

Average nearest neighbor degrees in scale-free networks

Dong Yao¹, Pim van der Hoorn² and Nelly Litvak^{3,4}

¹Duke University, North Carolina, United States

²Northeastern University, Boston, United States

³University of Twente, Enschede, the Netherlands

⁴Eindhoven University of Technology, Eindhoven, the Netherlands

January 1, 2018

Abstract

The average nearest neighbor degree (ANND) of a node of degree k is widely used to measure dependencies between degrees of neighbor nodes in a network. We formally analyze ANND in undirected random graphs when the graph size tends to infinity. The limiting behavior of ANND depends on the variance of the degree distribution. When the variance is finite, the ANND has a deterministic limit. When the variance is infinite, the ANND scales with the size of the graph, and we prove a corresponding central limit theorem in the configuration model (CM, a network with random connections). As ANND proved uninformative in the infinite variance scenario, we propose an alternative measure, the average nearest neighbor rank (ANNR). We prove that ANNR converges to a deterministic function whenever the degree distribution has finite mean. We then consider the erased configuration model (ECM), where self-loops and multiple edges are removed, and investigate the well-known ‘structural negative correlations’, or ‘finite-size effects’, that arise in simple graphs, such as ECM, because large nodes can only have a limited number of large neighbors. Interestingly, we prove that for any fixed k , ANNR in ECM converges to the same limit as in CM. However, numerical experiments show that finite-size effects occur when k scales with n .

1 Introduction

The goal of this paper is to analytically derive the limiting properties of the average nearest neighbor degree (ANND) in a general class of random graphs. The motivation for this analysis is that the ANND is one of the commonly accepted measures for dependencies between degrees of neighbor nodes. Such dependencies are called degree-degree correlations or network assortativity. A network is said to be assortative, if the correlation between degrees of neighbor nodes is positive and disassortative when it is negative. In assortative networks, nodes of high degree have a preference to connect to nodes of high degree. When the network is disassortative, nodes of high degree have a connection preference for nodes of low degree [23]. If there is no connection preference, the network is said to have neutral mixing.

Currently, degree-degree correlations are part of the standard set of properties used to characterize the structure of networks. See [24] for a survey of the work on network assortativity. The effect of degree-degree correlations on disease spreading in networks has been extensively addressed in the literature, cf. [2, 4, 5, 6]. For instance, it was shown that disassortative networks are easier to immunize and a disease takes longer to spread in assortative networks [11]. In the field of neuroscience, it was shown that assortative brain networks are better suited for signal processing [26], while assortative neural networks are more robust to random noise [12]. Under attacks, when edges or vertices are removed, assortative networks appear to be more resilient than disassortive networks [23, 27]. On the other hand, when different networks interact, assortativity actually decreases the robustness of the whole system [30].

A well established measure for degree-degree correlation computes Pearson’s correlation coefficient on the joint data of the degrees at both sides of an edge [22, 23]. However, this measure has been shown to depend on the size of the network and it converges to zero when the degree distribution has infinite variance [14, 20]. To remedy this, new measures have been introduced [14, 17] that are based on rank-correlations such as Spearman’s rho and Kendall’s tau. These measures are shown to converge to a proper limit when the size of the network tends to infinity [14, 16], see also [15] for an overview.

Both Pearson’s correlation coefficient and rank-correlation measures represent the full degree-degree correlation structure of a network as *one* number and hence, are unable to capture local correlation structures. Instead, another common approach in the literature is to compute for a node of degree k , the average degree of its neighbors and then take the average of these values over all nodes of degree k . This gives us a function of k called the Average Nearest Neighbor Degree (ANND) [8, 25]. For assortative networks the ANND is an increasing function of k , while it is decreasing in k for disassortative networks and constant when the network has neutral mixing.

Networks with neutral mixing are an essential tool in the analysis of degree-degree correlations. An important example of such networks are given by the configuration model [7, 9, 21, 29], which generates networks with a given degree sequence and lets the nodes connect randomly to each other. It is well known that the ANND in the configuration model is a constant, given by the empirical second moment of the degrees, divided by the empirical first moment. However, until now, no rigorous proof of this result was known, and the asymptotic behavior of ANND in the infinite graph size limit has not been analyzed. Given the wide application of ANND, this is an essential gap especially in the most realistic scenario when degrees have infinite variance. In this case, the empirical second moment grows with the size of the network and hence the ANND diverges as the network size tends to infinity. This means that, similarly to the Pearson’s correlation coefficient, the ANND is not a consistent measure for degree-degree correlations in networks.

In this paper we address the convergence of the ANND in random graphs with given joint degree distribution of neighbor nodes, and in the configuration model. In Section 4 we prove that, when the degree distribution has a finite variance, the ANND converges point-wise in probability in the general case. Moreover, for the configuration model the rate of convergence is uniform. In Section 5 we turn to the important case of infinite variance and focus on the configuration model. We prove a central limit theorem for the ANND where the limiting random variable has a stable distribution with infinite variance. In Section 6 we apply the strategy of using ranks instead of degrees and propose the Average Nearest Neighbor Rank (ANNR) correlation measure. We prove that ANNR converges point-wise for any joint degree distribution. In the configuration model the limit is a constant, which is determined by the size-biased degree distribution. Moreover, the limit is preserved, for any fixed k , in the erased configuration model, which is a simple graph obtained by removing self-loops and multiple edges.

Numerical experiments in Section 7 illustrate our results for ANND and ANNR. One striking observation is that the empirical average of ANND and ANNR over several networks rapidly decrease with k , when k is large. We explain this by the fact that for sufficiently large k , with positive probability, there is no node of such degree. When correcting by this probability, ANND and ANNR in the configuration model no longer decrease. Moreover, ANNR shows remarkably small fluctuations around its theoretical value even in small networks.

In the erased configuration model, for very large degrees k , which scale with the size of the graph, we observe a decline of ANNR. This phenomenon, called ‘structural correlations’, is well known in the literature [1, 8]. It is explained by the fact that the graph is simple, therefore, nodes of large degrees are forced to connect to nodes of smaller degrees. Structural correlations for the rank-based correlation measure Spearman’s rho have been observed before [18]. Here we see that this phenomenon holds for the ANNR as well.

We start with introducing definitions and notations in Section 2. Then in Section 3, before analyzing ANND, we establish an upper bound K_n depending on n , such that all values $1, 2, \dots, K_n$ are present, with high probability, in an i.i.d. degree sequence of sequence length n , sampled from a regularly-varying distribution. To the best of our knowledge, this result is new and can be of independent interest.

2 Notations and definitions

2.1 Graphs and degree distributions

Let $G_n = ([n], E_n)$ denote an undirected graph of size n , with the nodes labeled from 1 to n , and E_n the set of edges. We write D_i for the degree of node i and call $\mathbf{D}_n = (D_1, \dots, D_n)$ the degree sequence of G_n . We use the term ‘degree sequence’ to refer to any sequence (D_1, \dots, D_n) for which the sum $L_n = \sum_{i=1}^n D_i$ is even.

For computations, it is convenient to replace each edge between nodes i and j by two directed edges $i \rightarrow j$ and $j \rightarrow i$. We denote by G_{ij} the number of edges $i \rightarrow j$, and note that G_{ij} can be larger than 1 because we allow for multi-graphs. By construction we have $G_{ij} = G_{ji}$. Furthermore, G_{ii} is *twice* the number of self-loops of node i . With these notation the degree of a node is given as $D_i = \sum_{j=1}^n G_{ij}$, which is the number of half edges attached to the node. We will use the notation $\sum_{i \rightarrow j}$ to denote the summation over all edges $i \rightarrow j$ in the graph (note that each pair of connected nodes in the undirected graph is counted twice in such summation).

In this paper, we use i.i.d. degree sequences, formally defined as follows [13]. Let \mathcal{D} be a positive integer-valued random variable, with cumulative distribution function F and probability density function f . Let $D_1, D_2, \dots, D_{n-1}, d_n$, be i.i.d. samples from F , and define:

$$D_n = d_n + \mathbb{1}_{\{(\sum_{i=1}^{n-1} D_i + d_n) \text{ is odd}\}}.$$

We will write $\mathbf{D}_n = \text{IID}(\mathcal{D})$ for such sequence. Observe that the correction term in D_n is uniformly bounded in n , therefore, it does not contribute to the asymptotic behavior of random graphs with degree sequence $\text{IID}(\mathcal{D})$. Hence, without loss of generality, we will consider the degrees D_1, D_2, \dots, D_n as i.i.d. samples from \mathcal{D} .

Throughout the paper we will denote the empirical degree density function by

$$f_n(k) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{D_i=k\}}, \quad k = 0, 1, \dots,$$

and the size-biased empirical degree density function will be denoted by

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

In addition, we will denote by F_n and F_n^* , the cumulative distribution functions corresponding to f_n and f_n^* , respectively.

2.2 Distances between probability measures

In order to describe convergence of empirical distributions to their limits, we will use two different distances between probability measures: the total variation distance and the Kantorovich-Rubinstein distance.

Let f and g be two probability density functions on the non-negative integers and let F and G denote their respective cumulative distribution functions.

The total variation distance d_{tv} is defined as

$$d_{tv}(f, g) = \frac{1}{2} \sum_{k=0}^{\infty} |f(k) - g(k)|.$$

Using the definition

$$F(k) = \sum_{l \leq k} f(l),$$

we immediately see that

$$\sup_{k \geq 0} |F(k) - G(k)| \leq 2d_{tv}(f, g). \quad (1)$$

The Kantorovich-Rubinstein distance d_1 between f and g , is defined as follows:

$$d_1(f, g) = \sum_{k \geq 0} |F(k) - G(k)|. \quad (2)$$

Let X and Y be non-negative integer-valued random variables with distributions F and G respectively, and assume that F and G have a finite mean. Then it follows from

$$\mathbb{E}(X) = \sum_{k \geq 0} \mathbb{P}(X > k) \quad (3)$$

and the triangle inequality that

$$|\mathbb{E}(X) - \mathbb{E}(Y)| \leq d_1(F, G).$$

In particular, convergence in d_1 implies convergence of the means.

We will usually use the Kantorovich-Rubinstein distance when the means are finite, and the total variation distance when the means are infinite.

2.3 Regularly-varying distributions

An important feature shared by many real-world networks is that their degree distribution is scale-free. This is often visualized by showing that $\mathbb{P}(D > k)$ behaves as an inverse power of k . Mathematically we can model this using regularly-varying distributions. In this paper, we assume that \mathcal{D} has a regularly-varying density, i.e.

$$\mathbb{P}(\mathcal{D} = k) = l(k)k^{-\gamma-1}, \quad k = 1, 2, \dots \quad (4)$$

The parameter $\gamma > 1$ is called the *exponent* of the distribution and $l(x)$ is a slowly varying function, which means that for every $\lambda > 0$,

$$\lim_{x \rightarrow \infty} \frac{l(\lambda x)}{l(x)} = 1.$$

We will furthermore assume that the slowly-varying function $l(x)$ is eventually monotone.

Observe that if \mathcal{D} satisfies (5) then $\mathbb{E}[\mathcal{D}^p] < \infty$ for all $p < \gamma$. We refer to [3] for a thorough treatment of regular variation. We do want to point out that due to Karamata's theorem, it follows from (4) that

$$\mathbb{P}(\mathcal{D} > t) = \tilde{l}(t)t^{-\gamma} \quad \text{for all } t > 0, \quad (5)$$

where $\tilde{l}(x) \sim l(x)/\gamma$ as $x \rightarrow \infty$. See Lemma 9.1 from more details. Due to this asymptotic equivalence we shall slightly abuse notation and use $l(x)$ to denote both the slowly-varying function associated with $\mathbb{P}(\mathcal{D} = k)$ and $\mathbb{P}(\mathcal{D} > t)$. When \mathcal{D} satisfies (4) we say that \mathcal{D} is regularly varying with exponent γ .

2.4 Stable distributions

Stable distributions are important for us, since they come up as the limit distribution for central limit theorems involving regularly-varying distributions. A random variable X is said to have a stable distribution if for every $n \geq 2$ there exists constants a_n and b_n such that for any sequence X_1, \dots, X_n of independent copies of X ,

$$X_1 + X_2 + \dots + X_n \stackrel{d}{=} a_n X + b_n.$$

The characteristic function of stable distribution can be classified using four parameters α, β, σ and μ , see [26, Definition 1.1.6], and the corresponding random variable is denoted by $S_\alpha(\sigma, \beta, \mu)$. The parameter α is called the stability index and is the most important parameter for our purposes, since it relates to the exponent γ of the regularly-varying distribution. The Stable Law CLT

([28, Theorem 4.5.1]), states that for a sequence $(X_i)_{i \geq 1}$ of i.i.d. copies of a regularly-varying random variable with exponent $\gamma > 1$ there exist a slowly-varying function $l_0(n)$ such that

$$\frac{n^{-\frac{1}{\gamma}}}{l_0(n)} \sum_{i=1}^n X_i \xrightarrow{d} S_\gamma(1, \beta, 0) \quad (6)$$

When the X_i are non-negative, as is the case for degrees, $\beta = 1$. Hence we will omit the dependence on the other three parameters and write S_γ for a stable distribution with stability index γ .

2.5 Average nearest neighbor degree

The Average Nearest Neighbor Degree (ANND) of nodes of degree k in a graph G_n , of size n , is formally defined on simple graphs as (see also [8])

$$\Phi_n(k) = \mathbb{1}_{\{f_n(k) > 0\}} \sum_{\ell > 0} \ell P(\ell|k), \quad (7)$$

where $P(\ell|k)$ is defined as the probability that a node of degree k is connected to a node of degree ℓ . The indicator $\mathbb{1}_{\{f_n(k) > 0\}}$ stands for the event that at least one node of degree k exists in the network, otherwise, $\Phi_n(k)$ is set to zero. Let us define

$$h_n(k, \ell) = \frac{1}{L_n} \sum_{i \rightarrow j} \mathbb{1}_{\{D_i = k, D_j = \ell\}}, \quad (8)$$

as the empirical joint distribution of the degrees on both sides of a randomly sampled edge in the graph G_n . We shall refer to $h_n(k, \ell)$ as the *joint degree distribution*. Next, we note that $P(\ell|k)$ is equivalent to the probability that a randomly selected edge $i \rightarrow j$, conditioned on $D_i = k$, satisfies $D_j = \ell$, i.e. $P(\ell|k) = h_n(k, \ell) / f_n^*(k)$. Using this, we can extend (7) to the setting of arbitrary (multi)graphs as

$$\Phi_n(k) = \mathbb{1}_{\{f_n(k) > 0\}} \frac{\sum_{\ell > 0} h_n(k, \ell) \ell}{f_n^*(k)}. \quad (9)$$

2.6 Configuration model

The configuration model (CM) [7, 9, 21, 29] is an important model for generating graphs G_n of size n , with a specific degree sequence. Since it generates graphs with neutral mixing (see e.g. [14, 15]), it is also a crucial model for the analysis of degree-degree correlations.

Given degree sequence \mathbf{D}_n , we assign D_i stubs (half-edges) to each node i . Then we randomly pair the L_n stubs to obtain a graph (possibly a multi-graph) G_n with the given degree sequence. This procedure can be extended to generate graphs with a specific degree distribution. Let \mathcal{D} be a non-negative, integer-valued random variable with probability density f . When $\mathbf{D}_n = \text{IID}(\mathcal{D})$, the configuration model generates random graphs, in which the empirical degree distribution f_n converges to f as $n \rightarrow \infty$.

CM can be adjusted to generate simple graphs. One approach is to simply repeat the wiring of the graph until the resulting graph is simple. This is the repeated configuration model (RCM). The RCM can be applied successfully only if the probability to obtain a simple graph converges to a nonzero value as $n \rightarrow \infty$. It is well-known (see [13, Chapter 7]) that this is indeed the case if and only if \mathcal{D} has finite second moment. Another way to obtain a simple graph is to simply remove all self-loops and replace all multiple edges between i and j by a single edge. This model, called the erased configuration model (ECM), generates a simple graph with correct asymptotic degree distribution whenever $\mathbf{D}_n = \text{IID}(\mathcal{D})$ and \mathcal{D} has a finite mean (see [13, Theorem 7.10]).

3 Sampled degrees

Recall that in (9) we set $\Phi_n(k) = 0$ if the degree sequence contains no nodes of degree k . Therefore, before we start our analysis of the behavior of $\Phi_n(k)$, we need to understand which k 's are present in the degree sequence $\mathbf{D}_n = \text{IID}(\mathcal{D})$. The following result for regularly-varying distributions is, to the best of our knowledge, not known in the literature and can be of independent interest.

Theorem 3.1. *Let $\mathbf{X}_n = \{X_1, \dots, X_n\}$, be independent copies of an integer-valued random variable X , with regularly-varying probability mass function f with exponent $\gamma > 1$ and suppose that $f(k) > 0$ for all $k > 0$. Then for any $0 < a < \frac{1}{\gamma+1}$*

$$\lim_{n \rightarrow \infty} \mathbb{P}(\{1, 2, \dots, \lceil n^a \rceil\} \subseteq \mathbf{X}_n) = 1.$$

On the other hand, if $a > \frac{1}{\gamma+1}$, then

$$\lim_{n \rightarrow \infty} \mathbb{P}(\lceil n^a \rceil \in \mathbf{X}_n) = 0.$$

Proof. Throughout the proof $l(k)$ denotes the slowly-varying function associated with the probability mass function of X . We start with the first statement.

Since $a < 1/(\gamma + 1)$, it follows from Potter's bounds for slowly varying functions, that there exist $\gamma + 1 < b < \frac{1}{a}$, $C > 0$ and $K \geq 1$, such that $f(k) \geq Ck^{-b}$ for all $k \geq K$. Then, for sufficiently large n ,

$$\begin{aligned} & \mathbb{P}(\{1, 2, \dots, \lceil n^a \rceil\} \subseteq \{X_1, X_2, \dots, X_n\}) \\ & \geq 1 - \sum_{k=1}^{\lceil n^a \rceil} \mathbb{P}(k \notin \{X_1, X_2, \dots, X_n\}) \\ & \geq 1 - \sum_{k=1}^{K-1} (1 - \mathbb{P}(X = k))^n - \sum_{k=K}^{\lceil n^a \rceil} (1 - Ck^{-b})^n \\ & \geq 1 - (K-1)(1 - \mathbb{P}(X = K-1))^n - (n^a + 1)(1 - Cn^{-ab})^n. \end{aligned}$$

Because $\mathbb{P}(X = K-1) < 1$ we have $(K-1)(1 - \mathbb{P}(X = K-1))^n \rightarrow 0$ as $n \rightarrow \infty$, while $ab < 1$ implies that $(n^a + 1)(1 - Cn^{-ab})^n \rightarrow 0$ and hence $\mathbb{P}(\{1, 2, \dots, \lceil n^a \rceil\} \subseteq \{X_1, X_2, \dots, X_n\}) \rightarrow 1$.

We follow a similar approach for the second statement. Since $\frac{1}{a} > \gamma + 1$, it follows from Potter's bounds that there exist $\frac{1}{a} < b' < \gamma + 1$, $C' > 0$ and $K \geq 1$, such that $f(k) \leq C'k^{-b'}$ for all $k \geq K$. Then, for sufficiently large n we obtain

$$\begin{aligned} \mathbb{P}(\lceil n^a \rceil \in \{X_1, \dots, X_n\}) &= 1 - (1 - f(\lceil n^a \rceil))^n \\ &\leq 1 - (1 - Cn^{-ab'})^n. \end{aligned}$$

Since $ab' > 1$, we have $\lim_{n \rightarrow \infty} (1 - Cn^{-ab'})^n = 1$, which gives the result. \square

Remark 3.2. *Theorem 3.1 can be adjusted in a straight-forward way to the case where the probability density function f of X has infinite support but is zero on some finite subset of the positive integers.*

It is interesting to notice that the degrees higher than $n^{\frac{1}{\gamma+1}}$ will appear in \mathbf{D}_n , in particular the maximal degree scales as $n^{\frac{1}{\gamma}}$. However, only up to n^a with $a < \frac{1}{\gamma+1}$ one can guarantee that *all* degrees between 1 and n^a will participate in \mathbf{D}_n . In other words, for $\frac{1}{\gamma+1} < a < b < \frac{1}{\gamma}$ some values $k \in [n^a, n^b]$ will be missing in \mathbf{D}_n .

