Similarly,

$$\begin{aligned}\mathbb{P}(D\in[a,b]\mu n/k) &\ge (1+\mathrm{o}(1))\int_{\lceil a\mu n/k\rceil}^{\lfloor b\mu n/k\lfloor+1} cx^{-\tau}\,\mathrm{d}x\\ &\ge (1+\mathrm{o}(1))\int_{a\mu n/k+1}^{b\mu n/k} cx^{-\tau}\,\mathrm{d}x\\ &= (1+\mathrm{o}(1))\int_{a\mu n/k}^{b\mu n/k} cx^{-\tau}\,\mathrm{d}x,\end{aligned}$$

Downloaded from https://www.cambridge.org/core. University of Bath, on 09 Feb 2020 at 11:24:30, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2019.45

so that

$$\mathbb{P}(D \in [a, b]\mu n/k) = (1 + o(1)) \int_{a\mu n/k}^{b\mu n/k} cx^{-\tau}\,\mathrm{d}x \quad \text{when } k \gg n^{(\tau-2)/(\tau-1)}.$$

Define the random measure

$$M^{(n)}[a, b] = \frac{1}{\mu^{1-\tau} n^{2-\tau} k^{\tau-1}} \sum_{u\in[n]} \mathbf{1}_{\{D_u \in [a,b]\mu n/k\}}. \tag{26}$$

Since the degrees are i.i.d. samples from (3), the number of vertices with degrees in interval $[a, b]$ is binomially distributed. Then,

$$\begin{aligned} M^{(n)}[a, b] &= \frac{1}{\mu^{1-\tau} n^{2-\tau} k^{\tau-1}} |\{u \colon D_u \in [a, b]\mu n/k\}| \\ &= (1 + o_{\mathbb{P}}(1)) \frac{\mathbb{P}(D \in [a, b]\mu n/k)}{\mu^{1-\tau} n^{2-\tau} k^{\tau-1}} \\ &= \frac{1 + o_{\mathbb{P}}(1)}{(\mu n)^{1-\tau} k^{\tau-1}} \int_{a\mu n/k}^{b\mu n/k} cx^{-\tau}\,\mathrm{d}x \xrightarrow{\mathbb{P}} \int_a^b cy^{-\tau}\,\mathrm{d}y \\ &=: \lambda[a, b], \end{aligned}$$

where we used the change of variables $y = xk/(\mu n)$. By Lemma 2,

$$\begin{aligned} \mathbb{E}_n[a_{\varepsilon_n}(k, W_n^k(\delta))] &= \frac{\sum_{u\in W_n^k(\delta)} D_u(1 - \mathrm{e}^{-D_u k/L_n})}{k}(1 + o(1)) \\ &= \frac{\mu n}{k} \frac{\sum_{u\in W_n^k(\delta)} (D_u k/(\mu n))(1 - \mathrm{e}^{-D_u k/(\mu n)})}{k}(1 + o(1)) \\ &= \frac{\mu^{2-\tau} n^{3-\tau}}{k^{3-\tau}} \int_\delta^{1/\delta} t(1 - \mathrm{e}^{-t})\,\mathrm{d}M^{(n)}(t)(1 + o(1)). \end{aligned} \tag{27}$$

Fix $\eta > 0$. Since $t(1 - \mathrm{e}^{-t})$ is bounded and continuous on $[\delta, 1/\delta]$, we can find $m < \infty$, disjoint intervals $(B_i)_{i\in[m]}$ and constants $(b_i)_{i\in[m]}$ such that $\cup B_i = [\delta, 1/\delta]$ and

$$\left| t(1 - \mathrm{e}^{-t}) - \sum_{i=1}^m b_i \mathbf{1}_{\{t\in B_i\}} \right| < \eta/\lambda([\delta, 1/\delta]),$$

for all $t \in [\delta, 1/\delta]$. Because $M^{(n)}(B_i) \xrightarrow{\mathbb{P}} \lambda(B_i)$ for all $i$,

$$\lim_{n\to\infty} \mathbb{P}(|M^{(n)}(B_i) - \lambda(B_i)| > \eta/(mb_i)) = 0.$$

Furthermore,

$$\begin{aligned} &\left| \int_\delta^{1/\delta} t(1 - \mathrm{e}^{-t})\,\mathrm{d}M^{(n)}(t) - \int_\delta^{1/\delta} t(1 - \mathrm{e}^{-t})\,\mathrm{d}\lambda(t) \right| \\ &\quad\le \left| \int_\delta^{1/\delta} t(1 - \mathrm{e}^{-t}) - \sum_{i=1}^m b_i \mathbf{1}_{\{t\in B_i\}}\,\mathrm{d}M^{(n)}(t) \right| \end{aligned}$$

Downloaded from https://www.cambridge.org/core. University of Bath, on 09 Feb 2020 at 11:24:30, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2019.45

$$+\left|\int_\delta^{1/\delta} t(1-\mathrm{e}^{-t})-\sum_{i=1}^m b_i\mathbf{1}_{\{t\in B_i\}}\,\mathrm{d}\lambda(t)\right|$$
$$+\left|\int_\delta^{1/\delta}\sum_{i=1}^m b_i\mathbf{1}_{\{t\in B_i\}}\,\mathrm{d}M^{(n)}(t)-\int_\delta^{1/\delta}\sum_{i=1}^m b_i\mathbf{1}_{\{t\in B_i\}}\,\mathrm{d}\lambda(t)\right|. \tag{28}$$

Using the fact that

$$\int_\delta^{1/\delta}\mathbf{1}_{\{t\in B_i\}}\,\mathrm{d}M^{(n)}(t)=M^{(n)}(B_i)$$

yields

$$\begin{aligned}&\left|\int_\delta^{1/\delta}\sum_{i=1}^m b_i\mathbf{1}_{\{t\in B_i\}}\,\mathrm{d}M^{(n)}(t)-\int_\delta^{1/\delta}\sum_{i=1}^m b_i\mathbf{1}_{\{t\in B_i\}}\,\mathrm{d}\lambda(t)\right|\\&\quad=\left|\sum_{i=1}^m b_i(M^{(n)}(B_i)-\lambda(B_i))\right|\\&\quad=\left|\sum_{i=1}^m \mathrm{o}_{\mathbb{P}}(\eta/m)\right|=\mathrm{o}_{\mathbb{P}}(\eta).\end{aligned}$$

