Universität Osnabrück

Dissertation zur Erlangung des Doktorgrades (Dr. rer. nat.)
des Fachbereichs Mathematik/Informatik
der Universität Osnabrück

February 2019

Limit theorems in preferential attachment random graphs

vorgelegt von

Carina Betken

Acknowledgments

First of all I wish to thank my supervisor Hanna Döring for giving me the opportunity to write this thesis, for having confidence in me and for being a constant source of motivation and support. I thank Marcel Ortgiese for the productive and inspiring collaboration especially during my time in Bath. I highly appreciate the support of the *Deutsche Forschungsgmeinschaft* via *Research Training Group 1916-Combinatorial Structures in Geometry* for the financial aids which made this research stay possible. I thank Adrian Röllin for the appraisal of this thesis.

I wish to thank all my current and former colleagues in Osnabrück for the numerous mathematical and non-mathematical talks during the coffee breaks and the pleasant atmosphere at the institute which no doubt contributed to the success of my PhD studies. I am thankful to all my friends for their support and encouragement whenever needed. Among these, I am deeply grateful to Stephan Bussmann and Tobias Soethe for their careful proofreading of this manuscript.
I owe the deepest gratitude to my parents, grandparents and siblings for countless reasons: Thank you for your encouragement through all the years. Finally, I thank Paul for his exceptional love, patience and support.

Contents

1 Introduction **1**

2 Preliminaries **4**

2.1 Stein's method 4

2.1.1 Normal approximation via Stein's method and size-bias coupling 5

2.1.2 Stein's method in a nutshell 10

2.1.3 The generator approach 11

2.2 Coupling 13

3 Degree evolutions in preferential attachment models via Stein's method **17**

3.1 Preferential attachment models 17

3.2 Stein's method for preferential attachment models 22

3.2.1 Stein-operator 22

3.2.2 Stein solution and bounds 25

3.2.3 Results for general preferential attachment models without loops 29

3.2.4 Results for the Barabási-Albert model 39

3.3 Rates of convergence for the out-degree 49

4 Rates of convergence via coupling **53**

4.1 Coupling for general models 53

4.2 Coupling with the Barabási-Albert graph 74

5 A central limit theorem for the number of isolated vertices **83**

5.1 Size-bias construction 84

5.2 Proof of Theorem 5.1 87

List of notations **131**

List of Figures **131**

Bibliography **132**

1. Introduction

In many structures in science and nature one can observe components interacting with one another, e.g. molecules in metabolisms, agents in technological systems or people in social networks. Such structures can be modeled and analyzed with the help of random networks, where components are considered as nodes and edges represent relations between them. [Hof17] provides an excellent overview over this research field. One of the first and probably most studied models of random graphs is the Erdös-Rényi graph, which was introduced in [ER59] at the end of the 1950s. In this graph model each edge exists independent of all others with fixed probability p. However, it does not explain the structures observed in many real world networks such as the World Wide Web, social interaction or biological neural networks (see [HK18], [G$^+$10] and [C$^+$06] respectively). More precisely, such networks exhibit so-called power-laws as degree distributions, i.e. one finds

$$\mathbb{P}(\deg(v) = k) \propto k^{-\gamma},$$

where $\deg(v)$ denotes the degree of some vertex v in the network and γ is a constant which typically lies in the interval $(2, 3)$. The principle of *preferential attachment* has become a well-known concept to explain the occurrence of such scale-invariant distributions reasonably well. The two typical characteristics of preferential attachment models are that they are dynamic in the sense that vertices are successively added over time and that new vertices prefer to connect to those of the older vertices, which are already well connected in the existing network. Due to the latter, preferential attachment is also referred to as the *rich-get-richer paradigm*.
The construction rules for the network can be made precise in various ways, so that starting with the pioneering work [BA99] of Barabási and Albert various different models of preferential attachment random graphs have appeared in the scientific literature in recent years. In [KR01] and [DM09] the authors suggest to look at rather general models, in which the probability of attaching a new vertex to a current one is a function f of its degree, called the attachment rule. The choice of this function has a huge impact on the qualitative features of the network. One can distinguish three main regimes for f: the superlinear regime, where $f(k) \gg k$, the linear regime, where $f(k) \approx k$, and the sublinear regime, where $f(k) \ll k$. In the superlinear case there emerges a dominant vertex which attracts most of the

edges, so that after the insertion of n vertices the degree of that vertex is of order n. For more details see for instance [OS05]. The behaviour for linear attachment functions is probably the most studied case. This is not very surprising as the first mathematical rigorous work on preferential attachment random graphs by Barabási and Albert was concerned with such networks. Beyond that, the dynamics of the preferential attachment model in this regime give a plausible explanation for the occurrence of power-law distributions which, as mentioned before, is a striking feature of many real-world networks. The linear along with the sublinear regime were for example studied in [DM09] and [RTV07]. In both papers almost sure convergence of the empirical degree distribution towards a power-law distribution in the linear and stretched exponential distributions in the sublinear regime, is shown for slightly different models. The results in these works are of asymptotic nature, thus holding for the number of vertices tending to infinity. However, real-world networks are always of finite size, though indeed often rather huge. This is the reason why we are interested in the distributional distance between the empirical distributions of different random quantities and the known limiting distributions in the presence of n vertices. For the linear case with fixed outdegree [PRR13] and [PRR17] study the rescaled degree of a finite number of fixed vertices and in [Ros13] the author shows that empirical indegree distribution converges to a mixed binomial distribution. All three works deduce rates of convergence, where [Ros13] and [PRR13] make use of Stein's method. We complement these results by deducing error bounds on the distributional distance between the indegree distribution and the limiting distribution provided by [DM09] by a new variation of Stein's method in a more general model, including the sublinear as well as the linear case and comprising models with fixed as well as random outdegree.

The aim of this thesis is to deduce error bounds on the distributional distances between the laws of three random quantities in a graph on n vertices and their respective limiting distributions. The random quantities we consider are the indegree of a uniformly chosen vertex, the outdegree of a vertex and the rescaled number of isolated vertices. The manuscript is structured as follows: chapter 2 introduces the two main methods of proof used in this thesis: Stein's method and coupling. In chapter 3 we first of all introduce a general preferential attachment model which includes the sublinear as well as the linear case and comprises models with fixed as well as random outdegree. In section 3.2 we derive error bounds on the distributional distance between the indegree of a uniformly chosen vertex and a given limiting distribution μ. In particular we deal with the model introduced by Barabási and Albert. To do so, we develop a new variation of Stein's method for a new class of limiting distributions. Therefore, we will use the fact that the limiting distribution is the stationary distribution of a Markov chain together with the generator method of Barbour. Our principal interest concerns the asymptotic evolution of the indegree distribution, since the outdegree of every vertex is fixed after the time step in which it was inserted into the network. However, in Theorem 3.20 we also give a limit result for the outdegree distribution in a preferential attachment model with random outdegree using well-developed results on Poisson-approximation via Stein's method. In chapter 4 we derive similar results as in the preceding chapter but through the use of coupling techniques this time. In section 4.1 we take advantage of the fact that the discrete-time Markov chain introduced in chapter 3 resembles the dynamics of a continuous-time Markov chain with generator given by the Stein operator deduced before. In the following section 4.2 we then couple the general preferential attachment model introduced in chapter 3 to the Barabási-Albert model, which likewise allows us to deduce rates of convergence for this model with the help of coupling. Chapter 5 shows that for some class of attachment functions, the number of isolated vertices asymptotially follows a standard normal distribution. Here we use Stein's method for approximation by a standard normal distribution in combination with size-bias coupling.

2. Preliminaries

The following two sections provide an overview of the concepts and methods needed to follow the thesis at hand. However, the reader is assumed to have basic knowledge in probability and measure theory, as the fundamental notions and results of these fields will not be recalled.

2.1. Stein's method

Stein's method is a well-known tool to derive error bounds between the law of a random variable of interest and a known and a known target distribution, which is usually better understood. It was first developed for the approximation of sums of dependent random variables by the normal distribution in [Ste72] and has been adopted to various other target distributions, including Poisson in [Che75] and [BH84], geometric in [Pek96], negative binomial in [BP99] and [Ros13], exponential in [PR11] and [CFR11], and many more.

When bounding the mentioned error through the help of Stein's method one converts the original problem of bounding the distributional distance of the two random variables under consideration into a problem of bounding the expectation of some functional operator applied to the random variable whose distribution is to be approximated. Finding this operator and linking it to a notion of distributional distance (by solving Stein equation) solely depends on the known limiting distribution, whereas bounding the expectation afterwards only requires information on the random variable of interest (though information on the target distribution is of course contained in the operator). There is a collection of techniques both for the first and the second part of this setup. We will present one technique for each of the two steps: in section 2.1.3 we introduce the so-called generator approach which we will use to link the total-variation distance to the mentioned expectation in chapter 3, and the concept of size-bias coupling to bound the expectation in the case of normal approximation in chapter 5.

Section 2.1.1 introduces Stein's method for approximation by a standard normal distribution, section 2.1.2 gives the general setup for arbitrary limiting distributions.

2.1.1. Normal approximation via Stein's method and size-bias coupling

As mentioned before, Stein's method was developed as a tool to prove central limit theorems for sums of dependent random variables and was first published in 1972 in [Ste72]. The crucial observation lying at the heart of this method is the following lemma:

Lemma 2.1 (Stein's lemma). *Define the functional operator*

$$\mathcal{A}f = f'(x) - xf(x).$$

(i) For $Z \sim \mathcal{N}(0,1)$ we have

$$\mathbb{E}\left[\mathcal{A}f(Z)\right] = 0$$

for all absolutely continuous functions f such that $\mathbb{E}\left[f'(Z)\right] < \infty$.

(ii) If X is a random variable such that for all continuously differentiable functions f

$$\mathbb{E}\left[\mathcal{A}f(X)\right] = 0,$$

then $X \sim \mathcal{N}(0,1)$.

Before we prove this lemma we show how this characterization can be used to determine error bounds in the Kolmogorov metric for the approximation of the law of a random variable by the normal distribution. To do so, we have to solve the so-called *Stein equation*

$$wf(w) - f'(w) = h(w) - \Phi(z) \tag{2.1}$$

for every $f \in \mathcal{F}_K = \{\mathbb{1}_{(-\infty,z]}, z \in \mathbb{R}\}$, where Φ denotes the cumulative distribution function of $Z \sim \mathcal{N}(0,1)$. Having solved (2.1) then yields

$$d_K(X,Z) = \sup_{h \in \mathcal{F}_K} \left|\mathbb{E}\left[h(X)\right] - \mathbb{E}\left[h(Z)\right]\right| = \sup_{z \in \mathbb{R}} \mathbb{E}\left[\mathcal{A}f_z(X)\right].$$

Here, f_z denotes the unique bounded solution of (2.1) for the test function $\mathbb{1}\{\cdot \in (-\infty, z]\}$.

Lemma 2.2 (Properties of the solution). *For fixed $z \in \mathbb{R}$ the unique bounded solution $f_z(w)$ of the equation*

$$wf(w) - f'(w) = \mathbb{1}\{w \le z\} - \Phi(Z)$$

is given by

$$f_z(w) = \begin{cases} \sqrt{2\pi} e^{w^2/2} \Phi(w) \left(1 - \Phi(z)\right) & \text{if } w \le z, \\ \sqrt{2\pi} e^{w^2/2} \Phi(z) \left(1 - \Phi(w)\right) & \text{if } w > z. \end{cases} \tag{2.2}$$

Moreover

$$\|f_z\| \le \sqrt{\frac{\pi}{2}} \text{ and } ||f_z'|| \le 2. \tag{2.3}$$

Proof. To show that the solution in (2.2) solves the Stein equation, we first multiply both sides of (2.1) with $e^{-w^2/2}$ to obtain

$$\left(e^{-w^2/2} f_z(w)\right)' = e^{-w^2/2} \left(\mathbb{1}\{w \le z\} - \Phi(z)\right).$$

Integrating both sides gives

$$e^{-w^2/2} f_z(w) = \int_{-\infty}^{w} e^{-x^2/2} \left(\mathbb{1}\{x \le z\} - \Phi(z)\right) \mathrm{d}x,$$

so that

$$\begin{aligned} f_z(w) &= e^{w^2/2} \int_{-\infty}^{w} e^{-x^2/2} \left(\mathbb{1}\{x \le z\} - \Phi(z)\right) \mathrm{d}x \\ &= -e^{w^2/2} \int_{w}^{\infty} e^{-x^2/2} \left(\mathbb{1}\{x \le z\} - \Phi(z)\right) \mathrm{d}x, \end{aligned}$$

which gives (2.2). We have thus found a specific solution to (2.1). To obtain the general solution, we have to add a constant multiple of the solution to the homogeneous equation

$$f'(w) - wf(w) = 0,$$

which is $ce^{w^2/2}$ for any $c \in \mathbb{R}$. Hence (2.2) is the only bounded solution to (2.1), where the boundedness is shown by (2.3). However, we refer to [CGS10, Appendix chapter 2] for a proof of (2.3), since it is rather technical and does not provide insights into the techniques necessary to understand this thesis. □

Proof of Lemma 2.1. For *(i)* we need to show that $\mathbb{E}[\mathcal{A}f(Z)] = 0$ for all f such that $\mathbb{E}[f'(Z)] < \infty$ if Z has a standard normal distribution. Therefore, note that

$$\begin{aligned} &\mathbb{E}[f'(Z)] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f'(w) e^{-w^2/2} \,\mathrm{d}w \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} f'(w) \left(\int_{-\infty}^{w} -xe^{-x^2/2} \,\mathrm{d}x\right) \mathrm{d}w + \frac{1}{\sqrt{2\pi}} \int_{0}^{\infty} f'(w) \left(\int_{w}^{\infty} xe^{-x^2/2} \,\mathrm{d}x\right) \mathrm{d}w. \end{aligned}$$

By Fubini's theorem it follows that

$$\begin{aligned} \mathbb{E}[f'(Z)] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} \left(\int_{x}^{0} f'(w) \,\mathrm{d}w\right) (-x) e^{-x^2/2} \,\mathrm{d}x \\ &\quad + \frac{1}{\sqrt{2\pi}} \int_{0}^{\infty} \left(\int_{0}^{x} f'(w) \,\mathrm{d}w\right) xe^{-x^2/2} \,\mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left(f(x) - f(0)\right) xe^{-x^2/2} \,\mathrm{d}x \\ &= \mathbb{E}[Zf(Z)]. \end{aligned}$$

To show (ii) note that for any $z \in \mathbb{R}$ the function f_z in (2.2) is clearly continuous and piecewise continuously differentiable. Furthermore, due to (2.3) we know that it is also bounded. Thus

$$0 = \mathbb{E}\left[\mathcal{A} f_z(X)\right] = \mathbb{E}\left[f_z'(X) - x f_z(X)\right] = \mathbb{E}\left[\mathbb{1}\{X \leq z\} - \Phi(z)\right] = \mathbb{P}(X \leq z) - \Phi(z).$$

Hence, X has a standard normal distribution.

□

Remark 2.3. *We can write* (2.1) *more generally as*

$$f'(w) - w f(w) = h(w) - \mathbb{E}\left[h(Z)\right]$$

for any measurable function h such that $\mathbb{E}[h(Z)] < \infty$, where $Z \sim \mathcal{N}(0,1)$. The solution to this equation is then given by

$$f_h(w) = -e^{w^2/2} \int_w^\infty \left(h(x) - \mathbb{E}\left[h(Z)\right]\right) e^{-x^2/2} \,\mathrm{d}x$$

with

$$\|f_h\| \leq \sqrt{\frac{\pi}{2}} \left\|h(\cdot) - \mathbb{E}\left[h(Z)\right]\right\| \text{ and } ||f_h'|| \leq 2 \left\|h(\cdot) - \mathbb{E}\left[h(Z)\right]\right\|,$$

see for example [CGS10] for more details.

Now, the next step is to bound the right-hand side of (2.1) in order to obtain the desired error bounds. One approach to do so is via size-bias couplings, which were first used by Goldstein and Rinott in the context of Stein's method, see [GR96]. We will make use of this technique in chapter 5. The subsequent results of the remaining part of this section (except for some minor adjustments) can all be found in [Ros11].

Definition 2.4. *For a random variable $X \geq 0$ with $\mathbb{E}[X] = \mu < \infty$, we say that the random variable X^s has the size-bias distribution with respect to X if for f such that $\mathbb{E}[X f(X)] < \infty$ we have*

$$\mathbb{E}\left[X f(X)\right] = \mu \mathbb{E}\left[f(X^s)\right].$$

An equivalent characterization of the size-bias distribution is that of F^s being absolutely continuous with respect to the distribution F of X with Radon-Nikodým density

$$\frac{\mathrm{d}F^s(x)}{\mathrm{d}F(x)} = \frac{x}{\mu}.$$

Corollary 2.5. *If $X \geq 0$ is a random variable with $\mathbb{E}[X] = \mu < \infty$, then the random variable X^s with the size-bias distribution of X is such that*

$$\mathbb{P}(X^s = k) = \frac{k\mathbb{P}(X = k)}{\mu}.$$

The next result, which is the essential part of our proof of a central limit theorem for the rescaled number of isolated vertices in chapter 5, was first proven in [GR96] and is also formulated in [Ros11, Theorem 3.20]. Here we state a slightly modified version of it, which is already adapted to the context of random graphs.

Theorem 2.6. *For a random graph $\mathcal{G}_n$ let $W_n \geq 0$ be some $\sigma(\mathcal{G}_n)$-measurable random variable with $W_n \geq 0$, $\mathbb{E}[W_n] = \mu_n < \infty$ and $\mathbb{V}(W_n) = \sigma_n^2$. Let W_n^s be defined on the same space as W_n and have the size-bias distribution with respect to W_n. If $\tilde{W}_n = \frac{W_n - \mu_n}{\sigma_n}$ and $Z \sim \mathcal{N}(0,1)$, then*

$$d_W(\tilde{W}_n, Z) \leq \frac{\mu_n}{\sigma_n^2}\sqrt{\frac{2}{\pi}}\sqrt{\mathbb{V}(\mathbb{E}[W_n^s - W_n|\mathcal{G}_n])} + \frac{\mu_n}{\sigma_n^3}\mathbb{E}\left[(W_n^s - W_n)^2\right]. \tag{2.4}$$

Proof. Since we are using Stein's method the strategy is to bound

$$\left|\mathbb{E}\left[f'(\tilde{W}_n) - \tilde{W}_n f(\tilde{W}_n)\right]\right|$$

for bounded f with bounded first and second derivative. Using the definition of the size-bias distribution we obtain

$$\begin{aligned}\mathbb{E}\left[\tilde{W}_n f(\tilde{W}_n)\right] &= \mathbb{E}\left[\frac{W_n - \mu_n}{\sigma_n} f\left(\frac{W_n - \mu_n}{\sigma_n}\right)\right] \\ &= \frac{\mu_n}{\sigma_n}\mathbb{E}\left[f\left(\frac{W_n^s - \mu_n}{\sigma_n}\right) - f\left(\frac{W_n - \mu_n}{\sigma_n}\right)\right]\end{aligned}$$

so that by a Taylor expansion we get

$$\mathbb{E}\left[f(\tilde{W}_n)\right] = \frac{\mu_n}{\sigma_n}\mathbb{E}\left[\frac{W_n^s - W_n}{\sigma_n} f'\left(\frac{W_n - \mu_n}{\sigma_n}\right) + \frac{(W_n^s - W_n)^2}{2\sigma_n^2} f''\left(\frac{W_n^* - \mu_n}{\sigma_n}\right)\right],$$

for some W_n^* in the interval with endpoints W_n and W_n^s. Using the definition of $\tilde{W}_n$ in the previous expression, we obtain

$$\begin{aligned}\left|\mathbb{E}\left[\tilde{W}_n f'\left(\tilde{W}_n\right) - \tilde{W}_n f(\tilde{W}_n)\right]\right| \leq &\left|\mathbb{E}\left[f'(\tilde{W}_n)\left(1 - \frac{\mu_n}{\sigma_n^2}(W_n^s - W_n)\right)\right]\right| \\ &+ \frac{\mu_n}{2\sigma_n^3}\left|\mathbb{E}\left[f''\left(\frac{X_n^* - \mu_n}{\sigma_n}\right)(W_n^s - W_n)^2\right]\right|.\end{aligned}$$

Since according to (2.3) the solutions f to the Stein equation fulfill $\|f'\| \leq \sqrt{\frac{2}{\pi}}$ and $\|f''\| \leq 2$ we have

$$\frac{\mu_n}{2\sigma_n^3}\left|\mathbb{E}\left[f''\left(\frac{X_n^* - \mu_n}{\sigma_n}\right)(W_n^s - W_n)^2\right]\right| \leq \frac{\mu_n}{\sigma_n^3}\mathbb{E}\left[(W_n^s - W_n)^2\right],$$

which is the second term appearing in the error bound (2.4). To bound the first term note that by the definition of W_n^s we have

$$\mathbb{E}\left[W_n^s\right] = \frac{1}{\mu_n}\mathbb{E}\left[W_n^2\right] = \frac{\sigma_n^2 + \mu_n^2}{\mu_n}$$

so that by the properties of the Stein solution mentioned above, the law of iterated expectations and the Cauchy-Schwarz inequality we obtain

$$\begin{aligned}\left|\mathbb{E}\left[f'(\tilde{W}_n)\left(1 - \frac{\mu_n}{\sigma_n^2}(W_n^s - W_n)\right)\right]\right| &\leq \sqrt{\frac{2}{\pi}}\mathbb{E}\left[\left|1 - \frac{\mu_n}{\sigma_n^2}\mathbb{E}\left[W_n^s - W_n|\mathcal{G}_n\right]\right|\right] \\ &\leq \frac{\mu_n}{\sigma_n^2}\sqrt{\frac{2}{\pi}}\sqrt{\mathbb{V}(\mathbb{E}\left[W_n^s - W_n|\mathcal{G}_n\right])}.\end{aligned}$$

□

We now give a general construction for a size-bias version of some random variable W_n which can be written as a sum of random variables. Just as the previous result, this construction was first given in [GR96]. The version we put is taken from [Ros11].

Coupling construction
For the case in which $W_n = \sum_{i=1}^n X_i$ with $X_i \geq 0$ and $\mathbb{E}\left[X_i\right] = \nu_i$, section 3.4.1 in [Ros11] provides the following construction of a size-bias version of W_n:

1. For each $i = 1, \ldots n$, let X_i^s have the size-bias distribution of X_i independent of $(X_j)_{j\neq i}$ and $(X_j^s)_{j\neq i}$. Given $X_i^s = x$ define the vector $(X_j^{(i)})_{j\neq i}$ to have distribution of $(X_j)_{j\neq i}$ conditional on $X_i = x$.

2. Choose a random summand X_I, where the index I is chosen proportional to μ_i and independent of all else. Specifically, $\mathbb{P}(I = i) = \frac{\nu_i}{\mu_n}$, where $\mu_n = \mathbb{E}\left[W_n\right]$.

3. Define $W_n^s = \sum_{j\neq I} X_j^{(I)} + X_I^s$.

Proposition 2.7. *Let $W_n = \sum_{i=1}^n X_i$ with $X_i \geq 0$, $\mathbb{E}\left[X_i\right] = \nu_i$ and $\mu_n = \mathbb{E}\left[W_n\right]$. If W_n^s is constructed according to items 1 - 3 above, then W_n^s has the size-bias distribution of W_n.*

Proof. To prove the result it is enough to show that

$$\mathbb{E}\left[W_n f(W_n)\right] = \mu_n \mathbb{E}\left[f(W_n^I)\right],$$

where $W_n^I = \sum_{j\neq I} X_j^{(I)} + X_I^s$ as given above. If we can show that

$$\mathbb{E}\left[X_i f(W_n)\right] = \nu_i \mathbb{E}\left[f(W_n^i)\right]. \tag{2.5}$$

it follows that

$$\mathbb{E}\left[W_n f(W_n)\right] = \sum_{i=1}^{n} \mathbb{E}\left[X_i f(W_n)\right] = \sum_{i=1}^{n} \nu_i \mathbb{E}\left[f(W_n^i)\right] = \mu_n \mathbb{E}\left[f(W_n^I)\right]$$

since $\mathbb{P}(I = i) = \frac{\nu_i}{\mu_n}$. To show (2.5), note that for $h(X_i) = \mathbb{E}\left[f(W_n)|X_i\right]$ we have

$$\begin{aligned}\mathbb{E}\left[X_i f(W_n)\right] &= \mathbb{E}\left[X_i \mathbb{E}\left[f(W_n)|X_i\right]\right] = \mathbb{E}\left[X_i h(X_i)\right] \\ &= \nu_i \mathbb{E}\left[h(X_i^s)\right] = \nu_i \mathbb{E}\left[\mathbb{E}\left[f(W_n)|X_i^s\right]\right] = \nu_i \mathbb{E}\left[f(W_n^i)\right].\end{aligned}$$

□

Corollary 2.8. *Let $X_1, \ldots X_n$ be zero-one random variables and let $p_i := \mathbb{P}(X_i = 1)$. For each $i = 1, \ldots, n$ let $(X_j^{(i)})_{j \neq i}$ have the distribution of $(X_j)_{j \neq i}$ conditional on $X_i = 1$. If $W_n = \sum_{i=1}^n X_i$, $\mu_n = \mathbb{E}\left[W_n\right]$, and I is chosen independent of all else with $\mathbb{P}(I = i) = \frac{p_i}{\mu_n}$, then $W_n^s = \sum_{j \neq I} X_j^{(I)} + 1$ has the size-bias distribution of W_n.*

Proof. The result follows directly from Proposition 2.7 and the observation that for a Bernoulli random variable X, its size-bias version is given by $X^s = 1$, as for all f such that $\mathbb{E}\left[X f(X)\right] < \infty$ we have $\mathbb{E}\left[X f(X)\right] = p_i f(1)$. □

2.1.2. Stein's method in a nutshell

As mentioned at the beginning of this chapter Stein's method can be generalized to limiting distributions other than normal. The general procedure to follow when developing a Stein's method for a new target distribution is given by the following four steps:

1. Decide on a suitable limiting distribution for the random variable of interest.
2. Find a characterizing operator $\mathcal{A}$ of the target distribution μ, in the sense that for all functions g in the domain of $\mathcal{A}$

$$\mathbb{E}\left[\mathcal{A}g(W)\right] = 0 \Leftrightarrow W \sim \mu.$$

3. Find a solution g_h to the Stein-equation

$$h(k) - \int h \, \mathrm{d}\mu = \mathcal{A}g(k), \tag{2.6}$$

for each h in a measure-determining class of functions $\mathcal{F}$. This yields

$$\mathrm{d}_{\mathcal{F}}(W, X) = \sup_{h \in \mathcal{F}} \left|\mathbb{E}\left[h(W)\right] - \mathbb{E}\left[h(X)\right]\right| = \mathbb{E}\left[\mathcal{A}g_h(X)\right]. \tag{2.7}$$

For different families $\mathcal{F}$ of test functions one gets different probability metrics. Consider for example the following classes of test functions

- for $\mathcal{F} = \{h : \mathbb{R} \to \mathbb{R} : |h(x) - h(y)| \leq |x - y|\}$ we get $\mathrm{d}_{\mathcal{F}}(W, X) = d_W(W, X)$, where d_W denotes the Wasserstein distance.
- for $\mathcal{F} = \{\mathbb{1}_{(-\infty,z]}, z \in \mathbb{R}\}$ we get $\mathrm{d}_{\mathcal{F}}(W, X) = d_K(W, X)$, where d_K denotes the Kolmogorov distance (see the previous section).
- for $\mathcal{F} = \{\mathbb{1}_A(\cdot), A \in \mathrm{Borel}(\mathbb{R})\}$ we get $\mathrm{d}_{\mathcal{F}}(W, X) = d_{TV}(W, X)$, where d_{TV} denotes the total variation distance.

4. Bound the right-hand side of (2.7) to get bounds on the distributional distance of W and X. Usually one needs to find bounds on either $g_h^{(k)}$ (in the continuous case) or $\Delta g_h(k) := g_h(k+1) - g_h(k)$ (in the discrete case) to do so. Uniform bounds on these quantities are then often referred to as Stein's "magic factors".

2.1.3. The generator approach

A rather general approach for finding suitable Stein-operators and corresponding solutions is the generator-approach introduced by Barbour [Bar88], who applied it to multivariate Poisson approximation, and which was developed by Götze in [Göt91] for multivariate normal approximation. The basic idea behind this approach is to find a Markov process (X_t) with invariant distribution μ. We write

$$T_t f(x) := \mathbb{E}\left[g(X_t)|X_0 = x\right]$$

for the operator semigroup corresponding to (X_t), which acts on $\mathcal{L}^2(\mu)$. From general Markov theory we now know that the infinitesimal generator $\mathcal{A}$ corresponding to the process X_t is given by the following limit in $\mathcal{L}^2(\mu)$

$$\mathcal{A}g := \lim_{t \searrow 0} \frac{T_t g - g}{t} \tag{2.8}$$

for $g \in dom(\mathcal{A})$, the class of $\mathcal{L}^2(\mu)$-functions for which the limit exists.
Under certain conditions on the operator $\mathcal{A}$, the theory of operator semigroups (e.g. Proposition 1.5 in [EK86]) yields

$$\int_0^t T_s g \,\mathrm{d}s \in dom(\mathcal{A}) \text{ for all } g \in \mathcal{L}^2(\mu) \text{ and } \mathcal{A}\left(\int_0^t T_s g \,\mathrm{d}s\right) = T_t g - g, \tag{2.9}$$

where the integral is a $\mathcal{L}^2(\mu)$ valued Riemann-integral. We now get the following result which shows that a generator $\mathcal{A}$ of a Markov process as given in (2.8) defines a Stein operator for the invariant distribution μ of the process.

Proposition 2.9. *A $(\mathcal{X}, \mathcal{B})$-valued random variable X has distribution μ with generator $\mathcal{A}$ if and only if for all functions g in $dom(\mathcal{A})$ we have $\mathbb{E}[\mathcal{A}g(X)] = 0$. Thus $\mathcal{A}$ as in* (2.8) *defines a Stein operator for μ.*

Proof. Let X have distribution μ and take $g \in \mathit{dom}(\mathcal{A})$. As $(X_t)_{t\geq 0}$ has invariant distribution μ we get

$$\int_{\mathcal{X}} T_t g \,\mathrm{d}\mu = \mathbb{E}_\mu \left[T_t g(X)\right] = \mathbb{E}_\mu \left[\mathbb{E}\left[g(X_t)|X_0 = X\right]\right] = \mathbb{E}_\mu \left[g(X_t)\right] = \int_{\mathcal{X}} g \,\mathrm{d}\mu$$

so that

$$\int_{\mathcal{X}} T_t g - g \,\mathrm{d}\mu = 0 \text{ and thus } \int_{\mathcal{X}} \frac{T_t g - g}{t} \,\mathrm{d}\mu = 0$$

for all $t \geq 0$. As

$$\mathcal{A}g = \lim_{t \searrow 0} \frac{T_t g - g}{t} \in \mathcal{L}^2(\mu)$$

the theorem of dominated convergence yields

$$0 = \lim_{t \searrow 0} \int_{\mathcal{X}} \frac{T_t g - g}{t} \,\mathrm{d}\mu = \int_{\mathcal{X}} \lim_{t \searrow 0} \frac{T_t g - g}{t} \,\mathrm{d}\mu = \int_{\mathcal{X}} \mathcal{A}g \,\mathrm{d}\mu = \mathbb{E}\left[\mathcal{A}g(X)\right]$$

so that

$$X \sim \mu \Rightarrow \mathbb{E}\left[\mathcal{A}g(X)\right] = 0.$$

For the converse take $g \in \mathcal{L}^2(\mu)$ and $t > 0$. Due to (2.9) we have

$$\mathbb{E}\left[T_t g(X) - g(X)\right] = \mathbb{E}\left[\mathcal{A}\left(\int_0^t T_s g \,\mathrm{d}s\right)(X)\right] = 0$$

and thus

$$\mathbb{E}\left[T_t g(X)\right] = \mathbb{E}\left[g(X)\right]$$

for all $t > 0$, which yields that $\mathcal{L}(X)$ is an invariant distribution for $(X_t)_{t\geq 0}$. As $(X_t)_{t\geq 0}$ is an ergodic process, we obtain that $\mathcal{L}(X) = \mu$. □

To find a solution to the Stein equation (2.6), we can apply (2.9) to the function $g = h - \mu(h)$ with $h \in \mathcal{F}$. This yields

$$\mathcal{A}\left(-\int_0^t (T_s h - \mu(h)) \,\mathrm{d}s\right) = h - T_t h$$

and as T_t is the corresponding operator semigroup to the process $(X_t)_{t\geq 0}$ with invariant distribution μ, we obtain

$$\mathcal{A}\left(-\int_0^\infty (T_s h - \mu(h)) \,\mathrm{d}s\right) = h - \mu(h)$$

for $t \to \infty$, if the integral exists. Thus a solution to (2.9) is given by

$$g_h(k) := -\int_0^\infty (T_s h(k) - \mu(h)) \,\mathrm{d}s.$$

Consequently, one can deduce a Stein-operator and a corresponding solution to the Stein equation with the help of the generator approach.

2.2. Coupling

The coupling method, which was introduced by Wolfgang Döblin in the late 1930s [Döb38], is a powerful technique in probability theory which allows to compare two probability measures defined on the same measurable space. This technique is so useful because a comparison between distributions is reduced to a comparison between random variables and thus provides an effective method of obtaining upper bounds on distributional distances. The following section captures the most important results from the works [dH12] and [LPW06] which are relevant for the thesis at hand. We start by defining the coupling of two probability measures:

Definition 2.10. *A coupling of two probability measures μ and ν on the same measurable space $(\Omega, \mathcal{A})$ is any probability measure $\mathbb{P}$ on the product measurable space $(\Omega \times \Omega, \mathcal{A} \otimes \mathcal{A})$ (where $\mathcal{A} \otimes \mathcal{A}$ is the smallest σ-algebra containing $\mathcal{A} \times \mathcal{A}$) whose marginals are μ and ν, i.e.*

$$\mu(x) = \sum_{y \in \Omega} \mathbb{P}(x, y) \text{ and } \nu(y) = \sum_{x \in \Omega} \mathbb{P}(x, y) \text{ for all } x, y \in \Omega.$$

Coupling of two random variables are defined in a similar way.

Definition 2.11. *A coupling of two random variables X and Y taking values in $(\Omega, \mathcal{A})$ is any pair of random variables (X', Y') taking values in $(\Omega \times \Omega, \mathcal{A} \otimes \mathcal{A})$ whose marginals have the same distribution as X and Y, i.e.*

$$X' \stackrel{\mathcal{D}}{=} X \text{ and } Y' \stackrel{\mathcal{D}}{=} Y.$$

Note that the law $\mathbb{P}$ of (X', Y') is a coupling of the laws of μ and ν of X and Y respectively in the sense of Definition 2.10.

The following proposition shows the close connection between couplings and the total variation distance of two random variables. Remember that for any two probability measures μ and ν we have

$$\|\mu - \nu\|_{TV} := \sup_{A \in \mathcal{A}} |\mu(A) - \nu(A)|.$$

For the rest of this section we will restrict to the case of discrete probability spaces, e.g. we assume Ω to be finite or countably infinite. This is sufficient for our purposes as we will use the given results for Markov chains X and Y taking values in $\mathbb{N}_0$.

Proposition 2.12. *Let μ and ν be two probability distributions on Ω. Then*

$$\|\mu - \nu\|_{TV} = \frac{1}{2} \sum_{x \in \Omega} |\mu(x) - \nu(x)|.$$

Proof. Let $A \subset \Omega$ be an event and set $B = \{x \in \Omega : \mu(x) \geq \nu(x)\}$. We then get

$$\begin{aligned}\mu(A) - \nu(A) &= \mu(A \cap B) + \mu(A \cap B^c) - \nu(A)\\ &\leq \mu(A \cap B) + \nu(A \cap B^c) - \nu(A)\\ &= \mu(A \cap B) - \nu(A \cap B)\end{aligned}$$

since for all $x \in A \cap B^c$ we have $\nu(x) > \mu(x)$. Moreover,

$$\mu(A \cap B) - \nu(A \cap B) = \sum_{x \in A \cap B} \mu(x) - \nu(x) \leq \sum_{x \in B} \mu(x) - \nu(x) = \mu(A) - \nu(B),$$

since for all $x \in B$ we have $\mu(x) - \nu(x) \geq 0$. Hence

$$\mu(A) - \nu(A) \leq \mu(B) - \nu(B). \tag{2.10}$$

Since the result above holds for all $A \subset \Omega$ we also have

$$\nu(A) - \mu(A) = \mu(A^c) - \nu(A^c) \leq \mu(B) - \nu(B) = \nu(B^c) - \mu(B^c). \tag{2.11}$$

The last equality shows that the upper bounds (2.10) and (2.11) actually coincide. Furthermore, for $A = B$ (or B^c) we obtain equality in the calculations above, so that the maximal difference is achieved in this case. With these considerations we get

$$\begin{aligned}\|\mu - \nu\|_{TV} = \sup_{A \subset \Omega} |\mu(A) - \nu(A)| &= \frac{1}{2}\left(\mu(B) - \nu(B) + \nu(B^c) - \mu(B^c)\right)\\ &= \frac{1}{2}\sum_{x \in \Omega} |\mu(x) - \nu(x)|.\end{aligned}$$

□

Remark 2.13. *The proof of the previous proposition shows in particular that*

$$\|\mu - \nu\|_{TV} = \sum_{\substack{x \in \Omega \\ \mu(x) \geq \nu(x)}} \left(\mu(x) - \nu(x)\right). \tag{2.12}$$

Proposition 2.14. *Let X, Y be two random variables with probability distributions μ and ν. Then, for any coupling (X', Y') of X and Y, we have*

$$\|\mu - \nu\|_{TV} \leq \mathbb{P}(X' \neq Y'). \tag{2.13}$$

In fact, we have

$$\|\mu - \nu\|_{TV} = \inf\{\mathbb{P}(X' \neq Y') : (X', Y') \textit{ is a coupling of } \mu \textit{ and } \nu\}. \tag{2.14}$$

Proof. To show (2.13) note that

$$\mu(A) - \nu(A) = \mathbb{P}(X \in A) - \mathbb{P}(Y \in A) \le \mathbb{P}(X \in A, Y \notin A) \le \mathbb{P}(X \neq Y)$$

so that definitely

$$\mu(A) - \nu(A) \le \inf\{\mathbb{P}(X' \neq Y') : (X', Y') \text{ is a coupling of } \mu \text{ and } \nu\}.$$

We will now construct a coupling (X', Y') with $\|\mu(A) - \nu(A)\| = \mathbb{P}(X' \neq Y')$. We will do so by forcing X and Y to be equal as often as they possibly can be. To generate X and Y let

$$p = \sum_{x\in\Omega} \mu(x) \wedge \nu(x) = \sum_{\substack{x\in\Omega \\ \mu(x)\le\nu(x)}} \mu(x) + \sum_{\substack{x\in\Omega \\ \nu(x)<\mu(x)}} \nu(x).$$

Adding and subtracting $\sum_{\substack{x\in\Omega \\ \mu(x)>\nu(x)}} \mu(x)$ to the right-hand side yields

$$p = \sum_{x\in\Omega} \mu(x) \wedge \nu(x) = 1 - \sum_{\substack{x\in\Omega \\ \nu(x)<\mu(x)}} (\mu(x) - \nu(x)) = 1 - \|\mu - \nu\|_{TV}, \tag{2.15}$$

where we used Proposition 2.12. We now flip a coin which shows heads with probability p. If the coin comes up heads, we choose a value $Z \in \Omega$ according to the probability distribution

$$\alpha(x) = \frac{\mu(x) \wedge \nu(x)}{p}$$

and set $X = Y = Z$. If the coin comes up tails, we choose X according to the probability distribution

$$\beta(x) = \begin{cases} \frac{\mu(x)-\nu(x)}{\|\mu-\nu\|_{TV}} & \text{if } \mu(x) > \nu(x), \\ 0, & \text{otherwise.} \end{cases}$$

Then we independently choose Y according to the probability distribution

$$\gamma(x) = \begin{cases} \frac{\mu(x)-\nu(x)}{\|\mu-\nu\|_{TV}} & \text{if } \nu(x) > \mu(x), \\ 0, & \text{otherwise.} \end{cases}$$

Note that due to (2.12) and (2.15) α, β and γ all define probability distributions. With these considerations we obtain

$$\begin{aligned} \mathbb{P}(X = x) &= p\alpha(x) + (1-p)\beta(x) \\ &= \mu(x) \wedge \nu(x) + (\mu(x) - \nu(x))\, \mathbb{1}\{\mu(x) > \nu(x)\} = \mu(x) \end{aligned}$$

and

$$\begin{aligned}\mathbb{P}(Y = x) &= p\alpha(x) + (1-p)\gamma(x) \\ &= \mu(x) \wedge \nu(x) + (\nu(x) - \mu(x))\, \mathbb{1}\{\nu(x) > \mu(x)\} = \nu(x)\end{aligned}$$

for all $x \in \Omega$, meaning that X has distribution μ and Y has distribution ν respectively. Note that in the case in which the coin comes up tails, we have $X \neq Y$, so that $X = Y$ if and only if the coin shows heads. Hence

$$\mathbb{P}(X \neq Y) = 1 - p = \|\mu - \nu\|_{TV}$$

by (2.15). □

3. Degree evolutions in preferential attachment models via Stein's method

The main purpose of this chapter is to give rates of convergence in total variation between the indegree distribution of a uniformly chosen vertex in a general preferential attachment random graph, based on the model introduced by Dereich and Mörters [DM09] in 2009, and its limiting distribution with Stein's method. The first section states assumptions on the model under which the results, that we will prove, hold and gives a more detailed description of one of the models, which will be relevant in section 3.3 and chapter 5. In section 3.2 we develop Stein's method for limiting distributions of the form

$$\mu_k = \frac{1}{1+f(k)} \prod_{i=0}^{k-1} \frac{f(i)}{1+f(i)},$$

where f is some monotonically increasing (sub-)linear function in the preferential attachment model. After that we use this method to derive rates of convergence of the indegree distribution of a uniformly chosen vertex towards μ.
Finally, section 3.3 proves that the outdegree in a more specific model described in example 3.1 is approximately Poisson-distributed and also gives error bounds on this approximation.

The results presented in this section, except for those in section 3.2.4, can be found in the preprint [BDO19].

3.1. Preferential attachment models

We study a general preferential attachment model, based on the model introduced by Dereich and Mörters in [DM09], where the connection probabilities are given by a general (sub-)linear function of the old degree. Though our methods do not rely on the explicit details of the model under consideration, we will first give a detailed introduction to the model in [DM09] as this will also be the relevant model analyzed in section 3.3 and chapter 5. Afterwards we will state the assumptions on a preferential attachment model that need to be met for our results to hold, highlight some other models that fit in this class and then state our results in full generality.

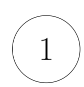
Example 3.1 (Preferential attachment with random outdegree)**.**

Take any $f : \mathbb{N}_{\{0\}} \to (0, \infty)$ with $f(n) \leq n + 1$. The network we consider is built according to the following rules. Each graph $\mathcal{G}_n$ consists of n vertices labeled $\{1, \ldots, n\}$ and a random number of edges. We start with the graph $\mathcal{G}_1$ consisting of one vertex and no edges.
At time $n + 1$ we insert vertex $n + 1$ into the graph $\mathcal{G}_n$ and independently for each $k \in [n]$ we add a directed edge from $n + 1$ to k with probability

$$\frac{f(\deg_n^-(k))}{n}, \tag{3.1}$$

where $\deg_n^-(k)$ denotes the indegree of vertex k at time n. Due to the assumptions imposed on our attachment function f it is guaranteed that in each evolution step (3.1) in fact lies between zero and one.
In contrast to many other models, like for instance those considered in [BA99], [KR01] [OS05], [Ros13] and [RTV07], the outdegree of every vertex in this model is random and can be zero. In many applications (like for example collaboration networks) this seems to be a more reasonable assumption than a fixed outdegree. Note that the outdegree of every vertex is fixed after the time step in which it was inserted into the network. Formally we are dealing with a directed network. However, by construction, edges are always pointing from younger to older vertices, so that the directions can be recreated from the undirected labeled graph.
An example of the first evolutionary steps of such a preferential attachment graph is visualized in the following figures.

$\mathcal{G}_1$:



$\mathcal{G}_2$:

$\mathcal{G}_3$:

$\mathcal{G}_4$:

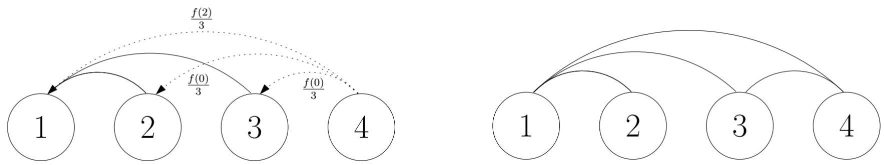


Figure 3.1: Evolution of $\mathcal{G}_n$

Clearly, the topological structure of the network crucially depends on the attachment function f. Some examples of networks observed after $n = 50$ time steps are depicted in Figures 3.2 -3.5.

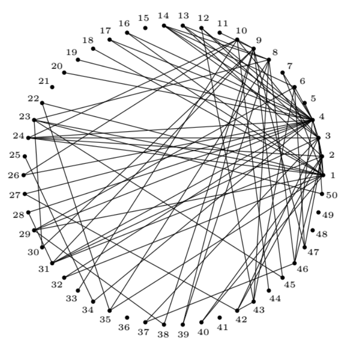


Figure 3.2: $f(k) = 0.8k + 0.6$

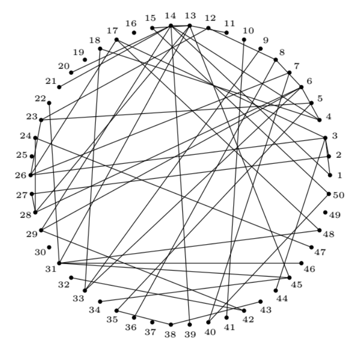


Figure 3.3: $f(k) = \sqrt{k+1}$

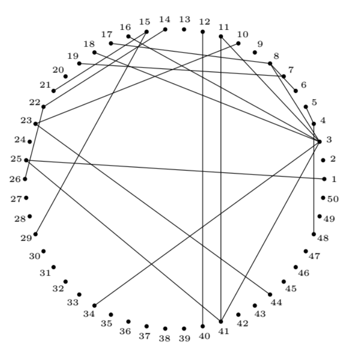


Figure 3.4: $f(k) = 0.4k + 0.3$

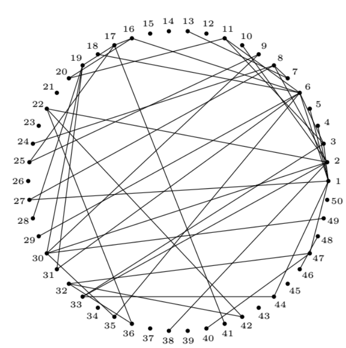


Figure 3.5: $f(k) = 0.8\sqrt{k} + 0.6$

Though this framework is already sufficiently general to lead to typical degree distributions that can be power-laws, but also stretched exponential distributions, we can prove our results in this section for an even larger class of preferential attachment models, fulfilling the following assumptions:

Assumptions (A). Fix $d_0 \in \mathbb{N}_0$ and let $f : \mathbb{N}_0 \to (0,\infty)$ such that $f(n) \leq \max\{n+1-d_0, 1\}$. We assume that $(\mathcal{G}_n)_{n\geq 1}$ is a sequence of directed random graphs with vertex set $[n] := \{1,\dots,n\}$. The initial graph $\mathcal{G}_1$ consists of one vertex, labeled 1, and d_0 (directed) self-edges. For any $n \geq 1$, at time $n+1$ we add a vertex, which we label $n+1$, to the vertex set for each $j \in [n]$ we insert at most one directed edge from $n+1$ to j such that

$$\mathbb{P}(n+1 \text{ connects to } j \,|\, \mathcal{G}_n) = \frac{f(\deg_n^-(j))}{n}. \tag{3.2}$$

Here $\deg_n^-(j)$ denotes the indegree of vertex j after the n-th vertex has been inserted. Note that by construction we have that $\deg_n^-(j) \leq d_0 + n - 1$, so that by the condition on f the right hand side of (3.2) is indeed ≤ 1.

These assumptions do not completely specify the model: they allow for deterministic as well as random outdegree and also only impose conditions on the marginal probabilities of connecting $n+1$ to j. In particular, additionally to the model introduced in Example 3.1 the following models are included.

Example 3.2 (Preferential attachment with fixed outdegree)**.** Start with $\mathcal{G}_1$ consisting of vertex 1 and a (directed) self-loop. At time $n+1$ insert vertex $n+1$ and connect it to exactly one previous vertex $j \in [n]$ with probability

$$\frac{\deg_n(j) + \delta}{n(2+\delta)},$$

where $\deg_n(j)$ denotes the total degree of vertex j at time n and $\delta > -1$ is a parameter of the model. Noticing that $\deg_n(j) = \deg_n^-(j) + 1$, this fits into our framework with $f(k) = \frac{k+(1+\delta)}{2+\delta}$ and $d_0 = 1$. This model almost coincides with the one proposed in [BRST01] (there however $\delta = 0$ and they allow for self-loops) and it is very closely related to what is referred to as model (b) in [Hof17, Chapter 8.2].

Example 3.3 (Spatial preferential attachment model)**.** In [ACJP08], the authors introduce the following spatial random graph model. Let S be the unit hypercube in $\mathbb{R}^m$. The initial graph consists of vertex 1 that is placed uniformly at random into S and no edges. For each vertex i we define *the sphere of influence* $S(i,n)$ *of* i as the ball (in the torus metric induced by the Euclidean metric) that has volume $\frac{A_1 \deg_n^-(i)+A_2}{n}$ centered at the position of i, where $A_1, A_2 \geq 0$. Fix a parameter $p \in [0,1]$. Then, at time $n+1$, we insert vertex $n+1$ at a position that is chosen uniformly at random in S. Now, independently for each vertex j such that the position of $n+1$ is in $S(j,n)$ insert an edge from $n+1$ to j with probability p. In particular, we get

$$\mathbb{P}(n+1 \text{ connects to } j \,|\, \mathcal{G}_n) = p \cdot \frac{A_1 \deg_n^-(j) + A_2}{n}.$$

Thus, this model fits into our framework if we choose $f(k) = pA_1k + pA_2$ and assume that the constants are chosen such that $pA_1, pA_2 \leq 1$.

The following theorem from [DM09] shows that in the model described in Example 3.1 the empirical indegree distribution converges almost surely in total variation norm to a distribution μ and the outdegree of every vertex is asymptotically Poisson distributed.

Theorem 3.4. *(a) Let*

$$\mu_k = \frac{1}{1+f(k)} \prod_{i=0}^{k-1} \frac{f(i)}{1+f(i)}, \quad k = 0, 1, \ldots, \tag{3.3}$$

which is a sequence of probability weights. Then, almost surely,

$$\lim_{n\to\infty} X(n) = \mu$$

in total variation norm, where $X(n) = (X_k(n) : k \in \mathbb{N}_0)$ *and* $X_k(n) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}\{deg_n^-(i) = k\}$.

(b) If f *satisfies* $f(k) \leq \eta k+1$ *for some* $\eta \in (0,1)$*, then the conditional distribution of the outdegree of* $(n+1)$*-th incoming node (given the graph at time* n*) converges almost surely in the total variation norm to the Poisson distribution with parameter* $\lambda := \langle \mu, f \rangle = \sum_{k\geq 0} f(k)\mu_k$.

As a side effect our results will also show that part a) of the previous theorem not only holds for the preferential attachment model described in Example 3.1, but for all models satisfying Assumptions **(A)**.

Remark 3.5. *Following [DM09, Example 1.3.] for* $f(k) = \gamma k + \eta$ *with fixed* $\gamma, \eta \in (0,1]$ *for all* $k \in \mathbb{N}_0$*, we have*

$$\mu_k = \frac{1}{\gamma} \frac{\Gamma(k + \frac{\eta}{\gamma})\Gamma(\frac{\eta+1}{\gamma})}{\Gamma(k + \frac{1+\eta+\gamma}{\gamma})\Gamma(\frac{\eta}{\gamma})},$$

so that by Stirling's formula, $\Gamma(t+a)\backslash\Gamma(t) \sim t^a$ *for* t *tending to infinity, we obtain*

$$\mu_k \sim -\frac{\Gamma(\frac{\eta+1}{\gamma})}{\gamma\,\Gamma(\frac{\eta}{\gamma})}\, k^{-(1+\frac{1}{\gamma})}. \tag{3.4}$$

Therefore, our framework allows for models with power-law distribution with tail exponent $1+1/\gamma \in [2,\infty)$*. Furthermore, if* $f(k) \sim \gamma k^\alpha$ *with* $0 < \alpha < 1$*,* $\gamma > 0$*, then*

$$\log \mu_k \sim \frac{1}{\gamma(1-\alpha)} k^{1-\alpha}, \quad \text{for } k \to \infty,$$

so that we obtain a limiting distribution with stretched exponential tails.

3.2. Stein's method for preferential attachment models

The main purpose of this chapter is to deduce rates of convergence of $\mathcal{L}(deg_n(I_n))$ in preferential attachment models satisfying **(A)**, where $I_n \sim \mathcal{U}\{1,\dots,n\}$ and $\deg_n(i)$ denotes the indegree of vertex i in $\mathcal{G}_n$. Our main result here is the following theorem

Theorem 3.6. *Let W_n denote the indegree of a uniformly chosen vertex at time n in a preferential attachment model satisfying Assumptions **(A)**. Suppose further that there exists $k_* \in \mathbb{N}_0$ such that $f(k) > k$ for all $k < k_*$ and $f(k) \leq k$ for all $k \geq k_*$. Then, there exists a constant $C > 0$ such that for all $n \geq 2$*

$$d_{\mathrm{TV}}(W_n, W) \leq C\ \frac{\log(n)}{n}, \tag{3.5}$$

where $W \sim \mu$ and μ as in (3.3).

The condition that there exists $k_* \in \mathbb{N}_0$ such that $f(k) > k$ for all $k < k_*$ and $f(k) \leq k$ for all $k \geq k_*$ is for example fulfilled for all sublinear models such that $\max_k \Delta f(k) = \max_k(f(k+1) - f(k)) < 1$, which is a popular condition in the setting of Example 3.1, see e.g. [DM13].

The next Theorem gives a weaker result in the regime where $f(k) \in [k, k+\gamma]$ for all k and $\gamma \in (0,1)$. This is not surprising since for example in the case $f(k) = k+\gamma$, the distribution has power law exponent 2 and does no longer have a finite mean, see (3.4).

Theorem 3.7. *Let W_n denote the indegree of uniformly chosen vertex at time n in a preferential attachment model satisfying Assumptions **(A)**. Suppose further that $f(k) \in [k, k+\gamma]$ for all $k \in \mathbb{N}_0$ and some $\gamma \in (0,1)$. Then, there exists a constant $C > 0$ such that for all $n \geq 1$,*

$$d_{\mathrm{TV}}(W_n, W) \leq C\ n^{-(1-\gamma)},$$

where $W \sim \mu$ and μ as in (3.3).

The proof of these two results uses Stein's method, which to our knowledge has not been developed for a general class of limiting distributions as given in (3.3). The following three paragraphs will now deal with the steps of Stein's method as outlined in section 2.1.2.