4 Limiting behavior of Φ_n in graphs with general joint degree distributions and finite second degree

In this section we analyze the graphs with given limiting joint degree distribution $h(k, \ell)$ and finite second moment. Our main result (Theorem 4.2) proves the convergence of the average nearest neighbor degree Φ_n to its limit

$$\Phi(k) = \mathbb{1}_{\{f^*(k) > 0\}} \frac{\sum_{\ell=1}^{\infty} h(k, \ell) \ell}{f^*(k)}, \quad (10)$$

where

$$f^*(k) = \frac{kf(k)}{\mathbb{E}[\mathcal{D}]}$$

is the size-biased probability density function f^* of a non-negative integer random variable \mathcal{D} . Note that for $k \geq 1$, $f(k) > 0$ if and only if $f^*(k) > 0$.

As commonly accepted in the random graph literature, we impose regularity assumptions on the degree sequences. Assumption 4.1 uses the distance d_1 to simultaneously state the convergence of f_n , f_n^* and their expectations to the corresponding limiting values. Note that convergence of the expectation of f_n^* is equivalent to the convergence of the second moment of f_n . We denote by A^c the event complementary to the event A .

Assumption 4.1 (Regularity of the degrees). *There exist a probability density f on the non-negative integers with size biased version f^* and some $\alpha, \varepsilon > 0$, such that if we set*

$$\Omega_n = \{\max\{d_1(f_n, f), d_1(f_n^*, f^*)\} \leq n^{-\varepsilon}\},$$

then, as $n \rightarrow \infty$,

$$\mathbb{P}(\Omega_n^c) = O(n^{-\alpha}).$$

Assumption 4.1 looks slightly stronger than just convergence because it requires that the distances between the empirical and limiting distributions are not larger than a negative power of n with high probability. However, this is not restrictive. In particular, Theorem 4.1 below states that Assumption 4.1 holds for $\mathbf{D}_n = \text{IID}(\mathcal{D})$ whenever \mathcal{D} has a finite $(2 + \eta)$ -moment for some $\eta > 0$. The proof, which can be found in Section 9.2, is a straightforward extension of the proof of [15, Theorem 3.1].

Theorem 4.1. *Suppose that $\mathbb{E}[\mathcal{D}^{2+\eta}] < \infty$ for some $0 < \eta < 1$. Let $\mathbf{D}_n = \text{IID}(\mathcal{D})$, $0 < \varepsilon \leq \frac{\eta}{2(\eta+2)}$ and Ω_n be defined as in Assumption 4.1. Then, as $n \rightarrow \infty$,*

$$\mathbb{P}(\Omega_n^c) = O(n^{-\varepsilon})$$

so that, in particular, \mathbf{D}_n satisfies Assumption 4.1 with $\alpha = \varepsilon$.

We remark that the above result still holds if, instead of Ω_n , we consider the event

$$\{\max\{d_1(f_n, f), d_1(f_n^*, f^*)\} \leq Kn^{-\varepsilon}\},$$

for any $K > 0$.

The second regularity assumption imposes, in a similar fashion, the convergence of the joint degree distribution $h_n(k, \ell)$, of the degrees at both ends of a randomly selected edge.

Assumption 4.2 (Regularity of the joint distribution). *There exists a joint probability density function $h(k, \ell)$ on the positive integers and some $\kappa > 0$, such that if*

$$\Gamma_n = \left\{ \sum_{k, \ell=1}^{\infty} |h_n(k, \ell) - h(k, \ell)| \mathbb{1}_{\{f_n(k) > 0\}} \leq n^{-\kappa} \right\},$$

then we have $\lim_{n \rightarrow \infty} \mathbb{P}(\Gamma_n) = 1$.

Assumption 4.2 is satisfied for $\mathbf{D}_n = \text{IID}(\mathcal{D})$ in several models, for example, the Configuration Model (see [15, Proposition 6.2]) and the Maximally Disassortative Graph Algorithm (see [19, Theorem 3.3]).

We are now ready to state the main result of this section.

Theorem 4.2. *Let $\{G_n\}_{n \geq 1}$ be a sequence of graphs, which satisfies Assumptions 4.1 and 4.2 and assume that the limiting distribution f has a finite $(2 + \eta)$ -moment for some $0 < \eta < 1$. Then, for each fixed k such that $f(k) > 0$ and each $0 < \delta < \min \left\{ \varepsilon, \frac{\kappa \eta}{\eta + 1} \right\}$,*

$$\lim_{n \rightarrow \infty} \mathbb{P} (|\Phi_n(k) - \Phi(k)| \leq n^{-\delta}) = 1.$$

The proof is based on splitting of $|\Phi_n(k) - \Phi(k)|$ in several terms and bounding each of them separately, using Assumptions 4.1 and 4.2. We give the proof in Section 9.3.

5 Limiting behavior of Φ_n in Configuration Models

In [14, 16] it was shown that different measures for degree-degree correlations converge to zero, as $n \rightarrow \infty$, for both the multi-graph configuration model, as well as the repeated and erased versions. Therefore, one would expect that $\Phi_n(k)$ converges to some constant, independent of k . We will show that this is indeed the case for all three models when we consider $\mathbf{D}_n = \text{IID}(\mathcal{D})$, where \mathcal{D} has finite $(2 + \eta)$ moment. When the second moment is infinite and \mathcal{D} is regularly varying, we establish a central limit theorem, where the limiting random variable has a stable distribution.

5.1 Multi-graphs, finite variance of the degrees

We first consider the case when \mathcal{D} has finite $(2 + \eta)$ -moment. In [15] it is proven that the directed version of the configuration model satisfies Assumption 4.2, where $h(k, \ell) = f^*(k)f^*(\ell)$. The proof can be adjusted in a straight-forward manner to the undirected case. If we plug this result for $h(k, \ell)$ into (10) and define $\nu_p = \mathbb{E}[\mathcal{D}^p]$ for $p = 1, 2$, we get

$$\Phi(k) = \frac{\sum_{\ell > 0} f^*(k)f^*(\ell)\ell}{f^*(k)} = \sum_{\ell > 0} \frac{f(\ell)\ell^2}{\nu_1} = \frac{\nu_2}{\nu_1}.$$

Hence, it follows from Theorem 4.2 that

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\left| \Phi_n(k) - \frac{\nu_2}{\nu_1} \right| > n^{-\delta} \right) = 0. \quad (11)$$

With a little extra work, we can prove the following stronger result, which states that the convergence in probability is uniform in k , and gives an upper bound on the speed of convergence. The proof exploits the construction of CM and is provided in Section 9.4.

Theorem 5.1. *Let \mathcal{D} be an integer-valued random variable which satisfies $\mathbb{E}[\mathcal{D}^{2+\eta}] < \infty$, for some $0 < \eta < 1$, and let $\{G_n\}_{n \geq 1}$ be a sequence of graphs generated by CM with $\mathbf{D}_n = \text{IID}(\mathcal{D})$. Then, for any $0 < \delta < \frac{\eta}{2(\eta+2)}$,*

$$\sup_{k \geq 0} \mathbb{P} \left(\left| \Phi_n(k) - \frac{\nu_2}{\nu_1} \right| \mathbf{1}_{\{f_n(k) > 0\}} > n^{-\delta} \right) = O \left(n^{-\frac{\delta}{2}} + n^{\delta - \frac{\eta}{2(\eta+2)}} \right).$$

5.2 Multi-graph, infinite variance of the degrees

It is important to note that the finite second moment condition in Theorem 5.1 can not be relaxed to the case where \mathcal{D} has only finite mean, since then $\sum_{\ell > 0} \ell^2 f(\ell)$ is no longer finite and hence $\sum_{\ell > 0} \ell^2 f_n(\ell)$ diverges as $n \rightarrow \infty$. Consequently $\Phi_n(k)$ will increase as $n \rightarrow \infty$.

In order to understand how $\Phi_n(k)$ scales with n observe that for CM we have

$$\Phi_n(k) = \frac{\sum_{\ell>0} h_n(k, \ell) \ell}{f_n^*(k)} \approx \frac{1}{L_n} \sum_{\ell>0} \ell^2 f_n(\ell) = \sum_{i=1}^n \frac{D_i^2}{L_n}.$$

When D_i are sampled from a regularly-varying random variable with exponent $1 < \gamma < 2$, it follows that $\sum_{i=1}^n D_i^2$ scales as $n^{2/\gamma}$ and $L_n = \sum_{i=1}^n D_i \approx n\mathbb{E}[\mathcal{D}]$. It now follows that $\Phi_n(k)$ scales as $n^{2/\gamma-1}$. These scaling terms can be made exact by adding slowly-varying functions, and we show in Theorem 5.3 that $\Phi_n(k)$ rescaled with $n^{2/\gamma-1}$ converges to a random variable with a stable distribution.

Before formulating the result, we need a weaker version of Assumption 4.1, for this does not hold anymore when the second moment of the degrees is infinite.

Assumption 5.1. *There exist a probability density f on non-negative integers with size biased version f^* and some $\alpha, \varepsilon > 0$, such that if we set*

$$\Omega_n := \{\max\{d_1(f_n, f), d_{tv}(f_n^*, f^*)\} \leq n^{-\varepsilon}\},$$

then, as $n \rightarrow \infty$,

$$\mathbb{P}(\Omega_n^c) = O(n^{-\alpha}).$$

Observe that this assumption resembles Assumption 4.1 with the only exception that we replaced the Kantorovich-Rubenstein distance d_1 for the size-biased degree distribution f^* by the total variation distance d_{tv} . This is because the convergence of the Kantorovich-Rubenstein distance implies the convergence of the first moment, which does not hold for the size-biased distribution when \mathcal{D} has infinite variance. The next theorem from [15] shows that Assumption 5.1 holds in CM with $\mathbf{D}_n = \text{IID}(\mathcal{D})$.

Theorem 5.2 ([15] Theorem 3.1). *Suppose that $\mathbb{E}(\mathcal{D}^{1+\eta}) < \infty$, for some $0 < \eta < 1$. Let $\mathbf{D}_n = \text{IID}(\mathcal{D})$, $0 < \varepsilon \leq \frac{\eta}{4(\eta+2)}$ and Ω_n as defined in Assumption 5.1. Then, as $n \rightarrow \infty$,*

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

so that, in particular, \mathbf{D}_n satisfies Assumption 5.1 with $\alpha = \varepsilon$.

Similar to Theorem 4.1, this result still holds if we consider the upper bound $Kn^{-\varepsilon}$ in Ω_n , for some $K > 0$.

With this result we can now state the following central limit theorem for Φ_n in the configuration model with regularly-varying degrees.

Theorem 5.3. *Let \mathcal{D} be an integer-valued regular varying-random variable with exponent $1 < \gamma < 2$ and let $\{G_n\}_{n \geq 1}$ be a sequence of graphs generated by CM with $\mathbf{D}_n = \text{IID}(\mathcal{D})$. Assume $f(k) > 0$ for all $k > 0$. Then there exists a slowly varying function $l_0(n)$, such that $\frac{\Phi_n(k)}{l_0(n)n^{\frac{2}{\gamma}-1}}$ converges in distribution to $S_{\gamma/2}$, having stable distribution with stability index $\frac{\gamma}{2}$. More precisely, for any $\tau < \frac{1}{\gamma+1}$ and any bounded Lipschitz function g ,*

$$\lim_{n \rightarrow \infty} \sup_{1 \leq k \leq n^\tau} \left| \mathbb{E} \left[g \left(\frac{\Phi_n(k)}{l_0(n)n^{\frac{2}{\gamma}-1}} \right) - g(S_{\gamma/2}) \right] \right| = 0.$$

Remark 5.4. *Although the ANND does not converge to a constant for CM with infinite variance degrees, it is worth noting that the distribution of the limit random variable $S_{\gamma/2}$ is independent of k . This reflects the independence between the degree of connected nodes in the CM.*

The proof of Theorem 5.3 is given in Section 9.4. This result has two important consequences. First, it shows that, up to some slowly-varying function, $\Phi_n(k)$ scales as $n^{\frac{2}{\gamma}-1}$ and hence $\Phi_n(k) \rightarrow \infty$ with n . Second, since the rescaled limit is a random variable, even for large n the value of the rescaled $\Phi_n(k)$ can be significantly different for different graphs generated by the configuration model. Therefore, we conclude that the ANND is not a proper measure for graphs where the degree distribution has infinite second moment.

5.3 Repeated configuration model

Recall that RCM repeatedly generates CM graphs until the resulting graph is simple. Let S_n denote the event that the graph G_n , generated by CM with $\mathbf{D}_n = \text{IID}(\mathcal{D})$, is simple. Then $\mathbb{P}(S_n)$ converges to a positive number whenever \mathcal{D} has finite second moment (Theorem 7.12 in [14]). Therefore, in the rest of this section we assume that $\mathbb{E}(\mathcal{D}^{2+\eta}) < \infty$ for some $\eta > 0$. Since for all k with $f(k) > 0$ it holds that

$$\mathbb{P}\left(\left|\Phi_n(k) - \frac{\nu_2}{\nu_1}\right| > n^{-\delta} \mid S_n\right) \leq \frac{\mathbb{P}\left(\left|\Phi_n(k) - \frac{\nu_2}{\nu_1}\right| > n^{-\delta}\right)}{\mathbb{P}(S_n)},$$

the result of Theorem 5.1 can be extended in a trivial way to the RCM. We state it here for completeness, where we exclude the speed of convergence since it depends on the convergence of $\mathbb{P}(S_n)$.

Theorem 5.5. *Let \mathcal{D} be an integer-valued random variable which satisfies $\mathbb{E}[\mathcal{D}^{2+\eta}] < \infty$, for some $0 < \eta < 1$, and let $\{G_n\}_{n \geq 1}$ be a sequence of graphs generated by the RCM with $\mathbf{D}_n = \text{IID}(\mathcal{D})$. Then, for any $0 < \delta < \frac{\eta}{2(\eta+2)}$,*

$$\lim_{n \rightarrow \infty} \sup_k \mathbb{P}\left(\left|\Phi_n(k) - \frac{\nu_2}{\nu_1}\right| \mathbf{1}_{\{f_n(k) > 0\}} > n^{-\delta}\right) = 0.$$

5.4 Erased configuration model

Unlike the RCM, the ECM is well-defined and has the correct limiting degree distribution when degrees have only finite expectation. Therefore, the ECM is used for generating simple graphs with infinite variance of the degrees. Moreover, due to computational advantages of the ECM over the RCM, the ECM is also preferred when the degree distribution has finite variance. In this section we consider both finite and infinite variance scenario.

Because of the removal of edges, the eventual degree sequence in the ECM is, in general, different from $\mathbf{D}_n = \text{IID}(\mathcal{D})$. We will use symbols with hats to denote characteristics of the ECM. For instance, \widehat{D}_i denotes the degree of node i in the ECM, while D_i denotes its degree before the removal of edges. Similarly, $\widehat{\Phi}_n(k)$ is the ANND in the ECM, while $\Phi_n(k)$ is computed on the multi-graph CM.

Our main result in this section establishes the scaling of the error between Φ_n and $\widehat{\Phi}_n$. The proof is provided in Section 9.5.

Theorem 5.6. *Let \mathcal{D} be regularly varying with exponent $\gamma > 1$ and let $\{G_n\}_{n \geq 1}$ be a sequence of graphs generated by the ECM with $\mathbf{D}_n = \text{IID}(\mathcal{D})$. Set $a = (\gamma - 1)^2 / (2\gamma) > 0$ and let k be such that $f^*(k) > 0$. If $\gamma \in (1, 2)$, then*

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\left|\widehat{\Phi}_n(k) - \Phi_n(k)\right| > n^{\frac{2}{\gamma}-1-a}\right) = 0,$$

while for $\gamma > 2$ it holds that

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\left|\widehat{\Phi}_n(k) - \Phi_n(k)\right| > n^{-a}\right) = 0.$$

The first result states that the difference $\left|\widehat{\Phi}_n(k) - \Phi_n(k)\right|$, although not necessarily vanishing, is of the smaller order than $\Phi_n(k)$ in Theorem 5.3. Therefore, we can extend the latter theorem to the ECM as follows.

Theorem 5.7. *Let \mathcal{D} be regularly varying with exponent $1 < \gamma < 2$ and let $\{G_n\}_{n \geq 1}$ be a sequence of graphs generated by the ECM with $\mathbf{D}_n = \text{IID}(\mathcal{D})$. Then there exists a slowly-varying function $l_0(n)$ and random variable $S_{\gamma/2}$ having a stable distribution with shape parameter $\frac{\gamma}{2}$, such that for each k with $f^*(k) > 0$,*

$$\frac{\widehat{\Phi}_n(k)}{l_0(n)n^{\frac{2}{\gamma}-1}} \xrightarrow{d} S_{\gamma/2} \quad \text{as } n \rightarrow \infty.$$

The second part of Theorem 5.6 states that the difference $|\widehat{\Phi}_n(k) - \Phi_n(k)|$ goes to zero as $n \rightarrow \infty$, thus, using (11), we obtain the following result.

Theorem 5.8. *Let \mathcal{D} be regularly varying with exponent $\gamma > 2$ and let $\{G_n\}_{n \geq 1}$ be a sequence of graphs generated by ECM with $\mathbf{D}_n = \text{IID}(\mathcal{D})$. Then for any k such that $f^*(k) > 0$ and any $0 < \delta < \frac{\gamma-1}{2(\gamma+3)}$, we have*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\left| \widehat{\Phi}_n(k) - \frac{\nu_2}{\nu_1} \right| > n^{-\delta} \right) = 0.$$

The proof of Theorem 5.8 follows directly by splitting

$$\left| \widehat{\Phi}_n(k) - \frac{\nu_2}{\nu_1} \right| \leq \left| \Phi_n(k) - \frac{\nu_2}{\nu_1} \right| + \left| \widehat{\Phi}_n(k) - \Phi_n(k) \right|$$

and then using Theorem 5.2 and Theorem 5.6. Observe that when $\gamma > 2$, we have $\eta := (\gamma-2)/2 > 0$ and $\mathbb{E} [\mathcal{D}^{2+\eta}] < \infty$, which yields the bound for δ in the above theorem.

Interestingly, in the literature, it has been observed that the ANND in the ECM decreases for large k . This phenomenon, termed as ‘structural correlations’ or ‘finite-size effects’ has been ascribed to the fact that the graph is simple, thus larger nodes do not have sufficient number of large neighbors [1, 8]. Theorem 5.6 says that asymptotically the difference between ANND in CM and ECM is vanishing for any *fixed* k . However, when k scales as a positive power of n , the difference between ANND in CM and ECM might be indeed non-vanishing because of the more significant contribution of self-loops and multiple edges. In numerical experiments we observed that there is indeed a difference between the range of ANND in CM and ECM. However, the numerical results showed enormous fluctuations even in the case of finite variance of the degrees, see e.g. Figure 3. Therefore, the numerical comparison of CM and ECM was inconclusive and is omitted in this paper. Exact evaluation of the finite-size effects remains an interesting open question.

6 Average nearest neighbor rank

Many real networks exhibit power law degree distributions with infinite variance. As we saw above, the ANND functional in CM with $\mathbf{D}_n = \text{IID}(\mathcal{D})$ degree sequences has the obvious disadvantage that it scales as the network size becomes larger, which makes it difficult to compare networks that have similar structure but different sizes. Also, in the infinite variance case, the scaled limit is a proper random variable, which can vary as the degree sample changes.