Thus, (3.3.2) results in

$$\left|\int_\delta^{1/\delta} t(1-\mathrm{e}^{-t})\,\mathrm{d}M^{(n)}(t)-\int_\delta^{1/\delta} t(1-\mathrm{e}^{-t})\,\mathrm{d}\lambda(t)\right|\le\eta\frac{M^{(n)}([\delta,1/\delta])}{\lambda([\delta,1/\delta])}+\eta+\mathrm{o}_{\mathbb{P}}(\eta).$$

Using the fact that $M^{(n)}([\delta,1/\delta])=\mathrm{O}_{\mathbb{P}}(\lambda([\delta,1/\delta]))$ proves that

$$\int_\delta^{1/\delta} t(1-\mathrm{e}^{-t})\,\mathrm{d}M^{(n)}(t)\xrightarrow{\mathbb{P}}\int_\delta^{1/\delta} t(1-\mathrm{e}^{-t})\,\mathrm{d}\lambda(t)=c\int_\delta^{1/\delta} x^{1-\tau}(1-\mathrm{e}^{-x})\,\mathrm{d}x, \tag{29}$$

which together with (27) proves the lemma. □

3.3.3. *Conditional variance of* $a(k)$. We now show that the variance of $a_{\varepsilon_n}(k, W_n^k(\delta))$ is small when conditioning on the degree sequence, so that $a_{\varepsilon_n}(k, W_n^k(\delta))$ concentrates around its expected value computed in Lemma 2. Define the event

$$\mathcal{K}_n=\{|W_n^k(\delta)|\le n(n/k)^{1-\tau}\log(n)\}.$$

Since the degrees are i.i.d. samples from (3), $|W_n^k(\delta)|$ is distributed as a binomial with parameters $(n, C(n/k)^{1-\tau})$ for some constant $C$. Therefore, $\mathbb{P}(\mathcal{K}_n)\to 1$.

**Leema 4.** *When* $n^{(\tau-2)/(\tau-1)}\ll k\ll n^{1/(\tau-1)}/\log(n)$, *on the event*

$$(D_i)_{i\in[n]}\in\Lambda_n\cap\mathcal{D}_n(\mu n/k)\cap\mathcal{W}_n\cap\mathcal{K}_n,$$

*we have*

$$\frac{\mathrm{Var}_n(a_{\varepsilon_n}(k, W_n^k(\delta)))}{\mathbb{E}_n[a_{\varepsilon_n}(k, W_n^k(\delta))]^2}\to 0.$$

Downloaded from https://www.cambridge.org/core. University of Bath, on 09 Feb 2020 at 11:24:30, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2019.45

*Proof.* We write the variance of $a_{\varepsilon_n}(k, W_n^k(\delta))$ as

$$\mathrm{Var}_n(a_{\varepsilon_n}(k, W_n^k(\delta))) = \frac{1}{k^2|M_{\varepsilon_n}(k)|^2} \sum_{i,j\in M_{\varepsilon_n}(k)} \sum_{u,v\in W_n^k(\delta)} D^{(\mathrm{er})}{}_u D^{(\mathrm{er})}{}_w \\ \times (\mathbb{P}_n(X_{iu} = X_{jv} = 1) - \mathbb{P}_n(X_{iu} = 1)\mathbb{P}_n(X_{jv} = 1)) \qquad (30)$$

Equation (30) splits into various cases, depending on the size of $\{i, j, u, v\}$. We denote the contribution of $|\{i, j, u, v\}| = r$ to the variance by $V^{(r)}(k)$. We first consider $V^{(4)}(k)$. We can write

$$\mathbb{P}_n(X_{iu} = X_{jv} = 0) = \mathbb{P}_n(X_{iu} = 0)\mathbb{P}_n(X_{jv} = 0 \mid X_{iu} = 0).$$

For the second term, we first pair all half-edges adjacent to vertex $i$, conditionally on not pairing to vertex $u$. Then the second term can be interpreted as the probability that vertex $j$ does not pair to vertex $v$ in a new configuration model with $\hat{L}_n = L_n - D_i = L_n(1 + \mathrm{O}(n^{-(\tau-2)/(\tau-1)} \log(n)))$ half-edges, where the new degree of vertex $j$ is reduced by the number of half-edges from vertex $i$ that paired to $j$, $\hat{X}_{ij}$. Similarly, the new degree of vertex $v$ is reduced by the amount of half-edges from vertex $i$ that paired to $v$, $\hat{X}_{iv}$. Since $\hat{X}_{ij}$ can be dominated by a binomial random variable with parameters $D_i$, $D_j/L_n$,

$$\mathbb{P}(\hat{X}_{ij} > D_j n^{-(\tau-2)/(\tau-1)}) \le \mathrm{e}^{-D_j n^{-(\tau-2)/(\tau-1)}/3},$$

and a similar statement holds for $\hat{X}_{iv}$. Let $\mathcal{T}_n$ denote the event that $\hat{X}_{ij} \le D_j n^{-(\tau-2)/(\tau-1)}$ and $\hat{X}_{iv} \ge D_v n^{-(\tau-2)/(\tau-1)}$.

$$\mathbb{P}_n(X_{jv} = 0 \mid X_{iu} = 0) = \mathbb{P}_n(X_{jv} = 0 \mid X_{iu} = 0, \mathcal{T}_n)\mathbb{P}_n(\mathcal{T}_n) + \mathrm{O}(\mathbb{P}_n(\mathcal{T}_n^c)).$$

On the event $\mathcal{T}_n$ the new degree of vertex $j$ is $\hat{D}_j = D_j(1 + \mathrm{O}(n^{-(\tau-2)/(\tau-1)}))$, and a similar statement holds for vertex $v$. Furthermore, since $i, j \in M_{\varepsilon_n}(k)$, $n^{(\tau-2)/(\tau-1)} \ll D_i, D_j \ll n^{1/(\tau-1)}$ and since $u, v \in W_n^k(\delta)$, $n^{(\tau-2)/(\tau-1)} \ll D_u, D_v \ll n^{1/(\tau-1)}$ as well. Thus,

$$\mathbb{P}_n(\mathcal{T}_n^c) = \mathrm{O}(\mathrm{e}^{-D_j n^{-(\tau-2)/(\tau-1)}} + \mathrm{e}^{-D_v n^{-(\tau-2)/(\tau-1)}}) = \mathrm{o}(\mathrm{e}^{-D_j D_v/L_n})$$