3.2.1. Stein-operator

Lemma 3.8. *Let μ be given by* (3.3)*. Then, μ is a probability distribution and any $\mathbb{N}_0$-valued random variable W satisfies $W \sim \mu$ if and only if*

$$\mathbb{E}[\mathcal{A}g(W)] = 0,$$

for all $g : \mathbb{N}_0 \to \mathbb{R}$ such that $\mathbb{E}[g(Y)] < \infty$, where

$$\mathcal{A}g(k) := f(k)\big(g(k+1) - g(k)\big) + g(0) - g(k).$$

Proof. Let N_t be a process that starts in 0 and jumps from i to $i+1$ at rate $f(i)$. Then the time of the k-th jump is given in distribution by $S_k = \sum_{i=0}^{k-1} \frac{1}{f(i)} E_i$, where $E_0, E_1, \dots$ is an i.i.d. sequence of exponentials with rate 1. Now (see also Cor. 50 in [Bha07]) let Y be an independent exponential random variable with parameter 1, then

$$\mathbb{P}\left(N_Y \geq k\right) = \mathbb{P}\left(Y \geq S_k\right).$$

By first conditioning on S_k, we get

$$\begin{aligned}
\mathbb{P}\left(Y \geq S_k\right) &= \mathbb{E}\big[\mathbb{E}[\mathbb{1}\{Y \geq S_k\} \mid S_k]\big] = \mathbb{E}[e^{-S_k}] = \mathbb{E}[e^{-\sum_{i=0}^{k-1} \frac{1}{f(i)} E_i}] \\
&= \prod_{i=0}^{k-1} \mathbb{E}[e^{-E_i/f(i)}] = \prod_{i=0}^{k-1} \frac{f(i)}{1+f(i)},
\end{aligned} \tag{3.6}$$

using that $\mathbb{E}[e^{-\lambda E_i}] = \frac{1}{1+\lambda}$ in the last step. We have

$$\begin{aligned}
\mathbb{P}(N_Y = k) &= \prod_{i=0}^{k-1} \frac{f(i)}{1+f(i)} - \prod_{i=0}^{k} \frac{f(i)}{1+f(i)} \\
&= \left(1 - \frac{f(k)}{1+f(k)}\right) \prod_{i=0}^{k-1} \frac{f(i)}{1+f(i)} = \frac{1}{1+f(k)} \prod_{i=0}^{k-1} \frac{f(i)}{1+f(i)} = \mu_k.
\end{aligned}$$

In particular this shows that μ defines a probability measure on $\mathbb{N}_0$ and (3.6) gives

$$\mu([k,\infty)) = \prod_{i=0}^{k-1} \frac{f(i)}{1+f(i)} = f(k-1)\mu_{k-1}. \tag{3.7}$$

Using this connection to a jump process we obtain

$$\begin{aligned}
\sum_{k=0}^{\infty} g(k)\mu_k &= \mathbb{E}\big[g\big(N_Y\big)\big] = \int_0^{\infty} \mathbb{E}\big[g\big(N_s\big)\big] e^{-s}\, \mathrm{d}s \\
&= g(0) + \int_0^{\infty} \mathbb{E}\Big[\Big(g\big(N_s+1\big) - g\big(N_s\big)\Big) f\big(N_s\big)\Big] e^{-s}\, \mathrm{d}s \\
&= g(0) + \mathbb{E}\Big[\Big(g\big(N_Y+1\big) - g\big(N_Y\big)\Big) f\big(N_Y\big)\Big],
\end{aligned}$$

where we used an integration by parts formula for the third equality. More precisely, the fact that N_t is an inhomogeneous Poisson point process yields

$$\frac{d}{ds}\mathbb{E}\left[g(N_s)\right] = \mathbb{E}\left[f(N_s)(g(N_s+1) - g(N_s))\right],$$

where we used [Çın11, Theorem 6.11].

So we have shown that if $W \sim \mu$ then $\mathbb{E}\left[\big(g(W+1) - g(W)\big)f(W) + g(0) - g(W)\right] = 0$ for all g such that $\mathbb{E}\left[g(W)\right] < \infty$. Conversely, let

$$\sum_{k \geq 0} \big((g(k+1) - g(k))f(k) + g(0) - g(k)\big)\mathbb{P}(W = k) = 0$$

for all g such that $\sum_{k\geq 0} g(k)\mathbb{P}\left(W=k\right) < \infty$. We now have to show that $W \sim \mu$. Choose a class of functions $(g_i)_{i\geq 1}$ with $g_i(k) := k\mathbb{1}\{k \leq i\}$. Then $\sum_{k\geq 0} g_i(k)\mathbb{P}(W = k) = \sum_{k=0}^{i} k\mathbb{P}(W=k) < \infty$, as for every i, this is a finite sum. Now

$$\begin{aligned}0 = \mathbb{E}\left[g_1(W)\right] &= ((g_1(1)-g_1(0))f(0) + (g_1(0)-g_1(0)))\mathbb{P}(W=0)\\ &\quad + ((g_1(2)-g_1(1))f(1) + (g_1(0)-g_1(1)))\mathbb{P}(W=1)\\ &= f(0)\mathbb{P}(W=0) - (1+f(1))\mathbb{P}(W=1)\end{aligned}$$

so that

$$(1+f(1))\mathbb{P}(W=1) = f(0)\mathbb{P}(W=0). \tag{3.8}$$

Using (3.8) in the analogous calculations for g_2 gives

$$\begin{aligned}&(1+f(2))\mathbb{P}(W=2) = f(1)\mathbb{P}(W=1)\\ &\qquad \Rightarrow \mathbb{P}(W=2) = \frac{f(1)}{1+f(2)}\mathbb{P}(W=1) = \frac{1}{1+f(2)}\frac{f(1)}{1+f(1)}f(0)\mathbb{P}(W=0).\end{aligned}$$

Iterating the whole procedure yields

$$\mathbb{P}(W=k) = \frac{1}{1+f(k)}\prod_{i=1}^{k-1}\frac{f(i)}{1+f(i)}f(0)\mathbb{P}(W=0) = \mu_k\mathbb{P}(W=0)(1+f(0)).$$

Now as $(\mu_k)_{k\geq 0}$ as given in (3.3) is a probability distribution we get

$$\begin{aligned}&\mathbb{P}(W=0) = 1 - (1+f(0))\mathbb{P}(W=0)\sum_{k\geq 1}\mu_k\\ \Leftrightarrow\ &\mathbb{P}(W=0)\Big(1 + (1+f(0))\sum_{k\geq 1}\mu_k\Big) = 1\\ \Leftrightarrow\ &\mathbb{P}(W=0)\Big(1+(1+f(0))(1-\mu_0)\Big) = 1\\ \Leftrightarrow\ &\mathbb{P}(W=0)\left(1+(1+f(0))\frac{f(0)}{1+f(0)}\right) = 1\\ \Leftrightarrow\ &\mathbb{P}(W=0) = \frac{1}{1+f(0)}.\end{aligned}$$

Consequently $\mathbb{P}(W=k) = \mu_k\ \forall k$ and we have thus found the Stein operator

$$\mathcal{A}g(k) = \big(g(k+1)-g(k)\big)f(k) + g(0) - g(k) \tag{3.9}$$

for μ. □

As outlined in section 2.1 the next step is now to solve the so-called Stein equation

$$\mathcal{A}g(k) = h(k) - \int h \, \mathrm{d}\mu \tag{3.10}$$

for all functions in some suitable class of test functions $\mathcal{H}$. As we want to show rates of convergence in total variation norm we are interested in the case where $\mathcal{H}$ is the class of all indicator functions $h_A(\cdot) = \mathbb{1}\{\cdot \in A\}$ with $A \in \mathcal{B}(\mathbb{R})$. After solving (3.10) and taking expectations on both sides, we then get

$$d_{TV}(W, X) = \mathbb{E}\left[\mathcal{A}g_h(X)\right], \tag{3.11}$$

see section 2.1, and thus bounding the right-handside of (3.11) uniformly in A gives bounds on the total variation distance.

3.2.2. Stein solution and bounds

This section is concerned with the second step of Stein's method as described in 2.1.2. It only depends on the operator $\mathcal{A}$ and the class of test functions inducing the probability metric and is thus independent of preferential attachment models. We first derive solutions to the Stein equation (3.11) in the subsequent lemma and consequently deduce smoothness estimates of these in Lemma 3.10.

Lemma 3.9. *The unique solution of the Stein equation for μ, i.e.*

$$\mathcal{A}g = h - \mu(h), \tag{3.12}$$

for any $h(\cdot) = \mathbb{1}\{\cdot \in A\} := \mathbb{1}_A(\cdot)$, with $A \subset \mathbb{N}_0$ is given by

$$g_h(k) = -\int_0^\infty \left(\mathbb{E}_k h(Z_t) - \int h \, \mathrm{d}\mu\right) \mathrm{d}t, \tag{3.13}$$

where (Z_t) is a continuous-time Markov process with generator $\mathcal{A}$ and $\mathbb{E}_k$ denotes the expectation with respect to $\mathbb{P}(\cdot | Z_0 = k)$.

Proof. Let $h = \mathbb{1}_A$ for some $A \subset \mathbb{N}_0$. One can check that the Markov chain (Z_t) with generator $\mathcal{A}$ is irreducible, non-explosive and has invariant distribution μ. It follows that $\mathbb{E}_k h(Z_t) \to \mu(h)$ as $t \to \infty$.

By Kolmogorov's backward equation for the Markov chain $(Z_t)_{t \geq 0}$, see e.g. [Nor98, Thm. 2.8.4.], we have for any $k, j \in \mathbb{N}_0$,

$$\begin{aligned} &\mathbb{P}_k(Z_t = j) - \mathbb{1}_{\{k\}}(j) = \mathbb{P}_k(Z_t = j) - \mathbb{P}_k(Z_0 = j) \\ &= \int_0^t \Big\{ f(k)(\mathbb{P}_{k+1}(Z_s = j) - \mathbb{P}_k(Z_s = j)) + \mathbb{P}_0(Z_s = j) - \mathbb{P}_k(Z_s = j) \Big\} \, \mathrm{d}s. \end{aligned}$$

In particular, as h as above is bounded, we have by Fubini

$$\mathbb{E}_k[h(Z_t)] - h(k) = \int_0^t f(k)(\mathbb{E}_{k+1}[h(Z_s)] - \mathbb{E}_k[h(Z_s)]) + \mathbb{E}_0[h(Z_s)] - \mathbb{E}_k[h(Z_s)] \,\mathrm{d}s.$$

Now, define

$$g_h^t(k) := -\int_0^t (\mathbb{E}_k h(Z_s) - \mu(h)) \,\mathrm{d}s.$$

Thus, we can deduce that

$$\mathbb{E}_k[h(Z_t)] - h(k) = -(\mathcal{A}g_h^t)(k).$$

We note that the left hand side converges to $\mu(h) - h(k)$ as μ is the invariant distribution of $(Z_t)_{t\geq 0}$. By the definition of $\mathcal{A}$ it is straightforward to see that the operator $\mathcal{A}$ is closed under pointwise convergence. Hence, it suffices to show that $g_h^t(k)$ converges to $g_h(k)$ as defined in (3.13) in order to conclude that g_h solves (3.10).

To see the latter, we can estimate since $h = \mathbb{1}_A$,

$$\begin{aligned}\int_0^\infty \left|\mathbb{E}_k[h(Z_t)] - \mu(h)\right| \mathrm{d}t &= \int_0^\infty \left|\mathbb{E}\left[h(Z_t^{(k)})\right] - \mathbb{E}\left[h(Z_t^{(\mu)})\right]\right| \mathrm{d}t \\ &\leq \int_0^\infty \min_{(X_t,Y_t) \text{ coupling of } (Z_t^{(k)},Z_t^{(\mu)})} \mathbb{P}(X_t \neq Y_t)\,\mathrm{d}t,\end{aligned}$$

where we denote by $Z_t^{(k)}$ and $Z_t^{(\mu)}$ the Markov chains started in k and μ respectively. We will now construct a coupling (X_t, Y_t) of $Z_t^{(k)}$ and $Z_t^{(\mu)}$ in the following way: let $X_0 = k$ and choose Y_0 according to μ. We let both chains evolve independently until Y falls to zero at a random time τ. We then force X to fall to zero as well and from that point on the two chains evolve together, so that $X_t = Y_t$ for $t \geq \tau$. One can easily check that this defines a coupling of $Z_t^{(k)}$ and $Z_t^{(\mu)}$. As Y falls down to zero at rate one we get

$$\mathbb{P}(X_t \neq Y_t) \leq \mathbb{P}(\tau \geq t) = e^{-t}$$

so that

$$\int_0^\infty \left|\mathbb{E}_k[h(Z_t)] - \mu(h)\right| \mathrm{d}t \leq \int_0^\infty e^{-t}\,\mathrm{d}t = 1.$$

Hence, $g_h^t(k) \to g_h(k)$ by dominated convergence, which completes the proof of the lemma. □

In this section, we exploit the connection to the Markov process with generator $\mathcal{A}$ to find bounds on the Stein solution $g_A := g_{\mathbb{1}_A}$ for $A \subset \mathbb{N}_0$. As mentioned before, this part of the proof only depends on the limiting distribution μ and does not make use of the preferential attachment setting. More precisely, we will show that $v_A(k) := f(k)g_A(k)$ is uniformly bounded in k. We have

$$\begin{aligned} g_A(k) &= -\int_0^\infty \big(\mathbb{E}\big[\mathbb{1}\{Y_t^{(k)} \in A\}\big] - \mu(A)\big)\, dt \\ &= -\int_0^\infty \Big(\mathbb{E}\Big[\sum_{j\in A} \mathbb{1}\{Y_t^{(k)} = j\}\Big] - \sum_{j\in A}\mu_j\Big)\, dt \\ &= \sum_{j\in A}\Big(-\int_0^\infty \big(\mathbb{E}\big[\mathbb{1}\{Y_t^{(k)} = j\}\big] - \mu_j\big)\, dt\Big) \\ &= \sum_{j\in A} g_{\mathbb{1}\{j\}}(k), \end{aligned} \tag{3.14}$$

thus g_A can be written as a sum of functions of the form $g_{\mathbb{1}\{j\}}$ and we start by calculating the latter.

Lemma 3.10. *For $g_j := g_{\mathbb{1}\{\cdot = j\}}$ as in (3.13) we have*

$$\Delta g_j(k) = g_j(k+1) - g_j(k) = \begin{cases} -\frac{\mu_j}{\mu_k}\frac{1}{f(k)(1+f(k))}, & \text{for } j \geq k+1, \\ \frac{1}{1+f(j)}, & \text{for } j = k, \\ 0, & \text{for } j \leq k-1. \end{cases} \tag{3.15}$$

Proof. We apply the techniques used in [BX01] and adapt them to our Markov process Z_t. Therefore define

$$\tau_{k,k+1} = \inf\{t : Z_t^{(k)} = k+1\},$$

where as before, $Z_t^{(k)}$ denotes a Markov process with generator $\mathcal{A}$ starting in k. Then, for $k \leq j-1$ we obtain via the representation (3.13) of the Stein solution,

$$\begin{aligned} g_j(k) &= -\int_0^\infty \big(\mathbb{E}\big[\mathbb{1}_{\{j\}}(Z_t^{(k)})\big] - \mu_j\big) dt \\ &= -\mathbb{E}\Big[\int_0^{\tau_{k,k+1}} \Big(\mathbb{1}_{\{j\}}(Z_t^{(k)}) - \mu_j\Big) dt\Big] - \mathbb{E}\Big[\int_{\tau_{k,k+1}}^\infty \Big(\mathbb{1}_{\{j\}}(Z_t^{(k)}) - \mu_j\Big) dt\Big] \\ &= \mu_j \mathbb{E}\left[\tau_{k,k+1}\right] + g_j(k+1), \end{aligned}$$

where the last equality uses the strong Markov property of Z_t. Rearranging yields

$$g_j(k) - g_j(k+1) = \mu_j \mathbb{E}\left[\tau_{k,k+1}\right] \geq 0$$

for $k \leq j-1$. Following the same procedure for $k \geq j+1$ and $\tau_{k,0} = \inf\{t : Z_t^{(k)} = 0\}$, we get

$$g_j(k) = -\int_0^\infty \big(\mathbb{E}\big[\mathbb{1}_{\{j\}}(Z_t^{(k)})\big] - \mu_j\big)\, dt$$

$$
\begin{aligned}
&= -\mathbb{E}\Big[\int_0^{\tau_{k,0}} (\mathbb{1}_{\{j\}}(Z_t^{(k)}) - \mu_j)\, dt\Big] - \mathbb{E}\Big[\int_{\tau_{k,0}}^{\infty} (\mathbb{1}_{\{j\}}(Z_t^{(k)}) - \mu_j)\, dt\Big] \\
&= \mu_j \mathbb{E}\left[\tau_{k,0}\right] + g_j(0)
\end{aligned}
$$

and thus

$$
g_j(k) - g_j(0) = \mu_j \mathbb{E}\left[\tau_{k,0}\right]
$$

for $k \geq j+1$. Since the rate by which the process Z_t moves to zero is 1, independently of the current state of the process, we notice that $\mathbb{E}\left[\tau_{k,0}\right] = 1$ for all k and therefore the equation above simplifies to

$$
g_j(k) - g_j(0) = \mu_j.
$$

This means that $g_j(k)$ is constant for $k \geq j+1$, so that

$$
g_j(k+1) - g_j(k) = 0 \text{ for } k \geq j+1.
$$

Furthermore, we have

$$
\begin{aligned}
g_j(j+1) &= -\int_0^{\infty} \big(\mathbb{E}\big[1_{\{j\}}(Z_t^{(j+1)})\big] - \mu_j\big)\, dt \\
&= -\mathbb{E}\Big[\int_0^{\tau_{j+1,j}} (\mathbb{1}_{\{j\}}(Z_t^{(j+1)}) - \mu_j)\, dt\Big] - \mathbb{E}\Big[\int_{\tau_{j+1,j}}^{\infty} (\mathbb{1}_{\{j\}}(Z_t^{(j+1)}) - \mu_j)\, dt\Big] \\
&= \mu_j \mathbb{E}\left[\tau_{j+1,j}\right] + g_j(j) \\
&= \mu_j \big(\mathbb{E}[\tau_{j+1,0}] + \mathbb{E}[\tau_{0,j}]\big) + g_j(j) \\
&= \mu_j \left(1 + \mathbb{E}\left[\tau_{0,j}\right]\right) + g_j(j),
\end{aligned}
$$

yielding

$$
g_j(j+1) - g_j(j) = \mu_j \left(1 + \mathbb{E}\left[\tau_{0,j}\right]\right) = \mu_j \mathbb{E}\left[\tau_{j,j}\right] \geq 0,
$$

where $\tau_{j,j}$ defines the first return time to j of the Markov chain started at j. Thus the following equations hold for the Stein solution g_j:

$$
g_j(k+1) - g_j(k) = \begin{cases} -\mu_j \mathbb{E}\left[\tau_{k,k+1}\right], & \text{for } j \geq k+1, \\ \mu_j \mathbb{E}\left[\tau_{j,j}\right], & \text{for } j = k, \\ 0, & \text{for } j \leq k-1. \end{cases} \tag{3.16}
$$

We can simplify this expression further as follows. Let S_k be the first jump time of $Z_t^{(k)}$. Then, by definition of the Markov chain, we have that $S_k \sim \text{Exp}(1 + f(k))$. Since $Z_t^{(k)}$ jumps to $k+1$ at rate $f(k)$ and to zero at rate 1 we obtain

$$
\mathbb{E}\left[\tau_{k,k+1}\right] = \mathbb{E}[S_k] + \frac{1}{1+f(k)} \mathbb{E}[\tau_{0,k+1}] = \frac{1}{1+f(k)} (1 + \mathbb{E}[\tau_{0,k}] + \mathbb{E}[\tau_{k,k+1}])
$$

and thus

$$\mathbb{E}\left[\tau_{k,k+1}\right] = \frac{1}{f(k)}\big(1+\mathbb{E}\left[\tau_{0,k}\right]\big) = \frac{1}{f(k)}\mathbb{E}\left[\tau_{k,k}\right]. \tag{3.17}$$

The classic theory of Markov chains, see e.g. [Nor98, Thm. 3.6.3], yields that

$$\mu_j = \frac{1}{(1+f(j))\mathbb{E}\left[\tau_{j,j}\right]}. \tag{3.18}$$

Rearranging (3.18) and combining it with (3.16) and (3.17) yields the statement of the lemma. □

Proposition 3.11. *For any $k \in \mathbb{N}_0$ and $A \subset \mathbb{N}_0$, we have*

$$|v_A(k)| \leq 1. \tag{3.19}$$

Proof. By Lemma 3.10 and (3.14) we get

$$v_A(k) = \sum_{j\in A} f(k)\Delta g_j(k) = -\frac{f(k)}{\mu_k f(k)(1+f(k))}\mu(A\cap[k,\infty)) + \frac{f(k)}{1+f(k)}\mathbb{1}\{k\in A\}.$$

Using the identity $\mu([k,\infty)) = \prod_{i=0}^{k-1}\frac{f(i)}{1+f(i)} = (1+f(k))\mu_k$ from (3.7), we obtain

$$v_A(k) = -\frac{1}{\mu([k,\infty))}\mu(A\cap[k,\infty)) + \frac{f(k)}{1+f(k)}\mathbb{1}\{k\in A\},$$

so that the proposition follows immediately. □

3.2.3. Results for general preferential attachment models without loops

We will now derive the claimed error bounds from the right hand side of (2.7). Therefore we use the following dynamic way of generating a uniform random variable on $[n]$ (cf. [For09]): let J_n be a Markov chain with $J_1 = 1$ and such that

$$\mathbb{P}(J_{n+1} = J_n \mid J_n) = \frac{n}{n+1} \quad \text{and} \quad \mathbb{P}(J_{n+1} = n+1 \mid J_n) = \frac{1}{n+1}.$$

Then, we have that J_n is uniformly distributed on $[n]$ for every n (cf. Lemma 3.12). In particular, we know that we can generate the indegree of a uniform vertex as $X_n := \deg_n^-(J_n)$ and moreover, (X_n) turns out to be a Markov chain. In a first step, additionally to Assumptions (**A**) we assume that $d_0 = 0$, and use the above Markov structure to show in Lemma 3.13 that

$$\mathbb{E}\left[\mathcal{A}g_A(X_{n+1})\right] = \frac{1}{n+1}\Big(\sum_{\ell=1}^{n}\sum_{k=0}^{\ell-1}\Delta v_A(k)h(k,\ell) + v_A(0)\Big), \tag{3.20}$$

where

$$v_A(k) := f(k)\Delta g_A(k) \text{ and } h(k,\ell) := f(k)\mathbb{P}(X_\ell = k) - \mathbb{P}(X_\ell \geq k+1).$$

Then, since we already have the smoothness estimate (3.19), it remains to analyse $h(k,\ell)$ and show that these terms are small. The corresponding analysis is carried out in Proposition 3.14. We first show inductively, that under the conditions of Theorems 3.6 and 3.7, for fixed ℓ the functions $k \mapsto h(k,\ell)$ are first increasing and then decreasing, which ultimately allows us to deal with the inner sum over k. Finally, we show that for suitable constants $C > 0$, $k \leq \ell - 1$,

$$h(k,\ell) \leq \begin{cases} C\,\ell^{-1} & \text{under the assumptions of Thm. 3.6,} \\ C\,\ell^{-(1-\gamma)} & \text{under the assumptions of Thm. 3.7.} \end{cases}$$

Via (3.20) these bounds lead to the error bounds of $\log(n)/n$ and $n^{-(1-\gamma)}$ in Theorems 3.6 and 3.7. Throughout the proofs we will always assume that $d_0 = 0$, the case $d_0 > 0$ will be shown using an easy coupling argument.
We use the following Markov chain to describe the evolution of the indegree of a uniform vertex. Similar ideas were also used in [DM09, For09].

Lemma 3.12. *Let X_n be a Markov chain with $\mathbb{P}(X_1 = 0) = 1$ and transition probabilities given for any $i \geq 1$ as*

$$\mathbb{P}(X_{n+1} = j \,|\, X_n = i) = \begin{cases} \frac{f(i)}{n+1} & \text{if } j = i+1, \\ \frac{n-f(i)}{n+1} & \text{if } j = i, \\ \frac{1}{n+1} & \text{if } j = 0 \end{cases}$$

and

$$\mathbb{P}(X_{n+1} = j \,|\, X_n = 0) = \begin{cases} \frac{f(0)}{n+1} & \text{if } j = 1, \\ 1 - \frac{f(0)}{n+1} & \text{if } j = 0. \end{cases}$$

*Then $\mathcal{L}(X_n) = \mathcal{L}(W_n)$, where W_n denotes the indegree of a uniformly chosen vertex in any preferential attachment model at time n satisfying Assumptions **(A)** with $d_0 = 0$.*

Note that the Markov chain starts at time 1 to match the index of the random graph.

Proof. Consider the Markov chain $(J_n)_{n\in\mathbb{N}}$ starting in 1, e.g. $J_1 = 1$, and such that for $n \geq 1$

$$\mathbb{P}(J_{n+1} = J_n | J_n) = \frac{n}{n+1} \quad \text{and} \quad \mathbb{P}(J_{n+1} = n+1 | J_n) = \frac{1}{n+1}.$$

Then it is straightforward to check by induction that J_n is uniformly distributed on $\{1, \ldots, n\}$. We now set $X_n := \deg_n^-(J_n)$, so that in particular it follows $\mathcal{L}(W_n) = \mathcal{L}(X_n)$. Then, using the dynamics of the preferential attachment model and the tower property, the following transition probabilities hold for X_n and $1 \leq j \leq n$:

$$\begin{aligned}\mathbb{P}(X_{n+1} = j+1 | X_n = j) &= \frac{f(j)}{n} \cdot \frac{n}{n+1} = \frac{f(j)}{n+1},\\ \mathbb{P}(X_{n+1} = j | X_n = j) &= (1 - \frac{f(j)}{n}) \cdot \frac{n}{n+1} = \frac{n - f(j)}{n+1},\\ \mathbb{P}(X_{n+1} = 0 | X_n = j) &= \frac{1}{n+1}.\end{aligned}$$

Moreover, for $j = 0$,

$$\begin{aligned}\mathbb{P}(X_{n+1} = 1 | X_n = 0) &= \frac{f(0)}{n} \cdot \frac{n}{n+1} = \frac{f(0)}{n+1},\\ \mathbb{P}(X_{n+1} = 0 | X_n = 0) &= (1 - \frac{f(0)}{n}) \cdot \frac{n}{n+1} + \frac{1}{n+1} = 1 - \frac{f(0)}{n+1}.\end{aligned}$$

This completes the proof of the lemma. □

Following the discussion in section 2.1.2, the next step is to find a bound on

$$\begin{aligned}\mathbb{E}\left[\mathcal{A}g_A(W_{n+1})\right] &= \mathbb{E}\left[\mathcal{A}g_A(X_{n+1})\right]\\ &= \mathbb{E}\left[f(X_{n+1})\Delta g_A(X_{n+1}) + (g_A(0) - g_A(X_{n+1}))\right],\end{aligned}$$

where we recall that $\Delta g_A(k) := g_A(k+1) - g_A(k)$.

Lemma 3.13. *For the Markov chain $(X_n)_{n\geq 0}$ defined in Lemma 3.12 and $v_A(k) := f(k)\Delta g_A(k)$, with $A \subset \mathbb{N}_0$, we have*

$$\begin{aligned}&\mathbb{E}\left[\mathcal{A}g_A(X_{n+1})\right]\\ &= \frac{1}{n+1}\Big(\Big(\sum_{\ell=1}^{n}\sum_{k=0}^{\ell-1}\Delta v_A(k)\big(f(k)\mathbb{P}(X_\ell = k) - \mathbb{P}(X_\ell \geq k+1)\big)\Big) + v_A(0)\Big). \quad (3.21)\end{aligned}$$

Proof. Let $h : \mathbb{N}_0 \to \mathbb{R}$ be such that $h(0) = 0$. Then,

$$\begin{aligned}\mathbb{E}\left[h(X_{n+1})\right] &= \mathbb{E}\left[\mathbb{E}\left[h(X_{n+1}) \,|\, X_n\right]\right]\\ &= \mathbb{E}\Big[h(X_n)\,\frac{n - f(X_n)}{n+1} + h(X_n+1)\,\frac{f(X_n)}{n+1}\Big]\\ &= \frac{n}{n+1}\mathbb{E}[h(X_n)] + \frac{1}{n+1}\mathbb{E}[f(X_n)\Delta h(X_n)].\end{aligned}$$

Using the fact that $X_1 = 0$, we iteratively obtain

$$\mathbb{E}\left[h(X_{n+1})\right] = \frac{1}{n+1}\sum_{\ell=1}^{n}\mathbb{E}[f(X_\ell)\Delta h(X_\ell)].$$

Define

$$h(k) = \mathcal{A}g_A(k) - v_A(0) = v_A(k) + (g_A(0) - g_A(k)) - v_A(0).$$

Thus, we get

$$\begin{aligned}
\mathbb{E}[\mathcal{A}g_A(X_{n+1})] &= \mathbb{E}[h(X_{n+1})] + v_A(0) \\
&= \frac{1}{n+1}\sum_{\ell=1}^{n}\Big(\mathbb{E}[f(X_\ell)\Delta v_A(X_\ell)] - \mathbb{E}[f(X_\ell)\Delta g_A(X_\ell)]\Big) + v_A(0) \\
&= \frac{1}{n+1}\sum_{\ell=1}^{n}\sum_{k=0}^{\ell-1} f(k)\Delta v_A(k)\mathbb{P}(X_\ell = k) \\
&\quad - \frac{1}{n+1}\sum_{\ell=1}^{n}\sum_{k=0}^{\ell-1}(v_A(k) - v_A(0))\mathbb{P}(X_\ell = k) + \frac{v_A(0)}{n+1}.
\end{aligned} \tag{3.22}$$

For the second sum, we write

$$\begin{aligned}
\sum_{k=0}^{\ell-1}(v_A(k) - v_A(0))\mathbb{P}(X_\ell = k) &= \sum_{k=0}^{\ell-1}\sum_{i=0}^{k-1}\Delta v_A(i)\mathbb{P}(X_\ell = k) \\
&= \sum_{i=0}^{\ell-1}\Delta v_A(i)\sum_{k=i+1}^{\ell-1}\mathbb{P}(X_\ell = k) = \sum_{i=0}^{\ell-1}\Delta v_A(i)\mathbb{P}(X_\ell \geq i+1).
\end{aligned}$$

Combining the latter with (3.22) yields the statement of the lemma. □

Now, the next proposition gives the desired results on $h(k,l)$ as mentioned above.

Proposition 3.14. *Suppose f satisfies $f(k) \leq k+1$. Define*

$$h(k,\ell) := f(k)\mathbb{P}(X_\ell = k) - \mathbb{P}(X_\ell \geq k+1),$$

where $(X_\ell)_{\ell\geq 1}$ is the Markov chain from Lemma 3.12.

(i) Then, for any $k \in \mathbb{N}_0, \ell \in \mathbb{N}$ we have

$$h(k,\ell) \geq 0,$$

and moreover for $k \geq \ell$, we have $h(k,\ell) = 0$.

(ii) *Suppose there exists K such that $k \le f(k)$ for all $0 \le k \le K$, then we have*

$$\Delta^{(1)} h(k,\ell) := h(k+1,\ell) - h(k,\ell) \ge 0 \tag{3.23}$$

for all $\ell \le K+1,\ k \le \ell - 2$.

(iii) *Assume there exist $k_* \in \mathbb{N}_0$ such that $f(k) \le k$ for all $k \ge k_*$ and $f(k) > k$ for $k < k_*$. Then, for all $\ell \in \mathbb{N}$ there exists $I(\ell) \in \{0, \dots, \ell-1\}$ such that*

$$h(k+1,\ell) - h(k,\ell) \begin{cases} \ge 0 & \text{if } k < I(\ell), \\ \le 0 & \text{if } k \ge I(\ell). \end{cases} \tag{3.24}$$

Moreover, $I(\ell+1) \in \{I(\ell), I(\ell)+1\}$.

(iv) *Assume there exists $k_* \in \mathbb{N}_0$ such that $f(k) \le k$ for all $k \ge k_*$. Then, there exists a constant $C > 0$ such that for all $k \in \mathbb{N}_0, \ell \in \mathbb{N}$,*

$$h(k,\ell) \le \frac{C}{\ell}.$$

(v) *If there exists $\gamma \in (0,1)$ such that $f(k) \in [k, k+\gamma]$ for all $k \in \mathbb{N}_0$, then*

$$\sup_{k \in \mathbb{N}_0} h(k,\ell) \le C \ell^{-(1-\gamma)}.$$

Before we start with the proof of the proposition we derive a recursive formula for the coefficients h in the ℓ coordinate as well as for its increments in the k coordinate.

Lemma 3.15. *Let h be defined as in Proposition 3.14, then $h(k,\ell) = 0$ for all $k \ge \ell$ and for all $\ell \in \mathbb{N}$, $k \in \mathbb{N}_0$, we have*

$$h(k,\ell+1) = \left(\frac{\ell}{\ell+1} - \frac{f(k)}{\ell+1} \right) h(k,\ell) + \frac{f(k)}{\ell+1} h(k-1,\ell), \tag{3.25}$$

where we define $h(-1,\ell) = 0$. Moreover, if we define $\Delta^{(1)} h(k,\ell) := h(k+1,\ell) - h(k,\ell)$, we have that for all $\ell \in \mathbb{N}$ and $k \le \ell - 1$

$$\Delta^{(1)} h(k,\ell+1) = \left(\frac{\ell}{\ell+1} - \frac{f(k+1)}{\ell+1} \right) \Delta^{(1)} h(k,\ell) + \frac{f(k)}{\ell+1} \Delta^{(1)} h(k-1,\ell). \tag{3.26}$$

Proof. Note that since $X_\ell \le \ell - 1$ $\mathbb{P}$-a.s, we have $h(k,\ell) = 0$ for any $k \ge \ell$. Moreover, by the definition of the Markov chain (X_n), for $k \ge 1$,

$$\begin{aligned} h(k,\ell+1) &= f(k)\mathbb{P}(X_{\ell+1} = k) - \mathbb{P}(X_{\ell+1} \ge k+1) \\ &= f(k)\Big(\Big(1 - \frac{f(k)+1}{\ell+1}\Big) \mathbb{P}(X_\ell = k) + \frac{f(k-1)}{\ell+1} \mathbb{P}(X_\ell = k-1) \Big) \end{aligned}$$

$$
\begin{aligned}
&\quad - \Big(\frac{\ell}{\ell+1}\mathbb{P}(X_\ell \geq k+1) + \frac{f(k)}{\ell+1}\mathbb{P}(X_\ell = k)\Big) \\
&= \left(\frac{\ell}{\ell+1} - \frac{f(k)}{\ell+1}\right) h(k,\ell) - \frac{f(k)}{\ell+1}\mathbb{P}(X_\ell \geq k+1) \\
&\quad + \frac{f(k)}{\ell+1}\big(f(k-1)\mathbb{P}(X_\ell = k-1) - \mathbb{P}(X_\ell = k)\big) \\
&= \left(\frac{\ell}{\ell+1} - \frac{f(k)}{\ell+1}\right) h(k,\ell) + \frac{f(k)}{\ell+1} h(k-1,\ell).
\end{aligned}
$$

Note, for $k = 0$, we have

$$
\begin{aligned}
h(0,\ell+1) &= f(0)\mathbb{P}(X_{\ell+1} = 0) - \mathbb{P}(X_{\ell+1} \geq 1) \\
&= f(0)\Big(\Big(1 - \frac{f(0)}{\ell+1}\Big)\mathbb{P}(X_\ell = 0) + \frac{1}{\ell+1}\mathbb{P}(X_\ell \geq 1)\Big) \\
&\qquad - \frac{f(0)}{\ell+1}\mathbb{P}(X_\ell = 0) - \Big(1 - \frac{1}{\ell+1}\Big)\mathbb{P}(X_\ell \geq 1) \\
&= \Big(1 - \frac{f(0)+1}{\ell+1}\Big)h(0,\ell) = \Big(\frac{\ell}{\ell+1} - \frac{f(0)}{\ell+1}\Big)h(0,\ell).
\end{aligned}
$$

Therefore, the identity (3.25) also holds for $k = 0$ since we defined $h(-1,\ell) = 0$ for all $\ell \in \mathbb{N}$. By (3.25)

$$
\begin{aligned}
&h(k+1,l+1) - h(k,l+1) \\
&= \left(\frac{\ell}{\ell+1} - \frac{f(k+1)}{\ell+1}\right) h(k+1,\ell) + \frac{f(k+1)}{\ell+1} h(k,\ell) \\
&\quad - \left(\frac{\ell}{\ell+1} - \frac{f(k)}{\ell+1}\right) h(k,\ell) - \frac{f(k)}{\ell+1} h(k-1,\ell) \\
&= \left(\frac{\ell}{\ell+1} - \frac{f(k+1)}{\ell+1}\right) \left(h(k+1,\ell) - h(k,\ell)\right) + \frac{f(k)}{\ell+1} \left(h(k,\ell) - h(k-1,\ell)\right),
\end{aligned}
$$

which proves (3.26). □

Proof of Proposition 3.14. Before we start with the proof, note that for $\ell = 2$, we have

$$
\begin{aligned}
h(0,2) &= f(0)\mathbb{P}(X_2 = 0) - \mathbb{P}(X_2 \geq 1) \\
&= f(0)\left(1 - \frac{f(0)}{2}\right) - \frac{f(0)}{2} = \frac{f(0)(1-f(0))}{2}.
\end{aligned} \tag{3.27}
$$

Moreover,

$$
h(1,2) = f(1)\mathbb{P}(X_2 = 1) = \frac{f(1)f(0)}{2}. \tag{3.28}
$$

(i) We now show that for any $\ell \in \mathbb{N}$: $h(k,\ell) \geq 0$ for any $k \in \mathbb{N}_0$ by induction on ℓ. Note that for $\ell = 1$ the statement holds trivially and for $\ell = 2$ the base case follows from (3.27) and (3.28) since $f(0) \le 1$. We now assume that the statement holds for some ℓ, then from (3.25) and for $k \le \ell - 1$ we have

$$h(k, \ell+1) = \Big(\frac{\ell}{\ell+1} - \frac{f(k)}{\ell+1}\Big) h(k,\ell) + \frac{f(k)}{\ell+1} h(k-1,\ell),$$

which is nonnegative due to the induction hypothesis and the condition that $f(k) \le k+1 \le \ell$. For $k = \ell$ we have $h(\ell, \ell+1) = f(\ell)\mathbb{P}(X_{\ell+1}) \ge 0$. This implies the induction step since all other terms are 0.

(ii) Now suppose that there exists K such that $k \le f(k) \le k+1$ for all $k \le K$. As before we will use induction on ℓ to show the stated result. For $\ell = 2$ we get from (3.27) and (3.28)

$$h(1,2) - h(0,2) = \frac{f(0)}{2}(f(0) + f(1) - 1) \ge 0,$$

as $f(0) \ge 0$ and $f(1) \ge 1$ by assumption.

Suppose that statement (3.23) is true for some $\ell \le K$. By (3.26) we obtain

$$\Delta^{(1)} h(k, \ell+1) = \left(\frac{\ell}{\ell+1} - \frac{f(k+1)}{\ell+1}\right) \Delta^{(1)} h(k,\ell) + \frac{f(k)}{\ell+1} \Delta^{(1)} h(k-1,\ell),$$

which is nonnegative for $k \le \ell - 2$ due to the induction hypothesis and since

$$f(k) \le f(k+1) \le f(\ell-1) \le \ell \le \ell+1.$$

It remains to show that $\Delta^{(1)} h(\ell-1, \ell+1) \ge 0$. Again by (3.26) and using that $\Delta^{(1)} h(\ell-1,\ell) = -h(\ell-1,\ell)$, we get that

$$\begin{aligned}\Delta^{(1)} h(\ell-1, \ell+1) &= \left(\frac{\ell}{\ell+1} - \frac{f(\ell)}{\ell+1}\right)(-h(\ell-1,\ell)) + \frac{f(\ell-1)}{\ell+1} \Delta^{(1)} h(\ell-2,\ell) \\ &= \frac{f(\ell)-\ell}{\ell+1} h(\ell-1,\ell) + \frac{f(\ell-1)}{\ell+1} \Delta^{(1)} h(\ell-2,\ell),\end{aligned}$$

which is nonnegative by induction hypothesis and since $f(\ell) \ge \ell$.

(iii) Again we show by induction on ℓ that there exists $I(\ell) \in \{0, \dots, \ell-1\}$ such that (3.24) is valid. Note that for $\ell = 2$ the statement holds trivially. Moreover, for $\ell \le k_*$ the statement holds by (ii) with $I(\ell) = \ell - 1$. Suppose statement (3.24) is true for some $\ell \ge k_*$. By (3.26) we obtain

$$\Delta^{(1)} h(k, \ell+1) = \left(\frac{\ell}{\ell+1} - \frac{f(k+1)}{\ell+1}\right) \Delta^{(1)} h(k,\ell) + \frac{f(k)}{\ell+1} \Delta^{(1)} h(k-1,\ell).$$

From this we can deduce that if $k < I(\ell)$, then since $I(\ell) \le \ell - 1$, we have $f(k+1) \le k+2 \le \ell$, so that $\Delta^{(1)} h(k, \ell+1) \ge 0$. Conversely, if $k > I(\ell)$ and

$k \le \ell - 2$, we get by a similar argument that $\Delta^{(1)} h(k, \ell+1) \le 0$. Note the case $k = \ell$ holds since $\Delta^{(1)} h(\ell, \ell+1) = -h(\ell, \ell+1)$.

By the recursion (3.25) together with (3.28), we get

$$h(\ell, \ell+1) = \frac{f(\ell)}{\ell+1} h(\ell-1, \ell) = f(\ell) \prod_{i=1}^{\ell} \frac{f(i-1)}{i+1}. \tag{3.29}$$

It remains to show that $\Delta h^{(1)}(\ell-1, \ell+1) \le 0$. Since $\Delta^{(1)} h(\ell-2, \ell) \le 0$, we have again by (3.25),

$$\begin{aligned}
&\Delta^{(1)} h(\ell-1, \ell+1) \\
&\quad = h(\ell, \ell+1) - \Big(\frac{\ell}{\ell+1} - \frac{f(\ell-1)}{\ell+1} \Big) h(\ell-1, \ell) - \frac{f(\ell-1)}{\ell+1} h(\ell-2, \ell) \\
&\quad \le h(\ell, \ell+1) - \frac{\ell}{\ell+1} h(\ell-1, \ell) \\
&\quad = f(\ell) \prod_{j=1}^{\ell} \frac{f(j-1)}{j+1} - \frac{\ell}{\ell+1} f(\ell-1) \prod_{j=1}^{\ell-1} \frac{f(j-1)}{j+1} \\
&\quad = \frac{f(\ell-1)}{\ell+1} (f(\ell) - \ell) \prod_{j=1}^{\ell-1} \frac{f(j-1)}{j+1},
\end{aligned}$$

which is negative as $f(\ell) \le \ell$ since $\ell \ge k_*$. In particular, we have seen that $I(\ell+1) \in \{I(\ell), I(\ell)+1\}$.

(iv) Define

$$C := f(0) \Big(1 \vee \max_{1 \le k \le k^*} \prod_{i=1}^{k} \frac{f(i)}{i} \Big).$$

Then, we will show inductively that for all $\ell \in \mathbb{N}$ and $k \le \ell - 1$:

$$h(k, \ell) \le \frac{C}{\ell}. \tag{3.30}$$

For $\ell = 1$ we have

$$h(0,1) = f(0) \mathbb{P}(X_1 = 0) = f(0) \le C.$$

Now, assume that (3.30) holds for some $\ell \in \mathbb{N}$. Using the identity (3.25) we obtain for $k \le \ell - 1$

$$\begin{aligned}
h(k, \ell+1) &= \left(\frac{\ell}{\ell+1} - \frac{f(k)}{\ell+1} \right) h(k, \ell) + \frac{f(k)}{\ell+1} h(k-1, \ell) \\
&\le \left(\frac{\ell}{\ell+1} - \frac{f(k)}{\ell+1} \right) \frac{C}{\ell} + \frac{f(k)}{\ell+1} \frac{C}{\ell} = \frac{C}{\ell+1},
\end{aligned}$$

where we used that $f(k) \leq k+1 \leq \ell$ in the second step. For $k = \ell$, we have by (3.29)

$$h(\ell, \ell+1) = f(\ell) \prod_{i=1}^{\ell} \frac{f(i-1)}{i+1} = \frac{f(0)}{\ell+1} \prod_{i=1}^{\ell} \frac{f(i)}{i}.$$

Then, if $\ell \leq k^*$, this is trivially bounded by $C/(\ell+1)$. Furthermore, if $\ell > k^*$, then

$$h(\ell, \ell+1) = \frac{f(0)}{\ell+1} \prod_{i=1}^{\ell} \frac{f(i)}{i} \leq \frac{C}{\ell+1} \prod_{i=k^*+1}^{k} \frac{f(i)}{i} \leq \frac{C}{\ell+1},$$

since $f(i) \leq i$ for all $i \geq k^*$. This completes the induction step.

(v) Note that by (ii), $k \mapsto h(k, \ell)$ is increasing for $k \leq \ell - 1$. In particular, from (i) it follows that

$$\sup_{k \in \mathbb{N}_0} h(k, \ell) = \sup_{k \leq \ell-1} h(k, \ell) = h(\ell-1, \ell).$$

By (3.29), we get that

$$h(\ell-1, \ell) = f(\ell-1) \prod_{i=1}^{\ell-1} \frac{f(i-1)}{i+1} \leq \frac{\prod_{i=0}^{\ell-1}(i+\gamma)}{\ell!} = \frac{1}{\Gamma(\gamma)} \frac{\Gamma(\ell+\gamma)}{\Gamma(\ell+1)} \sim \frac{1}{\Gamma(\gamma)} \ell^{\gamma-1},$$

using the asymptotics of the Gamma function. This immediately gives statement (v).

□

We can now combine our previous estimates to prove the two main theorems simultaneously.

Proofs of Theorems 3.6 and 3.7. We first consider the case in which the preferential attachment model satisfies Assumption **(A)** with $d_0 = 0$, so that we can generate the indegree of a uniform vertex using the Markov chain $(X_\ell)_{\ell \geq 1}$ defined in Lemma 3.12. Using the notation $h(k, \ell) = f(k)\mathbb{P}(X_\ell = k) - \mathbb{P}(X_\ell \geq k+1)$, we have from Lemma 3.13, that for any $A \subset \mathbb{N}_0$,

$$\mathbb{E}[\mathcal{A} g_A(X_{n+1})] = \frac{1}{n+1} \Big(\Big(\sum_{\ell=1}^{n} \sum_{k=0}^{\ell-1} \Delta v_A(k) h(k, \ell) \Big) + v_A(0) \Big).$$

Using a discrete integration by parts formula and the fact that $h(\ell, \ell) = 0$, we can rewrite the inner sum as

$$\sum_{k=0}^{\ell-1} \Delta v_A(k) h(k, \ell) = -v_A(0) h(0, \ell) - \sum_{k=0}^{\ell-1} v_A(k+1) \Delta^{(1)} h(k, \ell).$$

Under the assumptions of Theorem 3.6 and 3.7, respectively, there exists $I(\ell)$ such that $\Delta^{(1)}h(k,\ell) \geq 0$ for $k < I(\ell)$ and $\Delta^{(1)}h(k,\ell) \leq 0$ for $k \geq I(\ell)$. In the case of Theorem 3.6 this follows from Proposition 3.14 (iii), for Theorem 3.7 we used part (ii) of Proposition 3.14.

In particular, we have that

$$\begin{aligned}
&\Big|\sum_{k=0}^{\ell-1} v_A(k+1)\Delta^{(1)}h(k,\ell)\Big| \\
&\qquad \leq \sup_k |v_A(k)| \Big(h(I(\ell),\ell) - h(0,\ell) + h(I(\ell),\ell) - h(\ell,\ell)\Big) \\
&\qquad \leq 2\sup_k |v_A(k)| \sup_{k\leq \ell-1} h(k,\ell).
\end{aligned}$$

Therefore, we get

$$\begin{aligned}
&\big|\mathbb{E}[\mathcal{A}g_A(X_{n+1})]\big| \\
&\quad \leq \frac{|v_A(0)|}{n+1} + \Big|\frac{1}{n+1}\sum_{\ell=1}^{n}\Big(\Big(\sum_{k=0}^{\ell-1} v_A(k+1)\Delta^{(1)}h(k,\ell)\Big) + v_A(0)h(0,\ell)\Big)\Big| \\
&\quad \leq \frac{|v_A(0)|}{n+1} + \frac{2}{n+1}\sup_k |v_A(k)| \sum_{\ell=1}^{n} \sup_{k\leq \ell-1} h(k,\ell).
\end{aligned}$$

Hence, if we combine this estimate with Proposition 3.11, we obtain that

$$d_{\mathrm{TV}}(W_{n+1},W) = \sup_{A\subset\mathbb{N}_0} \big|\mathbb{E}[\mathcal{A}g_A(X_{n+1})]\big| \leq \frac{1}{n+1} + \frac{2}{n+1}\sum_{\ell=1}^{n}\sup_{k\leq \ell-1} h(k,\ell).$$

Finally, we note that in the case of Theorem 3.6 we can apply Proposition 3.14 (iv) to deduce that there exists a constant $C > 0$ such that

$$d_{\mathrm{TV}}(W_{n+1},W) \leq \frac{1}{n+1} + \frac{2\,C}{n+1}\sum_{\ell=1}^{n}\frac{1}{\ell},$$

which immediately produces the required bound. In the case of Theorem 3.7, we can instead apply Proposition 3.14 (v) to get a constant $C > 0$ such that

$$d_{\mathrm{TV}}(W_{n+1},W) \leq \frac{1}{n+1} + \frac{2\,C}{n+1}\sum_{\ell=1}^{n}\ell^{-(1-\gamma)},$$

which again yields the statement of the theorem. Finally, we consider the case in which the model satisfies Assumptions **(A)** with $d_0 > 0$. In this case, by the same argument as in Lemma 3.12, the indegree of a uniformly chosen vertex has the same distribution as a Markov chain $(\tilde{X}_n)_{n\geq 1}$ with $\tilde{X}_1 = d_0$, but the same transition probabilities as $(X_n)_{n\geq 1}$. Let $\tau = \inf\{k \geq 2 : \tilde{X}_k = 0\}$. We can couple $(X_n), (\tilde{X}_n)$ by first letting $\tilde{X}_n$ evolve and then letting X_n evolve independently until time τ. Further, we set $X_k := \tilde{X}_k$ for all $k \geq \tau$. By the characterization of d_{TV} in terms of couplings, we thus have

$$d_{\mathrm{TV}}(X_n, \tilde{X}_n) \leq \mathbb{P}(\tau > n) = \prod_{i=1}^{n} \Big(1 - \frac{1}{i+1}\Big) = \frac{1}{n+1}.$$

By the first part of the proof for $d_0 = 0$, this completes the proof also for $d_0 > 0$. □

3.2.4. Results for the Barabási-Albert model

It is straightforward to see that graph models allowing for self-loops do not satisfy Assumptions **(A)**. Unfortunately this class of graph models includes the probably most prominent model appearing in the literature. It was first introduced by Barabási and Albert in [BA99]. A generalized version of this model can be described as follows: we start with a graph consisting of one vertex and a single self-loop. Now at each discrete time step n we insert a new vertex, which we label n, together with a single edge, which connects to one of the present vertices according to the following probabilities

$$\begin{aligned} \mathbb{P}(n+1 \to n+1 | PA_n^{1,\delta}) &= \frac{1+\delta}{n(2+\delta)+(1+\delta)} = \frac{f(0)}{n+\alpha}, \\ \mathbb{P}(n+1 \to i | PA_n^{1,\delta}) &= \frac{\delta + D_n(i)}{n(2+\delta)+(1+\delta)} = \frac{f(\deg_n^-(i))}{n+\alpha} \text{ for } i \leq n, \end{aligned} \tag{3.31}$$

where $f(k) = \frac{k}{2+\delta} + \alpha$, with $\alpha = \frac{1+\delta}{2+\delta}$. Here $D_n(i)$ denotes the total degree of vertex i, $\deg_n^-(i)$ refers to the indegree of vertex i a time n, $PA_n^{1,\delta}$ denotes the graph at time n and $\delta \geq -1$ is a parameter of the model, which, for $\delta = 0$, yields the classical Barabási-Albert model (cf. [BA99]). In [Ros13] the author deduces rates of convergence for the degree of a uniformly chosen vertex to a mixed binomial distribution via Stein's method. Theorem 3.16 gives the same rates of convergence towards the limiting distribution μ, however, though we also use Stein's method, the proofs are fundamentally different.

Theorem 3.16. *Let W_n denote the indegree of a uniformly chosen vertex at time n in the (generalized) Barabási-Albert model. Then there exists a constant $C > 0$ such that for all $n \geq 2$*

$$d_{\mathrm{TV}}(W_n, W) \leq C\, \frac{\log(n)}{n}, \tag{3.32}$$

where $W \sim \mu$ and μ as in (3.3).

The proof follows the same structure as those of Theorems 3.6 and 3.7. Most of the calculations are very similar, but as the Markov process can now reach states 0 and 1 from any other state, some additional terms turn up and we have to alter the statements of Proposition 3.14 to some extent.

Let J_n be defined as in Lemma 3.12 and set $Y_n = \deg_n^{BA,-}(J_n) := \deg_n^-(J_n)$, where $\deg_n^{BA,-}(i)$ refers to the indegree of vertex i in the Barabási-Albert model at time n. We get the following transition probabilities for the Markov chain Y_n:

$$\begin{aligned}
\mathbb{P}(Y_{n+1} = j+1|Y_n = j) &= \frac{n}{n+1} \cdot \frac{f(j)}{n+\alpha} \quad \text{for } j \neq 0,\\
\mathbb{P}(Y_{n+1} = 1|Y_n = 0) &= \frac{f(0)}{n+\alpha},\\
\mathbb{P}(Y_{n+1} = j|Y_n = j) &= \frac{n}{n+1}\left(1 - \frac{f(j)}{n+\alpha}\right) \quad \text{for } j \geq 2,\\
\mathbb{P}(Y_{n+1} = 1|Y_n = 1) &= \frac{n}{n+1}\left(1 - \frac{f(1)}{n+\alpha}\right) + \frac{1}{n+1} \cdot \frac{f(0)}{n+\alpha},\\
\mathbb{P}(Y_{n+1} = 0|Y_n = 0) &= \frac{n}{n+1}\left(1 - \frac{f(0)}{n+\alpha}\right) + \frac{1}{n+1}\left(1 - \frac{f(0)}{n+\alpha}\right) = 1 - \frac{f(0)}{n+\alpha},\\
\mathbb{P}(Y_{n+1} = 0|Y_n = j) &= \frac{1}{n+1}\left(1 - \frac{f(0)}{n+\alpha}\right) \quad \text{for } j \neq 0,\\
\mathbb{P}(Y_{n+1} = 1|Y_n = j) &= \frac{1}{n+1} \cdot \frac{f(0)}{n+\alpha} \quad \text{for } j \geq 2.
\end{aligned}$$

Lemma 3.17. *For the Markov chain $(Y_n)_{n\geq 0}$ defined above and $v_A(k) := f(k)\Delta g_A(k)$, with $A \subset \mathbb{N}_0$, we have*

$$\begin{aligned}
&|\mathbb{E}\left[\mathcal{A}g_A(Y_{n+1})\right]|\\
&\leq \left|\frac{1}{n+1}\Big(\Big(\sum_{\ell=1}^{n}\sum_{k=0}^{\ell-1} \Delta v_A(k)\big(f(k)\mathbb{P}(Y_\ell = k) - \mathbb{P}(Y_\ell \geq k+1)\big)\Big)\Big)\right| + C\,\frac{\log(n)}{n},
\end{aligned} \tag{3.33}$$

for some constant $C > 0$.