Therefore, in this section we introduce a new measure that is suitable in the infinite variance case. A classical approach is to use rank-based measures. Thus we introduce the average nearest neighbor rank (ANNR), denoted by $\Theta_n(k)$, which is defined as follows:

$$\Theta_n(k) = \mathbb{1}_{\{f_n(k) > 0\}} \frac{\sum_{\ell > 0} h_n(k, \ell) F_n^*(\ell)}{f_n^*(k)}.$$

The difference with $\Phi_n(k)$ is that we replace ℓ in the summation on the right-hand side of (9) by $F_n^*(\ell)$. Recall that F_n^* is the cumulative size-biased degree distribution, i.e.,

$$F_n^*(\ell) = \frac{1}{L_n} \sum_{i=1}^n D_i \mathbb{1}_{\{D_i \leq \ell\}}.$$

Hence, the value $F_n^*(\ell)$ can be seen as assigning a rank to edges involving a node of degree ℓ . For instance, for the largest degree D_{\max} , $F_n^*(D_{\max}) = 1$, while for the smallest degree D_{\min} , $F_n^*(D_{\min})$ is the fraction of edges in the network with one of the nodes having degree D_{\min} . Recall that the total fraction of edges with one node having degree k equals $f_n^*(k)$. Therefore, recalling that $P(\ell|k) = h_n(k, \ell)/f_n^*(k)$ is the probability that a node of degree k is connected to a node of

degree ℓ , the function $\Theta_n(k)$ computes that average value of F_n^* (rank) over all neighbors of nodes with degree k .

Note that $F_n^*(\ell)$ is a *non-decreasing* function of the degree, that is the degrees are ranked from low to high. Then, just as $\Phi_n(k)$, the ANNR $\Theta_n(k)$ tends to be increasing for assortative networks, and decreasing for disassortative networks. When the wiring is neutral, such as in the configuration model, the ANNR should be a constant, the value of which we will derive below.

Note that due to the straightforward equality $\sum_{\ell>0} h_n(k, \ell) = f_n^*(k)$, the value of $\Theta_n(k)$ always lies in the interval $[0, 1]$. This solves the scaling problem. Further, the results in this section require only that \mathcal{D} has finite $(1 + \eta)$ -moment.

Similarly to the ANND, we define the limiting version of the ANNR:

$$\Theta(k) = \mathbb{1}_{\{f(k)>0\}} \frac{\sum_{\ell>0} h(k, \ell) F^*(\ell)}{f^*(k)}. \quad (12)$$

The next theorem establishes convergence of $\Theta_n(k)$ to $\Theta(k)$. See Section 9.6 for the proof.

Theorem 6.1 (General convergence of ANNR). *Let $\{G_n\}_{n \geq 1}$ be any sequence of graphs of size n that satisfies Assumptions 4.2 and 5.1. Then, for each fixed k such that $f(k) > 0$ and each $0 < \delta < \min\{\frac{\eta}{8+4\eta}, \kappa\}$,*

$$\lim_{n \rightarrow \infty} \mathbb{P}(|\Theta_n(k) - \Theta(k)| > n^{-\delta}) = 0.$$

Remark 6.2. *Since $|\Theta_n(k)|, |\Theta(k)| \leq 1$, by dominated convergence we have that*

$$\lim_{n \rightarrow \infty} \mathbb{E}[|\Theta_n(k) - \Theta(k)|] = 0.$$

Comparing Theorem 6.1 to Theorem 4.2 we see that convergence of the ANNR holds for any random variable with finite $(1 + \eta)$ -moment, instead of finite $(2 + \eta)$ -moment of \mathcal{D} . This makes the ANNR a more suitable measure for degree-degree correlations than the ANND.

Now consider CM with $\mathbf{D}_n = \text{IID}(\mathcal{D})$. Then we have that $h(k, \ell) = f^*(k)f^*(\ell)$ and hence we can write

$$\Theta(k) = \sum_{\ell>0} f^*(\ell) F^*(\ell) = \mathbb{E}[F^*(\mathcal{D}^*)],$$

where \mathcal{D}^* has the cumulative distribution function F^* . Similarly to the ANND, we can directly apply Theorem 6.1 to get that $\Theta_n(k)$ in the CM converges in probability to $\mathbb{E}[F^*(\mathcal{D}^*)]$, for all k such that $f(k) > 0$. In addition, we can prove a stronger result for ANNR, which is similar to Theorem 5.1, with the exception of only needing a finite $(1 + \eta)$ -moment.

Theorem 6.3 (Convergence of ANNR in CM). *Let \mathcal{D} be an integer-valued random variable which satisfies $\mathbb{E}[D^{1+\eta}] < \infty$, for some $0 < \eta < 1$, and let $\{G_n\}_{n \geq 1}$ be a sequence of graphs generated by CM with $\mathbf{D}_n = \text{IID}(\mathcal{D})$. Then, for any $\delta < \frac{\eta}{8+4\eta}$, as $n \rightarrow \infty$,*

$$\sup_{k \geq 0} \mathbb{P}(|\Theta_n(k) - \mathbb{E}[F^*(\mathcal{D}^*)]| \mathbb{1}_{\{f_n(k)>0\}} > n^{-\delta}) = O(n^{-\delta}).$$

To establish convergence of the ANNR in the Erased Configuration Model we follow a similar approach as for the ANND and study the difference $|\widehat{\Theta}_n(k) - \Theta_n(k)|$. We will prove that this difference converges to zero, in probability.

Theorem 6.4. *Let \mathcal{D} be regularly varying with exponent $\gamma > 1$ and let $\{G_n\}_{n \geq 1}$ be a sequence of graphs generated by ECM with $\mathbf{D}_n = \text{IID}(\mathcal{D})$. Then, for any $\delta > 0$ and k such that $f^*(k) > 0$,*

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(|\widehat{\Theta}_n(k) - \Theta_n(k)| > \delta\right) = 0.$$

With this result we obtain the following version of Theorem 6.3 for the ECM.

Theorem 6.5 (Convergence of ANNR in ECM). *Let \mathcal{D} be regularly varying with exponent $\gamma > 1$ and let $\{G_n\}_{n \geq 1}$ be a sequence of graphs generated by ECM with $\mathbf{D}_n = \text{IID}(\mathcal{D})$. Then, for any $\delta > 0$ and k such that $f^*(k) > 0$,*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\left| \widehat{\Theta}_n(k) - \mathbb{E}[F^*(\mathcal{D}^*)] \right| \mathbb{1}_{\{f_n(k) > 0\}} > \delta \right) = 0.$$

Similar to Theorem 5.8, this result follows from Theorem 6.3 and 6.4 by splitting

$$\left| \widehat{\Theta}_n(k) - \mathbb{E}[F^*(\mathcal{D}^*)] \right| \leq |\Theta_n(k) - \mathbb{E}[F^*(\mathcal{D}^*)]| + \left| \widehat{\Theta}_n(k) - \Theta_n(k) \right|.$$

Note, however, that Theorem 6.5 is less strong than Theorem 5.8, since we only have convergence in probability. Replacing the lower bound δ inside the probability with $n^{-\delta}$ requires a better understanding of the scaling of the number of removed stubs of nodes, which falls outside the focus of this paper. Still, we show in the next section that the numerical results for ANNR in ECM are as strong as in CM.

7 Numerical experiments

In this section we will illustrate our results for ANND and ANNR in CM of different sizes n and regularly-varying degree distributions. For the sake of illustration, we take $\gamma = 2.5$ (finite variance) and $\gamma = 1.5$ (infinite variance). We generate 100 realizations of each graph with each γ .

Here we will show results for the multi-graph CM. We have also computed the ANND and the ANNR for ECM. However, for small values of k the results for CM and ECM were very close, as predicted by Theorem 5.8, and for large values of k the results were unstable and inconclusive because of the small sample size. Since our main point here is to demonstrate the advantages of the ANNR, we will omit the results for ECM and leave this topic for future research.

In order to generate power law degree sequences, let X be a Pareto random variable with exponent $\gamma > 1$,

$$\mathbb{P}(X > t) = t^{-\gamma} \quad \text{for all } t \geq 1.$$

Then $\mathcal{D} = \lfloor X \rfloor$, which is integer-valued, satisfies

$$\mathbb{P}(\mathcal{D} = k) = \mathbb{P}(k \leq X < k + 1) = k^{-\gamma} - (k + 1)^{-\gamma} \sim \gamma t^{-\gamma-1}. \quad (13)$$

In particular, it follows that $\mathcal{D} = \lfloor X \rfloor$ satisfies (4) with eventually monotone slowly-varying function. Hence it is regularly varying with exponent γ .

We can express the first and second moment of \mathcal{D} through the Riemann zeta function ζ as follows:

$$\nu_1 = \zeta(\gamma) \quad \text{and} \quad \nu_2 = 2\zeta(\gamma - 1) - \zeta(\gamma),$$

from which we obtain that $\nu_2/\nu_1 = 2\zeta(\gamma - 1)/\zeta(\gamma) - 1$. For specific numerical values see Table 1.

Now, for each value of γ and n , we generate 100 graphs G_n , using CM with $\mathbf{D}_n = \text{IID}(\mathcal{D})$. Then, for each k , we compute the average of $\Phi_n(k)$ and $\Theta_n(k)$, with respect to the 100 generated graphs and plot these as a function of k .

7.1 Computing $\mathbb{E}[F^*(\mathcal{D}^*)]$

In order to validate our results for the ANNR in CM we need to know the value of its limit $\mathbb{E}[F^*(\mathcal{D}^*)]$. Elementary computations yield

$$F^*(k) = \frac{1}{\zeta(\gamma)} \sum_{t=1}^k t^{-\gamma} - \frac{k(k+1)^{-\gamma}}{\zeta(\gamma)}.$$

Now, since \mathcal{D}^* has probability density function $f^*(k) = kf(k)/\zeta(\gamma)$ we have that

$$\begin{aligned}\mathbb{E}[F^*(\mathcal{D}^*)] &= \sum_{k=1}^{\infty} F^*(k)f^*(k) \\ &= \frac{1}{\zeta(\gamma)^2} \sum_{k=1}^{\infty} \left(kf(k) \sum_{t=1}^k t^{-\gamma} - k^2(k+1)^{-\gamma} f(k) \right) \\ &= \frac{1}{\zeta(\gamma)^2} \sum_{k=1}^{\infty} \left((k^{1-\gamma} - k(k+1)^{-\gamma}) \sum_{t=1}^k t^{-\gamma} - k^{2-\gamma}(k+1)^{-\gamma} + k^2(k+1)^{-2\gamma} \right).\end{aligned}$$

The above expression cannot be written in terms of known functions. Hence, we need to evaluate it numerically. To this end, we chose N large enough such that the first two digits of $\sum_{k=1}^M F^*(k)f^*(k)$ remain the same for $M \geq N$. We remark that for $\gamma = 1.5$ we had to take $N = 10^7$ while we could not achieve the precision up to the third digit even for $N = 5 \times 10^8$. The results are shown in Table 1.

γ	ν_2/ν_1	$\mathbb{E}[F^*(\mathcal{D}^*)]$
1.5	-	0.545542
1.8	-	0.592175
2.0	-	0.625316
2.2	6.502744	0.658512
2.5	2.894745	0.706477

Table 1: Numerical values of the limit of Φ_n and Θ_n as given by, respectively, Theorem 5.1 and Theorem 6.1, for \mathcal{D} with distribution (13) and different values of γ .

7.2 ANND and ANNR

We compute the ANND and the ANNR for 100 realizations of the sequence $\mathbf{D}_n = \text{IID}(\mathcal{D})$. Recall that when there is no degree k in \mathbf{D}_n , then $\Phi_n(k)$ and $\Theta_n(k)$ are set to zero. The values of $\Phi_n(k)$ and $\Theta_n(k)$ (including the zero values) are averaged over 100 CM graphs with $\mathbf{D}_n = \text{IID}(\mathcal{D})$. Figures 1 and 2 show the results for Φ_n and Θ_n , respectively.

The first observation is that the ANND clearly diverges when $\gamma < 2$, while this does not happen for the ANNR.

The second observation is that both functions decline sharply when k exceeds a certain threshold, which increases in n . This phenomenon is explained by Theorem 3.1. Indeed, according to this theorem, when k is of a larger order of magnitude than $n^{1/(\gamma+1)}$, then the probability that degree k is present in $\mathbf{D}_n = \text{IID}(\mathcal{D})$ is smaller than one and in fact converges to zero. In the numerical experiments, for large k , degree k is not present in all 100 sampled sequences. Then $\Phi_n(k)$ and $\Theta_n(k)$ are set to zero, and this artificially decreases the average value of $\Phi_n(k)$ and $\Theta_n(k)$ over the 100 samples. Note that in Figures 1 and 2 the decline indeed starts around the point $k = n^{1/(\gamma+1)}$.

We can correct for this by counting, for each k , the number of graphs that have nodes of degree k and then dividing by this number instead of dividing by the sample size 100. The result for the corrected ANND with $\gamma = 2.5$ is shown in Figure 3. Observe that the plot no longer shows the decrease with k , and for small k the value of $\Phi_n(k)$ is unchanged. However, for large k we obtain very unstable fluctuations due to the small sample size. This can be possibly remedied, at least in the finite variance scenario, by averaging over a larger number of graphs. How large the sample size should be to reach stability for large k , and under which conditions it is possible, is another interesting question for future research.

We have also obtained numerical results for the corrected $\widehat{\Phi}_n(k)$ in the ECM. These are omitted because the plots showed similar high fluctuations as in Figure 3, and numerical comparison between CM and ECM was inconclusive.

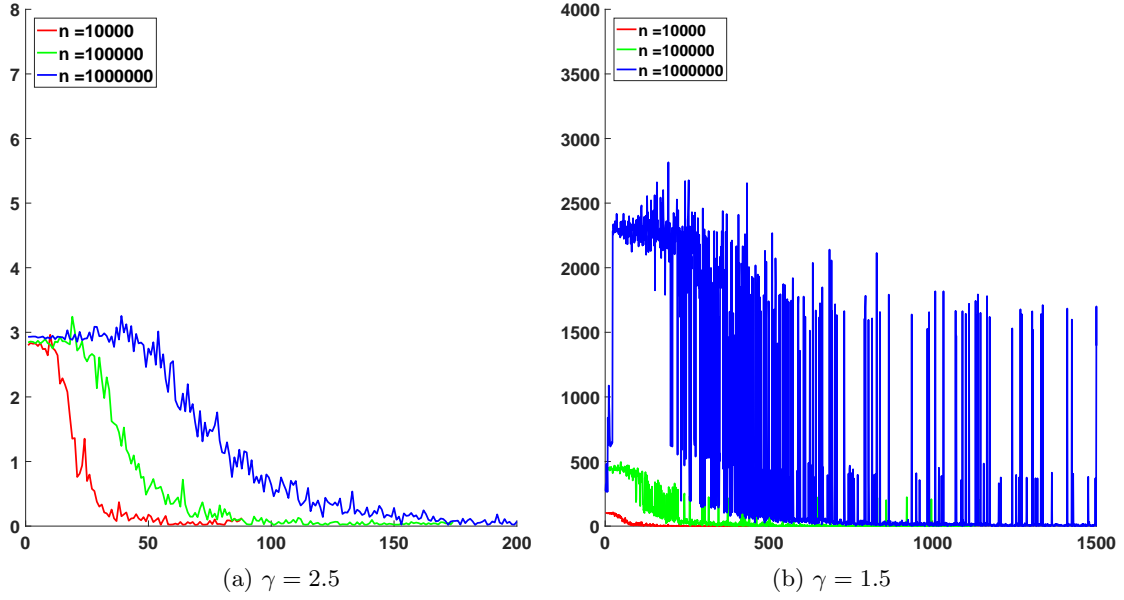


Figure 1: Plots for $\Phi_n(k)$, as a function of k , based on 100 CM graphs with $\mathbf{D}_n = \text{IID}(\mathcal{D})$.

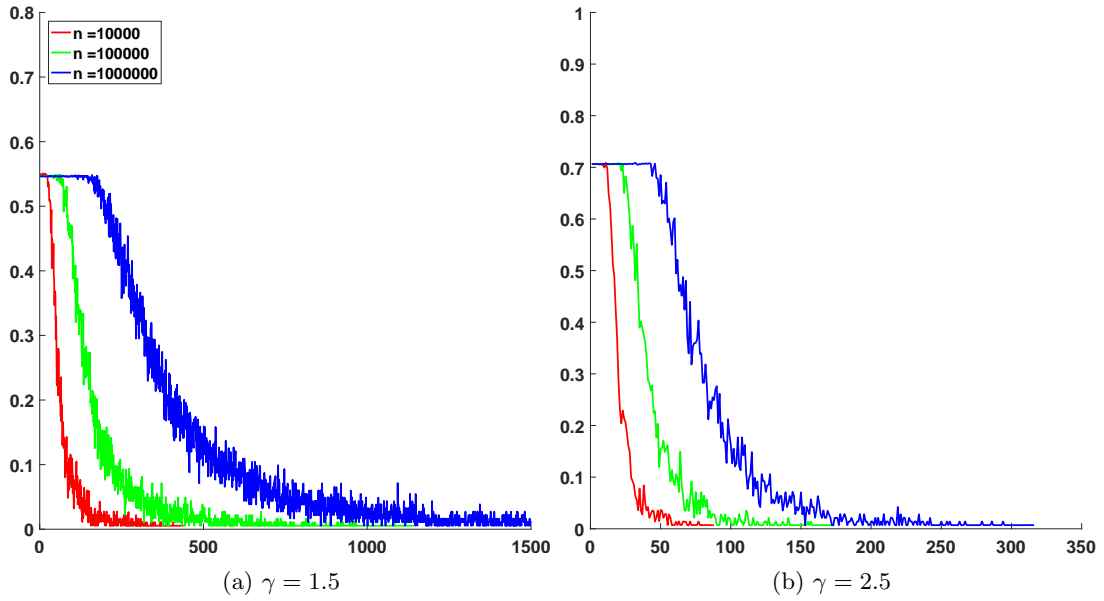


Figure 2: Plots for $\Theta_n(k)$, as a function of k , based on 100 CM graphs with $\mathbf{D}_n = \text{IID}(\mathcal{D})$.

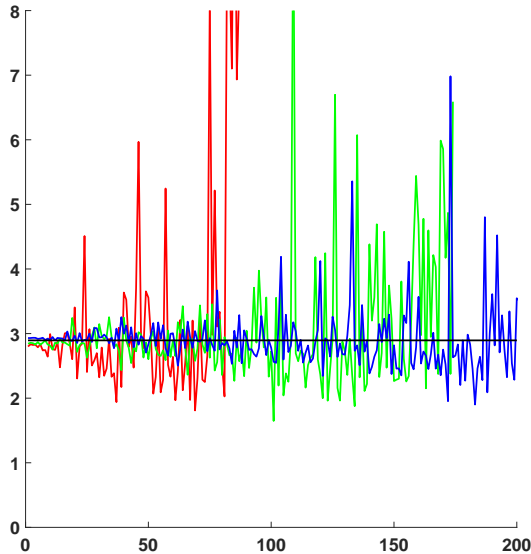


Figure 3: The corrected average of $\Phi_n(k)$, for 100 graphs with $\gamma = 2.5$. The black line represents the theoretical limit ν_2/ν_1 .

Since $\Theta_n(k) \leq 1$ for all k and n , ANNR remains stable after correction even in the infinite variance scenario. This is clearly observed in Figure 4 where we present the corrected plot of $\Theta_n(k)$ in CM, for $\gamma = 2.5$ and $\gamma = 1.5$. In both cases the plot very closely follows the straight line, which represents the limit value $\mathbb{E}[F^*(\mathcal{D})]$.

