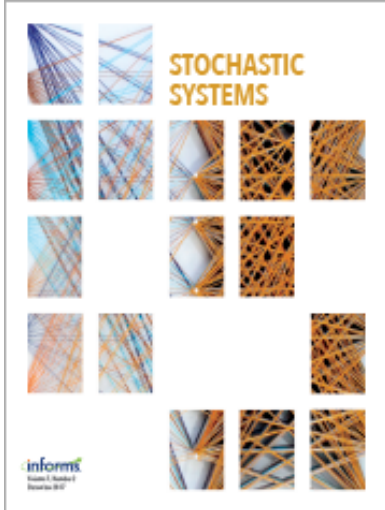




Stochastic Systems

Publication details, including instructions for authors and subscription information:
<http://pubsonline.informs.org>

Generating Maximally Disassortative Graphs with Given Degree Distribution

Pim van der Hoorn, Liudmila Ostroumova Prokhorenkova, Egor Samosvat

To cite this article:

Pim van der Hoorn, Liudmila Ostroumova Prokhorenkova, Egor Samosvat (2018) Generating Maximally Disassortative Graphs with Given Degree Distribution. *Stochastic Systems* 8(1):1-28. <https://doi.org/10.1287/stsy.2017.0006>

This work is licensed under a Creative Commons Attribution 4.0 International License. You are free to copy, distribute, transmit and adapt this work, but you must attribute this work as “*Stochastic Systems*. Copyright © 2017 The Author(s). <https://doi.org/10.1287/stsy.2017.0006>, used under a Creative Commons Attribution License: <https://creativecommons.org/licenses/by/4.0/>.”

Copyright © 2017, The Author(s)

Please scroll down for article—it is on subsequent pages



With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes. For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

Generating Maximally Disassortative Graphs with Given Degree Distribution

Pim van der Hoorn,^{a,b} Liudmila Ostroumova Prokhorenkova,^{c,d} Egor Samosvat^{c,d}

^a University of Twente; ^b DK-Lab, Northeastern University, Boston, Massachusetts 02115; ^c Moscow Institute of Physics and Technology; ^d Yandex

Contact: w.vanderhoorn@northeastern.edu,  <http://orcid.org/0000-0002-2250-1274> (PvdH); ostroumova-la@yandex.ru,  <http://orcid.org/0000-0002-1520-4167> (LOP); sameg@yandex-team.ru (ES)

Received: December 2016

Accepted: September 2017

Published Online in Articles in Advance:
January 31, 2018

MSC2010 Subject Classification: Primary:
05C80, 62H20

<https://doi.org/10.1287/stsy.2017.0006>

Copyright: © 2018 The Author(s)

Abstract. We present an algorithm that solves the problem of generating graphs, with a given degree distribution, that are maximally disassortative (with respect to Spearman's rank correlation). As a result, we obtain a general lower bound for Spearman's rho on graphs, which depends on the distribution of the probability mass between the head and tail of the size-biased degree distribution.

 **Open Access Statement:** This work is licensed under a Creative Commons Attribution 4.0 International License. You are free to copy, distribute, transmit and adapt this work, but you must attribute this work as "Stochastic Systems. Copyright © 2017 The Author(s). <https://doi.org/10.1287/stsy.2017.0006>, used under a Creative Commons Attribution License: <https://creativecommons.org/licenses/by/4.0/>."

Funding: This work is supported by Russian Science Foundation [Grant 16-11-10014]. This work is supported by ARO [Grant W911NF1610391].

Keywords: random graphs • degree distribution • degree-degree correlation • disassortativity

1. Introduction

An important second order characteristic of the topology of a graph, introduced in Newman (2002), is the correlation between the degrees at both sides of a randomly sampled edge, also called degree-degree correlation or degree assortativity. A graph is called assortative, or is said to have assortative mixing, if this correlation is positive and disassortative if it is negative. In assortative graphs, nodes of a certain degree have a preference to connect to nodes of similar degree, while in a disassortative graph the opposite is true, for instance, nodes of small degree connect to nodes with large degree. When the degrees of connected nodes are uncorrelated the graph is said to have neutral mixing, in which case there is no connection preference between nodes, based on degrees.

Recently, the problem of generating graphs with a given joint degree structure has been investigated. In Bassler et al. (2015) and Stanton and Pinar (2012) algorithms are introduced for constructing and sampling graphs with a given joint degree matrix J , where an entry J_{kl} denotes the number of edges between nodes of degrees k and l . An algorithm for generating random graphs whose joint degree distribution converges to a given limiting distribution is given in Deprez and Wüthrich (2015) and Hurd (2015) under the assumption that the degrees are uniformly bounded in the size of the graph.

A different branch of research is concerned with generating graphs that have extreme degree-degree correlation structure, either maximally assortative or disassortative, and analyzing structural properties of such graphs. One algorithm that is often used for this is the so-called edge swap algorithm (Kannan et al. 1999, Maslov and Sneppen 2002, Xulvi-Brunet and Sokolov 2004). In the context of degree-degree correlations, this algorithm starts from an initial graph, with a prescribed degree sequence, and in each step two edges are sampled and switched based on some rule, in order to obtain a maximally (dis)assortative graph. In Menche et al. (2010) this algorithm is used to obtain scaling results for Pearson's correlation coefficient, as introduced in Newman (2002) on maximally (dis)assortative graphs, where the degrees follow a scale-free distribution. The results from Menche et al. (2010) are extended in Yang et al. (2017), where a lower bound for Pearson's correlation coefficient is established in scale-free graphs.

One of the problems with the current analysis of graphs with extreme degree-degree correlation structure is the use of Pearson's correlation coefficient as a measure for assortativity, since this measure has been shown to be size-dependent when the degree distribution has infinite variance (Van Der Hofstad and Litvak 2014, Van Der Hoorn and Litvak 2015). In these papers, new, rank-based, correlation measures are introduced and it is shown that these measures converge to a proper limit, determined by the joint degree distribution, under very

standard assumptions, see Van Der Hofstad and Litvak (2014), Van Der Hoorn and Litvak (2014). Therefore, in this paper, we follow their suggestion and use a rank correlation measure related to Spearman's rho.

In addition, although many papers have addressed the problem of generating maximally disassortative graphs, none have provided significant insights into the connection structure of such graphs. We introduce a greedy algorithm for generating graphs with a given degree distribution that are maximally disassortative, with respect to the rank correlation measure Spearman's rho. The algorithm gives insights into the joint degree structure of these graphs and we use these to characterize the limiting joint degree distribution of maximally disassortative graphs, in terms of the size-biased degree distribution. Moreover, due to use of a general framework, describing the convergence of the empirical distributions, we are able to characterize the speed of the convergence.

An important consequence of the joint degree structure of maximally disassortative graphs is that the tail of the distribution alone does not affect the minimum value of Spearman's rho. Moreover, we are able to construct regularly varying distributions with a prescribed exponent, such that Spearman's rho on any graph with this degree distribution is bounded from below by a negative value that is arbitrary close to zero. This observation is crucial, since it tells us that the tail of the degree distribution by itself does not give any guarantees on the minimal value for degree-degree correlations and we should consider the full shape of the distribution instead.

We complement our theoretical results with simulations that show the concentration of Spearman's rho for graphs generated by our algorithm and illustrate how this measure is influenced by the shape of the size-biased degree distribution. We observe that the minimal value of Spearman's rho becomes larger when more mass is placed in the head of the degree distribution, i.e., when we increase the probability of small degrees. On the other hand, when we increase the probability for large degrees, which increases the mass in the tail of the distribution, the minimal value of Spearman's rho decreases.

The rest of this paper is structured as follows. We first introduce necessary notation and state our main results in the next section. In Section 3.1 we describe the algorithm for generating graphs that solves the optimization problem (12). A complete characterization of the empirical and limiting joint degree distribution is then given in Section 3.2 where we also discuss the proof of Theorem 2.2. We describe the construction of degree sequences with arbitrary small value of Spearman's rho in Section 4 and explain how Theorem 2.3 and Corollary 2.4 follow from this. In Section 5 we illustrate our results by providing simulations for maximally disassortative graphs where the degrees follow a scale-free and a Poisson distribution. Finally, Section 6 contains all the proofs of our results.

2. Notations and Results

We will start by introducing some notation and summarizing our main results.

2.1. Graphs and Degree Sequences

We call a sequence of non-negative integers $\mathbf{D}_n = \{D_1, D_2, \dots, D_n\}$ a degree sequence if $L_n = \sum_{i=1}^n D_i$ is even. In terms of graphs, L_n is the sum of the degrees and hence *twice* the number of edges in a graph with degree sequence $\mathbf{D}_n$. We further define the empirical and sized-biased degree distributions by, respectively,

$$f_n(k) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{D_i=k\}}, \quad (1)$$

$$f_n^*(k) = \frac{1}{L_n} \sum_{i=1}^n k \mathbb{1}_{\{D_i=k\}}, \quad (2)$$

and let F_n and F_n^* be the corresponding cumulative distribution functions.

For our results, we will consider sequences of degree sequences $\{\mathbf{D}_n\}_{n \geq 1}$ and require the empirical probability mass functions f_n and f_n^* to converge to certain limits f and f^* . Note however, that we made no assumptions on the origin of the degree sequences. That is, $\mathbf{D}_n$ might be fixed for each n , for instance when considering d -regular graphs, or it might be generated through some random process. When the degree sequence is random, so are the empirical functions f_n and f_n^* . Therefore, we will state our assumption in the following general way.

Assumption 2.1. *There exist probability mass functions f and f^* , on the non-negative integers, and an $\varepsilon > 0$, such that*

$$\sum_{k=0}^{\infty} k^{1+\varepsilon} f(k) < \infty \quad (3)$$

for some $\eta > 0$ and if we define,

$$\Omega_n = \left\{ \max \left\{ \sum_{k=0}^{\infty} \left| \sum_{t=0}^k f_n(t) - f(t) \right|, \sum_{k=0}^{\infty} |f_n^*(k) - f^*(k)| \right\} \leq n^{-\varepsilon} \right\},$$

then

$$\lim_{n \rightarrow \infty} \mathbb{P}(\Omega_n) = 1.$$

We will denote by F and F^* the cumulative distributions of f and f^* , respectively. Since we assume that the event Ω_n occurs, asymptotically, with probability one, we will often use the probability of its complement Ω_n^c to describe the speed of convergence in our results. In addition, for simplicity of notation, we will assume throughout this paper that $f(k), f^*(k) > 0$ for all $k \geq 0$. Our results extend in a straightforward manner to other cases, by considering only those k for which $f(k), f^*(k) > 0$.

To give some explanation regarding Assumption 2.1 we remark that the first expression in the maximum of the event Ω_n is related to the Kantorovich-Rubinstein distance or, equivalently, the Wasserstein metric of order one, between the distributions F_n and F . Convergence in this metric is equivalent to weak convergence as well as convergence of the first absolute moments, see Villani (2008) for more details. Hence, Assumption 2.1 describes the joint convergence of f_n to f and f_n^* to f^* in the Kantorovich-Rubinstein distance and the 1-norm, respectively. We used different metrics for the convergence of f_n and f_n^* , since the Wasserstein metric is only a true distance when the distributions have finite first absolute moment. We are not assuming that the distribution f^* has finite first absolute moment since we want to consider graphs whose degree distributions have infinite second moment, which implies that the size-biased degree distribution has infinite mean.

In order to state our results we will use the following definition

Definition 2.2. Let $\{\mathbf{D}_n\}_{n \geq 1}$ be a sequence of degree sequences. We say that $\{\mathbf{D}_n\}_{n \geq 1} \in \mathcal{D}_{\eta, \varepsilon}(f, f^*)$ if and only if $\{\mathbf{D}_n\}_{n \geq 1}$ satisfies Assumption 2.1 with density functions f and f^* and $\eta, \varepsilon > 0$. For a sequence of graphs $\{G_n\}_{n \geq 1}$ with a degree sequences $\{\mathbf{D}_n\}_{n \geq 1}$, we will write $\{G_n\}_{n \geq 1} \in \mathcal{G}_{\eta, \varepsilon}(f, f^*)$ if $\{\mathbf{D}_n\}_{n \geq 1} \in \mathcal{D}_{\eta, \varepsilon}(f, f^*)$.

We will often abuse notation and write $\mathbf{D}_n \in \mathcal{D}_{\eta, \varepsilon}(f, f^*)$ and similarly for G_n .

The simplest example of degree sequences that satisfy Assumption 2.1 is when we take $d = 2m$, for some integer m and consider the d -regular degree sequences ($\mathbf{D}_n = \{d, \dots, d\}$). In this case we have

$$f_n(k) = f(k) = \mathbb{1}_{\{k=d\}} = f_n^*(k) = f^*(k),$$

and hence, $\mathbf{D}_n \in \mathcal{D}_{\eta, \varepsilon}(f, f^*)$ for any $\eta, \varepsilon > 0$. Note that here the probability space is just the trivial space.

For an arbitrary probability density function f , a well known algorithm for generating degree sequences with the corresponding distribution is by sampling the degrees i.i.d. from the distribution and then increase the last degree by 1 if the sum is not even. We will refer to this as the IID algorithm. The following lemma states that when the distribution from which the degrees are sampled has just a bit more than finite mean, the resulting degree sequence satisfies Assumption 2.1 with the given probability density function f .

Lemma 2.1. Let D be an integer valued random variable with probability density function f , such that $\mathbb{E}[D^{1+\eta}] < \infty$ for some $\eta > 0$. Denote by μ the mean of D and define $f^*(k) = \mathbb{E}[D \mathbb{1}_{\{D=k\}}] / \mu$. Then if $\mathbf{D}_n$ is generated by the IID algorithm by sampling from f ,

$$\mathbf{D}_n \in \mathcal{D}_{\eta, \varepsilon}(f, f^*) \quad \text{for any } \varepsilon \leq \eta / (8 + 4\eta).$$

Moreover,

$$\mathbb{P}(\Omega_n) \geq 1 - O(n^{-\varepsilon}),$$

as $n \rightarrow \infty$.

2.3. Spearman's Rho on Graphs

Historically, Spearman's rho (Spearman 1904) is a measure of the correlation between two random variables X and Y , based on a set of joint observations $(X_i, Y_i)_{1 \leq i \leq m}$. For this we first rank both vectors $\{X_i\}_{1 \leq i \leq m}$ and $\{Y_i\}_{1 \leq i \leq m}$ and then compute Pearson's correlation coefficient on the joint vector of these ranks $(R_i^X, R_i^Y)_{1 \leq i \leq m}$, where R_i^X and R_i^Y are the ranks of, respectively, X_i and Y_i .

The general definition of Spearman's rho for two continuous random variables X and Y with, respectively, cumulative distribution F_X and F_Y and joint distribution H is Schweizer and Wolff (1981)

$$\rho(X, Y) = 12 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (H(x, y) - F_X(x)F_Y(y)) dF_X(x) dF_Y(y)$$

Now let X and Y be integer-valued with cumulative distribution. Then it follows from Proposition 2.4 in Mesfioui and Tajar (2005) that,

$$\rho(X, Y) = 3\mathbb{E}[\mathcal{F}_X(X)\mathcal{F}_Y(Y)] - 3, \quad (4)$$

where

$$\mathcal{F}_X(k) = F_X(k) + F_X(k-1), \quad \text{for all } k \in \mathbb{Z}, \quad (5)$$

and similar for $\mathcal{F}_Y$.

In graphs, Spearman's rho is used to compute the correlation between the degrees of connected nodes, called degree-degree correlations (Van Der Hofstad and Litvak 2014, Van Der Hoorn and Litvak 2015). To give a definition, it is convenient to consider directed edges. To make this work on undirected graphs we replace each edge $i-j$ by two edges, $i \rightarrow j$ and $j \rightarrow i$. We refer to this graph as the bi-directed version of the original graph. Although the graph on which Spearman's rho is computed is directed, we will not distinguish between this and the original undirected graph G_n . That is, we will write $i \rightarrow j \in G_n$ to mean that $i \rightarrow j$ is present in the bi-directed version of G_n , which is equivalent to stating that $i-j \in G_n$. We recall that L_n denotes the sum over all degrees, so that L_n is twice the number of undirected edges and equal to the corresponding number of directed edges in G_n .