Note that throughout this section C denotes a constant, which may vary from line to line, but is always independent of n.

Proof. As in the proof of Lemma 3.13 let $h : \mathbb{N}_0 \to \mathbb{R}$ be such that $h(0) = 0$. Then,

$$\begin{aligned}
\mathbb{E}\left[h(Y_{n+1})\right] &= \mathbb{E}\left[\mathbb{E}\left[h(Y_{n+1}) \,|\, Y_n\right]\right] \\
&= \mathbb{E}\Big[h(Y_n)\,\frac{n}{n+1}\left(1 - \frac{f(Y_n)}{n+\alpha}\right) + h(Y_n+1)\,\frac{n}{n+1}\frac{f(Y_n)}{n+\alpha} + h(1)\frac{f(0)}{(n+1)(n+\alpha)}\Big] \\
&= \frac{n}{n+1}\mathbb{E}[h(Y_n)] + \frac{1}{n+1}\mathbb{E}\left[\frac{n}{n+\alpha} f(Y_n)\Delta h(Y_n)\right] + \frac{f(0)}{(n+1)(n+\alpha)} h(1).
\end{aligned}$$

Through iteration, we get

$$\mathbb{E}\left[h(Y_{n+1})\right] = \frac{1}{n+1}\mathbb{E}\left[h(Y_1)\right] + \frac{1}{n+1}\sum_{\ell=1}^{n}\frac{\ell}{\ell+\alpha}\mathbb{E}[f(Y_\ell)\Delta h(Y_\ell)] + \frac{1}{n+1}\sum_{\ell=1}^{n}\frac{f(0)}{l+\alpha}h(1).$$

Again we define

$$h(k) = v_A(k) + (g_A(0) - g_A(k)) - v_A(0).$$

Using (3.11) the absolute value of the first term can be bounded by $\frac{3}{n+1}$ and the last one by $\frac{3\log(n)}{n}$. Having taken care of these two terms we obtain

$$\begin{aligned}
&|\mathbb{E}[\mathcal{A}g_A(Y_{n+1})]| = |\mathbb{E}[h(Y_{n+1}) + v_A(0)]| \\
&\quad\le \left|\frac{1}{n+1}\sum_{\ell=1}^{n}\frac{\ell}{\ell+\alpha}\Big(\mathbb{E}[f(Y_\ell)\Delta v_A(Y_\ell)] - \mathbb{E}[f(Y_\ell)\Delta g_A(Y_\ell)]\Big) + v_A(0)\right| + C\,\frac{\log(n)}{n} \\
&\quad\le \left|\frac{1}{n+1}\sum_{\ell=1}^{n}\Big(\mathbb{E}[f(Y_\ell)\Delta v_A(Y_\ell)] - \mathbb{E}[f(Y_\ell)\Delta g_A(Y_\ell)]\Big)\right| \\
&\quad+ \left|\frac{1}{n+1}\sum_{\ell=1}^{n}\frac{\alpha}{\ell+\alpha}\Big(\mathbb{E}[f(Y_\ell)\Delta v_A(Y_\ell)] - \mathbb{E}[f(Y_\ell)\Delta g_A(Y_\ell)]\Big) + v_A(0)\right| + C\,\frac{\log(n)}{n} \\
&\quad\le \left|\frac{1}{n+1}\sum_{\ell=1}^{n}\sum_{k=0}^{\ell} f(k)\Delta v_A(k)\mathbb{P}(Y_\ell = k) - \frac{1}{n+1}\sum_{\ell=1}^{n}\sum_{k=0}^{\ell}(v_A(k) - v_A(0))\mathbb{P}(Y_\ell = k)\right| \\
&\qquad+ \frac{|v_A(0)|}{n+1} + C\,\frac{1}{n+1}\sum_{\ell=1}^{n}\frac{\alpha}{\ell+\alpha}\mathbb{E}[f(Y_\ell)] + C\,\frac{\log(n)}{n}.
\end{aligned} \tag{3.34}$$

Now, since $Y_\ell \overset{d}{=} \deg_\ell^-(U_\ell)$, where U_ℓ denotes the uniform distribution on $[\ell]$, we get

$$\mathbb{E}\left[f(Y_\ell)\right] = \mathbb{E}\left[\frac{1}{\ell}\sum_{j=1}^{\ell}\left(\frac{\deg_\ell^-(j)}{2+\delta} + \alpha\right)\right] = \frac{1}{\ell(2+\delta)}(\ell + \ell(1+\delta)) = 1,$$

so that the last terms can all be bounded by $C\,\frac{\log(n)}{n}$. Moreover

$$\sum_{k=0}^{\ell}(v_A(k) - v_A(0))\mathbb{P}(Y_\ell = k) = \sum_{k=0}^{\ell}\sum_{i=0}^{k-1}\Delta v_A(i)\mathbb{P}(Y_\ell = k)$$

$$= \sum_{i=0}^{\ell-1} \Delta v_A(i) \sum_{k=i+1}^{\ell} \mathbb{P}(Y_\ell = k) = \sum_{i=0}^{\ell-1} \Delta v_A(i) \mathbb{P}(Y_\ell \geq i+1).$$

Plugging all this into (3.34) we obtain the desired statement. □

As in section 3.2.3 we define

$$h(k,\ell) := f(k)\mathbb{P}(Y_\ell = k) - \mathbb{P}(Y_\ell \geq k+1),$$

where $(Y_\ell)_{\ell\geq 1}$ is the Markov chain defined at the beginning of this section. The following lemma gives a similar recursive formula for h as can be found in Lemma 3.15.

Lemma 3.18. *Let h be as above, then $h(k,\ell) = 0$ for all $k \geq \ell+1$ and for all $k,\ell \in \mathbb{N}$, $k \geq 2$, we have*

$$h(k,\ell+1) = \frac{\ell}{\ell+1}\left(1 - \frac{f(k)}{\ell+\alpha}\right) h(k,\ell) + \frac{\ell}{\ell+1}\frac{f(k)}{\ell+\alpha} h(k-1,\ell) \tag{3.35}$$

and

$$h(0,\ell+1) = \frac{\ell}{\ell+1}\left(1 - \frac{f(0)}{\ell+\alpha}\right) h(0,\ell) - \frac{\alpha}{(\ell+1)(\ell+\alpha)}, \tag{3.36}$$

$$h(1,\ell+1) = \frac{\ell}{\ell+1}\left(1 - \frac{1}{\ell+\alpha}\right) h(1,\ell) + \frac{\ell}{(\ell+1)(\ell+\alpha)} h(0,\ell) + \frac{\alpha}{(\ell+1)(\ell+\alpha)} \tag{3.37}$$

for $k = 0,1$.
Moreover for $\Delta^{(1)}h(k,\ell) := h(k+1,\ell) - h(k,\ell)$, we get that for all $\ell \in \mathbb{N}$ and $2 \leq k \leq \ell-1$

$$\Delta^{(1)}h(k,\ell+1) = \frac{\ell}{\ell+1}\left(1 - \frac{f(k+1)}{\ell+\alpha}\right) \Delta^{(1)}h(k,\ell) + \frac{\ell}{\ell+1}\frac{f(k)}{\ell+\alpha}\Delta^{(1)}h(k-1,\ell) \tag{3.38}$$

as well as

$$\Delta^{(1)}h(0,\ell+1) = \frac{\ell}{\ell+1}\left(1 - \frac{1}{\ell+\alpha}\right) \Delta^{(1)}h(0,\ell) + \frac{\alpha}{\ell+\alpha}\frac{\ell}{\ell+1} \quad \textit{and}$$

$$\Delta^{(1)}h(1,\ell+1) = \frac{\ell}{\ell+1}\left(1 - \frac{f(2)}{\ell+\alpha}\right) \Delta^{(1)}h(1,\ell) + \frac{\ell}{(\ell+\alpha)(\ell+1)}\Delta^{(1)}h(0,\ell) - \frac{\alpha}{(\ell+1)(\ell+\alpha)}$$

Proof. Note that since $Y_\ell \leq \ell$ $\mathbb{P}$-a.s, we have $h(k,\ell) = 0$ for any $k \geq \ell+1$. Moreover, by the definition of the Markov chain (Y_n), for $k \geq 2$, we have

$$h(k,\ell+1) = f(k)\mathbb{P}(Y_{\ell+1} = k) - \mathbb{P}(Y_{\ell+1} \geq k+1)$$

$$
\begin{aligned}
&= f(k)\Big(\frac{\ell}{\ell+1}\Big(1-\frac{f(k)}{\ell+\alpha}\Big)\mathbb{P}(Y_\ell = k) + \frac{\ell}{\ell+1}\frac{f(k-1)}{\ell+\alpha}\mathbb{P}(Y_\ell = k-1)\Big)\\
&\quad - \Big(\frac{\ell}{\ell+1}\mathbb{P}(Y_\ell \geq k+1) + \frac{\ell}{\ell+1}\frac{f(k)}{\ell+\alpha}\mathbb{P}(Y_\ell = k)\Big)\\
&= \frac{\ell}{\ell+1}h(k,\ell) - \frac{\ell}{\ell+1}\frac{f(k)}{\ell+\alpha}f(k)\mathbb{P}(Y_\ell = k)\\
&\quad + \frac{\ell}{\ell+1}\frac{f(k)}{\ell+\alpha}\big(f(k-1)\mathbb{P}(Y_\ell = k-1) - \mathbb{P}(Y_\ell = k)\big)\\
&= \frac{\ell}{\ell+1}\left(1-\frac{f(k)}{\ell+\alpha}\right)h(k,\ell) + \frac{\ell}{\ell+1}\frac{f(k)}{\ell+\alpha}h(k-1,\ell),
\end{aligned}
$$

which is exactly (3.35). Using this result we obtain

$$
\begin{aligned}
\Delta^{(1)}&h(k,\ell+1) = h(k+1,\ell+1) - h(k,\ell+1)\\
&= \left(\frac{\ell}{\ell+1} - \frac{\ell}{\ell+1}\frac{f(k+1)}{\ell+\alpha}\right)h(k+1,\ell) + \frac{\ell}{\ell+1}\frac{f(k+1)}{\ell+\alpha}h(k,\ell)\\
&\quad - \left(\frac{\ell}{\ell+1} - \frac{\ell}{\ell+1}\frac{f(k)}{\ell+\alpha}\right)h(k,\ell) - \frac{\ell}{\ell+1}\frac{f(k)}{\ell+\alpha}h(k-1,\ell)\\
&= \frac{\ell}{\ell+1}\left(1-\frac{f(k+1)}{\ell+\alpha}\right)(h(k+1,\ell)-h(k,\ell))\\
&\quad + \frac{\ell}{\ell+1}\frac{f(k)}{\ell+\alpha}\left(h(k,\ell)-h(k-1,\ell)\right),
\end{aligned}
$$

for $k \geq 2$, which is (3.38). With $f(0) = \alpha$ we get

$$
\begin{aligned}
h(0,&\ell+1) = f(0)\mathbb{P}(Y_{\ell+1} = 0) - \mathbb{P}(Y_{\ell+1} \geq 1)\\
&= f(0)\left(\frac{1}{\ell+1}\left(1-\frac{f(0)}{\ell+\alpha}\right) + \frac{\ell}{\ell+1}\left(1-\frac{f(0)}{\ell+\alpha}\right)\mathbb{P}(Y_\ell = 0)\right)\\
&\quad - \left(\frac{\ell}{\ell+1}\mathbb{P}(Y_\ell \geq 1) + \frac{\ell}{\ell+1}\frac{f(0)}{\ell+\alpha}\mathbb{P}(Y_\ell = 0) + \frac{f(0)}{(\ell+1)(\ell+\alpha)}\right)\\
&= \frac{f(0)}{\ell+1}\left(1-\frac{f(0)}{\ell+\alpha}\right) + \frac{\ell}{\ell+1}h(0,\ell) - \frac{\ell}{\ell+1}\frac{f(0)}{\ell+\alpha}h(0,\ell)\\
&\quad - \frac{\ell}{\ell+1}\frac{f(0)}{\ell+\alpha} - \frac{f(0)}{(\ell+1)(\ell+\alpha)}\\
&= \frac{\ell}{\ell+1}\left(1\frac{f(0)}{\ell+\alpha}\right)h(0,\ell) - \frac{\alpha}{(\ell+1)(\ell+\alpha)},
\end{aligned}
$$

where we used that $\frac{f(0)}{\ell+1}\left(1-\frac{f(0)}{\ell+\alpha}\right) = \frac{\alpha\ell}{(\ell+1)(\ell+\alpha)}$. Similarly we obtain

$$
\begin{aligned}
h(1,\ell+1) &= f(1)\mathbb{P}(Y_{\ell+1}=1) - \mathbb{P}(Y_{\ell+1}\geq 2)\\
&= f(1)\left(\frac{1}{\ell+1}\frac{f(0)}{\ell+\alpha} + \frac{\ell}{\ell+1}\frac{f(0)}{\ell+\alpha}\mathbb{P}(Y_\ell=0) + \frac{\ell}{\ell+1}\left(1-\frac{f(1)}{\ell+\alpha}\right)\mathbb{P}(Y_\ell=1)\right)\\
&\quad -\left(\frac{\ell}{\ell+1}\mathbb{P}(Y_\ell\geq 2) + \frac{\ell}{\ell+1}\frac{f(1)}{\ell+\alpha}\mathbb{P}(Y_\ell=1)\right)\\
&= \frac{f(0)}{(\ell+1)(\ell+\alpha)} + \frac{\ell}{\ell+1}h(1,\ell) - \frac{\ell}{(\ell+1)(\ell+\alpha)}h(1,\ell) + \frac{\ell}{(\ell+1)(\ell+\alpha)}h(0,\ell)\\
&= \frac{\ell}{\ell+1}\left(1-\frac{1}{\ell+\alpha}\right)h(1,\ell) + \frac{\ell}{(\ell+1)(\ell+\alpha)}h(0,\ell) + \frac{f(0)}{(\ell+1)(\ell+\alpha)}, \qquad (3.39)
\end{aligned}
$$

where we used $f(1)=\frac{1}{2+\delta}+\frac{1+\delta}{2+\delta}=1$. This gives (3.36) and (3.37) Using the above identities as well as (3.35) for $h(2,\ell+1)$ yields the desired results. □

The following proposition gives the necessary results on $h(k,l)$, which slightly vary from the ones given in Proposition 3.14 for models fulfilling Assumptions **(A)**.

Proposition 3.19. *(i) For h as before and any $\ell\in\mathbb{N}$ we have*

$$h(0,\ell) = -\frac{1}{\ell}.$$

(ii) For any $k,\ell\in\mathbb{N}$ we have

$$|h(k,\ell)| \leq \frac{\hat{C}}{\ell},$$

with $\hat{C}=\max\{3,\prod_{i=1}^{k^*}\frac{f(i+1)}{i+\alpha}\}$, where $k^*\geq 1$ is such that

$$f(k+1)\geq k+\alpha \quad \forall k\leq k^* \text{ and } f(k+1)\leq k+\alpha \text{ for } k>k^*.$$

(iii) For any $\ell\in\mathbb{N}$ we get

$$\Delta^{(1)}h(0,\ell) = h(1,\ell)-h(0,\ell) \geq \frac{\alpha}{\ell}.$$

(iv) For $\ell\leq\frac{1}{1+\delta}, k\leq \ell-1$ we have

$$\Delta^{(1)}h(k,\ell) = h(k+1,\ell)-h(k,\ell)\geq 0.$$

(iv) For all $\ell\in\mathbb{N}$ there exists $I(\ell)\in\{0,\dots,\ell\}$ such that

$$\Delta^{(1)}h(k,\ell) = h(k+1,\ell)-h(k,\ell)\begin{cases}\geq 0 & \text{if } k<I(\ell),\\ \leq 0 & \text{if } k\geq I(\ell).\end{cases} \qquad (3.40)$$

Moreover, $I(\ell+1)\in\{I(\ell),I(\ell)+1\}$.

Proof. First note that

$$
\begin{aligned}
h(0,1) &= f(0)\mathbb{P}(Y_1 = 0) - \mathbb{P}(Y_1 \geq 1) = -\mathbb{P}(Y_1 = 1) = -1, \\
h(1,1) &= f(1)\mathbb{P}(Y_1 = 1) - \mathbb{P}(Y_1 \geq 2) = f(1)\mathbb{P}(Y_1 = 1) = 1, \\
h(0,2) &= f(0)\mathbb{P}(Y_2 = 0) - \mathbb{P}(Y_2 \geq 1) = \frac{f(0)}{2}\left(1 - \frac{f(0)}{1+\alpha}\right) - \left(1 - \frac{1}{2}\left(1 - \frac{f(0)}{1+\alpha}\right)\right) \\
&= \frac{1}{2}\left(f(0) - \frac{f(0)^2}{1+\alpha} - 1 - \frac{f(0)}{1+\alpha}\right) = -\frac{1}{2}, \\
h(1,2) &= f(1)\mathbb{P}(Y_2 = 1) - \mathbb{P}(Y_2 \geq 2) \\
&= f(1)\left(\frac{1}{2}\left(1 - \frac{f(1)}{1+\alpha}\right) + \frac{f(0)}{2(1+\alpha)}\right) - \frac{f(1)}{2(1+\alpha)} \\
&= \frac{1}{2}\left(1 + \frac{\alpha - 2}{1+\alpha}\right) = \frac{2\alpha - 1}{2(1+\alpha)}
\end{aligned}
$$

and

$$
h(2,2) = f(2)\mathbb{P}(Y_2 = 2) = \frac{f(2)}{2(1+\alpha)}.
$$

As in the proof of Proposition 3.14, we will use induction to show all of the stated results.

(i) From the calculations above we have

$$
h(0,1) = -1 \text{ and } h(0,2) = -\frac{1}{2}.
$$

Using the induction hypothesis $h(0,\ell) = -\frac{1}{\ell}$ and the recursive formula deduced in Lemma 3.18 we get

$$
h(0,\ell+1) = \left(\frac{\ell}{\ell+1} - \frac{\alpha\ell}{(\ell+1)(\ell+\alpha)}\right)\left(-\frac{1}{\ell}\right) - \frac{\alpha}{(\ell+1)(\ell+\alpha)} = -\frac{1}{\ell+1}.
$$

(ii) For $k = 0$ and all $\ell \geq 2$ the statement follows from (i). We now proceed via induction on ℓ. As $h(k,\ell) = 0$ for $k \geq \ell + 1$ we need to show that the result holds for $k \leq \ell$. For $\ell = 2$ we have

$$
|h(0,2)| \leq \frac{1}{2}, \quad |h(1,2)| \leq \frac{1}{2}\left|\frac{2\alpha - 1}{1+\alpha}\right| \leq \frac{1}{2}, \quad h(2,2) = \frac{f(2)}{2(1+\alpha)} \leq \frac{\hat{C}}{2},
$$

which proves the base clause.
Assuming that the statement holds for an $\ell \in \mathbb{N}$ and all $k \leq \ell$, (3.39) and (i) yield

$|h(1,\ell+1)|$

$$
\begin{aligned}
&= \left| \left(\frac{\ell}{\ell+1} - \frac{\ell}{(\ell+1)(\ell+\alpha)} \right) h(1,\ell) + \frac{\ell}{(\ell+1)(\ell+\alpha)} h(0,\ell) + \frac{f(0)}{(\ell+1)(\ell+\alpha)} \right| \\
&= \left| \left(\frac{\ell}{\ell+1} - \frac{\ell}{(\ell+1)(\ell+\alpha)} \right) h(1,\ell) - \frac{1}{(\ell+1)(\ell+\alpha)} + \frac{f(0)}{(\ell+1)(\ell+\alpha)} \right| \\
&\leq \left(\frac{\ell}{\ell+1} - \frac{\ell}{(\ell+1)(\ell+\alpha)} \right) \frac{\hat{C}}{\ell} + \frac{1-\alpha}{(\ell+\alpha)(\ell+1)} \\
&\leq \frac{\hat{C}}{\ell+1} \left(1 - \frac{1}{\ell+\alpha} + \frac{1-\alpha}{\ell+\alpha} \right) \leq \frac{\hat{C}}{\ell+1}
\end{aligned}
$$

and by (3.35)

$$
\begin{aligned}
|h(k,\ell+1)| &= \left| \left(\frac{\ell}{\ell+1} - \frac{\ell}{\ell+1} \frac{f(k)}{\ell+\alpha} \right) h(k,\ell) + \frac{\ell}{\ell+1} \frac{f(k)}{\ell+\alpha} h(k-1,\ell) \right| \\
&\leq \left(\frac{\ell}{\ell+1} - \frac{\ell}{\ell+1} \frac{f(k)}{\ell+\alpha} + \frac{\ell}{\ell+1} \frac{f(k)}{\ell+\alpha} \right) \frac{\hat{C}}{\ell+1} = \frac{\hat{C}}{\ell+1},
\end{aligned}
$$

for $2 \leq k \leq \ell$ as we proved the bound on $h(1,\ell)$ just beforehand. We still have to deal with the case $k = \ell + 1$:

$$
\begin{aligned}
|h(\ell+1,\ell+1)| = f(\ell+1)\mathbb{P}(Y_{\ell+1} = \ell+1) &= f(\ell+1) \prod_{i=1}^{\ell} \frac{i}{i+1} \frac{f(i)}{i+\alpha} \\
&= \frac{f(0)}{\ell+1} \prod_{i=1}^{\ell} \frac{f(i+1)}{i+\alpha} \leq \frac{\hat{C}}{\ell+1},
\end{aligned}
$$

due to the definition of $\hat{C}$.

(iii) For $\ell = 2$ we have

$$
\Delta^{(1)} h(0,2) = \frac{2\alpha-1}{2(1+\alpha)} + \frac{1}{2} = \frac{1}{2} \left(1 - \frac{1}{1+\alpha} + \frac{2\alpha}{1+\alpha} \right) = \frac{3\alpha}{2(1+\alpha)} \geq \frac{\alpha}{2}.
$$

Assuming that the statement is true for some $\ell \in \mathbb{N}$, we get

$$
\begin{aligned}
h(1,\ell+1) - h(0,\ell+1) &= \left(\frac{\ell}{\ell+1} - \frac{\ell}{(\ell+1)(\ell+\alpha)} \right) \Delta^{(1)} h(0,\ell) + \frac{\alpha}{\ell+\alpha} \frac{\ell}{\ell+1} \\
&\geq \left(\frac{\ell}{\ell+1} - \frac{\ell}{(\ell+1)(\ell+\alpha)} \right) \frac{\alpha}{\ell} + \frac{\alpha}{\ell+\alpha} \frac{\ell}{\ell+1} \\
&= \frac{\alpha}{\ell+1} + \frac{\alpha}{\ell+\alpha} \frac{\ell}{\ell+1} - \frac{\alpha}{(\ell+1)(\ell+\alpha)} \geq \frac{\alpha}{\ell+1}.
\end{aligned}
$$

(iv) Obviously $\Delta^{(1)} h(0,\ell) \geq 0$ for all $\ell \in \mathbb{N}$ due to (iii). Now

$$
\Delta^{(1)} h(1,2) = \frac{f(2)}{2(1+\alpha)} - \frac{2\alpha-1}{2(1+\alpha)} = \frac{1}{2(1+\alpha)} \left(\left(\frac{2}{2+\delta} + \alpha \right) + 1 - 2\alpha \right) \geq 0,
$$

as $0<\alpha<1$, yields the result for $\ell=2$. Assuming that the statement is true for some $\ell \leq \frac{1}{1+\delta}-1$, (iii) gives

$$\begin{aligned}
&\Delta^{(1)}h(1,\ell+1)\\
&=\left(\frac{\ell}{\ell+1}-\frac{\ell}{\ell+1}\frac{f(2)}{\ell+\alpha}\right)\Delta^{(1)}h(1,\ell)+\frac{\ell}{(\ell+\alpha)(\ell+1)}\Delta^{(1)}h(0,\ell)-\frac{\alpha}{(\ell+1)(\ell+\alpha)}\\
&\geq \frac{\ell}{(\ell+\alpha)(\ell+1)}\frac{\alpha}{\ell}-\frac{\alpha}{(\ell+1)(\ell+\alpha)}=0
\end{aligned}$$

and for $2\leq k\leq \ell-1$

$$\begin{aligned}
&h(k+1,\ell+1)-h(k,\ell+1)\\
&=\left(\frac{\ell}{\ell+1}-\frac{\ell}{\ell+1}\frac{f(k+1)}{\ell+\alpha}\right)\Delta^{(1)}h(k,\ell)+\frac{\ell}{\ell+1}\frac{f(k)}{\ell+\alpha}\Delta^{(1)}h(k-1,\ell),
\end{aligned}$$

which is non-negative due to the induction hypothesis as

$$f(k+1)\leq f(\ell)=\frac{\ell}{2+\delta}+\alpha\leq \ell+\alpha.$$

It remains to check the case $k=\ell$. We have

$$\begin{aligned}
&h(\ell+1,\ell+1)-h(\ell,\ell+1)\\
&\quad=f(\ell+1)\mathbb{P}(Y_{\ell+1}=\ell+1)-f(\ell)\mathbb{P}(Y_{\ell+1}=\ell)+\mathbb{P}(Y_{\ell+1}\geq \ell+1)\\
&\quad=f(\ell+1)\frac{\ell}{\ell+1}\frac{f(\ell)}{\ell+\alpha}\mathbb{P}(Y_\ell=\ell)-f(\ell)\left[\frac{\ell}{\ell+1}\left(1-\frac{f(\ell)}{\ell+\alpha}\right)\mathbb{P}(Y_\ell=\ell)\right.\\
&\qquad\left.+\frac{\ell}{\ell+1}\frac{f(\ell-1)}{\ell+\alpha}\mathbb{P}(Y_\ell=\ell-1)\right]+\frac{\ell}{\ell+1}\frac{f(\ell)}{\ell+\alpha}\mathbb{P}(Y_\ell=\ell)\\
&\quad=\frac{\ell}{\ell+1}f(\ell)\mathbb{P}(Y_\ell=\ell)\left(\frac{f(\ell+1)}{\ell+\alpha}-1+\frac{f(\ell)}{\ell+\alpha}\right)\\
&\qquad-\frac{\ell}{\ell+1}\frac{f(\ell)}{\ell+\alpha}\left(f(\ell-1)\mathbb{P}(Y_\ell=\ell-1)-\mathbb{P}(Y_\ell=\ell)\right)\\
&\quad=\frac{\ell}{\ell+1}\left(\frac{f(\ell+1)}{\ell+\alpha}-1\right)h(\ell,\ell)+\frac{\ell}{\ell+1}\frac{f(\ell)}{\ell+\alpha}(h(\ell,\ell)-h(\ell-1,\ell)),
\end{aligned}$$

which is non-negative due to $h(\ell,\ell)=f(\ell)\mathbb{P}(Y_\ell=\ell)>0$, the induction hypothesis and the fact that

$$\frac{f(\ell+1)}{\ell+\alpha}-1=\frac{\frac{\ell+1}{2+\delta}+\alpha-(\ell+\alpha)}{\ell+\alpha}\geq 0 \quad\Leftrightarrow\quad \ell\leq\frac{1}{1+\delta}.$$

Now suppose the statement is true for some $\ell > \frac{1}{1+\delta}$. The rest of the proof follows analogously to the case of models fulfilling Assumptions **(A)**, the only thing to verify is that for $I(\ell) < \ell$ we have $\Delta^{(1)}h(k, \ell+1) < 0$ for $k = \ell, \ell+1$. Indeed,

$$\Delta^{(1)}h(\ell+1, \ell+1) = -h(\ell+1, \ell+1) = -f(\ell+1)\mathbb{P}(Y_{\ell+1} = \ell+1) < 0$$

and from the calculations above we also get

$$\begin{aligned}\Delta^{(1)}h(\ell, \ell+1) &= \left(\frac{\ell}{\ell+1}\frac{f(\ell+1)}{\ell+\alpha} - \frac{\ell}{\ell+1}\right)h(\ell, \ell) + \frac{\ell}{\ell+1}\frac{f(\ell)}{\ell+\alpha}\Delta^{(1)}h(\ell-1, \ell)\\ &\leq 0,\end{aligned}$$

since $\ell > \frac{1}{1+\delta}$ so that $\frac{f(\ell+1)}{\ell+\alpha} < 1$ and $\Delta^{(1)}h(\ell-1, \ell) \leq 0$ as $I(\ell) < \ell$.

□

Proof of Theorem 3.16. By Lemma 3.17 it follows that

$$\mathbb{E}\left[\mathcal{A}g_A(Y_{n+1})\right] \leq \frac{1}{n+1}\Big(\sum_{\ell=1}^{n}\sum_{k=0}^{\ell-1}\Delta v_A(k)h(k, \ell)\Big) + C\,\frac{\log(n)}{n}.$$

As in the proof of Theorems 3.6 and 3.7 we use a discrete integration by parts formula to get

$$\sum_{k=0}^{\ell-1}\Delta v_A(k)h(k, \ell) = -v_A(0)h(0, \ell) - \sum_{k=0}^{\ell} v_A(k+1)\Delta^{(1)}h(k, \ell),$$

where we also exploited that $h(\ell+1, \ell) = 0$. By Proposition 3.19 (iv) we obtain

$$\begin{aligned}&\Big|\sum_{k=0}^{\ell-1} v_A(k+1)\Delta^{(1)}h(k, \ell)\Big|\\ &\qquad\leq \sup_k |v_A(k)|\Big(h(I(\ell), \ell) - h(0, \ell) + h(I(\ell)+1, \ell) - h(\ell, \ell)\Big)\\ &\qquad\leq 2\sup_k |v_A(k)| \sup_{k\leq \ell-1} h(k, \ell).\end{aligned}$$

Now Proposition 3.19 (ii) and Proposition 3.11 yield

$$\begin{aligned}&\big|\mathbb{E}[\mathcal{A}g_A(X_{n+1})]\big|\\ &\leq \Big|\frac{1}{n+1}\sum_{\ell=1}^{n}\Big(\Big(\sum_{k=0}^{\ell-1} v_A(k+1)\Delta^{(1)}h(k, \ell)\Big) + v_A(0)h(0, \ell)\Big)\Big| + C\,\frac{\log(n)}{n}\\ &\leq \frac{2}{n+1}\sup_k |v_A(k)| \sum_{\ell=1}^{n}\sup_{k\leq \ell-1} h(k, \ell) + C\,\frac{\log(n)}{n}\\ &\leq \frac{\hat{C}}{n+1}\sum_{\ell=1}^{n}\frac{1}{\ell} + C\,\frac{\log(n)}{n}.\end{aligned}$$

Hence,

$$d_{\mathrm{TV}}(W_{n+1}, W) = \sup_{A \subset \mathbb{N}_0} \left| \mathbb{E}[\mathcal{A} g_A(X_{n+1})] \right| \leq C \, \frac{\log(n)}{n},$$

for some constant $C \geq \hat{C}$. □

3.3. Rates of convergence for the out-degree

Theorems 3.6, 3.7 and 3.16 all deal with the indegree of a uniformly chosen vertex. However, for the model described in Example 3.1, an obvious question concerns the distribution of the random outdegree. In [DM09] the authors show that the outdegree is approximately Poisson distributed. The next theorem gives an error bound on this approximation.

Theorem 3.20. *Let D_n denote the outdegree of vertex n in the model described in Example 3.1 and suppose that for some $\gamma \in (0,1)$, we have $f(k) \leq \gamma k + 1$ for all $k \in \mathbb{N}_0$. Then there exists $C > 0$ such that*

$$d_{\mathrm{TV}}(D_n, Po(\lambda_n)) \leq C \begin{cases} \frac{1}{n+1}, & \text{for } 0 < \gamma < \frac{1}{2}, \\ \frac{\log(n)}{n}, & \text{for } \gamma = \frac{1}{2}, \\ n^{-2(1-\gamma)}, & \text{for } \frac{1}{2} < \gamma < 1, \end{cases}$$

where $Po(\lambda_n)$ denotes the Poisson distribution with parameter $\lambda_n = \mathbb{E}\left[f(W_{n-1})\right]$ and W_{n-1} has the distribution of the indegree of a uniformly chosen vertex at time $n-1$. Moreover, $\lambda_n \to \lambda := \mathbb{E}[f(W)]$, where $W \sim \mu$ as in (3.3). Finally, if $f(k) = \gamma k + \beta$ for $\gamma \in (0,1), \beta \in [0,1]$, then

$$|\lambda_n - \mathbb{E}\left[f(W)\right]| \leq n^{-1+\gamma}.$$

For the proof we will need the following moment bound.

Lemma 3.21. *For the preferential attachment model as in Example 3.1 with $f(k) \leq \gamma k + 1$ for all k and some $\gamma \in (0,1)$, we have, for all $n \in \mathbb{N}$,*

$$\mathbb{E}[f(\deg_n^-(i))] \leq \left(\frac{n}{i}\right)^{\gamma} \quad \text{for all } i \in [n].$$

Proof. For the linear attachment rule $f^{(\gamma)}(k) := \gamma \cdot k + 1$, the statement of the lemma is proved in [DM13, Lemma 2.7]. Denote the indegrees in the corresponding preferential attachment model by $\deg_n^{-,\gamma}(i)$ and consider general f and associated degrees $\deg_n^-(i)$. Then, since $f \leq f^{(\gamma)}$ we can couple the models so that $\deg_n^-(i) \leq \deg_n^{-,\gamma}(i)$ for all $i \in [n], n \in \mathbb{N}$. In particular, we have that

$$\mathbb{E}[f(\deg_n^-(i))] \leq \mathbb{E}[f^{(\gamma)}(\deg_n^-(i))] \leq \mathbb{E}[f^{(\gamma)}(\deg_n^{-,\gamma}(i))] \leq \left(\frac{n}{i}\right)^{\gamma},$$

as required. □

Using a result of [BH84] for Poisson approximation (again based on the Chen-Stein method), we can now prove Theorem 3.20.

Proof of Theorem 3.20. By the independence assumption for incoming edges, it follows that the indegree evolutions $(\deg_k^-(i))_{k\geq i}$ and $(\deg_k^-(j))_{k\geq j}$ are independent if $i \neq j$. In particular, if we write $X_{i,n} = \mathbb{1}\{\text{there is an edge from } n \text{ to } i\} = \deg_n^-(i) - \deg_{n-1}^-(i)$, then we can write the outdegree D_n of vertex n as

$$D_n = \sum_{i=1}^{n-1} X_{i,n},$$

i.e. as the sum of independent Bernoulli variables. Note that

$$p_{i,n} := \mathbb{P}(X_{i,n} = 1) = \mathbb{E}\left[\mathbb{E}\left[\deg_n^-(i) - \deg_{n-1}^-(i) | \mathcal{G}_{n-1}\right]\right] = \mathbb{E}\left[\frac{f(\deg_{n-1}(i))}{n-1}\right].$$

Therefore,

$$\lambda_n := \mathbb{E}\left[D_n\right] = \mathbb{E}\Big[\frac{1}{n-1}\sum_{i=1}^{n-1} f(\deg_{n-1}(i))\Big] = \mathbb{E}\left[f(W_{n-1})\right],$$

where W_{n-1} denotes the indegree of a uniformly chosen vertex after the insertion of vertex $n-1$. From the proof of Theorem 1.1 (b) in [DM09] we know that $\lambda_n \to \mathbb{E}[f(W)]$ if $W \sim \mu$. Applying [BH84, Thm. 1.1] we obtain that

$$d_{\mathrm{TV}}(D_n, Po(\lambda_n)) \leq \frac{1-e^{-\lambda_n}}{\lambda_n} \sum_{i=1}^{n-1} p_{i,n}^2 \leq \min\{1, \frac{1}{\lambda_n}\} \sum_{i=1}^{n-1} p_{i,n}^2. \tag{3.41}$$

It remains to control the sum on the right hand side. By Lemma 3.21, we have that

$$\sum_{i=1}^{n-1} p_{i,n}^2 = \frac{1}{(n-1)^2} \sum_{i=1}^{n-1} \mathbb{E}[f(\deg_{n-1}^-(i))]^2 \leq \frac{1}{(n-1)^2} \sum_{i=1}^{n-1} \left(\frac{n}{i}\right)^{2\gamma}.$$

Since $\lambda_n \to \lambda := \mathbb{E}[f(W)]$ we can deduce from (3.41) that

$$d_{\mathrm{TV}}(D_n, Po(\lambda_n)) \leq C \begin{cases} \frac{1}{n+1}, & \text{for } 0 < \gamma < \frac{1}{2}, \\ \frac{\log(n)}{n}, & \text{for } \gamma = \frac{1}{2}, \\ n^{-2(1-\gamma)}, & \text{for } \frac{1}{2} < \gamma < 1, \end{cases}$$

for a suitable constant $C > 0$, which proves the first part of Theorem 3.20.

For the final part, we assume that $f(k) = \gamma k + \beta$, for $\gamma \in (0,1), \beta \in [0,1]$. First note that in this case by (3.7)

$$\begin{aligned} \lambda = \mathbb{E}[f(W)] = \gamma\mathbb{E}[W] + \beta &= \gamma \sum_{k\geq 1} \mu([k,\infty)) + \beta \\ &= \gamma \sum_{k\geq 1} f(k-1)\mu_{k-1} + \beta = \gamma\lambda + \beta. \end{aligned}$$

In particular, $\lambda = \frac{\beta}{1-\gamma}$. Following a similar argument as in the proof of Theorem 1.1 (b) in [DM09], we have that

$$
\begin{aligned}
\mathbb{E}[f(W_{n+1})] &= \frac{1}{n+1}\sum_{i=1}^{n+1}\mathbb{E}\Big[\mathbb{E}[f(\deg^-_{n+1}(i))\,|\,\mathcal{G}_n]\Big] \\
&= \frac{1}{n+1}\sum_{i=1}^{n}\mathbb{E}\Big[\mathbb{E}[f(\deg^-_{n+1}(i)) - f(\deg^-_n(i))\,|\,\mathcal{G}_n]\Big] \\
&\quad + \frac{f(0)}{n+1} + \frac{1}{n+1}\sum_{i=1}^{n}\mathbb{E}\Big[\mathbb{E}[f(\deg^-_n(i))\,|\,\mathcal{G}_n]\Big] \\
&= \frac{1}{n+1}\left(\sum_{i=1}^{n}\mathbb{E}\Big[\mathbb{E}[\gamma(\deg^-_{n+1}(i) - \deg^-_n(i))\,|\,\mathcal{G}_n]\Big] + \beta + \sum_{i=1}^{n}\mathbb{E}[f(\deg^-_n(i))]\right) \\
&= \frac{1}{n+1}\sum_{i=1}^{n}\gamma\mathbb{E}\Big[\frac{f(\deg^-_n(i)}{n}\Big] + \frac{\beta}{n+1} + \frac{1}{n+1}\sum_{i=1}^{n}\mathbb{E}[f(\deg^-_n(i))] \\
&= (1 - \frac{1-\gamma}{n+1})\mathbb{E}[f(W_n)] + \frac{\beta}{n+1}.
\end{aligned}
$$

Using the fact that $\lambda = \frac{\beta}{1-\gamma}$, we obtain that for $\bar{\lambda}_{n+1} := \mathbb{E}[f(W_n)] - \lambda$,

$$
\bar{\lambda}_{n+1} = \Big(1 - \frac{1-\gamma}{n}\Big)\bar{\lambda}_n.
$$

Hence,

$$
|\bar{\lambda}_{n+1}| = \prod_{i=1}^{n}\Big(1 - \frac{1-\gamma}{i}\Big)|\bar{\lambda}_1| \le Cn^{-(1-\gamma)},
$$

for a suitable constant $C > 0$, as claimed. □

4. Rates of convergence via coupling

In the following chapter we look at the same preferential attachment models as were considered in the previous one. However, this time we will use a different technique to derive rates of convergence. Unfortunately, this techniques can only be applied to a subclass of the attachment functions considered in chapter 3 and the rates of convergence contain an additional factor of order $\log(n)$ compared to the bounds given in Theorem 3.7.

The coupling method, introduced by Wolfgang Döblin, is a powerful technique in probability theory which allows to compare two probability measures. One of the most prominent applications is in the theory of Markov processes. Here, one constructs two copies of a process, one of which is already in stationarity, on a joint probability space and shows that the two processes coincide with high probability after some random time. We will use this technique in an alternative proof to deduce rates of convergence for one of the random quantities already considered in section 3.2, namely the indegree of a uniformly chosen vertex.
In section 4.1 we begin by considering preferential attachment models satisfying our Assumptions **(A)** from section 3.1 and subsequently in section 4.2 deal with models that do not satisfy those assumptions.

4.1. Coupling for general models

Noticing that the behaviour of the discrete-time Markov chain $(X_n)_{n\geq 1}$ as given in Lemma 3.12 resembles that of a continuous-time Markov chain $(Z_t)_{t\geq 0}$ with generator $\mathcal{A}$, we can apply coupling techniques to deduce rates of convergence.

Our main result in this section gives bounds on the distance of W_n (the degree of a uniformly chosen vertex in $\mathcal{G}_n$) and its limiting distribution in the total variation metric for attachment functions f, such that the first and second moment exist. The result is formulated in the following theorem.

Theorem 4.1. *For every monotonically increasing attachment function $f : \mathbb{N}_0 \to (0, \infty)$ and $f \in \mathcal{L}^2(\mu)$, where μ is the measure defined in (3.3), there exists a constant $C > 0$ such that*

$$d_{TV}(W_n, W) \leq C \, \frac{\log(n+1)^2}{n},$$

where $W \sim \mu$ and W_n denotes the indegree of a uniformly chosen vertex at time n.

Remark 4.2. *Following [DM09] for $f(k) \sim \gamma k^\alpha$ with $0 < \alpha < 1$, we have*

$$\log(\mu_k) \sim -\frac{1}{\gamma}\frac{1}{1-\alpha}k^{1-\alpha},$$

such that in this case the assumptions of Theorem 4.1 are met. Furthermore, for $f(k) = \gamma k + \beta$ with $\gamma, \beta \in (0, 1]$, [DM09] gives

$$\mu_k \sim \frac{\Gamma(\frac{\beta+1}{\gamma})}{\gamma\Gamma(\frac{\beta}{\gamma})}k^{-(1+\frac{1}{\gamma})},$$

so in the case $\gamma > \frac{1}{2}$ Theorem 4.1 is not applicable. In particular this shows that Theorem 3.7 can be applied to a larger class of attachment functions.

In order for the reader not to get lost in the tedious calculations, we give a short overview of the proof. The main idea is to construct an explicit coupling (X'_n, Y'_n) of (X_n, Y_n), where Y_n and Y'_n respectively are observations of the continuous-time process Z_t at discrete time instances, and apply the coupling inequality

$$d_{TV}(X_n, Y_n) \leq \mathbb{P}(X'_n \neq Y'_n), \tag{4.1}$$

see Proposition 2.14. Thus, we will construct a coupling of the discrete- and the continuous-time process so that the two chains are in the same state with high probability. We will proceed similar as in the proof of Proposition 2.14. Due to the scaling of the transition probabilities of the discrete-time Markov chain, we will observe the continuous-time process Z at discrete time steps $\psi(n)$, where

$$\psi(n) = \sum_{i=1}^{n} \frac{1}{i},$$

so that $\psi(n+1) - \psi(n) = \frac{1}{n+1}$. We denote the observation of Z at the n-th point in time by Y_n, i.e.

$$Y_n = Z_{\psi(n)}.$$

As the transition probabilities of the two chains differ, we cannot construct a coupling in such a way, that the chains evolve together from some random point in time onwards. However, we can construct it in such a way that once the chains are in the same state, they stay together with maximal probability, e.g. if the chains are both in state k at time n they move together to state ℓ with probability $\min\{\mathbb{P}(X_{n+1} = \ell | X_n = k), \mathbb{P}(Y_{n+1} = \ell | Y_n = k)\}$. Lemma 4.3 gives the necessary inequalities to decide which transition probability is the smaller of the two and also shows that for fixed k these probabilities asymptotically coincide. In the proof of Lemma 4.5 we give the precise definition of a coupling (X_n', Y_n') of (X_n, Y_n). More precisely, we force the chains to meet, whenever the continuous-time process moves to 0 and then keep them together with maximal probability. Furthermore, we show that this really defines a coupling, and finally deduce an upper bound on the probability of the event that the chains drift apart, even though they have been in the same state before. As it is more likely for the discrete-time process to perform a jump to zero than for the continuous-time process (cf. Lemma 4.3), we show in Lemma 4.6 that the likelihood of X moving to zero without Y, is small. The proof of Theorem 4.1 uses these results to show that for n large, the probability that the chains are not in the same state, is close to zero.

Lemma 4.3. *For the time continuous process Y_n observed at discrete times we have*

$$\mathbb{P}(Y_{n+1} = 0 | Y_n = k) \leq \frac{1}{n+1} \text{ for } 0 < k \leq n, \tag{4.2}$$

$$\mathbb{P}(Y_{n+1} = k+1 | Y_n = k) \leq \frac{f(k)}{n+1}, \tag{4.3}$$

$$\mathbb{P}(Y_{n+1} = k | Y_n = k) \geq 1 - \frac{1+f(k)}{n+1} \text{ for } k \neq 0, \tag{4.4}$$

$$\mathbb{P}(Y_{n+1} = 0 | Y_n = 0) \geq 1 - \frac{f(0)}{n+1} \tag{4.5}$$

and

$$\mathbb{P}(Y_{n+1} = j | Y_n = k) \leq \frac{(1+f(k+1))^2}{2(n+1)^2} \tag{4.6}$$

for $n \geq 4$. Moreover

$$\sup_{j \geq 0} |\mathbb{P}(X_{n+1} = j | X_n = k) - \mathbb{P}(Y_{n+1} = j | Y_n = k)| \leq \frac{(1+f(k+1))^2}{(n+1)^2} \tag{4.7}$$

for $j \notin \{0, k, k+1\}$ and all $k \leq n$, where $(X_n)_{n \geq 1}$ denotes the discrete time Markov process introduced in the previous chapter.

Proof. Throughout the proof we will make extensive use of the following inequalities, resulting from bounds on the Lagrange remainder in the Taylor series of $\exp(-x)$:

$$x - \frac{x^2}{2} \leq 1 - \exp(-x) \leq x. \tag{4.8}$$

Let $A_{n+1}^{(\diamond)}$, with $\diamond \in \{>, \geq, =\}$, denote the events that the time continuous process Z_t moved more than once, at least once or exactly once respectively in the interval $J_n := (0, \frac{1}{n+1}]$. By B_k we denote the event that Y_n is in state k the next time we observe the process. With this notation we obtain

$$\mathbb{P}(Y_{n+1} = 0 | Y_n = k) = \mathbb{P}_k(A_{n+1}^{(=)}, B_0) + \mathbb{P}_k(A_{n+1}^{(>)}, B_0) \tag{4.9}$$

for $k \neq 0$, where we used $\mathbb{P}(Y_{n+1} = 0 | Y_n = k) := \mathbb{P}_k(Y_{n+1} = 0)$ for simplicity of notation. Recall that in the time-continuous process Z_t the times between two movements are exponentially distributed, where the parameter is given by the sum of all rates of jumps which are possible from the current state of the chain. Using

$$\mathbb{P}_k(Y_{n+1} = 0 | A_{n+1}^{(=)}) = \frac{1}{1 + f(k)} \text{ and } \mathbb{P}_k(Y_{n+1} = k+1 | A_{n+1}^{(=)}) = \frac{f(k)}{1 + f(k)},$$

we can then write the first probability in (4.9) as

$$\begin{aligned}
\mathbb{P}_k(A_{n+1}^{(=)}, B_0) &= \frac{1}{1 + f(k)} \left(\mathbb{P}_k(A_{n+1}^{(\geq)}) - \mathbb{P}_k(A_{n+1}^{(>)}) \right) \\
&= \frac{1}{1 + f(k)} \left(\mathbb{P}_k \left(\mathrm{Exp}(1 + f(k)) \leq \frac{1}{n+1} \right) - \mathbb{P}_k(A_{n+1}^{(>)}) \right) \\
&= \frac{1}{1 + f(k)} \left(\left(1 - \exp\left(-\frac{1 + f(k)}{n+1} \right) \right) - \mathbb{P}_k(A_{n+1}^{(>)}) \right) \\
&= \frac{1}{n+1} - \frac{1 + f(k)}{2(n+1)^2} + \sum_{j=2}^{\infty} (-1)^j \frac{(1 + f(k))^j}{(j+1)!(n+1)^{j+1}} - \frac{1}{1 + f(k)} \mathbb{P}_k(A_{n+1}^{(>)}),
\end{aligned} \tag{4.10}$$

where $\mathrm{Exp}(\alpha)$ denotes a random variable following an exponential distribution with parameter α. Plugging this into (4.9) we obtain

$$\begin{aligned}
&\mathbb{P}(Y_{n+1} = 0 | Y_n = k) \\
&= \left(\frac{1}{n+1} - \frac{1 + f(k)}{2(n+1)^2} + \sum_{j=1}^{\infty} (-1)^{j+1} \frac{(1 + f(k))^{j+1}}{(j+2)!(n+1)^{j+2}} \right) \\
&\quad - \left(\frac{1}{1 + f(k)} \mathbb{P}_k(A_{n+1}^{(>)}) - \mathbb{P}_k(A_{n+1}^{(>)}, B_0) \right) \\
&\leq \left(\frac{1}{n+1} - \frac{1 + f(k)}{2(n+1)^2} + \frac{(1 + f(k))^2}{6(n+1)^3} \right) - \left(\frac{1}{1 + f(k)} \mathbb{P}_k(A_{n+1}^{(>)}) - \mathbb{P}_k(A_{n+1}^{(>)}, B_0) \right).
\end{aligned}$$

We will now show that

$$-\frac{1 + f(k)}{2(n+1)^2} + \frac{(1 + f(k))^2}{6(n+1)^3} - \frac{1}{1 + f(k)} \mathbb{P}_k(A_{n+1}^{(>)}) + \mathbb{P}_k(A_{n+1}^{(>)}, B_0) < 0.$$

For that we denote by $D^{(\diamond)}_{n+1}$, with $\diamond \in \{>, \geq, =\}$, the event that Z_t falls down more than once, at least once or exactly once resepectively in the intervall J_n. This yields

$$\mathbb{P}_k\Big(A^{(>)}_{n+1}, B_0\Big) = \mathbb{P}_k\Big(A^{(>)}_{n+1}, B_0, D^{(=)}_{n+1}\Big) + \mathbb{P}_k\Big(A^{(>)}_{n+1}, B_0, D^{(>)}_{n+1}\Big).$$

Now

$$\begin{aligned}\mathbb{P}_k\Big(A^{(>)}_{n+1}, B_0, D^{(=)}_{n+1}\Big) &= \mathbb{P}_k\Big(B_0 \big| A^{(>)}_{n+1}, D^{(=)}_{n+1}\Big)\mathbb{P}_k\Big(D^{(=)}_{n+1}\big|A^{(>)}_{n+1}\Big)\mathbb{P}_k\Big(A^{(>)}_{n+1}\Big)\\ &\leq \frac{1}{1+f(k+1)}\mathbb{P}_k\big(A^{(>)}_{n+1}\big),\end{aligned}$$

as the process only falls in the last step and has moved before, so that it must fall down from a stage greater than k.
We now divide the interval J_n into $n+1$ equidistant intervalls I_ℓ of length $\frac{1}{(n+1)^2}$ and define the event

$$M_0^\ell := \{\text{first move occurs in } I_\ell\} = \{\text{no move up to } I_\ell\} \cap \{\text{process moves in } I_\ell\}.$$

Using total probability we then obtain

$$\begin{aligned}\mathbb{P}_k\Big(A^{(>)}_{n+1}, B_0, D^{(>)}_{n+1}\Big) = \mathbb{P}_k\Big(B_0, D^{(>)}_{n+1}\Big) &= \sum_{l=0}^{n} \mathbb{P}_k(B_0, D^{(>)}_{n+1}|M_0^\ell)\mathbb{P}_k(M_0^\ell)\\ &\leq \sum_{l=0}^{n} \mathbb{P}_0\big(A^{(>)}_{n+1}\big)\frac{1}{1+f(k)}\left(1-\exp\left(-\frac{1+f(k)}{(n+1)^2}\right)\right)\\ &\leq \sum_{l=0}^{n} \frac{(1+f(1))^2}{2(n+1)^2}\frac{1}{(n+1)^2}\\ &\leq \frac{(1+f(1))^2}{2(n+1)^3} \leq \frac{1+f(1)}{3(n+1)^2},\end{aligned} \tag{4.11}$$

for $n \geq 4$, since $f(k) \leq k+1$ by Assumptions **(A)**. Here we used that the probability that the process started in 0 moves at least two times in J_n is smaller than the probability that in a Poisson process of intensity $(1+f(1))$ on $\mathbb{R}$ at least two events occur in J_n. More generally for $\Pi(\frac{1+f(k+1)}{n+1})$ denoting a Poisson process of intensity $\frac{1+f(k+1)}{n+1} =: \lambda$, we get

$$\begin{aligned}\mathbb{P}_k\big(A^{(>)}_{n+1}\big) \leq \mathbb{P}\left(\Pi\left(\frac{1+f(k+1)}{n+1}\right) \geq 2\right) &= \sum_{l\geq 2}\lambda^l\frac{e^{-\lambda}}{l!} = \lambda^2 e^{-\lambda}\sum_{l\geq 0}\frac{\lambda^l}{l!}\frac{1}{(l+1)(l+2)}\\ &\leq \frac{\lambda^2}{2}e^{-\lambda}\sum_{l\geq 0}\frac{\lambda^l}{l!} = \frac{\lambda^2}{2} = \frac{(1+f(k+1))^2}{2(n+1)^2}.\end{aligned} \tag{4.12}$$

We have

$$\frac{1+f(k)}{2(n+1)^2} - \frac{(1+f(k))^2}{6(n+1)^3} = \frac{1+f(k)}{2(n+1)^2}\Big(1 - \underbrace{\frac{(1+f(k))}{3(n+1)}}_{\le \frac{1}{3}}\Big) \ge \frac{1+f(k)}{3(n+1)^2},$$

for $1 \le k \le n-1$, which in combination with (4.11) gives (4.2) for n large enough.