Finally, in Figure 5 we plot the corrected version of ANNR for the ECM. We see that the ANNR in ECM again shows a great stability, and we clearly observe the point-wise convergence to the right constant for each fixed k , as stated in Theorems 6.3 and 6.4. We also clearly observe structural correlations, or finite-size effects, for top values of the degrees, especially for $\gamma = 1.5$. Since the rank correlations are not affected by high dispersion in the values of the degrees, these finite-size effects can only be explained by simplicity of the graph. This is in agreement with previous work [18], where we observed structural correlations in another rank-based dependency measure – Spearman’s rho. We see that for ANNR, these effects appear only when k is very large, say, greater than some $k_{critical}(n)$. One can expect that $k_{critical}(n)$ scales as a positive power of n , however, establishing the exact scaling for $k_{critical}(n)$ seems to be a difficult problem.

8 Discussion

The most important implication of our results is that the ANND cannot be used in the case when degrees have infinite variance. This is a very similar situation as was observed before for Pearson’s correlation coefficient [14]. In particular, the ANND scales with the graph size, which makes it not suitable for comparison of networks of the same structure but different sizes. Moreover, even when rescaled, the ANND converges to a random variable instead of a number, and therefore it is impossible to establish any meaningful relation between the scaled ANND and the network’s assortativity. Therefore, the use of the ANNR is strongly recommended.

In addition, we would like to mention two interesting open problems, stemming from this research, that we will address in the near future. First of all, in fact, the ANND deals with ‘double sampling’ as follows. 1) In order to create a graph of size n , the degree sequence $\text{IID}(\mathcal{D})$ is sampled. As we proved in Theorem 3.1, each specific sufficiently large value k will appear in such sequence only with a small probability, therefore it will take some time to sample several sequences that include such degree. 2) Each node in the network samples its neighbors. This double sampling gives rise to vast fluctuations of the ANND for large values of k even in the CM

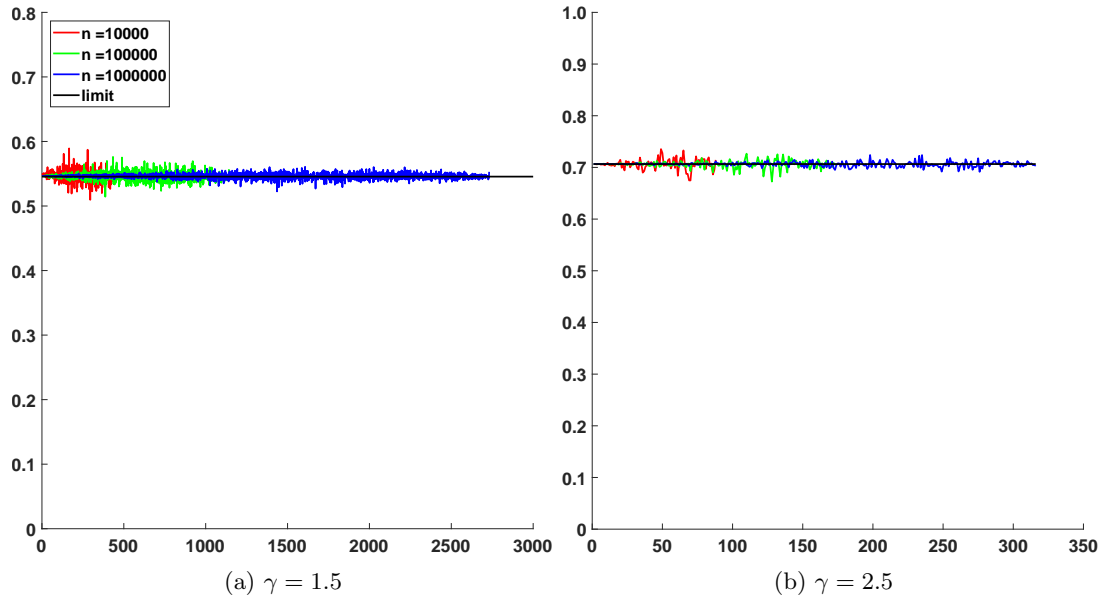


Figure 4: Corrected average of Θ_n in CM, for 100 graphs with $\gamma = 2.5$ and $\gamma = 1.5$. The black line represents the theoretical limit $\mathbb{E}[F^*(\mathcal{D}^*)]$.

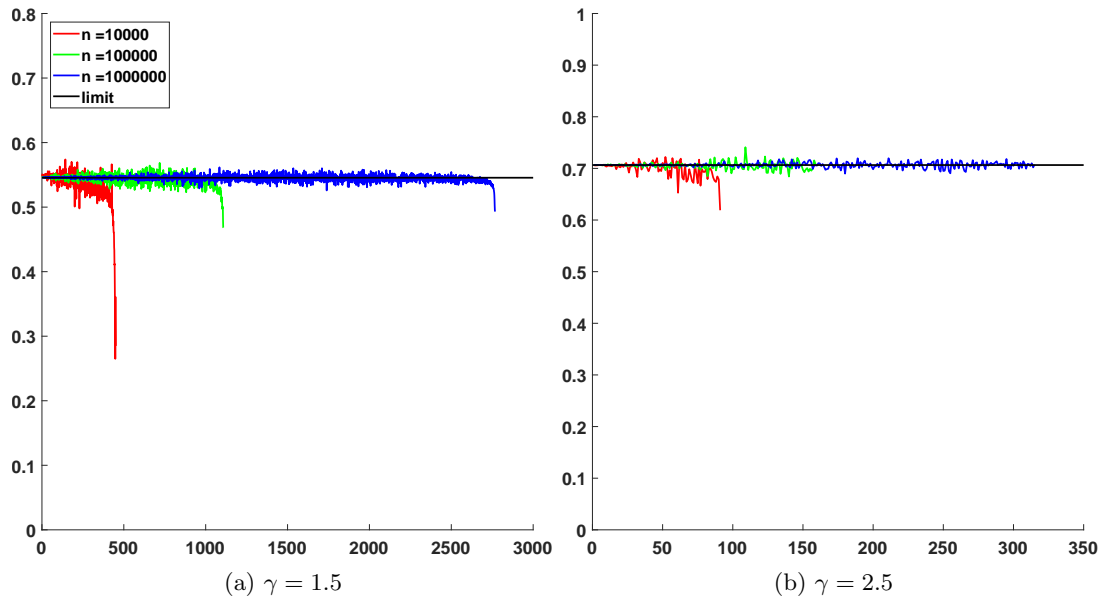


Figure 5: Corrected average of Θ_n in ECM, for 100 graphs with $\gamma = 2.5$ and $\gamma = 1.5$. The black line represents the theoretical limit $\mathbb{E}[F^*(\mathcal{D}^*)]$.

with finite variance of the degrees, despite of the convergence of the ANND to a deterministic (constant) limit, that has been proved in Theorem 5.1. We observed this phenomenon in Figure 3. The magnitude of these fluctuations and the number of graphs necessary to obtain convergence of the ANND for all k have not been addressed in this paper, and are the focus of future research.

Second, it has been observed in the literature [1,8] that the ANND in the ECM decreases for large values of k . This phenomenon has been ascribed to the simplicity of the graph, because large nodes do not have other large nodes to connect to, so they must create disassortative connections. These ‘structural negative correlations’, or ‘finite-size effects’ are broadly recognized in the literature, see [15] and references therein. Interestingly, our results state that ANND in the CM and the ECM are asymptotically equivalent for any fixed k , and the same holds for the ANNR. However, numerical results suggest that ANND and even ANNR are subject to the finite-size effects for k that scales as a power of n . The exact mathematical characterization of the finite-size effects in terms of ANND and ANNR remains an interesting open problem.

9 Proofs

In this section, we give the proofs of the results in the paper.

9.1 Regularly-varying degrees

We start with a small result, relating regular variation of $\mathbb{P}(X = k)$, for some integer-valued random variable, to that of the inverse cdf $\mathbb{P}(X > t)$.

Lemma 9.1. *Suppose that X is an integer-valued random variable with probability function*

$$\mathbb{P}(X = k) = l(k)k^{-\gamma-1}, \quad k = 1, 2, \dots,$$

with $\gamma > 1$ and some slowly-varying function $l(x)$ that is eventually monotone. Then, as $t \rightarrow \infty$,

$$\mathbb{P}(X > t) \sim \frac{l(t)}{\gamma} t^{-\gamma}.$$

In particular this implies that

$$\mathbb{P}(X > t) = \tilde{l}(t)t^{-\gamma},$$

with $\tilde{l}(x) \sim l(x)/\gamma$.

Proof. To prove the result we will bound the sums $\sum_{k>t}^{\infty} \mathbb{P}(X = k)$ by integrals, using that $l(x)$ is eventually monotone. We will assume that $l(x)$ is eventually monotonic decreasing. The proof for monotonic increasing $l(x)$ is similar.

First observe that

$$\mathbb{P}(X > t) \sim \frac{l(t)}{\gamma} t^{-\gamma},$$

implies that $\tilde{l}(x) = x^\gamma \mathbb{P}(X > x)$ is slowly varying and $\tilde{l}(x) \sim l(x)/\gamma$, which proves the second statement.

For the first statement, let K be the smallest integer such that for all $y \geq x \geq K$, $l(x) \geq l(y)$. In addition define for all $x \in [1, \infty)$, $f(x) = l(x)x^{-\gamma-1}$. Then for all $t \geq K + 1$

$$\begin{aligned} \sum_{k=\lceil t \rceil}^{\infty} \mathbb{P}(X = k) &= \mathbb{P}(X = \lceil t \rceil) + \sum_{k=\lceil t \rceil+1}^{\infty} \int_{k-1}^k \mathbb{P}(X = k) dx \\ &\leq \mathbb{P}(X = \lceil t \rceil) + \sum_{k=\lceil t \rceil+1}^{\infty} \int_{k-1}^k f(x) dx \\ &= \mathbb{P}(X = \lceil t \rceil) + \int_{\lceil t \rceil}^{\infty} f(x) dx. \end{aligned}$$

Similarly, we have for all $t \geq K + 1$,

$$\sum_{k=\lceil t \rceil}^{\infty} \mathbb{P}(X = k) = \sum_{k=\lceil t \rceil}^{\infty} \int_k^{k+1} \mathbb{P}(X = k) dx \geq \int_{\lceil t \rceil}^{\infty} f(x) dx.$$

We therefore obtain

$$1 \leq \frac{\sum_{k=\lceil t \rceil}^{\infty} \mathbb{P}(X = k)}{\int_{\lceil t \rceil}^{\infty} f(x) dx} \leq 1 + \frac{\mathbb{P}(X = \lceil t \rceil)}{\int_{\lceil t \rceil}^{\infty} f(x) dx}. \quad (14)$$

By Karamata's theorem ([3, 1.5.11]) it follows that, as $t \rightarrow \infty$,

$$\int_t^{\infty} f(x) dx \sim \frac{l(t)t^{-\gamma}}{\gamma}. \quad (15)$$

Since $\mathbb{P}(X = \lceil t \rceil) = o(l(t)t^{-\gamma})$, it follows from (14) and (15) that, as $t \rightarrow \infty$,

$$\frac{\gamma \mathbb{P}(X > t)}{l(t)t^{-\gamma}} = \left(\frac{\gamma \int_{\lceil t \rceil}^{\infty} f(x) dx}{l(t)t^{-\gamma}} \right) \left(\frac{\sum_{k=\lceil t \rceil}^{\infty} \mathbb{P}(X = k)}{\int_{\lceil t \rceil}^{\infty} f(x) dx} \right) \sim 1,$$

which finishes the proof. \square

9.2 Proof of Theorem 4.1

The proof we present here is an adjustment of [15, Theorem 3.1] to the setting with finite variance. It uses the following technical lemma, which is a direct consequence of the Burkholder's inequality. See for instance [15] for a short proof.

Lemma 9.2 (Lemma 3.5 [15]). *Let $\{X_i\}_{i \geq 1}$ be a sequence of i.i.d. zero mean random variables such that $\mathbb{E}[|X_1|^p] < \infty$, for some $1 < p < 2$. Then there exists a constant $K > 0$, which only depends on p , such that*

$$\mathbb{E} \left[\left| \sum_{i=1}^n X_i \right|^p \right] \leq Kn \mathbb{E}[|X_1|^p].$$

Proof of Theorem 4.1. Let $\varepsilon \leq \eta/(4 + 2\eta)$, define the events

$$A_n = \{d_1(f_n, f) \leq n^{-\varepsilon}\}, \quad B_n = \{d_1(f_n^*, f^*) \leq n^{-\varepsilon}\}$$

and note that $\mathbb{P}(\Omega_n^c) \leq \mathbb{P}(A_n^c) + \mathbb{P}(A_n \cap B_n^c)$. Therefore it is enough to show that

$$\mathbb{P}(A_n^c) + \mathbb{P}(A_n \cap B_n^c) = O(n^{-\varepsilon}),$$

as $n \rightarrow \infty$.

Since $\mathbb{E}[\mathcal{D}^{1+\eta}] < \infty$ we use [10, Proposition 4], with $\alpha = 1 + \eta$, to obtain

$$\mathbb{P}(A_n^c) \leq n^\varepsilon \mathbb{E}[d_1(f_n, f)] \leq n^{\varepsilon-1+1/(1+\eta)} \left(\frac{2(1+\eta)}{\eta} + \frac{2}{1-\eta} \right) \mathbb{E}[\mathcal{D}^{1+\eta}] = O(n^{-\varepsilon}),$$

where the last part follows from the fact that $\varepsilon - 1 + 1/(1 + \eta) < -\varepsilon$.

For the remaining probability, define

$$X_{ik} = D_i \mathbb{1}_{\{D_i > k\}} - \mathbb{E}[\mathcal{D} \mathbb{1}_{\{\mathcal{D} > k\}}],$$

and write

$$d_1(f_n^*, f^*) = \sum_{k=0}^{\infty} \left| \frac{1}{L_n} \sum_{i=1}^n D_i \mathbb{1}_{\{D_i > k\}} - \mathbb{E}[\mathcal{D} \mathbb{1}_{\{\mathcal{D} > k\}}] \right|$$

$$\begin{aligned}
&\leq \frac{1}{\nu n} \sum_{k=0}^{\infty} \left| \sum_{i=1}^n X_{ik} \right| + \left| \frac{1}{L_n} - \frac{1}{\nu n} \right| \sum_{k=0}^{\infty} D_i \mathbb{1}_{\{D_i \geq k\}} \\
&\leq \frac{1}{\nu n} \sum_{k=0}^{\infty} \left| \sum_{i=1}^n X_{ik} \right| + \frac{|L_n - \nu n|}{\nu n}.
\end{aligned}$$

Since on the event A_n we have $|L_n - \nu n| \leq n^{1-\varepsilon}$ we get

$$\mathbb{P} \left(\frac{|L_n - \nu n|}{\nu n} > \frac{2n^{-\varepsilon}}{\nu + 1}, A_n \right) = \mathbb{P} \left(\frac{n^{-\varepsilon}}{\nu} > \frac{2n^{-\varepsilon}}{\nu + 1}, A_n \right) = 0,$$

and hence

$$\begin{aligned}
\mathbb{P}(A_n \cap B_n^c) &= \mathbb{P}(d_1(f_n^*, f^*) > n^{-\varepsilon}, A_n) \\
&\leq \mathbb{P} \left(\frac{1}{\nu n} \sum_{k=0}^{\infty} \left| \sum_{i=1}^n X_{ik} \right| > \frac{(\nu - 1)n^{-\varepsilon}}{\nu + 1} \right) + \mathbb{P} \left(\frac{|L_n - \nu n|}{\nu n} > \frac{2n^{-\varepsilon}}{\nu + 1}, A_n \right) \\
&= \mathbb{P} \left(\frac{1}{\nu n} \sum_{k=0}^{\infty} \left| \sum_{i=1}^n X_{ik} \right| > \frac{(\nu - 1)n^{-\varepsilon}}{\nu + 1} \right).
\end{aligned}$$

For the last term let $p = 1/(1 - 2\varepsilon)$ and note that since $0 < \eta < 1$ and $0 < \varepsilon \leq \eta/(2\eta + 4)$ we have $1 < p \leq 1 + \eta/2$. Therefore, by first applying Markov's inequality, followed by Hölder's inequality and then Lemma 9.2, we get for some $K > 0$ depending only on ε ,

$$\begin{aligned}
\mathbb{P} \left(\frac{1}{\nu n} \sum_{k=0}^{\infty} \left| \sum_{i=1}^n X_{ik} \right| > \frac{\nu n^{-\varepsilon}}{\nu + 1} \right) &\leq \frac{\nu + 1}{\nu^2 n^{1-\varepsilon}} \sum_{k=0}^{\infty} \mathbb{E} \left[\left| \sum_{i=1}^n X_{ik} \right| \right] \\
&\leq \frac{\nu + 1}{\nu^2 n^{1-\varepsilon}} \sum_{k=0}^{\infty} \mathbb{E} \left[\left| \sum_{i=1}^n X_{ik} \right|^p \right]^{\frac{1}{p}} \\
&\leq \frac{\nu + 1}{\nu^2 n^{1-\varepsilon}} \sum_{k=0}^{\infty} (K n \mathbb{E} [|X_{1k}|^p])^{\frac{1}{p}} \\
&\leq \frac{K^{\frac{1}{p}} n^{\frac{1}{p}} (\nu + 1)}{\nu^2 n^{1-\varepsilon}} \sum_{k=0}^{\infty} \mathbb{E} [|X_{1k}|^p]^{\frac{1}{p}} \\
&\leq \frac{2K^{\frac{1}{p}} n^{\frac{1}{p}} (\nu + 1)}{\nu^2 n^{1-\varepsilon}} \sum_{k=0}^{\infty} \mathbb{E} [\mathcal{D}^p \mathbb{1}_{\{\mathcal{D} > k\}}]^{\frac{1}{p}}.
\end{aligned}$$

To finish the argument we write

$$\begin{aligned}
\mathbb{E} [\mathcal{D}^p \mathbb{1}_{\{\mathcal{D} > k\}}] &= \mathbb{E} [\mathcal{D}^p \mathbb{1}_{\{\mathcal{D} > k\}}] \mathbb{1}_{\{k < 1\}} + \mathbb{E} [\mathcal{D}^p \mathbb{1}_{\{\mathcal{D} > k\}}] \mathbb{1}_{\{k \geq 1\}} \\
&\leq \mathbb{E} [\mathcal{D}^{p+1}] \mathbb{1}_{\{k < 1\}} + k^{-p} \mathbb{E} [\mathcal{D}^{2p}] \mathbb{1}_{\{k \geq 1\}},
\end{aligned}$$

so that, using $\Gamma(s)$ to denote the Gamma function,

$$\begin{aligned}
\frac{2K^{\frac{1}{p}} n^{\frac{1}{p}} (\nu + 1)}{\nu^2 n^{1-\varepsilon}} \sum_{k=0}^{\infty} \mathbb{E} [\mathcal{D}^p \mathbb{1}_{\{\mathcal{D} > k\}}]^{\frac{1}{p}} &\leq \frac{2K^{\frac{1}{p}} n^{\frac{1}{p}} (\nu + 1)}{\nu^2 n^{1-\varepsilon}} \left(\mathbb{E} [\mathcal{D}^{p+1}] + \mathbb{E} [\mathcal{D}^{2p}] \sum_{k=1}^{\infty} k^{-p} \right) \\
&= n^{\frac{1}{p} + \varepsilon - 1} \frac{2K^{\frac{1}{p}} (\nu + 1)}{\nu^2} (\mathbb{E} [\mathcal{D}^{p+1}] + \mathbb{E} [\mathcal{D}^{2p}] \Gamma(p)) \\
&= O \left(n^{\frac{1}{p} + \varepsilon - 1} \right) = O(n^{-\varepsilon}),
\end{aligned}$$

where we used that $\mathbb{E} [\mathcal{D}^{p+1}] \leq \mathbb{E} [\mathcal{D}^{2p}] \leq \mathbb{E} [\mathcal{D}^{2+\eta}] < \infty$ and $1/p + \varepsilon - 1 = -\varepsilon$. \square

