A LAW OF LARGE NUMBERS AND CENTRAL LIMIT THEOREM FOR THE LEAVES IN A RANDOM GRAPH MODEL

MAXIMILIAN WIMMER

M.S. CREATIVE COMPONENT; SUPERVISOR SUNDER SETHURAMAN

Abstract. We are calculating the expectation and variance of the number of leaves in the scale-free network model suggested in [CL06]. Afterwards, we prove a Law of Large Numbers and a Central Limit Theorem for the number of leaves using the Martingale Central Limit Theorem. Finally, we check our results by simulation.

1. Introduction

Random graphs are the key tool in mathematics for modeling large real world networks. The classical model for random graphs was introduced by Erdős and Rényi in their groundbreaking papers [ER59, ER60] in the late 1950s. They start with a graph with n vertices. Each of the $\binom{n}{2}$ pairs of vertices will be connected by an edge with a fixed probability p, i.e. the appearance of one edge is independent to the appearance of the other edges.

Among questions like connectivity, size of the maximum component and average path length, one important property of random graphs is the distribution of the vertex degrees, or shortly the degree distribution of the graph. The Erdős-Rényi model yields a degree distribution which is approximately Poisson with a tail decreasing faster then exponentially, as shown in [Bol01].

In 1999, Barabási and Albert studied the degree distribution in various large real world networks like the internet graph. They discovered that these surprisingly follow a power law distribution. The relatively heavy tail in such a distribution results in the existence of a few vertices with high degree. However, this observation conflicted with the classical Erdős-Rényi model for random graphs. Therefore they conjectured two main features of real world networks that were not incorporated in the Erdős-Rényi model:

- Growth. Networks in the real world usually develop over time. They start with a relatively low number of vertices, whereas new vertices are added later with time.
- Preferential attachment. It is also reasonable to assume that a new vertex is more likely to connect to a big existing vertex than to a small existing vertex.

Finally, Barabási and Albert proposed a new model for random graphs in [BA99]:

Date: July 6, 2006.

2000 *Mathematics Subject Classification.* Primary 05C80, Secondary 60G42.

> ...starting with a small number (m_0) of vertices, at every time step we add a new vertex with $m(\leq m_0)$ edges that link the new vertex to m different vertices already present in the system. To incorporate preferential attachment, we assume that the probability Π that a new vertex will be connected to vertex i depends on the connectivity k_i of that vertex, so that $\Pi(k_i) = k_i / \sum_j k_j$.

Bollobás et al. clarified this model in [BRST01] and provided a rigorous proof that this model yields a power law distribution $p_k \sim Ck^{-3}$ as $k \to \infty$. However, the analysis of real data suggested that the exponent of the power law distribution does not necessarily have to be exactly -3. Table 1 shows the exponents in various measurements of real world networks. Notice that these are often directed graphs, whereas we will only deal with undirected graphs in this paper.

Network	Size	$\langle d \rangle$	$-\gamma_{\text{out}}$	$-\gamma_{\text{in}}$	l	Reference
WWW	325,729	4.51	2.45	2.1	11.2	Albert et al., 1999
WWW	$4 \cdot 10^7$	7	2.38	2.1		Kumar et al., 1999
WWW	$2 \cdot 10^8$	7.5	2.72	2.1	16	Broder et al., 1999
Internet, domain	3,015–4,389	3.42–3.76	2.1–2.2	2.1–2.2	4	Faloutsos, 1999
Internet, router	3,888	2.57	2.48	2.48	12.15	Faloutsos, 1999
Internet, router	150,000	2.66	2.4	2.4	11	Govindan, 2000
Movie actors	212,250	28.78	2.3	2.3	4.54	Barabási and Albert, 1999
Co-authors, SPIRES	56,527	173	1.2	1.2	4	Newmann, 2001b
Co-authors, neuro.	209,293	11.54	2.1	2.1	6	Barabási et al., 2001
Co-authors, math.	70,975	3.9	2.5	2.5	9.5	Barabási et al., 2001
Sexual contacts	2,810		3.4	3.4		Liljeros et al., 2001
Metabolic, E. coli	778	7.4	2.2	2.2	3.2	Jeong et al., 2000
Protein, S. cerev.	1,870	2.39	2.4	2.4		Jeong, Mason et al., 2001
Ythan estuary	134	8.7	1.05	1.05	2.43	Montoya and Solé, 2000
Silwood Park	154	4.75	1.13	1.13	3.4	Montoya and Solé, 2000
Citation	783,339	8.75		3		Redner, 1998
Phone call	$53 \cdot 10^6$	3.16	2.1	2.1		Aiello et al., 2000
Words, co-occurrence	460,902	70.13	2.7	2.7		Ferrer and Solé, 2001
Words, synonyms	22,311	13.48	2.8	2.8		Yook et al., 2001b