Given a graph G_n we consider the joint measurements of the degrees on both sides of the edges $(D_i, D_j)_{i \rightarrow j}$ and compute Spearman's rho on them. For this we need to rank the degree vectors, given by the list of edges, which contain many ties. Consider, for example, any edge $i \rightarrow j \in G$. Then the value D_i will be present in the list $(D_i)_{i \rightarrow j}$ for each edge coming out of node i , resulting in at least D_i tied values. There are several ways to resolve ties and in this paper we resolve them at random as described in Van Der Hofstad and Litvak (2014) and Van Der Hoorn and Litvak (2015). We do this by adding to each value of the degree, independently of everything else, a value sampled uniformly at random from $(0, 1)$. After this procedure, with probability one, all entries will be distinct and we can rank them. To make this more precise, let $(U_{i \rightarrow j}, W_{i \rightarrow j})$ to be a vector of independent uniform random variables $U_{i \rightarrow j}$ and $W_{i \rightarrow j}$ on $(0, 1)$, for each edge $i \rightarrow j \in G$. Then, instead of $(D_i, D_j)_{i \rightarrow j}$ we consider the list of joint measurements $(D_i + U_{i \rightarrow j}, D_j + W_{i \rightarrow j})_{i \rightarrow j}$. The ranking functions $R_*(i \rightarrow j)$ and $R^*(i \rightarrow j)$, corresponding to the ranks of, respectively, $D_i + U_{i \rightarrow j}$ and $D_j + W_{i \rightarrow j}$ are then given by

$$R_*(i \rightarrow j) = \sum_{s \rightarrow t \in G} \mathbb{1}_{\{D_s + U_{s \rightarrow t} \geq D_i + U_{i \rightarrow j}\}},$$

$$R^*(i \rightarrow j) = \sum_{s \rightarrow t \in G} \mathbb{1}_{\{D_t + W_{s \rightarrow t} \geq D_j + W_{i \rightarrow j}\}},$$

where we let $\sum_{i \rightarrow j \in G}$ denote the sum over all edges $i \rightarrow j$ in the graph G . With these definitions, Spearman's rho is given by, see Van Der Hofstad and Litvak (2014), Van Der Hoorn and Litvak (2015),

$$\rho(G_n) = \frac{12 \sum_{i \rightarrow j \in G} R_*(i \rightarrow j) R^*(i \rightarrow j) - 3L_n(L_n + 1)^2}{L_n^3 - L_n}. \quad (6)$$

To link $\rho(G_n)$ to (4), let h_n denote the empirical joint probability density function of the degrees on both sides of a random edge, i.e.,

$$h_n(k, l) = \frac{1}{L_n} \sum_{i \rightarrow j \in G} \mathbb{1}_{\{D_i = k\}} \mathbb{1}_{\{D_j = l\}}.$$

Then, if h_n converges to some limiting joint probability density function h , it follows from Theorem 3.2 in Van Der Hofstad and Litvak (2014) that

$$\rho(G_n) \xrightarrow{\mathbb{P}} \rho(X, Y) \quad \text{as } n \rightarrow \infty,$$

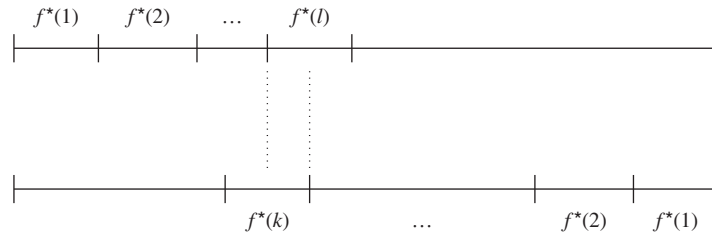
where $\xrightarrow{\mathbb{P}}$ denotes convergence in probability and X and Y have joint distribution h . In other words, $\rho(G_n)$ is a consistent estimator of $\rho(X, Y)$. It is important to note that this convergence holds whenever the degree distribution has finite mean. In contrast, Pearson's correlation coefficient converges only if the degree distribution has finite third moment. Moreover, for any sequence of graphs with a scale-free distribution with infinite second moment, any limit point of Pearson is positive, see Theorem 3.1 in Van Der Hofstad and Litvak (2014), hence it cannot classify large disassortative graphs. This is why we work with Spearman's rho.

In Van Der Hoorn and Litvak (2015) it is shown that $\rho(G_n)$ is asymptotically equivalent to

$$\tilde{\rho}(G_n) = \frac{3}{L_n} \sum_{i \rightarrow j} \mathcal{F}_n^*(D_i) \mathcal{F}_n^*(D_j) - 3, \quad (7)$$

see also Theorem 5.3 in Van Der Hoorn (2016) for a stronger result. Since this expression is easier to analyze mathematically, we will use this measures in our statements. We show with numerical experiments in Section 5 that our results also hold for the original expression (6).

Figure 1. Illustration of the functions ψ and $\mathcal{E}$.



2.4. Main Results

In order to state the first result we define, for any $k, l \geq 1$, the functions

$$\psi(k, l) = \mathbb{1}_{\{1-F^*(k) < F^*(l)\}} \mathbb{1}_{\{1-F^*(k-1) > F^*(l-1)\}}, \quad (8)$$

$$\mathcal{E}(k, l) = \min(1 - F^*(k-1), F^*(l)) - \max(1 - F^*(k), F^*(l-1)). \quad (9)$$

These functions can be understood as follows. Consider the partition of the interval $[0, 1]$ given by the sequence $\{f^*(1), f^*(2), \dots\}$. Now take a copy of this partitioned interval, reverse it and place it below the original interval, see Figure 1. Then $\psi(k, l)$ is the indicator of the event that the interval corresponding to $f^*(l)$ on the top intersects with the interval corresponding to $f^*(k)$ at the bottom, while $\mathcal{E}(k, l)$ is the size of this intersection.

With these functions we now define

$$h(k, l) = \psi(k, l) \mathcal{E}(k, l), \quad k, l = 1, 2, \dots \quad (10)$$

By definition of ψ and $\mathcal{E}$, and their representation in Figure 1, it immediately follows that h is a joint probability density function with marginals f^* . To see this, fix a k and note that the only non-zero contributions to $h(k, l)$ are obtained by those values of l for which the interval corresponding to $f^*(l)$ in the top line intersects with $f^*(k)$ on the bottom line. Since for each of these values $\mathcal{E}(k, l)$ is the length of the intersection of these intervals, by summing over l , we uncover the full size of the interval corresponding to $f^*(k)$ which by construction equals $f^*(k)$. Since h is symmetric we therefore obtain

$$\sum_{l \geq 1} h(k, l) = f^*(k) \quad \text{and} \quad \sum_{k \geq 1} h(k, l) = f^*(l).$$

Our main result states that if X and Y have joint probability density h , then Spearman's rho on graphs with a degree sequence satisfying Assumption 2.1 is bounded from below by $\rho(X, Y)$, and that this minimum is attained for a specific sequence of graphs.

Theorem 2.2. Let $G_n \in \mathcal{G}_{n, \varepsilon}(f, f^*)$ and let D_*, D^* be random variables with joint probability density h as defined in (10). Then, for any $0 < \delta < \min(\varepsilon, 1/2)$ and $K > 0$,

$$\mathbb{P}(\tilde{\rho}(G_n) \geq \rho(D_*, D^*) - Kn^{-\delta}) \geq 1 - O(n^{-\varepsilon + \kappa} + \mathbb{P}(\Omega_n^c)), \quad (11)$$

as $n \rightarrow \infty$, where

$$\kappa = \frac{\varepsilon + \delta}{2}.$$

Moreover, there exist graphs $\hat{G}_n$ with the same degree sequence as G_n , such that, as $n \rightarrow \infty$,

$$\mathbb{P}(|\tilde{\rho}(\hat{G}_n) - \rho(D_*, D^*)| > Kn^{-\delta}) \leq O(n^{-\varepsilon + \kappa} + \mathbb{P}(\Omega_n^c)).$$

This result can be understood in terms of the following optimization problem. Given degree sequences $\mathbf{D}_n \in \mathcal{D}_{n, \varepsilon}(f, f^*)$, define

$$\mathcal{F}_n^*(k) = F_n^*(k) + F_n^*(k-1).$$

and consider, for fixed n , the following objective function

$$\min_{G \in \mathcal{G}(\mathbf{D}_n)(f, f^*)} \frac{1}{L_n} \sum_{i \rightarrow j \in G} \mathcal{F}_n^*(D_i) \mathcal{F}_n^*(D_j), \quad (12)$$

where the minimum is understood to be taken over all graphs G_n with degree sequences satisfying Assumption 2.1 with density functions f and f^* . Then Theorem 2.2 states that with high probability

$$\min_{G_n \in \mathcal{G}_{n,\varepsilon}(f, f^*)} \rho(G_n) = \rho(D_*, D^*),$$

where $\rho(D_*, D^*)$ is given by, see (4),

$$\rho(D_*, D^*) = 3 \sum_{k,l=0}^{\infty} \mathcal{F}^*(k) \mathcal{F}^*(l) \psi(k, l) \mathcal{E}(k, l) - 3,$$

with ψ and $\mathcal{E}$ as defined in (8) and (9), respectively. Moreover, Theorem 2.2 provides a sequence of graphs $\hat{G}_n$ that attains this minimum, i.e., a sequence of maximally disassortative graphs with the degree distribution f . These graphs will be generated by our algorithm, which we will present in Section 3.

We remark that although Theorem 2.2 solves the minimization problem of degree-degree correlations in undirected graphs by giving the asymptotic minimum $\rho(D_*, D^*)$ on Spearman's rho, this minimum is, in general, hard to derive since it depends on the full size-biased limit density f^* . However, for specific cases it can be computed numerically by computing

$$\sum_{k=0}^K \sum_{l=0}^L \mathcal{F}^*(k) \mathcal{F}^*(l) \psi(k, l) \mathcal{E}(k, l),$$

for certain upper bounds K and L .

Part of the proof of Theorem 2.2 consists of showing that h is the limiting probability density function of the joint degree distribution of maximally disassortative graphs. From the interpretation of the functions ψ and $\mathcal{E}$, it follows that for all $k \geq K$, for some threshold K , all intervals corresponding to $f^*(k)$ on the top will be contained in the interval $f^*(1)$ at the bottom and vice versa. This implies that the large degree nodes will, asymptotically, all be connected to nodes with degree one. As a consequence we have that the tail of the distribution f^* , and hence also that of f , plays a negligible role in the lower bound of Spearman's rho. Therefore, we can construct degree distributions with specified tail behavior so that Spearman's rho on such graphs approaches zero from below, with arbitrary precision. This is an important result, since a measured correlation close to zero is seen as an indication that the graph has neutral mixing of degrees, while due to Theorem 2.2 we know that these graphs are maximally disassortative.

Theorem 2.3. *Let f be any probability density function with support on the non-negative integers, mean μ and*

$$\sum_{k=0}^{\infty} k^{1+\eta} f(k) < \infty,$$

for some $\eta > 0$. Then, for any $-1 < \rho < 0$, there exists a probability density function f_ρ on the non-negative integers with mean μ_ρ , which satisfies

$$\lim_{k \rightarrow \infty} \frac{1 - F_\rho(k)}{1 - F(k)} = \frac{\mu_\rho}{\mu}.$$

Moreover, for any sequence of graphs $G_n \in \mathcal{G}_{n,\varepsilon}(f_\rho, f_\rho^)$, where $f_\rho^*(k) = k f_\rho(k) / \mu_\rho$, we have*

$$\mathbb{P}(\tilde{\rho}(G_n) > \rho) \geq 1 - O(n^{-1+\kappa} + n^{-\varepsilon+3\kappa/4} + \mathbb{P}(\Omega_n^c)),$$

as $n \rightarrow \infty$, where $\kappa = \min(\varepsilon, 1/2)$.

The main message of Theorem 2.3 is that it is not the tail of the degree distribution that is crucial for the minimal value of $\tilde{\rho}(G_n)$.

The characterization of the tail of the degree distribution is most prominently present in the analysis of so-called scale-free networks. These are graphs where the limiting degree distribution F satisfies

$$1 - F(k) = \mathcal{L}(k) k^{-\gamma}, \quad \gamma > 1, \tag{13}$$

for some slowly varying function $\mathcal{L}$. The exponent γ is referred to as the tail exponent. As a corollary to Theorem 2.3 we obtain the following result which states that knowledge of only the tail exponent does not give any guarantees on the minimum value of Spearman's rho.

Corollary 2.4. *For any $-1 < \rho < 0$ and $\gamma > 1$, there exist probability density functions f and f^* , where F satisfies (13), such that for any sequence of graphs $G_n \in \mathcal{G}_{n,\varepsilon}(f, f^*)$*

$$\mathbb{P}(\rho(G_n) > \rho) \geq 1 - O(n^{-1+\kappa} + n^{-\varepsilon+3\kappa/4} + \mathbb{P}(\Omega_n^c)),$$

as $n \rightarrow \infty$, where $\kappa = \min(\varepsilon, 1/2)$.

3. Generating Maximally Disassortative Graphs

We will describe an algorithm, called the Disassortative Graph Algorithm (DGA), that solves (12).

3.1. The Disassortative Graph Algorithm

Any degree sequence $\mathbf{D}_n$ can be represented by a list of stubs, where for each node i we have D_i stubs labeled i . A graph with degree sequence $\mathbf{D}_n$ is then completely determined by the pairing of the stubs. In order to describe our algorithm, let N_k denote the number of nodes with degree k and let z_n be the unique integer satisfying

$$\sum_{t=1}^{z_n} tN_t \geq \frac{L_n}{2} \quad \text{and} \quad \sum_{t=1}^{z_n-1} tN_t < \frac{L_n}{2}. \quad (14)$$

The idea of the Disassortative Graph Algorithm is to use z_n to divide the stubs in two columns. In the left column S_n we add the stubs belonging to nodes with high degree ($D_i \geq z_n$), in descending order. The right column T_n will be filled with stubs that belong to nodes with small degree ($D_i \leq z_n$) in ascending order. After this ordering we start pairing stubs from the left column to stubs in the right column, until we reach the first pair (i, j) for which $D_i = z_n = D_j$. We are now left with stubs belonging to nodes with degree z_n , hence the value of Spearman's rho (7) will not be influenced by the specific way in which we connect them. This means that we can, in principle, use any algorithm to connect these medium degree nodes. We will use the configuration model (Bollobás 1980, Molloy and Reed 1995, Newman et al. 2001), more specifically the repeated configuration model, see Section 7.4 in Van Der Hofstad (2016). The full algorithm is described in detail below.