9.3 ANND in general graphs

Proof of Theorem 4.2. Let Ω_n and Γ_n be as defined in Assumptions 4.1 and 4.2, respectively. Define $\Lambda_n = \Omega_n \cap \Gamma_n$. Then by the assumptions we have that

$$\lim_{n \rightarrow \infty} \mathbb{P}(\Lambda_n) = 1, \quad (16)$$

and hence it is enough to prove the result conditioned on the event Λ_n . For this we first split $|\Phi(k) - \Phi_n(k)|$ into two terms as follows:

$$\begin{aligned} |\Phi(k) - \Phi_n(k)| \mathbb{1}_{\{f_n(k) > 0\}} &\leq \mathbb{1}_{\{f_n(k) > 0\}} \left| \frac{1}{f_n^*(k)} - \frac{1}{f^*(k)} \right| \sum_{\ell=1}^{\infty} h_n(k, \ell) \ell \\ &\quad + \mathbb{1}_{\{f^*(k) > 0\}} \frac{1}{f^*(k)} \left| \sum_{\ell=1}^{\infty} h_n(k, \ell) \ell - h(k, \ell) \ell \right| := \Xi_n^{(1)} + \Xi_n^{(2)} \end{aligned}$$

We obtain a bound for $\Xi_n^{(1)}$ on Λ_n by bounding both multiplicative terms in $\Xi_n^{(1)}$. To this end, first, on Γ_n we obtain

$$\begin{aligned} \sum_{\ell=1}^{\infty} h_n(k, \ell) \ell &\leq \sum_{k, \ell=1}^{\infty} h(k, \ell) \ell + \left| \sum_{k, \ell=1}^{\infty} h_n(k, \ell) \ell - h(k, \ell) \ell \right| \\ &\leq \sum_{\ell=1}^{\infty} f^*(\ell) \ell + d_1(f_n^*, f^*) = \mathbb{E}[\mathcal{D}^2] + n^{-\varepsilon}. \end{aligned} \quad (17)$$

Next we see that on Ω_n ,

$$\left| \frac{1}{f_n^*(k)} - \frac{1}{f^*(k)} \right| = \frac{|f_n^*(k) - f^*(k)|}{f_n^*(k) f^*(k)} \leq \frac{d_1(f_n^*, f^*)}{f^*(k)^2 - f^*(k) n^{-\varepsilon}} \leq \frac{n^{-\varepsilon}}{f^*(k)^2 - f^*(k) n^{-\varepsilon}}. \quad (18)$$

Combining (17) and (18), we obtain

$$\Xi_n^{(1)} \leq \frac{n^{-\varepsilon}}{f^*(k)^2 - f^*(k) n^{-\varepsilon}} (\mathbb{E}[\mathcal{D}^2] + n^{-\varepsilon}) = O(n^{-\varepsilon}). \quad (19)$$

In order to bound $\Xi_n^{(2)}$, we use a cut-off technique. Choose $w_n = \lfloor n^p \rfloor$, where p is a positive constant to be determined. Then we write

$$\Xi_n^{(2)} \leq \frac{\mathbb{1}_{\{f_n(k) > 0\}}}{f^*(k)} \left| \sum_{\ell=1}^{w_n} h_n(k, \ell) \ell - h(k, \ell) \ell \right| + \frac{1}{f^*(k)} \left| \sum_{\ell > w_n} h_n(k, \ell) \ell - h(k, \ell) \ell \right| := \Xi_n^{(3)} + \Xi_n^{(4)}.$$

To control $\Xi_n^{(3)}$, we use that on Γ_n , $|\sum_{\ell=1}^{w_n} h_n(k, \ell) - h(k, \ell)| \leq n^{-\kappa}$, so that on Λ_n ,

$$\Xi_n^{(3)} \leq \frac{w_n |\sum_{\ell=1}^{w_n} h_n(k, \ell) - h(k, \ell)|}{f^*(k)} \leq \frac{n^{-\kappa+p}}{f^*(k)}. \quad (20)$$

For $\Xi_n^{(4)}$, we use that on Ω_n we have

$$\frac{1}{f^*(k)} \left| \sum_{\ell > w_n} h_n(k, \ell) \ell - h(k, \ell) \ell \right| \leq \frac{w_n^{-\eta}}{f^*(k)} \left(\sum_{\ell > w_n} h_n(k, \ell) \ell^{1+\eta} + h(k, \ell) \ell^{1+\eta} \right).$$

Now, we obtain

$$\sum_{\ell > w_n} h(k, \ell) \ell^{1+\eta} \leq \sum_{k, \ell > 0} h(k, \ell) \ell^{1+\eta} = \sum_{\ell > 0} f^*(\ell) \ell^{1+\eta} = \frac{1}{\nu_1} \sum_{\ell=0}^{\infty} f(\ell) \ell^{2+\eta} = \frac{\mathbb{E}[\mathcal{D}^{2+\eta}]}{\nu_1}.$$

Next, on Ω_n we have that

$$\begin{aligned} \sum_{\ell > w_n} h_n(k, \ell) \ell^{1+\eta} &\leq \sum_{k, \ell > 0} h_n(k, \ell) \ell^{1+\eta} = \sum_{\ell > 0} f_n^*(\ell) \ell^{1+\eta} \\ &= \frac{1}{L_n} \sum_{\ell > 0} \ell^{2+\eta} \sum_{i=1}^n \mathbb{1}_{\{D_i = \ell\}} \leq \frac{1}{\nu_1 n - n^{1-\varepsilon}} \sum_{i=1}^n D_i^{2+\eta}. \end{aligned}$$

Therefore, using Markov's inequality, we have, for all $0 < c < 1$,

$$\begin{aligned} \mathbb{P}\left(\Xi_n^{(4)} > cn^{-\delta}, \Lambda_n\right) &\leq \frac{n^\delta w_n^{-\eta}}{c f_n^*(k)} \left(\frac{\mathbb{E}[\mathcal{D}^{2+\eta}]}{\nu_1 - n^{-\varepsilon}} + \frac{\mathbb{E}[\mathcal{D}^{2+\eta}]}{\nu_1} \right) \\ &= \frac{n^\delta w_n^{-\eta}}{c f_n^*(k)(\nu_1 - n^{-\varepsilon})} (2\mathbb{E}[\mathcal{D}^{2+\eta}] + n^{-\varepsilon}) = O(n^{\delta-p\eta}). \end{aligned} \quad (21)$$

We now use a standard bound to obtain

$$\begin{aligned} \mathbb{P}\left(|\Phi(k) - \Phi_n(k)| > n^{-\delta}, \Lambda_n\right) &\leq \mathbb{P}\left(\Xi_n^{(1)} > \frac{1}{3} n^{-\delta}, \Lambda_n\right) \\ &\quad + \mathbb{P}\left(\Xi_n^{(3)} > \frac{1}{3} n^{-\delta}, \Lambda_n\right) + \mathbb{P}\left(\Xi_n^{(4)} > \frac{1}{3} n^{-\delta}, \Lambda_n\right). \end{aligned}$$

It follows from (19), (20) and (21) that whenever $p < \kappa$ and

$$0 < \delta < \min\{\varepsilon, \kappa - p, p\eta\}, \quad (22)$$

we have

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(|\Phi(k) - \Phi_n(k)| > n^{-\delta}, \Lambda_n\right) = 0. \quad (23)$$

Finally, set $p = \kappa/(\eta + 1) > 0$ so that $\kappa - p = p\eta$. Then (22) holds for all $\delta < \min\left\{\varepsilon, \frac{\kappa\eta}{\eta+1}\right\}$, and the result follows from (23) and (16). \square

9.4 ANND in the configuration model

Proof of Theorem 5.1. Let Ω_n be defined as in Assumption 4.1 with $\varepsilon = \delta/2 \leq \eta/(4 + 2\eta)$.

First, observe that

$$\frac{\nu_2}{\nu_1} = \sum_{\ell=1}^{\infty} f^*(\ell) \ell.$$

Hence we get

$$\begin{aligned} \left| \Phi_n(k) - \frac{\nu_2}{\nu_1} \right| \mathbb{1}_{\{f_n(k) > 0\}} &= \frac{1}{f_n^*(k)} \left| \sum_{\ell > 0=1}^{\infty} (h_n(k, \ell) - f_n^*(k) f^*(\ell)) \ell \right| \mathbb{1}_{\{f_n(k) > 0\}} \\ &\leq \frac{1}{f_n^*(k)} \left| \sum_{\ell=1}^{\infty} (h_n(k, \ell) - f_n^*(k) f^*(\ell)) \ell \right| \mathbb{1}_{\{f_n(k) > 0\}} \end{aligned} \quad (24)$$

$$+ \left| \sum_{\ell=1}^{\infty} (f_n^*(\ell) - f^*(\ell)) \ell \right|. \quad (25)$$

Term (25) is independent of k and satisfies

$$\left| \sum_{\ell=1}^{\infty} (f_n^*(\ell) - f^*(\ell)) \ell \right| \leq d_1(f_n^*, f^*).$$

Therefore we have that

$$\mathbb{P}\left(\left|\sum_{\ell=1}^{\infty}(f_n^*(\ell) - f^*(\ell))\ell\right| > \frac{n^{-\delta}}{2}, \Omega_n\right) = O(n^{-\delta}).$$

We are now left with (24), which requires a bit more work. We will prove that there exists a constant C , independent of n and k , such that

$$\mathbb{P}\left(\left|\Phi_n(k) - \frac{\nu_2}{\nu_1}\right| \mathbb{1}_{\{f_n(k) > 0\}} > n^{-\delta}, \Omega_n\right) \leq C\left(n^{\delta - \frac{\eta}{2(\eta+2)}}\right). \quad (26)$$

This together with Theorem 4.1 will give the result.

Recall that G_{ij} is the number of edges from i to j where self-loops are counted twice. Let us now define

$$X_{ij}(k, \ell) = \mathbb{1}_{\{D_i=k, D_j=\ell\}} \left(\frac{G_{ij}}{L_n} - \frac{D_i D_j}{L_n^2} \right). \quad (27)$$

Then we have that

$$|h_n(k, \ell) - f_n^*(k)f_n^*(\ell)| = \left| \sum_{i,j=1}^n X_{ij}(k, \ell) \right|. \quad (28)$$

Now we will again use a cut-off technique. Let $p = 1/(2(\eta + 2))$, denote $w_n = \lfloor n^p \rfloor$ and use (28) to bound (24) as follows:

$$\begin{aligned} & \frac{1}{f_n^*(k)} \left| \sum_{\ell=1}^{\infty} (h_n(k, \ell) - f_n^*(k)f_n^*(\ell))\ell \right| \mathbb{1}_{\{f_n(k) > 0\}} \\ & \leq \frac{\mathbb{1}_{\{f_n(k) > 0\}}}{f_n^*(k)} \sum_{\ell=1}^{\infty} \left| \sum_{i,j=1}^n \ell X_{ij}(k, \ell) \right| \\ & \leq \frac{\mathbb{1}_{\{f_n(k) > 0\}}}{f_n^*(k)} \sum_{\ell=1}^{w_n} \left| \sum_{i,j=1}^n \ell X_{ij}(k, \ell) \right| + \frac{\mathbb{1}_{\{f_n(k) > 0\}}}{f_n^*(k)} \sum_{\ell > w_n} \sum_{i,j=1}^n D_j |X_{ij}(k, \ell)| \\ & := \Xi_n^{(1)} + \Xi_n^{(2)}. \end{aligned}$$

We will bound the probability $\mathbb{P}\left(\Xi_n^i > \frac{n^{-\delta}}{4}, \Omega_n\right)$ for $i = 1, 2$, separately, starting with $\Xi_n^{(2)}$. Observe that, if \mathbb{E}_n denotes the conditional expectation given the degree sequence, then it follows that

$$\mathbb{E}_n(G_{ij}) = \begin{cases} \frac{D_i D_j}{L_n - 1}, & \text{if } i \neq j, \\ \frac{D_i(D_i - 1)}{L_n - 1}, & \text{if } i = j, \end{cases}$$

and hence by (27) we get

$$\mathbb{E}_n(|X_{ij}(k, \ell)|) \leq 4 \frac{D_i D_j}{L_n^2} \mathbb{1}_{\{D_i=k, D_j=\ell\}}. \quad (29)$$

Therefore we obtain

$$\begin{aligned} \mathbb{P}\left(\Xi_n^{(2)} > \frac{n^{-\delta}}{4}, \Omega_n\right) &= \mathbb{P}\left(\frac{\mathbb{1}_{\{f_n(k) > 0\}}}{f_n^*(k)} \sum_{\ell > w_n} \sum_{i,j=1}^n D_j |X_{ij}(k, \ell)| > \frac{n^{-\delta}}{4}, \Omega_n\right) \\ &\leq 4n^\delta \sum_{\ell > w_n} \mathbb{E} \left[\frac{1}{f_n^*(k)} \mathbb{1}_{\{f_n(k) > 0\}} \left\| \sum_{i,j=1}^n X_{ij}(k, \ell) D_j \right\| \middle| \Omega_n \right] \end{aligned}$$

$$\begin{aligned}
&\leq 16n^\delta \sum_{i,j=1}^n \mathbb{E} \left[\frac{1}{f_n^*(k)} \mathbb{1}_{\{f_n(k)>0\}} \frac{D_i D_j^2}{L_n^2} \mathbb{1}_{\{D_j>w_n\}} \mathbb{1}_{\{D_i=k\}} \middle| \Omega_n \right] \\
&= 16n^\delta \sum_{j=1}^n \mathbb{E} \left[\frac{D_j^2}{L_n} \mathbb{1}_{\{D_j>w_n\}} \middle| \Omega_n \right] \\
&= 16n^{\delta+1} \mathbb{E} \left[\frac{D_1^2}{L_n} \mathbb{1}_{\{D_1>w_n\}} \middle| \Omega_n \right] \\
&\leq \frac{16n^\delta}{\nu_1 - n^{-\varepsilon}} \mathbb{E} [D^2 \mathbb{1}_{\{D>w_n\}}] \\
&\leq \frac{16}{\nu_1 - n^{-\varepsilon}} \mathbb{E} [D^{2+\eta}] n^\delta w_n^{-\eta}.
\end{aligned}$$

From this it follows that there exists a constant $C_1 > 0$, independent of k , such that

$$\mathbb{P} \left(\Xi_n^2 > \frac{n^{-\delta}}{4}, \Omega_n \right) \leq C_1 n^{\delta-p\eta} = C_1 n^{\delta - \frac{\eta}{2(\eta+2)}}. \quad (30)$$

Finally we deal with $\Xi_n^{(1)}$. Using Markov's inequality and Cauchy-Schwartz inequality, we get

$$\begin{aligned}
\mathbb{P} \left(\Xi_n^{(1)} > \frac{n^{-\delta}}{4}, \Omega_n \right) &= \mathbb{P} \left(\frac{\mathbb{1}_{\{f_n(k)>0\}}}{f_n^*(k)} \sum_{\ell=1}^{w_n} \left| \sum_{i,j=1}^n \ell X_{ij}(k, \ell) \right| > \frac{n^{-\delta}}{4}, \Omega_n \right) \\
&\leq 4n^\delta \sum_{\ell=1}^{w_n} \mathbb{E} \left[\frac{1}{f_n^*(k)} \mathbb{1}_{\{f_n(k)>0\}} \left| \sum_{i,j=1}^n X_{ij}(k, \ell) \ell \right| \middle| \Omega_n \right] \\
&= 4n^\delta \sum_{\ell=1}^{w_n} \ell \mathbb{E} \left[\frac{1}{f_n^*(k)} \mathbb{1}_{\{f_n(k)>0\}} \mathbb{E}_n \left[\left| \sum_{i,j=1}^n X_{ij}(k, \ell) \right| \right] \middle| \Omega_n \right] \\
&\leq 4n^\delta \sum_{\ell=1}^{w_n} \ell \mathbb{E} \left[\frac{1}{f_n^*(k)} \mathbb{1}_{\{f_n(k)>0\}} \mathbb{E}_n \left[\left| \sum_{i,j=1}^n X_{ij}(k, \ell) \right|^2 \right]^{1/2} \middle| \Omega_n \right]
\end{aligned}$$

In order to bound $\mathbb{E}_n \left[\left| \sum_{i,j=1}^n X_{ij}(k, \ell) \right|^2 \right]^{1/2}$ we use Lemma 9.3 and, in particular, its Corollary 9.4 below. Although Lemma 9.3 is crucial for the current proof, its own proof is quite technical. Hence, we will postpone it till the end of this section.

Invoking the result from Corollary 9.4 we get a constant C_2 , independent of k such that

$$\begin{aligned}
&n^\delta \sum_{\ell=1}^{w_n} \ell \mathbb{E} \left[\frac{1}{f_n^*(k)} \mathbb{1}_{\{f_n(k)>0\}} \mathbb{E}_n \left[\left| \sum_{i,j=1}^n X_{ij}(k, \ell) \right|^2 \right]^{1/2} \middle| \Omega_n \right] \\
&\leq C_2 n^\delta \sum_{\ell=1}^{w_n} \ell \mathbb{E} \left[\frac{1}{\sqrt{L_n}} \middle| \Omega_n \right] \leq \frac{C_2 n^{\delta-\frac{1}{2}} w_n^2}{\sqrt{\nu_1 - n^{-\varepsilon}}}.
\end{aligned}$$

Hence, there exists a constant $C_3 > 0$, independent of k such that

$$\mathbb{P} \left(\Xi_n^{(1)} > \frac{n^{-\delta}}{4}, \Omega_n \right) \leq C_3 n^{\delta+2p-\frac{1}{2}} = C_3 n^{\delta - \frac{\eta}{2(\eta+2)}}. \quad (31)$$

Combining (31) and (30) we get such that

$$\mathbb{P} \left(\left| \Phi_n(k) - \frac{\nu_2}{\nu_1} \right| > n^{-\delta}, \Omega_n \right) \leq (C_1 + C_3) n^{\delta - \frac{\eta}{2(\eta+2)}},$$

which proves (26). \square

We now proceed with the proof of Lemma 9.3 which is a stronger version of Lemma 6.3 in [15].

Lemma 9.3. *Let $X_{ij}(k, \ell)$ be as defined in (27). Then there exist a constant $C > 0$ such that, for any $k, \ell > 0$,*

$$\mathbb{E}_n \left[\left(\sum_{i,j=1}^n X_{ij}(k, \ell) \right)^2 \right] \leq \frac{C f_n^*(k)^2 f_n^*(\ell)^2}{L_n}. \quad (32)$$

Using that $f_n^*(\ell) \leq 1$, we immediately get the next corollary, which we used in the proof of Theorem 5.1.