Table 1. Different studies of real world networks. We indicate the size of the network and its average degree $\langle d \rangle$. Since most real world networks are directed, we separately list the exponents of the outdegree γ_{out} and indegree γ_{in}. Moreover, we compute the average path length l. The references can be found in the paper [AB02].

In order to get different exponents for the power law distribution, Krapivsky et al. generalized this model in [KRL00]. Here, vertices of degree k are chosen with probability proportional to $f(k)$ for some (increasing) function f. They analyzed different choices of f that lead to different results:

- $f(k) = k$ yields exactly the Barabási-Albert model.
- $f(k) = k^\alpha$, $\alpha < 1$, yields $p_k \approx \mu k^{-\alpha} \exp(-ck^{1-\alpha})$.

- $f(k) = k^\alpha$, $\alpha > 1$, yields a model where there is only one vertex with degree of order k, whereas all other vertices have degree of order 1 (cf. [OS05]).
- $f(k) = k + \beta$, $\beta > -1$, yields $p_k \sim Ck^{-3-\beta}$.

Especially the last case seems interesting, since it yields a power law distribution with arbitrary exponent < -2. For this case, a Law of Large Numbers and a Central Limit Theorem for the number of vertices with low degree has been proven by Móri in [Mór02].

There are various similar models around. Mitzenmacher gives an overview of many of them in his survey [Mit04]. Each of those models has its unique properties regarding to degree distribution, connectivity, average path length between two vertices, etc. Since none of the models can match all available sample data, there is no general consensus about a "best" model.

We will analyze one of these models, suggested very recently by Chung and Lu in [CL06]. This model is very similar to a generalized Pólya urn scheme, as described in [CHJ03].

2. Model Definition

We start at time $t = 1$ with a graph consisting of two vertices that are connected by a single edge. At each step $t = 2, 3, \ldots$, we first flip a coin. With probability $p > 0$, we add a new vertex to the graph and connect it with one new edge to an existing vertex. The latter is selected randomly with probability proportional to its degree. With probability $1 - p$, we add a new edge between two existing vertices. These are selected randomly with probability proportional to their degree, whereas self-loops are permitted.

This model yields a power law degree distribution with exponent $-1 - \frac{2}{2-p}$, as shown in [CF03, Mit04]. If $p = 1$, this model is identical to Barabási-Albert model with $m_0 = 2$ and $m = 1$.

Throughout this paper, let Z_n denote the number of leaves, i.e. the number of vertices with degree 1, after the n-th step. Moreover, $\mathcal{F}_n$ denotes the σ-algebra generated by the edges after the n-th step. It is clear, that we have exactly n edges at time n. Therefore, the sum of the degrees of all vertices at time n is always $2n$.

3. Expectation and Higher Moments

We follow the same idea as in [Szy05] to compute the moments. There are only four possibilities how the number of leaves changes at each step. It can increase by one, if we add a new vertex to the graph and attach it to an existing vertex with degree ≥ 2. It remains, if we add a new vertex and attach it to a previous leaf. It also remains, if we only add a new edge between two vertices with degree ≥ 2. It can decrease by one, if we only add a new edge, where one of the vertices had degree 1. Finally, it can decrease by 2 if we only add a new edge between two previous leaves. Therefore we have

$$Z_{n+1} = \begin{cases} Z_n + 1 & \text{with probability } p\left(1 - \frac{Z_n}{2n}\right) \\ Z_n & \text{with probability } p\frac{Z_n}{2n} + (1-p)\left(1 - \frac{Z_n}{2n}\right)^2 \\ Z_n - 1 & \text{with probability } 2(1-p)\frac{Z_n}{2n}\left(1 - \frac{Z_n}{2n}\right) \\ Z_n - 2 & \text{with probability } (1-p)\left(\frac{Z_n}{2n}\right)^2 \end{cases}.$$