Furthermore, by (24),

$$\mathbb{P}_n(X_{jv} = 0 \mid X_{iu} = 0, \mathcal{T}_n) = \mathrm{e}^{-\hat{D}_j\hat{D}_v/\hat{L}_n} + \mathrm{O}(n^{-(\tau-2)/(\tau-1)} \log(n)) \\ = \mathrm{e}^{-D_j D_v/L_n}(1 + \mathrm{o}(1)),$$

so that

$$\mathbb{P}_n(X_{iu} = X_{jv} = 0) \quad = \mathrm{e}^{-D_i D_u/L_n}\, \mathrm{e}^{-D_j D_v/L_n}(1 + \mathrm{o}(1)),$$

using the fact that $D_i D_u, D_j D_v = \mathrm{O}(n)$. This results in

$$\mathbb{P}_n(X_{iu} = X_{jv} = 1) = 1 - \mathbb{P}_n(X_{iu} = 0) - \mathbb{P}_n(X_{jv} = 0) + \mathbb{P}_n(X_{iu} = X_{jv} = 0) \\ = 1 + (-\mathrm{e}^{-D_u k/L_n} - \mathrm{e}^{-D_v k/L_n} + \mathrm{e}^{-D_u k/L_n - D_v k/L_n})(1 + \mathrm{o}(1)) \\ = (1 - \mathrm{e}^{-D_u k/L_n})(1 - \mathrm{e}^{-D_v k/L_n})(1 + \mathrm{o}(1)),$$

Downloaded from https://www.cambridge.org/core. University of Bath, on 09 Feb 2020 at 11:24:30, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2019.45

where the last equality holds because $D_u k = \Theta(n)$ and $D_v k = \Theta(n)$ for $u, v \in W_n^k(\delta)$. Since $(1-\mathrm{e}^{-D_u k/L_n})(1-\mathrm{e}^{-D_v k/L_n}) \in [\delta f(n), f(n)/\delta]$ for some $f(n)$ when $D_u, D_v \in W_n^k(\delta)$, these error terms are uniform in $i, j \in M_{\varepsilon_n}(k)$ as well as $u, v \in W_n^k(\varepsilon)$. Therefore

$$\begin{aligned} V^{(4)}(k) &= \frac{1}{|M_{\varepsilon_n}(k)|^2 k^2} \sum_{i,j\in M_{\varepsilon_n}(k)} \sum_{u,v\in W_n^k(\delta)} D^{(\mathrm{er})}{}_u D^{(\mathrm{er})}{}_v (1-\mathrm{e}^{-D_u k/L_n})(1-\mathrm{e}^{-D_v k/L_n})(1+\mathrm{o}(1)) \\ &\quad - D^{(\mathrm{er})}{}_u D^{(\mathrm{er})}{}_v (1-\mathrm{e}^{-D_u k/L_n})(1-\mathrm{e}^{-D_u k/L_n})(1+\mathrm{o}(1)) \\ &= \sum_{u,v\in W_n^k(\delta)} \mathrm{o}(k^{-2} D_u D_v (1-\mathrm{e}^{-D_u k/L_n})(1-\mathrm{e}^{-D_v k/L_n})) \\ &= \mathrm{o}(\mathbb{E}_n[a_{\varepsilon_n}(k, W_n^k(\delta))]^2), \end{aligned}$$

where we replaced $D^{(\mathrm{er})}$ with $D$ as in (25) and the last equality follows from Lemma 2. Since there are no overlapping edges when $\{i, j, u, v\} = 3$, $V^{(3)}(k)$ can be bounded similarly.

We then consider the contribution from $V^{(2)}$, which is the contribution where the two edges are the same. By Lemma 3, we have to show that this contribution is small compared to $n^{6-2\tau} k^{2\tau-6}$. We bound the summand in (30) as

$$D_u^2(\mathbb{P}_n(X_{iu}=1) - \mathbb{P}_n(X_{iu}=1)^2) \le D_u^2.$$

Thus, using the fact that on $\mathcal{A}_n$, $|M_{\varepsilon_n}(k)| \ge 1$, $V^{(2)}$ can be bounded as

$$V^{(2)} \le \frac{1}{k^2|M_{\varepsilon_n}(k)|^2} \sum_{i\in M_{\varepsilon_n}(k)} \sum_{u\in W_n^k(\delta)} D_u^2 = \frac{1}{k^2|M_{\varepsilon_n}(k)|} \sum_{u\in W_n^k(\delta)} D_u^2 = \mathrm{O}(n^2k^{-4})|W_n^k(\delta)|.$$

Thus, on the event $\mathcal{K}_n$,

$$V^{(2)} = \mathrm{O}_{\mathbb{P}}(n^{4-\tau} k^{\tau-5}),$$

which is smaller than $n^{6-2\tau} k^{2\tau-6}$ when $k \gg n^{(\tau-2)/(\tau-1)}$, as required. □

*Proof of Proposition 2.* Since $\mathbb{P}(\Lambda_n \cap \mathcal{D}_n(\mu n/k) \cap \mathcal{W}_n \cap \mathcal{K}_n) \to 1$, Lemma 4 together with the Chebyshev inequality shows that

$$\frac{a_{\varepsilon_n}(k, W_n^k(\delta))}{\mathbb{E}_n[a_{\varepsilon_n}(k, W_n^k(\delta))]} \xrightarrow{\mathbb{P}} 1.$$

Combining this with Lemmas 2 and 3 yields

$$\frac{a_{\varepsilon_n}(k, W_n^k(\delta))}{n^{3-\tau}k^{\tau-3}} \xrightarrow{\mathbb{P}} c\mu^{2-\tau} \int_\delta^{1/\delta} x^{1-\tau}(1-\mathrm{e}^{-x})\,\mathrm{d}x.$$

□

3.3.4. *Contributions outside* $W_n^k(\delta)$. In this section, we prove Proposition 1 and show that the contribution to $a_{\varepsilon_n}(k)$ outside of the major contributing regimes as described in (21) is negligible.