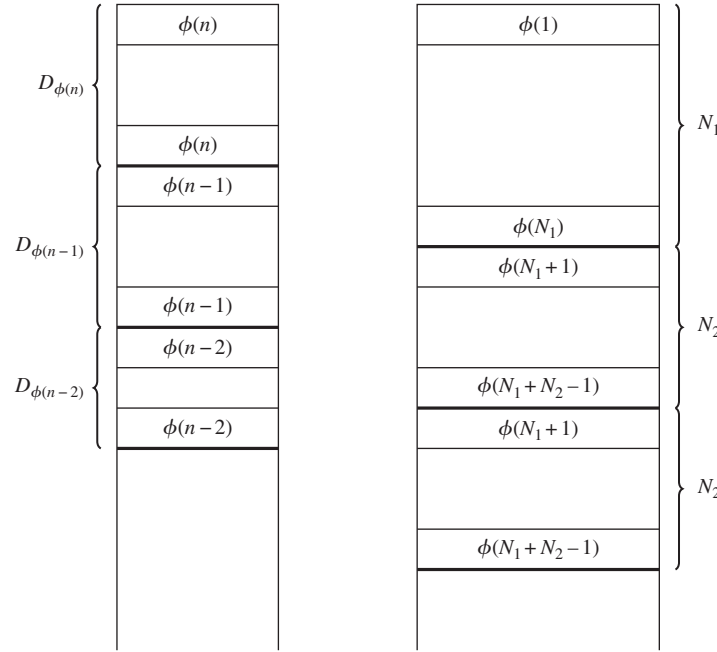
Algorithm 1 (Disassortative graph algorithm)

- 1: Input: A degree sequence $\mathbf{D}_n$.
- 2: Rank the nodes by their degrees in ascending order and let $\varrho(k)$ and denote the node with rank k ,
i.e., $D_{\varphi(n)} \geq D_{\varphi(n-1)} \geq \dots \geq D_{\varphi(2)} \geq D_{\varphi(1)}$.
- 3: Create two empty lists S_n and T_n .
- 4: Set $i = n$ and $j = 1$.
- 5: **while** $D_{\varphi(i)} \geq z_n$ **do**
- 6: Fill the next $D_{\varphi(i)}$ slots of S_n with stubs labeled $\varphi(i)$.
- 7: Set $i = i - 1$.
- 8: **end while**
- 9: **while** $D_{\varphi(j)} \leq z_n$ **do**
- 10: Add to T_n , $D_{\varphi(j)}$ copies of stubs labeled: $\varphi(j), \dots, \varrho(j + N_{D_{\varphi(j)}} - 1)$.
- 11: Set $j = j + N_{D_{\varphi(j)}}$.
- 12: **end while**
- 13: Set $t = 1$, $i = S_n[t]$ and $j = T_n[t]$
- 14: **while not** $D_i = z_n = D_j$ **do**
- 15: Add edge $i - j$ to G_n .
- 16: Set $t = t + 1$, $i = S_n[t]$ and $j = T_n[t]$.
- 17: **end while**
- 18: Set $\mathbf{D}_n^z$ to be the degree sequence corresponding to the remaining unpaired stubs.
- 19: Pair the stubs in $\mathbf{D}_n^z$ using the configuration model.
- 20: Output: G_n .

We will denote by G_n^* the induced sub-graph that has been created at the end of step 17 and let the complement $G_n^z = G_n \setminus G_n^*$ denote the graph generated by the configuration model in step 19. In addition we will write $G_n = \text{DGA}(\mathbf{D}_n)$ if G_n is generated by the Disassortative Graph Algorithm with degree sequence $\mathbf{D}_n$ as input. An illustration of the lists S_n and T_n is displayed in Figure 2.

It is important to note that, in principle, the DGA generates multi-graphs, i.e., graphs with self-loops and multiple edges between nodes. However, as we will discuss in Section 3.3, when combined with the IID algorithm, it can be used to generate simple graphs with a given degree distribution, since the probability that the resulting graph is simple converges to a positive constant, see Corollary 3.5 and Algorithm 2.

We first illustrate the DGA on the simple degree sequence $\{1, 2, 2, 3\}$, see Figure 3. Observe that in this case $z_n = 2$. Figure 3(a) shows the initialization state where we have created the lists S_n and T_n and no stubs have been connected. We start, Figure 3(b), by connecting the nodes at the top of the lists, 4 and 1. Then we move down the lists, Figure 3(c), and connect 4 and 2. The next step, Figure 3(d), is where the specific way the algorithm ordered the stubs in both lists comes into play.

Figure 2. Top part of the two lists of stubs.

There is one stub left on the node with the largest degree, node 4. The smallest degree among the still available nodes is two. Therefore we want to connect node 4 to a node with degree two which are 2 and 3. However, since there is already an edge between 4 and 2, connecting them again will create multi-edges between these nodes. The ordering of the lists resolves this by making sure we first connected to each different node with the same degree before we can create an edge between two nodes that have already been connected. In this example we therefore connect 4 and 3.

After this step the algorithm reaches a pair of nodes that both have degree $z_n = 2$, Figure 3(e). This is where we stop and pair the remaining stubs using the configuration model. Since in this specific example only nodes 2 and 3 have a stub left, we connect these, Figure 3(f).

3.2. Joint Degree Distribution of Maximally Disassortative Graphs

Before we turn to analysis of the DGA it is useful to look at the empirical joint degree distribution of graphs generated by the algorithm. We will give a complete characterization of both the empirical and limiting joint degree distributions in Proposition 3.1 and Theorem 3.2, respectively.

In order to analyze the structure of the joint degree distribution we approach the algorithm from a different angle. First observe that if we are only interested in the degrees D_i and D_j for an undirected edge $i - j$, then the specific way in which the stubs are ordered by the algorithm is irrelevant for $h_n(k, l)$. This means we do not have to consider the label of the nodes to which stubs belong, only their degree. Note that the number of stubs belonging to nodes of degree k equals kN_k . Moreover, due to the symmetry in the transition to directed edges, by replacing an edge $i - j$ with edges $i \rightarrow j$ and $j \rightarrow i$, the directed structure of the graph generated by DGA can be seen as follows.

Consider the partition of the set $\{1, 2, \dots, L_n\}$ given by kN_k for $k = 0, 1, \dots$, represented as a line of length L_n partitioned into intervals of size kN_k . Now take a copy of this partitioned line, reverse it and place it below the original one, see Figure 4. Both lines can be seen as the lists of all stubs, ordered by the degree of the nodes to which they belong. For the top line the stubs are ordered, from left to right, in increasing order of the degree, while for the bottom line the degrees are in decreasing order. Then the DGA can be seen as creating directed edges $i \rightarrow j$ between the nodes i corresponding to the stubs on the bottom line and nodes j corresponding to stubs in the top line.

From this representation we observe that an edge $i \rightarrow j$ between nodes of degree $D_i = k$ and $D_j = l$ exists if and only if the interval corresponding to kN_k in the partitioned bottom line has an intersection with the interval

Figure 3. Example of the DGA on a simple degree sequence with $z_n = 2$.

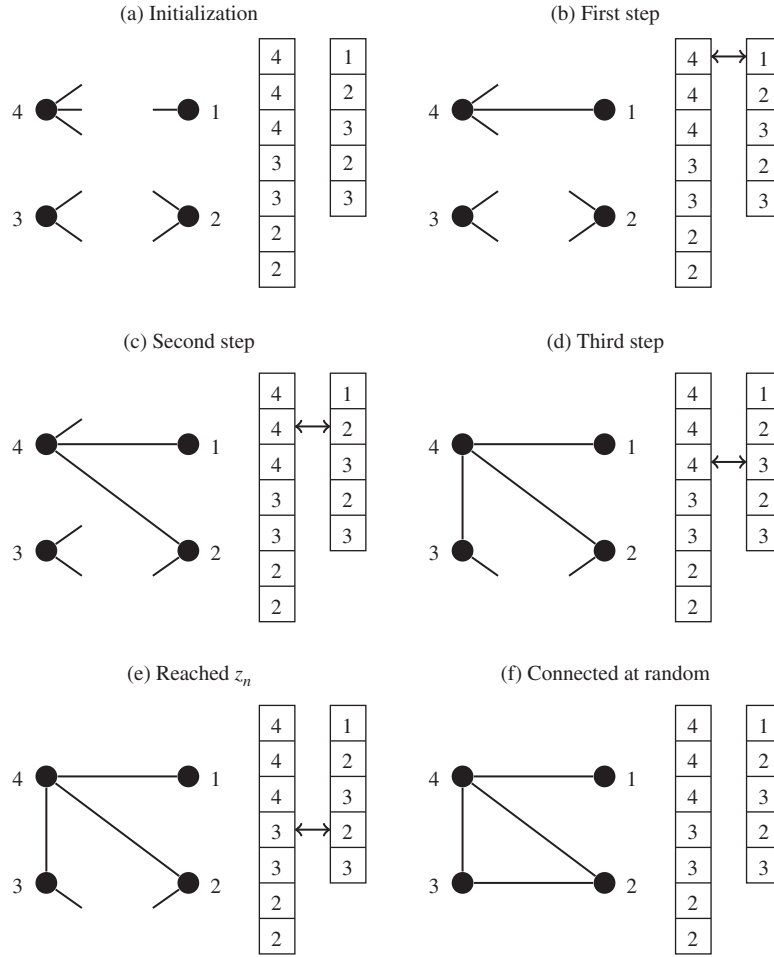
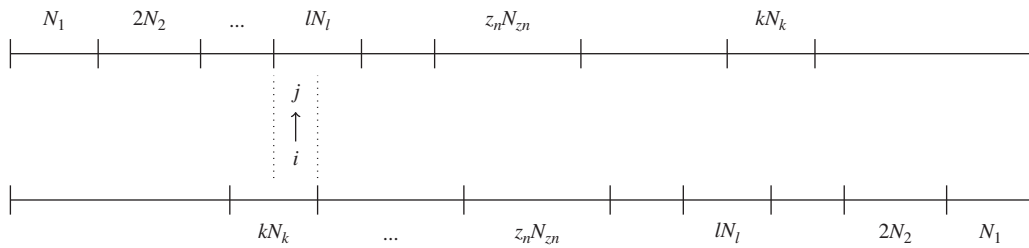


Figure 4. Partition representation of connections in the DGA.



corresponding to lN_l in the partitioned upper line. In terms of N_t this holds, if and only if,

$$\sum_{t=k+1}^{\infty} tN_t < \sum_{t=1}^l tN_t \quad \text{and} \quad \sum_{t=1}^{l-1} tN_t < \sum_{t=k}^{\infty} tN_t. \quad (15)$$

Moreover, the number of edges that connect nodes of degree k and l is equal to the size of the intersection,

$$\min \left\{ \sum_{t=k}^{\infty} tN_t, \sum_{t=1}^l tN_t \right\} - \max \left\{ \sum_{t=k+1}^{\infty} tN_t, \sum_{t=1}^{l-1} tN_t \right\}. \quad (16)$$

This partitioned representation of both the DGA and the joint degree structure, as displayed in Figure 4, will be crucial for the analysis of the structure of maximally disassortative graphs.

First let $\psi_n(k, l)$ denote the indicator that there exists a directed edge $i \rightarrow j$ with $D_i = k$ and $D_j = l$. Then since for any $k \geq 0$,

$$\frac{1}{L_n} \sum_{t=0}^k t N_t = \frac{1}{L_n} \sum_{t=0}^k t \sum_{i=1}^n \mathbb{1}_{\{D_i=t\}} = F_n^*(k),$$

it follows from (15) that

$$\psi_n(k, l) := \mathbb{1}_{\{1-F_n^*(k) < F_n^*(l)\}} \mathbb{1}_{\{1-F_n^*(k-1) > F_n^*(l-1)\}}. \quad (17)$$

Moreover, if we let $\mathcal{E}_n(k, l)$ denote the average number of edges between nodes of degree k and l , then (16) implies that

$$\mathcal{E}_n(k, l) = \min(1 - F_n^*(k-1), F_n^*(l)) - \max(1 - F_n^*(k), F_n^*(l-1)). \quad (18)$$

Summarizing we therefore have the following result.

Proposition 3.1. *Let $G_n = \text{DGA}(\mathbf{D}_n)$, for some degree sequence $\mathbf{D}_n$ and define the functions ψ_n and $\mathcal{E}_n$, on the positive integers by*

$$\psi_n(k, l) = \mathbb{1}_{\{1-F_n^*(k) < F_n^*(l)\}} \mathbb{1}_{\{1-F_n^*(k-1) > F_n^*(l-1)\}} \quad \text{and} \quad (19)$$

$$\mathcal{E}_n(k, l) = \min(1 - F_n^*(k-1), F_n^*(l)) - \max(1 - F_n^*(k), F_n^*(l-1)). \quad (20)$$

Then,

$$h_n(k, l) = \psi_n(k, l) \mathcal{E}_n(k, l).$$

From Proposition 3.1 we obtain that when $\mathbf{D}_n \in \mathcal{D}_{\eta, \varepsilon}(f, f^*)$, the limiting joint degree probability density function for graphs generated by the DGA is given by (10).

Theorem 3.2. *Let $\mathbf{D}_n \in \mathcal{D}_{\eta, \varepsilon}(f, f^*)$ and $G_n = \text{DGA}(\mathbf{D}_n)$. In addition let $h(k, l)$ be as defined in (10), take $0 < \delta < \varepsilon$, $K > 0$ and define the event*

$$\Xi_n = \left\{ \sum_{k, l=0}^{\infty} |h_n(k, l) - h(k, l)| \leq K n^{-\delta} \right\}.$$

Then

$$\mathbb{P}(\Xi_n) \geq 1 - O(n^{-\varepsilon+\delta} + \mathbb{P}(\Omega_n^c)),$$

as $n \rightarrow \infty$.

We will use Theorem 3.2 in Section 6.5 to prove our main result, Theorem 2.2.

3.3. Properties of the Disassortative Graph Algorithm

We will now address several properties of the Disassortative Graph Algorithm. The first is concerned with the optimization problem (12).

Theorem 3.3. *The Dissassortative Graph Algorithm solves (12).*

This result can be explained as follows. Let $\mathbf{a}_{L_n}$ be the list of degrees with respect to the labels of the stubs, ordered in descending order. That is

$$\mathbf{a}_{2L_n} = \left(\underbrace{D_{\phi(n)}, \dots, D_{\phi(n)}}_{D_{\phi(n)}}, D_{\phi(n-1)}, \dots, \underbrace{D_{\phi(N_1)}, \dots, D_{\phi(1)}}_{N_1} \right).$$

Then the DGA pairs the degrees a_i and a_{L_n+1-i} , which minimizes the sum $\sum_{i \rightarrow j \in G} \mathcal{F}_n^*(D_i) \mathcal{F}_n^*(D_j)$ and hence the DGA minimizes Spearman's rho $\bar{\rho}(G_n)$. Observe that, in addition, the algorithm minimizes $\sum_{i \rightarrow j \in G} D_i D_j$ so that we also obtain the minimum for the s metric of the graph G , $s_{\min}$, as considered in Alderson and Li (2007). In particular, this implies that the DGA also minimizes Pearson's correlation coefficient. Moreover, the fact that we could use an arbitrary algorithm to connect the nodes of degree z_n confirms the observation in Alderson and Li (2007) that graphs with minimal s metric are not unique with respect to their structure.

Proving Theorem 2.2. From Theorem 3.2 it follows that if $\mathbf{D}_n \in \mathcal{D}_{\eta, \varepsilon}(f, f^*)$, the empirical joint density function of the degrees of graphs generated by the DGA converges to the joint density function defined by (10). As a result, Spearman's rho on these graphs, $\tilde{\rho}$, converges to Spearman's rho of two random variables with this joint density, which proves the second part of Theorem 2.2. The first part follows because Theorem 3.3 states that $\tilde{\rho}$ is minimized by the DGA. The details of the proof can be found in Section 6.5.

As we have already mentioned, the joint degree structure, and hence the optimality of the DGA, depends only on the degree of nodes that are connected and not on their labels. In the algorithm, however, we use an ordering for filling the lists of stubs S_n and T_n . This is to make sure that the probability that G_n^* is simple, i.e., it has no self-loops and no more than one edge between nodes i and j , converges to one as $n \rightarrow \infty$.

To understand the intuition behind the proof, consider the first time the algorithm sees a stub belonging to a node i in the list S_n with degree $D_i > z_n$. Then node i will be connected to the nodes corresponding to the next D_i stubs in T_n . Now consider such a stub, belonging to node j . Then there will be more than one edge $i - j$ if and only if there is more than one stub belonging to node j among the D_i stubs in T_n , which can only happen when $D_i > N_{D_j}$. Since the degree of nodes in T_n is bounded by z_n , we have that N_{D_j} scales as n , while the maximal degree is $o(n)$, since f has finite mean. Therefore, the event $D_i > N_{D_j}$ for $D_i > z_n$ and $D_j \leq z_n$ has vanishing probability. We hence have the following result, the details of the proof can be found in Section 6.3.