Using (4.10) and the fact that $\sum_{j=2}^{\infty}(-1)^j \frac{(1+f(k))^j}{(j+1)!(n+1)^{j+1}} > 0$ we get

$$\begin{aligned}
\mathbb{P}_k(Y_{n+1}=0) \ge \mathbb{P}(A_{n+1}^{(=)}, B_0) &\ge \frac{1}{n+1} - \frac{(1+f(k))}{2(n+1)^2} - \frac{1}{1+f(k)}\mathbb{P}_k\big(A_{n+1}^{(>)}\big) \\
&\ge \frac{1}{n+1} - \frac{(1+f(k))}{2(n+1)^2} - \frac{(1+f(k+1))^2}{2(n+1)^2} \\
&\ge \frac{1}{n+1} - \frac{(1+f(k+1))^2}{(n+1)^2}
\end{aligned}$$

and thus

$$|\mathbb{P}_k(X_{n+1}=0) - \mathbb{P}_k(Y_{n+1}=0)| \le \frac{(1+f(k+1))^2}{(n+1)^2}. \tag{4.13}$$

Note that since $\mathbb{P}_k(X_{n+1}=0) = \frac{1}{n+1}$ for all $k \ge 1$, we also have

$$|\mathbb{P}_\ell(X_{n+1}=0) - \mathbb{P}_k(Y_{n+1}=0)| \le \frac{(1+f(k+1))^2}{(n+1)^2} \quad \forall \ell \ge 1.$$

Using (4.8) and an adapted version of (4.9), we similarly obtain

$$\begin{aligned}
&\mathbb{P}_k(Y_{n+1}=k+1) \\
&= \frac{f(k)}{1+f(k)}\left(\left(1-\exp\left(-\frac{1+f(k)}{n+1}\right)\right) - \mathbb{P}_k\big(A_{n+1}^{(>)}\big)\right) + \mathbb{P}_k\big(A_{n+1}^{(>)}, B_{k+1}\big) \\
&\le \frac{f(k)}{1+f(k)}\left(\frac{1+f(k)}{n+1} - \mathbb{P}_k\big(A_{n+1}^{(>)}\big)\right) + \mathbb{P}_k\big(A_{n+1}^{(>)}\big)\mathbb{P}_k\big(B_{k+1}|A_{n+1}^{(>)}\big) \\
&= \frac{f(k)}{n+1} + \mathbb{P}_k\big(A_{n+1}^{(>)}\big)\left(\mathbb{P}_k\big(B_{k+1}|A_{n+1}^{(>)}\big) - \frac{f(k)}{1+f(k)}\right) \\
&\le \frac{f(k)}{n+1} = \mathbb{P}_k(X_{n+1}=k+1),
\end{aligned}$$

since $\mathbb{P}_k\Big(B_{k+1}\big|A_{n+1}^{(>)}\Big) \le \mathbb{P}_0\big(A_{n+1}^{(\ge)}\big) \le \frac{1+f(0)}{n+1} \le \frac{f(k)}{1+f(k)}$ for all k and $n \ge 4$ by assumption. With (4.8) and (4.12) we get the following lower bound

$$\begin{aligned}
&\mathbb{P}_k(Y_{n+1} = k+1) \\
&\qquad \geq \frac{f(k)}{n+1} - \frac{f(k)(1+f(k))}{2(n+1)^2} + \mathbb{P}_k(A^{(>)}_{n+1}) \underbrace{\Big(\mathbb{P}_k\big(B_{k+1} | A^{(>)}_{n+1}\big) - \frac{f(k)}{1+f(k)} \Big)}_{\geq -1} \\
&\qquad \geq \frac{f(k)}{n+1} - \frac{f(k)(1+f(k))}{2(n+1)^2} - \mathbb{P}_k\big(A^{(>)}_{n+1}\big) \\
&\qquad \geq \frac{f(k)}{n+1} - \frac{(1+f(k+1))^2}{(n+1)^2},
\end{aligned}$$

so that

$$|\mathbb{P}_k(X_{n+1} = k+1) - \mathbb{P}_k(Y_{n+1} = k+1)| \leq \frac{(1+f(k+1))^2}{(n+1)^2}. \tag{4.14}$$

Now for $k \neq 0$,

$$\begin{aligned}
\mathbb{P}(Y_{n+1} = k | Y_n = k) \geq 1 - \mathbb{P}_k\big(A^{(\geq)}_{n+1}\big) &= \exp\left(-\frac{1+f(k)}{n+1} \right) \\
&\geq 1 - \frac{1+f(k)}{n+1} = \mathbb{P}_k(X_{n+1} = k),
\end{aligned} \tag{4.15}$$

but also

$$\begin{aligned}
\mathbb{P}(Y_{n+1} = k | Y_n = k) &= 1 - \mathbb{P}_k\big(A^{(\geq)}_{n+1}\big) + \mathbb{P}_k(A^{(\geq)}_{n+1}, B_k) \\
&= \exp\left(-\frac{1+f(k)}{n+1} \right) + \mathbb{P}_k(A^{(\geq)}_{n+1}, B_k) \\
&\leq 1 - \frac{1+f(k)}{n+1} + \frac{(1+f(k))^2}{2(n+1)^2} + \mathbb{P}_k(A^{(>)}_{n+1}) \\
&\leq 1 - \frac{1+f(k)}{n+1} + \frac{(1+f(k+1))^2}{(n+1)^2}
\end{aligned}$$

so that

$$|\mathbb{P}_k(X_n = k) - \mathbb{P}_k(Y_n = k)| \leq \frac{(1+f(k+1))^2}{(n+1)^2} \tag{4.16}$$

for $k \geq 1$. For $k = 0$ we get

$$\mathbb{P}(Y_{n+1} = 0 | Y_n = 0) \geq 1 - \mathbb{P}_0(A^{(\geq)}_{n+1})$$

and can improve the lower bound (4.15) in this case since the time until the process moves out of 0 is exponentially distributed with parameter $f(0)$ because in contrast to (4.15) the process does not move away from 0 if it falls. So we have

$$\mathbb{P}_0(A^{(\geq)}_{n+1}) = \mathbb{P}\left(\mathrm{Exp}(f(0)) \leq \frac{1}{n+1} \right) = 1 - \exp\left(-\frac{f(0)}{n+1} \right)$$

and with (4.8) we get

$$\mathbb{P}(Y_{n+1}=0|Y_n=0) \ge \exp\left(-\frac{f(0)}{n+1}\right) \ge 1-\frac{f(0)}{n+1} = \mathbb{P}(X_{n+1}=0|X_n=0).$$

Analogously to the calculations in the case $k \neq 0$ we get

$$\begin{aligned}\mathbb{P}(Y_{n+1}=0|Y_n=0) &= 1-\mathbb{P}_0\big(A_{n+1}^{(\ge)}\big)+\mathbb{P}_0(A_{n+1}^{(\ge)},B_0)\\ &\le 1-\frac{f(0)}{n+1}+\frac{(1+f(1))^2}{2(n+1)^2}+\mathbb{P}_0(A_{n+1}^{(>)})\\ &\le 1-\frac{f(0)}{n+1}+\frac{(1+f(1))^2}{(n+1)^2},\end{aligned}$$

so that also for $k=0$ we obtain

$$|\mathbb{P}_0(X_n=0)-\mathbb{P}_0(Y_n=0)| \le \frac{(1+f(1))^2}{(n+1)^2}. \tag{4.17}$$

Observing that with (4.12)

$$|\mathbb{P}_k(X_{n+1}=j)-\mathbb{P}_k(Y_{n+1}=j)| = \mathbb{P}_k(Y_{n+1}=j) \le \mathbb{P}_k(A_{n+1}^{(>)}) \le \frac{(1+f(k+1))^2}{2(n+1)^2},$$

for $j \notin \{0,k,k+1\}$, we get (4.6). (4.13), (4.14) and (4.17) together yield (4.7). □

Remark 4.4. *As $\mathbb{P}(X_{n+1}=0|X_n=k)=\frac{1}{n+1}$ is independent of k for $k \neq 0$, we also have*

$$|\mathbb{P}(X_{n+1}=0|X_n=\ell)-\mathbb{P}(Y_{n+1}=0|Y_n=k)| \le \frac{(1+f(k+1))^2}{(n+1)^2} \tag{4.18}$$

for all $\ell \neq 0$.

For the construction of a coupling (X_n', Y_n') of X_n and Y_n we will use the following abbreviatory notations

$$\begin{aligned} p_{n+1}^X(k|m) &:= \mathbb{P}(X_{n+1}=k|X_n=m),\\ p_{n+1}^Y(j|\ell) &:= \mathbb{P}(Y_{n+1}=j|Y_n=\ell),\\ p_{n+1}^{X'}(k|m,\ell) &:= \mathbb{P}(X_{n+1}'=k|X_n'=m,Y_n'=\ell),\\ p_{n+1}^{Y'}(j|m,\ell) &:= \mathbb{P}(Y_{n+1}'=j|X_n'=m,Y_n'=\ell) \text{ and}\\ \hat{p}_{n+1}(k,j|m,\ell) &:= \mathbb{P}(X_{n+1}'=k,Y_{n+1}'=j|X_n'=m,Y_n'=\ell).\end{aligned}$$

Lemma 4.5. *There exists a (Markovian) coupling* $(X_n^{'}, Y_n^{'})$ *of* X_n *and* Y_n *such that*

$$\mathbb{P}(X_{k+1}^{'} \neq Y_{k+1}^{'} | X_k^{'} = Y_k^{'}) \leq 2 \ \frac{\mathbb{E}\left[(f(Y_k+1)+1)^2\right]}{(k+1)^2}, \tag{4.19}$$

for all $k \geq 0$ *.*

Proof. As stated before we will couple the chains so that they stay together with maximal probability once they have met. This idea is made precise in the construction of an optimal coupling described in [LPW06]. Following that construction, a step is performed by both chains with the minimal of the two probabilities to do so. Thus the chain with the higher probability to move in such a way, can only perform that step without the other one with probability given by the difference of the two probabilities. If these differences converge to zero for all possible movements, the chains will stay in the same place with probability one if time tends to ∞. We define our coupling $(X_n^{'}, Y_n^{'})$ on $\mathbb{N} \times \mathbb{N}$ as follows:

Let $J_n' \sim \mathcal{U}\{1, \ldots, n\}$ for all n and put $X_n' = \deg_n^-(J_n')$. For $\boldsymbol{\ell \neq m}$ we let the two chains evolve independently until Y_n falls down to 0. In this case we put $J_n' = n$ so that $X_n' = 0$. For $\ell \neq 0$ we now define

$$\hat{p}_{n+1}(k, j | m, \ell) = \frac{p_{n+1}^X(k|m) \cdot p_{n+1}^Y(j|\ell)}{1 - p_{n+1}^Y(0|\ell)} \quad \text{for } k, j \neq 0,$$

$$\hat{p}_{n+1}(0, 0 | m, \ell) = p_{n+1}^Y(0|\ell),$$

$$\hat{p}_{n+1}(k, 0 | m, \ell) = 0 \quad \text{for } k \neq 0.$$

Furthermore, for $j \neq 0$, we set

$$\hat{p}_{n+1}(0, j, J_{n+1}' = n+1 | m, \ell) = \frac{\left(\frac{1}{n+1} - p_{n+1}^Y(0|\ell)\right) p_{n+1}^Y(j|\ell)}{1 - p_{n+1}^Y(0|\ell)},$$

$$\hat{p}_{n+1}(0, j, J_{n+1}' \neq n+1 | m, \ell) = 0, \quad \text{for } m \neq 0$$

and

$$\hat{p}_{n+1}(0, j, J_{n+1}' \neq n+1 | 0, \ell) = \frac{(p_{n+1}^X(0|0) - \frac{1}{n+1}) p_{n+1}^Y(j|\ell)}{1 - p_{n+1}^Y(0|\ell)},$$

so that

$$\hat{p}_{n+1}(0, j | m, \ell) = \frac{(p_{n+1}^X(0|m) - p_{n+1}^Y(0|\ell)) p_{n+1}^Y(j|\ell)}{1 - p_{n+1}^Y(0|\ell)}$$

for $j \neq 0$. For $\ell = 0$ (and consequently $m \neq 0$) we set

$$\hat{p}_{n+1}(k, j|m, 0) = p^X_{n+1}(k|m) \cdot p^Y_{n+1}(j|0).$$

If the chains are already in the same state, i.e. $\boldsymbol{\ell} = \boldsymbol{m}$, we define the coupling in such a way that they stay together with maximal probability. Thus, with Lemma 4.3, we get

$$\hat{p}_{n+1}(0, 0|m, m) = \min\{p^X_{n+1}(0|m), p^Y_{n+1}(0|m)\} = p^Y_{n+1}(0|m) \text{ for } m \neq 0,$$

$$\hat{p}_{n+1}(m, m|m, m) = \min\{p^X_{n+1}(m|m), p^Y_{n+1}(m|m)\} = p^X_{n+1}(m|m) \text{ and}$$

$$\hat{p}_{n+1}(m+1, m+1|m, m) = \min\{p^X_{n+1}(m+1|m), p^Y_{n+1}(m+1|m)\} = p^Y_{n+1}(m+1|m).$$

For $k = 0$ and $m \neq 0$ we put

$$\hat{p}_{n+1}(0, m|m, m) = \frac{\left(p^X_{n+1}(0|m) - p^Y_{n+1}(0|m)\right)\left(p^Y_{n+1}(m|m) - p^X_{n+1}(m|m)\right)}{1 - \left(p^Y_{n+1}(0|m) + p^Y_{n+1}(m+1|m) + p^X_{n+1}(m|m)\right)},$$

$$\hat{p}_{n+1}(0, m+1|m, m) = 0,$$

and

$$\hat{p}_{n+1}(0, j|m, m) = \frac{\left(p^X_{n+1}(0|m) - p^Y_{n+1}(0|m)\right) p^Y_{n+1}(j|m)}{1 - \left(p^Y_{n+1}(0|m) + p^Y_{n+1}(m+1|m) + p^X_{n+1}(m|m)\right)},$$

for $j \notin \{0, m, m+1\}$. For $k = m$ we put

$$\hat{p}_{n+1}(m, j|m, m) = 0, \tag{4.20}$$

for all $j \neq m$. For $k = m + 1$ we let

$$\hat{p}_{n+1}(m+1, 0|m, m) = 0 \quad \text{for } m \neq 0,$$

$$\hat{p}_{n+1}(m+1, m|m, m) = \frac{\left(p^X_{n+1}(m+1|m) - p^Y_{n+1}(m+1|m)\right)\left(p^Y_{n+1}(m|m) - p^X_{n+1}(m|m)\right)}{1 - (p^Y_{n+1}(0|m) + p^Y_{n+1}(m+1|m) + p^X_{n+1}(m|m))}$$

and

$$\hat{p}_{n+1}(m+1, j|m, m) = \frac{\left(p^X_{n+1}(m+1|m) - p^Y_{n+1}(m+1|m)\right) p^Y_{n+1}(j|m)}{1 - (p^Y_{n+1}(0|m) + p^Y_{n+1}(m+1|m) + p^X_{n+1}(m|m))},$$

for $j \notin \{0, m, m+1\}$.

Note that the coupling is constructed in such a way that whenever the two chains are coupled, X cannot stay in any state without Y staying as well and Y can neither move to 0 without X nor can it move one state up without X following.

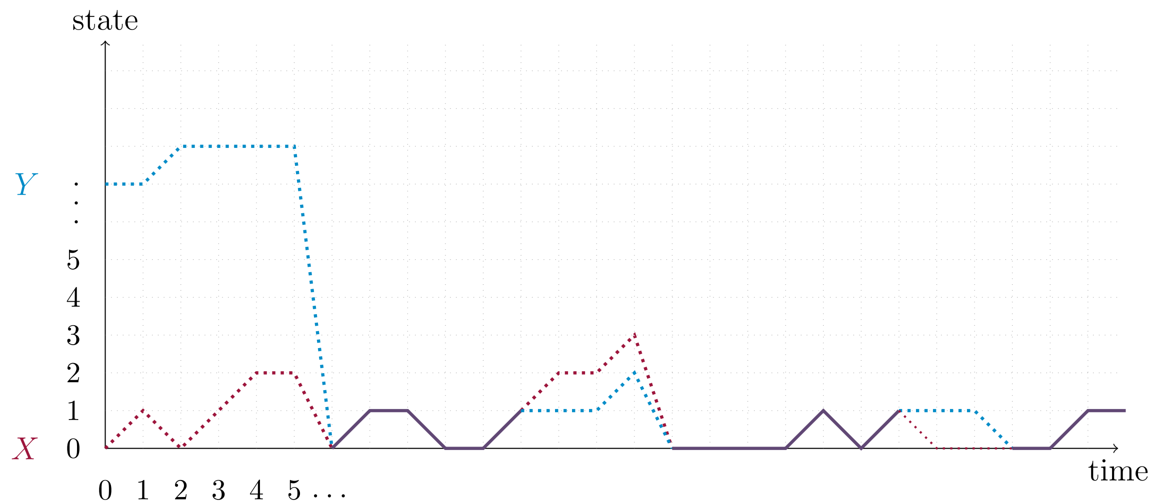


Figure 4.1: The paths of X and Y are depicted as a dotted red, resp. blue lines. At times at which the two chains are moving together, their path is displayed as a continuous violet line.

We now show that this in fact defines a coupling of the two chains. For $\boldsymbol{\ell \neq m, \ell \neq 0}$ we have

$$\begin{aligned} p_{n+1}^{X'}(0|m,\ell) &= \sum_{j\geq 0} \hat{p}_{n+1}(0,j|m,\ell) \\ &= p_{n+1}^{Y}(0|\ell) + \left(p_{n+1}^{X}(0|m) - p_{n+1}^{Y}(0|\ell)\right) \sum_{j\geq 1} \frac{p_{n+1}^{Y}(j|\ell)}{1-p_{n+1}^{Y}(0|\ell)} \\ &= p_{n+1}^{X}(0|m) \end{aligned}$$

and for $k \neq 0$

$$\begin{aligned} p_{n+1}^{X'}(k|m,\ell) = \sum_{j\geq 0} \hat{p}_{n+1}(k,j|m,\ell) &= \sum_{j\geq 1} \hat{p}_{n+1}(k,j|m,\ell) \\ &= \sum_{j\geq 1} \frac{p_{n+1}^{X}(k|m)\cdot p_{n+1}^{Y}(j|\ell)}{1-p_{n+1}^{Y}(0|\ell)}, \\ &= p_{n+1}^{X}(k|m). \end{aligned}$$

We now look at the case $\boldsymbol{\ell \neq m}$ but $\boldsymbol{\ell = 0}$ (e.g. $m \neq 0$) and obtain

$$p_{n+1}^{X}(k|m,0) = \sum_{j\geq 0} \hat{p}_{n+1}(k,j|m,0) = p_{n+1}^{X}(k|m) \sum_{j\geq 0} p_{n+1}^{Y}(j|0) = p_{n+1}^{X}(k|m).$$

We can now proceed analogously for Y. If $\boldsymbol{j, \ell \neq 0}$ we have

$$
\begin{aligned}
&p^{Y'}_{n+1}(j|m,\ell) \\
&= \sum_{k\geq 0} \hat{p}_{n+1}(k,j|m,\ell) = \hat{p}_{n+1}(0,j|m,\ell) + \sum_{k\geq 1} \hat{p}_{n+1}(k,j|m,\ell) \\
&= \left(p^X_{n+1}(0|m) - p^Y_{n+1}(0|\ell)\right) \cdot \frac{p^Y_{n+1}(j|\ell)}{1-p^Y_{n+1}(0|\ell)} + \sum_{k\geq 1} \frac{p^X_{n+1}(k|m)\cdot p^Y_{n+1}(j|\ell)}{1-p^Y_{n+1}(0|\ell)} \\
&= p^Y_{n+1}(j|\ell)\Big(\frac{p^X_{n+1}(0|m) - p^Y_{n+1}(0|\ell)}{1-p^Y_{n+1}(0|\ell)} + \frac{1-p^X_{n+1}(0|m)}{1-p^Y_{n+1}(0|\ell)}\Big) \\
&= p^Y_{n+1}(j|\ell).
\end{aligned}
$$

For $j = 0$ and still $\boldsymbol{\ell \neq 0}$, we get

$$
\begin{aligned}
p^{Y'}_{n+1}(0|m,\ell) &= \sum_{k\geq 0} \hat{p}_{n+1}(k,0|m,\ell) \\
&= \hat{p}_{n+1}(0,0|m,\ell) = p^Y_{n+1}(0|\ell).
\end{aligned}
$$

In the case $\boldsymbol{\ell = 0}$ we have

$$
p^{Y'}_{n+1}(j|m,0) = \sum_{k\geq 0} \hat{p}_{n+1}(k,j|m,0) = \sum_{k\geq 0} p^X_{n+1}(k|m) p^Y_{n+1}(j|0) = p^Y_{n+1}(j|0).
$$

For X and the case $\boldsymbol{\ell = m \neq 0}$ the situation is the following:

$$
\begin{aligned}
&p^{X'}_{n+1}(0|m,m) = \sum_{j\geq 0} \hat{p}_{n+1}(0,j|m,m) \\
&= p^Y_{n+1}(0|m) + \frac{\left(p^X_{n+1}(0|m) - p^Y_{n+1}(0|m)\right)\left(p^Y_{n+1}(m|m) - p^X_{n+1}(m|m)\right)}{1-\left(p^Y_{n+1}(0|m) + p^Y_{n+1}(m+1|m) + p^X_{n+1}(m|m)\right)} \\
&\quad + \sum_{j\notin\{0,m,m+1\}} \frac{\left(p^X_{n+1}(0|m) - p^Y_{n+1}(0|m)\right) p^Y_{n+1}(j|m)}{1-\left(p^Y_{n+1}(0|m) + p^Y_{n+1}(m+1|m) + p^X_{n+1}(m|m)\right)}
\end{aligned}
$$

$$
\begin{aligned}
&= p^Y_{n+1}(0|m) \\
&\quad + \left(p^X_{n+1}(0|m) - p^Y_{n+1}(0|m)\right)\Bigg(\frac{p^Y_{n+1}(m|m) - p^X_{n+1}(m|m)}{1-\left(p^Y_{n+1}(0|m) + p^Y_{n+1}(m+1|m) + p^X_{n+1}(m|m)\right)}
\end{aligned}
$$

$$+\frac{1-(p^Y_{n+1}(0|m)+p^Y_{n+1}(m|m)+p^Y_{n+1}(m+1|m))}{1-\left(p^Y_{n+1}(0|m)+p^Y_{n+1}(m+1|m)+p^X_{n+1}(m|m)\right)}\Bigg)$$

$$=p^X_{n+1}(0|m).$$

Furthermore, for all $m \geq 0$, we have

$$p^{X'}_{n+1}(m|m,m)=\hat{p}_{n+1}(m,m|m,m)=p^X_{n+1}(m|m)$$

by (4.20), as well as

$$p^{X'}_{n+1}(m+1|m,m)=\sum_{j\geq 0}\hat{p}_{n+1}(m+1,j|m,m)$$

$$=\frac{\left(p^X_{n+1}(m+1|m)-p^Y_{n+1}(m+1|m)\right)\left(p^Y_{n+1}(m|m)-p^X_{n+1}(m|m)\right)}{1-\left(p^Y_{n+1}(0|m)+p^Y_{n+1}(m+1|m)+p^X_{n+1}(m|m)\right)}+p^Y_{n+1}(m+1|m)$$

$$+\sum_{j\notin\{0,m,m+1\}}\frac{\left(p^X_{n+1}(m+1|m)-p^Y_{n+1}(m+1|m)\right)p^Y_{n+1}(j|m)}{1-\left(p^Y_{n+1}(0|m)+p^Y_{n+1}(m+1|m)+p^X_{n+1}(m|m)\right)}$$

$$=p^Y_{n+1}(m+1|m)+\left(p^X_{n+1}(m+1|m)-p^Y_{n+1}(m+1|m)\right)$$

$$=p^X_{n+1}(m+1|m).$$

Performing the analogous calculations for Y and $m \neq 0$ yields

$$p^{Y'}_{n+1}(0|m,m)=\hat{p}_{n+1}(0,0|m,m)=p^Y_{n+1}(0|m)$$

as well as for $m \geq 0$

$$p^{Y'}_{n+1}(m+1|m,m)=\hat{p}_{n+1}(m+1,m+1|m,m)=p^Y_{n+1}(m+1|m),$$

$$p^{Y'}_{n+1}(m|m,m)=\sum_{k\in\{0,m,m+1\}}\hat{p}_{n+1}(k,m|m,m)$$

$$=p^X_{n+1}(m|m)+\frac{\left(p^X_{n+1}(0|m)-p^Y_{n+1}(0|m)\right)\left(p^Y_{n+1}(m|m)-p^X_{n+1}(m|m)\right)}{1-\left(p^Y_{n+1}(0|m)+p^Y_{n+1}(m+1|m)+p^X_{n+1}(m|m)\right)}$$

$$+\frac{\left(p^X_{n+1}(m+1|m)-p^Y_{n+1}(m+1|m)\right)\left(p^Y_{n+1}(m|m)-p^X_{n+1}(m|m)\right)}{1-\left(p^Y_{n+1}(0|m)+p^Y_{n+1}(m+1|m)+p^X_{n+1}(m|m)\right)}$$

$$=p^X_{n+1}(m|m)+\left(p^Y_{n+1}(m|m)-p^X_{n+1}(m|m)\right)\cdot 1=p^Y_{n+1}(m|m)$$

and for $j \notin \{0,m,m+1\}$

$$p^{Y'}_{n+1}(j|m,m)=\sum_{k\in\{0,m,m+1\}}\hat{p}_{n+1}(k,j|m,m)$$

$$= \frac{\left(p^X_{n+1}(0|m) - p^Y_{n+1}(0|m)\right) p^Y_{n+1}(j|m)}{1 - \left(p^Y_{n+1}(0|m) + p^Y_{n+1}(m+1|m) + p^X_{n+1}(m|m)\right)} + 0$$

$$+ \frac{\left(p^X_{n+1}(m+1|m) - p^Y_{n+1}(m+1|m)\right) p^Y_{n+1}(j|m)}{1 - \left(p^Y_{n+1}(0|m) + p^Y_{n+1}(m+1|m) + p^X_{n+1}(m|m)\right)}$$

$$= p^Y_{n+1}(j|m).$$

Thus $(X_n^{'}, Y_n^{'})$ is a Markovian coupling of X_n and Y_n on $\mathbb{N} \times \mathbb{N}$. □

As mentioned before, we need to bound the probability that the coupled processes are not in the same state at a given time in order to prove Theorem 4.1. To do so, we establish a bound on the likelihood that the chains do not meet, even though the discrete time chain X moves to 0, in the following lemma.

Lemma 4.6. *Let the birthtime of the vertex we look at time n be given by the Markov chain J_n with $J_1 = 1$ and*

$$\mathbb{P}(J_{n+1} = J_n \mid J_n) = \frac{n}{n+1} \quad \textit{and} \quad \mathbb{P}(J_{n+1} = n+1 \mid J_n) = \frac{1}{n+1},$$

just as in the proof of Lemma 3.12. Furthermore we define F_k as the event that we look at the indegree of vertex k at time k and $G_{k,n}$ as the event that we look at the same vertex from time k to time n, i.e.

$$F_k := \{J_k = k\} \quad \textit{and} \quad G_{k,n} := \{J_j = J_k \ \forall k+1 \le j \le n\}..$$

We then get

$$\mathbb{P}(X_{k+1}^{'} \neq Y_{k+1}^{'} | F_{k+1}, G_{k+1,n}) \le 3 \ \frac{\mathbb{E}\left[(f(Y_k+1)+1)^2\right]}{k+1}. \tag{4.21}$$

Proof. Since for all $k \ge 1$ the random variable J_k is independent of the states the processes are in at time $k-1$, we have

$$\begin{aligned}
&\mathbb{P}(X_{k+1}^{'} \neq Y_{k+1}^{'} | F_{k+1}, G_{k+1,n}) = \mathbb{P}(Y_{k+1}^{'} \neq 0 | F_{k+1}) \\
&= \sum_{m=1}^{k} \sum_{l=0}^{\infty} \mathbb{P}(Y_{k+1}^{'} \neq 0 | F_{k+1}, X_{k+1}^{'} = 0, Y_k^{'} = l, X_k^{'} = m) \mathbb{P}(Y_k^{'} = l, X_k^{'} = m | F_{k+1}) \\
&\quad + \sum_{l=0}^{\infty} \mathbb{P}(Y_{k+1}^{'} \neq 0 | F_{k+1}, X_{k+1}^{'} = 0, Y_k^{'} = l, X_k^{'} = 0) \mathbb{P}(X_k^{'} = 0, Y_k^{'} = \ell) \\
&:= A_1 + A_2,
\end{aligned}$$

where we used that F_{k+1} implies $\{X_{k+1} = 0\}$ and that J_{k+1} is independent of $X_k^{'}$ and $Y_k^{'}$. Using the mentioned independence again, we obtain

$$A_1 = \sum_{m=1}^{k} \sum_{l=0}^{\infty} \frac{\mathbb{P}(Y_{k+1}^{'} \neq 0, X_{k+1}^{'} = 0 | Y_k^{'} = l, X_k^{'} = m)}{\mathbb{P}(F_{k+1} | Y_k^{'} = l, X_k^{'} = m)} \mathbb{P}(Y_k^{'} = l, X_k^{'} = m)$$

$$
\begin{aligned}
&= \sum_{m=1}^{k}\sum_{l=0}^{\infty} \mathbb{P}(Y_{k+1}^{'} \neq 0, X_{k+1}^{'} = 0 | Y_k^{'} = l, X_k^{'} = m)(k+1)\mathbb{P}(Y_k^{'} = l, X_k^{'} = m) \\
&= (k+1)\sum_{m=1}^{k}\sum_{l=0}^{\infty}\sum_{j=1}^{\infty} \hat{p}_{k+1}(0, j|m, l)\mathbb{P}(X_k^{'} = m, Y_k^{'} = \ell) \\
&= (k+1)\sum_{m=1}^{k}\sum_{\substack{\ell=0,\\ \ell\neq m}}^{\infty}\sum_{j=1}^{\infty} \hat{p}_{k+1}(0, j|m, \ell)\mathbb{P}(X_k^{'} = m, Y_k^{'} = \ell) \\
&\quad + (k+1)\sum_{m=1}^{k}\sum_{j=1}^{\infty} \hat{p}_{k+1}(0, j|m, m)\mathbb{P}(X_k^{'} = m, Y_k^{'} = m) \\
&:= A_{1,1} + A_{1,2}.
\end{aligned}
$$

We can now bound both terms by exploiting the construction of our coupling described in Lemma 4.5. Hence,

$$
\begin{aligned}
A_{1,1} &= (k+1)\sum_{m=1}^{k}\sum_{\ell=0,\ell\neq m}^{\infty}\sum_{j=1}^{\infty} \hat{p}_{k+1}(0, j|m, \ell)\mathbb{P}(X_k^{'} = m, Y_k^{'} = \ell) \\
&= (k+1)\sum_{m=1}^{k}\Bigg(\sum_{\ell=1,\ell\neq m}^{\infty}\sum_{j=1}^{\infty} \hat{p}_{k+1}(0, j|m, \ell)\mathbb{P}(X_k^{'} = m, Y_k^{'} = \ell) \\
&\qquad\qquad + \sum_{j=1}^{\infty} \hat{p}_{k+1}(0, j|m, 0)\mathbb{P}(X_k^{'} = m, Y_k^{'} = 0)\Bigg) \\
&= (k+1)\sum_{m=1}^{k}\Big(\sum_{\substack{\ell=1,\\ \ell\neq m}}^{\infty}\sum_{j=1}^{\infty} \frac{(p_{k+1}^{X}(0|m) - p_{k+1}^{Y}(0|\ell))p_{k+1}^{Y}(j|\ell)}{1 - p_{k+1}^{Y}(0|\ell)}\mathbb{P}(X_k^{'} = m, Y_k^{'} = \ell) \\
&\qquad + \sum_{j=1}^{\infty} p_{k+1}^{X}(0|m)p_{k+1}^{Y}(j|0)\mathbb{P}(X_k^{'} = m, Y_k^{'} = 0)\Big) \\
&= (k+1)\sum_{m=1}^{k}\Big(\sum_{\ell=1,\ell\neq m}^{\infty} (p_{k+1}^{X}(0|m) - p_{k+1}^{Y}(0|\ell))\mathbb{P}(X_k^{'} = m, Y_k^{'} = \ell) \\
&\qquad\qquad + p_{k+1}^{X}(0|m)(1 - p_{k+1}^{Y}(0|0))\mathbb{P}(X_k^{'} = m, Y_k^{'} = 0)\Big) \\
&\leq (k+1)\sum_{m=1}^{k}\sum_{\ell=1,\ell\neq m}^{\infty} \frac{(f(\ell+1)+1)^2}{(k+1)^2}\mathbb{P}(X_k^{'} = m, Y_k^{'} = \ell) \\
&\quad + (k+1)\sum_{m=1}^{k}\frac{1}{k+1}\left(1 - \left(1 - \frac{f(0)}{k+1}\right)\right)\mathbb{P}(X_k^{'} = m, Y_k^{'} = 0)
\end{aligned}
$$

$$\leq \frac{\mathbb{E}\left[(f(Y_k+1)+1)^2\right]}{k+1}, \tag{4.22}$$

where we used (4.7). Now using the transition probabilities of the coupling in the case $\ell = m$ and (4.7) again gives

$$\begin{aligned}
A_{1,2} &= (k+1)\sum_{m=1}^{k} \mathbb{P}(X_k^{'} = m, Y_k^{'} = m) \\
&\quad \cdot \left(\sum_{j \notin \{0,m,m+1\}} \frac{(p^X_{n+1}(0|m) - p^Y_{n+1}(0|m))p^Y_{n+1}(j|m)}{1-(p^Y_{n+1}(0|m) + p^Y_{n+1}(m+1|m) + p^X_{n+1}(m|m))} \right. \\
&\quad \left. + \frac{(p^X_{n+1}(0|m) - p^Y_{n+1}(0|m))(p^Y_{n+1}(m|m) - p^X_{n+1}(m|m))}{1-(p^Y_{n+1}(0|m) + p^Y_{n+1}(m+1|m) + p^X_{n+1}(m|m))} \right) \\
&= (k+1)\sum_{m=1}^{k}(p^X_{n+1}(0|m) - p^Y_{n+1}(0|m))\mathbb{P}(X_k^{'} = m, Y_k^{'} = m) \\
&\leq (k+1)\sum_{m=1}^{k} \frac{(f(m+1)+1)^2}{(k+1)^2}\mathbb{P}(Y_k^{'} = m) \\
&\leq \frac{\mathbb{E}\left[(f(Y_k+1)+1)^2\right]}{k+1}.
\end{aligned} \tag{4.23}$$

It remains to deal with the case $m = 0$, thus to bound A_2. As a result of the case distinction for the transition probability $\hat{p}_{n+1}(0,j|0,\ell)$ outlined in the proof of Lemma 4.5 we obtain

$$\begin{aligned}
A_2 &= \sum_{l\geq 1}^{\infty}\sum_{j\geq 1} \mathbb{P}(Y_{k+1}^{'} = j | F_{k+1}, X_{k+1}^{'} = 0, Y_k^{'} = l, X_k^{'} = 0)\mathbb{P}(X_k^{'} = 0, Y_k^{'} = \ell) \\
&\quad + \sum_{j\geq 1} \mathbb{P}(Y_{k+1}^{'} = j | F_{k+1}, X_{k+1}^{'} = 0, Y_k^{'} = 0, X_k^{'} = 0)\mathbb{P}(X_k^{'} = 0, Y_k^{'} = 0) \\
&\leq (k+1)\sum_{l\geq 1}\sum_{j\geq 1} \mathbb{P}(X_{k+1}^{'} = 0, Y_{k+1}^{'} = j, F_{k+1} | Y_k^{'} = l, X_k^{'} = 0)\mathbb{P}(X_k^{'} = 0, Y_k^{'} = \ell) \\
&\quad + (k+1)\sum_{j\geq 1} \mathbb{P}(X_{k+1}^{'} = 0, Y_{k+1}^{'} = j, F_{k+1} | Y_k^{'} = 0, X_k^{'} = 0)\mathbb{P}(X_k^{'} = 0, Y_k^{'} = 0) \\
&\leq (k+1)\sum_{l\geq 1}\sum_{j\geq 1} \frac{(\frac{1}{k+1} - p^Y_{k+1}(0|\ell))p^Y_{k+1}(j|\ell)}{1-p^Y_{k+1}(0|\ell)}\mathbb{P}(X_k^{'} = 0, Y_k^{'} = \ell) \\
&\quad + (k+1)\sum_{j\geq 1}\hat{p}_{k+1}(0,j|0,0)\mathbb{P}(X_k^{'} = 0, Y_k^{'} = 0) \\
&= (k+1)\sum_{l\geq 1}(\frac{1}{k+1} - p^Y_{k+1}(0|\ell))\mathbb{P}(X_k^{'} = 0, Y_k^{'} = \ell)
\end{aligned}$$

$$
\begin{aligned}
&+ (k+1)\sum_{j\geq 2} \hat{p}_{k+1}(0,j|0,0)\mathbb{P}(X_k^{'}=0, Y_k^{'}=0) \\
&= (k+1)\sum_{l\geq 1}(p_{k+1}^X(0|l) - p_{k+1}^Y(0|\ell))\mathbb{P}(X_k^{'}=0, Y_k^{'}=\ell) \\
&+ (k+1)\sum_{j\geq 2}(p_{k+1}^X(0,|0) - p_{k+1}^Y(0|0))p_{k+1}^Y(j|0)\mathbb{P}(X_k^{'}=0, Y_k^{'}=0) \\
&\leq (k+1)\sum_{l\geq 1}\frac{(f(\ell+1)+1)^2}{(k+1)^2}\mathbb{P}(X_k^{'}=0, Y_k^{'}=\ell) \\
&+ (k+1)\frac{(f(1)+1)^2}{(k+1)^2}\mathbb{P}(X_k^{'}=0, Y_k^{'}=0) \\
&\leq \frac{\mathbb{E}\left[(f(Y_k+1)+1)^2\right]}{k+1},
\end{aligned}
$$

where we used (4.7). In combination with (4.22) and (4.23), this yields the desired result. □

Proof of Theorem 4.1. As $Z \sim W$ and Y_n denotes the observation of Z at deterministic time instances we have

$$
d_{TV}\left(W_n, W\right) = d_{TV}\left(X_n, Z\right) = d_{TV}\left(X_n, Y_n\right).
$$

Furthermore, remember that by equation 4.1 we have

$$
d_{TV}(X_n, Y_n) \leq \mathbb{P}(X_n^{'} \neq Y_n^{'})
$$

for any coupling $(X_n^{'}, Y_n^{'})$ of X_n and Y_n. In particular, this is valid for the coupling described in the previous section.
Using the construction of J_n we can handle the events defined in Lemma 4.6. We have

$$
\mathbb{P}(F_k) = \frac{1}{k} \quad \text{and} \quad \mathbb{P}(G_{k,n}) = \prod_{j=k+1}^{n}(1-\frac{1}{j}) = \frac{k}{n}
$$

and for $E_{k,n} := F_k \cap G_{k,n}$ we then obtain

$$
\mathbb{P}(E_{k,n}) = \mathbb{P}(F_k \cap G_{k,n}) = \mathbb{P}(F_k)\mathbb{P}(G_{k,n}) = \frac{1}{n},
$$

where we used that J_k and J_i are independent for $i \neq k$.

With Lemma 4.6 we obtain

$$
\mathbb{P}(X_n^{'} \neq Y_n^{'}) = \sum_{k=2}^{n}\mathbb{P}(X_n^{'} \neq Y_n^{'}, E_{k,n}) + \mathbb{P}(X_n^{'} \neq Y_n^{'}, G_{1,n})
$$

$$
\begin{aligned}
&\leq \frac{1}{n}\sum_{k=2}^{n}\Big(\mathbb{P}(X_n^{'} \neq Y_n^{'}|E_{k,n}, X_k^{'} = Y_k^{'})\mathbb{P}(X_k^{'} = Y_k^{'}|E_{k,n}) \\
&\quad +\mathbb{P}(X_n^{'} \neq Y_n^{'}|E_{k,n}, X_k^{'} \neq Y_k^{'})\mathbb{P}(X_k^{'} \neq Y_k^{'}|E_{k,n})\Big) + \frac{1}{n} \\
&\leq \frac{1}{n}\sum_{k=2}^{n}\mathbb{P}(X_n^{'} \neq Y_n^{'}|E_{k,n}, X_k^{'} = Y_k^{'})\mathbb{P}(X_k^{'} = Y_k^{'}|E_{k,n}) \\
&\quad + \frac{1}{n}\sum_{k=2}^{n}\frac{\mathbb{E}\left[(f(Y_k+1)+1)^2\right]}{k} + \frac{1}{n} \\
&:= R_1 + R_2 + \frac{1}{n}. \qquad (4.24)
\end{aligned}
$$

In order to deal with R_1, we define $H_{k,\ell}$ as the event that the processes evolve together from time k to ℓ and drift apart afterwards, hence

$$
H_{k,\ell} = \{X_i^{'} = Y_i^{'} \ \forall k \leq i \leq \ell, X_{\ell+1}^{'} \neq Y_{\ell+1}^{'}\}.
$$

With this notation we get

$$
\begin{aligned}
&\mathbb{P}(X_n^{'} \neq Y_n^{'}|E_{k,n}, X_k^{'} = Y_k^{'})\mathbb{P}(X_k^{'} = Y_k^{'}|E_{k,n}) \\
&\leq \sum_{l=k}^{n-1}\mathbb{P}(X_n^{'} \neq Y_n^{'}, H_{k,\ell}|G_{k,n}, F_k, X_k^{'} = Y_k^{'} = 0) \\
&\leq \sum_{\ell=k}^{n-1}\mathbb{P}(H_{k,\ell}|G_{k,l}, X_k^{'} = Y_k^{'} = 0) \\
&\leq \sum_{\ell=k}^{n-1}\frac{\mathbb{P}(H_{k,\ell}|X_k^{'} = Y_k^{'} = 0)}{\mathbb{P}(G_{k,\ell}|X_k^{'} = Y_k^{'} = 0)} \\
&\leq \sum_{\ell=k}^{n-1}\frac{\ell}{k}\mathbb{P}(X_{\ell+1}^{'} \neq Y_{\ell+1}^{'}|X_\ell^{'} = Y_\ell^{'}, X_k^{'} = Y_k^{'} = 0),
\end{aligned}
$$

where we used again that J_j is independent of $X_k^{'}$ and $Y_k^{'}$ for $k+1 \leq j \leq n$.

Furthermore,

$$
\begin{aligned}
&\mathbb{P}(X_{\ell+1}^{'} \neq Y_{\ell+1}^{'}|X_\ell^{'} = Y_\ell^{'}, X_k^{'} = Y_k^{'} = 0) \\
&= \sum_{m=0}^{\infty}\mathbb{P}(X_{\ell+1}^{'} \neq Y_{\ell+1}^{'}|X_\ell^{'} = Y_\ell^{'} = m, X_k^{'} = Y_k^{'} = 0)\mathbb{P}(Y_\ell^{'} = m|X_\ell^{'} = Y_\ell^{'}, X_k^{'} = Y_k^{'} = 0) \\
&= \sum_{m=0}^{\infty}\mathbb{P}(Y_\ell^{'} = m|X_\ell^{'} = Y_\ell^{'}, X_k^{'} = Y_k^{'} = 0)
\end{aligned}
$$

$$
\begin{aligned}
&\cdot\Big(\mathbb{P}(Y'_{\ell+1}\neq 0|X'_{\ell+1}=0,X'_\ell=Y'_\ell=m)p^{X'}_{\ell+1}(0|m,m)\\
&\quad+\mathbb{P}(Y'_{\ell+1}\neq m|X'_{\ell+1}=m,X'_\ell=Y'_\ell=m)p^{X'}_{\ell+1}(m|m,m)\\
&\quad+\mathbb{P}(Y'_{\ell+1}\neq m+1|X'_{\ell+1}=m+1,X'_\ell=Y'_\ell=m)p^{X'}_{\ell+1}(m+1|m,m)\Big)\\
=&\sum_{m=0}^{\infty}\mathbb{P}(Y'_\ell=m|X'_\ell=Y'_\ell,X'_k=Y'_k=0)\\
&\cdot\Big(\Big(1-\mathbb{P}(Y'_{\ell+1}=0|X'_{\ell+1}=0,X'_\ell=Y'_\ell=m)\Big)p^{X'}_{\ell+1}(0|m,m)+0\\
&+\Big(1-\mathbb{P}(Y'_{\ell+1}=m+1|X'_{\ell+1}=m+1,X'_\ell=Y'_\ell=m)\Big)p^{X'}_{\ell+1}(m+1|m,m)\Big)\\
=&\sum_{m=0}^{\infty}\mathbb{P}(Y'_\ell=m|X'_\ell=Y'_\ell,X'_k=Y'_k=0)\\
&\cdot\Big(p^{X'}_{n+1}(0|m)-\hat{p}_{\ell+1}(0,0|m,m)+p^{X'}_{n+1}(m+1|m)-\hat{p}_{\ell+1}(m+1,m+1|m,m)\Big)\\
=&\sum_{m=1}^{\infty}\mathbb{P}(Y'_\ell=m|X'_\ell=Y'_\ell,X'_k=Y'_k=0)\\
&\cdot\Big(\big(p^{X}_{\ell+1}(0|m)-p^{Y}_{\ell+1}(0|m)\big)+\big(p^{X}_{\ell+1}(m+1|m)-p^{Y}_{\ell+1}(m+1|m)\big)\Big)\\
&+\mathbb{P}(Y'_\ell=0|X'_\ell=Y'_\ell,X'_k=Y'_k=0)(p^{X}_{\ell+1}(1|0)-p^{Y}_{\ell+1}(1|0))\\
\leq&\sum_{m=0}^{\infty}\mathbb{P}(Y'_\ell=m|X'_\ell=Y'_\ell,X'_k=Y'_k=0)\cdot 2\max_j\big|p^{X}_{\ell+1}(j|m)-p^{Y}_{\ell+1}(j|m)\big|\\
\leq&\, 2\sum_{m=0}^{\infty}\mathbb{P}(Y'_\ell=m|X'_\ell=Y'_\ell,X'_k=Y'_k=0)\frac{(f(m+1)+1)^2}{(\ell+1)^2}\\
=&\, 2\,\frac{\mathbb{E}_{\ell,k}\left[(f(Y'_\ell+1)+1)^2\right]}{(\ell+1)^2}\leq 2\,\frac{\mathbb{E}\left[(f(Y'_\ell+1)+1)^2\right]}{(\ell+1)^2}=2\,\frac{\mathbb{E}\left[(f(Y_\ell+1)+1)^2\right]}{(\ell+1)^2},
\end{aligned}
$$

where $\mathbb{E}_{\ell,k}$ denotes the expectation with respect to the probability measure $\mathbb{P}(\cdot|X'_\ell=Y'_\ell,X'_k=Y'_k=0)$ and the last inequality follows from the fact that changing to the newest vertex is independent of the stage the processes are in, so that conditioning on $X'_k=Y'_k=0$ decreases the expectation.
Remembering that Z denotes the continuous-time process corresponding to Y and that $\psi(\ell)=\sum_{m=1}^{\ell}\frac{1}{m}$, we define the random variable

$$M^{\psi}_\ell:=\mathbb{1}\{Z\text{ does not fall between }\psi(\ell)\text{ and }\psi(\ell+1)\}\leq\mathbb{1}\{Z_{\psi(\ell+1)}\geq Z_{\psi(\ell)}\}$$

and get

$$\mathbb{E}\left[\int_{\psi(\ell)}^{\psi(\ell+1)} f(Z_s)^2\,\mathrm{d}s\right] \geq \mathbb{E}\left[\mathbb{1}\{Z_{\psi(\ell+1)} \geq Z_{\psi(\ell)}\}\int_{\psi(\ell)}^{\psi(\ell+1)} f(Z_s)^2\,\mathrm{d}s\right]$$
$$\geq \frac{1}{\ell+1}\mathbb{E}\left[M_\ell^\psi\; f(Z_{\psi(\ell)})^2\,\mathrm{d}s\right]$$
$$= \frac{1}{\ell+1}\mathbb{E}\left[M_\ell^\psi\; f(Y_\ell)^2\,\mathrm{d}s\right]$$
$$\geq \frac{1}{\ell+1}\left(1-\frac{1}{\ell+1}\right)\mathbb{E}\left[f(Y_\ell)^2\right]$$
$$= \frac{\ell}{(\ell+1)^2}\mathbb{E}\left[f(Y_\ell)^2\right], \tag{4.25}$$

where the last inequality uses that the process always falls at rate 1, independent of the state the process is in. By Lemma 4.3

$$\mathbb{P}(Y_{\ell+1} \neq 0 | Y_\ell = k) \geq 1 - \frac{1}{\ell+1}$$

for $k \neq 0$, which we can assume without loss of generality, as in the case $\psi(\ell) = 0$ we have $\mathbb{P}(\psi(\ell+1) \geq \psi(\ell)) = 1$ and the previous inequalities remain valid. (4.25) now yields

$$\sum_{l=k}^{n} \frac{\mathbb{E}\left[f(Y_\ell)^2\right]}{(\ell+1)k} \leq \frac{1}{k}\left(\frac{\ell+1}{\ell}\right)\sum_{l=k}^{n}\mathbb{E}\left[\int_{\psi(\ell)}^{\psi(\ell+1)} f(Z_s)^2\,\mathrm{d}s\right] \leq \frac{2}{k}\mathbb{E}\left[\int_{\psi(k)}^{\psi(n+1)} f(Z_s)^2\,\mathrm{d}s\right].$$

We now have

$$\int_{\psi(k)}^{\psi(n+1)} f(Z_s)^2 ds \leq \int_0^{\log(n+1)} f(Z_s)^2\,\mathrm{d}s$$
$$= \log(n+1)\left(\frac{1}{\log(n+1)}\int_0^{\log(n+1)} f(Z_s)^2\,\mathrm{d}s\right)$$
$$\leq C\log(n+1),$$

since

$$\frac{1}{\log(n+1)}\int_0^{\log(n+1)} f(Z_s)^2\,\mathrm{d}s \longrightarrow_{n\to\infty} \mathbb{E}\left[f(Z)^2\right] < \infty$$

by the Ergodic Theorem for stationary Markov processes in continuous-time and $\mathbb{E}\left[f(Z)^2\right] < \infty$ by assumption. In a similar way it follows, that $\int_{\psi(1)}^{\psi(n)} f(Z_s)\,\mathrm{d}s \leq C\log(n+1)$ as $\mathbb{E}\left[f(Z)\right] < \infty$. Consequently

$$\frac{1}{k}\sum_{\ell=k}^{n}\frac{\mathbb{E}\left[(f(Y_\ell+1)+1)^2\right]}{\ell+1} \leq \frac{1}{k}\sum_{\ell=k}^{n}\frac{\mathbb{E}\left[(f(Y_\ell)+2)^2\right]}{\ell+1}$$

$$
\begin{aligned}
&= \frac{1}{k} \sum_{\ell=k}^{n} \frac{\mathbb{E}\left[f(Y_\ell)^2 + 4f(Y_\ell)\right] + 4}{\ell + 1} \\
&\leq \frac{2}{k}\, \mathbb{E}\left[\int_{\psi(k)}^{\psi(n+1)} f(Z_s)^2 ds + 4 \int_{\psi(k)}^{\psi(n+1)} f(Z_s) ds + 4\right] \\
&\leq C\ \frac{\log(n+1)}{k}.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
R_1 &\leq \frac{2}{n} \sum_{k=1}^{n} \sum_{\ell=k}^{n} \frac{\mathbb{E}\left[(f(Y_\ell + 1) + 1)^2\right]}{\ell^2} \\
&\leq \frac{C \log(n+1)}{n} \sum_{k=1}^{n} \frac{1}{k} \leq C\ \frac{\log(n+1)^2}{n}.
\end{aligned} \tag{4.26}
$$

Similarly we obtain

$$
\frac{1}{n} \sum_{k=1}^{n} \frac{\mathbb{E}\left[f(Y_k)^2\right]}{k} \leq \frac{1}{n} \mathbb{E}\left[\int_{\psi(1)}^{\psi(n+1)} f(Z_s)^2 ds\right] \leq C\ \frac{\log(n+1)}{n}
$$

and thus

$$
R_2 \leq C\ \frac{\log(n+1)}{n}. \tag{4.27}
$$

Putting (4.26) and (4.27) into (4.24) proves the Theorem. □

4.2. Coupling with the Barabási-Albert graph

Additionally to the Barabási-Albert model ($PA_n^{m,\delta}$) described in section 3.2.4, a second, rather similar, model is described in [Hof17]. In the following we will refer to this as *model b* and denote by $PA_n^{m,\delta}(b)$ the graph at time n corresponding to this construction where $\delta \geq -1$ is a parameter of the model, allowing to change the attachment function. This model does not allow for self loops and satisfies our Assumptions **(A)** with attachment function $f(k) = k + \eta$ for some $\eta \in (0, 1)$. Thus Theorem 3.7 can be applied to model b. Again, we will use coupling techniques to obtain rates of convergence for the Barabási-Albert model.
First, let us give a more detailed description of the second model. We can restrict to the case $m = 1$, as for $m > 1$ the models are defined in terms of the model with $m = 1$. We start with a graph consisting of two vertices and two edges connecting them. As we do not allow for self-loops, connection rule (3.31) changes to

$$\mathbb{P}(n + 1 \to i | PA_n^{1,\alpha}(b)) = \frac{\alpha + D_n^{(b)}(i)}{n(2 + \alpha)} \text{ for } i \leq n, \tag{4.28}$$

where $\alpha \geq -1$ is the parameter of the model. By choosing α for each n in a suitable way (cf. exercise 8.7 in [Hof17]) we are able to construct a coupling, such that we can transfer our results to the Barabási-Albert model via the triangle inequality. Unfortunately, the coupling will turn out to be of order $\frac{\log(n)^2}{n}$, so that we do not get the same order of the rate as before. To formulate this result let $X_n^\alpha = \deg_n^{(b,\alpha)}(U_n)$ and $Y_n^\delta = \deg_n^{BA,\delta}(\tilde{U}_n)$, where $\deg_n^{(b,\alpha)}(i)$ denotes the indegree of vertex i in model b with parameter α, $\deg_n^{BA,\delta}(i)$ refers to the indegree of vertex i in the Barabási-Albert model with parameter δ and both U_n and $\tilde{U}_n$ are uniformly distributed on $[n]$. Note that $D_n^{(b)}(i) = \deg_n^{(b)}(i) + 1$ and $D_n^{BA}(i) = \deg_n^{BA}(i) + 1$, respectively.
We get the following theorem on the distance of X_n^α and Y_n^δ in the total variation metric.

Theorem 4.7. *For each $\delta \geq -1$ there exists a sequence of parameters $(\alpha_n)_{n\geq 1}$ and a coupling $(\hat{X}_n, \hat{Y}_n)$ of $X_n = X_n^{\alpha_n}$ and $Y_n = Y_n^\delta$ such that*

$$d_{TV}(X_n, Y_n) \leq \mathbb{P}(\hat{X}_n \neq \hat{Y}_n) \leq C \ \frac{\log(n)^2}{n}.$$

Following [Hof17, Exercise 8.7] we choose $\alpha_n = \frac{2\delta(n+1)}{2n+1}$ so that

$$\alpha_n - \delta = \frac{\delta}{2n + 1} \text{ and } \alpha_n \to \delta \text{ for } n \to \infty.$$

Remark 4.8. *In order to guarantee that* (4.28) *in fact defines a probability we need $\alpha_n = \frac{2\delta(n+1)}{2n+1} \geq -1$, so our coupling does not work in the case $\delta = -1$. For $-1 < \delta < 0$ we can couple the chains for n large enough so that $\alpha_n \geq -1$.*

The structure of the proof is similar to the proof of Theorem 4.1: first we compare the transition probabilities of the two Markov chains X_n and Y_n (see Lemma 4.9), so that in the proof of Theorem 4.7 we can define a Markovian coupling of the two chains, such that they stay together with maximal probability once they have met. In contrast to the previous section we only deal with discrete-time Markov chains, so that we can define our coupling, such that we look at vertices with the same birth-times in both models. More precisely, we define two random variables J_n and $\tilde{J}_n$, which are uniformly distributed on $[n]$, in a dynamic way, just as in Lemma 3.12 and put $X_n = \deg_n^-(J_n)$ and $Y_n = \deg_n^-(\tilde{J}_n)$, where $J_n = \tilde{J}_n$ for all $n \in \mathbb{N}$.