*Proof of Proposition 1.* We use the fact that

$$\mathbb{P}_n(X_{ij}=1) \le \min\left(1, \frac{D_i D_l}{L_n}\right).$$

Downloaded from https://www.cambridge.org/core. University of Bath, on 09 Feb 2020 at 11:24:30, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2019.45

This yields

$$\begin{aligned}\mathbb{E}[a_{\varepsilon_n}(k,\bar{W}_n^k(\delta))] &= \mathbb{E}[\mathbb{E}_n[a_{\varepsilon_n}(k,\bar{W}_n^k(\delta))]] \\ &\leq \frac{n}{k}\mathbb{E}\Big[D\min\Big(1,\frac{kD}{L_n}\Big)\mathbf{1}_{\{D\in\bar{W}_n^k(\delta)\}}\Big] \\ &\leq K\frac{n}{k}\int_0^{\delta\mu n/k} x^{1-\tau}\min\Big(1,\frac{kx}{\mu n}\Big)\,\mathrm{d}x + K\frac{n}{k}\int_{\mu n/(\delta k)}^{\infty} x^{1-\tau}\min\Big(1,\frac{kx}{\mu n}\Big)\,\mathrm{d}x, \quad (31)\end{aligned}$$

for some $K > 0$. For ease of notation, we assume that $\mu = 1$. We have to show that the contribution to (31) from vertices $u$ such that $D_u < \delta n/k$ or $D_u > n/(\delta k)$ is small. First, we study the contribution to (31) for $D_u < \delta n/k$. We can bound this contribution by taking the second term of the minimum, which bounds the contribution as

$$\int_0^{\delta n/k} x^{2-\tau}\,\mathrm{d}x = \frac{\delta^{3-\tau}}{\tau-3}(k/n)^{\tau-3}.$$

Then, we study the contribution for $D_u > n/(k\varepsilon)$. This contribution can be bounded by taking 1 for the minimum in (31):

$$\frac{n}{k}\int_{n/(\delta k)}^{\infty} x^{1-\tau}\,\mathrm{d}x = \frac{\delta^{\tau-2}}{\tau-2}(k/n)^{\tau-3}.$$

Taking $\kappa = \min(\tau-2, 3-\tau) > 0$ then proves the proposition. □

### 3.4. Expected average nearest-neighbour degree

As in (22), we can write

$$\mathbb{E}[a(k,G_n)] = \mathbb{E}[a(k,W_n^k(\delta))] + \mathbb{E}[a(k,\bar{W}_n^k(\delta))]. \quad (32)$$

By Proposition 1,

$$\mathbb{E}[a(k,\bar{W}_n^k(\delta))]/(n/k)^{\tau-3} = \mathrm{O}(\delta^{\kappa}).$$

We now focus on the first term. The expected degree of a neighbour of a randomly chosen vertex of degree $k$ can be written as

$$\begin{aligned}\mathbb{E}[a(k,W_n^k(\delta))] &= \mathbb{E}[D^{(\mathrm{er})}{}_{\mathcal{N}_{V_k}(U)}\mathbf{1}_{\{D_{\mathcal{N}_{V_k}(U)}\in[\delta,1/\delta]\mu n/k\}}] \\ &= \mathbb{E}[D_{\mathcal{N}_{V_k}(U)}\mathbf{1}_{\{D_{\mathcal{N}_{V_k}(U)}\in[\delta,1/\delta]\mu n/k\}}](1+\mathrm{o}(1)),\end{aligned}$$

where $\mathcal{N}_{V_k}(U)$ denotes a uniformly chosen neighbour of a vertex of degree $k$. By (24), we can write the connection probability between a vertex of degree $k$ and a neighbour of degree $d \in [\delta, 1/\delta]\mu n/k$ as $1-\mathrm{e}^{-kd/(\mu n)}(1+\mathrm{o}(1))$. Therefore

$$\begin{aligned}\mathbb{E}[a(k,W_n^k(\delta))] &= (1+\mathrm{o}(1))\int_{\delta\mu n/k}^{\mu n/(\delta k)} cx^{1-\tau}(1-\mathrm{e}^{-xk/(\mu n)})\,\mathrm{d}x \\ &= (1+\mathrm{o}(1))(n/k)^{3-\tau}c\mu^{2-\tau}\int_{\delta}^{1/\delta} x^{1-\tau}(1-\mathrm{e}^{-x})\,\mathrm{d}x.\end{aligned}$$

Combining this with (32) and Proposition 1, and letting first $n \to \infty$ and then $\delta \to 0$, proves (11).

Downloaded from https://www.cambridge.org/core. University of Bath, on 09 Feb 2020 at 11:24:30, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2019.45

## 4. Proofs of Theorem 2 and 3

We now briefly show how the proof of Theorem 1 can be adapted for the rank-1 inhomogeneous random graph and the hyperbolic random graph to prove Theorems 2 and 3. We let $\mathbb{P}_n$ denote the probability conditioned on the weights in the rank-1 inhomogeneous random graph or conditioned on the radial coordinates in the hyperbolic model.

### 4.1. Inhomogeneous random graph

First, we show how to prove Theorem 2(i). In the rank-1 inhomogeneous random graph, the degree of vertex $i$ with weight $h_i$ satisfies $D_i = h_i(1 + o_{\mathbb{P}}(1))$ when $h_i \gg 1$ [41]. We condition on the event that the largest weight is smaller than $n^{1/(\tau-1)} \log(n)$, which happens with high probability. Then, when $h \ll n^{(\tau-2)/(\tau-1)}/\log(n)$, $p(h, h') = hh'/(\mu n)$ for all vertices. When $u \in M_{\varepsilon_n}(k)$, $h_u = k(1 + o_{\mathbb{P}}(1))$, so that conditionally on the weight sequence

$$\begin{aligned} a_{\varepsilon_n}(k) &= \frac{1}{k|M_{\varepsilon_n}(k)|} \sum_{u \in M_{\varepsilon_n}(k)} \sum_{i \in [n]} D_i \mathbb{P}_n(X_{iu} = 1) \\ &= (1 + o_{\mathbb{P}}(1)) \frac{1}{k} \sum_{i \in [n]} h_i \frac{h_i k}{\mu n} \\ &= (1 + o_{\mathbb{P}}(1)) \sum_{i \in [n]} \frac{h_i^2}{\mu n}, \end{aligned}$$

which is equivalent to (20) because the weights are also sampled from (3). This proves Theorem 2(i).

Similarly to (21), we define for the rank-1 inhomogeneous random graph

$$W_n^{k,\mathrm{HVM}}(\delta) = \{u \colon h_u \in [\delta \mu n/k, \mu n/(\delta k)]\}. \tag{33}$$

Then it is easy to show that Proposition 1 also holds for the rank-1 inhomogeneous random graph with (33) instead of $W_n^k(\delta)$. Because the weights are sampled from (3) and $\mathbb{P}_n(X_{ij} = 1) = \min(h_i h_j/(\mu n), 1)$, (31) and therefore also Proposition 1 hold for the rank-1 inhomogeneous random graph as well.