Corollary 9.4. *Let $X_{ij}(k, \ell)$ be as defined in (27). Then there exists constant $C > 0$, such that for any $k, \ell > 0$,*

$$\mathbb{E}_n \left[\left(\sum_{i,j=1}^n X_{ij}(k, \ell) \right)^2 \right] \leq C \frac{f_n^*(k)^2}{L_n}.$$

Proof of Lemma 9.3. Note we can assume that $f_n(k) > 0$, otherwise by definition both sides of (32) equal zero. To proceed, define

$$Y_{ijst} = \mathbb{E}_n [G_{ij}G_{st}] - \mathbb{E}_n [G_{ij}] \frac{D_s D_t}{L_n} - \mathbb{E}_n [G_{st}] \frac{D_i D_j}{L_n} + \frac{D_i D_j D_s D_t}{L_n^2},$$

so that

$$\mathbb{E}_n \left(\left(\sum_{i,j=1}^n X_{ij}(k, \ell) \right)^2 \right) = \frac{1}{L_n^2} \sum_{i,j=1}^n \sum_{s,t=1}^n 1_{\{D_i=k\}} 1_{\{D_j=\ell\}} 1_{\{D_s=k\}} 1_{\{D_t=\ell\}} Y_{ijst}.$$

We will prove the result by considering the different cases for the indices i, j, s and t . In the rest of the proof we will write $i \neq j \neq s$ to denote that all indices in such inequality are pairwise distinct. Similarly, we write $i = j \neq s \neq t$ to denote all indices (i, j, s, t) , where $i = j$ and the three indices i, s, t are pairwise distinct. With this notation we define the sets

$$\begin{aligned} I_1 &= \{i, j, s, t : i \neq j \neq s \neq t\} \\ I_2 &= \{i, j, s, t : i = j \neq s \neq t\} \\ I_3 &= \{i, j, s, t : i = j = s \neq t\} \\ I_4 &= \{i, j, s, t : i = j \neq s = t\} \\ I_5 &= \{i, j, s, t : i = s \neq j \neq t\} \\ I_6 &= \{i, j, s, t : i = s \neq j = t\} \\ I_7 &= \{i, j, s, t : i = j = s = t\} \end{aligned}$$

Due to symmetry, we can omit other possible cases (e.g. the set $i = j \neq s \neq t$ is equivalent to $i \neq j \neq s = t$). Therefore we only need to consider the sets given above. Moreover, we can assume without loss of generality that $L_n \geq 4$ so that

$$L_n - 1 \geq \frac{L_n}{2}, \quad L_n - 3 \geq \frac{L_n}{4} \quad \text{and} \quad 2L_n + 3 \leq 3L_n. \quad (33)$$

Case 1: $i \neq j \neq s \neq t$

Since all indices are distinct have that

$$\mathbb{E}_n [G_{ij}G_{st}] = \frac{D_i D_j D_s D_t}{(L_n - 1)(L_n - 3)},$$

and hence

$$Y_{ijst} = \frac{(2L_n + 3)D_i D_j D_s D_t}{L_n^2 (L_n - 1)(L_n - 3)}.$$

Using the above and (33) we obtain

$$\begin{aligned}
& \frac{1}{L_n^2} \sum_{i,j,s,t \in I_1} \mathbb{1}_{\{D_i=k, D_j=\ell, D_s=k, D_t=\ell\}} Y_{ijst} \\
& \leq \frac{1}{L_n^2} \sum_{i,j,s,t=1}^n \mathbb{1}_{\{D_i=k, D_j=\ell, D_s=k, D_t=\ell\}} \frac{(2L_n + 3)D_i D_j D_s D_t}{L_n^2 (L_n - 1)(L_n - 3)} \\
& \leq 24 \sum_{i,j,s,t=1}^n \mathbb{1}_{\{D_i=k, D_j=\ell, D_s=k, D_t=\ell\}} \frac{k^2 \ell^2}{L_n^5} \\
& = 24 \frac{1}{L_n^5} \left(\sum_{i=1}^n \mathbb{1}_{\{D_i=k\}} k^2 \right) \left(\sum_{j=1}^n \mathbb{1}_{\{D_j=\ell\}} \ell^2 \right) \\
& \leq 24 \frac{1}{L_n^5} \left(\sum_{i=1}^n \mathbb{1}_{\{D_i=k\}} k \right)^2 \left(\sum_{j=1}^n \mathbb{1}_{\{D_j=\ell\}} \ell \right)^2 \\
& = 24 \frac{f_n^*(k)^2 f_n^*(\ell)^2}{L_n}.
\end{aligned}$$

Case 2: $i = j \neq s \neq t$

In this case we have

$$\mathbb{E}_n [G_{ij} G_{st}] = \frac{D_i(D_i - 1)D_s D_t}{(L_n - 1)(L_n - 3)} \quad \text{and} \quad Y_{ijst} = \frac{3D_i(D_i - 1)D_s D_t}{L_n(L_n - 1)(L_n - 3)} - \frac{D_i D_j D_s D_t}{L_n^2(L_n - 1)}.$$

In addition, since $i = j$ it follows that the sum over I_2 is non-zero if and only if $k = \ell$ and hence

$$\begin{aligned}
& \frac{1}{L_n^2} \sum_{i,j,s,t \in I_2} \mathbb{1}_{\{D_i=k, D_j=\ell, D_s=k, D_t=\ell\}} Y_{ijst} \\
& \leq \sum_{i,s,t=1}^n \mathbb{1}_{\{D_i=k, D_s=k, D_t=\ell\}} \frac{3D_i(D_i - 1)D_s D_t}{L_n^3(L_n - 1)(L_n - 3)} \mathbb{1}_{\{k=\ell\}} \\
& \leq \sum_{i,s,t=1}^n \mathbb{1}_{\{D_i=D_s=D_t=k=\ell\}} \frac{24D_i^2 D_s D_t}{L_n^5} \\
& = \sum_{i,s,t=1}^n \mathbb{1}_{\{D_i=D_s=D_t=k=\ell\}} \frac{24k^2 \ell^2}{L_n^5} \\
& \leq \frac{24f_n^*(k)^2 f_n^*(\ell)^2}{L_n} \mathbb{1}_{\{k=\ell\}}.
\end{aligned}$$

The computations for the other cases follows in a similar way, from which the result follows. \square

Remark 9.5. Note that we have only proved a pointwise convergence for every k . Unfortunately, we cannot directly extend the above method to derive a scaling for $\mathbb{P} \left(\sup_{f^*(k) > 0} \left| \Phi_n(k) - \frac{\mu_2}{\mu_1} \right| > n^{-\delta} \right)$ because the obtained upper bounds, which are based on expectations, are not sharp enough to imply such uniform convergence.

We proceed with the proof of the central limit theorem for the ANND in the configuration model with regularly-varying degree distribution. In what follows we use that if \mathcal{D} is regularly varying with exponent $1 < \gamma$, then $\mathbb{E} [\mathcal{D}^{1+\eta}] < \infty$, for $\eta = (\gamma - 1)/2$. In particular we have the a degree sequence \mathbf{D}_n generated by IID(\mathcal{D}) satisfies Assumption 5.1 with $\varepsilon \leq (\gamma - 1)/4(\gamma + 3)$.

Proof of Theorem 5.3. By the stable-law CLT, there exists a slowly-varying function $l_0(n)$ such that

$$\frac{\sum_{i=1}^n D_i^2}{l_0(n)n^{\frac{2}{\gamma}}} \xrightarrow{d} S_{\gamma/2}, \quad \text{as } n \rightarrow \infty, \quad (34)$$

where $S_{\gamma/2}$ is a $\gamma/2$ -stable random variable.

Let Ω_n be as in Assumption 5.1, with $\varepsilon = (\gamma - 1)/4(\gamma + 3)$. Define the events

$$A_{nk} = \{f_n(k) > 0\}, \quad B_n = \left\{ \sum_{i=1}^n D_i^2 \leq n^{\frac{2}{\gamma} + \frac{\varepsilon}{2}} \right\}.$$

and let $\Lambda_{nk} = A_{nk} \cap B_n \cap \Omega_n$. Then $\lim_{n \rightarrow \infty} \sup_{k \leq n^\tau} \mathbb{P}(A_{nk}^c) = 0$ by Theorem 3.1, and it follows from (34), c.f. [15, Proposition 2.5], that $\mathbb{P}(B_n) \rightarrow 1$, as $n \rightarrow \infty$. Together with Theorem 5.2 this implies that

$$\lim_{n \rightarrow \infty} \sup_{k \leq n^\tau} \mathbb{P}(\Lambda_{nk}^c) = 0. \quad (35)$$

We now split the main term into three terms as follows:

$$\begin{aligned} & \left| \mathbb{E} \left[g \left(\frac{\nu_1 \Phi_n(k)}{l_0(n)n^{\frac{2}{\gamma}-1}} \right) - g(S_{2/\gamma}) \right] \right| \leq \left| \mathbb{E} \left[g \left(\frac{\sum_{i=1}^n D_i^2}{l(n)n^{\frac{2}{\gamma}}} \right) - g(S_{2/\gamma}) \right] \right| \\ & + \left| \mathbb{E} \left[g \left(\frac{\nu_1 \sum_{\ell > 0} f_n^*(\ell)\ell}{l_0(n)n^{\frac{2}{\gamma}-1}} \right) - g \left(\frac{\sum_{i=1}^n D_i^2}{l_0(n)n^{\frac{2}{\gamma}}} \right) \right] \right| \\ & + \left| \mathbb{E} \left[g \left(\frac{\nu_1 \Phi_n(k)}{l_0(n)n^{\frac{2}{\gamma}-1}} \right) - g \left(\frac{\nu_1 \sum_{\ell > 0} f_n^*(\ell)\ell}{l_0(n)n^{\frac{2}{\gamma}-1}} \right) \right] \right| := \Xi_n^{(1)} + \Xi_n^{(2)} + \Xi_n^{(3)}, \end{aligned}$$

and we will show that all three terms converge to zero. We remark that by Potter's bounds we have $\lim_{n \rightarrow \infty} l_0(n)n^{-\delta} = 0$, for any $\delta > 0$.

It follows from (34) that

$$\lim_{n \rightarrow \infty} \sup_{k \leq n^\tau} \Xi_n^{(1)} = 0.$$

We proceed with $\Xi_n^{(2)}$. Note that on the event Ω_n we have $|L_n - \nu_1 n| \leq n^{1-\varepsilon}$. In addition,

$$\sum_{\ell > 0} f_n^*(\ell)\ell = \frac{1}{L_n} \sum_{i=1}^n D_i^2.$$

Next, since g is bounded and Lipschitz continuous, there exists $C_0 > 0$ such that

$$\begin{aligned} \Xi_n^{(2)} & \leq \frac{C_0}{l_0(n)n^{\frac{2}{\gamma}}} \mathbb{E} \left[\sum_{i=1}^n D_i^2 \left| \frac{\nu_1 n}{L_n} - 1 \right| \mathbb{1}_{\{\Lambda_{nk}\}} \right] + 2C_0 \mathbb{P}(\Lambda_{nk}^c) \\ & \leq \frac{C_0 n^{\frac{\varepsilon}{2}}}{l_0(n)} \mathbb{E} \left[\frac{|\nu_1 n - L_n|}{L_n} \mathbb{1}_{\{\Lambda_{nk}\}} \right] + 2C_0 \mathbb{P}(\Lambda_{nk}^c) \leq \frac{C_1 n^{-\frac{\varepsilon}{2}}}{l_0(n)} + 2C_0 \mathbb{P}(\Lambda_{nk}^c) \end{aligned}$$

for some constant $C_1 > 0$. The second term in the last expression converges to zero due to (35). It follows that

$$\lim_{n \rightarrow \infty} \sup_{k \leq n^\tau} \Xi_n^{(2)} = 0.$$

Finally, we turn to $\Xi_n^{(3)}$. Using again that g is bounded and Lipschitz continuous, we get

$$\Xi_n^{(3)} \leq \frac{\nu_1 C_0}{l_0(n)n^{\frac{2}{\gamma}-1}} \mathbb{E} \left[\left| \Phi_n(k) - \sum_{\ell > 0} f_n^*(\ell)\ell \right| \mathbb{1}_{\{\Lambda_n\}} \right] + 2C_0 \mathbb{P}(\Lambda_n^c).$$

Again, the last term converges to zero due to (35). For the other term, similarly to the proof of Theorem 5.1, we define

$$X_{ij}(k, \ell) = \mathbb{1}_{\{D_i=k, D_j=\ell\}} \left(\frac{G_{ij}}{L_n} - \frac{D_i D_j}{L_n^2} \right)$$

and use that on A_{nk} we have $f_n^*(k) > 0$, to write

$$\left| \Phi_n(k) - \sum_{\ell>0} f_n^*(\ell)\ell \right| = \frac{|\sum_{\ell>0} h_n(k, \ell)\ell - f_n^*(k)f_n^*(\ell)\ell|}{f_n^*(k)} \leq \frac{\sum_{\ell>0} \ell \left| \sum_{i,j=1}^n X_{ij}(k, \ell) \right|}{f_n^*(k)}.$$

Taking the expectation, conditioned on the degree sequence, and using Lemma 9.3 we then obtain

$$\begin{aligned} \mathbb{E}_n \left[\left| \Phi_n(k) - \sum_{\ell>0} f_n^*(\ell)\ell \right| \right] &\leq \frac{\sum_{\ell>0} \ell}{f_n^*(k)} \mathbb{E}_n \left[\left| \sum_{i,j=1}^n X_{ij}(k, \ell) \right| \right] \\ &\leq \frac{\sum_{\ell>0} \ell}{f_n^*(k)} \mathbb{E}_n \left[\left(\sum_{i,j=1}^n X_{ij}(k, \ell) \right)^2 \right]^{1/2} \\ &\leq \frac{C}{L_n^{1/2}} \sum_{\ell>0} f_n^*(\ell)\ell = \frac{C}{L_n^{3/2}} \sum_{i=1}^n D_i^2, \end{aligned}$$

for some constant $C > 0$. Using again that on the event Ω_n we have $|L_n - \nu_1 n| \leq n^{1-\varepsilon}$, we obtain

$$\begin{aligned} \frac{\nu_1 C_0}{l_0(n)n^{\frac{2}{\gamma}-1}} \mathbb{E} \left[\left| \Phi_n(k) - \sum_{\ell>0} f_n^*(\ell)\ell \right| \mathbb{1}_{\{\Lambda_n\}} \right] &\leq \frac{\nu_1 C_0}{l_0(n)n^{\frac{2}{\gamma}-1}} \mathbb{E} \left[\frac{C}{L_n^{3/2}} \sum_{i=1}^n D_i^2 \mathbb{1}_{\{\Lambda_{kn}\}} \right] \\ &\leq \frac{\nu_1 C_0 C n^{\frac{2}{\gamma} + \frac{\varepsilon}{2} - \frac{3}{2} - \frac{2}{\gamma} + 1}}{(\nu_1 - n^{-\varepsilon})^{3/2} l_0(n)} = \frac{\nu_1 C_0 C n^{-\frac{1-\varepsilon}{2}}}{(\nu_1 - n^{-\varepsilon})^{3/2} l_0(n)}. \end{aligned}$$

Since we chose $\varepsilon = (\gamma - 1)/4(\gamma + 3) < 1$, it follows that

$$\lim_{n \rightarrow \infty} \sup_{k \leq n^\tau} \Xi_n^{(3)} = 0,$$

which proves the last statement. \square

9.5 Erased configuration model

Denote by Y_i the number of erased stubs of vertex i and let $E_n = \sum_{i=1}^n Y_i = L_n - \widehat{L}_n$ be the total number of erased stubs, which is twice the number of erased undirected edges. We start with two technical results on the relation between the empirical densities $f_n(k)$ and $\widehat{f}_n(k)$.

Lemma 9.6. *Let \mathcal{D} be regularly varying with exponent $1 < \gamma < 2$ and $\{G_n\}_{n \geq 1}$ be generated by ECM with $\mathbf{D}_n = \text{IID}(\mathcal{D})$. Then, for any $K, \delta > 0$*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\sum_{k=0}^{\infty} \left| \widehat{f}_n(k) - f_n(k) \right| > K n^{1-\gamma+\delta} \right) = 0.$$

When $\gamma > 2$, we have

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\sum_{k=0}^{\infty} \left| \widehat{f}_n(k) - f_n(k) \right| > K n^{-1+\delta} \right) = 0.$$

Proof. We write

$$\sum_{k=0}^{\infty} \left| \widehat{f}_n(k) - f_n(k) \right| \leq \sum_{k=0}^{\infty} \frac{1}{n} \sum_{i=1}^n \left| \mathbb{1}_{\{\widehat{D}_i=k\}} - \mathbb{1}_{\{D_i=k\}} \right|$$

$$\begin{aligned}
&\leq \sum_{k=0}^{\infty} \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{Y_i > 0\}} \left(\mathbb{1}_{\{\widehat{D}_i = k\}} + \mathbb{1}_{\{D_i = k\}} \right) \\
&= \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{Y_i > 0\}} \sum_{k=0}^{\infty} \left(\mathbb{1}_{\{\widehat{D}_i = k\}} + \mathbb{1}_{\{D_i = k\}} \right) = \frac{2}{n} \sum_{i=1}^n \mathbb{1}_{\{Y_i > 0\}} \leq \frac{2E_n}{n}.
\end{aligned}$$

Therefore, using [15, Theorem 8.13], we get for $1 < \gamma < 2$,

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\sum_{k=0}^{\infty} \left| \widehat{f}_n(k) - f_n(k) \right| > Kn^{1-\gamma+\delta} \right) \leq \lim_{n \rightarrow \infty} \mathbb{P} \left(E_n > \frac{Kn^{2-\gamma+\delta}}{2} \right) = 0,$$

while for $\gamma > 2$ we have

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\sum_{k=0}^{\infty} \left| \widehat{f}_n(k) - f_n(k) \right| > Kn^{-1+\delta} \right) \leq \lim_{n \rightarrow \infty} \mathbb{P} \left(E_n > \frac{Kn^\delta}{2} \right) = 0,$$

□

In the above proof we bounded the number of nodes which had a stub removed, $\sum_{i=1}^n \mathbb{1}_{\{Y_i > 0\}}$, with the total number of erased stubs E_n . Although this is not a tight bound, it is sufficient for our results and we use this bound several times in proofs regarding the ECM. The next lemma controls the event on which the empirical degree and size-biased degree distribution are non-zero, given that the limit distribution is non-zero.

Lemma 9.7. *Let \mathcal{D} be regularly varying with exponent $\gamma > 1$ and $\{G_n\}_{n \geq 1}$ be generated by ECM with $\mathbf{D}_n = \text{IID}(\mathcal{D})$. Then, if we denote by $f(k)$ the probability density function of \mathcal{D} , and let k be such that $f(k) > 0$,*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\widehat{f}_n(k) > \frac{f(k)}{2}, f_n(k) > \frac{f(k)}{2} \right) = 1,$$

and in particular

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\widehat{f}_n(k) > 0, f_n(k) > 0 \right) = 1.$$

Proof. For the first result we show that

$$\lim_{n \rightarrow \infty} 1 - \mathbb{P} \left(\widehat{f}_n(k) > \frac{f(k)}{2}, f_n(k) > \frac{f(k)}{2} \right) = 0,$$

which also implies the second result since $f(k) > 0$. By splitting the probability we get

$$\begin{aligned}
&1 - \mathbb{P} \left(\widehat{f}_n(k) > \frac{f(k)}{2}, f_n(k) > \frac{f(k)}{2}, \Lambda_n \right) \\
&\leq \mathbb{P} \left(\widehat{f}_n(k) \leq \frac{f(k)}{2} \right) + \mathbb{P} \left(f_n(k) \leq \frac{f(k)}{2} \right) \\
&\leq \mathbb{P} \left(|\widehat{f}_n(k) - f(k)| \geq \frac{f(k)}{2} \right) + \mathbb{P} \left(|f_n(k) - f(k)| \geq \frac{f(k)}{2} \right) \\
&\leq \mathbb{P} \left(|\widehat{f}_n(k) - f_n(k)| \geq \frac{f(k)}{4} \right) + 2\mathbb{P} \left(|f_n(k) - f(k)| \geq \frac{f(k)}{4} \right).
\end{aligned}$$

Since the first probability converges to zero by Lemma 9.6 while this holds the second term by Theorem 5.2, the result follows. □

Proof of Theorem 5.6. Recall that $a = (\gamma-1)^2/(2\gamma)$ and note that $f^*(k) > 0$ implies that $f(k) > 0$ and $k \geq 1$. Let Ω_n be as in Assumption 5.1 with $\varepsilon = (\gamma-1)/(4(\gamma+3))$ and define the events

$$A_n = \left\{ \max_{1 \leq i \leq n} D_i \leq n^{\frac{1}{\gamma} + \frac{a}{4}} \right\},$$

$$B_n = \{\mathbb{E}_n[E_n] \leq n^{2-\gamma+\frac{a}{4}}\},$$

$$C_n = \left\{ \widehat{f}_n(k) > \frac{f(k)}{2}, f_n(k) > \frac{f(k)}{2} \right\}.$$

Then $\mathbb{P}(A_n^c) \rightarrow 0$, since D_i are i.i.d. samples from \mathcal{D} , while $\mathbb{P}(B_n^c) \rightarrow 0$ by Theorem 8.13 in [15]. Finally, Lemma 9.7 implies that $\mathbb{P}(C_n) \rightarrow 1$ so that if we define $\Lambda_n = A_n \cap B_n \cap C_n \cap \Omega_n$, then

$$\mathbb{P}(\Lambda_n^c) \leq \mathbb{P}(\Omega_n^c) + \mathbb{P}(A_n^c) + \mathbb{P}(B_n^c) + \mathbb{P}(C_n^c) \rightarrow 0,$$

as $n \rightarrow \infty$. Hence it is sufficient to prove the result conditioned on the event Λ_n .