Lemma 4.9. *There exists a constant C, depending only on δ, such that*

$$\max_k |p^X_{n+1}(k|j) - p^Y_{n+1}(k|j)| \le C\,\frac{j + \mathbb{1}\{j=0\}}{n^2}, \tag{4.29}$$

where p^X_{n+1} and p^Y_{n+1} denote the transition probabilities of the chains X_n and Y_n defined in the previous paragraph.

Proof. In order to compare the transition probabilities of X_n and Y_n, we need to give their exact expressions. To shorten notation we write

$$p^X_{n+1}(k|j) := \mathbb{P}(X_{n+1} = k | X_n = j) \text{ and } p^Y_{n+1}(\ell|m) := \mathbb{P}(Y_{n+1} = \ell | Y_n = m).$$

For the chain X_n we obtain

$$\begin{aligned}
p^X_{n+1}(j+1|j) &= \frac{n}{n+1}\frac{j+1+\alpha_n}{(2+\alpha_n)n} = \frac{j+1+\alpha_n}{(2+\alpha_n)(n+1)},\\
p^X_{n+1}(j|j) &= \frac{n}{n+1}\Big(1 - \frac{j+1+\alpha_n}{n(2+\alpha_n)}\Big) \text{ for } j \neq 0,\\
p^X_{n+1}(0|0) &= \frac{n}{n+1}\Big(1 - \frac{1+\alpha_n}{n(2+\alpha_n)}\Big) + \frac{1}{n+1},\\
p^X_{n+1}(0|j) &= \frac{1}{n+1} \text{ for } j \neq 0
\end{aligned}$$

and the analogous calculations for Y_n yield

$$\begin{aligned}
p^Y_{n+1}(j+1|j) &= \frac{n}{n+1}\left(\frac{j+1+\delta}{(2+\delta)n + (1+\delta)}\right) \text{ for } j \neq 0,\\
p^Y_{n+1}(1|0) &= \frac{n}{n+1}\left(\frac{1+\delta}{(2+\delta)n + (1+\delta)}\right) + \frac{1+\delta}{(n+1)(n(2+\delta) + (1+\delta))}\\
&= \frac{1+\delta}{(2+\delta)n + (1+\delta)},
\end{aligned}$$

$$p_{n+1}^Y(j|j) = \frac{n}{n+1}\left(1 - \frac{j+1+\delta}{n(2+\delta)+(1+\delta)}\right) \text{ for } j \geq 2,$$

$$p_{n+1}^Y(1|1) = \frac{n}{n+1}\left(1 - \frac{2+\delta}{n(2+\delta)+(1+\delta)}\right) + \frac{1+\delta}{(n+1)(n(2+\delta)+(1+\delta))},$$

$$p_{n+1}^Y(0|0) = \frac{n}{n+1}\left(1 - \frac{1+\delta}{n(2+\delta)+(1+\delta)}\right) + \frac{1}{n+1}\Big(1 - \frac{1+\delta}{n(2+\delta)+(1+\delta)}\Big),$$

$$p_{n+1}^Y(0|j) = \frac{1}{n+1}\left(1 - \frac{1+\delta}{n(2+\delta)+(1+\delta)}\right) \text{ for } j \neq 0,$$

$$p_{n+1}^Y(1|j) = \frac{1+\delta}{(n+1)(n(2+\delta)+(1+\delta))} \text{ for } j \geq 2.$$

We can now compare these transition probabilities. For $j \neq 0$ we have

$$\begin{aligned} 0 < p_{n+1}^X(0|j) - p_{n+1}^Y(0|j) &= \frac{1}{n+1} - \frac{1}{n+1}\left(1 - \frac{1+\delta}{n(2+\delta)+(1+\delta)}\right) \\ &= \frac{1+\delta}{(n(2+\delta)+(1+\delta))(n+1)} \leq \frac{1}{n^2}. \end{aligned} \tag{4.30}$$

Next we deal with

$$p_{n+1}^X(j+1|j) - p_{n+1}^Y(j+1|j) = \frac{j+1+\alpha_n}{(2+\alpha_n)(n+1)} - \frac{n}{n+1}\left(\frac{j+1+\delta}{(2+\delta)n+(1+\delta)}\right).$$

Note that for $\delta \geq -1$, we have

$$\begin{aligned} (2+\alpha_n)(n+1) &= (2+\delta+\frac{\delta}{2n+1})(n+1) = (2+\delta)n + (2+\delta) + \frac{n+1}{2n+1}\delta \\ &\geq (2+\delta)n + (1+\delta), \end{aligned} \tag{4.31}$$

but

$$\begin{aligned} (2+\alpha_n)(n+1) &= \left(2+\delta+\frac{\delta}{2n+1}\right)(n+1) = (2+\delta)(n+1) + \frac{n+1}{2n+1}\delta \\ &\leq (2+\delta)(n+1) + \frac{n+1}{n}(1+\delta). \end{aligned}$$

Therefore

$$\begin{aligned} p_{n+1}^X(j+1|j) - p_{n+1}^Y(j+1|j) &\geq \frac{\alpha_n - \delta}{(2+\delta)(n+1) + \frac{n+1}{n}(1+\delta)} \\ &= \frac{n}{n+1}\frac{\frac{\delta}{2n+1}}{(2+\delta)n+(1+\delta)} \geq 0, \end{aligned}$$

for $\delta \geq 0$ and for $\delta \in (-1, 0)$

$$\begin{aligned}
&\frac{j+1+\alpha_n}{(2+\alpha_n)(n+1)} - \frac{n}{n+1}\left(\frac{j+1+\delta}{(2+\delta)n+(1+\delta)}\right) \\
&\geq (j+1)\left(\frac{1}{(2+\alpha_n)(n+1)} - \frac{n}{n+1}\frac{1}{(2+\delta)n+1+\delta}\right) \\
&\quad + \delta\left(\frac{1+\frac{1}{2n+1}}{(2+\alpha_n)(n+1)} - \frac{n}{n+1}\frac{1}{(2+\delta)n+1+\delta}\right) \\
&\geq (j+1)\left(\frac{1}{(2+\alpha_n)(n+1)} - \frac{n}{n+1}\frac{1}{(2+\delta)n+1+\delta}\right) \\
&\quad + 2\delta\left(\frac{1}{(2+\alpha_n)(n+1)} - \frac{n}{n+1}\frac{1}{(2+\delta)n+1+\delta}\right) \\
&\geq (j+1+2\delta)\left(\frac{1}{(2+\alpha_n)(n+1)} - \frac{n}{n+1}\frac{1}{(2+\delta)n+1+\delta}\right),
\end{aligned}$$

which is non-negative as we are in the case $j \geq 1$. Hence, by (4.31) we obtain

$$\begin{aligned}
\left|p^X_{n+1}(j+1|j) - p^Y_{n+1}(j+1|j)\right| &= p^X_{n+1}(j+1|j) - p^Y_{n+1}(j+1|j) \\
&= \frac{j+1+\alpha_n}{(2+\alpha_n)(n+1)} - \frac{n}{n+1}\left(\frac{j+1+\delta}{(2+\delta)n+(1+\delta)}\right) \\
&= \frac{(j+1+\alpha_n)(2+\delta+(1+\delta)/n) - (j+1+\delta)(2+\alpha_n)}{(n+1)(2+\alpha_n)(2+\delta+(1+\delta)/n)} \\
&= \frac{(j+1+\frac{1+\delta}{n})(\delta-\alpha_n) + 2(\alpha_n-\delta) + \alpha_n(1+\delta)/n}{(n+1)(2+\alpha_n)(2+\delta+(1+\delta)/n)} \\
&\leq \frac{(j-1)\frac{\delta-\alpha_n}{2+\delta} + \frac{\alpha_n(1+\delta)}{(2+\alpha_n)(2+\delta)n} + \frac{(j+1)(1+\delta)}{n(2+\delta)}}{n+1} \\
&\leq \frac{\frac{j-1}{n} + \frac{j+1}{n} + \frac{1}{n}}{n+1} \leq \frac{2j+1}{n(n+1)}. \qquad (4.32)
\end{aligned}$$

In the case $j = 0$ we obtain

$$\begin{aligned}
p^X_{n+1}(1|0) - p^Y_{n+1}(1|0) &= \frac{1+\alpha_n}{(2+\alpha_n)(n+1)} - \left(\frac{1+\delta}{(2+\delta)n+(1+\delta)}\right) \\
&= \frac{(1+\alpha_n)((2+\delta)n+(1+\delta)) - (1+\delta)(2+\alpha_n)(n+1)}{(2+\alpha_n)(n+1)(2+\delta)n+(1+\delta)} \\
&= \frac{n(\alpha_n-\delta) - (1+\delta)}{(2+\alpha_n)(n+1)(2+\delta)n+(1+\delta)}
\end{aligned}$$

$$= \frac{\frac{\delta n}{2n+1} - (1+\delta)}{(2+\alpha_n)(n+1)((2+\delta)n + (1+\delta))} < 0$$

for $\delta \geq -1$ and thus

$$|p^X_{n+1}(1|0) - p^Y_{n+1}(1|0)| = p^X_{n+1}(1|0) - p^Y_{n+1}(1|0) \leq \frac{2}{n(n+1)} \leq \frac{2}{n^2}. \tag{4.33}$$

Looking at the calculations for $p^X_{n+1}(j+1|j) - p^Y_{n+1}(j+1|j)$ in the case $j \geq 1$, we directly see that

$$p^X_{n+1}(j|j) - p^Y_{n+1}(j|j) < 0,$$

for $\delta \geq -1$ and $j \geq 2$. It follows that

$$\left|p^X_{n+1}(j|j) - p^Y_{n+1}(j|j)\right| = p^Y_{n+1}(j|j) - p^X_{n+1}(j|j) \leq 4\, \frac{j}{(n+1)^2}. \tag{4.34}$$

With the same argument in the case $j = 1$ we obtain

$$p^X_{n+1}(1|1) - p^Y_{n+1}(1|1) < 0$$

and consequently

$$\left|p^X_{n+1}(1|1) - p^Y_{n+1}(1|1)\right| = p^Y_{n+1}(1|1) - p^X_{n+1}(1|1) \leq \frac{4}{(n+1)^2}. \tag{4.35}$$

For $j = 0$ the situation is slightly different. Indeed, for both chains the probability to stay in state 0 is greater than for any other state, but for X_n the deviation is larger, since conditioned on $J_n = n$ we have $X_n = 0$ $\mathbb{P}$-a.s., but $\mathbb{P}(Y_n = 1|J_n = n) > 0$. More precisely

$$\begin{aligned} p^X_{n+1}(0|0) - p^Y_{n+1}(0|0) &= \frac{n}{n+1}\Big(1 - \frac{1+\alpha_n}{n(2+\alpha_n)}\Big) + \frac{1}{n+1} \\ &- \left(\frac{n}{n+1}\Big(1 - \frac{1+\delta}{n(2+\delta)+(1+\delta)}\Big) + \frac{1}{n+1}\Big(1 - \frac{1+\delta}{n(2+\delta)+(1+\delta)}\Big)\right) \\ &= \frac{1+\delta}{n(2+\delta)+(1+\delta)} - \frac{1+\alpha_n}{(n+1)(2+\alpha_n)} \geq 0, \end{aligned}$$

which can be seen from the calculations conducted for $p^X_{n+1}(1|0) - p^Y_{n+1}(1|0)$. Consequently

$$\left|p^X_{n+1}(0|0) - p^Y_{n+1}(0|0)\right| = p^X_{n+1}(0|0) - p^Y_{n+1}(0|0) \leq \frac{2}{n^2}. \tag{4.36}$$

Combining (4.30), (4.32), (4.33), (4.34), (4.35) and (4.36) yields the desired result. □

We can now proof Theorem 4.7 by first defining a coupling of the two chains and then using the previous result, that the chains stay together with high probability once they meet.

Proof of Theorem 4.7. We now define a coupling $(\hat{X}_n, \hat{Y}_n)$ of the two chains in a similar way as in section 4.1. Again we let the two chains evolve independently, but whenever $J_n = n$, we put $\tilde{J}_n = n$ as well. Consequently $J_n = \tilde{J}_n$ for all n. As before we define the coupling in such a way that whenever $X_n = Y_n$ the chains stay together with maximal probability. We can then bound their distance in total variation via

$$d_{TV}(X_n, Y_n) = \min_{(X_n', Y_n') \text{ coupling of } (X_n, Y_n)} \mathbb{P}(X_n' \neq Y_n') \leq \mathbb{P}(\hat{X}_n, \hat{Y}_n).$$

Just as in section 4.1 we define

$$\begin{aligned}\mathbb{P}(\hat{X}_{n+1} = \ell, \hat{Y}_{n+1} = \ell | \hat{X}_n = m, \hat{Y}_n = m) &= \hat{p}_{n+1}(\ell, \ell | m, m) \\ &:= \min(p^X_{n+1}(\ell|m), p^Y_{n+1}(\ell|m))\end{aligned}$$

and for $\ell \neq j$ we set

$$\hat{p}_{n+1}(k, j|m, m) = \frac{(p^X_{n+1}(k|m) - \hat{p}_{n+1}(k, k|m, m))(p^Y_{n+1}(j|m) - \hat{p}_{n+1}(j, j|m, m))}{1 - \sum_{i \geq 0} \hat{p}_{n+1}(i, i|m, m)}.$$

For $k \neq m$ we put

$$\hat{p}_{n+1}(k, j|m, \ell) = \frac{p^X_{n+1}(\ell|m) p^Y_{n+1}(j|k)}{1 - p^Y_{n+1}(0|k)} \text{ for } \ell, j \neq 0,$$

$$\hat{p}_{n+1}(0, 1|m, \ell) = p^X_{n+1}(0|m) - p^Y_{n+1}(0|\ell),$$

$$\hat{p}_{n+1}(0, j|m, \ell) = p^Y_{n+1}(j|\ell) \frac{p^X_{n+1}(0|m) - p^Y_{n+1}(0|\ell)}{1 - p^Y_{n+1}(0|\ell)} \text{ for } j \geq 2,$$

$$\hat{p}_{n+1}(\ell, 0|m, \ell) = 0, \text{ for } \ell \neq 0 \text{ and}$$

$$\hat{p}_{n+1}(0, 0|m, \ell) = p^Y_{n+1}(0, \ell).$$

To show that this really defines a coupling of the two chains one can easily adjust the proof of Theorem 4.1 to this setting.
Remember that we defined

$$E_{k,n} := F_k \cap G_{k,n} := \{J_k = k\} \cap \{J_\ell \leq k, k \leq \ell \leq n\}$$

and that we have

$$\mathbb{P}(\hat{X}_n \neq \hat{Y}_n) \leq \frac{1}{n} \sum_{k=2}^{n} \mathbb{P}(\hat{X}_n \neq \hat{Y}_n | E_{k,n}) + \frac{1}{n}$$

as well as

$$\begin{aligned}\mathbb{P}(\hat{X}_n \neq \hat{Y}_n | E_{k,n}) &= \mathbb{P}(\hat{X}_n \neq \hat{Y}_n | E_{k,n}, \hat{X}_k = \hat{Y}_k)\mathbb{P}(\hat{X}_k = \hat{Y}_k | E_{k,n}) \\ &\quad + \mathbb{P}(\hat{X}_n \neq \hat{Y}_n | E_{k,n}, \hat{X}_k \neq \hat{Y}_k)\mathbb{P}(\hat{X}_k \neq \hat{Y}_k | E_{k,n}) \\ &:= R_{1,k} + R_{2,k}.\end{aligned}$$

Again we will first bound the probability that the two chains do not meet even though $J_k = \tilde{J}_k = k$:

$$\begin{aligned}\mathbb{P}(\hat{X}_k \neq \hat{Y}_k | E_{k,n}) &= \mathbb{P}(\hat{Y}_k = 1 | \tilde{J}_k = k) \\ &= \frac{1+\delta}{(k-1)(2+\delta)+(1+\delta)} \leq \frac{1}{k},\end{aligned}$$

so that

$$\frac{1}{n}\sum_{k=2}^{n} R_{2,k} \leq \frac{1}{n}\sum_{k=2}^{n}\frac{1}{k} \leq C\ \frac{\log(n)}{n}. \tag{4.37}$$

Note that throughout the proof C always denotes a constant only dependent on δ, but which may vary from step to step.
From the proof of Theorem 4.1 we know that

$$\begin{aligned}&\mathbb{P}(\hat{X}_n \neq \hat{Y}_n | E_{k,n}, \hat{X}_k = \hat{Y}_k)\mathbb{P}(\hat{X}_k = \hat{Y}_k | E_{k,n}) \\ &\leq \sum_{\ell=k}^{n-1} \frac{\ell}{k}\mathbb{P}(\hat{X}_{\ell+1} \neq \hat{Y}_{\ell+1} | \hat{X}_\ell = \hat{Y}_\ell, \hat{X}_k = \hat{Y}_k = 0).\end{aligned}$$

By the construction of the coupling as well as by the fact that $\hat{X}_n$ and $\hat{Y}_n$ are Markov chains, we get

$$\begin{aligned}&\mathbb{P}(\hat{X}_{\ell+1} \neq \hat{Y}_{\ell+1} | \hat{X}_\ell = \hat{Y}_\ell, \hat{X}_k = \hat{Y}_k = 0) := \mathbb{P}_k(\hat{X}_{\ell+1} \neq \hat{Y}_{\ell+1} | \hat{X}_\ell = \hat{Y}_\ell) \\ &= \sum_{m=0}^{\ell-1} \mathbb{P}_k(\hat{X}_{\ell+1} \neq \hat{Y}_{\ell+1} | \hat{X}_\ell = \hat{Y}_\ell = m)\mathbb{P}_k(\hat{X}_\ell = m | \hat{X}_\ell = \hat{Y}_\ell) \\ &= \mathbb{P}(\hat{X}_{\ell+1} = 0, \hat{Y}_{\ell+1} \neq 0 | \hat{X}_\ell = \hat{Y}_\ell = 0)\mathbb{P}_k(\hat{X}_\ell = 0 | \hat{X}_\ell = \hat{Y}_\ell) \\ &\quad + \mathbb{P}(\hat{X}_{\ell+1} = 0, \hat{Y}_{\ell+1} \neq 0 | \hat{X}_\ell = \hat{Y}_\ell = 1)\mathbb{P}_k(\hat{X}_\ell = 1 | \hat{X}_\ell = \hat{Y}_\ell) \\ &\quad + \mathbb{P}(\hat{X}_{\ell+1} = 2, \hat{Y}_{\ell+1} \neq 2 | \hat{X}_\ell = \hat{Y}_\ell = 1)\mathbb{P}_k(\hat{X}_\ell = 1 | \hat{X}_\ell = \hat{Y}_\ell) \\ &\quad + \sum_{m=2}^{\ell-1}\Big(\mathbb{P}(\hat{X}_{\ell+1} = m+1, \hat{Y}_{\ell+1} \neq m+1 | \hat{X}_\ell = \hat{Y}_\ell = m)\mathbb{P}_k(\hat{X}_\ell = m | \hat{X}_\ell = \hat{Y}_\ell)\Big) \\ &\quad + \sum_{m=2}^{\ell-1}\Big(\mathbb{P}(\hat{X}_{\ell+1} = 0, \hat{Y}_{\ell+1} \neq 0 | \hat{X}_\ell = \hat{Y}_\ell = m)\mathbb{P}_k(\hat{X}_\ell = m | \hat{X}_\ell = \hat{Y}_\ell)\Big)\end{aligned}$$

$$
\begin{aligned}
&= (p^X_{\ell+1}(0|0) - \hat{p}_{\ell+1}(0,0|0,0))\mathbb{P}_k(\hat{X}_\ell = 0|\hat{X}_\ell = \hat{Y}_\ell) \\
&\quad + (p^X_{\ell+1}(0|1) - \hat{p}_{\ell+1}(0,0|1,1))\mathbb{P}_k(\hat{X}_\ell = 1|\hat{X}_\ell = \hat{Y}_\ell) \\
&\quad + (p^X_{\ell+1}(2|1) - \hat{p}_{\ell+1}(2,2|1,1))\mathbb{P}_k(\hat{X}_\ell = 2|\hat{X}_\ell = \hat{Y}_\ell) \\
&\quad + \sum_{m=2}^{\ell-1}(p^X_{\ell+1}(m+1|m) - \hat{p}_{\ell+1}(m+1,m+1|m,m))\mathbb{P}_k(\hat{X}_\ell = m|\hat{X}_\ell = \hat{Y}_\ell) \\
&\quad + \sum_{m=2}^{\ell-1}(p^X_{\ell+1}(0|m) - \hat{p}_{n+1}(0,0|m,m))\mathbb{P}_k(\hat{X}_\ell = m|\hat{X}_\ell = \hat{Y}_\ell) \\
&\le 2\sum_{m=0}^{\ell-1}\max_j \left|p^X_{\ell+1}(j|m) - \hat{p}_{\ell+1}(j,j|m,m)\right| \mathbb{P}_k(\hat{X}_\ell = m|\hat{X}_\ell = \hat{Y}_\ell) \\
&\le 2\sum_{m=0}^{\ell-1}\max_j \left|p^X_{\ell+1}(j|m) - p^Y_{\ell+1}(j|m)\right| \mathbb{P}_k(\hat{X}_\ell = m|\hat{X}_\ell = \hat{Y}_\ell) \\
&\le C\sum_{m=0}^{\ell-1}\frac{m + \mathbb{1}\{m=0\}}{\ell^2}\mathbb{P}_k(\hat{X}_\ell = m|\hat{X}_\ell = \hat{Y}_\ell).
\end{aligned}
$$

Hence, we obtain

$$
\begin{aligned}
\sum_{k=2}^{n} R_{1,k} &\le C \ \sum_{k=2}^{n}\frac{1}{k}\sum_{\ell=k}^{n-1}\sum_{m=0}^{\ell-1}\frac{m+\mathbb{1}\{m=0\}}{\ell}\mathbb{P}_k(\hat{X}_\ell = m|\hat{X}_\ell = \hat{Y}_\ell) \\
&\le C \ \left(\sum_{k=2}^{n}\frac{1}{k}\sum_{\ell=k}^{n-1}\frac{\mathbb{E}_k\left[\hat{X}_\ell\right]}{\ell} + \sum_{k=2}^{n}\frac{1}{k}\sum_{\ell=k}^{n-1}\frac{\mathbb{P}_k(\hat{X}_\ell = 0|\hat{X}_\ell = \hat{Y}_\ell)}{\ell}\right)
\end{aligned}
$$

Conditioning on the event that the chain was in state 0 at time k can only decrease the expectation, as the rate by which the process moves to zero is always one, independently of the state the process is in. Thus, we get

$$
\mathbb{E}_k\left[\hat{X}_{\ell+1}\right] \le \mathbb{E}\left[\hat{X}_{\ell+1}\right]
$$

For simplicity of notation we write $\deg_\ell := \deg_\ell^{b,\alpha_\ell}$. Now for all $\ell \ge 1$

$$
\begin{aligned}
\mathbb{E}\left[\hat{X}_\ell\right] = \mathbb{E}\left[\deg_\ell(U_n)\right] &= \frac{1}{\ell}\mathbb{E}\left[\sum_{[j=1]}^{\ell}\deg_\ell(j)\right] \\
&= \frac{1}{\ell}\mathbb{E}\left[\sum_{j=1}^{\ell}\deg_\ell^-(j)\right] = \frac{1}{\ell}\cdot\ell = 1,
\end{aligned}
$$

where U_n is uniformly distributed on $[n]$ and $\deg_\ell^-(j)$ denotes the outdegree of vertex j at time ℓ, which equals 1 for all j and ℓ by the definition of the model. Consequently,

$$\sum_{k=2}^{n} R_{1,k} \leq C \left(\sum_{k=2}^{n} \frac{1}{k} \sum_{\ell=k}^{n-1} \frac{\mathbb{E}_k\left[\hat{X}_\ell\right]}{\ell} + \sum_{k=2}^{n} \frac{1}{k} \sum_{\ell=k}^{n-1} \frac{\mathbb{P}_k(\hat{X}_\ell = 0 | \hat{X}_\ell = \hat{Y}_\ell)}{\ell} \right)$$

$$\leq C \sum_{k=2}^{n} \frac{1}{k} \sum_{\ell=k}^{n-1} \frac{\mathbb{E}_k\left[\hat{X}_\ell\right]}{\ell} \leq C \log(n)^2.$$

Combining this with (4.37) yields

$$\mathbb{P}(\hat{X}_n \neq \hat{Y}_n) \leq \frac{1}{n} \sum_{k=2}^{n} (R_{1,k} + R_{2,k}) + \frac{1}{n} \leq C \frac{\log(n)^2}{n},$$

as desired. □

5. A central limit theorem for the number of isolated vertices

In this chapter we consider the distribution of the number of isolated vertices in the preferential attachment model introduced in chapter 3. Here we call a vertex isolated if it has neither incoming nor outgoing edges. In fact, we show that for a certain class of attachment functions the properly rescaled number of isolated vertices fulfills a central limit theorem. More precisely, we show the following:

Theorem 5.1. *Let W_n denote the number of isolated vertices in the preferential attachment graph $\mathcal{G}_n$ with attachment function f. For the rescaled version*

$$\tilde{W}_n = \frac{W_n - \mu_n}{\sigma_n}$$

of W_n, where $\mu_n := \mathbb{E}[W_n]$ and $\sigma_n = \sqrt{\mathbb{V}[W_n]}$, we have that

$$d_W(\tilde{W}_n, Z) \leq C \begin{cases} \frac{1}{\sqrt{n}} & \text{for } \gamma < \frac{1}{2}, \\ \frac{\log(n)^2}{\sqrt{n}} & \text{for } \gamma = \frac{1}{2}, \\ n^{4\gamma - \frac{5}{2}} & \text{for } \gamma > \frac{1}{2}, \end{cases}$$

where $C > 0$ is a constant independent of n, Z denotes a random variable with $Z \sim \mathcal{N}(0,1)$ and $\gamma := \max_{k \geq 0}(f(k+1) - f(k))$.

Example 5.2. The class of attachment functions for which Theorem 5.1 provides a central limit theorem comprises, for example, functions of the form $f(k) = \gamma k^\alpha + \eta$ with $\gamma \in (0, \frac{5}{8}), \alpha \in (0,1]$ and $\eta \in (0,1]$ or $f(k) = (k+1)^\alpha$ with $\alpha \leq 0.7$.

It is not very suprising to see that the rates vary according to the maximal increase $\gamma := \max_{k\geq 0} \Delta f(k)$ of the attachment function. It is clear that for concave functions f we have that the larger γ the more likely it is for edges to emerge, so that connectivity of the network increases with γ. In fact, [DM13, Theorem 1.6] shows that the network topology changes for $\gamma^- := \lim_{n\to\infty} \frac{f(n)}{n} = \frac{1}{2}$ in the sense that for $\gamma^- \geq \frac{1}{2}$ a giant connected component emerges, that is, a connected component comprising a positive fraction of all vertices present. However, Theorem 5.1

only yields a limit result for attachment functions f such that $\gamma < \frac{5}{8}$. Since our methods of proof do not exhibit sharp bounds, we cannot conclude that the central limit theorem does not hold for $\gamma \geq \frac{5}{8}$. Nevertheless, the fact that with increasing γ isolated vertices become less frequent suggests that at some point the number of isolated vertices perhaps rather follows a Poisson than a standard normal distribution.
The main idea of the proof is to use Theorem 2.6, which gives the distance between the law of a non-negative random variable and the standardnormal distribution using a size-bias coupling. In our setting we let

$$X := W_n = \sum_{i=1}^{n} X_{i,n},$$

where $X_{i,n} = \mathbb{1}\{\text{vertex } i \text{ is isolated in } \mathcal{G}_n\}$. Now, for each fixed n, we have $\mathbb{E}[X] = \mathbb{E}[W_n] \leq n < \infty$.

Therefore we will need to construct a random variable W_n^s having size-bias distribution of W_n. We will do so by applying the general construction which we described in chapter 2 (see also [Ros11, section 3.4.1]) to our setting.

5.1. Size-bias construction

Following corollary 2.8 we construct a random variable having the size-bias distribution of W_n as follows: we choose one of the vertices in $\mathcal{G}_n$ proportional to $\frac{\mathbb{E}[X_{i,n}]}{\mathbb{E}[W_n]}$, delete all its adjacent edges and adjust the remaining summands. More formally, for I chosen proportional to $\vartheta_{i,n} := \mathbb{E}[X_{i,n}]$, independent of all else, we put $X_I^s = 1$ and $(X_j^I)_{j \neq I}$ with distribution conditional on $X_I = 1$. For $I = i$ we generate X_j^i for all $i \neq I$ by reconsidering every edge present in $\mathcal{G}_n$ and deleting it with probability

$$\mathbb{P}(\text{edge } \{k \to \ell\} \text{ is deleted} | \{k \to \ell\} \text{ is in } \mathcal{G}_n) = 1 - \frac{\mu_{k-1}^f(\ell, i)}{\mu_{k-1}^f(\ell)},$$

where $\mu_{k-1}^f(\ell) := \mathbb{E}\left[f(\deg_{k-1}^-(\ell))\right]$ and $\tilde{\mu}_{k-1}^f(\ell, i) = \mathbb{E}\left[f(\deg_{k-1}^-(\ell)) | X_{i,n} = 1\right]$. We will denote the resulting graph by $\mathcal{G}_n^{(i)}$. As connections only depend on the indegree of the older of the two vertices, we have

$$\begin{aligned}\mu_{k-1}^f(\ell, i) = \mathbb{E}\left[f(\deg_{k-1}^-(\ell)) | X_{i,n} = 1\right] &= \sum_{m=0}^{k-\ell+1} f(m) \mathbb{P}(\deg_{k-1}^-(\ell) = m | X_{i,n} = 1) \\ &= \sum_{m=0}^{k-\ell+1} f(m) \mathbb{P}(\deg_{k-1}^-(\ell) = m | i \nrightarrow \ell) \\ &= \mathbb{E}\left[f(\deg_{k-1}^-(\ell)) | i \nrightarrow \ell\right] := \tilde{\mu}_{k-1}^f(\ell, i)\end{aligned}$$

for $i > \ell$ since the isolation of vertex i only affects the out- but not the indegree of vertex ℓ. For $i < \ell$ we have

$$\begin{aligned}\mathbb{E}\left[f(\deg^-_{k-1}(\ell))|X_{i,n}=1\right] &= \sum_{m=0}^{k-\ell+1} f(m)\mathbb{P}(\deg^-_{k-1}(\ell)=m|X_{i,n}=1)\\ &= \sum_{m=0}^{k-\ell+1} f(m)\mathbb{P}(\deg^-_{k-1}(\ell)=m)\\ &= \mathbb{E}\left[f(\deg^-_{k-1}(\ell))\right]\end{aligned}$$

as connections from any $k > \ell$ to ℓ only depend on $\deg^-_{k-1}(\ell)$, which is independent of $\deg^+(\ell)$. This shows that for $\ell > i$ the isolation of vertex i does not affect connections $\{\ell \leftrightarrow k\}$ with $k > i$ so that in order for ℓ to be isolated in $\mathcal{G}_n^{(i)}$ those edges cannot be present in $\mathcal{G}_n$. More general, the isolation of vertex i does not affect edges $\{\ell \to k\}$ if both k and ℓ emerged later than time i, see also Figure 5.4 for a visualization of this effect. Figures 5.1-5.4 show how $\mathcal{G}_n^{(i)}$ is constructed from $\mathcal{G}_n$.

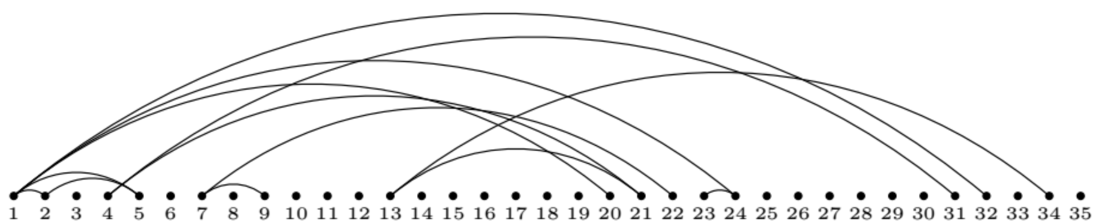


Figure 5.1: Preferential attachment graph for attachment function $f(k) = \frac{3}{10}\,k^{\frac{1}{5}} + \frac{2}{5}$

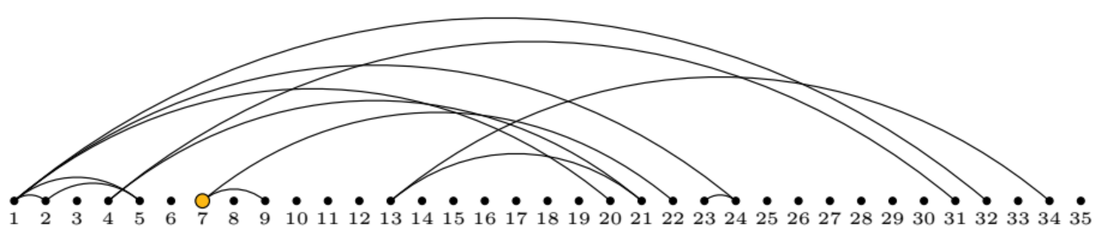


Figure 5.2: Choose one of the vertices according to $\mathbb{P}(I=i) = \frac{\vartheta_{i,n}}{\mu_n}$. Here: $i = 7$.

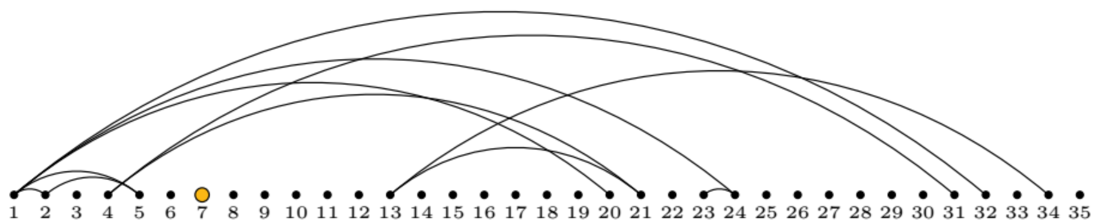


Figure 5.3: Remove all adjacent edges.

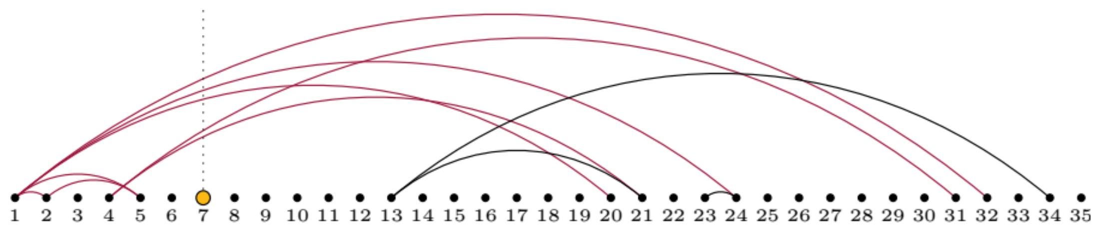


Figure 5.4: Remove all other edges with probability $1 - \frac{\mu^f_{k-1}(\ell,i)}{\mu^f_{k-1}(\ell)}$. Only the red edges are affected by the isolation of vertex 7, since for all others the probability to be deleted is 0.

To show that the resulting graph has distribution conditioned on $X_{i,n} = 1$ we first introduce some notation: For the event that there exists an edge between vertices $\ell < k$ in $\mathcal{G}_n$ we write $\{k \to \ell\}$ and accordingly $\{k \xrightarrow{i} \ell\}$ for the event that there is an edge pointing from k to ℓ in $\mathcal{G}_n^{(i)}$. Every edge $k \to \ell$ in $\mathcal{G}_n^{(i)}$ is now present with probability given by

$$\begin{aligned}\mathbb{P}(k \xrightarrow{i} \ell) &= \mathbb{P}(k \xrightarrow{i} \ell | k \to \ell)\mathbb{P}(k \to \ell)\\ &= \frac{\tilde{\mu}^f_{k-1}(\ell,i)}{\mu^f_{k-1}(\ell)}\frac{\mu^f_{k-1}(\ell)}{k} = \frac{\tilde{\mu}^f_{k-1}(\ell,i)}{k}\\ &= \mathbb{P}(k \to \ell | i \nrightarrow \ell) = \mathbb{P}(k \to \ell | X_{i,n} = 1).\end{aligned}$$

Following Proposition 2.7 (see also [Ros11, Proposition 3.21]) the number of isolated vertices $W_n^s = \sum_{i=1, i\neq I}^n X_j^I + 1$ in $\mathcal{G}_n^I$ has distribution given by the size-bias distribution of W_n. Additionally to $\tilde{\mu}^f_{k-1}(\ell,i)$ defined above, we introduce

$$\hat{\mu}^f_{k-1}(\ell,i) := \mathbb{E}\left[f(\deg^-_{k-1}(\ell))|i \to \ell\right]$$

so that

$$\mu^f_{k-1}(\ell) = \mathbb{P}(i \nrightarrow \ell)\tilde{\mu}^f_{k-1}(\ell,i) + \mathbb{P}(i \to \ell)\hat{\mu}^f_{k-1}(\ell,i).$$

Note that

$$\mathbb{P}(k \overset{i}{\not\to} \ell | k \to \ell) = 1 - \frac{\tilde{\mu}^f_{k-1}(\ell, i)}{\mu^f_{k-1}(\ell)} = \frac{\mu^f_{k-1}(\ell) - \tilde{\mu}^f_{k-1}(\ell, i)}{\mu^f_{k-1}(\ell)}$$

and

$$\begin{aligned}
\mu^f_{k-1}(\ell) - \tilde{\mu}^f_{k-1}(\ell, i) &= (1 - \mathbb{P}(i \to \ell))\tilde{\mu}^f_{k-1}(\ell, i) + \mathbb{P}(i \to \ell)\hat{\mu}^f_{k-1}(\ell, i) - \tilde{\mu}^f_{k-1}(\ell, i) \\
&= \mathbb{P}(i \to \ell)(\hat{\mu}^f_{k-1}(\ell, i) - \tilde{\mu}^f_{k-1}(\ell, i)) \\
&\leq \frac{\mu^f_{i-1}(\ell)}{i} \hat{\mu}^f_{k-1}(\ell, i) \\
&\leq f(1) \frac{i^\gamma k^\gamma}{i \ell^{2\gamma}}, \qquad (5.1)
\end{aligned}$$

for all $\ell \in \{1, \ldots, j-2\}$ by Lemma 3.21. Lemma 2.8. and Lemma 2.10 in [DM13] yield that

$$\frac{\hat{\mu}^f_{k-1}(\ell, i)}{\mu^f_{k-1}(\ell)} \leq \frac{f(1)}{f(0)},$$

which gives

$$\mathbb{P}(k \overset{i}{\not\to} \ell | k \to \ell) \leq \frac{\mu^f_{i-1}(\ell)}{i} \frac{\hat{\mu}^f_{k-1}(\ell, i)}{\mu^f_{k-1}(\ell)} \leq \frac{f(1)}{f(0)} i^{\gamma-1} \ell^{-\gamma}. \qquad (5.2)$$

5.2. Proof of Theorem 5.1

To bound the distance between the law of W_n and the standard normal distribution we have to bound both terms on the right-hand side of (2.4). Lemma 5.3 gives the order of μ_n as well as a lower bound on σ_n^2. With the help of these results we can then deduce an upper bound on both terms, see Lemma 5.8, Lemma 5.9 and Lemma 5.10.

Lemma 5.3. *Let W_n denote the number of isolated vertices in the preferential attachment graph $\mathcal{G}_n$ described before. For $\mu_n = \mathbb{E}[W_n]$ and $\sigma_n^2 = \mathbb{V}[W_n]$ we then get*

$$\mu_n \sim n \qquad (5.3)$$

and

$$\sigma_n^2 \geq Cn, \qquad (5.4)$$

where $C > 0$ is a constant independent of n.

Here, by $g \sim f$ we mean that $\lim_{n\to\infty} \frac{g(n)}{f(n)} = c \in (0,\infty)$, or equivalently there exist constants b_1, $b_2 \in \mathbb{R}$ such that

$$b_1 f(n) \leq g(n) \leq b_2 f(n).$$

Before proving this Lemma we give three auxiliary results which we will use frequently throughout this chapter. First note that for $f(0) := \eta \in (0,1]$ and by the asymptotics of the gamma function we have

$$\begin{aligned}\prod_{i=2}^{n}\left(1-\frac{f(0)}{i}\right) &= \frac{\prod_{i=2}^{n}(i-\eta)}{n!} = \frac{\prod_{i=1}^{n-1}(i+1-\eta)}{n!} \\ &= \frac{\Gamma(n+1-\eta)}{\Gamma(2-\eta)\Gamma(n+1)} \sim n^{-\eta},\end{aligned} \tag{5.5}$$

since $\lim_{n\to\infty} \frac{\Gamma(n+\alpha)}{\Gamma(n)n^{\alpha}} = 1$. Furthermore, for $\eta > -1$ we have

$$d_1 n^{1+\eta} \leq \int_1^{n-1} j^{\eta}\,\mathrm{d}j \leq \sum_{j=2}^{n} j^{\eta} \leq \int_2^{n+1} j^{\eta}\,\mathrm{d}j \leq d_2 n^{1+\eta}$$

so that

$$\sum_{j=1}^{n} j^{\eta} \sim n^{1+\eta}. \tag{5.6}$$

The last auxiliary result concerns the probability of a vertex to have outdegree zero. First note that

$$\begin{aligned}\mathbb{P}(j \to k) &= \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}\{j \to k\}|\mathcal{G}_{j-1}\right]\right] \\ &= \mathbb{E}\left[\frac{f(\deg^{-}_{j-1}(k))}{j}\right] = \frac{\mu^{f}_{j-1}(k)}{j},\end{aligned}$$

so that

$$\begin{aligned}p_{n,0} &:= \mathbb{P}(\deg^{+}(n) = 0) = \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}\{\deg^{+}(n) = 0\}|\mathcal{G}_{n-1}\right]\right] \\ &= \prod_{j=1}^{n-1}\left(1 - \frac{\mathbb{E}\left[f(\deg_{n-1}(j))\right]}{n}\right) := \prod_{j=1}^{n-1}\left(1-\frac{\mu^{f}_{n-1}(j)}{n}\right) > 0,\end{aligned}$$

for every fixed n as $\mu^{f}_{n-1}(j) \leq \mu^{f}_{n-1}(1) \leq n-1$. On account of Theorem 3.20 we also have

$$\exp(-\lambda) - a(n) \leq p_{n,0} \leq \exp(-\lambda) + a(n),$$

for some positive sequence $(a(n))$ tending to zero. Here $\lambda = \mathbb{E}_\mu[f(X)]$, where $\mathbb{E}_\mu$ denotes the expectation with respect to the measure μ defined in (3.3). Thus, as $0 < p_{n,0} < 1$ for all n, there exist constants $c_1, c_2 \in (0,1)$ such that

$$c_1 \le \min_{n\in\mathbb{N}} p_{n,0} \le \max_{n\in\mathbb{N}} p_{n,0} \le c_2. \tag{5.7}$$

With these remarks we now turn to the proof of Lemma 5.3.

Proof of Lemma 5.3. Let $\{i_{1,n}, \ldots, i_{W_{n-1},n}\}$ denote the set of vertices that are isolated in $\mathcal{G}_{n-1}$. We then have

$$\begin{aligned}
&\mathbb{E}[W_n] = \mathbb{E}[\mathbb{E}[W_n|W_{n-1}]] \\
&= \mathbb{E}\left[\sum_{j=1}^{W_{n-1}} \mathbb{1}\{i_{j,n} \text{ is isolated in } \mathcal{G}_n | i_{j,n} \text{ isolated in } \mathcal{G}_{n-1}\} + \mathbb{1}\{\deg^+(n) = 0\}\right] \\
&= \mathbb{E}\left[\sum_{j=1}^{W_{n-1}} \left(1 - \frac{f(0)}{n}\right)\right] + p_{n,0} = \left(1 - \frac{f(0)}{n}\right)\mathbb{E}[W_{n-1}] + p_{n,0},
\end{aligned}$$

and by iteration

$$\mathbb{E}[W_n] = \prod_{i=2}^{n}\left(1 - \frac{f(0)}{i}\right)\mathbb{E}[W_1] + \sum_{j=2}^{n} p_{j,0} \prod_{k=j+1}^{n}\left(1 - \frac{f(0)}{k}\right).$$

With the auxiliary results mentioned above this yields

$$\begin{aligned}
\mathbb{E}[W_n] &= \prod_{i=2}^{n}\left(1 - \frac{f(0)}{i}\right) + \sum_{j=2}^{n}\prod_{k=j+1}^{n}\left(1 - \frac{f(0)}{k}\right)p_{j,0} \\
&\sim \frac{1}{n^\eta} + \sum_{j=2}^{n}\frac{j^\eta}{n^\eta}p_{j,0} \sim \frac{1}{n^\eta} + \frac{1}{n^\eta}\sum_{j=2}^{n} j^\eta \sim n,
\end{aligned}$$

where we used (5.6) in the last step. This shows (5.3). We now turn to the lower variance bound given in (5.4). We have

$$\begin{aligned}
\mathbb{V}[W_n] &= \sum_{i=1}^{n}\mathbb{V}[X_{i,n}] + 2\sum_{i=1}^{n}\sum_{j=1}^{i-1}\operatorname{Cov}[X_{i,n}, X_{j,n}] \\
&= \sum_{i=1}^{n}\mathbb{P}(X_{i,n} = 1)(1 - \mathbb{P}(X_{i,n} = 1)) \\
&\quad + 2\sum_{i=1}^{n}\sum_{j=1}^{i-1}\mathbb{P}(X_{i,n} = 1)(\mathbb{P}(X_{j,n} = 1|X_{i,n} = 1) - \mathbb{P}(X_{j,n} = 1)).
\end{aligned}$$

Note that the outdegree is fixed from time i onwards and that connections formed afterwards only rely on the indegree of i, i.e. $\deg_n^-(i)$ and $\deg^+(i)$ are independent random variables. To shorten notation we put

$$\begin{aligned}&\mathbb{P}(\deg^+(i) = i_1, \deg_n^-(i) = i_2, \deg^+(j) = j_1, \deg_n^-(j) = j_2)\\ &:= \mathbb{P}(d_n(i) : (i_1, i_2), d_n(j) : (j_1, j_2))\end{aligned}$$

and accordingly

$$\mathbb{P}(\deg^+(i) = i_1, \deg_n^-(i) = i_2) = \mathbb{P}(d_n(i) : (i_1, i_2)).$$

Setting $a_n := \prod_{i=2}^n \left(1 - \frac{f(0)}{i}\right)$ we have

$$\begin{aligned}\mathbb{P}(\deg_n^-(i) = 0) &= \prod_{\ell=i+1}^{n} \mathbb{P}(\{\ell \nrightarrow i\} | \bigcap_{r=i+1}^{\ell-1} \{r \nrightarrow i\})\\ &= \prod_{\ell=i+1}^{n} \left(1 - \frac{f(0)}{\ell}\right) = \frac{a_n}{a_i} \sim \left(\frac{i}{n}\right)^\eta \end{aligned} \tag{5.8}$$

and thus

$$\sum_{i=1}^n \mathbb{P}(X_{i,n} = 1)(1 - \mathbb{P}(X_{i,n} = 1)) \sim \sum_{i=1}^n p_{i,0} \left(\frac{i}{n}\right)^\eta (1 - p_{i,0} \left(\frac{i}{n}\right)^\eta) \sim \sum_{i=1}^n \left(\frac{i}{n}\right)^\eta \sim n,$$

where we used (5.7) and (5.6). Furthermore we have

$$\begin{aligned}&\mathbb{P}(X_{i,n} = 1, X_{j,n} = 1) = \mathbb{P}(d_n(j) : (0,0) | d_n(i) : (0,0))\mathbb{P}(d_n(i) : (0,0))\\ &= \mathbb{P}(\deg^+(j) = 0 | \deg^+(i) = 0)\mathbb{P}(\deg_n^-(j) = 0 | \deg^+(i) = 0)\mathbb{P}(\deg^+(i) = 0)\mathbb{P}(\deg_n^-(i) = 0)\\ &\geq \mathbb{P}(\deg^+(j) = 0)\mathbb{P}(\deg_n^-(j) = 0)\mathbb{P}(\deg^+(i) = 0)\mathbb{P}(\deg_n^-(i) = 0)\\ &= \mathbb{P}(X_{i,n=1})\mathbb{P}(X_{j,n} = 1),\end{aligned} \tag{5.9}$$

which shows that $X_{i,n}$ and $X_{j,n}$ are positively correlated. Here we used that in- and outdegree of a fixed vertex are independent as we had noticed before. We will use this result as well as the fact that $\deg^+(j)$ is independent of $\deg_n^-(i)$ for $i > j$ to find a lower bound on $\mathrm{Cov}[X_{i,n}, X_{j,n}]$. More precisely

$$\begin{aligned}&\sum_{i=1}^n \sum_{j=1}^{i-1} \mathbb{P}(X_{i,n} = 1)(\mathbb{P}(X_{j,n} = 1 | X_{i,n} = 1) - \mathbb{P}(X_{j,n} = 1))\\ &= \sum_{i=1}^n \sum_{j=1}^{i-1} \mathbb{P}(X_{i,n} = 1)(\mathbb{P}(d_n(j) : (0,0)) | d_n(i) : (0,0)) - \mathbb{P}(d_n(j) : (0,0)))\end{aligned}$$

$$\geq C\sum_{i=1}^{n}\sum_{j=1}^{i-1}\left(\frac{i}{n}\right)^{\eta}\Big(\mathbb{P}(\deg^+(j)=0|\deg^+(i)=0)\mathbb{P}(\deg^-_n(j)=0|\deg^+(i)=0)$$

$$-\,\mathbb{P}(\deg^+(j)=0)\mathbb{P}(\deg^-_n(j)=0)\Big)$$

$$\geq C\sum_{i=1}^{n}\sum_{j=1}^{i-1}\left(\frac{i}{n}\right)^{\eta}\mathbb{P}(\deg^+(j)=0)\mathbb{P}(\deg^-_n(j)=0)\left(\frac{1}{\mathbb{P}(i\nrightarrow j)}-1\right)$$

$$\geq C\sum_{i=1}^{n}\sum_{j=1}^{i-1}\left(\frac{i}{n}\right)^{\eta}\left(\frac{j}{n}\right)^{\eta}\frac{\mu^f_{j-1}(i)}{j-\mu^f_{j-1}(i)}$$

$$\geq C\sum_{i=1}^{n}\sum_{j=1}^{i-1}\left(\frac{i}{n}\right)^{\eta}\left(\frac{j}{n}\right)^{\eta}\frac{f(0)}{j}\sim n$$

by (5.6) and consequently

$$\mathbb{V}\,[W_n]\geq C\,n,$$

where $C>0$ is independent of n. □

Before we start proving Theorem 5.1 we give the following auxiliary result which shows how the non-existence of an edge influences the degree evolution of the older of the two vertices forming that edge.

Proposition 5.4. *For $a_{j,\ell}:=\mathbb{P}(j\nrightarrow\ell)$ and $a^{(i)}_{j,\ell}:=\mathbb{P}(j\nrightarrow\ell|i\nrightarrow\ell)$ we have*

$$\prod_{\ell=m}^{k}a^{(i)}_{j,\ell}-\prod_{\ell=m}^{k}a_{j,\ell}\leq f(1)\frac{2^{1-\gamma}}{2^{1-\gamma}-1}\sum_{\ell=m}^{k}j^{\gamma-1}i^{\gamma-1}\ell^{-2\gamma}:=\xi^k_m(j,i)$$

for all k,m with $m<k\leq i-1$. It follows

$$\xi^k_m(j,i)\leq C\begin{cases}j^{\gamma-1}i^{\gamma-1}k^{1-2\gamma} & \text{for } \gamma<\frac{1}{2},\\ j^{\gamma-1}i^{\gamma-1}\log(k) & \text{for } \gamma=\frac{1}{2},\\ j^{\gamma-1}i^{\gamma-1}m^{-2\gamma+1} & \text{for } \gamma>\frac{1}{2}.\end{cases}\tag{5.10}$$

Furthermore, for any r with $m\leq r\leq k-1$ we have

$$\prod_{\substack{\ell=m\\ \ell\neq r}}^{i-1}a^{(i)}_{j,\ell}-\prod_{\substack{\ell=m\\ \ell\neq r}}^{i}a_{j,\ell}\leq C\,\xi^{i-1}_m(j,i)\leq C\begin{cases}j^{\gamma-1}i^{-\gamma} & \text{for } \gamma<\frac{1}{2},\\ j^{\gamma-1}i^{\gamma-1}\log(i) & \text{for } \gamma=\frac{1}{2},\\ j^{\gamma-1}i^{\gamma-1}m^{-2\gamma+1} & \text{for } \gamma>\frac{1}{2}.\end{cases}\tag{5.11}$$

Proof. We proceed via induction on the number N of factors. For any $m \leq k \leq i-1$ we have

$$\begin{aligned} a^{(i)}_{j,m} - a_{j,m} &= \mathbb{P}(j \nrightarrow m | i \nrightarrow m) - \mathbb{P}(j \nrightarrow m) \\ &= \left(1 - \frac{\tilde{\mu}^f_{j-1}(m,i)}{j}\right) - \left(1 - \frac{\mu^f_{j-1}(m)}{j}\right) \\ &= \frac{\mu^f_{j-1}(m) - \tilde{\mu}^f_{j-1}(m,i)}{j} \\ &\leq f(1) j^{\gamma-1} i^{\gamma-1} m^{-2\gamma}, \end{aligned}$$

by (5.1). This proves the base clause for $N = k-m+1 = 1$. For the induction step note that $a^{(i)}_{j,\ell}$ and $a_{j,\ell}$ are asymptotically equivalent. More specifically we have

$$\begin{aligned} \frac{a^{(i)}_{j,\ell}}{a_{j,\ell}} &= \frac{\mathbb{P}(j \nrightarrow \ell | i \nrightarrow \ell)}{\mathbb{P}(j \nrightarrow \ell)} = \frac{1 - \frac{\tilde{\mu}^f_{j-1}(\ell,i)}{j}}{1 - \frac{\mu^f_{j-1}(\ell)}{j}} = 1 + \frac{\mu^f_{j-1}(\ell) - \tilde{\mu}^f_{j-1}(\ell,i)}{j - \mu^f_{j-1}(\ell)} \\ &\leq 1 + \frac{f(1) i^{\gamma-1} \ell^{-2\gamma} j^{\gamma}}{j - j^{\gamma}} \leq 1 + f(1) \frac{2^{1-\gamma}}{2^{1-\gamma}-1} j^{\gamma-1} i^{\gamma-1} \ell^{-2\gamma} \\ &:= 1 + C_f j^{\gamma-1} i^{\gamma-1} \ell^{-2\gamma} \end{aligned}$$

since $\mu^f_{j-1}(\ell) \leq j^{\gamma}$ by Lemma 3.21. Due to the induction hypothesis we then get

$$\begin{aligned} \prod_{\ell=m}^{k} a^{(i)}_{j,\ell} - \prod_{\ell=m}^{k} a_{j,\ell} &\leq \prod_{\ell=m}^{k-1} a^{(i)}_{j,\ell} (1 + C_f j^{\gamma-1} i^{\gamma-1} k^{-2\gamma}) a_{j,k} - \prod_{\ell=m}^{k} a_{j,\ell} \\ &= a_{j,k} \left(\prod_{\ell=m}^{k-1} a^{(i)}_{j,\ell} - \prod_{\ell=m}^{k-1} a_{j,\ell} \right) + \prod_{\ell=m}^{k-1} a^{(i)}_{j,\ell} \cdot a_{j,k} \cdot C_f j^{\gamma-1} i^{\gamma-1} k^{-2\gamma} \\ &\leq C_f \sum_{\ell=m}^{k-1} j^{\gamma-1} i^{\gamma-1} \ell^{-2\gamma} + C_f j^{\gamma-1} i^{\gamma-1} k^{-2\gamma} \\ &= C_f \sum_{\ell=m}^{k} j^{\gamma-1} i^{\gamma-1} \ell^{-2\gamma}, \end{aligned}$$

which proves the claim. Consequently

$$\prod_{\ell=m}^{k} a^{(i)}_{j,\ell} - \prod_{\ell=m}^{k} a_{j,\ell} \leq C \begin{cases} j^{\gamma-1} i^{\gamma-1} k^{1-2\gamma} & \text{for } \gamma < \frac{1}{2}, \\ j^{\gamma-1} i^{\gamma-1} \log(k) & \text{for } \gamma = \frac{1}{2}, \\ j^{\gamma-1} i^{\gamma-1} m^{-2\gamma+1} & \text{for } \gamma > \frac{1}{2}. \end{cases}$$

To prove the second part, note that with the definition of $a_{j,i}$ we can write

$$\begin{aligned}\prod_{\ell=m}^{i-1} a_{j,\ell}^{(i)} - \prod_{\ell=m}^{i} a_{j,\ell} &= \prod_{\ell=1}^{i-1} a_{j,\ell}^{(i)} - \prod_{\ell=1}^{i-1} a_{j,\ell} \cdot \mathbb{P}(j \nrightarrow i) \\ &= \prod_{\ell=m}^{i-1} a_{j,\ell}^{(i)} - \prod_{\ell=m}^{i-1} a_{j,\ell} \left(1 - \frac{\mu_{j-1}^{f}(i)}{j}\right) \\ &\leq C_f \sum_{\ell=m}^{i-1} j^{\gamma-1} i^{\gamma-1} \ell^{-2\gamma} + \frac{j^{\gamma-1}}{i^{\gamma}} \\ &\leq 2C_f \sum_{\ell=m}^{i-1} j^{\gamma-1} i^{\gamma-1} \ell^{-2\gamma}, \end{aligned} \tag{5.12}$$

so that the terms only differ by a constant factor. The proof is completed by showing that the bound also holds if we omit one of the factors in each of the products. Noting that $a_{j,\ell}^{(i)} \geq a_{j,\ell}$ we get

$$\begin{aligned}\prod_{\substack{\ell=m \\ \ell\neq r}}^{i-1} a_{j,\ell}^{(i)} - \prod_{\substack{\ell=m \\ \ell\neq r}}^{i} a_{j,\ell} &= \frac{1}{a_{j,r}} \left(\prod_{\substack{\ell=m \\ \ell\neq r}}^{i-1} a_{j,\ell}^{(i)} a_{j,r} - \prod_{\ell=m}^{i} a_{j,\ell} \right) \\ &\leq \frac{1}{1 - \frac{\mu_{j-1}^{f}(r)}{j}} \left(\prod_{\ell=m}^{i-1} a_{j,\ell}^{(i)} - \prod_{\ell=m}^{i} a_{j,\ell} \right) \\ &\leq \frac{1}{1 - j^{\gamma-1}} \left(\prod_{\ell=m}^{i-1} a_{j,\ell}^{(i)} - \prod_{\ell=m}^{i} a_{j,\ell} \right) \\ &\leq \frac{1}{1 - 2^{\gamma-1}} \left(\prod_{\ell=m}^{i-1} a_{j,\ell}^{(i)} - \prod_{\ell=m}^{i} a_{j,\ell} \right), \end{aligned}$$

so that (5.12) also holds in this case, with an additional factor $(1 - 2^{\gamma-1})^{-1}$.