We now sketch how to adjust the proof of Proposition 2 to prove an analogous version for the rank-1 inhomogeneous random graph, which states that

$$\frac{a_{\varepsilon_n}(k, W_n^{k\mathrm{HVM}}(\delta))}{(n/k)^{3-\tau}} \xrightarrow{\mathbb{P}} c\mu^{2-\tau} \int_\delta^{1/\delta} x^{1-\tau} \min(x, 1)\,\mathrm{d}x. \tag{34}$$

Following the proofs of Lemmas 2–4, we see that these lemmas also hold for the rank-1 inhomogeneous random graph if we replace the connection probability of the erased configuration model of $1 - \mathrm{e}^{-D_i D_j/L_n}$ with $\min(h_i h_j/\mu n, 1)$. Note that for the rank-1 inhomogeneous random graph the contribution to (30) from three or four different vertices is 0, because the edge probabilities in the rank-1 inhomogeneous random graph conditioned on the weights are independent. From these lemmas, (34) follows. This then shows similarly to (23) that

$$\frac{a_{\varepsilon_n}(k)}{(n/k)^{3-\tau}} \xrightarrow{\mathbb{P}} c\mu^{2-\tau} \int_0^\infty x^{1-\tau} \min(x, 1)\,\mathrm{d}x = \frac{c\mu^{2-\tau}}{(3-\tau)(\tau-2)},$$

which proves Theorem 2(ii).

Downloaded from https://www.cambridge.org/core. University of Bath, on 09 Feb 2020 at 11:24:30, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2019.45

### 4.2. Hyperbolic random graph

We first provide a lemma that gives the connection probabilities conditioned on the radial coordinates in the hyperbolic random graph. Similarly to (13), let $\mathcal{A}_n$ denote the event that the maximal type is smaller than $n^{1/(\tau-1)} \log(n)$. Because $t(u)$ is distributed as (10), $\mathbb{P}(\mathcal{A}_n) \to 1$.

**Leema 5.** *Denote*

$$g(x) = \begin{cases} \dfrac{2}{\pi} \sin^{-1}(x) & x < 1, \\ 1 & x \geq 1. \end{cases}$$

*Then the probability that u and v are connected in a hyperbolic random graph conditionally on the radial coordinates and on the event $\mathcal{A}_n$, equals*

$$\mathbb{P}_n(X_{uv} = 1) = g(\nu t(u)t(v)/n)(1 + \mathrm{o}(1)), \tag{35}$$

*where the* $\mathrm{o}(1)$ *term is uniform in u and v.*

*Proof.* Suppose $\nu t(u)t(v)/n \geq 1$. Then,

$$1 \leq \frac{\nu t(u)t(v)}{n} = \frac{\nu\, \mathrm{e}^R\, \mathrm{e}^{-(r_u+r_v)/2}}{n} = \frac{n}{\nu}\, \mathrm{e}^{-(r_u+r_v)/2},$$

so that $r_u + r_v \leq 2\log(n/\nu) = R$. Thus, by (9),

$$\begin{aligned} \cosh(d(u, v)) &= \cosh(r_u)\cosh(r_v) - \sinh(r_u)\sinh(r_v)\cos(\theta_{uv}) \\ &\leq \cosh(r_u + r_v) \leq \cosh(R), \end{aligned}$$

so that the distance between $u$ and $v$ is less than $R$ and $u$ and $v$ are connected.

Now suppose that $\nu t(u)t(v)/n < 1$, so that $r_u + r_v > R$. We calculate the maximal value of $\theta_{uv}$ such that $u$ and $v$ are connected, which we denote by $\theta^*_{uv}$. When the angle between $u$ and $v$ equals $\theta^*_{uv}$, the hyperbolic distance between $u$ and $v$ is precisely $R$. Thus, using the definition of the hyperbolic sine and cosine, we obtain

$$\frac{\mathrm{e}^R - \mathrm{e}^{-R}}{2} = \frac{\mathrm{e}^{r_u} - \mathrm{e}^{-r_u}}{2}\frac{\mathrm{e}^{r_v} - \mathrm{e}^{-r_v}}{2} - \frac{\mathrm{e}^{r_u} + \mathrm{e}^{-r_u}}{2}\frac{\mathrm{e}^{r_v} + \mathrm{e}^{-r_v}}{2}\cos(\theta^*_{uv}). \tag{36}$$

Since, on the event $\mathcal{A}_n$, the maximal type is at most $n^{1/(\tau-1)}\log(n)$,

$$\mathrm{e}^{r_u - r_v} = (t(v)/t(u))^2 = \mathrm{O}(n^{2/(\tau-1)}\log^2(n))$$

uniformly in $u$ and $v$. Similarly,

$$\mathrm{e}^{r_v - r_u} = \mathrm{O}(n^{2/(\tau-1)}\log^2(n)) \quad \text{and} \quad \mathrm{e}^{-r_u - r_v} \leq \mathrm{e}^{r_u - r_v} = \mathrm{O}(n^{2/(\tau-1)}\log^2(n))$$

uniformly in $u$ and $v$. Furthermore, $\mathrm{e}^{-R} = \mathrm{O}(n^{-2})$, so that (36) becomes

$$\frac{1}{2}\mathrm{e}^R + \mathrm{O}(n^{-2}) = \frac{1}{4}\mathrm{e}^{r_u + r_v}(1 - \cos(\theta^*_{uv})) + \mathrm{O}(n^{2/(\tau-1)}\log^2(n)). \tag{37}$$

By the definitions of $t(u)$, $t(v)$ and $R$

$$\mathrm{e}^{r_u + r_v} = \mathrm{e}^R(\mathrm{e}^{(r_u + r_v - R)/2})^2 = \mathrm{e}^R\left(\frac{n}{\nu t(u)t(v)}\right)^2.$$

Downloaded from https://www.cambridge.org/core. University of Bath, on 09 Feb 2020 at 11:24:30, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2019.45

This yields for (37) that, uniformly in $u$ and $v$,

$$1-\cos(\theta^*_{uv}) = 2\left(\frac{\nu t(u)t(v)}{n}\right)^2 + \mathrm{O}\left(n^{2/(\tau-1)-2}\log^2(n)\frac{\nu^2 t(u)^2 t(v)^2}{n^2}\right)$$
$$= 2\left(\frac{\nu t(u)t(v)}{n}\right)^2 (1+O(n^{-2(\tau-2)/(\tau-1)}\log^2(n))),$$

so that

$$\theta^*_{uv} = \cos^{-1}(1-2(\nu t(u)t(v)/n)^2)(1+\mathrm{o}(1)).$$

Because $u$ and $v$ are connected if their relative angle is at most $\theta^*_{uv}$ and the angular coordinates of $u$ and $v$ are sampled uniformly, we obtain that

$$\mathbb{P}_n(X_{uv}=1) = \frac{1}{\pi}\cos^{-1}(1-2(\nu t(u)t(v)/n)^2)(1+\mathrm{o}(1))$$
$$= \frac{2}{\pi}\sin^{-1}(\nu t(u)t(v)/n)(1+\mathrm{o}(1)).$$

□

Using this lemma, we now prove Theorem 3.