Proposition 3.4. Let $\mathbf{D}_n \in \mathcal{D}_{\eta, \varepsilon}(f, f^*)$, $G_n = \text{DGA}(\mathbf{D}_n)$ and denote by $\mathcal{S}_n^*$ the event that G_n^* is simple, then

$$\mathbb{P}(\mathcal{S}_n^*) \geq 1 - O(n^{-\varepsilon} + n^{-1/2} + n^{-\eta/2} + \mathbb{P}(\Omega_n^c)),$$

as $n \rightarrow \infty$.

This proposition implies that the simplicity of the graph G_n , generated by the Disassortative Graph Algorithm, solely depends on the simplicity of G_n^z , constructed in Step 19. Now consider the degree sequence $\mathbf{D}_n^z$ corresponding to the remaining stubs, obtained in Step 18 and observe that these degrees are uniformly bounded by z_n . Take $\mathbf{D}_n \in \mathcal{D}_{\eta, \varepsilon}(f, f^*)$ and let z be the median of F^* , i.e., the unique integer such that

$$F^*(z) \geq \frac{1}{2} \quad \text{and} \quad F^*(z-1) < \frac{1}{2}. \quad (21)$$

We can show that

$$\lim_{n \rightarrow \infty} \mathbb{P}(z_n \leq z+1) = 1,$$

see the proof of Proposition 3.4 in Section 6.3. Therefore, if we define the event $A_n = \{z_n \leq z+1\}$, then conditioned on A_n the degrees in $\mathbf{D}_n^z$ are bounded by $z+1$. Hence, if we connect these stubs using the configuration model, and let $\mathcal{S}_n^z$ denote the event that G_n^z is simple, then it follows, see e.g., Van Der Hofstad (2016) Theorem 7.12, that there exist a constant $\bar{s} > 0$, such that

$$\lim_{n \rightarrow \infty} \mathbb{P}(\mathcal{S}_n) = \lim_{n \rightarrow \infty} \mathbb{P}(\mathcal{S}_n^z, A_n) + \mathbb{P}(A_n^c) = \bar{s}.$$

From this and Proposition 3.4 we obtain the following corollary.

Corollary 3.5. Let $\mathbf{D}_n \in \mathcal{D}_{\eta, \varepsilon}(f, f^*)$, $G_n = \text{DGA}(\mathbf{D}_n)$ and denote by $\mathcal{S}_n$ the event that G_n is simple. Then there exists a constant $\bar{s} > 0$ such that

$$\lim_{n \rightarrow \infty} \mathbb{P}(\mathcal{S}_n) = \bar{s}.$$

Note that by Lemma 2.1 it follows that if D is an integer valued random variable that satisfies, for some $\eta > 0$,

$$v := \mathbb{E}[D] < \infty \quad \text{and} \quad \mathbb{E}[D^{1+\eta}] < \infty,$$

then a degree sequence $\mathbf{D}_n$ generated by the IID algorithm satisfies $\mathbf{D}_n \in \mathcal{D}_{\eta, \varepsilon}$ for any $0 < \varepsilon < \eta/(8+4\eta)$, while

$$\mathbb{P}(\Omega_n) \geq 1 - O(n^{-\varepsilon}),$$

as $n \rightarrow \infty$. Therefore, if we want to generate maximally disassortative graphs with limit degree density f , we can first generate a degree sequence using the IID algorithm, by sampling from f , and then connect the nodes using the DGA. From Corollary 3.5 it now follows that, in order to generate maximally disassortative simple graphs, we could simply repeat this procedure until the resulting graph is simple. We summarized this procedure in Algorithm 2.

Algorithm 2 (Disassortative simple graph algorithm)

- 1: Input: A density function f and large integer n .
- 2: **repeat**
- 3: Sample a degree sequence $\mathbf{D}_n$ from f using the IID algorithm
- 4: Set $G_n = \text{DGA}(\mathbf{D}_n)$
- 5: **until** G_n is not simple
- 6: Output: G_n .

4. Spearman's Rho and the Tail of the Degree Distribution

We now investigate the influence of the degree distribution on the value of Spearman's rho, on maximally disassortative graphs, i.e., graphs generated by the DGA. We show that the tail of the distribution does not influence this value. This is achieved by transforming a given degree distribution, such that the asymptotic behavior of the tail of this distribution is preserved, while we increase the probability mass of the corresponding size-biased degree density at one.

Let us start by considering a degree distribution f , for which the size-biased distribution f^* satisfies $f^*(1) \geq 1/2$, i.e., f^* has median 1. Observe that in this case we have

$$h(k, l) = \begin{cases} f^*(k) & \text{if } k > 1 \text{ and } l = 1 \\ 2f^*(1) - 1 & \text{if } k = 1 \text{ and } l = 1 \\ f^*(l) & \text{if } k = 1 \text{ and } l > 1 \\ 0 & \text{else.} \end{cases}$$

Hence, if D_*, D^* have joint distribution h , as defined in (10), then

$$\begin{aligned} \mathbb{E}[\mathcal{F}^*(D_*)\mathcal{F}^*(D^*)] &= \sum_{k,l=1}^{\infty} \mathcal{F}^*(k)\mathcal{F}^*(l)h(k, l) \\ &= f^*(1)^2(2f^*(1) - 1) + 2f^*(1) \sum_{l=2}^{\infty} \mathcal{F}^*(l)f^*(l) \\ &\geq f^*(1)^2(2f^*(1) - 1) + 4f^*(1)^2 \sum_{l=2}^{\infty} f^*(l) \\ &= f^*(1)^2(2f^*(1) - 1) + 4f^*(1)^2(1 - f^*(1)) \\ &= 3f^*(1)^2 - 2f^*(1)^3, \end{aligned}$$

where we used that $\mathcal{F}^*(l) \geq 2f^*(1)$ for all $l > 1$. From this it follows that whenever $f^*(1) \geq 1/2$ and D_*, D^* have joint distribution h ,

$$3\mathbb{E}[\mathcal{F}^*(D_*)\mathcal{F}^*(D^*)] - 3 \geq 9f^*(1)^2 - 6f^*(1)^3 - 3. \quad (22)$$

Since the function on the right side of (22) is strictly monotonically increasing and is 0 when $f^*(1) = 1$, it follows that the limit of Spearman's rho on maximally disassortative graphs can be bounded from below by a value that is arbitrary close to 0, if $f^*(1)$ is large enough. Moreover, using that h is the joint degree distribution of graphs with minimal value of $\tilde{\rho}$, we have the following result.

Proposition 4.1. *Let f and f^* be such that $f^*(1) \geq \frac{1}{2}$ and $G_n \in \mathcal{G}_{n,\varepsilon}(f, f^*)$. Then, for any $0 < \delta < \min(\varepsilon, 1/2)$ and $K > 0$,*

$$\mathbb{P}(\tilde{\rho}(G_n) \geq 9f^*(1)^2 - 6f^*(1)^3 - 3 - Kn^{-\delta}) \geq 1 - O(n^{-\varepsilon+\delta} + \mathbb{P}(\Omega_n^c)),$$

as $n \rightarrow \infty$.

Given $0 < \omega < 1$, we now describe a construction that transforms any given distribution f with support on the positive integers, into a distribution f_ω , with support on the positive integers, such that $f_\omega(1) = \omega$ and

$$\lim_{k \rightarrow \infty} \frac{1 - F_\omega(k)}{1 - F(k)} = 1, \quad (23)$$

where F and F_ω are the cumulative distribution functions of f and f_ω , respectively. Observe that this means that the tail behavior of F_ω is the same as that of F .

Algorithm 3 (ω -transform of a probability density f)

- 1: Input: a probability density f , corresponding cdf F and $0 < \omega < 1$
- 2: Let K_ω be the smallest integer such that $F(K_\omega) > \omega$.
- 3: Set $x = F(K_\omega) - \omega > 0$ and $f_\omega(1) = \omega$.
- 4: **if** $K_\omega = 1$ **then**
- 5: Set $f_\omega(2) = f(2) + x$
- 6: **else**
- 7: **for** $2 \leq k \leq K_\omega$ **do**
- 8: Set $f_\omega(k) = x/(K_\omega - 1)$
- 9: **end for**
- 10: Set $f_\omega(K_\omega + 1) = f(K_\omega + 1)$
- 11: **end if**
- 12: **for** $k > K_\omega$ **do**
- 13: Set $f_\omega(k) = f(k)$
- 14: **end for**
- 15: Output: Probability density f_ω .

To see that f_ω defines a probability density function we compute

$$\begin{aligned} \sum_{t=1}^{\infty} f_\omega(t) &= \omega + \sum_{t=2}^{K_\omega+1} f_\omega(t) + \sum_{t=K_\omega+2}^{\infty} f(t) \\ &= \omega + x + f(K_\omega + 1) + 1 - F(K_\omega + 1) = 1. \end{aligned}$$

Moreover, since $f_\omega(k) = f(k)$ for all $k > K_\omega$ it follows that F_ω satisfies (23). We refer to f_ω as the ω -transform of f .

Proving Theorem 2.3. Let $-1 < \rho < 0$. Then with this transformation we can now transform a given distribution f , to get a distribution f_ρ whose size-biased distribution f_ρ^* satisfies

$$9f_\rho^*(1)^2 - 6f_\rho^*(1)^3 - 3 > \rho, \quad (24)$$

without affecting the asymptotic behavior of the tail of the original distribution f . Since the left hand side of (24) is monotonically increasing as a function of $f_\rho^*(1)$ and equal to -1 for some $f_\rho^*(1) \geq 1/2$, this implies that any $f_\rho^*(1)$ which solves (24), satisfies $f_\rho^*(1) \geq 1/2$. It then follows from Proposition 4.1 that for any sequence of graphs $G_n \in \mathcal{G}_{\eta, \varepsilon}(f_\rho, f_\rho^*)$,

$$\lim_{n \rightarrow \infty} \mathbb{P}(\tilde{\rho}(G_n) > \rho) = 1,$$

which proves Theorem 2.3. The details can be found in Section 6.5.

The construction we use for creating the adversary degree distribution f_ρ has one downside. In order to construct degree distributions such that $\tilde{\rho}(G_n)$ is arbitrary close to zero, the value of $f^*(1)$ should be arbitrary close to 1. Therefore, these distributions might not resemble real-world situations. The reason for this downside is that the construction is based on the very crude lower bound (22) on Spearman's rho, for which we had to assume $f^*(1) \geq 1/2$. This bound can be improved to an arbitrary value for $f^*(1)$ by carefully considering the contributions of different terms. However, for our purpose $f^*(1) \geq 1/2$ suffices.

Proving Corollary 2.4. As we mentioned in Section 2.4, Theorem 2.3 states that the minimal value of Spearman's rho is not determined by the tail of the distribution. Now let F be regularly varying with exponent $\gamma > 1$ and slowly varying function $\mathcal{L}$, see (13). Pick any $-1 < \rho < 0$ and let F_ρ be the transformed distribution, given by Theorem 2.3. We show that F_ρ is again regularly varying with exponent γ . Note that for this it is enough to show that $(1 - F_\rho(x))x^\gamma$ is slowly varying. To this end fix $t > 0$ and write

$$\begin{aligned} \frac{1 - F_\rho(tk)}{t^{-\gamma}(1 - F_\rho(k))} &= \left(\frac{1 - F_\rho(tk)}{1 - F(tk)} \right) \left(\frac{1 - F(k)}{1 - F_\rho(k)} \right) \left(\frac{1 - F(tk)}{t^{-\gamma}(1 - F(k))} \right) \\ &= \left(\frac{1 - F_\rho(tk)}{1 - F(tk)} \right) \left(\frac{1 - F(k)}{1 - F_\rho(k)} \right) \frac{\mathcal{L}(tk)}{\mathcal{L}(k)}. \end{aligned}$$

The product of the first two terms converge to 1, as $k \rightarrow \infty$, by Theorem 2.3, while this holds for the last term since $\mathcal{L}$ is slowly varying. Summarizing, we have

$$\lim_{k \rightarrow \infty} \frac{1 - F_\rho(tk)}{t^{-\gamma}(1 - F_\rho(k))} = 1,$$

which proves that $(1 - F_\rho(x))x^\gamma$ is slowly varying and hence F_ρ is regularly varying with exponent γ . This proves Corollary 2.4.

5. Spearman's Rho on Maximal Disassortative Graphs

In this section we use numerical experiments to illustrate the behavior of Spearman's rho for two types of degree distributions, regularly varying and Poisson. Each of these types has a parameter that can serve as a proxy for the way in which the mass of the probability density function is distributed over its support. For the regularly varying distribution this is the exponent γ , while for the Poisson distribution it is the mean λ . We will refer to these as the parameters of the distribution.

For the simulations we generated degree sequences $\mathbf{D}_n$ by sampling from the given distribution, using the IID algorithm, for different sizes n and values for the parameters. We then generated graphs G_n using the DGA. For each combination of size and parameter, we generated 10^3 graphs in this manner and computed $\rho(G_n)$, as defined in (6), on each of them. This gives us 10^3 samples of Spearman's rho on maximal disassortative graphs with the given size and degree distribution. We chose to compute $\rho(G_n)$ instead of $\tilde{\rho}(G_n)$ to show how our theoretical results for the latter translate to the former, which it approximates. The results for $\tilde{\rho}(G_n)$ are almost indistinguishable from those for $\rho(G_n)$ and hence we do not show them.

To analyze the speed of convergence of $\rho(G_n)$ we computed for each combination of size and parameter

$$X_n := |\rho(G_n) - \mathbb{E}'[\rho(G_n)]|,$$

where $\mathbb{E}'$ denotes the empirical mean, based on the 10^3 realizations per such combination. We then plotted the empirical inverse cumulative distribution of X_n for different sizes $n = 10^4, 10^5, 10^6$ and 10^7 . The results are shown in Figures 5 and 6.

In addition, to investigate the limit of Spearman's rho in maximally disassortative graphs, we computed $\mathbb{E}'[\rho(G_n)]$, with $n = 10^7$, for several values of the parameter of the distribution. We then plotted these values with respect to the parameter in Figure 7.

We now describe the specific distributions we used for the simulations and discuss the results.

5.1. Scale-Free Degree Distribution

Let X have a Pareto distribution with scale 1 and shape $\gamma > 1$, i.e.,

$$f_X(t) = \begin{cases} \gamma t^{-1-\gamma} & \text{if } t \geq 1 \\ 0 & \text{else,} \end{cases} \quad 1 - F_X(t) = \begin{cases} t^{-\gamma} & \text{if } t \geq 1 \\ 1 & \text{else,} \end{cases}$$

and define $D = \lfloor X \rfloor$. Then we have that $1 - F(k) = 1 - F_X(k+1)$, so that F is regularly varying with exponent $\gamma > 1$, while

$$f(k) = F(k) - F(k-1) = k^{-\gamma} - (k+1)^{-\gamma}. \quad (25)$$

Standard calculations yield that $\sum_{k=0}^{\infty} k f(k) = \zeta(\gamma)$, where ζ is the Riemann zeta function. Therefore we have that

$$f^*(k) = \frac{k f(k)}{\zeta(\gamma)},$$

so that $f^*(1) = (1 - 2^{-\gamma})/\zeta(\gamma)$ which is increasing in γ . Moreover $9f^*(1)^2 - 6f^*(1)^3 - 3 > -1$ for all $\gamma \geq 2.5$, which places it in the class of adversary distributions we considered in the previous section.

From Figure 5 we see that $\rho(G_n)$ is already strongly concentrated around its mean when $n = 10^5$. Even when the degree distribution has infinite variance ($\gamma = 1.5$) we have that $X_n \leq 0.025$, with reasonable high probability, for graphs of size $n = 10^5$. This shows, complementary to Theorem 2.2, that the DGA performs very well in practice, with respect to the convergence of Spearman's rho to the minimal achievable value $\rho(D_*, D^*)$.

Interestingly, the simulations suggest that the concentration of $\rho(G_n)$ around its mean for graphs of small size becomes tighter when γ decreases. Compare, for instance, the plots for $n = 10^4$ in the Figures 5(a)–5(c).