□

We are now ready to deduce an upper bound on the first term appearing in (2.4). Therefore, first note that

$$W_n^s - W_n = D_{n,I} + \mathbb{1}\{d_I > 0\} + R_{n,I}, \tag{5.13}$$

where $\mathcal{D}_{n,I}$ denotes the set of neighbours of vertex I with total degree one (to wit: I is their only neighbour),$D_{n,I} = |\mathcal{D}_{n,I}|$, $d_I = \deg_n^-(I) + \deg^+(I)$ denotes the total degree of vertex I and $R_{n,I}$ denotes the random variable which conditioned on the graph gives the number of vertices not in $\mathcal{D}_{n,I}$ that get isolated due to the isolation of vertex I. Remember that

$$\mathbb{P}(I = i) = \frac{\mathbb{E}\left[X_{i,n}\right]}{\mathbb{E}\left[W_n\right]} = \frac{\vartheta_{i,n}}{\mu_n}$$

and thus

$$\mathbb{V}\left[\mathbb{E}\left[W_n^s - W_n|\mathcal{G}_n\right]\right] = \mathbb{V}\left[\frac{1}{\mathbb{E}\left[W_n\right]}\sum_{i=1}^{n}\vartheta_{i,n}\mathbb{E}\left[W_{n,i}^s - W_n|\mathcal{G}_n\right]\right]$$

$$= \frac{1}{\mathbb{E}\left[W_n\right]^2}\mathbb{V}\left[\sum_{i=1}^{n}\vartheta_{i,n}\left(D_{n,i} + \mathbb{1}\{\text{degree}(i) > 0\} + R_{n,i}\right)\right]$$

$$\leq \frac{3}{\mathbb{E}\left[W_n\right]^2}\left(\mathbb{V}\left[\sum_{i=1}^{n}\vartheta_{i,n}D_{n,i}\right] + \mathbb{V}\left[\sum_{i=1}^{n}\vartheta_{i,n}\mathbb{1}\{\text{degree}(i) > 0\}\right] + \mathbb{V}\left[\sum_{i=1}^{n}\vartheta_{i,n}R_{n,i}\right]\right).$$

as $2\ \mathrm{Cov}[X,Y] \leq \mathbb{V}[X] + \mathbb{V}[Y]$. By (5.9) we know that $\mathrm{Cov}[X_{i,n}, X_{j,n}] > 0$ so that also $\mathrm{Cov}[(\mathbb{1}\{\text{degree}(i) > 0\}, \mathbb{1}\{\text{degree}(j) > 0\})] > 0$ and thus,

$$\mathbb{V}\left[\sum_{i=1}^{n}\vartheta_{i,n}\mathbb{1}\{\text{degree}(i) > 0\}\right] \leq \mathbb{V}\left[\sum_{i=1}^{n}\mathbb{1}\{\text{degree}(i) > 0\}\right] = \mathbb{V}\left[n - W_n\right]$$
$$= \mathbb{V}[W_n] = \sigma_n^2.$$

Hence

$$\mathbb{V}\left[\mathbb{E}\left[W_n^s - W_n|\mathcal{G}_n\right]\right] \leq \frac{4}{\mathbb{E}\left[W_n\right]^2}\left(\sigma_n^2 + \mathbb{V}\left[\sum_{i=1}^{n}\vartheta_{i,n}D_{n,i}\right] + \mathbb{V}\left[\sum_{i=1}^{n}\vartheta_{i,n}R_{n,i}\right]\right). \quad (5.14)$$

To prove an upper bound on this expression we proceed in three steps: Lemma 5.5 gives a bound on $\mathbb{V}\left[\sum_{i=1}^{n}\vartheta_{i,n}D_{n,i}\right]$, Lemma 5.7 and Lemma 5.8 give bounds on $\mathbb{V}[R_{n,i}]$ and $\mathrm{Cov}[R_{n,i}, R_{n,j}]$ respectively.

Lemma 5.5. *As before, let $D_{n,i}$ denote the number of neighbours of vertex i with total degree one in $\mathcal{G}_n$. We then have*

$$\mathbb{V}\left[\sum_{i=1}^{n}\vartheta_{i,n}D_{n,i}\right] \leq C\begin{cases} n & \text{for } \gamma < \frac{1}{2}, \\ \log(n)n & \text{for } \gamma = \frac{1}{2}, \\ n^{2\gamma} & \text{for } \gamma > \frac{1}{2}. \end{cases}$$

Proof. To deal with $\mathbb{V}\left[\sum_{i=1}^{n}\vartheta_{i,n}D_{n,i}\right]$ we define $Y_{n,i} := \mathbb{1}\{\text{vertex } i \text{ has degree 1 in } \mathcal{G}_n\}$. We then get

$$\mathbb{V}\left[\sum_{i=1}^{n}\vartheta_{i,n}D_{n,i}\right] = \mathbb{V}\left[\sum_{i=1}^{n}\vartheta_{i,n}^{*}Y_{n,i}\right]$$
$$\leq \sum_{i=1}^{n}\mathbb{V}\left[Y_{n,i}\right] + 2\sum_{j=1}^{n}\sum_{i=1}^{j-1}\mathrm{Cov}[Y_{n,j}, Y_{n,i}]\mathbb{1}\{\mathrm{Cov}[Y_{n,j}, Y_{n,i}] > 0\}$$

$$\leq \frac{n}{4} + 2\sum_{j=1}^{n}\sum_{i=1}^{j-1} \mathrm{Cov}[Y_{n,j}, Y_{n,i}])\mathbb{1}\{\mathrm{Cov}[Y_{n,j}, Y_{n,i}] > 0\},$$

where $\vartheta^*_{i,n} = \vartheta_{j,n}$ if j is the unique neighbour of vertex i and $\vartheta^*_{i,n} = 0$ if i does not have a unique neighbour. Now, the most involved part of the proof is to deal with the subtle dependencies between $Y_{n,i}$ and $Y_{n,j}$ in order to estimate $\mathrm{Cov}[Y_{n,i}, Y_{n,j}]$. We have

$$\begin{aligned}\mathbb{P}(Y_{n,i} = 1, Y_{n,j} = 1) = &\ \mathbb{P}(d_n(i) : (1,0), d_n(j) : (1,0)) + \mathbb{P}(d_n(i) : (1,0), d_n(j) : (0,1)) \\ &+ \mathbb{P}(d_n(i) : (0,1), d_n(j) : (0,1)) + \mathbb{P}(d_n(i) : (0,1), d_n(j) : (1,0))\end{aligned}$$

as well as

$$\mathbb{P}(Y_{n,i} = 1) = \mathbb{P}(d_n(i) : (1,0)) + \mathbb{P}(d_n(i) : (0,1)).$$

Plugging this into

$$\mathrm{Cov}[Y_{n,j}, Y_{n,i}] = \mathbb{P}(Y_{n,i} = 1, Y_{n,j} = 1) - \mathbb{P}(Y_{n,i} = 1)\mathbb{P}(Y_{n,j} = 1)$$

and rearranging yields

$$\begin{aligned}\mathrm{Cov}[Y_{n,j}, Y_{n,i}] &= \sum_{\substack{i_1+i_2=1,\\ j_1+j_2=1}} \mathbb{P}(d_n(i) : (i_1,i_2), d_n(j) : (j_1,j_2)) - \mathbb{P}(d_n(i) : (i_1,i_2))\mathbb{P}(d_n(j) : (j_1,j_2)) \\ &= \sum_{\substack{i_1+i_2=1,\\ j_1+j_2=1}} \mathbb{P}(d_n(i) : (i_1,i_2))\Big(\mathbb{P}(d_n(j) : (j_1,j_2) | d_n(i) : (i_1,i_2)) - \mathbb{P}(d_n(j) : (j_1,j_2))\Big) \\ &\leq C \sum_{\substack{i_1+i_2=1,\\ j_1+j_2=1}} \left(\frac{i}{n}\right)^{\gamma} \Big(\mathbb{P}(d_n(j) : (j_1,j_2) | d_n(i) : (i_1,i_2)) - \mathbb{P}(d_n(j) : (j_1,j_2))\Big). \qquad (5.15)\end{aligned}$$

To see that $\mathbb{P}(d_n(i) : (i_1,i_2)) \leq C\left(\frac{i}{n}\right)^{\gamma}$ note that

$$\mathbb{P}(d_n(i) : (i_1,i_2)) \leq \max\{\mathbb{P}(\deg^-_n(i) = 0), \mathbb{P}(\deg^-_n(i) = 1)\}$$

and by (5.8) we get

$$\mathbb{P}(\deg^-_n(i) = 0) = \prod_{j=i+1}^{n}\left(1 - \frac{f(0)}{j}\right) \leq C\ \left(\frac{i}{n}\right)^{\eta}. \qquad (5.16)$$

In order to bound $\mathbb{P}(\deg^-_n(i) = 1)$ we define the event

$$\begin{aligned}D_i^{(k)} &= \{\text{up to time } n \text{ the only incoming edge of } i \text{ is provided by vertex } k\} \\ &= \bigcap_{\substack{\ell=i+1\\ \ell\neq k}}^{n} \{k \nrightarrow \ell\} \cap \{k \to i\},\end{aligned}$$

so that

$$\mathbb{P}(\deg_n^-(i)=1) = \sum_{k=i+1}^{n} \mathbb{P}(D_i^{(k)}) = \sum_{k=i+1}^{n} \prod_{\ell=i+1}^{k-1} \left(1-\frac{f(0)}{\ell}\right) \frac{f(0)}{k} \prod_{\ell=k+1}^{n} \left(1-\frac{f(1)}{\ell}\right)$$
$$\leq \; C \sum_{k=i+1}^{n} \left(\frac{i+1}{k-1}\right)^{\eta} \frac{1}{k} \left(\frac{k+1}{n}\right)^{f(1)}$$
$$\leq C \; \frac{i^{\eta}}{n^{f(1)}} \sum_{k=i+1}^{n} k^{f(1)-\eta-1}$$
$$\leq C \; \left(\frac{i}{n}\right)^{\eta}. \tag{5.17}$$

as $f(1) \geq f(0) = \eta$. Hence

$$\mathbb{P}(d_n(i) : (i_1, i_2)) \leq C \left(\frac{i}{n}\right)^{\eta}$$

for $i_1, i_2 \in \{0, 1\}$. We will now bound each of the differences appearing in (5.15) individually, as the dependencies crucially rely on the exact formation of in- and outgoing edges of the two vertices.

Before we start, remember that by the definition of the model not only $\deg_n^-(j)$ and $\deg^+(j)$ are independent for every fixed vertex j and any n, but also $\deg_n^-(j)$ is independent of $\deg_n^-(i)$ and $\deg^+(i)$ for every older vertex $i < j$.
For $(i_1, i_2) = (j_1, j_2) = (1, 0)$, we obtain

$$\begin{aligned} &\mathbb{P}(d_n(j) : (1,0) | d_n(i) : (1,0)) - \mathbb{P}(d_n(j) : (1,0)) \\ &\qquad = \mathbb{P}(\deg_n^-(j) = 0) \left(\mathbb{P}(\deg^+(j) = 1 | d_n(i) : (1,0)) - \mathbb{P}(\deg^+(j) = 1)\right) \\ &\qquad \leq C \; \left(\frac{j}{n}\right)^{\eta} \left(\mathbb{P}(\deg^+(j) = 1 | d_n(i) : (1,0)) - \mathbb{P}(\deg^+(j) = 1)\right). \end{aligned} \tag{5.18}$$

according to (5.16). To bound the remaining difference let $C_j^{(k)}$ denote the event that vertex j only connects to vertex k when inserted into the network. More precisely,

$$C_j^{(k)} = \Big\{ \bigcap_{\substack{\ell=1,\\ \ell\neq k}}^{j-1} \{j \nrightarrow \ell\}, \{j \to k\} \Big\}.$$

With this definition we obtain

$$\begin{aligned} \mathbb{P}(\deg^+(j) = 1 | d_n(i) : (1,0)) &= \mathbb{P}(\deg^+(j) = 1 | \deg^+(i) = 1, \{j \nrightarrow i\}) \\ &= \sum_{\substack{k=1,\\ k\neq i}}^{j-1} \mathbb{P}(C_j^{(k)} | \deg^+(i) = 1, \{j \nrightarrow i\}) \end{aligned}$$

$$= \sum_{\substack{k=1,\\ k\neq i}}^{j-1} \sum_{r=1}^{i-1} \mathbb{P}(C_j^{(k)} | C_i^{(r)}, \{j \nrightarrow i\}) \mathbb{P}(C_i^{(r)} | \deg^+(i) = 1, \{j \nrightarrow i\})$$

$$= \sum_{\substack{k=1,\\ k\neq i}}^{j-1} \sum_{r=1}^{i-1} \frac{1}{\mathbb{P}(\{j \nrightarrow i\})} \mathbb{P}(C_j^{(k)} | C_i^{(r)}) \mathbb{P}(C_i^{(r)} | \deg^+(i) = 1).$$

Again, the conditional probability $\mathbb{P}(C_j^{(k)}|C_i^{(r)})$ depends crucially on the configuration of the graph. For $k \neq r$ and $k \leq i-1$ we have

$$\mathbb{P}(C_j^{(k)}|C_i^{(r)}) = \mathbb{P}\Big(\bigcap_{\substack{\ell=1\\ \ell\neq k}}^{j-1} \{j \nrightarrow \ell\}, \{j \to k\} \Big| \bigcap_{\substack{m=1\\ m\neq r}}^{i-1} \{i \nrightarrow m\}, \{i \to r\} \Big)$$

$$= \prod_{\substack{\ell=1\\ \ell\neq k,r}}^{i-1} \mathbb{P}(j \nrightarrow \ell | i \nrightarrow \ell)\ \mathbb{P}(j \nrightarrow r | i \to r) \mathbb{P}(j \to k | i \nrightarrow k) \prod_{\ell=i}^{j-1} \mathbb{P}(j \nrightarrow \ell)$$

$$\leq \prod_{\substack{\ell=1\\ \ell\neq k}}^{i-1} \mathbb{P}(j \nrightarrow \ell | i \nrightarrow \ell) \mathbb{P}(j \to k) \prod_{\ell=i}^{j-1} \mathbb{P}(j \nrightarrow \ell).$$

If $k \geq i+1$ (so in particular $k \neq r$) we obtain

$$\mathbb{P}(C_j^{(k)}|C_i^{(r)}) = \prod_{\substack{\ell=1\\ \ell\neq r}}^{i-1} \mathbb{P}(j \nrightarrow \ell | i \nrightarrow \ell) \mathbb{P}(j \nrightarrow r | i \to r) \mathbb{P}(j \to k) \prod_{\substack{\ell=i\\ \ell\neq k}}^{j-1} \mathbb{P}(j \nrightarrow \ell)$$

$$\leq \prod_{\ell=1}^{i-1} \mathbb{P}(j \nrightarrow \ell | i \nrightarrow \ell) \mathbb{P}(j \to k) \prod_{\substack{\ell=i\\ \ell\neq k}}^{j-1} \mathbb{P}(j \nrightarrow \ell).$$

In the last case, namely $k = r$ and still $k \leq i-1$, we get

$$\mathbb{P}(C_j^{(k)}|C_i^{(k)}) = \prod_{\substack{\ell=1\\ \ell\neq k}}^{i-1} \mathbb{P}(j \nrightarrow \ell | i \nrightarrow \ell) \mathbb{P}(j \to k | i \to k) \prod_{\ell=i}^{j-1} \mathbb{P}(j \nrightarrow \ell)$$

Furthermore, we can use the definition of $C_j^{(k)}$ to write

$$\mathbb{P}(\deg^+(j) = 1) = \sum_{k=1}^{j-1} \mathbb{P}(C_j^{(k)}) = \sum_{k=1}^{j-1} \prod_{\substack{\ell=1\\ \ell\neq k}}^{j-1} \mathbb{P}(j \nrightarrow \ell) \mathbb{P}(j \to k).$$

Putting all these results into (5.18) yields

$$
\begin{aligned}
&\mathbb{P}(\deg^+(j) = 1 | d_n(i) : (1,0)) - \mathbb{P}(\deg^+(j) = 1) \\
&\leq \sum_{k=1}^{i-1} \Big(\mathbb{P}(j \to k) \prod_{\substack{\ell=1 \\ \ell \neq k}}^{i-1} \mathbb{P}(j \nrightarrow \ell | i \nrightarrow \ell) \prod_{\ell=i+1}^{j-1} \mathbb{P}(j \nrightarrow \ell) \Big) \sum_{\substack{r=1 \\ r \neq k}}^{i-1} \mathbb{P}(C_i^{(r)} | \deg^+(i) = 1) \\
&\quad + \sum_{k=1}^{i-1} \Big(\mathbb{P}(j \to k | i \to k) \prod_{\substack{\ell=1 \\ \ell \neq k}}^{i-1} \mathbb{P}(j \nrightarrow \ell | i \nrightarrow \ell) \prod_{\ell=i+1}^{j-1} \mathbb{P}(j \nrightarrow \ell) \Big) \mathbb{P}(C_i^{(k)} | \deg^+(i) = 1) \\
&\quad + \sum_{k=i+1}^{j-1} \Big(\mathbb{P}(j \to k) \prod_{\ell=1}^{i-1} \mathbb{P}(j \nrightarrow \ell | i \nrightarrow \ell) \prod_{\substack{\ell=i+1 \\ \ell \neq k}}^{j-1} \mathbb{P}(j \nrightarrow \ell) \Big) \sum_{r=1}^{i-1} \mathbb{P}(C_i^{(r)} | \deg^+(i) = 1) \\
&\quad - \sum_{k=1}^{j-1} \prod_{\substack{\ell=1 \\ \ell \neq k}}^{j-1} \mathbb{P}(j \nrightarrow \ell) \mathbb{P}(j \to k) \\
&= \sum_{k=1}^{i-1} \Big(\mathbb{P}(j \to k) \prod_{\substack{\ell=1 \\ \ell \neq k}}^{i-1} \mathbb{P}(j \nrightarrow \ell | i \nrightarrow \ell) \prod_{\ell=i+1}^{j-1} \mathbb{P}(j \nrightarrow \ell) \Big) \Big(1 - \mathbb{P}(C_i^{(k)} | \deg^+(i) = 1) \Big) \\
&\quad + \sum_{k=1}^{i-1} \Big(\mathbb{P}(j \to k | i \to k) \prod_{\substack{\ell=1 \\ \ell \neq k}}^{i-1} \mathbb{P}(j \nrightarrow \ell | i \nrightarrow \ell) \prod_{\ell=i+1}^{j-1} \mathbb{P}(j \nrightarrow \ell) \Big) \mathbb{P}(C_i^{(k)} | \deg^+(i) = 1) \\
&\quad + \sum_{k=i+1}^{j-1} \Big(\mathbb{P}(j \to k) \prod_{\ell=1}^{i-1} \mathbb{P}(j \nrightarrow \ell | i \nrightarrow \ell) \prod_{\substack{\ell=i+1 \\ \ell \neq k}}^{j-1} \mathbb{P}(j \nrightarrow \ell) \Big) \\
&\quad - \Bigg(\sum_{k=1}^{i-1} \prod_{\substack{\ell=1 \\ \ell \neq k}}^{j-1} \mathbb{P}(j \nrightarrow \ell) \mathbb{P}(j \to k) + \sum_{k=i}^{j-1} \prod_{\substack{\ell=1 \\ \ell \neq k}}^{j-1} \mathbb{P}(j \nrightarrow \ell) \mathbb{P}(j \to k) \Bigg) \\
&\leq \sum_{k=1}^{i-1} \mathbb{P}(j \to k) \prod_{\ell=i+1}^{j-1} a_{j,\ell} \Big(\prod_{\substack{\ell=1 \\ \ell \neq k}}^{i-1} \mathbb{P}(j \nrightarrow \ell | i \nrightarrow \ell) - \prod_{\substack{\ell=1 \\ \ell \neq k}}^{i} \mathbb{P}(j \nrightarrow \ell) \Big) \\
&\quad + \sum_{k=1}^{i-1} \mathbb{P}(C_i^{(k)} | \deg^+(i) = 1) \prod_{\substack{\ell=1 \\ \ell \neq k}}^{i-1} a_{j,\ell}^{(i)} \prod_{\ell=i+1}^{j-1} a_{j,\ell} \Big(\mathbb{P}(j \to k | i \to k) - \mathbb{P}(j \to k \Big) \\
&\quad + \sum_{k=i+1}^{j-1} \mathbb{P}(j \to k) \prod_{\substack{\ell=i+1 \\ \ell \neq k}}^{j-1} a_{j,\ell} \Big(\prod_{\ell=1}^{i-1} \mathbb{P}(j \nrightarrow \ell | i \nrightarrow \ell) - \prod_{\ell=1}^{i} \mathbb{P}(j \nrightarrow \ell) \Big),
\end{aligned}
$$

where the inequality stems from the fact that we omitted the term for $k = i$ in the subtrahend. The first and third term in the expression above can be bounded using

Proposition 5.4. For the second term we use the fact that

$$\mathbb{P}(j \to k | i \to k) - \mathbb{P}(j \to k) \leq \frac{\hat{\mu}^f_{j-1}(k,i) - \mu^f_{j-1}(k)}{j} \leq j^{\gamma-1}k^{-\gamma}$$

to obtain

$$\begin{aligned}
&\sum_{k=1}^{i-1} \mathbb{P}(j \to k) \prod_{\ell=i+1}^{j-1} a_{j,\ell} \Big(\prod_{\substack{\ell=1\\ \ell\neq k}}^{i-1} a^{(i)}_{j,\ell} - \prod_{\substack{\ell=1\\ \ell\neq k}}^{i} a_{j,\ell} \Big) \\
&\quad + \sum_{k=1}^{i-1} \mathbb{P}(C_i^{(k)} | \deg^+(i) = 1) \prod_{\substack{\ell=1\\ \ell\neq k}}^{i-1} a^{(i)}_{j,\ell} \prod_{\ell=i+1}^{j-1} a_{j,\ell} \Big(\mathbb{P}(j \to k | i \to k) - \mathbb{P}(j \to k) \Big) \\
&\quad + \sum_{k=i+1}^{j-1} \mathbb{P}(j \to k) \prod_{\substack{\ell=i+1\\ \ell\neq k}}^{j-1} a_{j,\ell} \Big(\prod_{\ell=1}^{i-1} a^{(i)}_{j,\ell} - \prod_{\ell=1}^{i} a_{j,\ell} \Big) \\
&\leq C \left(\sum_{\substack{k=1\\ k\neq i}}^{j-1} \frac{j^{\gamma-1}}{k^\gamma} \xi_1^{i-1}(j,i) + \sum_{k=1}^{i-1} \mathbb{P}(C_i^{(k)} | \deg^+(i) = 1) \frac{j^{\gamma-1}}{k^\gamma} \right) \\
&\leq C \left(\sum_{\substack{k=1\\ k\neq i}}^{j-1} \frac{j^{\gamma-1}}{k^\gamma} \xi_1^{i-1}(j,i) + \sum_{k=1}^{i-1} \frac{\mathbb{P}(i \to k)}{\mathbb{P}(\deg^+(i) = 1)} \frac{j^{\gamma-1}}{k^\gamma} \right) \\
&\leq C \left(\sum_{\substack{k=1\\ k\neq i}}^{j-1} \frac{j^{\gamma-1}}{k^\gamma} \xi_1^{i-1}(j,i) + \sum_{k=1}^{i-1} \frac{1}{p_{1,i}} \frac{i^{\gamma-1}}{k^\gamma} \frac{j^{\gamma-1}}{k^\gamma} \right) \\
&\leq C \left(\sum_{\substack{k=1\\ k\neq i}}^{j-1} \frac{j^{\gamma-1}}{k^\gamma} \xi_1^{i-1}(j,i) + \sum_{k=1}^{i-1} \frac{i^{\gamma-1} j^{\gamma-1}}{k^{2\gamma}} \right),
\end{aligned}$$

where the last inequality uses the fact that due to Theorem 3.20 there exist constants $C_1, C_2 \in (0,1)$ such that $C_1 \leq p_{1,i} \leq C_2$ for all $i \in \mathbb{N}$ (cf. (5.7)). Looking at the different regimes for $\xi_1^{i-1}(j,i)$ we obtain

$$\begin{aligned}
\mathbb{P}(\deg^+(j) = 1 | d_n(i) : (1,0)) - \mathbb{P}(\deg^+(j) = 1) &\leq C \left(\sum_{\substack{k=1\\ k\neq i}}^{j-1} \frac{j^{\gamma-1}}{k^\gamma} \xi_1^{i-1}(j,i) + \sum_{k=1}^{i-1} \frac{i^{\gamma-1} j^{\gamma-1}}{k^{2\gamma}} \right) \\
&\leq C \begin{cases} j^{\gamma-1} i^{-\gamma} & \text{for } \gamma < \frac{1}{2}, \\ j^{-\frac{1}{2}} i^{-\frac{1}{2}} \log(i) & \text{for } \gamma = \frac{1}{2}, \\ j^{\gamma-1} i^{\gamma-1} & \text{for } \gamma > \frac{1}{2}, \end{cases} \qquad (5.19)
\end{aligned}$$

where we used (5.6) to deduce the orders of the sums. For the second case, i.e. $(i_1, i_2) = (1, 0)$ and $(j_1, j_2) = (0, 1)$, we need to consider

$$\begin{aligned}
\mathbb{P}(d_n(j) :&(0,1)|d_n(i) : (0,1)) - \mathbb{P}(d_n(j) : (0,1)) \\
&= \mathbb{P}(\deg_n^-(j) = 1) \left(\mathbb{P}(\deg^+(j) = 0|d_n(i) : (0,1)) - \mathbb{P}(\deg^+(j) = 0)\right).
\end{aligned}$$

We can proceed in a similar way as in the previous case to obtain

$$\begin{aligned}
&\mathbb{P}(\deg^+(j) = 0|d_n(i) : (0,1)) = \mathbb{P}(\deg^+(j) = 0|\{j \nrightarrow i\}, \deg^+(i) = 1) \\
&\quad = \sum_{r=1}^{i-1} \mathbb{P}(\deg^+(j) = 0|C_i^{(r)}, \{j \nrightarrow i\})\ \mathbb{P}(C_i^{(r)}|\deg^+(i) = 1) \\
&\quad = \sum_{r=1}^{i-1} \mathbb{P}\Big(\bigcap_{\ell=1}^{j-1}\{j \nrightarrow \ell\}\Big|\{i \to r\}, \bigcap_{\substack{m=1\\ m\neq r}}^{i-1}\{i \nrightarrow m\}, \{j \nrightarrow i\}\Big)\mathbb{P}(C_i^{(r)}|\deg^+(i) = 1) \\
&\quad = \sum_{r=1}^{i-1} \frac{1}{\mathbb{P}(j \nrightarrow i)} \prod_{\substack{\ell=1\\ \ell\neq r}}^{i-1} \mathbb{P}(j \nrightarrow \ell|i \nrightarrow \ell)\mathbb{P}(j \nrightarrow r|i \to r) \prod_{\ell=i}^{j-1} \mathbb{P}(j \nrightarrow \ell)\mathbb{P}(C_i^{(r)}|\deg^+(i) = 1) \\
&\quad \leq \prod_{\ell=1}^{i-1} \mathbb{P}(j \nrightarrow \ell|i \nrightarrow \ell) \prod_{\ell=i+1}^{j-1} \mathbb{P}(j \nrightarrow \ell),
\end{aligned}$$

where we used that $\mathbb{P}(j \nrightarrow i|\{i \to r\}, \bigcap_{\substack{m=1\\ m\neq r}}^{i-1}\{i \nrightarrow m\}) = \mathbb{P}(j \nrightarrow i)$ due to the independence of in- and outdegree of i. Since $\mathbb{P}(\deg^+(j) = 0) = \prod_{\ell=1}^{j-1} \mathbb{P}(j \nrightarrow \ell)$ we get

$$\begin{aligned}
\mathbb{P}(\deg^+(j) = 0|d_n(i) : (0,1)) - \mathbb{P}(\deg^+(j) = 0)& \\
\leq \prod_{\ell=i+1}^{j-1} \mathbb{P}(j \nrightarrow \ell) &\left(\prod_{\ell=1}^{i-1} \mathbb{P}(j \nrightarrow \ell|i \nrightarrow \ell) - \prod_{\ell=1}^{i} \mathbb{P}(j \nrightarrow \ell)\right) \\
\leq C &\begin{cases} j^{\gamma-1} i^{-\gamma} & \text{for } \gamma < \frac{1}{2}, \\ j^{-\frac{1}{2}} i^{-\frac{1}{2}} \log(i) & \text{for } \gamma = \frac{1}{2}, \\ j^{\gamma-1} i^{\gamma-1} & \text{for } \gamma > \frac{1}{2}. \end{cases}
\end{aligned}$$

on account of (5.11). In the case $(i_1, i_2) = (0, 1)$ and $(j_1, j_2) = (1, 0)$ we need to bound

$$\mathbb{P}(d_n(j) : (1,0)|d_n(i) : (0,1)) - \mathbb{P}(d_n(j) : (1,0)).$$

Due to the fact that the indegree of vertex j is independent of its outdegree as well as of the in- and outdegree of vertex i (for $i < j$), the expression can be rewritten as

$$\begin{aligned}&\mathbb{P}(d_n(j):(1,0)|d_n(i):(0,1))-\mathbb{P}(d_n(j):(1,0))\\&=\mathbb{P}(\deg_n^-(j)=0)\left(\mathbb{P}(\deg^+(j)=1|\deg^+(i)=0,\deg_n^-(i)=1)-\mathbb{P}(\deg^+(j)=1)\right).\end{aligned}$$

The conditional probability can be handled in much the same way as in the first case. We first condition on the single outgoing edge of j and get

$$\begin{aligned}\mathbb{P}(\deg^+(j)=1|\deg^+(i)=0,\deg_n^-(i)=1)&=\sum_{k=1}^{j-1}\mathbb{P}(C_j^{(k)}|\deg^+(i)=0,\deg_n^-(i)=1)\\&=\sum_{k=1}^{j-1}\mathbb{P}\Big(\bigcap_{\substack{\ell=1\\ \ell\neq k}}^{j-1}\{j\nrightarrow\ell\},\{j\to k\}|\bigcap_{\ell=1}^{i-1}\{i\nrightarrow\ell\},\deg_n^-(i)=1\Big)\\&=\sum_{k=1}^{j-1}\prod_{\substack{\ell=1\\ \ell\neq k}}^{i-1}\mathbb{P}(j\nrightarrow\ell|i\nrightarrow\ell)\mathbb{P}(j\to k|i\nrightarrow k)\mathbb{P}(j\nrightarrow i|\deg_n^-(i)=1)\prod_{\ell=i+1}^{j-1}\mathbb{P}(j\nrightarrow\ell)\\&\quad+\prod_{\ell=1}^{i-1}\mathbb{P}(j\nrightarrow\ell|i\nrightarrow\ell)\mathbb{P}(j\to i|\deg_n^-(i)=1)\prod_{\ell=i+1}^{j-1}\mathbb{P}(j\nrightarrow\ell)\\&\quad+\sum_{k=i+1}^{j-1}\prod_{\ell=1}^{i-1}\mathbb{P}(j\nrightarrow\ell|i\nrightarrow\ell)\mathbb{P}(j\nrightarrow i|\deg_n^-(i)=1)\mathbb{P}(j\to k)\prod_{\substack{\ell=i+1\\ \ell\neq k}}^{j-1}\mathbb{P}(j\nrightarrow\ell).\end{aligned}$$

We can now compare these terms with the expression of $\mathbb{P}(\deg^+(j)=1)$ given in (5.17) and since $\mathbb{P}(j\to k|i\nrightarrow k)\leq\mathbb{P}(j\to k)$ we get

$$\begin{aligned}&\mathbb{P}(\deg^+(j)=1|\deg^+(i)=0,\deg_n^-(i)=1)-\mathbb{P}(\deg^+(j)=1)\\&\leq\sum_{k=1}^{i-1}\mathbb{P}(j\to k)\prod_{\ell=i+1}^{j-1}\mathbb{P}(j\nrightarrow\ell)\Big(\prod_{\substack{\ell=1\\ \ell\neq k}}^{i-1}a_{j,\ell}^{(i)}-\prod_{\substack{\ell=1\\ \ell\neq k}}^{i}a_{j,\ell}\Big)\\&\quad+\prod_{\ell=i+1}^{j-1}\mathbb{P}(j\nrightarrow\ell)\Big(\prod_{\ell=1}^{i-1}a_{j,\ell}^{(i)}\,\mathbb{P}(j\to i|\deg_n^-(i)=1)-\prod_{\ell=1}^{i-1}a_{j,\ell}\,\mathbb{P}(j\to i)\Big)\\&\quad+\sum_{k=i+1}^{j-1}\mathbb{P}(j\to k)\prod_{\substack{\ell=i+1\\ \ell\neq k}}^{j-1}\mathbb{P}(j\nrightarrow\ell)\Big(\prod_{\ell=1}^{i-1}a_{j,\ell}^{(i)}-\prod_{\ell=1}^{i}a_{j,\ell}\Big).\end{aligned}$$

The first and third term already appeared in the first case, thus we can bound them by (5.19). To estimate the second term note that

$$\mathbb{P}(j\to i|\deg_n^-(i)=1)=\sum_{k=i+1}^{n}\mathbb{P}(j\to i|D_i^{(k)})\mathbb{P}(D_i^{(k)})=\mathbb{P}(j\to i|D_i^{(j)})\mathbb{P}(D_i^{(j)})\leq\frac{f(0)}{j}$$

and thus

$$\prod_{\ell=1}^{i-1} a_{j,\ell}^{(i)}\, \mathbb{P}(j \to i|\deg_n^-(i) = 1) - \prod_{\ell=1}^{i-1} a_{j,\ell}\mathbb{P}(j \to i) \le \prod_{\ell=1}^{i-1} a_{j,\ell}^{(i)} \frac{f(0)}{j} - \prod_{\ell=1}^{i-1} a_{j,\ell} \frac{\mu_{j-1}^f(i)}{j}$$
$$\le \frac{f(0)}{j}\Big(\prod_{\ell=1}^{i-1} a_{j,\ell}^{(i)} - \prod_{\ell=1}^{i-1} a_{j,\ell}\Big)$$

which can be bounded using Proposition 5.4. Hence we obtain

$$\mathbb{P}(d_n(j) :(1,0)|d_n(i) : (0,1)) - \mathbb{P}(d_n(j) : (1,0))$$
$$\le C \cdot \mathbb{P}(\deg_n^-(j) = 0) \begin{cases} j^{\gamma-1} i^{-\gamma} & \text{for } \gamma < \frac{1}{2}, \\ j^{-\frac{1}{2}} i^{-\frac{1}{2}} \log(i) & \text{for } \gamma = \frac{1}{2}, \\ j^{\gamma-1} i^{\gamma-1} & \text{for } \gamma > \frac{1}{2}. \end{cases}$$

We continue in this fashion to obtain the following result for $(i_1, i_2) = (j_1, j_2) = (0,1)$:

$$\mathbb{P}(\deg^+(j) = 0|d_n(i) : (0,1)) = \mathbb{P}\Big(\bigcap_{\ell=1}^{j-1}\{j \nrightarrow \ell\}\Big| \bigcap_{m=1}^{i-1}\{i \nrightarrow \ell\}, \deg_n^-(i) = 1\Big)$$
$$= \prod_{\ell=1}^{i-1} \mathbb{P}(j \nrightarrow \ell | i \nrightarrow \ell)\mathbb{P}(j \nrightarrow i|\deg_n^-(i) = 1) \prod_{\ell=i+1}^{j-1} \mathbb{P}(j \nrightarrow \ell)$$
$$\le \prod_{\ell=1}^{i-1} \mathbb{P}(j \nrightarrow \ell | i \nrightarrow \ell) \prod_{\ell=i+1}^{j-1} \mathbb{P}(j \nrightarrow \ell)$$

so that

$$\mathbb{P}(\deg^+(j) = 0|d_n(i) : (0,1)) - \mathbb{P}(\deg^+(j) = 0) \le \prod_{\ell=i+1}^{j-1} \mathbb{P}(j \nrightarrow \ell) \left(\prod_{\ell=1}^{i-1} a_{j,\ell}^{(i)} - \prod_{\ell=1}^{i} a_{j,\ell}\right)$$
$$\le \begin{cases} j^{\gamma-1} i^{-\gamma} & \text{for } \gamma < \frac{1}{2}, \\ j^{\gamma-1} i^{-\frac{1}{2}} \log(i) & \text{for } \gamma = \frac{1}{2}, \\ j^{\gamma-1} i^{\gamma-1} & \text{for } \gamma > \frac{1}{2} \end{cases}$$

according to Proposition 5.4. Taking all four cases together we get

$$\mathbb{P}(\deg^+(j) = j_1|d_n(i) : (i_1, i_2)) - \mathbb{P}(\deg^+(j) = j_1) \le \begin{cases} j^{\gamma-1} i^{-\gamma} & \text{for } \gamma < \frac{1}{2}, \\ j^{\gamma-1} i^{-\frac{1}{2}} \log(i) & \text{for } \gamma = \frac{1}{2}, \\ j^{\gamma-1} i^{\gamma-1} & \text{for } \gamma > \frac{1}{2}, \end{cases}$$

for all $(i_1, i_2), (j_1, j_2) \in \{0,1\}^2$ such that $i_1 + i_2 = j_1 + j_2 = 1$. Combing this result with (5.8) and (5.17) gives

$$\begin{aligned}
&\mathbb{P}(d_n(i) : (i_1, i_2), d_n(j) : (j_1, j_2)) - \mathbb{P}(d_n(i) : (i_1, i_2))\mathbb{P}(d_n(j) : (j_1, j_2)) \\
&\quad = \mathbb{P}(d_n(i) : (i_1, i_2)) \left(\mathbb{P}(d_n(j) : (j_1, j_2) | d_n(i) : (i_1, i_2)) - \mathbb{P}(d_n(j) : (j_1, j_2))\right) \\
&\quad \leq C\ \mathbb{P}(\deg_n^-(i) = i_1)\mathbb{P}(\deg_n^-(j) = j_1) \begin{cases} j^{\gamma-1} i^{-\gamma} & \text{for } \gamma < \frac{1}{2}, \\ j^{-\frac{1}{2}} i^{-\frac{1}{2}} \log(i) & \text{for } \gamma = \frac{1}{2}, \\ j^{\gamma-1} i^{\gamma-1} & \text{for } \gamma > \frac{1}{2}, \end{cases} \\
&\quad \leq \frac{C}{n^{2\eta}} \begin{cases} j^{\gamma+\eta-1} i^{-\gamma+\eta} & \text{for } \gamma < \frac{1}{2}, \\ j^{-\frac{1}{2}+\eta} i^{-\frac{1}{2}+\eta} \log(i) & \text{for } \gamma = \frac{1}{2}, \\ j^{\gamma+\eta-1} i^{\gamma+\eta-1} & \text{for } \gamma > \frac{1}{2} \end{cases}
\end{aligned}$$

and by (5.6)

$$\sum_{j=1}^{n} \sum_{i=1}^{j-1} \mathbb{1}\{\mathrm{Cov}[Y_{n,j}, Y_{n,i}] > 0\} \mathrm{Cov}[Y_{n,j}, Y_{n,i}] \leq C \begin{cases} n & \text{for } \gamma < \frac{1}{2}, \\ \log(n) n & \text{for } \gamma = \frac{1}{2}, \\ n^{2\gamma} & \text{for } \gamma > \frac{1}{2}, \end{cases}$$

so that finally

$$\mathbb{V}\left[\sum_{i=1}^{n} \vartheta_{i,n} D_{n,i}\right] \leq C \begin{cases} n & \text{for } \gamma < \frac{1}{2}, \\ \log(n) n & \text{for } \gamma = \frac{1}{2}, \\ n^{2\gamma} & \text{for } \gamma > \frac{1}{2}. \end{cases}$$

□

The last part to deal with is $\mathbb{V}\left[\sum_{i=1}^{n} \vartheta_{i,n} R_{n,i}\right]$. We have

$$\mathbb{V}\left[\sum_{i=1}^{n} \vartheta_{i,n} R_{n,i}\right] = \sum_{i=1}^{n} \vartheta_{i,n}^2 \mathbb{V}\left[R_{n,i}\right] + 2 \sum_{i=1}^{n} \sum_{j=1}^{i-1} \vartheta_{i,n} \vartheta_{j,n} \mathrm{Cov}\left[R_{n,i}, R_{n,j}\right].$$

We consider both sums separately. The corresponding results are formulated in Lemma 5.7 and Lemma 5.8 respectively. In order to prove these recall that $\mathcal{D}_{n,i}$ denotes the set of vertices who are only connected to vertex i in $\mathcal{G}_n$. We now define the random variables

$$\begin{aligned}
Z_{n,\ell}^{(i)} &= \mathbb{1}\{\ell \text{ is isolated in } \mathcal{G}_n^{(i)} \text{ but not in } \mathcal{G}_n \cap \ell \notin \mathcal{D}_{n,i}\} \\
&= \mathbb{1}\{N^{\neq i}(\ell) \cap E_+^{(i)}(\ell) \cap E_-^{(i)}(\ell)\}
\end{aligned}$$

with

$$N^{\neq i}(\ell) := \{\exists k \neq i : \{\ell \leftrightarrow k\}\},\ E_+^{(i)}(\ell) := \bigcap_{r=1}^{\ell-1}\{\ell \overset{i}{\not\twoheadrightarrow} r\},\ E_-^{(i)}(\ell) := \bigcap_{r=\ell+1}^{n}\{r \overset{i}{\not\twoheadrightarrow} \ell\},$$

so that $R_{n,i} = \sum_\ell^n Z_{n,\ell}^{(i)}$. We now recall the dependencies and independencies in the network that should be kept in mind since the subsequent calculations crucially rely on these. One important feature of the network, which we already mentioned before, is the independence of in- and outdegree of a fixed vertex as well as the independence of outgoing edges of a fixed vertex. Moreover it is useful to remember that younger vertices only contribute to the indegree of older vertices and older vertices can only contribute to the outdegree of younger vertices. More precisely, the event $\{Z_{n,\ell}^{(i)} = 1\}$ only has an impact on $\deg^+(k)$ but not on $\deg_n^-(k)$ if $k > \ell$. Furthermore, remember that by the construction of $\mathcal{G}_n^i$ the isolation of vertex i does not affect edges $\{r \to \ell\}$ if $\ell, r > i$, since neither of the two vertices involved depend on the in- or the outdegree of i, so the edge remains unaffected by the isolation of i, i.e

$$\{r \overset{i}{\not\twoheadrightarrow} \ell\} = \{r \not\twoheadrightarrow \ell\} \text{ for } r, \ell > i.$$

In particular this means that in the case that $\ell > i$ for the event $\{Z_{n,\ell}^{(i)} = 1\}$ to occur ℓ might neither have any incoming edges nor any outgoing edges to vertices younger than i, as these connections are unaffected by the isolation of i, so that

$$E_-^{(i)}(\ell) = \bigcap_{r=\ell+1}^{n}\{r \overset{i}{\not\twoheadrightarrow} \ell\} = \bigcap_{r=\ell+1}^{n}\{r \not\twoheadrightarrow \ell\} = \{\deg_n^-(\ell) = 0\}$$

and

$$E_+^{(i)}(\ell) = \bigcap_{r=1}^{\ell-1}\{\ell \overset{i}{\not\twoheadrightarrow} r\} = \bigcap_{r=1}^{i-1}\{\ell \overset{i}{\not\twoheadrightarrow} r\} \cap \bigcap_{r=i+1}^{\ell-1}\{r \not\twoheadrightarrow \ell\} \text{ for } \ell > i.$$

The following proposition states that the likelihood of an edge to exist decreases if the isolation of vertex I leads to the isolation of vertices not in $\mathcal{D}_{n,I}$, i.e. $Z_{n,\ell}^{(I)} = 1$ for some $i \notin \mathcal{D}_{n,I} \cup I$. An intuative example for this might be given by the fact that more than just $D_{n,I}$ vertices loose all their present connections due to the isolation of I hints at a rather sparse graph, since the probability for the deletion of an edge is rather small (cf. (5.2)).