By definition, and since $\widehat{f}_n^*(k), f^*(k) > 0$ on the event Λ_n , we have that

$$\widehat{\Phi}_n(k) = \frac{\sum_{i,j=1}^n \widehat{G}_{ij} \widehat{D}_j \mathbb{1}_{\{\widehat{D}_i=k\}}}{k \sum_{i=1}^n \mathbb{1}_{\{\widehat{D}_i=k\}}},$$

and we split $|\widehat{\Phi}_n(k) - \Phi_n(k)|$ in four terms as follows

$$\left| \widehat{\Phi}_n(k) - \Phi_n(k) \right| \leq \left| \frac{1}{k \sum_{i=1}^n \mathbb{1}_{\{\widehat{D}_i=k\}}} - \frac{1}{k \sum_{i=1}^n \mathbb{1}_{\{D_i=k\}}} \right| \sum_{i,j=1}^n \widehat{G}_{ij} \widehat{D}_j \mathbb{1}_{\{\widehat{D}_i=k\}} \quad (36)$$

$$+ \frac{1}{k \sum_{i=1}^n \mathbb{1}_{\{D_i=k\}}} \sum_{i,j=1}^n \widehat{G}_{ij} \widehat{D}_j \left| \mathbb{1}_{\{\widehat{D}_i=k\}} - \mathbb{1}_{\{D_i=k\}} \right| \quad (37)$$

$$+ \frac{1}{k \sum_{i=1}^n \mathbb{1}_{\{D_i=k\}}} \sum_{i,j=1}^n \widehat{G}_{ij} \mathbb{1}_{\{D_i=k\}} \left| \widehat{D}_j - D_j \right| \quad (38)$$

$$+ \frac{1}{k \sum_{i=1}^n \mathbb{1}_{\{D_i=k\}}} \sum_{i,j=1}^n D_j \mathbb{1}_{\{D_i=k\}} \left| \widehat{G}_{ij} - G_{ij} \right|. \quad (39)$$

The remainder of the proof consists of bounding each of the terms (36)–(39) by an expression of the form $cn^{\frac{1}{\gamma}+\frac{a}{4}-1}E_n$. This will give the result because for any constant $c > 0$ we have

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P} \left(cn^{\frac{1}{\gamma}+\frac{a}{4}-1}E_n > n^{\frac{2}{\gamma}-1-a}, \Lambda_n \right) &\leq \lim_{n \rightarrow \infty} cn^{\frac{5a}{4}-\frac{1}{\gamma}} \mathbb{E}[\mathbb{E}_n[E_n] | \Lambda_n] \\ &\leq \lim_{n \rightarrow \infty} cn^{\frac{3a}{2}-\gamma+2-\frac{1}{\gamma}} = 0, \end{aligned} \quad (40)$$

where the convergence to zero holds because

$$\frac{3a}{2} - \gamma + 2 - \frac{1}{\gamma} = \frac{3a}{2} - \frac{(\gamma-1)^2}{\gamma} = \frac{3a}{2} - 2a = -\frac{a}{2} < 0.$$

In order to bound (36), we first use that $|\widehat{f}_n(k) - f_n(k)| \leq E_n/n$ to obtain

$$\left| \frac{1}{k \sum_{i=1}^n \mathbb{1}_{\{\widehat{D}_i=k\}}} - \frac{1}{k \sum_{i=1}^n \mathbb{1}_{\{D_i=k\}}} \right| = \frac{|f_n(k) - \widehat{f}_n(k)|}{nk \widehat{f}_n(k) f_n(k)} \leq \frac{2E_n}{n^2 \widehat{f}_n^*(k) f_n^*(k)}.$$

Then, on Λ_n and using that $\widehat{D}_i \leq D_i$, we get

$$\begin{aligned} &\left| \frac{1}{k \sum_{i=1}^n \mathbb{1}_{\{\widehat{D}_i=k\}}} - \frac{1}{k \sum_{i=1}^n \mathbb{1}_{\{D_i=k\}}} \right| \sum_{i,j=1}^n \widehat{G}_{ij} \widehat{D}_j \mathbb{1}_{\{\widehat{D}_i=k\}} \\ &\leq \frac{2E_n n^{\frac{1}{\gamma}+\frac{a}{4}}}{n^2 k \widehat{f}_n(k) f_n(k)} \sum_{i,j=1}^n \widehat{G}_{ij} \mathbb{1}_{\{\widehat{D}_i=k\}} \end{aligned}$$

$$\begin{aligned}
&= \frac{2E_n n^{\frac{1}{\gamma} + \frac{\alpha}{4}}}{n^2 k \widehat{f}_n(k) f_n(k)} \sum_{i=1}^n \widehat{D}_i \mathbb{1}_{\{\widehat{D}_i=k\}} \\
&= \frac{2E_n n^{\frac{1}{\gamma} + \frac{\alpha}{4} - 1}}{f_n(k)} \leq \frac{4}{f(k)} n^{\frac{1}{\gamma} + \frac{\alpha}{4} - 1} E_n.
\end{aligned}$$

For (37), on the event Λ_n , using that $\widehat{G}_{ij} \leq G_{ij} \wedge 1$, we get

$$\begin{aligned}
&\frac{1}{k \sum_{i=1}^n \mathbb{1}_{\{D_i=k\}}} \sum_{i,j=1}^n \widehat{G}_{ij} D_j \left| \mathbb{1}_{\{\widehat{D}_i=k\}} - \mathbb{1}_{\{D_i=k\}} \right| \\
&\leq \frac{1}{nk f_n(k)} n^{\frac{1}{\gamma} + \frac{\alpha}{4}} \sum_{i,j=1}^n \widehat{G}_{ij} \left(\mathbb{1}_{\{\widehat{D}_i=k\}} + \mathbb{1}_{\{D_i=k\}} \right) \mathbb{1}_{\{Y_i > 0\}} \\
&\leq \frac{1}{nk f_n(k)} 2kn^{\frac{1}{\gamma} + \frac{\alpha}{4}} \sum_{i=1}^n \mathbb{1}_{\{Y_i > 0\}} \leq \frac{4}{f(k)} n^{\frac{1}{\gamma} + \frac{\alpha}{4} - 1} E_n.
\end{aligned}$$

In order to bound (38), note that $|\widehat{D}_i - D_i| = Y_i$. Hence, since $k \geq 1$, we have

$$\begin{aligned}
&\frac{1}{k \sum_{i=1}^n \mathbb{1}_{\{D_i=k\}}} \sum_{i,j=1}^n \widehat{G}_{ij} \mathbb{1}_{\{D_i=k\}} \left| \widehat{D}_j - D_j \right| = \frac{1}{nk f_n(k)} \sum_{i,j=1}^n \widehat{G}_{ij} Y_j \\
&\leq \frac{1}{nk f_n(k)} \sum_{j=1}^n D_j Y_j \leq \frac{2}{f(k)} n^{\frac{1}{\gamma} + \frac{\alpha}{4} - 1} E_n,
\end{aligned}$$

Finally, for (39), on the event Λ_n , we obtain

$$\frac{1}{k \sum_{i=1}^n \mathbb{1}_{\{D_i=k\}}} \sum_{i,j=1}^n D_j \mathbb{1}_{\{D_i=k\}} \left| \widehat{G}_{ij} - G_{ij} \right| \leq \frac{1}{nk f_n(k)} n^{\frac{1}{\gamma} + \frac{\alpha}{4}} \sum_{i=1}^n Y_i \leq \frac{2}{f(k)} n^{\frac{1}{\gamma} + \frac{\alpha}{4} - 1} E_n.$$

It follows that the sum of the terms (36)–(39) is bounded from above by $cn^{\frac{1}{\gamma} + \frac{\alpha}{4} - 1} E_n$, where

$$c = \frac{4}{f(k)} + \frac{4}{f(k)} + \frac{2}{f(k)} + \frac{2}{f(k)} = \frac{12}{f(k)}.$$

The result now follows from (40). \square

9.6 Average nearest neighbor rank

Proof of Theorem 6.1. By Assumption 4.2 there exists $\kappa > 0$ such that if

$$\Gamma_n = \left\{ \sum_{k,\ell > 0} |h_n(k, \ell) - h(k, \ell)| \mathbb{1}_{\{f_n(k) > 0\}} \leq n^{-\kappa} \right\},$$

then $\mathbb{P}(\Gamma_n) \rightarrow 1$. Now let Ω_n denote the event from Assumption 5.1. Then, if we define $\Lambda_n = \Gamma_n \cap \Omega_n$, it follows that $\mathbb{P}(\Lambda_n) \rightarrow 1$. Therefore, we only need to prove the result, conditioned on Λ_n .

By definition of $\Theta_n(k)$ and $\Theta(k)$ we have

$$\begin{aligned}
|\Theta_n(k) - \Theta(k)| \mathbb{1}_{\{f_n(k) > 0\}} &\leq \sum_{\ell > 0} h_n(k, \ell) F_n^*(\ell) \left| \frac{1}{f_n^*(k)} - \frac{1}{f^*(k)} \right| \mathbb{1}_{\{f_n(k) > 0\}} \\
&\quad + \frac{1}{f^*(k)} \left| \sum_{\ell > 0} h_n(k, \ell) F_n^*(\ell) - h(k, \ell) F^*(\ell) \right| \mathbb{1}_{\{f_n(k) > 0\}}
\end{aligned}$$

$$\begin{aligned}
&\leq \frac{\mathbb{1}_{\{f_n(k)>0\}}}{f_n^*(k)f^*(k)} \sum_{\ell>0} h_n(k,\ell)F_n^*(\ell) |f_n^*(k) - f^*(k)| \\
&+ \frac{1}{f^*(k)} \sum_{\ell>0} |h_n(k,\ell) - h(k,\ell)| F_n^*(\ell) \mathbb{1}_{\{f_n(k)>0\}} + \frac{1}{f^*(k)} \sum_{\ell>0} h(k,\ell) |F_n^*(\ell) - F^*(\ell)| \\
&:= \Xi_n^{(1)} + \Xi_n^{(2)} + \Xi_n^{(3)}.
\end{aligned}$$

We will show that

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\Xi_n^{(i)} > \frac{n^{-\delta}}{3}, \Lambda_n \right) = 0 \quad \text{for } i = 1, 2, 3.$$

The result then follows because

$$\begin{aligned}
&\mathbb{P} (|\Theta_n(k) - \Theta(k)| \mathbb{1}_{\{f_n(k)>0\}} > n^{-\delta}, \Lambda_n) \\
&\leq \mathbb{P} \left(\Xi_n^{(1)} > \frac{n^{-\delta}}{3}, \Lambda_n \right) + \mathbb{P} \left(\Xi_n^{(2)} > \frac{n^{-\delta}}{3}, \Lambda_n \right) + \mathbb{P} \left(\Xi_n^{(3)} > \frac{n^{-\delta}}{3}, \Lambda_n \right).
\end{aligned}$$

We start with $\Xi_n^{(1)}$. On the event Ω_n we have that $|f_n^*(k) - f^*(k)| \leq n^{-\varepsilon}$ and hence, since

$$\sum_{\ell>0} h_n(k,\ell)F_n^*(\ell) \leq 1,$$

we have

$$\Xi_n^{(1)} \leq \frac{\mathbb{1}_{\{f_n(k)>0\}} |f_n^*(k) - f^*(k)|}{f^*(k)(f^*(k) - n^{-\varepsilon})} \sum_{\ell>0} h_n(k,\ell)F_n^*(\ell) \leq \frac{d_{tv}(f_n^*, f^*)}{f^*(k)(f^*(k) - n^{-\varepsilon})} \leq \frac{n^{-\varepsilon}}{f^*(k)(f^*(k) - n^{-\varepsilon})}.$$

Therefore, since $\delta < \varepsilon$ it follows that

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\Xi_n^{(1)} > \frac{n^{-\delta}}{3}, \Lambda_n \right) = 0.$$

For $\Xi_n^{(2)}$ we note that

$$\Xi_n^{(2)} \leq \frac{1}{f^*(k)} \sum_{k,\ell>0} |h_n(k,\ell) - h(k,\ell)| \mathbb{1}_{\{f_n(k)>0\}}$$

and hence, since $\delta < \kappa$,

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\Xi_n^{(2)} > \frac{n^{-\delta}}{3}, \Lambda_n \right) = 0.$$

Finally, for $\Xi_n^{(3)}$ we use (1) to get

$$\Xi_n^{(3)} \leq \frac{\sup_{\ell>0} |F_n^*(\ell) - F^*(\ell)|}{f^*(k)} \sum_{\ell>0} h(k,\ell) \leq \frac{2d_{tv}(f_n^*, f^*)}{f^*(k)} \leq \frac{2n^{-\varepsilon}}{f^*(k)}$$

which implies that

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\Xi_n^{(3)} > \frac{n^{-\delta}}{3}, \Lambda_n \right) = 0.$$

□

Proof of Theorem 6.3. Recall that

$$\mathbb{E} [F^*(\mathcal{D}^*)] = \sum_{\ell>0} f^*(\ell)F^*(\ell).$$

Then we have

$$\begin{aligned}
|\Theta_n(k) - \mathbb{E}[F^*(\mathcal{D}^*)] \mathbb{1}_{\{f_n(k) > 0\}}| &= \frac{\mathbb{1}_{\{f_n(k) > 0\}}}{f_n^*(k)} \left| \sum_{\ell > 0} h_n(k, \ell) F_n^*(\ell) - f_n^*(k) f^*(\ell) F^*(\ell) \right| \\
&\leq \frac{\mathbb{1}_{\{f_n(k) > 0\}}}{f_n^*(k)} \left| \sum_{\ell > 0} (h_n(k, \ell) - f_n^*(k) f_n^*(\ell)) F_n^*(\ell) \right| \\
&\quad + \left| \sum_{\ell > 0} f_n^*(\ell) F_n^*(\ell) - f^*(\ell) F^*(\ell) \right| \\
&\leq \frac{\mathbb{1}_{\{f_n(k) > 0\}}}{f_n^*(k)} \left| \sum_{\ell > 0} (h_n(k, \ell) - f_n^*(k) f_n^*(\ell)) F_n^*(\ell) \right| \\
&\quad + \sum_{\ell > 0} |f_n^*(\ell) - f^*(\ell)| + \sum_{\ell > 0} |F_n^*(\ell) - F^*(\ell)| f^*(\ell) \\
&:= \Xi_n^{(1)} + \Xi_n^{(2)} + \Xi_n^{(3)}.
\end{aligned}$$

First, note that $\Xi_n^{(2)}, \Xi_n^{(3)}$ are independent of k and

$$\Xi_n^{(2)} = 2d_{tv}(f_n^*, f^*) \quad \text{and} \quad \Xi_n^{(3)} \leq \sup_{\ell > 0} |F_n^*(\ell) - F^*(\ell)| \leq d_{tv}(f_n^*, f^*).$$

Now let Ω_n is be defined as in Assumption 5.1 with $\varepsilon = \eta/(8 + 4\eta)$. Then, by Theorem 5.2

$$\mathbb{P}(\Omega_n^c) = O(n^{-\varepsilon}) = O(n^{-\delta}),$$

with all constants independent of k . In particular, since $\delta < \eta/(8 + 4\eta)$, this implies that

$$\sup_{k > 0} \mathbb{P}\left(\Xi_n^{(2)} + \Xi_n^{(3)} > \frac{n^{-\delta}}{2}\right) \leq 2\mathbb{P}\left(d_{tv}(f_n^*, f^*) > \frac{n^{-\delta}}{4}\right) = O(n^{-\delta}).$$

For $\Xi_n^{(1)}$ we will show that

$$\sup_{k > 0} \mathbb{P}\left(\Xi_n^{(1)} > \frac{n^{-\delta}}{2}, \Omega_n\right) = O(n^{\delta - \frac{1}{2}}) = O(n^{-\delta}),$$

since $\delta < \eta/(8 + 4\eta) < 1/4$. For this we will use the same approach as for the proof of Theorem 5.1. That is, we split the summation into the regimes $\ell \leq w_n$ and $\ell > w_n$, with $w_n = \lfloor n^p \rfloor$, and invoke Lemma 9.3. Recall the definition of $X_{ij}(k, \ell)$,

$$X_{ij}(k, \ell) = \mathbb{1}_{\{D_i=k, D_j=\ell\}} \left(\frac{G_{ij}}{L_n} - \frac{D_i D_j}{L_n^2} \right).$$

Then, since $f_n^*(\ell) \leq 1$,

$$\begin{aligned}
\Xi_n^{(1)} &\leq \frac{\mathbb{1}_{\{f_n(k) > 0\}}}{f_n^*(k)} \left| \sum_{\ell > 0} h_n(k, \ell) - f_n^*(k) f_n^*(\ell) \right| \\
&\leq \frac{\mathbb{1}_{\{f_n(k) > 0\}}}{f_n^*(k)} \sum_{\ell \leq w_n} \left| \sum_{i,j=1}^n X_{ij}(k, \ell) \right| + \frac{\mathbb{1}_{\{f_n(k) > 0\}}}{f_n^*(k)} \sum_{\ell > w_n} \left| \sum_{i,j=1}^n X_{ij}(k, \ell) \right|.
\end{aligned}$$

Taking the conditional expectation and using Lemma 9.3, in a similar way as in the proof of Theorem 5.1, we obtain

$$\mathbb{E}_n \left[\Xi_n^{(1)} \right] \leq C \sum_{\ell \leq w_n} \frac{f_n^*(\ell)}{L_n^{1/2}} + 2 \sum_{\ell > w_n} \sum_{i=1}^n \mathbb{1}_{\{D_i=\ell\}} \frac{D_i}{L_n}$$

$$\begin{aligned}
&= \frac{C}{L_n^{3/2}} \sum_{i=1}^n D_i \mathbb{1}_{\{D_i \leq w_n\}} + \frac{2}{L_n} \sum_{i=1}^n D_i \mathbb{1}_{\{D_i > w_n\}} \\
&\leq \frac{C}{L_n^{1/2}} + \frac{2}{L_n} \sum_{i=1}^n D_i \mathbb{1}_{\{D_i > w_n\}},
\end{aligned}$$

for some constant $C > 0$, which is independent from k . On the event Ω_n we have that $L_n > \nu_1 n - n^{1-\varepsilon}$, while

$$\mathbb{E} [\mathcal{D} \mathbb{1}_{\{\mathcal{D} > w_n\}}] \leq w_n^{-\eta} \mathbb{E} [\mathcal{D}^{1+\eta}] = O(n^{-\eta p}).$$

Therefore we get

$$\begin{aligned}
\mathbb{P} \left(\Xi_n^{(1)} > \frac{n^{-\delta}}{2}, \Omega_n \right) &\leq 2n^\delta \mathbb{E} \left[\Xi_n^{(1)} \middle| \Omega_n \right] \leq \frac{2n^\delta}{\mathbb{P}(\Omega_n)} \mathbb{E} \left[\mathbb{E}_n \left[\Xi_n^{(1)} \right] \mathbb{1}_{\{\Omega_n\}} \right] \\
&\leq \frac{2Cn^\delta}{(\nu_1 n - n^{-\varepsilon})^{1/2} \mathbb{P}(\Omega_n)} + \frac{4n^\delta}{(\nu_1 - n^{-\varepsilon}) \mathbb{P}(\Omega_n)} \mathbb{E} [\mathcal{D} \mathbb{1}_{\{\mathcal{D} > w_n\}}] \\
&= O \left(n^{\delta - \frac{1}{2}} + n^{\delta - \eta p} \right),
\end{aligned}$$

and hence, by taking $p = 2\delta/\eta$, we obtain that for any $0 < \delta < \eta/(8 + 4\eta)$,

$$\sup_{k>0} \mathbb{P} (|\Theta_n(k) - \mathbb{E}[F^*(\mathcal{D}^*)]| > n^{-\delta}) = O \left(n^{-\delta/2} + n^{\delta - \frac{1}{2}} + n^{-\delta} \right) = O \left(n^{\delta - \frac{1}{2}} \right).$$

□

Proof of Theorem 6.4. Let Ω_n be defined as in Assumption 5.1 for some $\varepsilon > 0$ and define the events

$$A_n = \left\{ \mathbb{E}_n [E_n] \leq n^{\frac{3-\gamma}{2}} \right\}, \quad B_n = \left\{ \hat{f}_n(k) > \frac{f(k)}{2}, f_n(k) > \frac{f(k)}{2} \right\}.$$

Then, since $(3 - \gamma)/2 > 2 - \gamma$, $\mathbb{P}(A_n) \rightarrow 1$ by [15, Theorem 8.13], while $\mathbb{P}(B_n) \rightarrow 1$ by Lemma 9.7. Hence, if we define $\Lambda_n = A_n \cap B_n \cap \Omega_n$, then $\mathbb{P}(\Lambda_n) \rightarrow 1$, as $n \rightarrow \infty$ so that it is enough to prove the result conditioned on the event Λ_n .