In Figure 7(a) we plotted the empirical average of $\rho(G_n)$ against the parameter γ of the degree density (25). Observe that, in contrary to the lower bound related to $f^*(1)$, we clearly see that Spearman's rho is strongly increasing as a function of γ and $\rho(G_n) > -0.8$ for $\gamma > 2$. Therefore it follows from Theorem 2.2 that the rank-correlation measure Spearman's rho on any graph with degree distribution (25) and $\gamma > 2$ will not have a value smaller than -0.8 . Moreover, when $\gamma \geq 2.5$ we see that $\mathbb{E}'[\rho(G_n)] > -0.5$. Since this is a lower bound for Spearman's rho on any graph with degree density (25), a consequence is that even if such graphs have a very disassortative joint degree structure they could potentially be classified differently.

Figure 5. Plot of the inverse cdf of X_n for graphs of different sizes and degree distribution (25), generated by the DGA, for three different choices of γ .

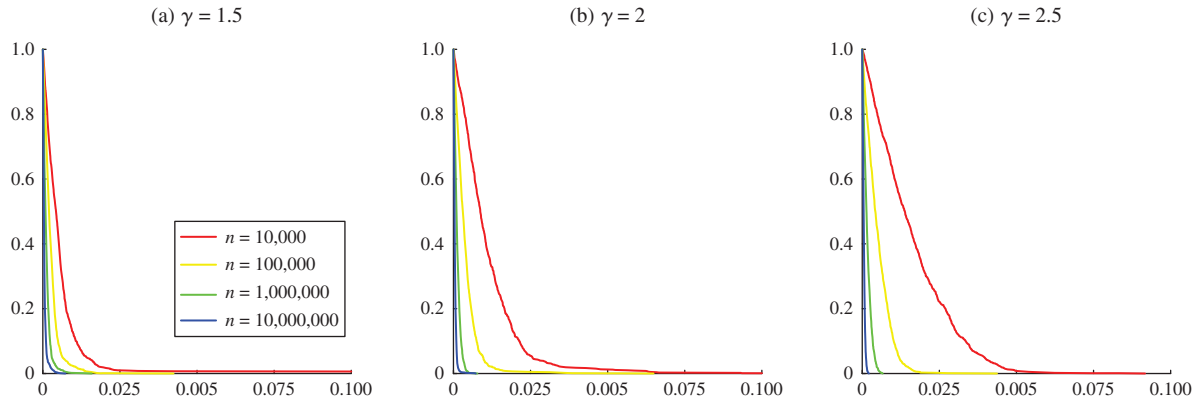


Figure 6. Plot of the inverse cdf of X_n for graphs of different sizes and Poisson degree distribution, generated by the DGA, for three different choices of λ .

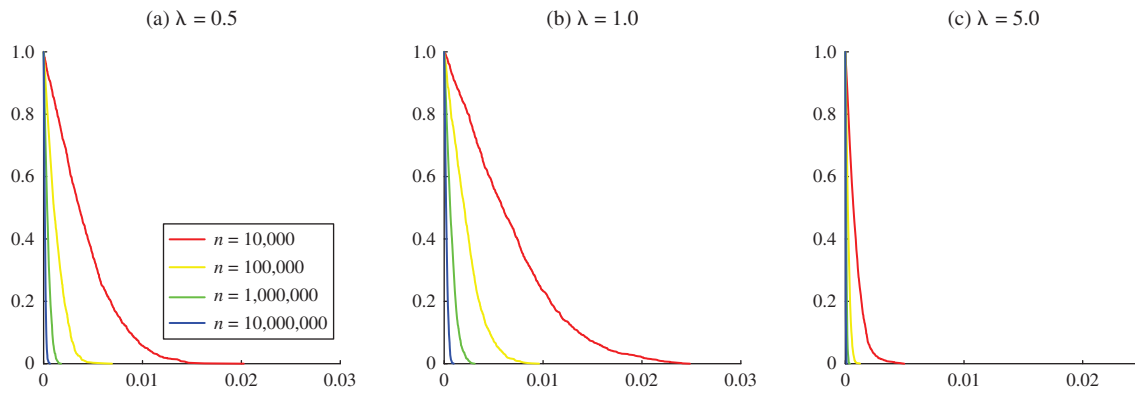
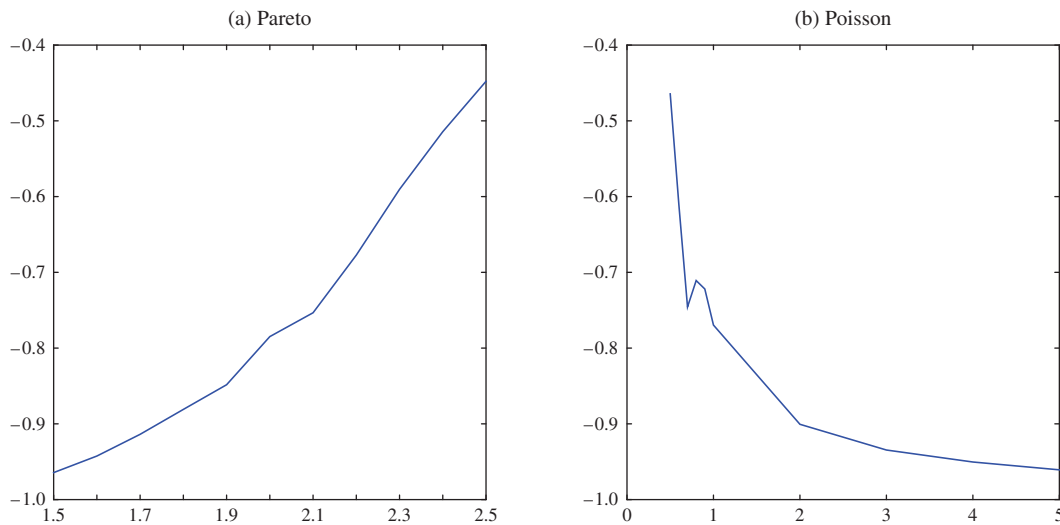


Figure 7. Plot of the empirical average of $\rho(G_n)$ for graph of size 10^7 and degree distributions (25) and Poisson, generated by the DGA, for different values of, respectively, γ and λ .



5.2. Poisson Degree Distribution

Let X be a Poisson random variable with mean λ and denote its probability density by f . Then it follows that $f^*(k) = f(k-1)$. Hence $f^*(1) = e^{-\lambda}$ is a decreasing function of λ and $9f^*(1)^2 - 6f^*(1)^3 - 3 > -1$ for at least all $\lambda \leq 0.4$. This is opposite to the degree distribution (25), where $f^*(1)$ was an increasing function of the parameter γ . This is reflected in Figure 7(b), where we see that $\mathbb{E}[\rho(G_n)]$ decreases with λ . Here we again see that the shape of the

degree distribution strongly influences the value of $\rho(G_n)$ for maximally disassortative graphs, and hence the minimal value that Spearman's rho can attain for any graph with this degree distribution. Note that, in contrast to the case with the regularly varying distribution, $\rho(G_n)$ is not monotonic with respect to λ . This could be due to the fact that the Poisson density is non-monotonic, while the density (25) is monotonically decreasing.

In addition we also observe that, similar to the previous setting, the DGA performs very well with respect to the convergence of $\rho(G_n)$. Already for very reasonable sizes, $n \geq 10^5$, the deviations around the mean are, with reasonable high probability, smaller than 0.02 for all three values of λ .

5.3. Important Observations and Insights

The simulations that we did show that the full shape of the degree distribution is of crucial importance for the minimal value that Spearman's rho can attain, which is in agreement with the theoretical results. However, in contrast to the artificial distribution we constructed in Theorem 2.3, which satisfied $f^*(1) \geq 1/2$, we observe that already for very reasonable degree distributions this minimum is much larger than -1 . Therefore, one should be careful when classifying a network as not being very disassortative when a small negative value of Spearman's rho is computed.

The simulations further suggest that the distribution of the mass of the probability density function f over its support is an important characteristic for determining the minimal value for Spearman's rho. To see this, consider the probability density (25) and observe that if we increase γ , then the value of $f(k)$ decreases for large k while $f(1)$ increases. The mean λ of the Poisson distribution can be used in a similar way, although in this case we need to decrease λ in order to increase $f(1)$ and decrease $f(k)$ for large values of k . For both distributions we see in Figure 7 that, as $f(1)$ increases (increasing γ /decreasing λ), the value of $\rho(G_n)$ in maximally disassortative graphs with this degree distribution increases and seems to approach 0. On the other hand, as we decrease $f(1)$ and the values of $f(k)$ for large k increase (decreasing γ /increasing λ) the minimal value of Spearman's rho decreases and seems to go to -1 . For instance, $|\mathbb{E}[\rho(G_n)]| \approx 6 \times 10^{-4}$ for degree distribution (25) with $\gamma = 10$ and -0.99 for $\gamma = 1.2$. For a Poisson degree distribution we have $|\mathbb{E}[\rho(G_n)]| \approx -0.027$ when $\lambda = 0.05$ and -0.978 when $\lambda = 10$. This demonstrates that the more mass the probability density function $f(k)$ has at large values of k , the more negative the minimal value for Spearman's rho. On the other hand, when $f(k)$ has more mass at small k this minimal value becomes much less negative and can approach zero.

6. Proofs

Here we prove the results stated in this paper.

6.1. Generating Degree Sequences $\mathbf{D}_n \in \mathcal{D}_{\eta, \varepsilon}(f, f^*)$

Proof of Lemma 2.1. We remark that altering the last degree by at most 1, to make the sum even, constitutes a correction term of order n^{-1} . Hence we will consider the degrees D_i as i.i.d. samples from D .

Now fix $\varepsilon \leq \eta/(8+4\eta)$ and define the events

$$A_n = \left\{ \sum_{k=0}^{\infty} \left| \sum_{t=0}^k f_n(t) - f(t) \right| \leq n^{-\eta/(2+2\eta)} \right\}$$

$$B_n = \left\{ \sum_{k=0}^{\infty} |f_n^*(k) - f^*(k)| \leq n^{-\varepsilon} \right\}$$

and notice that $\mathbb{P}(\Omega_n^c) \leq \mathbb{P}(A_n^c) + \mathbb{P}(A_n \cap B_n^c)$. For the first term we have, using Markov's inequality,

$$\mathbb{P}(A_n^c) \leq n^{\eta/(2+2\eta)} \mathbb{E}[d_1(f_n, f)] \leq O(n^{-\eta/(2+2\eta)}) = O(n^{-\varepsilon}),$$

as $n \rightarrow \infty$, where the second inequality follows from Chen and Olvera-Cravioto (2015, Proposition 4.2) and the last since $-\eta/(2+2\eta) < -\eta/(8+4\eta)$. Hence, we need to show that, as $n \rightarrow \infty$

$$\mathbb{P}(A_n \cap B_n^c) \leq O(n^{-\varepsilon}).$$

For this we compute that,

$$\begin{aligned} \sum_{k=0}^{\infty} |f_n^*(k) - f^*(k)| &= \sum_{k=0}^{\infty} \left| \frac{1}{L_n} \sum_{i=1}^n D_i \mathbb{1}_{\{D_i=k\}} - \frac{\mathbb{E}[D \mathbb{1}_{\{D=k\}}]}{\mu} \right| \\ &\leq \sum_{k=0}^{\infty} \left| \frac{1}{L_n} - \frac{1}{\mu n} \right| \sum_{i=1}^n D_i \mathbb{1}_{\{D_i=k\}} \end{aligned}$$

$$\begin{aligned} & + \sum_{k=0}^{\infty} \frac{1}{\mu n} \left| \sum_{i=1}^n D_i \mathbb{1}_{\{D_i=k\}} - \mathbb{E}[D \mathbb{1}_{\{D=k\}}] \right| \\ & = \frac{|L_n - \mu n|}{\mu n} + \sum_{k=0}^{\infty} \frac{1}{\mu n} \left| \sum_{i=1}^n X_{ik} \right|, \end{aligned}$$

where we defined $X_{ik} = D_i \mathbb{1}_{\{D_i=k\}} - \mathbb{E}[D \mathbb{1}_{\{D=k\}}]$. Now observe that conditioned on A_n we have that

$$\begin{aligned} \frac{|L_n - \mu n|}{\mu n^{1-\varepsilon}} & \leq \frac{n^{1-\eta/(2+2\eta)}}{\mu n^{1-\varepsilon}} \leq \mu^{-1} n^{\varepsilon-\eta/2(1+\eta)} \\ & \leq \mu^{-1} n^{-\eta/4(1+\eta)} \leq \mu^{-1} n^{-\eta/4(1+2\eta)}, \end{aligned}$$

so that

$$\mathbb{P}(A_n \cap B_n^c) \leq O(n^{-\varepsilon}) + \mathbb{P}\left(\sum_{k=0}^{\infty} \frac{1}{\mu n} \left| \sum_{i=1}^n X_{ik} \right| > n^{-\varepsilon}\right),$$

as $n \rightarrow \infty$.

To analyze the last probability take $a_n = \lfloor n^{2\varepsilon/\eta} \rfloor$. Then,

$$\begin{aligned} & \mathbb{P}\left(\sum_{k=0}^{\infty} \frac{1}{\mu n} \left| \sum_{i=1}^n X_{ik} \right| > n^{-\varepsilon}\right) \\ & \leq \frac{1}{\mu n^{1-\varepsilon}} \sum_{k=0}^{a_n} \mathbb{E}\left[\left|\sum_{i=1}^n X_{ik}\right|^2\right]^{1/2} + \frac{1}{\mu n^{1-\varepsilon}} \sum_{k=a_n+1}^{\infty} \mathbb{E}\left[\left|\sum_{i=1}^n X_{ik}\right|\right] \\ & \leq \frac{1}{\mu n^{1/2-\varepsilon}} \sum_{k=0}^{a_n} \text{Var}(X_{1k})^{1/2} + \frac{n^{\varepsilon}}{\mu} \sum_{k=a_n+1}^{\infty} \mathbb{E}[|X_{1k}|] \\ & \leq \frac{1}{\mu n^{1/2-\varepsilon}} \sum_{k=0}^{a_n} k + \frac{2n^{\varepsilon}}{\mu} \sum_{k=a_n+1}^{\infty} \mathbb{E}[D \mathbb{1}_{\{D=k\}}] \\ & \leq \frac{a_n(a_n+1)}{2\mu n^{1/2-\varepsilon}} + \frac{2n^{\varepsilon}}{\mu} \mathbb{E}[D \mathbb{1}_{\{D>a_n\}}] \\ & \leq \frac{a_n(a_n+1)}{2\mu n^{1/2-\varepsilon}} + 2n^{\varepsilon} a_n^{-\eta} \\ & = O(n^{-\varepsilon}), \end{aligned}$$

as $n \rightarrow \infty$. Here, for the last line, we used that

$$\frac{4\varepsilon}{\eta} - \frac{1}{2} + \varepsilon \leq \frac{1}{2+\eta} - \frac{1}{2} + \frac{\eta}{8+4\eta} = -\frac{\eta}{8+4\eta} \leq -\varepsilon,$$

so that

$$n^{\varepsilon-1/2} a_n^2 = O(n^{4\varepsilon/\eta-1/2+\varepsilon}) = O(n^{-\varepsilon}),$$

as $n \rightarrow \infty$. $\square$

6.2. Optimality of DGA

Theorem 3.3 is a consequence of the following lemma.