Proposition 5.6. *For the random variables $Z_{n,\ell}^{(i)}$ defined above, we have*

$$\mathbb{P}(m \to k | Z_{n,\ell}^{(i)} = 1) \leq \mathbb{P}(m \to k)$$

Proof. For $k > \ell$ (and thus $m > \ell$) the statement is true and equality holds, since $\{m \to k\}$ does not have any influence on connections of vertex ℓ. So we have to consider the case $k \leq \ell$. We have

$$\mathbb{P}(m \to k | Z_{n,\ell}^{(i)} = 1) = \mathbb{P}(m \to k | Z_{n,\ell}^{(i)} = 1, \ell \to k)\mathbb{P}(\ell \to k | Z_{n,\ell}^{(i)} = 1)$$

$$+ \mathbb{P}(m \to k | Z_{n,\ell}^{(i)} = 1, \ell \nrightarrow k)\mathbb{P}(\ell \nrightarrow k | Z_{n,\ell}^{(i)} = 1)$$

$$= \mathbb{P}(m \to k | \ell \to k)\mathbb{P}(\ell \to k | \ell \overset{i}{\nrightarrow} k)$$

$$+ \mathbb{P}(m \to k | \ell \nrightarrow k)\mathbb{P}(\ell \nrightarrow k | \ell \overset{i}{\nrightarrow} k)$$

and by Bayes' Theorem

$$\mathbb{P}(\ell \to k | \ell \overset{i}{\nrightarrow} k) = \frac{\mathbb{P}(\ell \overset{i}{\nrightarrow} k | \ell \to k)\mathbb{P}(\ell \to k)}{\mathbb{P}(\ell \overset{i}{\nrightarrow} k)}$$

$$= \mathbb{P}(\ell \to k) + \mathbb{P}(\ell \to k)\left(\frac{\mathbb{P}(\ell \overset{i}{\nrightarrow} k | \ell \to k)}{\mathbb{P}(\ell \overset{i}{\nrightarrow} k)} - 1\right)$$

as well as

$$\mathbb{P}(\ell \nrightarrow k | \ell \overset{i}{\nrightarrow} k) = \frac{\mathbb{P}(\ell \overset{i}{\nrightarrow} k | \ell \nrightarrow k)\mathbb{P}(\ell \nrightarrow k)}{\mathbb{P}(\ell \overset{i}{\nrightarrow} k)} = \frac{\mathbb{P}(\ell \nrightarrow k)}{\mathbb{P}(\ell \overset{i}{\nrightarrow} k)},$$

since edges not present in $\mathcal{G}_n$ cannot emerge in $\mathcal{G}_n^{(i)}$. We then obtain

$$\mathbb{P}(m \to k | Z_{n,\ell}^{(i)} = 1)$$

$$= \mathbb{P}(m \to k | \ell \to k)\mathbb{P}(\ell \to k) + \mathbb{P}(m \to k | \ell \to k)\mathbb{P}(\ell \to k)\left(\frac{\mathbb{P}(\ell \overset{i}{\nrightarrow} k | \ell \to k)}{\mathbb{P}(\ell \overset{i}{\nrightarrow} k)} - 1\right)$$

$$+ \mathbb{P}(m \to k | \ell \nrightarrow k)\mathbb{P}(\ell \nrightarrow k) + \mathbb{P}(m \to k | \ell \nrightarrow k)\mathbb{P}(\ell \nrightarrow k)\left(\frac{1}{\mathbb{P}(\ell \overset{i}{\nrightarrow} k)} - 1\right)$$

$$\leq \mathbb{P}(m \to k) + \mathbb{P}(m \to k | \ell \to k)\cdot$$

$$\left(\frac{\mathbb{P}(\ell \overset{i}{\nrightarrow} k | \ell \to k)\mathbb{P}(\ell \to k)}{\mathbb{P}(\ell \overset{i}{\nrightarrow} k)} - \mathbb{P}(\ell \to k) + \frac{\mathbb{P}(\ell \nrightarrow k)}{\mathbb{P}(\ell \overset{i}{\nrightarrow} k)} - \mathbb{P}(\ell \nrightarrow k)\right)$$

$$= \mathbb{P}(m \to k) + \mathbb{P}(m \to k | \ell \to k)\left(\frac{\mathbb{P}(\ell \overset{i}{\nrightarrow} k | \ell \to k)\mathbb{P}(\ell \to k) + \mathbb{P}(\ell \nrightarrow k)}{\mathbb{P}(\ell \overset{i}{\nrightarrow} k)} - 1\right).$$

By the construction of $\mathcal{G}_n^{(i)}$ from $\mathcal{G}_n$ we have

$$\mathbb{P}(\ell \overset{i}{\nrightarrow} k) = \mathbb{P}(\ell \nrightarrow k) + \mathbb{P}(\ell \to k)\mathbb{P}(\ell \overset{i}{\nrightarrow} k | \ell \to k),$$

which proves the assertion. □

This result now allows us to prove the following bound on $\sum_{i=1}^{n} R_{n,i}$:

Lemma 5.7. *For the random variable $R_{n,i}$ denoting the number of vertices not in $\mathcal{D}_{n,i}$ that lose all their connections in $\mathcal{G}_n$ due to the isolation of vertex i we have*

$$\sum_{i=1}^{n} \mathbb{V}\left[R_{n,i}\right] \leq \begin{cases} n & \text{for } \gamma < \frac{1}{2}, \\ n\log(n)^2 & \text{for } \gamma = \frac{1}{2}, \\ n^{2\gamma} & \text{for } \gamma > \frac{1}{2}. \end{cases}$$

Proof. With the definition of $Z_{n,\ell}^{(i)}$ we have

$$\begin{aligned}
\mathbb{V}\left[R_{n,i}\right] &= \mathbb{V}\left[\sum_{\ell=1}^{i-1} Z_{n,\ell}^{(i)} + \sum_{\ell=i+1}^{n} Z_{n,\ell}^{(i)}\right] \leq 2\left(\mathbb{V}\left[\sum_{\ell=1}^{i-1} Z_{n,\ell}^{(i)}\right] + \mathbb{V}\left[\sum_{\ell=i+1}^{n} Z_{n,\ell}^{(i)}\right]\right) \\
&= 2\sum_{\ell=1}^{i-1} \mathbb{V}\left[Z_{n,\ell}^{(i)}\right] + 4\sum_{\ell=1}^{i-1}\sum_{k=1}^{\ell-1} \mathrm{Cov}\left[Z_{n,\ell}^{(i)}, Z_{n,k}^{(i)}\right] \\
&\quad + 2\sum_{\ell=i+1}^{n} \mathbb{V}\left[Z_{n,\ell}^{(i)}\right] + 4\sum_{\ell=i+1}^{n}\sum_{k=\ell+1}^{n} \mathrm{Cov}\left[Z_{n,\ell}^{(i)}, Z_{n,k}^{(i)}\right].
\end{aligned}$$

Since

$$\mathbb{V}\left[Z_{n,\ell}^{(i)}\right] \leq \mathbb{P}(Z_{n,\ell}^{(i)} = 1)$$

we will now deal with the probability on the right-hand side. Therefor, we define the following events in order to condition on the first connection of a vertex:

$$\{\ell \overset{1.}{\leftrightarrow} m\} := \bigcap_{r=1}^{m-1} \{\ell \nrightarrow r\} \cap \{\ell \to m\} \quad \text{for } m < \ell$$

and

$$\{\ell \overset{1.}{\leftrightarrow} m\} := \bigcap_{r=1}^{\ell-1} \{\ell \nrightarrow r\} \cap \bigcap_{r=\ell+1}^{m-1} \{r \nrightarrow \ell\} \cap \{m \to \ell\} \quad \text{for } m > \ell.$$

Note that $Z_{n,\ell}^{(i)} = 0$ if ℓ is only connected to vertex i, so that in the case that $\ell \overset{1.}{\leftrightarrow} i$ the event $\{\ell \overset{1.}{\leftrightarrow} k\}$ refers to the first connection formed ignorant of all edges with endpoint in i.

$$\begin{aligned}
\mathbb{P}(Z_{n,\ell}^{(i)} = 1) &= \sum_{m=1}^{\ell-1} \mathbb{P}(Z_{n,\ell}^{(i)} = 1 | \ell \overset{1.}{\leftrightarrow} m)\mathbb{P}(\ell \overset{1.}{\leftrightarrow} m) + \sum_{\substack{m=\ell+1 \\ m \neq i}}^{n} \mathbb{P}(Z_{n,\ell}^{(i)} = 1 | \ell \overset{1.}{\leftrightarrow} m)\mathbb{P}(\ell \overset{1.}{\leftrightarrow} m) \\
&\leq \sum_{m=1}^{\ell-1} \frac{\mu_{\ell-1}(m)}{\ell} \mathbb{P}(\ell \overset{i}{\nrightarrow} m | \ell \to m) + \sum_{\substack{m=\ell+1 \\ m \neq i}}^{n} \frac{f(1)}{m} \prod_{\substack{k=\ell+1 \\ k \neq i}}^{m-1} \left(1 - \frac{f(0)}{k}\right) \mathbb{P}(m \overset{i}{\nrightarrow} \ell | m \to \ell)
\end{aligned}$$

$$\leq C \left(\sum_{m=1}^{\ell-1} \ell^{-1+\gamma} m^{-\gamma} i^{\gamma-1} m^{-\gamma} + \sum_{\substack{m=\ell+1 \\ m\neq i}}^{n} m^{-1} \left(\frac{\ell}{m}\right)^{\eta} \ell^{-1} i^{\gamma-1} \ell^{-\gamma} \right)$$

for $\ell < i$, where we used (5.6), (5.2) and the fact that $(1-\frac{f(0)}{i}) \geq \frac{1}{2}$ for $i \geq 2$. For $\ell > i$ we have that $\mathbb{P}(m \overset{i}{\not\to} \ell | m \to \ell) = 0$ and $\mathbb{P}(Z_{n,\ell}^{(i)} = 1 | \ell \overset{1.}{\leftrightarrow} m) = 0$ for $m \geq i+1$, since the isolation of vertex i only affects the indegree of vertices older than vertex i. Due to (5.2) it follows that

$$\begin{aligned}
\mathbb{P}(Z_{n,\ell}^{(i)} = 1) &= \sum_{m=1}^{i-1} \mathbb{P}(Z_{n,\ell}^{(i)} = 1 | \ell \overset{1.}{\leftrightarrow} m) \mathbb{P}(\ell \overset{1.}{\leftrightarrow} m) \\
&\leq \sum_{m=1}^{i-1} \mathbb{P}(\ell \overset{i}{\not\to} m | \ell \to m) \mathbb{P}(\ell \to m) \\
&\leq \sum_{m=1}^{i-1} \ell^{-1+\gamma} m^{-2\gamma} i^{\gamma-1}.
\end{aligned}$$

To summarize, we have

$$\mathbb{V}\left[Z_{n,\ell}^{(i)} = 1\right] \leq \mathbb{P}(Z_{n,\ell}^{(i)} = 1) \sim \begin{cases} i^{\gamma-1}\ell^{-\gamma} & \text{for } \gamma < \frac{1}{2}, \\ \ell^{-\frac{1}{2}} i^{-\frac{1}{2}} \log(\ell) & \text{for } \gamma = \frac{1}{2}, \\ \ell^{\gamma-1} i^{\gamma-1} & \text{for } \gamma > \frac{1}{2}, \end{cases} \tag{5.20}$$

for $\ell < i$ and

$$\mathbb{V}\left[Z_{n,\ell}^{(i)} = 1\right] \leq \mathbb{P}(Z_{n,\ell}^{(i)} = 1) \sim \begin{cases} i^{-\gamma}\ell^{\gamma-1} & \text{for } \gamma < \frac{1}{2}, \\ \ell^{-\frac{1}{2}} i^{-\frac{1}{2}} \log(i) & \text{for } \gamma = \frac{1}{2}, \\ \ell^{\gamma-1} i^{\gamma-1} & \text{for } \gamma > \frac{1}{2}, \end{cases} \tag{5.21}$$

for $\ell > i$. To deal with the covariances we write

$$\begin{aligned}
&\mathbb{P}(Z_{n,\ell}^{(i)} = 1, Z_{n,k}^{(i)} = 1) - \mathbb{P}(Z_{n,\ell}^{(i)} = 1)\mathbb{P}(Z_{n,k}^{(i)} = 1) \\
&= \mathbb{P}(Z_{n,\ell}^{(i)} = 1) \left(\mathbb{P}(Z_{n,k}^{(i)} = 1 | Z_{n,\ell}^{(i)} = 1) - \mathbb{P}(Z_{n,k}^{(i)} = 1) \right)
\end{aligned}$$

and again condition on the first connection of k in order to bound the conditional probability. We thus get

$$\mathbb{P}(Z_{n,k}^{(i)} = 1 | Z_{n,\ell}^{(i)} = 1) = \sum_{m=1}^{k-1} \mathbb{P}(Z_{n,k}^{(i)} = 1 | k \overset{1.}{\leftrightarrow} m, Z_{n,\ell}^{(i)} = 1) \mathbb{P}(k \overset{1.}{\leftrightarrow} m | Z_{n,\ell}^{(i)} = 1)$$

$$
\begin{aligned}
&+\sum_{\substack{m=k+1\\ m\neq i}}^{n} \mathbb{P}(Z_{n,k}^{(i)}=1|k \overset{1}{\leftrightarrow} m, Z_{n,\ell}^{(i)}=1)\mathbb{P}(k \overset{1}{\leftrightarrow} m|Z_{n,\ell}^{(i)}=1)\\
&\le \sum_{m=1}^{k-1} \mathbb{P}(k \overset{i}{\not\to} m|k \to m)\mathbb{P}(k \to m)\\
&+\sum_{\substack{m=k+1\\ m\neq i}}^{n} \mathbb{P}(m \overset{i}{\not\to} k|m \to k)\mathbb{P}(m \to k)\\
&\le \sum_{m=1}^{k-1} i^{\gamma-1}k^{\gamma-1}m^{-2\gamma} + \sum_{\substack{m=k+1\\ m\neq i}}^{n} i^{\gamma-1}m^{\gamma-1}k^{-2\gamma}\\
&\le C\begin{cases} n^{\gamma}i^{\gamma-1}k^{-2\gamma} & \text{for } \gamma<\frac{1}{2},\\ i^{-\frac{1}{2}}k^{-\frac{1}{2}}\log(k)+n^{\frac{1}{2}}i^{-\frac{1}{2}}k^{-1} & \text{for } \gamma=\frac{1}{2},\\ i^{\gamma-1}k^{\gamma-1}+n^{\gamma}i^{\gamma-1}k^{-2\gamma} & \text{for } \gamma>\frac{1}{2}.\end{cases}
\end{aligned}
\tag{5.22}
$$

By repeated use of (5.6) and with (5.20) we obtain

$$
\sum_{i=1}^{n}\sum_{\ell=1}^{i-1}\sum_{k=1}^{\ell-1} \operatorname{Cov}\left[Z_{n,\ell}^{(i)}, Z_{n,k}^{(i)}\right] \le C\begin{cases} n^{2\gamma} & \text{for } \gamma<\frac{1}{2},\\ n\log(n) & \text{for } \gamma=\frac{1}{2},\\ n^{2\gamma} & \text{for } \gamma>\frac{1}{2}.\end{cases}
$$

With (5.20) this yields

$$
\sum_{i=1}^{n} \mathbb{V}\left[\sum_{\ell=1}^{i-1} Z_{n,\ell}^{(i)}\right] \le C\begin{cases} n & \text{for } \gamma<\frac{1}{2},\\ n\log(n) & \text{for } \gamma=\frac{1}{2},\\ n^{2\gamma} & \text{for } \gamma>\frac{1}{2}.\end{cases}
\tag{5.23}
$$

For $\ell \ge i+1$ variance and covariance can be handled in much the same way with only minor differences in the precise calculations. In fact, we have

$$
\begin{aligned}
\mathbb{P}(Z_{n,k}^{(i)}=1|Z_{n,\ell}^{(i)}=1) &= \sum_{m=1}^{i-1} \mathbb{P}(Z_{n,k}^{(i)}=1|k \overset{1}{\leftrightarrow} m, Z_{n,\ell}^{(i)}=1)\mathbb{P}(k \overset{1}{\leftrightarrow} m|Z_{n,\ell}^{(i)}=1)\\
&= \sum_{m=1}^{i-1} \mathbb{P}(\deg_n^-(k)=0)\mathbb{P}(k \overset{i}{\not\to} m|k \to m)\prod_{r=m+1}^{i-1} \mathbb{P}(k \overset{i}{\not\to} r|\ell \overset{i}{\not\to} r)\\
&\quad \cdot \prod_{r=i+1}^{\ell-1} \mathbb{P}(k \not\to r|\ell \not\to r)\prod_{r=\ell+1}^{k-1} \mathbb{P}(k \overset{i}{\not\to} r)\mathbb{P}(k \overset{1}{\leftrightarrow} m|Z_{n,\ell}^{(i)}=1)\\
&\le \sum_{m=1}^{i-1} \mathbb{P}(\deg_n^-(k)=0)\mathbb{P}(k \overset{i}{\not\to} m|k \to m)\prod_{r=m+1}^{i-1} \mathbb{P}(k \overset{i}{\not\to} r|\ell \overset{i}{\not\to} r)
\end{aligned}
$$

$$\cdot \prod_{r=i+1}^{\ell-1} \mathbb{P}(k \nrightarrow r | \ell \nrightarrow r) \prod_{r=\ell+1}^{k-1} \mathbb{P}(k \overset{i}{\nrightarrow} r) \prod_{r=1}^{k-1} \mathbb{P}(k \nrightarrow r | \ell \nrightarrow r) \mathbb{P}(k \to m), \quad (5.24)$$

where we used that $\mathbb{P}(k \to m | Z_{n,\ell}^{(i)} = 1) \leq \mathbb{P}(k \to m)$ by Proposition 5.6 and $\mathbb{P}(k \nrightarrow r | Z_{n,\ell}^{(i)} = 1) \leq \mathbb{P}(k \nrightarrow r | \ell \nrightarrow r)$. In the same manner we can see that

$$\mathbb{P}(Z_{n,k}^{(i)} = 1) = \sum_{m=1}^{i-1} \mathbb{P}(Z_{n,k}^{(i)} = 1 | k \overset{1}{\leftrightarrow} m) \mathbb{P}(k \overset{1}{\leftrightarrow} m)$$
$$= \sum_{m=1}^{i-1} \mathbb{P}(\deg_n^-(k) = 0) \mathbb{P}(k \overset{i}{\nrightarrow} m | k \to m) \prod_{r=m+1}^{i-1} \mathbb{P}(k \overset{i}{\nrightarrow} r) \prod_{r=i+1}^{k-1} \mathbb{P}(k \nrightarrow r) \mathbb{P}(k \to m).$$

Proposition 5.4 gives

$$\prod_{r=i+1}^{\ell-1} \mathbb{P}(k \nrightarrow r | \ell \nrightarrow r) \leq \prod_{r=i+1}^{\ell-1} \mathbb{P}(k \nrightarrow r) + \xi_{i+1}^{\ell-1}(m, \ell) + \xi_{i+1}^{\ell-1}(k, \ell)$$

and putting this into (5.24) we observe that

$$\mathbb{P}(Z_{n,k}^{(i)} = 1 | Z_{n,\ell}^{(i)} = 1) - \mathbb{P}(Z_{n,k}^{(i)} = 1)$$
$$\leq \sum_{m=1}^{i-1} \mathbb{P}(\deg_n^-(k) = 0) \mathbb{P}(k \to m) \mathbb{P}(k \overset{i}{\nrightarrow} m | k \to m) \prod_{r=\ell+1}^{k-1} \mathbb{P}(k \nrightarrow r) \prod_{r=i+1}^{k-1} \mathbb{P}(k \nrightarrow r)$$
$$\cdot \left(\prod_{r=m+1}^{i-1} \mathbb{P}(k \overset{i}{\nrightarrow} r | \ell \overset{i}{\nrightarrow} r) - \prod_{r=m+1}^{i-1} \mathbb{P}(k \overset{i}{\nrightarrow} r) \right)$$
$$+ \sum_{m=1}^{i-1} \mathbb{P}(\deg_n^-(k) = 0) \mathbb{P}(k \to m) \mathbb{P}(k \overset{i}{\nrightarrow} m | k \to m)\, \xi_{i+1}^{\ell-1}(k, \ell).$$

Note that

$$\mathbb{P}(k \overset{i}{\nrightarrow} r | \ell \overset{i}{\nrightarrow} r) - \mathbb{P}(k \overset{i}{\nrightarrow} r)$$
$$= \frac{\tilde{\mu}_{m-1}^{f}(r, i)(\tilde{\mu}_{m-1}^{f}(r, \ell) - \mu_{m-1}^{f}(r))}{m \mu_{m-1}^{f}(r)} \leq \frac{(\tilde{\mu}_{m-1}^{f}(r, \ell) - \mu_{m-1}^{f}(r))}{m}$$
$$\leq m^{\gamma-1} \ell^{\gamma-1} r^{-2\gamma},$$

so we can proceed just as in the proof of Proposition 5.4 to show that for all $m < j < i - 1$

$$\prod_{r=m}^{j} \mathbb{P}(k \overset{i}{\nrightarrow} r | \ell \overset{i}{\nrightarrow} r) - \prod_{r=m}^{j} \mathbb{P}(k \overset{i}{\nrightarrow} r) \leq \xi_m^j(k, \ell)$$

and hence

$$\begin{aligned}
\mathbb{P}(Z_{n,k}^{(i)} = 1 | Z_{n,\ell}^{(i)} = 1) - \mathbb{P}(Z_{n,k}^{(i)} = 1)& \\
&\leq \sum_{m=1}^{i-1} \mathbb{P}(\deg_n^-(k) = 0)\mathbb{P}(k \to m)\mathbb{P}(k \overset{i}{\not\to} m | k \to m)\xi_{m+1}^{i-1}(k,\ell) \\
&\quad + \sum_{m=1}^{i-1} \mathbb{P}(\deg_n^-(k) = 0)\mathbb{P}(k \to m)\mathbb{P}(k \overset{i}{\not\to} m | k \to m)\ \xi_{i+1}^{\ell-1}(k,\ell) \\
&\leq 2\sum_{m=1}^{i-1} \mathbb{P}(\deg_n^-(k) = 0)\mathbb{P}(k \to m)\mathbb{P}(k \overset{i}{\not\to} m | k \to m)\ \xi_{m+1}^{\ell-1}(k,\ell) \\
&\leq C\sum_{m=1}^{i-1} \left(\frac{k}{n}\right)^{\eta} k^{\gamma-1}m^{-\gamma}i^{\gamma-1}m^{-\gamma}\xi_{m+1}^{\ell-1}(k,\ell)
\end{aligned}$$

for $m < j < i < \ell$. Here we used (5.2) and the fact that $\xi_{m+1}^{i-1}(k,\ell) \leq \xi_{m+1}^{\ell-1}(k,\ell)$ as well as $\xi_{i+1}^{\ell-1}(k,\ell) \leq \xi_{m+1}^{\ell-1}(k,\ell)$ by the definition of ξ. According to Proposition 5.4 we have

$$\xi_{m+1}^{\ell-1}(k,\ell) \leq C \begin{cases} k^{\gamma-1}\ell^{-\gamma} & \text{for } \gamma < \frac{1}{2}, \\ k^{-\frac{1}{2}}\ell^{-\frac{1}{2}}\log(\ell) & \text{for } \gamma = \frac{1}{2}, \\ k^{\gamma-1}\ell^{\gamma-1}m^{2\gamma-1} & \text{for } \gamma > \frac{1}{2}, \end{cases}$$

which in combination with (5.6) gives

$$\mathbb{P}(Z_{n,k}^{(i)} = 1 | Z_{n,\ell}^{(i)} = 1) - \mathbb{P}(Z_{n,k}^{(i)} = 1) \leq \frac{C}{n^{\eta}} \begin{cases} i^{-\gamma}k^{2\gamma+\eta-2}\ell^{-\gamma} & \text{for } \gamma < \frac{1}{2}, \\ i^{-\frac{1}{2}}\ell^{-\frac{1}{2}}k^{\eta-1}\log(i)\log(\ell) & \text{for } \gamma = \frac{1}{2}, \\ i^{\gamma-1}k^{\eta-1}\ell^{\gamma-1} & \text{for } \gamma > \frac{1}{2}. \end{cases}$$

Combining these results with (5.21) and repeatedly using (5.6) yields

$$\sum_{i=1}^{n}\sum_{\ell=i+1}^{n}\sum_{k=\ell+1}^{n} \operatorname{Cov}\left[Z_{n,\ell}^{(i)}, Z_{n,k}^{(i)}\right] \leq \begin{cases} n & \text{for } \gamma < \frac{1}{2}, \\ n\log(n)^2 & \text{for } \gamma = \frac{1}{2}, \\ n^{2\gamma} & \text{for } \gamma > \frac{1}{2}. \end{cases}$$

(5.21) and (5.6) eventually yield

$$\sum_{i=1}^{n} \mathbb{V}\left[\sum_{\ell=i+1}^{n} Z_{n,\ell}^{(i)}\right] \leq \begin{cases} n & \text{for } \gamma < \frac{1}{2}, \\ n\log(n)^2 & \text{for } \gamma = \frac{1}{2}, \\ n^{2\gamma} & \text{for } \gamma > \frac{1}{2}, \end{cases}$$

so that with (5.23) we finally obtain the desired result. □

Lemma 5.8. *For $R_{n,i}$ denoting the number of vertices in $\mathcal{G}_n^{(i)}$ that lose all their connections in $\mathcal{G}_n$ due to the isolation of vertex i and which are neither isolated in $\mathcal{G}_n$ nor contained in $\mathcal{D}_{n,i}$, we have*

$$\sum_{i=1}^{n}\sum_{j=1}^{i-1}\vartheta_{i,n}\vartheta_{j,n}\operatorname{Cov}\left[R_{n,i},R_{n,j}\right]\leq C\begin{cases}n & \text{for } \gamma<\frac{1}{2},\\ n\log(n)^4 & \text{for } \gamma=\frac{1}{2},\\ n^{6\gamma-2} & \text{for } \gamma>\frac{1}{2}.\end{cases}$$

Proof. As $R_{n,i}=\sum_{\ell=1}^{n}Z_{n,\ell}^{(i)}$ we get

$$\begin{aligned}\operatorname{Cov}\left[R_{n,i},R_{n,j}\right]&=\sum_{\ell=1}^{n}\sum_{m=1}^{n}\operatorname{Cov}\left[Z_{n,\ell}^{(i)},Z_{n,m}^{(j)}\right]\\&=\sum_{\ell=1}^{n}\operatorname{Cov}\left[Z_{n,\ell}^{(i)},Z_{n,\ell}^{(j)}\right]+\sum_{\ell=1}^{n}\sum_{\substack{m=1\\m\neq\ell}}^{n}\operatorname{Cov}\left[Z_{n,\ell}^{(i)},Z_{n,m}^{(j)}\right].\end{aligned}$$

Starting with the first sum (where $m=\ell$) we get

$$\operatorname{Cov}\left[Z_{n,\ell}^{(i)},Z_{n,\ell}^{(j)}\right]=\mathbb{P}(Z_{n,\ell}^{(i)}=1)\left(\mathbb{P}(Z_{n,\ell}^{(j)}=1|Z_{n,\ell}^{(i)}=1)-\mathbb{P}(Z_{n,\ell}^{(j)}=1)\right)$$

so that we have to bound $\mathbb{P}(Z_{n,\ell}^{(j)}=1|Z_{n,\ell}^{(i)}=1)-\mathbb{P}(Z_{n,\ell}^{(j)}=1)$. As mentioned before, edges of vertex i do not have an impact on edges $\{m\to\ell\}$ if $m,\ell>i$, so that the isolation of i only affects those edges with at least one endpoint older than i. Hence for $\ell\geq$ i+1 we have

$$\begin{aligned}\mathbb{P}(Z_{n,\ell}^{(j)}=1|Z_{n,\ell}^{(i)}=1)&=\sum_{k=1}^{j-1}\mathbb{P}(Z_{n,\ell}^{(j)}=1|Z_{n,\ell}^{(i)}=1,\ell\overset{1.}{\leftrightarrow}k)\mathbb{P}(\ell\overset{1.}{\leftrightarrow}k|Z_{n,\ell}^{(i)}=1)\\&\leq\sum_{k=1}^{j-1}\mathbb{P}(\ell\overset{j}{\not\to}k|\ell\to k)\mathbb{P}(\ell\to k|Z_{n,\ell}^{(i)}=1),\end{aligned}$$

and by Bayes' Theorem

$$\begin{aligned}\mathbb{P}(\ell\to k|Z_{n,\ell}^{(i)}=1)=\mathbb{P}(\ell\to k|\ell\overset{i}{\not\to}k)&=\frac{\mathbb{P}(\ell\overset{i}{\not\to}k|\ell\to k)}{\mathbb{P}(\ell\overset{i}{\not\to}k)}\mathbb{P}(\ell\to k)\\&\leq C\,i^{\gamma-1}k^{-\gamma}\ell^{\gamma-1}k^{-\gamma},\end{aligned}\tag{5.25}$$

as

$$\frac{1}{\mathbb{P}(\ell\overset{i}{\not\to}k)}\leq\frac{1}{\mathbb{P}(\ell\not\to k)}=\frac{\ell}{\ell-\mu_{\ell-1}^{f}(k)}\leq\frac{1}{1-\ell^{\gamma-1}}\leq\frac{1}{1-2^{\gamma-1}}.$$

With (5.2) we then get

$$\mathbb{P}(Z_{n,\ell}^{(j)} = 1 | Z_{n,\ell}^{(i)} = 1) \leq C \sum_{k=1}^{j-1} j^{\gamma-1} k^{-3\gamma} \ell^{\gamma-1} i^{\gamma-1}$$

$$\leq C \begin{cases} j^{-\gamma} \ell^{\gamma-1} i^{\gamma-1} & \text{for } \gamma < \frac{1}{2}, \\ \log(j) j^{-\frac{1}{2}} \ell^{-\frac{1}{2}} i^{-\frac{1}{2}} & \text{for } \gamma = \frac{1}{2}, \\ j^{\gamma-1} \ell^{\gamma-1} i^{\gamma-1} & \text{for } \gamma > \frac{1}{2} \end{cases}$$

and by (5.6) and (5.20) we obtain

$$\sum_{i=1}^{n} \sum_{j=1}^{i-1} \sum_{\ell=j+1}^{n} \operatorname{Cov}\left[Z_{n,\ell}^{(i)}, Z_{n,\ell}^{(j)}\right] \leq C \begin{cases} n^{\gamma} & \text{for } \gamma < \frac{1}{2}, \\ \log(n)^2 n^{\frac{1}{2}} & \text{for } \gamma = \frac{1}{2}, \\ n^{5\gamma-2} & \text{for } \gamma > \frac{1}{2}, \end{cases}$$

where we used that $\vartheta_{n,j} \sim \left(\frac{j}{n}\right)^{\eta} \leq \left(\frac{\ell}{n}\right)^{\eta}$. Using (5.25) in the case $\ell < j$ on the first sum we obtain

$$\begin{aligned} \mathbb{P}(Z_{n,\ell}^{(j)} = 1 | Z_{n,\ell}^{(i)} = 1) &\leq \sum_{k=1}^{\ell-1} \mathbb{P}(\ell \overset{j}{\not\to} k | \ell \to k) \mathbb{P}(\ell \to k | Z_{n,\ell}^{(i)} = 1) \\ &\quad + \sum_{k=\ell+1}^{n} \mathbb{P}(k \overset{j}{\not\to} \ell | k \to \ell) \mathbb{P}\Big(k \to \ell \Big| \bigcap_{m=\ell+1}^{k-1} \{m \not\to \ell\}, Z_{n,\ell}^{(i)} = 1\Big) \\ &\leq C \left(\sum_{k=1}^{\ell-1} j^{\gamma-1} k^{-2\gamma} \ell^{\gamma-1} i^{\gamma-1} + \sum_{k=\ell+1}^{n} j^{\gamma-1} \ell^{-2\gamma} k^{-1} i^{\gamma-1} \right) \\ &\leq C \log(n) \begin{cases} j^{\gamma-1} \ell^{-2\gamma} i^{\gamma-1} & \text{for } \gamma < \frac{1}{2}, \\ j^{-\frac{1}{2}} \ell^{-\frac{1}{2}} i^{-\frac{1}{2}} & \text{for } \gamma = \frac{1}{2}, \\ j^{\gamma-1} \ell^{\gamma-1} i^{\gamma-1} & \text{for } \gamma > \frac{1}{2}, \end{cases} \end{aligned}$$

and using (5.20) we thus obtain

$$\begin{aligned} &\sum_{i=1}^{n} \sum_{j=1}^{i-1} \sum_{\ell=1}^{j-1} \operatorname{Cov}\left[Z_{n,\ell}^{(i)}, Z_{n,\ell}^{(j)}\right] \\ &\quad = \sum_{i=1}^{n} \sum_{j=1}^{i-1} \sum_{\ell=1}^{j-1} \mathbb{P}(Z_{n,\ell}^{(i)} = 1) \Big(\mathbb{P}(Z_{n,\ell}^{(j)} = 1 | Z_{n,\ell}^{(i)} = 1) - \mathbb{P}(Z_{n,\ell}^{(j)} = 1) \Big) \end{aligned}$$

$$\leq C\ \log(n)\sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{\ell=1}^{j-1}\begin{cases} j^{\gamma-1}\ell^{-3\gamma}i^{2\gamma-2} & \text{for } \gamma<\frac{1}{2},\\ j^{-\frac{1}{2}}\ell^{-1}i^{-1}\log(\ell) & \text{for } \gamma=\frac{1}{2},\\ j^{\gamma-1}\ell^{2\gamma-2}i^{2\gamma-2} & \text{for } \gamma>\frac{1}{2},\end{cases}$$

$$\leq C\begin{cases} \log(n)n^{\gamma} & \text{for } \gamma<\frac{1}{2},\\ \log(n)^3\sqrt{n} & \text{for } \gamma=\frac{1}{2},\\ \log(n)n^{5\gamma-2} & \text{for } \gamma>\frac{1}{2},\end{cases}$$

where we used (5.6). It thus remains to deal with

$$\mathrm{Cov}\left[Z_{n,\ell}^{(i)},Z_{n,m}^{(j)}\right]=\mathbb{P}(Z_{n,\ell}^{(i)}=1)(\mathbb{P}(Z_{n,m}^{(j)}=1|Z_{n,\ell}^{(i)}=1)-\mathbb{P}(Z_{n,m}^{(j)}=1))$$

for $m\neq\ell$. To find a bound on $\mathbb{P}(Z_{n,\ell}^{(i)}=1|Z_{n,m}^{(j)}=1)-\mathbb{P}(Z_{n,\ell}^{(i)}=1)$ we proceed just as in the previous case and condition on the first edge connecting ℓ to some other vertex of the network, i.e. on the event $\{\ell\overset{1.}{\leftrightarrow}k\}$. For $\ell<m$ we get

$$\begin{aligned}\mathbb{P}(Z_{n,\ell}^{(i)}=1|Z_{n,m}^{(j)}=1)=&\sum_{k=1}^{(i-1)\wedge(\ell-1)}\mathbb{P}(Z_{n,\ell}^{(i)}=1|Z_{n,m}^{(j)}=1,\{\ell\overset{1.}{\leftrightarrow}k\})\mathbb{P}(\ell\overset{1.}{\leftrightarrow}k|Z_{n,m}^{(j)}=1)\\ &+\mathbb{1}\{\ell<i\}\sum_{k=\ell+1}^{n}\mathbb{P}(Z_{n,\ell}^{(i)}=1|Z_{n,m}^{(j)}=1,\{\ell\overset{1.}{\leftrightarrow}k\})\mathbb{P}(\ell\overset{1.}{\leftrightarrow}k|Z_{n,m}^{(j)}=1)\\ &:=T_{1,1}+T_{1,2},\end{aligned}$$

where $(i-1)\wedge(\ell-1)=\min\{i-1,\ell-1\}$. Analogously we get

$$\begin{aligned}\mathbb{P}(Z_{n,\ell}^{(i)}=1)=&\sum_{k=1}^{(i-1)\wedge(\ell-1)}\mathbb{P}(Z_{n,\ell}^{(i)}=1|\{\ell\overset{1.}{\leftrightarrow}k\})\mathbb{P}(\ell\overset{1.}{\leftrightarrow}k)\\ &+\mathbb{1}\{\ell<i\}\sum_{k=\ell+1}^{n}\mathbb{P}(Z_{n,\ell}^{(i)}=1|\{\ell\overset{1.}{\leftrightarrow}k\})\mathbb{P}(\ell\overset{1.}{\leftrightarrow}k)\\ &:=T_{2,1}+T_{2,2}.\end{aligned}$$

Remember that

$$\{Z_{n,\ell}^{(i)}=1\}=\bigcap_{r=1}^{\ell-1}\{\ell\overset{i}{\not\to}r\}\ \cap\bigcap_{r=\ell+1}^{n}\{r\overset{i}{\not\to}\ell\}\cap\{\ell\notin\mathcal{D}_{n,i}\}$$

and

$$\{\ell\overset{1.}{\leftrightarrow}k\}=\bigcap_{r=1}^{k-1}\{\ell\not\to r\}\cap\{\ell\to k\}$$

for $k < \ell$ and

$$\{\ell \overset{1.}{\leftrightarrow} k\} = \bigcap_{r=1}^{\ell-1} \{\ell \nrightarrow r\} \ \cap \bigcap_{r=\ell+1}^{k-1} \{\ell \nrightarrow r\} \cap \{k \to \ell\} \text{ for } k > \ell.$$

Thus, for $k \leq \ell - 1$ we obtain

$$\begin{aligned}
&\mathbb{P}\left(Z_{n,\ell}^{(i)} = 1 | Z_{n,m}^{(j)} = 1, \{\ell \overset{1.}{\leftrightarrow} k\}\right) \\
&= \mathbb{P}\Big(\bigcap_{r=k}^{\ell-1} \{\ell \overset{i}{\nrightarrow} r\}, \bigcap_{r=\ell+1}^{n} \{r \overset{i}{\nrightarrow} \ell\} \Big| \bigcap_{r=k+1}^{\ell} \{m \overset{j}{\nrightarrow} r\}, \{\ell \to k\}\Big) \\
&= \mathbb{P}\Big(\bigcap_{r=k}^{(\ell-1)\wedge(i-1)} \{\ell \overset{i}{\nrightarrow} r\}, \bigcap_{r=i+1}^{\ell-1} \{\ell \nrightarrow r\} \Big| \bigcap_{r=k+1}^{\ell-1} \{m \overset{j}{\nrightarrow} r\}, \{\ell \to k\}\Big) \mathbb{P}\Big(\bigcap_{r=\ell+1}^{n} \{r \overset{i}{\nrightarrow} \ell\} \Big| m \overset{j}{\nrightarrow} \ell\Big) \\
&= \mathbb{P}(\ell \overset{i}{\nrightarrow} k | \ell \to k) \prod_{r=k+1}^{(\ell-1)\wedge(i-1)} \mathbb{P}(\ell \overset{i}{\nrightarrow} r | m \overset{j}{\nrightarrow} r) \prod_{r=i+1}^{\ell-1} \mathbb{P}(\ell \nrightarrow r) \mathbb{P}\Big(\bigcap_{r=\ell+1}^{n} \{r \overset{i}{\nrightarrow} \ell\} \Big| m \overset{j}{\nrightarrow} \ell\Big),
\end{aligned}$$

where we used the independence of in- and outdegree of a fixed vertex, the fact that decisions for outgoing edges of a given vertex are made independently from each other as well as $\{\ell \overset{i}{\nrightarrow} k\} = \{\ell \nrightarrow k\}$ for $\ell, k > i$. As $k < \ell < m$ we furthermore get

$$\begin{aligned}
\mathbb{P}(\ell \overset{1.}{\leftrightarrow} k | Z_{n,m}^{(j)} = 1) &= \mathbb{P}\Big(\bigcap_{r=1}^{k-1} \{\ell \nrightarrow r\} \cap \{\ell \to k\} | \bigcap_{r=1}^{\ell-1} \{m \overset{j}{\nrightarrow} r\}\Big) \\
&= \prod_{r=1}^{k-1} \mathbb{P}(\ell \nrightarrow r | m \overset{j}{\nrightarrow} r) \mathbb{P}(\ell \to k | m \overset{j}{\nrightarrow} k) \\
&\leq \prod_{r=1}^{k-1} \mathbb{P}(\ell \nrightarrow r | m \nrightarrow r) \mathbb{P}(\ell \to k)
\end{aligned}$$

and for $k \geq \ell + 1$

$$\begin{aligned}
\mathbb{P}(\ell \overset{1.}{\leftrightarrow} k | Z_{n,m}^{(j)} = 1) \leq &\prod_{r=1}^{\ell-1} \mathbb{P}\left(\ell \nrightarrow r | m \nrightarrow r\right) \prod_{r=\ell+1}^{k-1} \mathbb{P}\Big(r \nrightarrow \ell \Big| m \overset{j}{\nrightarrow} \ell, \bigcap_{s=\ell+1}^{r-1} \{s \nrightarrow \ell\}\Big) \\
&\cdot \mathbb{P}(k \to \ell | \{m \overset{j}{\nrightarrow} \ell\}, \deg_{k-1}^{-}(\ell) = 0).
\end{aligned}$$

In the same manner we deduce

$$\mathbb{P}(Z_{n,\ell}^{(i)} = 1 | \ell \overset{1.}{\leftrightarrow} k) = \mathbb{P}(\ell \overset{i}{\nrightarrow} k | \ell \to k) \prod_{r=k+1}^{(\ell-1)\wedge(i-1)} \mathbb{P}(\ell \overset{i}{\nrightarrow} r) \prod_{r=i+1}^{\ell-1} \mathbb{P}(\ell \nrightarrow r) \mathbb{P}\left(\bigcap_{r=\ell+1}^{n} \{r \overset{i}{\nrightarrow} \ell\}\right),$$

$$\mathbb{P}(\ell \overset{1.}{\leftrightarrow} k) = \mathbb{P}\Big(\bigcap_{r=1}^{k-1} \{\ell \nrightarrow r\} \cap \{\ell \to k\}\Big) = \prod_{r=1}^{k-1} \mathbb{P}(\ell \nrightarrow r) \mathbb{P}(\ell \to k)$$

for $k \leq \ell$ and

$$\mathbb{P}(\ell \overset{1}{\leftrightarrow} k) = \prod_{r=1}^{\ell-1} \mathbb{P}(\ell \nrightarrow r) \prod_{r=\ell+1}^{k-1} \mathbb{P}\Big(r \nrightarrow \ell \Big| \bigcap_{s=\ell+1}^{r-1} \{s \nrightarrow \ell\}\Big) \mathbb{P}(k \to \ell | \deg^-_{k-1}(\ell) = 0)$$

for $k \geq \ell + 1$. One can see that the terms in the sums above differ in the three factors

$$\mathbb{P}(\ell \overset{1}{\leftrightarrow} k | Z^{(j)}_{n,m} = 1) \text{ and } \mathbb{P}(\ell \overset{1}{\leftrightarrow} k), \quad \prod_{r=k+1}^{(\ell-1)\wedge(i-1)} \mathbb{P}(\ell \overset{i}{\nrightarrow} r | m \overset{j}{\nrightarrow} r) \text{ and } \prod_{r=k+1}^{(\ell-1)\wedge(i-1)} \mathbb{P}(\ell \overset{i}{\nrightarrow} r)$$

$$\text{as well as } \mathbb{P}\Big(\bigcap_{r=\ell+1}^{n} \{r \overset{i}{\nrightarrow} \ell\} \Big| m \overset{j}{\nrightarrow} \ell \Big) \text{ and } \mathbb{P}\Big(\bigcap_{r=\ell+1}^{n} \{r \overset{i}{\nrightarrow} \ell\} \Big). \tag{5.26}$$

Due to Proposition 5.4 we have

$$\begin{aligned}
&\mathbb{P}(\ell \overset{1}{\leftrightarrow} k | Z^{(j)}_{n,m} = 1) - \mathbb{P}(\ell \overset{1}{\leftrightarrow} k) \\
&\leq \prod_{r=1}^{k-1} \mathbb{P}\left(\ell \nrightarrow r | m \overset{j}{\nrightarrow} r\right) \mathbb{P}\left(\ell \to k | m \overset{j}{\nrightarrow} k\right) - \prod_{r=1}^{k-1} \mathbb{P}\left(\ell \nrightarrow r\right) \mathbb{P}\left(\ell \to k\right) \\
&\leq \mathbb{P}\left(\ell \to k\right) \left(\prod_{r=1}^{k-1} \mathbb{P}\left(\ell \nrightarrow r | m \overset{j}{\nrightarrow} r\right) - \prod_{r=1}^{k-1} \mathbb{P}\left(\ell \nrightarrow r\right) \right) \\
&\leq \mathbb{P}\left(\ell \to k\right) \xi_1^{k-1}(\ell, m)
\end{aligned}$$

and noting that for any $r \leq \ell - 1$

$$\begin{aligned}
&\mathbb{P}(\ell \overset{i}{\nrightarrow} r | m \overset{j}{\nrightarrow} r) - \mathbb{P}(\ell \overset{i}{\nrightarrow} r) \leq \mathbb{P}(\ell \overset{i}{\nrightarrow} r | m \nrightarrow r) - \mathbb{P}(\ell \overset{i}{\nrightarrow} r) \\
&= \left(1 - \frac{\tilde{\mu}^f_{\ell-1}(r,m)}{\ell}\right) + \frac{\tilde{\mu}^f_{\ell-1}(r,m)}{\ell} \left(1 - \frac{\tilde{\mu}^f_{\ell-1}(r,i)}{\mu^f_{\ell-1}(r)}\right) - \left(1 - \frac{\tilde{\mu}^f_{\ell-1}(r,i)}{\ell}\right) \\
&\leq \frac{\mu^f_{\ell-1}(r) - \tilde{\mu}_{\ell-1}(r,m)}{\ell} \leq f(1)\ell^{\gamma-1} m^{\gamma-1} r^{-2\gamma},
\end{aligned}$$

we can again proceed just as in the proof of Proposition 5.4 to obtain

$$\prod_{r=k+1}^{(\ell-1)\wedge(i-1)} \mathbb{P}(\ell \overset{i}{\nrightarrow} r | m \overset{j}{\nrightarrow} r) - \prod_{r=k+1}^{(\ell-1)\wedge(i-1)} \mathbb{P}(\ell \overset{i}{\nrightarrow} r) \leq \xi_{k+1}^{(\ell-1)\wedge(i-1)}(\ell, m).$$

Using that

$$\mathbb{P}(m \overset{j}{\nrightarrow} \ell) \geq \mathbb{P}(m \nrightarrow \ell)$$

yields

$$\mathbb{P}\Big(\bigcap_{r=\ell+1}^{n}\{r \overset{i}{\not\to} \ell\}\Big| m \overset{j}{\not\to} \ell\Big) - \mathbb{P}\Big(\bigcap_{r=\ell+1}^{n}\{r \overset{i}{\not\to} \ell\}\Big) \le \mathbb{P}\Big(\bigcap_{r=\ell+1}^{n}\{r \overset{i}{\not\to} \ell\}\Big)\left(\frac{1}{\mathbb{P}(m \overset{j}{\not\to} \ell)} - 1\right)$$
$$\le \mathbb{P}\left(\bigcap_{r=\ell+1}^{n}\{r \overset{i}{\not\to} \ell\}\right)\left(\frac{1}{\mathbb{P}(m \not\twoheadrightarrow \ell)} - 1\right) = \mathbb{P}\left(\bigcap_{r=\ell+1}^{n}\{r \overset{i}{\not\to} \ell\}\right)\frac{\mu^f_{m-1}(\ell)}{m - \mu^f_{m-1}(\ell)}$$
$$\sim \mathbb{P}\left(\bigcap_{r=\ell+1}^{n}\{r \overset{i}{\not\to} \ell\}\right) m^{\gamma-1}\ell^{-\gamma}.$$

With these considerations we can now substitute the terms in (5.26) and rewrite $T_{1,1}$ given above.

$$T_{1,1} = \sum_{k=1}^{(i-1)\wedge(\ell-1)} \mathbb{P}(Z^{(i)}_{n,\ell} = 1 | Z^{(j)}_{n,m} = 1, \{\ell \overset{1}{\leftrightsquigarrow} k\})\mathbb{P}(\ell \overset{1}{\leftrightsquigarrow} k | Z^{(j)}_{n,m} = 1)$$
$$\le \sum_{k=1}^{(i-1)\wedge(\ell-1)} \mathbb{P}\big(\ell \overset{i}{\not\to} k \big| \ell \to k\big) \prod_{r=k+1}^{(\ell-1)\wedge(i-1)} \mathbb{P}(\ell \overset{i}{\not\to} r | m \overset{j}{\not\to} r) \prod_{r=i+1}^{\ell-1} \mathbb{P}(\ell \not\twoheadrightarrow r)$$
$$\cdot\, \mathbb{P}\Big(\bigcap_{r=\ell+1}^{n}\{r \overset{i}{\not\to} \ell\}\Big| m \overset{j}{\not\to} \ell\Big) \prod_{r=1}^{k-1} \mathbb{P}(\ell \not\twoheadrightarrow r | m \not\twoheadrightarrow r)\mathbb{P}(\ell \to k)$$
$$\le \sum_{k=1}^{(i-1)\wedge(\ell-1)} \mathbb{P}(\ell \overset{i}{\not\to} k | \ell \to k) \prod_{r=k+1}^{(\ell-1)\wedge(i-1)} \mathbb{P}(\ell \overset{i}{\not\to} r) \prod_{r=i+1}^{\ell-1} \mathbb{P}(\ell \not\twoheadrightarrow r)$$
$$\cdot\, \mathbb{P}\Big(\bigcap_{r=\ell+1}^{n}\{r \overset{i}{\not\to} \ell\}\Big| m \overset{j}{\not\to} \ell\Big) \prod_{r=1}^{k-1} \mathbb{P}(\ell \not\twoheadrightarrow r)\mathbb{P}(\ell \to k)$$
$$+ \sum_{k=1}^{(i-1)\wedge(\ell-1)} \xi_1^{k-1}(\ell,m)\mathbb{P}(\ell \overset{i}{\not\to} k | \ell \to k)\mathbb{P}(\ell \to k)\mathbb{P}\left(\bigcap_{r=\ell+1}^{n}\{r \overset{i}{\not\to} \ell\}\right)$$
$$+ \sum_{k=1}^{(i-1)\wedge(\ell-1)} \xi_{k+1}^{(i-1)\wedge(\ell-1)}(\ell,m)\mathbb{P}(\ell \overset{i}{\not\to} k | \ell \to k)\mathbb{P}(\ell \to k)\mathbb{P}\left(\bigcap_{r=\ell+1}^{n}\{r \overset{i}{\not\to} \ell\}\right)$$
$$+ \sum_{k=1}^{(i-1)\wedge(\ell-1)} m^{\gamma-1}\ell^{-\gamma}\mathbb{P}(\ell \overset{i}{\not\to} k | \ell \to k)\mathbb{P}(\ell \to k)\mathbb{P}\left(\bigcap_{r=\ell+1}^{n}\{r \overset{i}{\not\to} \ell\}\right)$$
$$\le T_{2,1} + 3 \cdot \sum_{k=1}^{(i-1)\wedge(\ell-1)} \psi(m,\ell,k) \cdot \mathbb{P}(\ell \overset{i}{\not\to} k | \ell \to k)\mathbb{P}(\ell \to k)\mathbb{P}\Big(\bigcap_{r=\ell+1}^{n}\{r \overset{i}{\not\to} \ell\}\Big),$$

where

$$\psi(m,\ell,k) = \max\{m^{\gamma-1}\ell^{-\gamma}, \xi_{k+1}^{(i-1)\wedge(\ell-1)}(\ell,m), \xi_1^{k-1}(\ell,m)\}$$
$$\le \begin{cases} \ell^{-\gamma}m^{\gamma-1} & \text{for } \gamma < \frac{1}{2}, \\ \log(\ell)m^{-\frac{1}{2}}\ell^{-\frac{1}{2}} & \text{for } \gamma = \frac{1}{2}, \\ m^{\gamma-1}\ell^{\gamma-1} & \text{for } \gamma > \frac{1}{2}, \end{cases}$$

with (5.10). Hence

$$T_{1,1} - T_{2,1} \le 3 \cdot \sum_{k=1}^{(i-1)\wedge(\ell-1)} \psi(m,\ell,k) \cdot \mathbb{P}(\ell \overset{i}{\not\to} k | \ell \to k)\mathbb{P}(\ell \to k)\mathbb{P}\Big(\bigcap_{r=\ell+1}^{n} \{r \overset{i}{\not\to} \ell\}\Big),$$

so that we have to consider $\mathbb{P}\Big(\bigcap_{r=\ell+1}^{n} \{r \overset{i}{\not\to} \ell\}\Big)$. We get

$$\begin{aligned}
\mathbb{P}\Big(\bigcap_{r=\ell+1}^{n} \{r \overset{i}{\not\to} \ell\}\Big) &= \mathbb{P}\Big(\bigcap_{r=\ell+1}^{n} \{r \overset{i}{\not\to} \ell\}\Big| \bigcap_{r=\ell+1}^{n} \{r \not\to \ell\}\Big)\mathbb{P}\Big(\bigcap_{r=\ell+1}^{n} \{r \not\to \ell\}\Big) \\
&\quad + \sum_{k=\ell+1}^{n} \mathbb{P}\Big(\bigcap_{r=\ell+1}^{n} \{r \overset{i}{\not\to} \ell\}\Big| \ell \overset{1.}{\leftrightarrow} k\Big)\mathbb{P}(\ell \overset{1.}{\leftrightarrow} k) \\
&\le \left(\frac{\ell}{n}\right)^{\eta} + \sum_{k=\ell+1}^{n} \mathbb{P}(k \overset{i}{\not\to} \ell | k \to \ell)\mathbb{P}(\ell \overset{1.}{\leftrightarrow} k) \\
&\le \left(\frac{\ell}{n}\right)^{\eta} + \sum_{k=\ell+1}^{n} \ell^{-\gamma} i^{\gamma-1} \left(\frac{\ell}{k}\right)^{\eta} \frac{f(0)}{k} \\
&\le \left(\frac{\ell}{n}\right)^{\eta} + \ell^{-\gamma} i^{\gamma-1} \le 2\max\Big\{\left(\frac{\ell}{n}\right)^{\eta}, \ell^{-\gamma} i^{\gamma-1}\Big\}
\end{aligned}$$

so that by (5.6)

$$T_{1,1} - T_{2,1} \le C\max\Big\{\left(\frac{\ell}{n}\right)^{\eta}, \ell^{-\gamma} i^{\gamma-1}\Big\} \sum_{k=1}^{(i-1)\wedge(\ell-1)} i^{\gamma-1}k^{-2\gamma}\ell^{\gamma-1} \cdot \psi(m,\ell,k)$$

$$\le C\max\left\{\left(\frac{\ell}{n}\right)^{\eta}, \frac{i^{\gamma-1}}{\ell^{\gamma}}\right\} \begin{cases} \ell^{-1}m^{\gamma-1}i^{\gamma-1}((i-1)\wedge(\ell-1))^{1-2\gamma} & \text{for } \gamma < \frac{1}{2}, \\ \log(n)\log(\ell)m^{-\frac{1}{2}}\ell^{-1}i^{-\frac{1}{2}} & \text{for } \gamma = \frac{1}{2}, \\ m^{\gamma-1}\ell^{2\gamma-2}i^{\gamma-1} & \text{for } \gamma > \frac{1}{2}. \end{cases} \quad (5.27)$$

We now proceed in a similar way to bound $T_{1,2} - T_{2,2}$. For $k \ge \ell + 1$ we get

$$
\begin{aligned}
\mathbb{P}(Z_{n,\ell}^{(i)} = 1 | Z_{n,m}^{(j)} = 1) &= \sum_{k=\ell+1}^{n} \mathbb{P}(Z_{n,\ell}^{(i)} = 1 | Z_{n,m}^{(j)} = 1, \{k \overset{1}{\leftrightarrow} \ell\}) \mathbb{P}(k \overset{1}{\leftrightarrow} \ell | Z_{n,m}^{(j)} = 1) \\
&\le \sum_{k=\ell+1}^{n} \mathbb{P}(k \overset{i}{\not\to} \ell | k \to \ell) \mathbb{P}\Big(\bigcap_{r=k+1}^{n} \{r \overset{i}{\not\to} \ell\} \Big| \deg_k^-(\ell) = 1, \{m \overset{j}{\not\to} \ell\} \Big) \\
&\quad \cdot \prod_{r=1}^{\ell-1} \mathbb{P}(\ell \nrightarrow r | m \overset{j}{\not\to} r) \mathbb{P}\Big(\bigcap_{r=\ell+1}^{k-1} \{r \nrightarrow \ell\} \Big| m \overset{j}{\not\to} \ell \Big) \mathbb{P}(k \to \ell | \deg_{k-1}^-(\ell) = 0)
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbb{P}(Z_{n,\ell}^{(i)} = 1) &= \sum_{k=\ell+1}^{n} \mathbb{P}(Z_{n,\ell}^{(i)} = 1 | \{k \overset{1}{\leftrightarrow} \ell\}) \mathbb{P}(k \overset{1}{\leftrightarrow} \ell) \\
&= \sum_{k=\ell+1}^{n} \mathbb{P}(k \overset{i}{\not\to} \ell | k \overset{1}{\leftrightarrow} \ell) \mathbb{P}\Big(\bigcap_{r=k+1}^{n} \{r \overset{i}{\not\to} \ell\} \Big| \deg_k^-(\ell) = 1 \Big) \\
&\quad \cdot \prod_{r=1}^{\ell-1} \mathbb{P}(\ell \nrightarrow r) \mathbb{P}\Big(\bigcap_{r=\ell+1}^{k-1} \{r \nrightarrow \ell\} \Big) \mathbb{P}(k \to \ell | \deg_{k-1}^-(\ell) = 0).
\end{aligned}
$$

Again, the summands differ in three of the factors, which are

$$
\mathbb{P}\Big(\bigcap_{r=k+1}^{n} \{r \overset{i}{\not\to} \ell\} \Big| \deg_k^-(\ell) = 1, \{m \overset{j}{\not\to} \ell\} \Big), \prod_{r=1}^{\ell-1} \mathbb{P}(\ell \nrightarrow r | m \overset{j}{\not\to} r),
$$
$$
\mathbb{P}\Big(\bigcap_{r=\ell+1}^{k-1} \{r \nrightarrow \ell\} \Big| m \overset{j}{\not\to} \ell \Big)
$$

and

$$
\mathbb{P}\Big(\bigcap_{r=k+1}^{n} \{r \overset{i}{\not\to} \ell\} \Big| \deg_k^-(\ell) = 1 \Big), \ \prod_{r=1}^{\ell-1} \mathbb{P}(\ell \nrightarrow r), \ \mathbb{P}\Big(\bigcap_{r=\ell+1}^{k-1} \{r \nrightarrow \ell\} \Big)
$$

respectively. By Proposition 5.4 we have

$$
\begin{aligned}
\prod_{r=1}^{\ell-1} \mathbb{P}(\ell \nrightarrow r | m \overset{j}{\not\to} r) - \prod_{r=1}^{\ell-1} \mathbb{P}(\ell \nrightarrow r) &\le \prod_{r=1}^{\ell-1} \mathbb{P}(\ell \nrightarrow r | m \nrightarrow r) - \prod_{r=1}^{\ell-1} \mathbb{P}(\ell \nrightarrow r) \\
&\le \xi_1^{\ell-1}(\ell, m) \le \begin{cases} \ell^{-\gamma} m^{\gamma-1} & \text{for } \gamma < \frac{1}{2}, \\ \ell^{\gamma-1} m^{\gamma-1} \log(\ell) & \text{for } \gamma = \frac{1}{2}, \\ \ell^{\gamma-1} m^{\gamma-1} & \text{for } \gamma > \frac{1}{2}. \end{cases}
\end{aligned}
$$

Analogous to the case $k \le \ell - 1$ we obtain

$$\mathbb{P}\Big(\bigcap_{r=\ell+1}^{k-1}\{r\not\to\ell\}\Big|m\overset{j}{\not\to}\ell\Big)-\mathbb{P}\Big(\bigcap_{r=\ell+1}^{k-1}\{r\not\to\ell\}\Big)\le\mathbb{P}\Big(\bigcap_{r=\ell+1}^{k-1}\{r\not\to\ell\}\Big)\left(\frac{1}{\mathbb{P}(m\overset{j}{\not\to}\ell)}-1\right)$$
$$\le\mathbb{P}\Big(\bigcap_{r=\ell+1}^{k-1}\{r\not\to\ell\}\Big)m^{\gamma-1}\ell^{-\gamma}.$$