*Proof of Theorem 2.* We first focus on $k \ll n^{(\tau-2)/(\tau-1)}/\log(n)$ and condition on the event $\mathcal{A}_n$. By [22, Section 4.3], for vertex $u$ with radial coordinate $r_u$,

$$\mathbb{E}_n[D_u] = (n-1)\frac{2\alpha \mathrm{e}^{-r_u/2}}{\pi(\alpha-1/2)}(1+\mathrm{O}(\mathrm{e}^{-r_u})) = \frac{2\nu(\tau-1)}{\pi(\tau-2)}t(u)(1+\mathrm{O}((t(u)/n)^2)), \tag{38}$$

where we used that $\alpha = (\tau-1)/2$, $t(u) = \mathrm{e}^{-(R-r_u)/2}$ and $R = \log(n/\nu)$. By [12, Theorem 2.7], for every $\tau$ and $\nu$, we can interpret the hyperbolic random graph as a variant of the geometric inhomogeneous random graph, defined in [12], where every vertex $u$ has weight $w_u = t(u)$. By [12, Lemma 3.5(ii)], $D_u = \mathbb{E}_n[D_u](1+\mathrm{o}_{\mathbb{P}}(1))$ in geometric inhomogeneous random graphs when $w_u = t(u) \gg 1$. Combining this with (38) we obtain that, in the hyperbolic random graph,

$$D_u = \frac{2\nu(\tau-1)}{\pi(\tau-2)}t(u)(1+\mathrm{o}_{\mathbb{P}}(1)) \tag{39}$$

when $1 \ll t(u) \ll n$. When $u \in M_{\varepsilon_n}(k)$, $D_u = k(1+\mathrm{o}(1))$. Using (39) then shows that

$$t(u) = \frac{\pi(\tau-2)}{2\nu(\tau-1)}k(1+\mathrm{o}_{\mathbb{P}}(1))$$

when $k \gg 1$ and $u \in M_{\varepsilon_n}(k)$. On the event $\mathcal{A}_n$, the largest type is $\mathrm{O}(n^{1/(\tau-1)}\log(n))$. Therefore, if $u \in M_{\varepsilon_n}(k)$, then $t(u)t(v)/n = \mathrm{o}(1)$ for all $v$. Applying that $\sin^{-1}(x) = x + \mathrm{O}(x^2)$ to (35) then shows that, for $u \in M_{\varepsilon_n}(k)$,

$$\mathbb{P}_n(X_{uv}=1) = \frac{2\nu t(u)t(v)}{\pi n}(1+\mathrm{o}(1)) = \frac{\tau-2}{\tau-1}\frac{kt(v)}{n}(1+\mathrm{o}_{\mathbb{P}}(1)).$$

Downloaded from https://www.cambridge.org/core. University of Bath, on 09 Feb 2020 at 11:24:30, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2019.45

Thus, conditionally on the types,

$$\begin{aligned}
a_{\varepsilon_n}(k) &= \frac{1}{k|M_{\varepsilon_n}(k)|}\sum_{u\in M_{\varepsilon_n}(k)}\sum_{v\in[n]} D_v\mathbb{P}_n(X_{uv}=1)\\
&= (1+\mathrm{o}_{\mathbb{P}}(1))\frac{2\nu(\tau-1)}{\pi(\tau-2)k}\sum_{v\in[n]} t(v)\frac{t(v)k(\tau-2)}{(\tau-1)n}\\
&= (1+\mathrm{o}_{\mathbb{P}}(1))\sum_{v\in[n]}\frac{2\nu t(v)^2}{\pi n}.
\end{aligned}$$

Combining this with the power-law distribution of the types (10) proves Theorem 3(i), which is the same as Theorem 2(i) where $\mu$ is replaced with $\pi/(2\nu)$ and $c/(\tau-1)$ with 1.

We now investigate the case $k \gg n^{(\tau-2)/(\tau-1)}$. Similarly to (21), we define for the hyperbolic random graph

$$W_n^{k,\mathrm{HRG}}(\delta)=\{u\colon t(u)\in[\delta\zeta n/k,\ \zeta n/(\delta k)]\}$$

with $\zeta=2(\tau-1)/(\pi(\tau-2))$. Using the fact that $2\sin^{-1}(x)/\pi\le x$ combined with Lemma 5, we obtain

$$\mathbb{P}_n(X_{uv}=1)\le\min\,(2\nu t(u)t(v)/(\pi n),\ 1).$$

Combining this with the fact that the $t(u)$s are sampled from a distribution similar to (3) shows that (31) also holds for the hyperbolic random graph, apart from a multiplicative constant. From there we can follow the same lines as the proof of Proposition 1, so that Proposition 1 also holds for the hyperbolic random graph.

We follow the lines of the proof of Lemma 2, replacing $1-\mathrm{e}^{-D_uD_v/L_n}$ with $g(\nu t(u)t(v)/n)$ and using (39) to show that

$$\begin{aligned}
\mathbb{E}_n[a_{\varepsilon_n}(k,\ W_n^k(\delta))] &= \frac{(1+\mathrm{o}(1))}{k|M_{\varepsilon_n}(k)|}\sum_{v\in M_{\varepsilon_n}(k)}\sum_{u\in W_n^k(\delta)} D_u g(\nu t(u)t(v)/n)\\
&= \frac{2\nu(\tau-1)}{k\pi(\tau-2)}\sum_{u\in W_n^k(\delta)} t(u)g\bigg(\frac{t(u)k\pi(\tau-2)}{2n(\tau-1)}\bigg)(1+\mathrm{o}_{\mathbb{P}}(1))\\
&= (1+\mathrm{o}_{\mathbb{P}}(1))\frac{\nu\zeta}{k}\sum_{u\in W_n^k(\delta)} t(u)g\bigg(\frac{t(u)k}{\zeta n}\bigg),
\end{aligned}\tag{40}$$

where the error term may be placed in front of the summation since the summand is in $n/k[f_1(\delta), f_2(\delta)]$ for some $0<f_1(\delta), f_2(\delta)<\infty$ not depending on $n$. We analyse this expression following the lines of the proof of Lemma 3. We define

$$M^{(n)}[a,b]=\frac{1}{\zeta^{1-\tau}n^{2-\tau}k^{\tau-1}}\sum_{u\in[n]}\mathbf{1}_{\{D_u\in[a,b]\zeta n/k\}}$$

similarly to (26). From there, we follow the lines of the proof of Lemma 3, again replacing the connection probability $1-\mathrm{e}^{-D_iD_j/(\mu n)}$ of the erased configuration model with $g(\nu t(u)t(v)/n)$ and replacing the constant $c$ from (3) with its equivalent constant for the hyperbolic model of