By definition we have that

$$\hat{\Theta}_n(k) = \frac{\sum_{i,j=1}^n \hat{G}_{ij} \hat{F}_n^*(\hat{D}_j) \mathbb{1}_{\{\hat{D}_i=k\}}}{k \sum_{i=1}^n \mathbb{1}_{\{\hat{D}_i=k\}}} \mathbb{1}_{\{\hat{f}_n(k) > 0\}},$$

where the hats denote expressions for the erased model. As in the proof of Theorem 5.6, we split $|\hat{\Theta}_n(k) - \Theta_n(k)|$ in four terms

$$\left| \hat{\Theta}_n(k) - \Theta_n(k) \right| \leq \left| \frac{\mathbb{1}_{\{\hat{f}_n(k) > 0\}}}{k \sum_{i=1}^n \mathbb{1}_{\{\hat{D}_i=k\}}} - \frac{\mathbb{1}_{\{f_n(k) > 0\}}}{k \sum_{i=1}^n \mathbb{1}_{\{D_i=k\}}} \right| \sum_{i,j=1}^n \hat{G}_{ij} \hat{F}_n^*(\hat{D}_j) \mathbb{1}_{\{\hat{D}_i=k\}} \quad (41)$$

$$+ \frac{\mathbb{1}_{\{f_n(k) > 0\}}}{k \sum_{i=1}^n \mathbb{1}_{\{D_i=k\}}} \sum_{i,j=1}^n \hat{G}_{ij} \hat{F}_n^*(\hat{D}_j) \left| \mathbb{1}_{\{\hat{D}_i=k\}} - \mathbb{1}_{\{D_i=k\}} \right| \quad (42)$$

$$+ \frac{\mathbb{1}_{\{f_n(k) > 0\}}}{k \sum_{i=1}^n \mathbb{1}_{\{D_i=k\}}} \sum_{i,j=1}^n F_n^*(D_j) \mathbb{1}_{\{D_i=k\}} \left| \hat{G}_{ij} - G_{ij} \right| \quad (43)$$

$$+ \frac{\mathbb{1}_{\{f_n(k) > 0\}}}{k \sum_{i=1}^n \mathbb{1}_{\{D_i=k\}}} \sum_{i,j=1}^n \hat{G}_{ij} \mathbb{1}_{\{D_i=k\}} \left| \hat{F}_n^*(\hat{D}_j) - F_n^*(D_j) \right| \quad (44)$$

and bound all four terms individually.

First, we bound the terms (41)–(43) by the expressions of the form $\frac{CE_n}{f(k)n}$ for some $C > 0$ and use that, for any $c > 0$ and $f(k) > 0$,

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{E_n}{n} > cf(k), \Lambda_n \right) \leq \lim_{n \rightarrow \infty} \frac{1}{cf(k)n} \mathbb{E} [\mathbb{E}_n [E_n] | A_n] = \lim_{n \rightarrow \infty} \frac{n^{-\frac{\gamma-1}{2}}}{cf(k)} = 0. \quad (45)$$

For (41), on Λ_n we have $\widehat{f}_n(k), f_n(k) > 0$ while $|f_n(k) - \widehat{f}_n(k)| \leq E_n/n$. Therefore we obtain

$$\left| \frac{1}{k \sum_{i=1}^n \mathbb{1}_{\{\widehat{D}_i=k\}}} - \frac{1}{k \sum_{i=1}^n \mathbb{1}_{\{D_i=k\}}} \right| = \frac{|f_n(k) - \widehat{f}_n(k)|}{nk \widehat{f}_n(k) f_n(k)} \leq \frac{E_n}{n^2 k \widehat{f}_n(k) f_n(k)}.$$

From this, and $\widehat{F}_n^* \leq 1$, it follows that

$$\begin{aligned} & \left| \frac{1}{k \sum_{i=1}^n \mathbb{1}_{\{\widehat{D}_i=k\}}} - \frac{1}{k \sum_{i=1}^n \mathbb{1}_{\{D_i=k\}}} \right| \sum_{i,j=1}^n \widehat{G}_{ij} \widehat{F}_n^*(\widehat{D}_j) \mathbb{1}_{\{\widehat{D}_i=k\}} \\ & \leq \frac{E_n}{n^2 k \widehat{f}_n(k) f_n(k)} \sum_{i,j=1}^n \widehat{G}_{ij} \mathbb{1}_{\{\widehat{D}_i=k\}} \\ & \leq \frac{E_n}{n^2 k \widehat{f}_n(k) f_n(k)} \sum_{i=1}^n \widehat{D}_i \mathbb{1}_{\{\widehat{D}_i=k\}} \\ & \leq \frac{E_n}{nf_n(k)} \leq \frac{2E_n}{nf(k)}. \end{aligned}$$

Continuing with (42), on the event Λ_n , using that $\widehat{G}_{ij} \leq G_{ij} \wedge 1$, we obtain

$$\begin{aligned} & \frac{1}{k \sum_{i=1}^n \mathbb{1}_{\{D_i=k\}}} \sum_{i,j=1}^n \widehat{G}_{ij} F_n^*(D_j) \left| \mathbb{1}_{\{\widehat{D}_i=k\}} - \mathbb{1}_{\{D_i=k\}} \right| \\ & \leq \frac{1}{kn f_n(k)} \sum_{i,j=1}^n \widehat{G}_{ij} \left(\mathbb{1}_{\{\widehat{D}_i=k\}} + \mathbb{1}_{\{D_i=k\}} \right) \mathbb{1}_{\{Y_i > 0\}} \\ & \leq \frac{2k}{kn f_n(k)} \sum_{i=1}^n \mathbb{1}_{\{Y_i > 0\}} \leq \frac{4E_n}{nf(k)}. \end{aligned}$$

For (43), we get, on the event Λ_n ,

$$\frac{1}{k \sum_{i=1}^n \mathbb{1}_{\{D_i=k\}}} \sum_{i,j=1}^n F_n^*(D_j) \mathbb{1}_{\{D_i=k\}} \left| \widehat{G}_{ij} - G_{ij} \right| \leq \frac{1}{nf_n(k)} \sum_{i=1}^n Y_i = \frac{E_n}{nf_n(k)} \leq \frac{2E_n}{nf(k)}.$$

Now, using (45), we obtain that

$$\lim_{n \rightarrow \infty} \mathbb{P} \left((41) + (42) + (43) > \frac{\delta}{2}, \Lambda_n \right) \leq \lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{E_n}{n} > \frac{\delta}{16} f(k), \Lambda_n \right) = 0.$$

It remains to bound (44). This requires a bit more work. First we split $|\widehat{F}_n^*(\widehat{D}_j) - F_n^*(D_j)|$ as

$$\left| \widehat{F}_n^*(\widehat{D}_j) - F_n^*(D_j) \right| \leq \left| \widehat{F}_n^*(\widehat{D}_j) - F_n^*(\widehat{D}_j) \right| + \left| F_n^*(\widehat{D}_j) - F_n^*(D_j) \right|$$

For the last term we have

$$\left| F_n^*(\widehat{D}_j) - F_n^*(D_j) \right| \leq \frac{1}{L_n} \sum_{i=1}^n D_i \left| \mathbb{1}_{\{D_i \leq \widehat{D}_j\}} - \mathbb{1}_{\{D_i \leq D_j\}} \right|$$

$$\leq \frac{2}{L_n} \sum_{i=1}^n D_i \mathbb{1}_{\{D_i \leq D_j\}} \mathbb{1}_{\{Y_j > 0\}} = 2F_n^*(D_j) \mathbb{1}_{\{Y_j > 0\}}.$$

For the other term we have

$$\begin{aligned} \left| \widehat{F}_n^*(\widehat{D}_j) - F_n^*(\widehat{D}_j) \right| &\leq \left| \frac{1}{\widehat{L}_n} - \frac{1}{L_n} \right| \left| \sum_{i=1}^n \widehat{D}_i \mathbb{1}_{\{\widehat{D}_i \leq \widehat{D}_j\}} + \frac{1}{L_n} \sum_{i=1}^n \left| \widehat{D}_i \mathbb{1}_{\{\widehat{D}_i \leq \widehat{D}_j\}} - D_i \mathbb{1}_{\{D_i \leq \widehat{D}_j\}} \right| \right| \\ &\leq \frac{E_n}{L_n} \widehat{F}_n^*(\widehat{D}_j) + \frac{1}{L_n} \sum_{i=1}^n Y_i \mathbb{1}_{\{\widehat{D}_i \leq \widehat{D}_j\}} + \frac{1}{L_n} \sum_{i=1}^n D_i \left| \mathbb{1}_{\{\widehat{D}_i \leq \widehat{D}_j\}} - \mathbb{1}_{\{D_i \leq \widehat{D}_j\}} \right| \\ &\leq \frac{2E_n}{L_n} + \frac{2}{L_n} \sum_{i=1}^n D_i \mathbb{1}_{\{Y_i > 0\}}. \end{aligned}$$

To summarize, we have

$$\left| \widehat{F}_n^*(\widehat{D}_j) - F_n^*(D_j) \right| \leq \frac{2E_n}{L_n} + 2F_n^*(D_j) \mathbb{1}_{\{Y_j > 0\}} + \frac{2}{L_n} \sum_{i=1}^n D_i \mathbb{1}_{\{Y_i > 0\}}.$$

Next we note that

$$\frac{1}{nkf_n(k)} \sum_{i,j=1}^n \widehat{G}_{ij} \mathbb{1}_{\{D_i=k\}} \leq \frac{1}{nkf_n(k)} \sum_{i=1}^n D_i \mathbb{1}_{\{D_i=k\}} = 1.$$

Hence, on the event Λ_n ,

$$\begin{aligned} &\frac{1}{k \sum_{i=1}^n \mathbb{1}_{\{D_i=k\}}} \sum_{i,j=1}^n \widehat{G}_{ij} \mathbb{1}_{\{D_i=k\}} \left| \widehat{F}_n^*(\widehat{D}_j) - F_n^*(D_j) \right| \\ &\leq \frac{2E_n}{L_n} + \frac{1}{L_n} \sum_{i=1}^n D_i \mathbb{1}_{\{Y_i > 0\}} + \frac{2}{nkf_n(k)} \sum_{i,j=1}^n \widehat{G}_{ij} F_n^*(D_j) \mathbb{1}_{\{Y_j > 0\}} \mathbb{1}_{\{D_i=k\}} \\ &\leq \frac{2E_n}{(\nu_1 n - n^{1-\varepsilon})} + \frac{1}{(\nu_1 n - n^{1-\varepsilon})} \sum_{i=1}^n D_i \mathbb{1}_{\{Y_i > 0\}} + \frac{4}{nf(k)} \sum_{j=1}^n D_j \mathbb{1}_{\{Y_j > 0\}} \\ &= \frac{2E_n}{(\nu_1 n - n^{1-\varepsilon})} + \left(\frac{1}{(\nu_1 n - n^{1-\varepsilon})} + \frac{4}{nf(k)} \right) \sum_{j=1}^n D_j \mathbb{1}_{\{Y_j > 0\}}. \end{aligned} \tag{46}$$

Using computations similar to those leading up to (45) we get

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{2E_n}{(\nu_1 n - n^{1-\varepsilon})} > \frac{\delta}{4}, \Lambda_n \right) = 0.$$

For the last term in (46), we note that the terms in front of the summation are $O(n^{-1})$. Hence it suffices to prove that for any $c > 0$

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{1}{n} \sum_{j=1}^n D_j \mathbb{1}_{\{Y_j > 0\}} > c \right) = 0. \tag{47}$$

For this we define $p = \frac{\gamma+1}{2} < \gamma$ and $q = p/(p-1)$. Then by first applying Markov's inequality and then Hölder's inequality we get

$$\mathbb{P} \left(\frac{1}{n} \sum_{j=1}^n D_j \mathbb{1}_{\{Y_j > 0\}} > c \right) \leq \frac{1}{c} \mathbb{E} [|D_1 \mathbb{1}_{\{Y_1 > 0\}}|] \leq \frac{\mathbb{E} [D_1^p]^{1/p}}{c} \mathbb{P}(Y_1 > 0)^{1/q}.$$

To finish the argument we use [9, Lemma 5.3], which states that the probability that the number of erased out-bound and in-bound stubs in the directed configuration model are positive, converges to zero. Their argument has a straightforward extension to the undirected case so that $\lim_{n \rightarrow \infty} \mathbb{P}(Y_1 > 0) = 0$, which implies (47) and hence

$$\lim_{n \rightarrow \infty} \mathbb{P} \left((44) > \frac{\delta}{2}, \Lambda_n \right) = 0,$$

which finishes the proof. □

References

- [1] Albert-László Barabási. *Network science*. Cambridge University Press, 2016.
- [2] Marc Barthélemy, Alain Barrat, Romualdo Pastor-Satorras, and Alessandro Vespignani. Dynamical patterns of epidemic outbreaks in complex heterogeneous networks. *Journal of theoretical biology*, 235(2):275–288, 2005.
- [3] Nicholas H Bingham, Charles M Goldie, and Jef L Teugels. *Regular variation*, volume 27. Cambridge university press, 1989.
- [4] Marián Boguná, Romualdo Pastor-Satorras, and Alessandro Vespignani. Absence of epidemic threshold in scale-free networks with degree correlations. *Physical review letters*, 90(2):028701, 2003.
- [5] Marián Boguná, Romualdo Pastor-Satorras, and Alessandro Vespignani. Epidemic spreading in complex networks with degree correlations. In *Statistical mechanics of complex networks*, pages 127–147. Springer, 2003.
- [6] Marián Boguná and Romualdo Pastor-Satorras. Epidemic spreading in correlated complex networks. *Physical Review E*, 66(4):047104, 2002.
- [7] Béla Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics*, 1(4):311–316, 1980.
- [8] Michele Catanzaro, Marián Boguná, and Romualdo Pastor-Satorras. Generation of uncorrelated random scale-free networks. *Physical Review E*, 71(2):27103, 2005.
- [9] Ningyuan Chen and Mariana Olvera-Cravioto. Directed random graphs with given degree distributions. *Stochastic Systems*, 3(1):147–186, 2013.
- [10] Ningyuan Chen and Mariana Olvera-Cravioto. Efficient simulation for branching linear recursions. In *Proceedings of the 2015 Winter Simulation Conference*, pages 1–12, Huntington Beach, CA, 2015.
- [11] Gregorio D’Agostino, Antonio Scala, Vinko Zlatić, and Guido Caldarelli. Robustness and assortativity for diffusion-like processes in scale-free networks. *EPL (Europhysics Letters)*, 97(6):68006, 2012.
- [12] Sebastiano de Franciscis, Samuel Johnson, and Joaquín J. Torres. Enhancing neural-network performance via assortativity. *Physical Review E*, 83(3):036114, 2011.
- [13] Remco van der Hofstad. *Random graphs and complex networks*, volume 1. Cambridge University Press, 2016.
- [14] Remco van der Hofstad and Nelly Litvak. Degree-degree dependencies in random graphs with heavy-tailed degrees. *Internet mathematics*, 10(3-4):287–334, 2014.

- [15] Pim van der Hoorn. *Asymptotic analysis of network structures: degree-degree correlations and directed paths*. PhD thesis, University of Twente, 2016.
- [16] Pim van der Hoorn and Nelly Litvak. Convergence of rank based degree-degree correlations in random directed networks. *Moscow Journal of Combinatorics and Number Theory*, 4(4):45–83, 2014.
- [17] Pim van der Hoorn and Nelly Litvak. Degree-Degree Dependencies in Directed Networks with Heavy-Tailed Degrees. *Internet Mathematics*, 11(2):155–179, 2015.
- [18] Pim van der Hoorn and Nelly Litvak. Phase transitions for scaling of structural correlations in directed networks. *Physical Review E*, 92(2):022803, 2015.
- [19] Pim van der Hoorn, Liudmila Ostroumova Prokhorenkova, and Egor Samosvat. Generating maximally disassortative graphs with given degree distribution. *arXiv preprint arXiv:1607.01742*, 2016.
- [20] Nelly Litvak and Remco van der Hofstad. Uncovering disassortativity in large scale-free networks. *Physical Review E*, 87(2):22801, 2013.
- [21] Michael Molloy and Bruce Reed. A critical point for random graphs with a given degree sequence. *Random structures & algorithms*, 6(2-3):161–180, 1995.
- [22] Mark E J Newman. Assortative mixing in networks. *Physical review letters*, 89(20):208701, 2002.
- [23] Mark E J Newman. Mixing patterns in networks. *Physical Review E*, 67(2):26126, 2003.
- [24] Rogier Noldus and Piet Van Mieghem. Assortativity in complex networks. *Journal of Complex Networks*, page cnv005, 2015.
- [25] Romualdo Pastor-Satorras, Alexei Vázquez, and Alessandro Vespignani. Dynamical and correlation properties of the internet. *Physical review letters*, 87(25):258701, 2001.
- [26] Christian Schmeltzer, Alexandre Hiroaki Kihara, Igor Michailovitsch Sokolov, and Sten Rüdiger. Degree correlations optimize neuronal network sensitivity to sub-threshold stimuli. *PloS one*, 10(6):e0121794, 2015.
- [27] Alexei Vázquez and Yamir Moreno. Resilience to damage of graphs with degree correlations. *Phys. Rev. E*, 67:015101, Jan 2003.
- [28] Ward Whitt. *Stochastic-Process Limits*. Springer New York, 2006.
- [29] Nicholas C Wormald. Some problems in the enumeration of labelled graphs. *Bulletin of the Australian Mathematical Society*, 21(01):159–160, 1980.
- [30] Di Zhou, H Eugene Stanley, Gregorio D’Agostino, and Antonio Scala. Assortativity decreases the robustness of interdependent networks. *Physical Review E*, 86(6):066103, 2012.