Lemma 6.1. Consider a sequence $0 \leq a_1 \leq \dots \leq a_m$ and let $\mathcal{P}_m$ denote the set of permutations of $\{1, \dots, m\}$. Then

$$\min_{\sigma \in \mathcal{P}_m} \sum_{k=1}^m a_k a_{\sigma(k)} = \sum_{k=1}^m a_k a_{m-k+1}$$

and this minimum is achievable for a permutation σ if and only if

$$a_{\sigma(1)} \geq a_{\sigma(2)} \geq \dots \geq a_{\sigma(m)}.$$

Proof. $[\Rightarrow]$ If $a_{\sigma(1)} \geq \dots \geq a_{\sigma(n)}$ then $\sum_{k=1}^m a_k a_{\sigma(k)} = \sum_{k=1}^m a_k a_{m-k+1}$
 $[\Leftarrow]$ Assume that $\sigma = \arg \min_{\sigma \in S_m} \sum_{k=1}^m a_k a_{\sigma(k)}$ but there exist $a_i < a_j$ such that $a_{\sigma(i)} < a_{\sigma(j)}$. Consider $\sigma^* = \sigma \cdot (ij)$ then $\sum_{k=1}^m a_k a_{\sigma(k)} - \sum_{k=1}^m a_k a_{\sigma^*(k)} = (a_i - a_j)(a_{\sigma(i)} - a_{\sigma(j)}) > 0$ which contradicts the initial assumption. $\square$

Proof of Theorem 3.3. Consider a degree sequence $\mathbf{D}_n$, rank it in ascending order and let $\phi(k)$ denotes the node with rank k among this degree sequence, as defined in the description of the DGA. Now define the sequence $\mathbf{a}_{L_n}$ by

$$a_k = \mathcal{F}_n^*(D_{\phi(i)}) \quad \text{for all } \sum_{t=1}^{i-1} D_{\phi(t)} < k \leq \sum_{t=1}^i D_{\phi(t)}, \quad (26)$$

where we use the convention that $\sum_{t=1}^0 D_{\phi(t)} = 0$. With this definition, the sequence $\mathbf{a}_{L_n}$ looks as follows

$$\begin{aligned} a_1 \leq \dots \leq a_k \leq \dots \\ = \underbrace{\mathcal{F}_n^*(D_{\phi(1)}) \leq \dots \leq \mathcal{F}_n^*(D_{\phi(1)})}_{D_{\phi(1)}} \leq \underbrace{\mathcal{F}_n^*(D_{\phi(2)}) \leq \dots \leq \mathcal{F}_n^*(D_{\phi(2)})}_{D_{\phi(2)}} \leq \dots \end{aligned}$$

Next, we note that for each graph $G \in \mathcal{G}(\mathbf{D}_n)$ there exists a permutation σ_G of such that

$$\sum_{i \rightarrow j \in G} \mathcal{F}_n^*(D_i) \mathcal{F}_n^*(D_j) = \sum_{k=1}^{L_n} a_k a_{\sigma_G(k)}.$$

Any directed graph, has a corresponding permutation σ of $\{1, \dots, L_n\}$ which defines how the outbound and inbound stubs of the bi-degree sequence are paired to obtain the graph. However, not every such permutation corresponds to a graph which is the bi-directed version of an undirected graph, i.e., for each edge $i \rightarrow j$ there is exactly one edge $j \rightarrow i$. Therefore let $\mathcal{P}(\mathbf{D}_n)$ denote the set of all permutations of $\{1, \dots, L_n\}$ which do corresponds to an undirected graph, in its directed representation. Then the optimization problem (12) is equivalent to the following problem

$$\min_{\sigma \in \mathcal{P}(\mathbf{D}_n)} \sum_{k=1}^{L_n} a_k a_{\sigma^*(k)}. \quad (27)$$

Now, recall the partitioned representation of the DGA we introduced in Section 3.2, see Figure 4. From this description of the algorithm it is not hard to see that, if $\mathbf{a}_{L_n}$ is defined by (26), then there exists a permutation σ^* with the property that

$$a_{\sigma^*(1)} \geq a_{\sigma^*(2)} \geq \dots \geq a_{\sigma^*(L_n)},$$

such that the DGA pairs the stubs corresponding to a_i and $a_{\sigma^*(i)}$. Therefore, Lemma 6.1 implies that

$$\sum_{k=1}^{L_n} a_k a_{\sigma^*(k)} = \min_{\sigma \in \mathcal{P}_{L_n}} \sum_{k=1}^{L_n} a_k a_{\sigma(k)},$$

where $\mathcal{P}_{L_n}$ denotes the set of all permutations of $\{1, \dots, L_n\}$. Since $\mathcal{P}(\mathbf{D}_n) \subseteq \mathcal{P}_{L_n}$, this implies that

$$\sum_{k=1}^{L_n} a_k a_{\sigma^*(k)} = \min_{\sigma \in \mathcal{P}(\mathbf{D}_n)} \sum_{k=1}^{L_n} a_k a_{\sigma(k)},$$

which proves that the DGA solves (27) and hence it solves (12). $\square$

6.3. Simplicity of G_n^*

Proof of Proposition 3.4. Let z_n and z be defined as in (14) and (21), respectively, and define the event

$$A_n = \{z_n \leq z + 1\}.$$

Then, by definition of z , we have that $F^*(z + 1) > 1/2$ and hence, by definition of z_n ,

$$\begin{aligned} \mathbb{P}(z_n > z + 1, \Omega_n) &\leq \mathbb{P}(F_n^*(z + 1) < \frac{1}{2}, \Omega_n) \\ &\leq \mathbb{P}(F^*(z + 1) - \frac{1}{2} < |F_n^*(z + 1) - F_n^*(z + 1)|, \Omega_n) \end{aligned}$$

$$\begin{aligned} &\leq \frac{\mathbb{E}[|F^*(z+1) - F_n^*(z+1)|\mathbb{1}_{\{\Omega_n\}}]}{F^*(z+1) - 1/2} \\ &\leq O(n^{-\varepsilon}), \end{aligned}$$

as $n \rightarrow \infty$. Therefore, if we define $\Lambda_n = A_n \cap \Omega_n$, it is enough to show that

$$1 - \mathbb{P}(\mathcal{S}_n, \Lambda_n) \leq O(n^{-\varepsilon} + n^{-1/2} + n^{-\eta/2}).$$

In order to analyze this probability, note that by construction there are no self-loops in G_n^* . Moreover, a node i with $D_i < z_n$ can only have more than one edge to a node j when $D_j > N_{D_i}$. Hence, when G_n^* is not simple it means that for some $1 \leq k \leq z_n$ we must have that $0 < N_k < \max_{1 \leq j \leq n} D_j$, hence

$$(\mathcal{S}_n^*)^c \subseteq \bigcup_{k=1}^{z_n} \left\{ 0 < N_k < \max_{1 \leq j \leq n} D_j \right\}.$$

Therefore, if we denote $f_{\min} = \min_{1 \leq k \leq z} f(k) > 0$, it follows from the union bound that

$$\begin{aligned} 1 - \mathbb{P}(\mathcal{S}_n, \Lambda_n) &\leq \sum_{k=1}^{z+1} \mathbb{P}\left(0 < N_k < \max_{1 \leq j \leq n} D_j, \Lambda_n\right) \\ &= \sum_{k=1}^{z+1} \mathbb{P}\left(0 < f_n(k) < \frac{\max_{1 \leq j \leq n} D_j}{n}, \Lambda_n\right) \\ &\leq \sum_{k=1}^{z+1} \left(f(k) - n^{-\varepsilon} < \frac{\max_{1 \leq j \leq n} D_j}{n}, \Lambda_n\right) \\ &\leq \sum_{k=1}^{z+1} \mathbb{P}\left(f_{\min} < \frac{\max_{1 \leq j \leq n} D_j}{n} + n^{-\varepsilon}, \Lambda_n\right) \\ &\leq \frac{(z+1)n^{-1}\mathbb{E}[\max_{1 \leq j \leq n} D_j \mathbb{1}_{\{\Lambda_n\}}] + (z+1)n^{-\varepsilon}}{f_{\min}} \\ &\leq \frac{(z+1)\mathbb{E}[\max_{1 \leq j \leq n} D_j \mathbb{1}_{\{D_j > \sqrt{n}\}} \mathbb{1}_{\{\Lambda_n\}}]}{nf_{\min}} \\ &\quad + \frac{(z+1)n^{-1/2}}{f_{\min}} + \frac{(z+1)n^{-\varepsilon}}{f_{\min}}. \end{aligned}$$

The last two terms are $O(n^{-1/2} + n^{-\varepsilon})$, as $n \rightarrow \infty$. We will now show that also the first term is $O(n^{-\varepsilon} + n^{-\eta/2})$. For this we note that

$$\frac{\max_{1 \leq j \leq n} D_j \mathbb{1}_{\{D_j > \sqrt{n}\}}}{n} \leq \frac{1}{n} \sum_{i=1}^n D_i \mathbb{1}_{\{D_i > \sqrt{n}\}} = \frac{L_n}{n} (1 - F_n^*(\sqrt{n})),$$

and

$$1 - F^*(\sqrt{n}) = \mathbb{E}[D \mathbb{1}_{\{D > \sqrt{n}\}}] \leq n^{-\eta/2} \mathbb{E}[D^{1+\eta}].$$

Therefore we obtain that, as $n \rightarrow \infty$,

$$\begin{aligned} &\frac{\mathbb{E}[\max_{1 \leq j \leq n} D_j \mathbb{1}_{\{\Lambda_n\}} \mathbb{1}_{\{D_j > \sqrt{n}\}}]}{nf_{\min}} \\ &\leq \frac{\mathbb{E}[L_n |F_n^*(\sqrt{n}) - F^*(\sqrt{n})| \mathbb{1}_{\{\Lambda_n\}}]}{nf_{\min}} + \frac{\mathbb{E}[2L_n(1 - F^*(\sqrt{n})) \mathbb{1}_{\{\Lambda_n\}}]}{nf_{\min}} \\ &\leq \frac{\mathbb{E}[(\nu n + n^{1-\varepsilon}) \sup_{k \geq 0} |F_n^*(k) - F^*(k)| \mathbb{1}_{\{\Lambda_n\}}]}{nf_{\min}} \\ &\quad + \frac{\mathbb{E}[2(\nu n + n^{1-\varepsilon})(1 - F^*(\sqrt{n})) \mathbb{1}_{\{\Lambda_n\}}]}{nf_{\min}} \\ &\leq \frac{(\nu + n^{-\varepsilon})n^{-\varepsilon}}{nf_{\min}} + \frac{2(\nu + n^{-\varepsilon})n^{-\eta/2} \mathbb{E}[D^{1+\eta}]}{f_{\min}} \\ &\leq O(n^{-\varepsilon} + n^{-\eta/2}), \end{aligned}$$

which completes the proof. $\square$

6.4. Joint Degree Distribution

Here we address the convergence of the empirical joint degree density $h_n(k, l)$, as defined in Proposition 3.1, to the density $h(k, l)$, as defined in (10).

We will use two technical lemmas, which deal with the difference between the functions ψ_n and ψ , and $\mathcal{E}_n$ and $\mathcal{E}$.

Lemma 6.2. *Let $\mathbf{D}_n \in \mathcal{D}_{\eta, \varepsilon}(f, f^*)$. Then, for any $k, l \geq 0$, $0 < \delta < \varepsilon$ and $K > 0$*

$$\mathbb{P}(|\psi_n(k, l) - \psi(k, l)| \mathcal{E}_n(k, l) > Kn^{-\delta}, \Omega_n) \leq O(n^{-\varepsilon+\delta}).$$

Lemma 6.3. *Let $\mathbf{D}_n \in \mathcal{D}_{\eta, \varepsilon}(f, f^*)$. Then, for any $k, l \geq 0$, $K > 0$ and $0 < \delta < \varepsilon$,*

$$\mathbb{P}(|\mathcal{E}(k, l) - \mathcal{E}_n(k, l)| > Kn^{-\delta}, \Omega_n) \leq O(n^{-\varepsilon+\delta}).$$

The proof of both lemmas is postponed till the end of this section. We will first give the proof of Theorem 3.2.

Proof of Theorem 3.2. Let p be the smallest integer satisfying

$$1 - F^*(p) < f^*(1), \quad (28)$$

and define p_n as the smallest integer that satisfies

$$1 - F_n^*(p_n) < f_n^*(1).$$

Then we have that $\mathbb{P}(p_n = p)$ converges to one, since

$$\begin{aligned} & \mathbb{P}(p \neq p_n, \Omega_n) \\ &= \mathbb{P}(1 - F_n^*(p) \geq f_n^*(1), \Omega_n) \\ &\leq \mathbb{P}((F^*(p) - F_n^*(p)) + (f^*(1) - f_n^*(1)) > f^*(1) - 1 + F^*(p), \Omega_n) \\ &\leq \mathbb{P}\left(|F^*(p) - F_n^*(p)| > \frac{f^*(1) - 1 + F^*(p)}{2}, \Omega_n\right) \\ &\quad + \mathbb{P}\left(|f^*(1) - f_n^*(1)| > \frac{f^*(1) - 1 + F^*(p)}{2}, \Omega_n\right) \\ &\leq \frac{2\mathbb{E}[|F^*(p) - F_n^*(p)| \mathbb{1}_{\{\Omega_n\}}]}{f^*(1) - 1 + F^*(p)} + \frac{2\mathbb{E}[|f^*(1) - f_n^*(1)| \mathbb{1}_{\{\Omega_n\}}]}{f^*(1) - 1 + F^*(p)} \\ &\leq \frac{4n^{-\varepsilon}}{f^*(1) - 1 + F^*(p)} \leq O(n^{-\varepsilon}), \end{aligned}$$

as $n \rightarrow \infty$, where we used that by definition of p it holds that $f^*(1) - 1 + F^*(p) > 0$. Therefore, if we define the event $P_n = \{p = p_n\}$ and let $\Lambda_n = P_n \cap \Omega_n$, then

$$\mathbb{P}(\Lambda_n) \geq 1 - O(n^{-\varepsilon} + \mathbb{P}(\Omega_n^c)),$$

so that for Theorem 3.2 it is enough to show that

$$\mathbb{P}(\Xi_n^c, \Lambda_n) \leq O(n^{-\varepsilon+\delta}), \quad (29)$$

as $n \rightarrow \infty$.

Now, observe that p_n is the smallest degree such that nodes i with degree $D_i > p_n$ will be connected to nodes with degree 1, by the DGA, while p is the corresponding degree for the limit distribution. Therefore we have

$$\psi(k, l) = \begin{cases} 1 & \text{for all } k > p \text{ and } l = 1 \\ 1 & \text{for all } k = 1 \text{ and } l > p \\ \psi(k, l) & \text{for all } k \leq p \text{ and } l \leq p \\ 0 & \text{else} \end{cases} \quad (30)$$

while, on the event P_n , the same relations hold for ψ_n . The idea of the proof is to split the analysis into the three regions

$$(k = 1, l > p), \quad (k, l \leq p) \quad \text{and} \quad (k > p, l = 1).$$

The hard work is in the second region. However, since on the event Λ_n all degree are bounded by p , it suffices to analyze individual terms

$$|\psi_n(k, l)\mathcal{E}_n(k, l) - \psi(k, l)\mathcal{E}(k, l)|,$$

instead of the full sum

$$\sum_{k, l=0}^p |\psi_n(k, l)\mathcal{E}_n(k, l) - \psi(k, l)\mathcal{E}(k, l)|.$$

Recall that

$$\Xi_n = \left\{ \sum_{k, l=0}^{\infty} |h_n(k, l) - h(k, l)| \leq Kn^{-\delta} \right\},$$

and let us bound the probability in (29) as follows,

$$\mathbb{P}(\Xi_n^c, \Lambda_n) \leq \mathbb{P}\left(\sum_{k, l=1}^{\infty} |\psi_n(k, l) - \psi(k, l)|\mathcal{E}_n(k, l) > \frac{Kn^{-\delta}}{2}, \Lambda_n\right) \quad (31)$$

$$+ \mathbb{P}\left(\sum_{k, l=1}^{\infty} \psi(k, l)|(\mathcal{E}(k, l) - \mathcal{E}_n(k, l))| > \frac{Kn^{-\delta}}{2}, \Lambda_n\right). \quad (32)$$

We will first deal with (31). By (30) and conditioned on Λ_n , we have that $|\psi_n(k, l) - \psi(k, l)| \neq 0$, only when $k, l \leq p$. Hence we get, using the union bound,

$$\begin{aligned} & \mathbb{P}\left(\sum_{k, l=1}^{\infty} |\psi_n(k, l) - \psi(k, l)|\mathcal{E}_n(k, l) > \frac{Kn^{-\delta}}{2}, \Lambda_n\right) \\ &= \mathbb{P}\left(\sum_{k, l=1}^p |\psi_n(k, l) - \psi(k, l)|\mathcal{E}_n(k, l) > \frac{Kn^{-\delta}}{2}, \Lambda_n\right) \\ &\leq \sum_{k, l=1}^p \mathbb{P}\left(|\psi_n(k, l) - \psi(k, l)|\mathcal{E}_n(k, l) > \frac{Kn^{-\delta}}{2p^2}, \Lambda_n\right) \\ &\leq O(n^{-\varepsilon+\delta}), \end{aligned}$$

where the last line follows from Lemma 6.2.