Downloaded from https://www.cambridge.org/core. University of Bath, on 09 Feb 2020 at 11:24:30, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2019.45

$\tau - 1$ (see (10)) and $\mu$ by $\zeta$. Then (40) results in

$$\begin{aligned}\mathbb{E}_n[a_{\varepsilon_n}(k, W_n^{k,\mathrm{HRG}}(\delta))] &= (1+\mathrm{o}_{\mathbb{P}}(1))\frac{\nu n\zeta^2}{k^2}\sum_{u\in W_n^k(\delta)}\frac{t(u)k}{n\zeta}g\bigg(\frac{t(u)k}{\zeta n}\bigg)\\ &= (1+\mathrm{o}_{\mathbb{P}}(1))\nu\bigg(\frac{n\zeta}{k}\bigg)^{3-\tau}\int_\delta^{1/\delta} tg(t)\,\mathrm{d}M^{(n)}(t).\end{aligned}$$

Similar steps that prove (29) then show that

$$\frac{\mathbb{E}_n[a_{\varepsilon_n}(k, W_n^{k,\mathrm{HRG}}(\delta))]}{(n/k)^{3-\tau}}\xrightarrow{\mathbb{P}}(\tau-1)\nu\zeta^{3-\tau}\int_\delta^{1/\delta} x^{1-\tau}g(x)\,\mathrm{d}x.$$

Furthermore, because conditionally on the radial coordinates, the probabilities that two distinct edges are present are independent, Lemma 4 also holds for the hyperbolic random graph. This proves an analogous proposition to Proposition 2, which states that

$$\frac{a_{\varepsilon_n}(k, W_n^{k,\mathrm{HRG}}(\delta))}{(n/k)^{3-\tau}}\xrightarrow{\mathbb{P}}(\tau-1)\nu\zeta^{3-\tau}\int_\delta^{1/\delta} x^{1-\tau}g(x)\,\mathrm{d}x.$$

Therefore, steps similar to those that led to (23) then show that

$$\frac{a_{\varepsilon_n}(k)}{(n/k)^{3-\tau}}\xrightarrow{\mathbb{P}}(\tau-1)\nu\zeta^{3-\tau}\int_0^{\infty} x^{1-\tau}g(x)\,\mathrm{d}x.$$

Finally,

$$\begin{aligned}\int_0^\infty x^{1-\tau}g(x)\,\mathrm{d}x &= \frac{2}{\pi}\int_0^1 x^{1-\tau}\sin^{-1}(x)\,\mathrm{d}x + \int_1^\infty x^{1-\tau}\,\mathrm{d}x\\ &= \frac{2}{\pi}\bigg[\frac{x^{2-\tau}\sin^{-1}(x)}{2-\tau}\bigg]_0^1 + \frac{1}{\tau-2}\int_0^1\frac{x^{2-\tau}}{\sqrt{1-x^2}}\,\mathrm{d}x + \frac{1}{\tau-2}\\ &= \frac{1}{\tau-2}\int_0^{1/(2\pi)}\sin(t)^{2-\tau}\,\mathrm{d}t,\end{aligned}$$

where the last equation uses the substitution $t = \sin(x)$. By [21, equation 3.621.5]

$$\frac{1}{\tau-2}\int_0^{1/(2\pi)}\sin(t)^{2-\tau}\,\mathrm{d}t = \frac{\Gamma((3-\tau)/2)\Gamma(1/2)}{2(\tau-2)\Gamma((4-\tau)/2)} = \frac{\sqrt{\pi}\,\Gamma((3-\tau)/2)}{2(\tau-2)\Gamma((4-\tau)/2)},$$

where $\Gamma$ denotes the gamma function, which finishes the proof of Theorem 3(ii). □

## Acknowledgements

This work was supported by NWO TOP grant 613.001.451. The author would like to thank Remco van der Hofstad and Johan S. H. van Leeuwaarden for their useful comments.

## References

[1] Albert, R., Jeong, H. and Barabási, A.-L. (1999). Internet: diameter of the world-wide web. *Nature* **401**, 130–131.
[2] Barabási, A.-L. (2016). *Network Science*. Cambridge University Press.

Downloaded from https://www.cambridge.org/core. University of Bath, on 09 Feb 2020 at 11:24:30, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2019.45

[3] Barrat, A., Barthélemy, M., Pastor-Satorras, R. and Vespignani, A. (2004). The architecture of complex weighted networks. *Proc. Natl Acad. Sci. USA* **101**, 3747–3752.

[4] Bhamidi, S., Dhara, S., van der Hofstad, R. and Sen, S. (2017). Universality for critical heavy-tailed network models: metric structure of maximal components. Available at arXiv:1703.07145.

[5] Bhamidi, S., van der Hofstad, R. and Hooghiemstra, G. (2017). Universality for first passage percolation on sparse random graphs. *Ann. Probab.* **45**, 2568–2630.

[6] Bhamidi, S., Sen, S. and Wang, X. (2017). Continuum limit of critical inhomogeneous random graphs. *Probab. Theory Related Fields* **169**, 565–641.

[7] Bode, M., Fountoulakis, N. and Müller, T. (2015). On the largest component of a hyperbolic model of complex networks. *Electron. J. Combin.* **22**, P3.24.

[8] Bode, M., Fountoulakis, N. and Müller, T. (2016). The probability of connectivity in a hyperbolic model of complex networks. *Random Structures Algorithms* **49**, 65–94.

[9] Boguñá, M. and Pastor-Satorras, R. (2003). Class of correlated random networks with hidden variables. *Phys. Rev.* E **68**, 036112.

[10] Boguñá, M., Pastor-Satorras, R. and Vespignani, A. (2003). Absence of epidemic threshold in scale-free networks with degree correlations. *Phys. Rev. Lett.* **90**, 028701.

[11] Bollobás, B. (1980). A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combin.* **1**, 311–316.

[12] Bringmann, K., Keusch, R. and Lengler, J. (2015). Sampling geometric inhomogeneous random graphs in linear time. In *25th European Symposium on Algorithms (ESA 2017)* (Leibniz International Proceedings in Informatics **87**), art. 20. Leibniz-Zentrum für Informatik, Schloss Dagstuhl.