In almost exactly the same way as in the previous case we can show

$$\begin{aligned}
&\mathbb{P}\Big(\bigcap_{r=k+1}^{n}\{r\overset{i}{\not\to}\ell\}\Big|\deg_k^-(\ell)=1,\{m\overset{j}{\not\to}\ell\}\Big)-\mathbb{P}\Big(\bigcap_{r=k+1}^{n}\{r\overset{i}{\not\to}\ell\}\Big|\deg_k^-(\ell)=1\Big)\\
&\le\mathbb{P}\Big(\bigcap_{r=k+1}^{n}\{r\overset{i}{\not\to}\ell\}\Big|\deg_k^-(\ell)=1,\{m\not\to\ell\}\Big)-\mathbb{P}\Big(\bigcap_{r=k+1}^{n}\{r\overset{i}{\not\to}\ell\}\Big|\deg_k^-(\ell)=1\Big)\\
&\le\mathbb{P}\Big(\bigcap_{r=k+1}^{n}\{r\overset{i}{\not\to}\ell\}\Big|\deg_k^-(\ell)=1\Big)\left(\frac{1}{\mathbb{P}(m\not\to\ell|\deg_k^-(\ell)=1)}-1\right)\\
&\le\mathbb{P}\Big(\bigcap_{r=k+1}^{n}\{r\overset{i}{\not\to}\ell\}\Big|\deg_k^-(\ell)=1\Big)m^{\gamma-1}\ell^{-\gamma},
\end{aligned}$$

because

$$\begin{aligned}
\mathbb{P}(m\not\to\ell|\deg_k^-(\ell)=1)&=1-\frac{\mathbb{E}\left[f(\deg_{m-1}^-(\ell))|\deg_k^-(\ell)=1\right]}{m}\\
&\ge1-\frac{\mathbb{E}\left[f(\deg_{m-1}^-(\ell))|\deg_{\ell+1}^-(\ell)=1\right]}{m}\\
&\ge1-\frac{f(1)m^\gamma\ell^{-\gamma}}{m}
\end{aligned}$$

by Lemma 3.21, so that

$$\begin{aligned}
\frac{1}{\mathbb{P}(m\not\to\ell|\deg_k^-(\ell)=1)}-1&\le\frac{m-(m-f(1)m^\gamma\ell^{-\gamma})}{m-f(1)m^\gamma\ell^{-\gamma}}=\frac{f(1)m^\gamma\ell^{-\gamma}}{m(1-f(1)m^{\gamma-1}\ell^{-\gamma})}\\
&\le C\ \frac{m^\gamma\ell^{-\gamma}}{m}=C\ m^{\gamma-1}\ell^{-\gamma},
\end{aligned}$$

where we used that $f(1)\le f(0)+1<2$ for all attachment functions considered in this chapter, so that $f(1)m^{\gamma-1}\ell^{-\gamma}<1$ for all $m>\ell\ge1$. Similar to the case $k\le(i-1)\wedge(\ell-1)$ we get

$$\begin{aligned}
T_{1,2}&=\sum_{k=\ell+1}^{n}\mathbb{P}(Z_{n,\ell}^{(i)}=1|Z_{n,m}^{(j)}=1,\{k\overset{1}{\leftrightarrow}\ell\})\mathbb{P}(k\overset{1}{\leftrightarrow}\ell|Z_{n,m}^{(j)}=1)\\
&\le\sum_{k=\ell+1}^{n}\mathbb{P}(k\overset{i}{\not\to}\ell|k\to\ell)\mathbb{P}\Big(\bigcap_{r=k+1}^{n}\{r\overset{i}{\not\to}\ell\}|\{m\overset{j}{\not\to}\ell\},\deg_k^-(\ell)=1\Big)\prod_{r=1}^{\ell-1}\mathbb{P}(\ell\not\to r|m\overset{j}{\not\to}r)\\
&\quad\cdot\mathbb{P}\Big(\bigcap_{r=\ell+1}^{k-1}\{r\not\to\ell\}|m\overset{j}{\not\to}\ell\Big)\mathbb{P}(k\to\ell|\deg_k^-(\ell)=0)
\end{aligned}$$

$$\leq T_{2,2} + 3 \sum_{k=\ell+1}^{n} \max\{m^{\gamma-1}\ell^{-\gamma}, \xi_1^{\ell-1}(\ell,m)\} \, \mathbb{P}\Big(\bigcap_{r=\ell+1}^{k-1} \{r \nrightarrow \ell\} \Big)$$

$$\cdot \, \mathbb{P}(k \overset{i}{\nrightarrow} \ell | k \to \ell) \mathbb{P}(k \to \ell | \deg^-_{k-1}(\ell) = 0)$$

$$\leq T_{2,2} + C \sum_{k=\ell+1}^{n} \left(\frac{\ell}{k}\right)^{\eta} i^{\gamma-1}\ell^{-\gamma} \frac{f(0)}{k} \cdot \begin{cases} \ell^{-\gamma}m^{\gamma-1} & \text{for } \gamma < \frac{1}{2}, \\ \ell^{-\frac{1}{2}}m^{-\frac{1}{2}}\log(\ell) & \text{for } \gamma = \frac{1}{2}, \\ \ell^{\gamma-1}m^{\gamma-1} & \text{for } \gamma > \frac{1}{2}, \end{cases}$$

so that

$$T_{1,2} - T_{2,2} \leq C \sum_{k=\ell+1}^{n} \left(\frac{\ell}{k}\right)^{\eta} i^{\gamma-1}\ell^{-\gamma} \frac{f(0)}{k} \cdot \begin{cases} \ell^{-\gamma}m^{\gamma-1} & \text{for } \gamma < \frac{1}{2}, \\ \ell^{-\frac{1}{2}}m^{-\frac{1}{2}}\log(\ell) & \text{for } \gamma = \frac{1}{2}, \\ \ell^{\gamma-1}m^{\gamma-1} & \text{for } \gamma > \frac{1}{2} \end{cases}$$

$$\leq C \begin{cases} \ell^{-2\gamma}m^{\gamma-1}i^{\gamma-1} & \text{for } \gamma < \frac{1}{2}, \\ \ell^{-1}m^{-\frac{1}{2}}i^{-\frac{1}{2}}\log(\ell) & \text{for } \gamma = \frac{1}{2}, \\ \ell^{-1}m^{\gamma-1}i^{\gamma-1} & \text{for } \gamma > \frac{1}{2}. \end{cases}$$

Combining this result with (5.27) yields

$$\mathbb{P}(Z^{(i)}_{n,\ell} = 1 | Z^{(j)}_{n,m} = 1) - \mathbb{P}(Z^{(i)}_{n,\ell} = 1)$$

$$\leq C \max\left\{ \left(\frac{\ell}{n}\right)^{\eta}, \ell^{-\gamma} i^{\gamma-1} \right\} \begin{cases} \ell^{-2\gamma}m^{\gamma-1}i^{\gamma-1} & \text{for } \gamma < \frac{1}{2}, \\ \log(n)m^{-\frac{1}{2}}\ell^{-1}i^{-\frac{1}{2}} & \text{for } \gamma = \frac{1}{2}, \\ m^{\gamma-1}\ell^{2\gamma-2}i^{\gamma-1} & \text{for } \gamma > \frac{1}{2} \end{cases}$$

and by repeated use of (5.6) we can now calculate

$$\sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{\substack{\ell,m \\ \ell<m}} \vartheta_{i,n}\vartheta_{j,n} \operatorname{Cov}[Z^{(i)}_{n,\ell}, Z^{(j)}_{n,m}]$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{\substack{\ell,m \\ \ell<m}} \vartheta_{i,n}\vartheta_{j,n} \mathbb{P}(Z^{(j)}_{n,m} = 1) \Big(\mathbb{P}(Z^{(i)}_{n,\ell} = 1 | Z^{(j)}_{n,m} = 1) - \mathbb{P}(Z^{(i)}_{n,\ell} = 1) \Big)$$

for the various constellations of i, j, m and ℓ. Before we begin recall (5.20) and (5.21), which we will use in all following cases. Also remember that

$$\vartheta_{i,n} \sim \left(\frac{i}{n}\right)^{\eta}$$

for all $i \leq n$. Starting with $\ell < m < j < i$ we get

$$
\begin{aligned}
&\sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{m=1}^{j-1}\sum_{\ell=1}^{m-1}\vartheta_{i,n}\vartheta_{j,n}\operatorname{Cov}[Z_{n,\ell}^{(i)},Z_{n,m}^{(j)}]\\
&=\sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{m=1}^{j-1}\sum_{\ell=1}^{m-1}\vartheta_{i,n}\vartheta_{j,n}\mathbb{P}(Z_{n,m}^{(j)}=1)\Big(\mathbb{P}(Z_{n,\ell}^{(i)}=1|Z_{n,m}^{(j)}=1)-\mathbb{P}(Z_{n,\ell}^{(i)}=1)\Big)\\
&\le\sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{m=1}^{j-1}\sum_{\ell=1}^{m-1}\begin{cases} j^{\gamma-1}m^{-\gamma}\cdot\ell^{-2\gamma}m^{\gamma-1}i^{\gamma-1} & \text{for } \gamma<\frac{1}{2},\\ \max\Big\{\log(\ell)\left(\frac{\ell}{n}\right)^{\eta},1\Big\}m^{-\frac{1}{2}}j^{-\frac{1}{2}}\log(m)\log(n)m^{-\frac{1}{2}}\ell^{-1}i^{-\frac{1}{2}} & \text{for } \gamma=\frac{1}{2},\\ m^{\gamma-1}j^{\gamma-1}\cdot m^{\gamma-1}\ell^{2\gamma-2}i^{\gamma-1} & \text{for } \gamma>\frac{1}{2},\end{cases}\\
&\le C\begin{cases}\sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{m=1}^{j-1}j^{\gamma-1}m^{-2\gamma}i^{\gamma-1} & \text{for } \gamma<\frac{1}{2},\\ \log(n)^{3}\sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{m=1}^{j-1}m^{-1}j^{-\frac{1}{2}}i^{-\frac{1}{2}} & \text{for } \gamma=\frac{1}{2},\\ \sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{m=1}^{j-1}m^{4\gamma-3}j^{\gamma-1}i^{\gamma-1} & \text{for } \gamma>\frac{1}{2},\end{cases}\\
&\le C\begin{cases}\sum_{i=1}^{n}\sum_{j=1}^{i-1}j^{-\gamma}i^{\gamma-1} & \text{for } \gamma<\frac{1}{2},\\ \log(n)^{4}\sum_{i=1}^{n}\sum_{j=1}^{i-1}i^{-\frac{1}{2}}j^{-\frac{1}{2}} & \text{for } \gamma=\frac{1}{2},\\ \sum_{i=1}^{n}\sum_{j=1}^{i-1}j^{5\gamma-3}i^{\gamma-1} & \text{for } \gamma>\frac{1}{2},\end{cases}\\
&\le C\begin{cases} n & \text{for } \gamma<\frac{1}{2},\\ n\log(n)^{4} & \text{for } \gamma=\frac{1}{2},\\ n^{6\gamma-2} & \text{for } \gamma>\frac{1}{2}.\end{cases}
\end{aligned}
$$

We will omit the calculations in the remaining five cases, however the procedure is always exactly the same. For $j<\ell<i<m$ we get

$$
\sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{\ell=j+1}^{i-1}\sum_{m=i+1}^{i-1}\vartheta_{i,n}\vartheta_{j,n}\operatorname{Cov}[Z_{n,\ell}^{(i)},Z_{n,m}^{(j)}]\le C\begin{cases} n & \text{for } \gamma<\frac{1}{2},\\ n\log(n)^{4} & \text{for } \gamma=\frac{1}{2},\\ n^{6\gamma-2} & \text{for } \gamma>\frac{1}{2}\end{cases}
$$

and

$$
\sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{\ell=j+1}^{i-1}\sum_{m=\ell+1}^{i-1}\vartheta_{i,n}\vartheta_{j,n}\operatorname{Cov}[Z_{n,\ell}^{(i)},Z_{n,m}^{(j)}]\le C\begin{cases} n & \text{for } \gamma<\frac{1}{2},\\ n\log(n)^{4} & \text{for } \gamma=\frac{1}{2},\\ n^{6\gamma-2} & \text{for } \gamma>\frac{1}{2}\end{cases}
$$

for $j < \ell < m < i$, where we used $\vartheta_{j,n} \sim \left(\frac{j}{n}\right)^\eta \le \left(\frac{\ell}{n}\right)^\eta$ in both cases. Similarly with $\vartheta_{j,n} \sim \left(\frac{j}{n}\right)^\eta \le \left(\frac{m}{n}\right)^\eta$ in the case $\ell < j < m < i$ we obtain

$$\sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{m=j+1}^{i-1}\sum_{\ell=1}^{j-1} \vartheta_{i,n}\vartheta_{j,n} \operatorname{Cov}[Z_{n,\ell}^{(i)}, Z_{n,m}^{(j)}] \le C \begin{cases} n & \text{for } \gamma < \frac{1}{2}, \\ n\log(n)^4 & \text{for } \gamma = \frac{1}{2}, \\ n^{4\gamma-1} & \text{for } \gamma > \frac{1}{2} \end{cases}$$

and

$$\sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{\ell=i+1}^{n}\sum_{m=\ell+1}^{n} \vartheta_{i,n}\vartheta_{j,n} \operatorname{Cov}[Z_{n,\ell}^{(i)}, Z_{n,m}^{(j)}] \le C \begin{cases} n & \text{for } \gamma < \frac{1}{2}, \\ n\log(n)^3 & \text{for } \gamma = \frac{1}{2}, \\ n^{6\gamma-2} & \text{for } \gamma > \frac{1}{2} \end{cases}$$

for $j < i < \ell < m$ and as $\vartheta_{i,n}\vartheta_{j,n} \le \frac{m^\eta \ell^\eta}{n^{2\eta}}$ in this case. The last case to consider is $\ell < j < i < m$, for which we obtain

$$\sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{\ell=1}^{j-1}\sum_{m=i+1}^{n} \vartheta_{i,n}\vartheta_{j,n} \operatorname{Cov}[Z_{n,\ell}^{(i)}, Z_{n,m}^{(j)}] \le C \begin{cases} n & \text{for } \gamma < \frac{1}{2}, \\ n\log(n)^4 & \text{for } \gamma = \frac{1}{2}, \\ n^{4\gamma-1} & \text{for } \gamma > \frac{1}{2}. \end{cases}$$

Hence

$$\sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{m=1}^{n}\sum_{\ell=1}^{m-1} \vartheta_{i,n}\vartheta_{j,n} \operatorname{Cov}[Z_{n,\ell}^{(i)}, Z_{n,m}^{(j)}] \le C \begin{cases} n & \text{for } \gamma < \frac{1}{2}, \\ n\log(n)^4 & \text{for } \gamma = \frac{1}{2}, \\ n^{6\gamma-2} & \text{for } \gamma > \frac{1}{2}. \end{cases}$$

All these calculations can be conducted in the exact same way for $m < \ell$ if we swap the roles of m and ℓ, i.e. we look at

$$\begin{aligned} &\sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{\ell=1}^{n}\sum_{m=1}^{\ell-1} \vartheta_{i,n}\vartheta_{j,n} \operatorname{Cov}[Z_{n,\ell}^{(i)}, Z_{n,m}^{(j)}] \\ &= \sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{\ell=1}^{n}\sum_{m=1}^{\ell-1} \vartheta_{i,n}\vartheta_{j,n} \mathbb{P}(Z_{n,\ell}^{(i)} = 1) \left(\mathbb{P}(Z_{n,m}^{(j)} = 1 | Z_{n,\ell}^{(i)} = 1) - \mathbb{P}(Z_{n,m}^{(j)} = 1)\right). \end{aligned}$$

Finally we obtain

$$\sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{m=1}^{n}\sum_{\ell=1}^{n} \vartheta_{i,n}\vartheta_{j,n} \operatorname{Cov}[Z_{n,\ell}^{(i)}, Z_{n,m}^{(j)}] \le C \begin{cases} n & \text{for } \gamma < \frac{1}{2}, \\ n\log(n)^4 & \text{for } \gamma = \frac{1}{2}, \\ n^{6\gamma-2} & \text{for } \gamma > \frac{1}{2}. \end{cases}$$

□

We can now combine Lemma 5.5, Lemma 5.7 and Lemma 5.8 to bound the first term appearing in (2.4).

Lemma 5.9. *For W_n^s having the size-bias distribution of W_n there exists a constant $C > 0$ independent of n such that*

$$\mathbb{V}[\mathbb{E}\left[W_n^s - W_n|\mathcal{G}_n\right]] \leq \left(\frac{2\sigma_n}{\mu_n}\right)^2 + \frac{C}{\mu_n^2}\begin{cases} n & \text{for } \gamma < \frac{1}{2}, \\ n\log(n)^4 & \text{for } \gamma = \frac{1}{2}, \\ n^{6\gamma-2} & \text{for } \gamma > \frac{1}{2}. \end{cases}$$

Proof. Remember that by (5.14) we have

$$\mathbb{V}[\mathbb{E}\left[W_n^s - W_n|\mathcal{G}_n\right]] \leq \frac{4}{\mu_n^2}\left(\sigma_n^2 + \mathbb{V}\left[\sum_{i=1}^n \vartheta_{i,n} D_{n,i}\right] + \mathbb{V}\left[\sum_{i=1}^n \vartheta_{i,n} R_{n,i}\right]\right).$$

Lemma 5.5 gives

$$\mathbb{V}\left[\sum_{i=1}^n \vartheta_{i,n} D_{n,i}\right] \leq C\begin{cases} n & \text{for } \gamma < \frac{1}{2}, \\ \log(n) n & \text{for } \gamma = \frac{1}{2}, \\ n^{2\gamma} & \text{for } \gamma > \frac{1}{2}, \end{cases}$$

whereas Lemma 5.7 and Lemma 5.8 yield

$$\mathbb{V}\left[\sum_{i=1}^n \vartheta_{i,n} R_{n,i}\right] \leq C\begin{cases} n & \text{for } \gamma < \frac{1}{2}, \\ n\log(n)^4 & \text{for } \gamma = \frac{1}{2}, \\ n^{6\gamma-2} & \text{for } \gamma > \frac{1}{2}, \end{cases}$$

which proves the claim. □

The next lemma now gives an upper bound on the second term in (2.4).

Lemma 5.10. *For W_n denoting the number of isolated vertice in a preferential attachment graph $\mathcal{G}_n$ and W_n^s having the size-bias distribution of W_n, there exists a constant C independent of n such that*

$$\mathbb{E}\left[(W_n^s - W_n)^2\right] \leq \frac{C}{\mu_n}\begin{cases} n & \text{for } \gamma < \frac{1}{2}, \\ n\log(n)^2 & \text{for } \gamma = \frac{1}{2}, \\ n^{4\gamma-1} & \text{for } \gamma > \frac{1}{2}. \end{cases}$$

Proof. First remember that to construct a graph for which the number of isolated vertices has the size-bias distribution of W_n, we choose one vertex I according to $\mathbb{P}(I = i) = \frac{\vartheta_{i,n}}{\mu_n}$. Furthermore, recall that $D_{i,n}$ denotes the number of neighbours of vertex i, which are only connected to i and d_i refers to the total degree of vertex i.

Consequently we have $\sum_{i=1}^n D_{i,n} \le n$. By conditioning on the graph $\mathcal{G}_n$ at time n and using (5.13), we obtain

$$\begin{aligned}
\mathbb{E}\left[(W_n^s - W_n)^2\right] &= \mathbb{E}\left[\frac{1}{\mu_n}\sum_{i=1}^n \vartheta_{i,n}\left(D_{i,n} + \mathbb{1}\{d_i > 0\} + R_{n,i}\right)^2\right] \\
&= \frac{1}{\mu_n}\left(\sum_{i=1}^n \vartheta_{i,n}\mathbb{E}\left[D_{i,n}^2\right] + 2\sum_{i=1}^n \vartheta_{i,n}\mathbb{E}\left[D_{i,n}\mathbb{1}\{d_i > 0\}\right] + \sum_{i=1}^n \vartheta_{i,n}\mathbb{E}\left[\mathbb{1}\{d_i > 0\}\right]\right. \\
&\qquad \left. + 2\sum_{i=1}^n \vartheta_{i,n}\mathbb{E}\left[D_{i,n}R_{n,i}\right] + 2\sum_{i=1}^n \vartheta_{i,n}\mathbb{E}\left[\mathbb{1}\{d_i > 0\}R_{n,i}\right] + \sum_{i=1}^n \vartheta_{i,n}\mathbb{E}\left[R_{n,i}^2\right]\right) \\
&\le \frac{1}{\mu_n}\left(\sum_{i=1}^n \mathbb{E}\left[D_{i,n}^2\right] + 2\sum_{i=1}^n \vartheta_{i,n}\mathbb{E}\left[D_{i,n}R_{n,i}\right] + 2\sum_{i=1}^n \vartheta_{i,n}\mathbb{E}\left[\mathbb{1}\{d_i > 0\}R_{n,i}\right]\right. \\
&\qquad \left. + \sum_{i=1}^n \vartheta_{i,n}\mathbb{E}\left[R_{n,i}^2\right]\right) + \frac{3n}{\mu_n}
\end{aligned}$$

and since $\mu_n \sim n$ according to Lemma 5.3 the last term is of constant order. To bound the remaining sums we define

$$\begin{aligned}
S_{ij}^{(n)} &:= \mathbb{1}\{i \text{ is the only neighbour of } j \text{ in } \mathcal{G}_n\} \\
&= \begin{cases} \bigcap_{\substack{r=1 \\ r\neq i}}^{j-1}\{j \nrightarrow r\} \cap \{j \to i\} \cap \bigcap_{r=j+1}^{n}\{r \nrightarrow j\} & \text{for } i < j, \\[2ex] \bigcap_{r=1}^{j-1}\{j \nrightarrow r\} \cap \{i \to j\} \cap \bigcap_{\substack{r=j+1 \\ r\neq i}}^{n}\{r \nrightarrow j\} & \text{for } i > j. \end{cases}
\end{aligned}$$

Thus $D_{i,n} = \sum_{j=1}^n S_{ij}^{(n)}$, where by construction of the network $S_{ii}^{(n)} = 0$. With these notations we get

$$D_{i,n}^2 = \sum_{j=1}^n S_{ij}^{(n)} + 2\sum_{k=1}^n \sum_{j=1}^{k-1} S_{ij}^{(n)} S_{ik}^{(n)}.$$

To calculate $\mathbb{E}\left[S_{ij}^{(n)}\right]$ we distinguish the two cases $i < j$ and $i > j$. For the latter we have

$$\begin{aligned}
\mathbb{E}\left[S_{ij}^{(n)}\right] &= \mathbb{P}\left(\bigcap_{k=1}^{j-1}\{j \nrightarrow k\}, \bigcap_{\substack{\ell=j+1 \\ \ell\neq i}}^{n}\{\ell \nrightarrow j\}, \{i \to j\}\right) \\
&= p_{j,0}\prod_{k=j+1}^{i-1}\left(1 - \frac{f(0)}{k}\right)\frac{f(0)}{i}\prod_{k=i+1}^{n}\left(1 - \frac{f(1)}{k}\right) \\
&\le c_2\left(\frac{j}{i-1}\right)^\eta \frac{f(0)}{i}\left(\frac{i}{n}\right)^\eta \le C\,\frac{j^\eta}{i n^\eta},
\end{aligned} \tag{5.28}$$

by the independence of in- and outdegree of a fixed vertex. For $i < j$ we get

$$\begin{aligned}\mathbb{E}\left[S_{ij}^{(n)}\right] &= \mathbb{P}\left(\bigcap_{k=1,k\neq i}^{j-1}\{j \nrightarrow k\}, \{j \to i\}, \bigcap_{\ell=j+1}^{n}\{\ell \nrightarrow j\}\right) \\ &= \prod_{k=1,k\neq i}^{j-1}\left(1 - \frac{\mu_{j-1}^{f}(k)}{j}\right)\frac{\mu_{j-1}^{f}(i)}{j}\prod_{\ell=j+1}^{n}\left(1 - \frac{f(0)}{\ell}\right) \\ &\leq \frac{1}{j}\left(\frac{j}{i}\right)^{\gamma}\left(\frac{j}{n}\right)^{\eta} = \frac{j^{-1+\gamma+\eta}}{i^{\gamma}n^{\eta}},\end{aligned} \tag{5.29}$$

where we used (5.5) in both calculations. Consequently

$$\begin{aligned}\frac{1}{\mu_n}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathbb{E}\left[S_{ij}^{(n)}\right] &= \frac{1}{\mu_n}\left(\sum_{i=1}^{n}\sum_{j=1}^{i-1}\mathbb{E}\left[S_{ij}^{(n)}\right] + \sum_{i=1}^{n}\sum_{j=i+1}^{n}\mathbb{E}\left[S_{ij}^{(n)}\right]\right) \\ &\leq \frac{C}{\mu_n}\left(\sum_{i=1}^{n}\sum_{j=1}^{i-1}\frac{j^{\eta}}{in^{\eta}} + \sum_{i=1}^{n}\sum_{j=i+1}^{n}\frac{j^{\gamma+\eta-1}}{i^{\gamma}n^{\eta}}\right) \\ &\leq 2\frac{C}{\mu_n}\frac{n^{1+\eta}}{n^{\eta}},\end{aligned}$$

where we used bounds of the form (5.6). As μ_n is of order n we obtain

$$\frac{1}{\mu_n}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathbb{E}\left[S_{ij}^{(n)}\right] \leq C \tag{5.30}$$

for some constant $C > 0$ independent of n. In order to bound $\mathbb{E}\left[S_{ij}^{(n)}S_{ik}^{(n)}\right]$ we consider the three cases $i < j < k$, $j < i < k$ and $j < k < i$. Since

$$\mathbb{E}\left[S_{ij}^{(n)}S_{ik}^{(n)}\right] = \mathbb{P}(S_{ij}^{(n)}S_{ik}^{(n)} = 1)$$

Thus, for $i < j < k$ and using the multiplication formula for conditional probabilities

$$\begin{aligned}&\mathbb{P}(S_{ij}^{(n)}S_{ik}^{(n)} = 1) \\ &= \mathbb{P}\Bigg(\bigcap_{m=k+1}^{n}\{m \nrightarrow k\}, \bigcap_{\ell=j+1}^{n}\{\ell \nrightarrow j\}, \bigcap_{\substack{m=1,\\ m\neq i,j}}^{k-1}\{k \nrightarrow m\}, \bigcap_{\substack{\ell=1,\\ \ell\neq i}}^{j-1}\{j \nrightarrow \ell\}, \{k \to i\}, \{j \to i\}\Bigg) \\ &= \mathbb{P}\Big(\bigcap_{m=k+1}^{n}\{m \nrightarrow k\}\Big)\mathbb{P}\Big(\bigcap_{\ell=j+1}^{n}\{\ell \nrightarrow j\}\Big)\mathbb{P}\Big(\bigcap_{\substack{m=1,\\ m\neq i,j}}^{k-1}\{k \nrightarrow m\}\Big|\bigcap_{\substack{\ell=1,\\ \ell\neq i}}^{j}\{j \nrightarrow \ell\}\Big)\end{aligned}$$

$$
\begin{aligned}
&\quad \cdot \mathbb{P}\Big(\bigcap_{\substack{\ell=1,\\ \ell\neq i}}^{j-1} \{j \nrightarrow \ell\}\Big) \mathbb{P}\left(k \to i | j \to i\right) \mathbb{P}\left(j \to i\right) \\
&\leq \prod_{m=k+1}^{n} \left(1 - \frac{f(0)}{m}\right) \prod_{\ell=j+1}^{n} \left(1 - \frac{f(0)}{\ell}\right) \mathbb{P}\left(k \to i | j \to i\right) \cdot \frac{\mu^f_{j-1}(i)}{j} \\
&\leq C \left(\frac{k}{n}\right)^{\eta} \left(\frac{j}{n}\right)^{\eta} \frac{k^{\gamma-1}}{i^{\gamma}} \frac{j^{\gamma-1}}{i^{\gamma}} = C\, k^{\eta+\gamma-1} j^{\eta+\gamma-1} n^{-2\eta} i^{-2\gamma}
\end{aligned}
$$

where we used [DM13, Lemma 2.10] and Lemma 3.21 to deduce that

$$
\mathbb{P}(k \to i | j \to i) = \frac{\hat{\mu}^f_{k-1}(i,j)}{k} \leq \frac{\hat{\mu}^f_{k-1}(i,i+1)}{k} \leq f(1) k^{\gamma-1} i^{-\gamma}
$$

and also exploited the dependency structure of the network. Similarly, for $j < i < k$,

$$
\begin{aligned}
&\mathbb{P}(S^{(n)}_{ij} S^{(n)}_{ik} = 1) \\
&= \mathbb{P}\Big(\bigcap_{m=k+1}^{n} \{m \nrightarrow k\}, \bigcap_{\substack{\ell=j+1,\\ \ell\neq i}}^{n} \{\ell \nrightarrow j\}, \bigcap_{\substack{m=1,\\ m\neq i,j}}^{k-1} \{k \nrightarrow m\}, \bigcap_{\ell=1}^{j-1} \{j \nrightarrow \ell\}, \{k \to i\}, \{i \to j\}\Big) \\
&= \mathbb{P}\Big(\bigcap_{m=k+1}^{n} \{m \nrightarrow k\}\Big) \mathbb{P}\Big(\bigcap_{\substack{\ell=j+1,\\ \ell\neq i}}^{n} \{\ell \nrightarrow j\} \Big| \{i \to j\}\Big) \mathbb{P}\Big(\bigcap_{\substack{m=1,\\ m\neq i,j}}^{k-1} \{k \nrightarrow m\} \Big| \bigcap_{\ell=1}^{j} \{j \nrightarrow \ell\} \Big) \\
&\quad \cdot \mathbb{P}\Big(\bigcap_{\substack{\ell=1,\\ \ell\neq i}}^{j-1} \{j \nrightarrow \ell\}\Big) \mathbb{P}\Big(k \to i\Big) \mathbb{P}\left(i \to j\right) \\
&\leq \prod_{m=k+1}^{n} \left(1 - \frac{f(0)}{m}\right) \prod_{\ell=j+1}^{i-1} \left(1 - \frac{f(0)}{j}\right) \prod_{\ell=i+1}^{n} \left(1 - \frac{f(1)}{j}\right) \frac{\mu^f_{k-1}(i)}{k} \frac{\mu^f_{i-1}(j)}{i} \\
&\leq C \left(\frac{k}{n}\right)^{\eta} \left(\frac{j}{n}\right)^{\eta} \frac{k^{\gamma}}{i^{\gamma} k} \frac{i^{\gamma}}{i j^{\gamma}} = C\, k^{-1+\eta+\gamma} j^{\eta-\gamma} n^{-2\eta} i^{-1}.
\end{aligned}
$$

In the last case, $j < k < i$, we obtain

$$
\begin{aligned}
&\mathbb{P}(S^{(n)}_{ij} S^{(n)}_{ik} = 1) \\
&= \mathbb{P}\Big(\bigcap_{\substack{m=k+1,\\ m\neq i}}^{n} \{m \nrightarrow k\}, \bigcap_{\substack{\ell=j+1,\\ \ell\neq i}}^{n} \{\ell \nrightarrow j\}, \bigcap_{\ell=1}^{j-1} \{j \nrightarrow \ell\}, \bigcap_{\substack{m=1,\\ m\neq j}}^{k-1} \{k \nrightarrow m\}, \{i \to k\}, \{i \to j\}\Big) \\
&= \mathbb{P}\Big(\bigcap_{\substack{m=k+1,\\ m\neq i}}^{n} \{m \nrightarrow k\} \Big| \{i \to k\}\Big) \mathbb{P}\Big(\bigcap_{\substack{\ell=j+1,\\ \ell\neq i}}^{n} \{\ell \nrightarrow j\} | \{i \to j\}\Big) \mathbb{P}\Big(\bigcap_{\substack{\ell=1,\\ \ell\neq i}}^{j-1} \{j \nrightarrow \ell\} \Big| \bigcap_{\substack{m=1,\\ m\neq i,j}}^{k-1} \{k \nrightarrow m\}\Big)
\end{aligned}
$$

$$\cdot \mathbb{P}\Big(\bigcap_{\substack{m=1,\\ m\neq i,j}}^{k-1} \{k \nrightarrow m\} \Big| \{i \to j\} \Big) \mathbb{P}(\{i \to k\}) \, \mathbb{P}(\{i \to j\})$$

$$\leq \prod_{m=k+1}^{i-1} \left(1 - \frac{f(0)}{m}\right) \prod_{m=i+1}^{n} \left(1 - \frac{f(1)}{m}\right) \prod_{\ell=j+1}^{i-1} \left(1 - \frac{f(0)}{\ell}\right) \prod_{\ell=i+1}^{n} \left(1 - \frac{f(1)}{\ell}\right)$$

$$\cdot \frac{\mu^f_{i-1}(k)}{i} \frac{\mu^f_{i-1}(j)}{i}$$

$$\leq C \left(\frac{k}{n}\right)^{\eta} \left(\frac{j}{n}\right)^{\eta} \frac{1}{i^2} \left(\frac{i}{k}\right)^{\gamma} \left(\frac{i}{j}\right)^{\gamma} = C\, n^{-2\eta} k^{\eta-\gamma} j^{\eta-\gamma} i^{-2+2\gamma}.$$

By repeated use of (5.6) we obtain

$$\begin{aligned} \sum_{i=1}^{n} \sum_{k=1}^{n} \sum_{j=1}^{k-1} \mathbb{E}\left[S^{(n)}_{ij} S^{(n)}_{ik}\right] &\leq C \Bigg(\sum_{i=1}^{n} \sum_{k=i+1}^{n} \sum_{j=i+1}^{k-1} k^{\eta+\gamma-1} j^{\eta+\gamma-1} n^{-2\eta} i^{-2\gamma} \\ &\quad + \sum_{i=1}^{n} \sum_{k=i+1}^{n} \sum_{j=1}^{i-1} k^{-1+\eta+\gamma} j^{\eta-\gamma} n^{-2\eta} i^{-1} \\ &\quad + \sum_{i=1}^{n} \sum_{k=1}^{i-1} \sum_{j=1}^{k} n^{-2\eta} k^{\eta-\gamma} j^{\eta-\gamma} i^{-2+2\gamma} \Bigg) \\ &\leq \frac{C}{n^{2\eta}} \left(n^{2\eta+1} + n^{2\eta+1} + n^{2\eta+1}\right) = C\, n, \end{aligned} \tag{5.31}$$

Combining (5.30) and (5.31) yields

$$\mathbb{E}\left[D^2_{i,n}\right] = \sum_{j=1}^{n} \mathbb{E}\left[S^{(n)}_{ij}\right] + 2 \sum_{k=1}^{n} \sum_{j=1}^{k-1} \mathbb{E}\left[S^{(n)}_{ij} S^{(n)}_{ik}\right] \leq C\ n. \tag{5.32}$$

With the definitions of $D_{i,n}$ and $R_{n,i}$ as well as the fact that $Z^{(i)}_{n,m} = 0$ for $S^{(n)}_{i,m} = 1$ we get

$$\mathbb{E}\left[D_{i,n} R_{n,i}\right] = \mathbb{E}\left[\sum_{j=1}^{n} \sum_{m=1}^{n} S^{(n)}_{i,j} Z^{(i)}_{n,m}\right] = \sum_{j=1}^{n} \sum_{\substack{m=1\\ m\neq j}}^{n} \mathbb{P}\left(S^{(n)}_{i,j} = 1\right) \mathbb{P}\left(Z^{(i)}_{n,m} = 1 | S^{(n)}_{i,j} = 1\right)$$

$$\leq \frac{C}{n^{\eta}} \sum_{j=1}^{i-1} \sum_{\substack{m=1\\ m\neq j}}^{n} \frac{j^{\eta}}{i} \mathbb{P}\left(Z^{(i)}_{n,m} = 1 | S^{(n)}_{i,j} = 1\right) + \frac{C}{n^{\eta}} \sum_{j=i+1}^{n} \sum_{\substack{m=1\\ m\neq j}}^{n} \frac{j^{\eta+\gamma-1}}{i^{\gamma}} \mathbb{P}\left(Z^{(i)}_{n,m} = 1 | S^{(n)}_{i,j} = 1\right),$$

where we used (5.28) and (5.29). For the conditional probability we obtain

$$\mathbb{P}(Z^{(i)}_{n,m} = 1 | S^{(n)}_{i,j} = 1) = \mathbb{P}\Big(Z^{(i)}_{n,m} = 1 | \bigcap_{r=1}^{j-1} \{r \nrightarrow r\}, \bigcap_{\substack{r=j+1\\ s\neq i}}^{n} \{s \nrightarrow j\}, \{i \to j\}\Big)$$

$$
\begin{aligned}
&= \sum_{r=1}^{m-1} \mathbb{P}(Z_{n,m}^{(i)} = 1 | S_{i,j}^{(n)} = 1, \{m \overset{1.}{\leftrightarrow} r\}) \mathbb{P}(m \overset{1.}{\leftrightarrow} r | S_{i,j}^{(n)} = 1) \\
&\quad + \mathbb{1}\{m < i\} \sum_{r=m+1}^{n} \mathbb{P}(Z_{n,m}^{(i)} = 1 | S_{i,j}^{(n)} = 1, \{r \overset{1.}{\leftrightarrow} m\}) \mathbb{P}(m \overset{1.}{\leftrightarrow} r | S_{i,j}^{(n)} = 1) \\
&\leq \sum_{r=1}^{m-1} i^{\gamma-1} r^{-2\gamma} m^{\gamma-1} + \mathbb{1}\{m < i\} \sum_{r=m+1}^{n} i^{\gamma-1} r^{\gamma-1} m^{-2\gamma} \\
&\leq C \begin{cases} i^{\gamma-1} m^{-\gamma} + n^{\gamma} i^{\gamma-1} m^{-2\gamma} & \text{for } \gamma < \frac{1}{2}, \\ \log(m) i^{\gamma-1} m^{\gamma-1} + n^{\gamma} i^{\gamma-1} m^{-2\gamma} & \text{for } \gamma = \frac{1}{2}, \\ i^{\gamma-1} m^{\gamma-1} + n^{\gamma} i^{\gamma-1} m^{-2\gamma} & \text{for } \gamma > \frac{1}{2} \end{cases} \\
&:= \Psi^f(n, m, i, j),
\end{aligned}
$$

where we used that

$$
\mathbb{P}(m \to r | S_{i,j}^{(n)} = 1) \leq \mathbb{P}(m \to r)
$$

and

$$
\mathbb{P}(r \to m | S_{i,j}^{(n)} = 1) \leq \mathbb{P}(r \to m),
$$

with equality holding for $j < r$ in the first, and $j < m$ in the second case. Hence we get

$$
\begin{aligned}
&\sum_{i=1}^{n} \vartheta_{i,n} \mathbb{E}\left[D_{i,n} R_{n,i}\right] \\
&\leq \frac{C}{n^{2\eta}} \sum_{i=1}^{n} \sum_{j=1}^{i-1} \sum_{\substack{m=1 \\ m \neq j}}^{n} \frac{j^{\eta}}{i^{1-\eta}} \Psi^f(n, m, i, j) + \frac{C}{n^{2\eta}} \sum_{i=1}^{n} \sum_{j=i+1}^{n} \sum_{\substack{m=1 \\ m \neq j}}^{n} \frac{j^{\eta+\gamma-1}}{i^{\gamma-\eta}} \Psi^f(n, m, i, j) \\
&\leq C \begin{cases} n & \text{for } \gamma < \frac{1}{2}, \\ \log(n) n & \text{for } \gamma = \frac{1}{2}, \\ n^{2\gamma} & \text{for } \gamma > \frac{1}{2}, \end{cases} \qquad (5.33)
\end{aligned}
$$

where we used (5.6) three times to obtain the last inequality. Using (5.20) and (5.21) it is straightforward to bound $\mathbb{E}\left[\sum_{i=1}^{n} \mathbb{1}\{d_i > 0\} R_{n,i}\right]$. More precisely, we have

$$
\mathbb{E}\left[\sum_{i=1}^{n} \mathbb{1}\{d_i > 0\} R_{n,i}\right] \leq \mathbb{E}\left[\sum_{i=1}^{n} R_{n,i}\right] = \sum_{i=1}^{n} \left(\sum_{j=1}^{i-1} \mathbb{P}(Z_{n,j}^{(i)} = 1) + \sum_{j=i+1}^{n} \mathbb{P}(Z_{n,j}^{(i)} = 1) \right)
$$

$$\leq C \begin{cases} n & \text{for } \gamma < \frac{1}{2} \\ n\log(n)^2 & \text{for } \gamma = \frac{1}{2} \\ n^{2\gamma} & \text{for } \gamma > \frac{1}{2}. \end{cases}$$

It now remains to deal with $\mathbb{E}\left[\sum_{i=1}^{n} \vartheta_{i,n} R_{n,i}^2\right]$. We have

$$\mathbb{E}\left[\sum_{i=1}^{n} \vartheta_{i,n} R_{n,i}^2\right] = \mathbb{E}\left[\sum_{i=1}^{n}\sum_{j=1}^{n} \vartheta_{i,n} Z_{n,j}^{(i)}\right] + \mathbb{E}\left[\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{\substack{m=1\\ m\neq j}}^{n} \vartheta_{i,n} Z_{n,j}^{(i)} Z_{n,m}^{(i)}\right]$$
$$= \sum_{i=1}^{n}\sum_{j=1}^{n} \vartheta_{i,n}\mathbb{P}(Z_{n,j}^{(i)} = 1) + \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{\substack{m=1\\ m\neq j}}^{n} \vartheta_{i,n}\mathbb{P}(Z_{n,j}^{(i)} = 1)\mathbb{P}(Z_{n,m}^{(i)} = 1 | Z_{n,j}^{(i)} = 1).$$

On account of (5.20) and (5.21) it is straightforward to see that

$$\sum_{i=1}^{n}\sum_{j=1}^{n} \vartheta_{i,n}\mathbb{P}(Z_{n,j}^{(i)} = 1) \leq C \begin{cases} n & \text{for } \gamma < \frac{1}{2}, \\ \log(n)n & \text{for } \gamma = \frac{1}{2}, \\ n^{2\gamma} & \text{for } \gamma > \frac{1}{2}. \end{cases}$$

Furthermore, (5.22) yields

$$\mathbb{P}(Z_{n,m}^{(i)} = 1 | Z_{n,j}^{(i)} = 1) \leq C \begin{cases} n^{\gamma} i^{\gamma-1} m^{-2\gamma} & \text{for } \gamma < \frac{1}{2}, \\ i^{-\frac{1}{2}} m^{-\frac{1}{2}} \log(m) + n^{\frac{1}{2}} i^{-\frac{1}{2}} m^{-1} & \text{for } \gamma = \frac{1}{2}, \\ i^{\gamma-1} m^{\gamma-1} + n^{\gamma} i^{\gamma-1} m^{-2\gamma} & \text{for } \gamma > \frac{1}{2}, \end{cases}$$

so that using (5.20), (5.21) and (5.6) again gives

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{\substack{m=1\\ m\neq j}}^{n} \vartheta_{i,n}\mathbb{P}(Z_{n,j}^{(i)} = 1)\mathbb{P}(Z_{n,m}^{(i)} = 1 | Z_{n,j}^{(i)} = 1) \leq \begin{cases} n & \text{for } \gamma < \frac{1}{2}, \\ n\log(n)^2 & \text{for } \gamma = \frac{1}{2}, \\ n^{4\gamma-1} & \text{for } \gamma > \frac{1}{2}. \end{cases}$$

Thus

$$\mathbb{E}\left[\sum_{i=1}^{n} \vartheta_{i,n} R_{n,i}^2\right] \leq \begin{cases} n & \text{for } \gamma < \frac{1}{2}, \\ n\log(n)^2 & \text{for } \gamma = \frac{1}{2}, \\ n^{4\gamma-1} & \text{for } \gamma > \frac{1}{2}. \end{cases} \tag{5.34}$$

Combining (5.32), (5.33),(5.34) and the fact that $\mu_n \sim n$ by Lemma 5.3, proves the assertion.

□

We can finally prove our main result, Theorem 5.1.

Proof of Theorem 5.1. Remember that due to Theorem 2.6

$$d_W(\tilde{W}_n, Z) \leq \frac{\mu_n}{\sigma_n^2}\sqrt{\frac{2}{\pi}}\sqrt{\mathbb{V}[\mathbb{E}\left[W_n^s - W_n|\mathcal{G}_n\right]]} + \frac{\mu_n}{\sigma_n^3}\mathbb{E}\left[(W_n^s - W_n]^2\right].$$

According to Lemma 5.3 σ_n^2 is at least of order n, so substituting the results given in Lemmas 5.3 and 5.9 into the first term yields

$$\frac{\mu_n}{\sigma_n^2}\sqrt{\frac{2}{\pi}}\sqrt{\mathbb{V}[\mathbb{E}\left[W_n^s - W_n|\mathcal{G}_n\right]]} \leq C\begin{cases} \frac{1}{\sqrt{n}} & \text{for } \gamma < \frac{1}{2}, \\ \frac{\log(n)^2}{\sqrt{n}} & \text{for } \gamma = \frac{1}{2}, \\ n^{3\gamma-2} & \text{for } \gamma > \frac{1}{2}. \end{cases}$$

Moreover Lemma 5.10 gives

$$\frac{\mu_n}{\sigma_n^3}\mathbb{E}\left[(W_n^s - W_n)^2\right] \leq \frac{C}{\sigma_n^3}\begin{cases} n & \text{for } \gamma < \frac{1}{2} \\ n\log(n)^2 & \text{for } \gamma = \frac{1}{2} \\ n^{4\gamma-1} & \text{for } \gamma > \frac{1}{2} \end{cases} \leq C\begin{cases} \frac{1}{\sqrt{n}} & \text{for } \gamma < \frac{1}{2}, \\ \frac{\log(n)^2}{\sqrt{n}} & \text{for } \gamma = \frac{1}{2}, \\ n^{4\gamma-\frac{5}{2}} & \text{for } \gamma > \frac{1}{2}. \end{cases}$$

Here, the last equality uses the fact that due to Lemma 5.3 we have $\sigma_n^2 \geq Cn$. Combining these two results proves Theorem 5.1.

□

List of notations

Events

$\{\ell \overset{1.}{\leftrightarrow} m\}$	first connection of ℓ is to m	106
$\{k \to \ell\}$	there exists an edge between vertices $\ell < k$ in $\mathcal{G}_n$	86
$\{k \overset{i}{\to} \ell\}$	there is an edge pointing from k to ℓ in $\mathcal{G}_n^{(i)}$	86
$A_{n+1}^{(\diamond)}$		56
B_k	Y_n is in state k at the next observation	56
$C_j^{(k)}$	j only connects to k when inserted into the network	96
$D_i^{(k)}$	up to time n the only incoming edge of i is provided by k	96
$D_{n+1}^{(\diamond)}$		57
$E_+^{(i)}(\ell)$	no outgoing edges of ℓ in $\mathcal{G}_n^{(i)}$	103
$E_-^{(i)}(\ell)$	no incoming edges of ℓ in $\mathcal{G}_n^{(i)}$	103
$E_{k,n}$	$F_k \cap G_{k,n}$	69
F_k	we look at the indegree of vertex k at time k	66
$G_{k,n}$	we look at the same vertex from time k to time n	66
$H_{k,\ell}$		70
M_0^{ℓ}	the first movement occurs in the interval I_ℓ	57
M_ℓ^{ψ}		72
$S_{ij}^{(n)}$	i is the unique neighbour of j in $\mathcal{G}_n$	124

Random Variables

$\mathcal{D}_{n,I}$		93
d_I		93
D_n	outdegree of vertex n	49
$d_n(i)$		90
$D_{n,I}$		93
$\deg^+(i)$	outdegree of vertex i	90
$\deg_n^-(k)$	indegree of vertex k in $\mathcal{G}_n$	18
$\deg_n^{BA,-}(i)$	indegree of vertex i in the Barabási-Albert model at time n	40
$\mathcal{G}_n$	preferential attachment graph on n vertices	18
$\mathcal{G}_n^{(i)}$	$\mathcal{G}_n$ conditioned on i being isolated	84
J_n	birth-time of vertex considered at time n	29

N_t		23
N_Y		23
$\Pi(\lambda)$	Poisson process of intensity λ	57
$PA_n^{m,\delta}(b)$		74
$PA_n^{1,\delta}$		39
$R_{n,I}$		93
W_n	§5 :number of isolated vertices in $\mathcal{G}_n$	83
W_n	§3: indegree of a uniformly chosen vertex in $\mathcal{G}_n$	22
X_n	Markov chain giving the indegree of a uniformly chosen vertex	30
$X_{i,n}$	§5: vertex i is isolated in $\mathcal{G}_n$	84
X^s	random variable having size-bias distribution with respect to X	8
Y_n	observation of Z_t at discrete time instances	54
$Y_{n,i}$	vertex i has degree 1 in $\mathcal{G}_n$	94
Z_t	continuous-time process with generator $\mathcal{A}$	25
$Z_{n,\ell}^{(i)}$	ℓ is isolated in $\mathcal{G}_n^{(i)}$ but not in $\mathcal{G}_n$	103
$Z_t^{(\mu)}$	Z_t started according to μ	26
$Z_t^{(k)}$	Z_t started in state k	26

Others

$\lVert\cdot\rVert_{TV}$	total variation metric	13
$\propto$	proportional to	1
$\vartheta_{i,n}$	$\mathbb{E}\left[X_{i,n}\right]$	84
g_h^t		26
$h(k,\ell)$		30
$a_{j,\ell}^{(i)}$		91
$a_{j,\ell}$		91
$\mathcal{A}$	Stein operator/ generator of a Markov process	5
C_f		92
Δg	$g(k+1)-g(k)$	27
$\Delta^{(1)}h(k,\ell)$	$h(k+1,\ell)-h(k,\ell)$	33
d_K	Kolmogorov distance	11
d_{TV}	Total variation distance	11
d_W	Wasserstein distance	11
$\mathbb{E}_k$	expectation conditional on starting the process in state k	25
$\mathrm{Exp}(\alpha)$	exponential distribution with parameter	28
Φ	distribution function of the standard normal distribution	6
γ		83
g_A		26
λ_n	$\mathbb{E}\left[f(W_{n-1})\right]$	49
$\hat{\mu}_{k-1}^f(\ell,i)$	$\mathbb{E}\left[f(\deg_{k-1}^-(\ell))\lvert i\to\ell\right]$	86
μ	limiting indegree distribution	21

$\mu^f_{k-1}(\ell)$	$\mathbb{E}\left[f(\deg^-_{k-1}(\ell))\right]$	84
μ_n	$\mathbb{E}\left[W_n\right]$	83
$\mu^f_{k-1}(\ell, i)$	$\mathbb{E}\left[f(\deg^-_{k-1}(\ell))\vert X_{i,n} = 1\right]$	84
$\hat{p}_{n+1}$		60
$\psi(m, \ell, k)$		117
$\Psi^f(n, m, i, j)$		128
$p^{(\cdot)}(\cdot\vert\cdot)$	transition probabilites for the coupling	60
$p_{n,0}$	$\mathbb{P}(\deg^+(n) = 0)$	88
T_t	operator semigroup corresponding to X_t	11
$x \wedge y$	minimum of x and y	113
$\xi^k_m(j, i)$		91

List of Figures

3.1 Evolution of $\mathcal{G}_n$ 19
3.2 Example of $\mathcal{G}_{50}$ for $f(k) = 0.8k + 0.6$ 19
3.3 Example of $\mathcal{G}_{50}$ for $f(k) = \sqrt{k+1}$ 19
3.4 Example of $\mathcal{G}_{50}$ for $f(k) = 0.4k + 0.3$ 19
3.5 Example of $\mathcal{G}_{50}$ for $f(k) = 0.8\sqrt{k} + 0.6$ 19

4.1 Coupling 63

5.1 Example of $\mathcal{G}_{35}$ for $f(k) = \frac{3}{10}k^{\frac{1}{5}} + \frac{2}{5}$ 85
5.2 Construction of $\mathcal{G}^{(I)}_{35}$ from $\mathcal{G}_{35}$ 85
5.3 Construction of $\mathcal{G}^{(I)}_{35}$ from $\mathcal{G}_{35}$ 86
5.4 $\mathcal{G}^{(I)}_{35}$ 86

Figures 3.2 – 3.5 and 5.1 – 5.4 have been generated by the use of a python script written by the author of the thesis.

Bibliography

[ACJP08] A. Aiello, W. Bonato, C. Cooper, J. Janssen, and P. Prałat. A spatial web graph model with local influence regions. *Internet Math.*, 5(1-2):175–196, 2008.

[BA99] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[Bar88] A. D. Barbour. Stein's method and Poisson process convergence. *J. Appl. Prob.*, Special Vol. 25A:175–184, 1988.

[BDO19] C. Betken, H. Döring, and M. Ortgiese. Fluctuations in a general preferential attachment model via Stein's method. *to appear in: Random Structures & Algorithms*, 2019+.

[BH84] A. D. Barbour and P. Hall. On the rate of poisson convergence. *Mathematical Proceedings of the Cambridge Philosophical Society*, 95(3):473–480, 1984.

[Bha07] S. Bhamidi. Universal techniques to analyze preferential attachment trees: Global and local analysis. *Preprint available at http://www.unc.edu/∼bhamidi/*, 2007.

[BP99] T. C. Brown and M. J. Phillips. Negative binomial approximation with Stein's method. *Methodology and Computing in Applied Probability*, 1(4):407–421, 1999.

[BRST01] B. Bollobás, O. Riordan, J. Spencer, and G. Tusnády. The degree sequence of a scale-free random graph process. *Random Structures & Algorithms*, 18(3):279–290, 2001.

[BX01] T. C. Brown and A. Xia. Stein's method and birth-death processes. *The Annals of Probability*, 29(3):1373–1403, 2001.

[C+06] A. Capocci et al. Preferential attachment in the growth of social networks: The internet encyclopedia wikipedia. *Phys. Rev. E*, 74:036116, Sep 2006.

[CFR11] S. Chatterjee, J. Fulman, and A. Röllin. Exponential approximation by Stein's method and spectral graph theory. *ALEA. Latin American Journal of Probability and Mathematical Statistics*, 8:197–223, 2011.

[CGS10] L.H.Y. Chen, L. Goldstein, and Q.M. Shao. *Normal Approximation by Stein's Method*. Springer Verlag, 2010.

[Che75] L. H. Y. Chen. Poisson approximation for dependent trials. *Ann. Probab.*, 3(3):534–545, 1975.

[Çın11] E. Çınlar. *Probability and Stochastics*. Graduate Texts in Mathematics. Springer New York, 2011.

[dH12] Frank den Hollander. Probability theory : The coupling method. 2012.

[DM09] S. Dereich and P. Mörters. Random networks with sublinear preferential attachment: degree evolutions. *Electron. J. Probab.*, 14:no. 43, 1222–1267, 2009.

[DM13] S. Dereich and P. Mörters. Random networks with sublinear preferential attachment: the giant component. *Ann. Probab.*, 41(1):329–384, 2013.

[Döb38] W. Döblin. Exposé de la Théorie des Chaînes simples constants de Markoff á un nombre fini d'États. *Revue Math. de l'Union Interbalkanique*, 2:77–105, 1938.

[EK86] S. N. Ethier and T. G. Kurtz. *Markov processes. Characterization and convergence*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986.

[ER59] P. Erdős and A. Rényi. On random graphs I. *Publicationes Mathematicae (Debrecen)*, 6:290–297, 1959.

[For09] E. Ford. *Barabási-Albert random graphs, scale-free distributions and bounds for approximation through Stein's method*. PhD thesis, Oxford, 2009.

[G$^+$10] P. Giammatteo et al. A PageRank-based preferential attachment model for the evolution of the world wide web. *EPL (Europhysics Letters)*, 91(1):18004, 2010.

[Göt91] F. Götze. On the rate of convergence in the multivariate CLT. *Ann. Probab.*, 19(2):724–739, 1991.

[GR96] L. Goldstein and Y. Rinott. Multivariate normal approximations by stein's method and size bias couplings. *Journal of Applied Probability*, 33(1):1–17, 1996.

[HK18] L. Hou and J. T. Kwok. Power law in sparsified deep neural networks. *CoRR*, abs/1805.01891, 2018.

[Hof17] R. van der Hofstad. *Random graphs and complex networks. Vol. 1.* Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2017.

[KR01] P. L. Krapivsky and S. Redner. Organization of growing random networks. *Phys. Rev. E*, 63:066123, May 2001.

[LPW06] D. A. Levin, Y. Peres, and E. L. Wilmer. *Markov chains and mixing times.* American Mathematical Society, 2006.

[Nor98] J. R. Norris. *Markov chains.* Cambridge series in statistical and probabilistic mathematics. Cambridge University Press, 1998.

[OS05] R. I. Oliveira and J. H. Spencer. Connectivity transitions in networks with super-linear preferential attachment. *Internet Mathematics*, 2:121–163, 2005.

[Pek96] E. A. Peköz. Stein's method for geometric approximation. *J. Appl. Prob.*, 33(3):707–713, 1996.

[PR11] E. A. Peköz and A. Röllin. New rates for exponential approximation and the theorems of Rényi and Yaglom. *Ann. Probab.*, 39(2):587–608, 03 2011.

[PRR13] E. A. Peköz, A. Röllin, and N. Ross. Degree asymptotics with rates for preferential attachment random graphs. *Ann. Appl. Probab.*, 23(3):1188–1218, 06 2013.

[PRR17] E. Peköz, A. Röllin, and N. Ross. Joint degree distributions of preferential attachment random graphs. *Advances in Applied Probability*, 49(2):368–387, 2017.

[Ros11] N. Ross. Fundamentals of Stein's method. *Probability Surveys*, 8:210–293, 2011.

[Ros13] N. Ross. Power laws in preferential attachment graphs and Stein's method for the negative binomial distribution. *Advances in Applied Probability*, 45(3):876–893, 2013.

[RTV07] A. Rudas, B. Tóth, and B. Valkó. Random trees and general branching processes. *Random Structures & Algorithms*, 31(2):186–202, 2007.

[Ste72] C. Stein. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Vol. II: Probability theory*, pages 583–602, 1972.