Next we consider (32). First we use (30) to bound the term inside the probability as follows

$$\sum_{k, l=1}^{\infty} \psi(k, l)|(\mathcal{E}(k, l) - \mathcal{E}_n(k, l))| \leq \sum_{k=1}^p \sum_{l=1}^p \psi(k, l)|\mathcal{E}(k, l) - \mathcal{E}_n(k, l)| \quad (33)$$

$$+ \sum_{l=p+1}^{\infty} |\mathcal{E}(1, l) - \mathcal{E}_n(1, l)| \quad (34)$$

$$+ \sum_{k=p+1}^{\infty} |\mathcal{E}(k, 1) - \mathcal{E}_n(k, 1)| \quad (35)$$

We will start by analyzing (34). For this we notice that $\mathcal{E}_n(1, l) - \mathcal{E}(1, l) = f_n^*(l) - f^*(l)$, so that

$$\sum_{l=p+1}^{\infty} \psi(1, l)|\mathcal{E}(1, l) - \mathcal{E}_n(1, l)| \leq \sum_{l=0}^{\infty} |f_n^*(l) - f^*(l)|.$$

The upper bound for (35) is the same. Therefore, again using the union bound, we have that

$$\begin{aligned} & \mathbb{P}\left(\sum_{k, l=0}^{\infty} \psi(k, l)|(\mathcal{E}(k, l) - \mathcal{E}_n(k, l))| > \frac{Kn^{-\delta}}{2}, \Lambda_n\right) \\ &\leq 2\mathbb{P}\left(|f_n^*(l) - f^*(l)| > \frac{Kn^{-\delta}}{6}, \Omega_n\right) \\ &\quad + \sum_{k, l=0}^p \mathbb{P}\left(|\mathcal{E}(k, l) - \mathcal{E}_n(k, l)| > \frac{Kn^{-\delta}}{6p^2}, \Omega_n\right) \\ &\leq O(n^{-\varepsilon+\delta}). \end{aligned}$$

Here we used Lemma 6.3 to bound the last probability in the second line.

With this final result we have proven (29) and hence Theorem 3.2. $\square$

All that is left is to prove the two technical Lemmas 6.2 and 6.3. Due to the use of both a minimum and maximum, in the definitions of $\mathcal{E}_n(k, l)$ and $\mathcal{E}(k, l)$ and the double cases in $\psi_n(k, l)$ and $\psi(k, l)$, the proofs consists of many case distinctions, where we have to bound each specific case. In order to improve the readability of the proofs we define, for any $k, l \geq 0$, the following events

$$\begin{aligned} A_n &= \{1 - F_n^*(k) < F_n^*(l)\} \\ B_n &= \{1 - F_n^*(k-1) > F_n^*(l-1)\}, \\ I_n &= \{1 - F_n^*(k-1) \leq F_n^*(l)\}, \\ J_n &= \{1 - F_n^*(k) \geq F_n^*(l-1)\}. \end{aligned}$$

With these definitions we have that $\psi_n(k, l) = \mathbb{1}_{\{A_n\}} \mathbb{1}_{\{B_n\}}$. Moreover since $A_n^c \cap B_n^c = \emptyset$ we have that

$$1 - \psi_n(k, l) = \mathbb{1}_{\{A_n\}} \mathbb{1}_{\{B_n^c\}} + \mathbb{1}_{\{A_n^c\}} \mathbb{1}_{\{B_n\}}. \quad (36)$$

Where the event A_n and B_n determine the value of $\psi_n(k, l)$, so do the events I_n and J_n define the expression for $\mathcal{E}_n(k, l)$, as follows:

$$\mathcal{E}_n(k, l) = \begin{cases} f_n^*(k) & \text{on the event } I_n \cap J_n \\ 1 - F_n^*(k-1) - F_n^*(l-1) & \text{on the event } I_n \cap J_n^c \\ F_n^*(k) + F_n^*(l) - 1 & \text{on the event } I_n^c \cap J_n \\ f_n^*(l) & \text{on the event } I_n^c \cap J_n^c. \end{cases} \quad (37)$$

Note that by their definitions,

$$0 \leq \psi_n(k, l), \psi(k, l), \mathcal{E}_n(k, l), \mathcal{E}(k, l) \leq 1,$$

for all $k, l \geq 0$. In addition we will often use the following result

Lemma 6.4. *Let $k, l \geq 0$ be such that $1 - F^*(k) < F^*(l)$. Then*

$$\mathbb{P}(1 - F_n^*(k) \geq F_n^*(l), \Omega_n) \leq O(n^{-\varepsilon}),$$

as $n \rightarrow \infty$.

If, on the other hand, $1 - F^*(k) > F^*(l)$, then

$$\mathbb{P}(1 - F_n^*(k) \leq F_n^*(l), \Omega_n) \leq O(n^{-\varepsilon}),$$

as $n \rightarrow \infty$.

Proof. We will prove the first statement, since the proof for the second is similar. First we write

$$\begin{aligned} &\mathbb{P}(1 - F_n^*(k) \geq F_n^*(l), \Omega_n) \\ &= \mathbb{P}((F^*(k) - F_n^*(k)) + 1 - F^*(k) \geq F^*(l) + (F_n^*(l) - F^*(l)), \Omega_n) \\ &= \mathbb{P}((F^*(k) - F_n^*(k)) + (F^*(l) - F_n^*(l)) \geq F^*(l) - 1 + F^*(k), \Omega_n). \end{aligned}$$

Next we use the union bound and Markov's inequality to obtain

$$\begin{aligned} &\mathbb{P}(1 - F_n^*(k) \geq F_n^*(l), \Omega_n) \\ &\leq \mathbb{P}\left(|F^*(k) - F_n^*(k)| \geq \frac{F^*(l) - 1 + F^*(k)}{2}, \Omega_n\right) \\ &\quad + \mathbb{P}\left(|F^*(l) - F_n^*(l)| \geq \frac{F^*(l) - 1 + F^*(k)}{2}, \Omega_n\right) \\ &\leq \frac{4\mathbb{E}[\sup_{k \geq 0} |F_n^*(k) - F^*(k)| \mathbb{1}_{\{\Omega_n\}}]}{F^*(l) - 1 + F^*(k)} \\ &\leq \frac{4\mathbb{E}[\sum_{k=0}^{\infty} |f_n^*(k) - f^*(k)| \mathbb{1}_{\{\Omega_n\}}]}{F^*(l) - 1 + F^*(k)} = O(n^{-\varepsilon}), \end{aligned}$$

as $n \rightarrow \infty$, where we used $1 - F^*(k) < F^*(l)$ for the last equality. $\square$

Proof of Lemma 6.2. Note that the specific expression of $\psi(k, l)$ depends on the ordering between

$$1 - F^*(k) \quad \text{and} \quad F^*(l),$$

and

$$1 - F^*(k-1) \quad \text{and} \quad F^*(l-1).$$

Therefore, we need to consider all different cases ($<$, $=$, $>$), where we treat equality as a separate case. This gives a total of nine cases. However, there are several combinations that do not need to be considered. For instance, $1 - F^*(k) > F^*(l)$ implies that $1 - F^*(k-1) \geq F^*(l-1)$. In the end, we are left with the following cases:

- (I) $1 - F^*(k) < F^*(l)$ and $1 - F^*(k-1) < F^*(l-1)$
- (II) $1 - F^*(k) < F^*(l)$ and $1 - F^*(k-1) = F^*(l-1)$
- (III) $1 - F^*(k) < F^*(l)$ and $1 - F^*(k-1) > F^*(l-1)$
- (IV) $1 - F^*(k) = F^*(l)$ and $1 - F^*(k-1) > F^*(l-1)$
- (V) $1 - F^*(k) > F^*(l)$ and $1 - F^*(k-1) > F^*(l-1)$

We will start with the first case.

(I) $1 - F^*(k) < F^*(l)$ and $1 - F^*(k-1) < F^*(l-1)$

First, note that in this case $\psi(k, l) = 0$. Moreover, since $F^*(l-1) > 1 - F^*(k-1)$, it follows from Lemma 6.4 that

$$\begin{aligned} \mathbb{P}(B_n) &\leq \mathbb{P}(B_n, \Omega_n^c) + \mathbb{P}(\Omega_n^c) \\ &\leq O(n^{-\varepsilon} + \mathbb{P}(\Omega_n^c)). \end{aligned}$$

Hence, since $\psi_n(k, l) = 0$ on the event B_n^c , we have

$$\begin{aligned} &\mathbb{P}(|\psi_n(k, l) - \psi(k, l)| \mathcal{E}_n(k, l) > Kn^{-\delta}) \\ &\leq \mathbb{P}(\psi_n(k, l) \mathcal{E}_n(k, l) > Kn^{-\delta}, B_n^c) + \mathbb{P}(B_n) \\ &\leq O(n^{-\varepsilon} + \mathbb{P}(\Omega_n^c)). \end{aligned}$$

(II) $1 - F^*(k) < F^*(l)$ and $1 - F^*(k-1) = F^*(l-1)$

In this case we have

$$1 - F^*(k) < 1 - F^*(k-1) = F^*(l-1) < F^*(l),$$

so that, by Lemma 6.4

$$\mathbb{P}(I_n^c) \leq O(n^{-\varepsilon} + \mathbb{P}(\Omega_n^c)) \quad \text{and} \quad \mathbb{P}(J_n) \leq O(n^{-\varepsilon} + \mathbb{P}(\Omega_n^c)).$$

Therefore, using (37) and $1 - F^*(k-1) = F^*(l-1)$, it follows that

$$\begin{aligned} &\mathbb{P}(|\psi_n(k, l) - \psi(k, l)| \mathcal{E}_n(k, l) > Kn^{-\delta}) \\ &\leq \mathbb{P}(2\mathcal{E}_n(k, l) > Kn^{-\delta}) \\ &\leq \mathbb{P}(2\mathcal{E}_n(k, l) > Kn^{-\delta}, I_n^c, J_n) + \mathbb{P}(I_n) + \mathbb{P}(J_n) + \mathbb{P}(I_n, J_n^c) \\ &\leq \mathbb{P}\left(|1 - F_n^*(k-1) - F_n^*(l-1)| > \frac{Kn^{-\delta}}{2}\right) + O(n^{-\varepsilon} + \mathbb{P}(\Omega_n^c)) \\ &\leq \mathbb{P}\left(|F^*(k-1) - F_n^*(k-1)| > \frac{Kn^{-\delta}}{4}\right) \\ &\quad + \mathbb{P}\left(|F^*(l-1) - F_n^*(l-1)| > \frac{Kn^{-\delta}}{4}\right) + O(n^{-\varepsilon} + \mathbb{P}(\Omega_n^c)) \\ &\leq O(n^{-\varepsilon+\delta} + \mathbb{P}(\Omega_n^c)), \end{aligned}$$

where for the fifth line we used that

$$\begin{aligned} &|1 - F_n^*(k-1) - F_n^*(l-1)| \\ &= |1 - F_n^*(k-1) - (1 - F^*(k-1)) + F^*(l-1) - F_n^*(l-1)| \\ &\leq |F^*(k-1) - F_n^*(k-1)| + |F^*(l-1) - F_n^*(l-1)|, \end{aligned}$$

since $1 - F^*(k-1) = F^*(l-1)$.

Case (IV) can be dealt with using arguments similar to case (II), while case (V) is similar to (I). Therefore, there is only one case left.

(III) $1 - F^*(k) < F^*(l)$ and $1 - F^*(k-1) > F^*(l-1)$

We first note that, since $1 - F^*(k) < F^*(l)$,

$$\mathbb{P}(A_n^c) \leq O(n^{-\varepsilon} + \mathbb{P}(\Omega_n^c)),$$

by Lemma 6.4, and similarly, since $1 - F^*(k-1) > F^*(l-1)$,

$$\mathbb{P}(B_n^c) \leq O(n^{-\varepsilon} + \mathbb{P}(\Omega_n^c)).$$

Since for this case $\psi(k, l) = 1$, we have,

$$\begin{aligned} & \mathbb{P}(|\psi_n(k, l) - \psi(k, l)| \mathcal{E}_n(k, l) > Kn^{-\delta}) \\ &= \mathbb{P}((1 - \psi_n(k, l)) \mathcal{E}_n(k, l) > Kn^{-\delta}) \\ &\leq \mathbb{P}((1 - \psi_n(k, l)) \mathcal{E}_n(k, l) > Kn^{-\delta}, A_n, B_n) \\ &\quad + \mathbb{P}(A_n^c) + \mathbb{P}(B_n^c) + \mathbb{P}(A_n^c, B_n^c) \\ &\leq O(n^{-\varepsilon} + \mathbb{P}(\Omega_n^c)), \end{aligned}$$

where we used (36) for the third line. $\square$

Proof of Lemma 6.3. Similar to the proof of Lemma 6.2 we will have to consider different cases. Here these are with respect to the different relations between

$$1 - F^*(k-1) \quad \text{and} \quad F^*(l),$$

and

$$1 - F^*(k) \quad \text{and} \quad F^*(l-1),$$

which determine the expression for $\mathcal{E}(k, l)$. To analyze each case we will also need to distinguish between the different expression of $\mathcal{E}_n(k, l)$, which are determined by the events I_n and J_n .