[13] Britton, T., Deijfen, M. and Martin-Löf, A. (2006). Generating simple random graphs with prescribed degree distribution. *J. Stat. Phys.* **124**, 1377–1397.

[14] Candellero, E. and Fountoulakis, N. (2014). Clustering and the hyperbolic geometry of complex networks. In *Algorithms and Models for the Web Graph (WAW 2014)* (Lecture Notes in Computer Science **8882**), pp. 1–12. Springer, Cham.

[15] Catanzaro, M., Boguñá, M. and Pastor-Satorras, R. (2005). Generation of uncorrelated random scale-free networks. *Phys. Rev.* E **71**, 027103.

[16] Chung, F. and Lu, L. (2002). The average distances in random graphs with given expected degrees. *Proc. Natl Acad. Sci. USA* **99**, 15879–15882.

[17] Colomer-de Simon, P. and Boguñá, M. (2012). Clustering of random scale-free networks. *Phys. Rev.* E **86**, 026120.

[18] van den Esker, H., van der Hofstad, R. and Hooghiemstra, G. (2008). Universality for the distance in finite variance random graphs. *J. Stat. Phys.* **133**, 169–202.

[19] Faloutsos, M., Faloutsos, P. and Faloutsos, C. (1999). On power-law relationships of the internet topology. In *ACM SIGCOMM Computer Communication Review*, vol. 29, pp. 251–262. ACM.

[20] Friedrich, T. and Krohmer, A. (2015). Cliques in hyperbolic random graphs. In *INFOCOM Proceedings 2015*, pp. 1544–1552. IEEE.

[21] Gradshteyn, I. and Ryzhik, I. (2015). *Table of Integrals, Series, and Products*. Elsevier.

[22] Gugelmann, L., Panagiotou, K. and Peter, U. (2012). Random hyperbolic graphs: degree sequence and clustering. In *ICALP Proceedings 2012, Part II*, pp. 573–585. Springer, Berlin and Heidelberg.

[23] van der Hofstad, R. (2017). *Random Graphs and Complex Networks*, vol. 1. Cambridge University Press.

[24] van der Hofstad, R., Hooghiemstra, G. and Van Mieghem, P. (2005). Distances in random graphs with finite variance degrees. *Random Structures Algorithms* **27**, 76–123.

[25] van der Hofstad, R., van der Hoorn, P., Litvak, N. and Stegehuis, C. (2017). Limit theorems for assortativity and clustering in the configuration model with scale-free degrees. Available at arxiv:1712.08097.

[26] van der Hofstad, R., Janssen, A. J. E. M., van Leeuwaarden, J. S. H. and Stegehuis, C. (2017). Local clustering in scale-free networks with hidden variables. *Phys. Rev.* E **95**, 022307.

[27] van der Hofstad, R., van Leeuwaarden, J. S. H. and Stegehuis, C. (2017). Optimal subgraph structures in scale-free configuration models. Available at arXiv:1709.03466.

[28] van der Hoorn, P. and Litvak, N. (2015). Upper bounds for number of removed edges in the erased configuration model. In *Algorithms and Models for the Web Graph (WAW 2015)* (Lecture Notes in Computer Science **9479**), pp. 54–65. Springer, Cham.

[29] Janson, S. and Luczak, M. J. (2007). A simple solution to the $k$-core problem. *Random Structures Algorithms* **30**, 50–62.

[30] Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N. and Barabási, A.-L. (2000). The large-scale organization of metabolic networks. *Nature* **407**, 651–654.

[31] Krioukov, D., Papadopoulos, F., Kitsak, M., Vahdat, A. and Boguná, M. (2010). Hyperbolic geometry of complex networks. *Phys. Rev.* E **82**, 036106.

Downloaded from https://www.cambridge.org/core. University of Bath, on 09 Feb 2020 at 11:24:30, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2019.45

[32] LESKOVEC, J. AND KREVL, A. (2014). SNAP Datasets: Stanford large network dataset collection. Available at http://snap.stanford.edu/data.
[33] MASLOV, S., SNEPPEN, K. AND ZALIZNYAK, A. (2004). Detection of topological patterns in complex networks: correlation profile of the internet. *Phys.* A **333**, 529–540.
[34] MAYO, M., ABDELZAHER, A. AND GHOSH, P. (2015). Long-range degree correlations in complex networks. *Comput. Social Networks* **2**, 4.
[35] NEWMAN, M. E. J. (2002). Assortative mixing in networks. *Phys. Rev. Lett.* **89**, 208701.
[36] OSTILLI, M. (2014). Fluctuation analysis in complex networks modeled by hidden-variable models: necessity of a large cutoff in hidden-variable models. *Phys. Rev.* E **89**, 022807.
[37] PARK, J. AND NEWMAN, M. E. J. (2003). Origin of degree correlations in the internet and other networks. *Phys. Rev.* E **68**, 026112.
[38] PASTOR-SATORRAS, R., VÁZQUEZ, A. AND VESPIGNANI, A. (2001). Dynamical and correlation properties of the internet. *Phys. Rev. Lett.* **87**, 258701.
[39] RAVASZ, E. AND BARABÁSI, A.-L. (2003). Hierarchical organization in complex networks. *Phys. Rev.* E **67**, 026112.
[40] SERRANO, M. Á. AND BOGUÑÁ, M. (2006). Percolation and epidemic thresholds in clustered networks. *Phys. Rev. Lett.* **97**, 088701.
[41] STEGEHUIS, C., VAN DER HOFSTAD, R., VAN LEEUWAARDEN, J. S. H. AND JANSSEN, A. J. E. M. (2017). Clustering spectrum of scale-free networks. *Phys. Rev.* E **96**, 042309.
[42] VÁZQUEZ, A. (2003). Growing network with local rules: preferential attachment, clustering hierarchy, and degree correlations. *Phys. Rev.* E **67**, 056104.
[43] VÁZQUEZ, A., PASTOR-SATORRAS, R. AND VESPIGNANI, A. (2002). Large-scale topological and dynamical properties of the internet. *Phys. Rev.* E **65**, 066130.
[44] WHITT, W. (2006). *Stochastic-Process Limits*. Springer, New York.
[45] YAO, D., VAN DER HOORN, P. AND LITVAK, N. (2018). Average nearest neighbor degrees in scale-free networks. *Internet Mathematics* **2018**. doi:10.24166/im.02.2018.

Downloaded from https://www.cambridge.org/core. University of Bath, on 09 Feb 2020 at 11:24:30, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/jpr.2019.45