We will consider the three cases where $1 - F^*(k) > F^*(l-1)$. The other six cases can be dealt with using similar arguments. First note that by Lemma 6.4

$$\mathbb{P}(J_n^c) \leq O(n^{-\varepsilon} + \mathbb{P}(\Omega_n^c)).$$

(I) $1 - F^*(k-1) < F^*(l)$ and $1 - F^*(k) > F^*(l-1)$

Similar to $\mathbb{P}(J_n^c)$, it follows from Lemma 6.4 that

$$\mathbb{P}(I_n^c) \leq O(n^{-\varepsilon} + \mathbb{P}(\Omega_n^c)).$$

Therefore, by conditioning on the different combinations of I_n and J_n , we get

$$\begin{aligned} & \mathbb{P}(|\mathcal{E}(k, l) - \mathcal{E}_n(k, l)| > Kn^{-\delta}) \\ &\leq \mathbb{P}(|\mathcal{E}(k, l) - \mathcal{E}_n(k, l)| > Kn^{-\delta}, I_n, J_n) + \mathbb{P}(J_n^c) + \mathbb{P}(I_n^c) + \mathbb{P}(J_n^c, I_n^c) \\ &\leq \mathbb{P}(|f^*(k) - f_n^*(k)| > Kn^{-\delta}, \Omega_n) + O(n^{-\varepsilon} + \mathbb{P}(\Omega_n^c)) \\ &\leq O(n^{-\varepsilon+\delta} + \mathbb{P}(\Omega_n^c)). \end{aligned}$$

(II) $1 - F^*(k-1) = F^*(l)$ and $1 - F^*(k) > F^*(l-1)$

Since $f^*(k) = F^*(k) - F^*(k-1)$,

$$\begin{aligned} |F_n^*(k) + F_n^*(l) - 1 - f^*(k)| &= |F_n^*(k) - F^*(k) + F_n^*(l) - 1 + F^*(k-1)| \\ &\leq |F_n^*(k) - F^*(k)| + |F_n^*(l) - F^*(l)|, \end{aligned}$$

from which it follows that

$$\begin{aligned} & \mathbb{P}(|\mathcal{E}(k, l) - \mathcal{E}_n(k, l)| > Kn^{-\delta}, I_n^c, J_n) \\ &\leq \mathbb{P}(|F_n^*(k) + F_n^*(l) - 1 - f^*(k)| > Kn^{-\delta}) \\ &\leq \mathbb{P}\left(|F_n^*(k) - F^*(k)| > \frac{Kn^{-\delta}}{2}\right) + \mathbb{P}\left(|F_n^*(l) - F^*(l)| > \frac{Kn^{-\delta}}{2}\right) \\ &\leq O(n^{-\varepsilon+\delta} + \mathbb{P}(\Omega_n^c)). \end{aligned}$$

Hence, we obtain

$$\begin{aligned} & \mathbb{P}(|\mathcal{E}(k, l) - \mathcal{E}_n(k, l)| > Kn^{-\delta}) \\ & \leq \mathbb{P}(|f^*(k) - f_n^*(k)| > Kn^{-\delta}) \\ & \quad + \mathbb{P}(|\mathcal{E}(k, l) - \mathcal{E}_n(k, l)| > Kn^{-\delta}, I_n^c, J_n) + 2\mathbb{P}(J_n^c) \\ & \leq O(n^{-\varepsilon+\delta} + \mathbb{P}(\Omega_n^c)). \end{aligned}$$

(III) $1 - F^*(k-1) > F^*(l)$ and $1 - F^*(k) > F^*(l-1)$

First we notice that in this case $\mathcal{E}(k, l) = F^*(l) + F^*(k) - 1$. Next, using Lemma 6.4, we have

$$\mathbb{P}(I_n) \leq O(n^{-\varepsilon+\delta} + \mathbb{P}(\Omega_n^c)).$$

Therefore it follows that

$$\begin{aligned} & \mathbb{P}(|\mathcal{E}(k, l) - \mathcal{E}_n(k, l)| > Kn^{-\delta}) \\ & \leq \mathbb{P}(|F^*(l) + F^*(k) - 1 - \mathcal{E}_n(k, l)| > Kn^{-\delta}, I_n^c, J_n) \\ & \quad + \mathbb{P}(I_n) + 2\mathbb{P}(J_n^c) \\ & \leq \mathbb{P}\left(|F_n^*(k) - F^*(k)| > \frac{Kn^{-\delta}}{2}\right) + \mathbb{P}\left(|F_n^*(l) - F^*(l)| > \frac{Kn^{-\delta}}{2}\right) \\ & \quad + \mathbb{P}(I_n) + 2\mathbb{P}(J_n^c) \\ & \leq O(n^{-\varepsilon+\delta} + \mathbb{P}(\Omega_n^c)). \quad \square \end{aligned}$$

6.5. Main Results

Here we will give the proofs of our two main results. We start with a useful result which we need to prove Theorem 2.2.

Proposition 6.5. *Let $G_n \in \mathcal{G}_{\eta, \varepsilon}(f, f^*)$ and let D_*, D^* be random variables with joint distribution h . Then, for any $0 < \delta < \varepsilon$ and $K > 0$,*

$$\mathbb{P}(|\tilde{\rho}(G_n) - \rho(D_*, D^*)| > Kn^{-\delta}) = O(n^{-\varepsilon+\delta} + \mathbb{P}(\Omega_n^c)),$$

as $n \rightarrow \infty$.

Proof. First we write

$$\begin{aligned} & |\tilde{\rho}(G_n) - \rho(D_*, D^*)| \\ & \leq 3 \left| \sum_{k, l=0}^{\infty} \mathcal{F}_n^*(k) \mathcal{F}_n^*(l) h_n(k, l) - \mathcal{F}^*(k) \mathcal{F}^*(l) h(k, l) \right| \\ & \leq 3 \left| \sum_{k, l=0}^{\infty} \mathcal{F}_n^*(k) \mathcal{F}_n^*(l) (h_n(k, l) - h(k, l)) \right| \\ & \quad + 3 \sum_{k, l=0}^{\infty} |\mathcal{F}_n^*(k) \mathcal{F}_n^*(l) - \mathcal{F}^*(k) \mathcal{F}^*(l)| h_n(k, l) \\ & \leq 12 \sum_{k, l=0}^{\infty} |h_n(k, l) - h(k, l)| + 24 \sup_k |F_n^*(k) - F^*(k)|. \end{aligned} \tag{38}$$

For the last inequality, we used

$$\begin{aligned} & \sum_{k, l=0}^{\infty} |\mathcal{F}_n^*(k) \mathcal{F}_n^*(l) - \mathcal{F}^*(k) \mathcal{F}^*(l)| h_n(k, l) \\ & \leq \sup_{k, l} |\mathcal{F}_n^*(k) \mathcal{F}_n^*(l) - \mathcal{F}^*(k) \mathcal{F}^*(l)| \\ & \leq \sup_{k, l} |\mathcal{F}_n^*(k) - \mathcal{F}^*(k)| \mathcal{F}_n^*(l) + \sup_{k, l} |\mathcal{F}_n^*(l) - \mathcal{F}^*(l)| \mathcal{F}^*(k) \\ & \leq 4 \sup_k |\mathcal{F}_n^*(k) - \mathcal{F}^*(k)| \leq 8 \sup_k |F_n^*(k) - F^*(k)|. \end{aligned}$$

Note that by Theorem 3.2

$$\mathbb{P}\left(12 \sum_{k,l=0}^{\infty} |h_n(k,l) - h(k,l)| > \frac{Kn^{-\delta}}{2}\right) = O(n^{-\varepsilon+\delta} + \mathbb{P}(\Omega_n^c)).$$

Moreover, on the event Ω_n ,

$$\sup_{k \geq 0} |F_n^*(k) - F^*(k)| \leq \sum_{k=0}^{\infty} |f_n^*(k) - f^*(k)| \leq n^{-\varepsilon}.$$

Hence, it follows from (38) and Markov's inequality that

$$\begin{aligned} & \mathbb{P}(|\rho(\tilde{G}_n) - \rho(D_*, D^*)| > Kn^{-\delta}) \\ & \leq \mathbb{P}\left(12 \sum_{k,l=0}^{\infty} |h_n(k,l) - h(k,l)| > \frac{Kn^{-\delta}}{2}\right) \\ & \quad + \mathbb{P}\left(24 \sup_k |F_n^*(k) - F^*(k)| > \frac{Kn^{-\delta}}{2}, \Omega_n\right) + O(\mathbb{P}(\Omega_n^c)) \\ & \leq O(n^{-\varepsilon+\delta} + \mathbb{P}(\Omega_n^c)) + \frac{48n^\delta}{K} \mathbb{E}\left[\sup_k |F_n^*(k) - F^*(k)| \mathbb{1}_{\{\Omega_n\}}\right] \\ & \leq O(n^{-\varepsilon+\delta} + \mathbb{P}(\Omega_n^c)). \quad \square \end{aligned}$$

We are now ready to give the proof of Theorem 2.2.

Proof of Theorem 2.2. Consider a graph $G_n \in \mathcal{G}_{\eta,\varepsilon}(f, f^*)$, denote its degree sequence by $\mathbf{D}_n$ and let $\hat{G}_n = \text{DGA}(\mathbf{D}_n)$. Then it follows from Proposition 6.5 that

$$\mathbb{P}(|\tilde{\rho}(\hat{G}_n) - \rho(D_*, D^*)| > Kn^{-\delta}) \leq O(n^{-\varepsilon+\delta} + \mathbb{P}(\Omega_n^c)),$$

which proves the second statement of the theorem.

For the first statement, note that by Theorem 3.3

$$\sum_{i \rightarrow j \in G_n} \mathcal{F}_n^*(D_i) \mathcal{F}_n^*(D_j) \geq \sum_{i \rightarrow j \in \hat{G}_n} \mathcal{F}_n^*(D_i) \mathcal{F}_n^*(D_j)$$

so that

$$\tilde{\rho}(G_n) \geq \tilde{\rho}(\hat{G}_n).$$

Therefore we have, as $n \rightarrow \infty$,

$$\begin{aligned} \mathbb{P}(\tilde{\rho}(G_n) < \rho(D_*, D^*) - Kn^{-\delta}) & \leq \mathbb{P}(\tilde{\rho}(\hat{G}_n) < \rho(D_*, D^*) - Kn^{-\delta}) \\ & \leq \mathbb{P}(|\tilde{\rho}(\hat{G}_n) - \rho(D_*, D^*)| > Kn^{-\delta}) \\ & \leq O(n^{-\varepsilon-\delta} + \mathbb{P}(\Omega_n^c)), \end{aligned}$$

which proves the first statement of the theorem.

We now move on to Theorem 2.3. We will follow the strategy described in Section 4, that is we will use the delta transformation to construct a degree distribution f_ρ for which $f_\rho^*(1)$ is large enough.

First observe that (22) together with Proposition 6.5 imply Proposition 4.1.

Proof of Theorem 2.3. Let ω be such that

$$9(\omega/2)^2 - 6(\omega/2)^3 - 3 = \rho + \epsilon, \quad (39)$$

for some $\epsilon > 0$, and denote by f^* the size-biased distribution of f . Now take f_ω^* to be the ω -transform of f^* and set

$$\mu_\rho = \left(\mu(1 - F(K_\omega)) + \sum_{t=1}^{K_\omega} \frac{f_\omega^*(t)}{t} \right)^{-1},$$

where K_ω was defined as the smallest integer such that $F^*(K_\omega) > \omega$. Now we define the function f_ρ by:

$$f_\rho(0) = \frac{\mu_\rho f_\omega^*(1)}{2} = f_\rho(1) = \frac{\mu_\rho f_\omega^*(1)}{2} \quad \text{and} \quad f_\rho(t) = \frac{\mu_\rho f_\omega^*(t)}{t} \quad \text{for } t \geq 2.$$

Then, since by construction $f_\rho^*(t) = f^*(t)$ for all $t > K_\omega$, it follows that

$$\begin{aligned} \sum_{t=0}^{\infty} f_\rho(t) &= \sum_{t=1}^{\infty} \frac{\mu_\rho f_\rho^*(t)}{t} \\ &= \mu_\rho \left(\sum_{t=1}^{K_\omega} \frac{f_\rho^*(t)}{t} + \sum_{t=K_\omega+1}^{\infty} \frac{f^*(t)}{t} \right) \\ &= \mu_\rho \left(\sum_{t=1}^{K_\omega} \frac{f_\rho^*(t)}{t} + \mu(1 - F(K_\omega)) \right) = 1, \end{aligned}$$

so that f_ρ is a probability density function. Moreover, since for all $k > K_\omega$

$$1 - F_\rho(k) = \sum_{t=k+1}^{\infty} f_\rho(t) = \mu_\rho \sum_{t=k+1}^{\infty} \frac{f_\omega^*(t)}{t} = \mu_\rho \sum_{t=k+1}^{\infty} \frac{f^*(t)}{t} = \frac{\mu_\rho}{\mu} \sum_{t=k+1}^{\infty} f(t),$$

it follows that $\sum_{k=0}^{\infty} t^{1+\eta} f_\rho(t) < \infty$ and

$$\lim_{k \rightarrow \infty} \frac{1 - F_\rho(k)}{1 - F(k)} = \frac{\mu_\rho}{\mu}.$$

Now let D have probability density f_ρ , and hence size-biased density $f_\rho^*(t) = t f_\rho(t) / \mu_\rho$, and let $\mathbf{D}_n$ be generated by the IID algorithm, by sampling from D . Then, by Lemma 2.1, $\mathbf{D}_n \in \mathcal{D}_{\eta, \varepsilon}(f_\rho, f_\rho^*)$ and since by construction of f_ρ we have that $f_\rho^*(1) = \omega/2$, it follows from (39) that

$$9f_\rho^*(1)^2 - 6f_\rho^*(1)^3 - 3 = \rho + \varepsilon.$$

Hence, if G_n is a graph with degree sequence $\mathbf{D}_n$, we have, by taking $\delta = \min(\varepsilon, 1/2)/2$ in Proposition 4.1, that as $n \rightarrow \infty$,

$$\begin{aligned} \mathbb{P}(\rho(G_n) > \rho) &\geq \mathbb{P}(\rho(G_n) > \rho + \varepsilon - n^{-\delta}) \\ &= \mathbb{P}(\rho(G_n) > 9f_\rho^*(1)^2 - 6f_\rho^*(1)^3 - 3 - n^{-\delta}) \\ &\geq 1 - O(n^{-\varepsilon+\delta} + \mathbb{P}(\Omega_n^c)). \quad \square \end{aligned}$$

Acknowledgments

This work was conducted during a two month visit of Pim van der Hoorn to Moscow Institute of Physics and Technology and Yandex.

References

- Alderson DL, Li L (2007) Diversity of graphs with highly variable connectivity. *Phys. Rev. E* 75(4):046102.
- Bassler KE, Del Genio CI, Erdős PL, Miklós I, Toroczkai Z (2015) Exact sampling of graphs with prescribed degree correlations. *New J. Phys.* 17(8):083052.
- Bollobás B (1980) A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *Eur. J. Combinatorics* 1(4):311–316.
- Chen N, Olvera-Cravioto M (2015) Efficient simulation for branching linear recursions. *Proc. Winter Simulation Conf., WSC'15* (IEEE, Piscataway, NJ), 2716–2727.
- Deprez P, Wüthrich MV (2015) Construction of directed assortative configuration graphs. Preprint arXiv:1510.00575.
- Hurd T (2015) The construction and properties of assortative configuration graphs. Preprint arXiv:1512.03084.
- Kannan R, Tetali P, Vempala S (1999) Simple markov-chain algorithms for generating bipartite graphs and tournaments. *Random Structures and Algorithms* 14(4):293–308.
- Maslov S, Sneppen K (2002) Specificity and stability in topology of protein networks. *Science* 296(5569):910–913.
- Menche J, Valleriani A, Lipowsky R (2010) Asymptotic properties of degree-correlated scale-free networks. *Phys. Rev. E* 81(4):046103.
- Mesfioui M, Tajar A (2005) On the properties of some nonparametric concordance measures in the discrete case. *Nonparametric Statist.* 17(5):541–554.
- Molloy M, Reed B (1995) A critical point for random graphs with a given degree sequence. *Random Structures and Algorithms* 6(2–3):161–180.
- Newman ME (2002) Assortative mixing in networks. *Phys. Rev. Lett.* 89(20):208701.
- Newman ME, Strogatz SH, Watts DJ (2001) Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E* 64(2):026118.
- Schweizer B, Wolff EF (1981) On nonparametric measures of dependence for random variables. *Ann. Statist.* 9(4):879–885.

- Spearman C (1904) The proof and measurement of association between two things. *Amer. J. Psych.* 15(1):72–101.
- Stanton I, Pinar A (2012) Constructing and sampling graphs with a prescribed joint degree distribution. *J. Experiment. Algorithmics* 17:3–5.
- Van Der Hofstad R (2016) *Random Graphs and Complex Networks*, Vol. 1 (Cambridge University Press, New York).
- Van Der Hofstad R, Litvak N (2014) Degree-degree dependencies in random graphs with heavy-tailed degrees. *Internet Math.* 10(3–4): 287–334.
- VanDer Hoorn P (2016) Asymptotic analysis of network structures: Degree-degree correlations and directed paths. Ph.D. thesis, University of Twente.
- Van Der Hoorn P, Litvak N (2014) Convergence of rank based degree-degree correlations in random directed networks. *Moscow J. Combinatorics and Number Theory* 4(4):45–83.
- Van Der Hoorn P, Litvak N (2015) Degree-degree dependencies in directed networks with heavy-tailed degrees. *Internet Math.* 11(2):155–179.
- Villani C (2008) *Optimal Transport: Old and New*, Vol. 338 (Springer, Berlin).
- Xulvi-Brunet R, Sokolov I (2004) Reshuffling scale-free networks: From random to assortative. *Phys. Rev. E* 70(6):066102.
- Yang D, Pan L, Zhou T (2017) Lower bound of assortativity coefficient in scale-free networks. *Chaos: An Interdisciplinary J. Nonlinear Sci.* 27(3):033113.