Therewith we can compute

$$\begin{aligned} \mathbb{E}[Z_{n+1} \mid \mathcal{F}_n] &= (Z_n+1)\cdot p\left(1 - \frac{Z_n}{2n}\right) \\ &\quad + Z_n \cdot \left(p\frac{Z_n}{2n} + (1-p)\left(1 - \frac{Z_n}{2n}\right)^2\right) \\ &\quad + (Z_n - 1)\cdot 2(1-p)\frac{Z_n}{2n}\left(1 - \frac{Z_n}{2n}\right) \\ &\quad + (Z_n - 2)\cdot (1-p)\left(\frac{Z_n}{2n}\right)^2 \\ &= \left(1 + \frac{p-2}{2n}\right) Z_n + p. \end{aligned}$$

Taking expectations on both sides we obtain the recurrence relation

$$\mathbb{E}Z_{n+1} = \left(1 + \frac{p-2}{2n}\right)\mathbb{E}Z_n + p,$$

which, with initial condition $Z_1 = 2$, can be solved (cf. Lemma A.1) as

$$\mathbb{E}Z_n = \frac{2p}{4-p}n + \frac{8\Gamma(\frac{1}{2}p + n - 1)}{(p-4)\Gamma(\frac{1}{2}p - 1)\Gamma(n)} \quad \sim n.$$

In order to compute the expected value of the 2nd moment $\mathbb{E}Z_n^2$, we have to solve the recurrence relation

$$\mathbb{E}Z_{n+1}^2 = \left(1 + \frac{-2+p}{n} + \frac{1-p}{2n^2}\right)\mathbb{E}Z_n^2 + \left(2p + \frac{2-3p}{2n}\right)\mathbb{E}Z_n + p.$$

This can be done as shown in Lemma A.3. The leading terms for $\mathbb{E}Z_n^2$ are

$$\mathbb{E}Z_n^2 = \frac{4p^2}{(4-p)^2}n^2 + \frac{2p(p^2-12p+12)}{(4-p)^2(3-p)}n + \text{l. o. t.} \quad \sim n^2.$$

Now we can write down the asymptotic distribution of the variance of Z_n,

$$\operatorname{Var} Z_n = \mathbb{E}Z_n^2 - (\mathbb{E}Z_n)^2 = \frac{2p(p^2-12p+12)}{(4-p)^2(3-p)}n + \text{l. o. t.} \quad \sim n.$$

Similarly we can compute the asymptotic behavior of the third and forth moment of Z_n. Here the recurrence relations become

$$\begin{aligned} \mathbb{E}Z_{n+1}^3 &= \left(1 + \frac{-6+3p}{2n} + \frac{3-3p}{2n^2}\right)\mathbb{E}Z_n^3 + \left(3p + \frac{6-9p}{2n} + \frac{-3+3p}{2n^2}\right)\mathbb{E}Z_n^2 + \left(3p + \frac{-2+p}{2n}\right)\mathbb{E}Z_n + p \\ \mathbb{E}Z_{n+1}^4 &= \left(1 + \frac{-4+2p}{n} + \frac{3-3p}{n^2}\right)\mathbb{E}Z_n^4 + \left(4p + \frac{6-9p}{n} + \frac{-6+6p}{n^2}\right)\mathbb{E}Z_n^3 \\ &\quad + \left(6p + \frac{-4+2p}{n} + \frac{7-7p}{2n^2}\right)\mathbb{E}Z_n^2 + \left(4p + \frac{2-3p}{2n}\right)\mathbb{E}Z_n + p. \end{aligned}$$

The leading terms can be computed as

$$\begin{aligned} \mathbb{E}Z_n^3 &= \frac{8p^3}{(4-p)^3}n^3 + \frac{12p^2(p^2-12p+12)}{(4-p)^3(3-p)}n^2 + \text{l. o. t.} \quad \sim n^3 \\ \mathbb{E}Z_n^4 &= \frac{16p^4}{(4-p)^4}n^4 + \frac{48p^3(p^2-12p+12)}{(4-p)^4(3-p)}n^3 + \text{l. o. t.} \quad \sim n^4. \end{aligned}$$

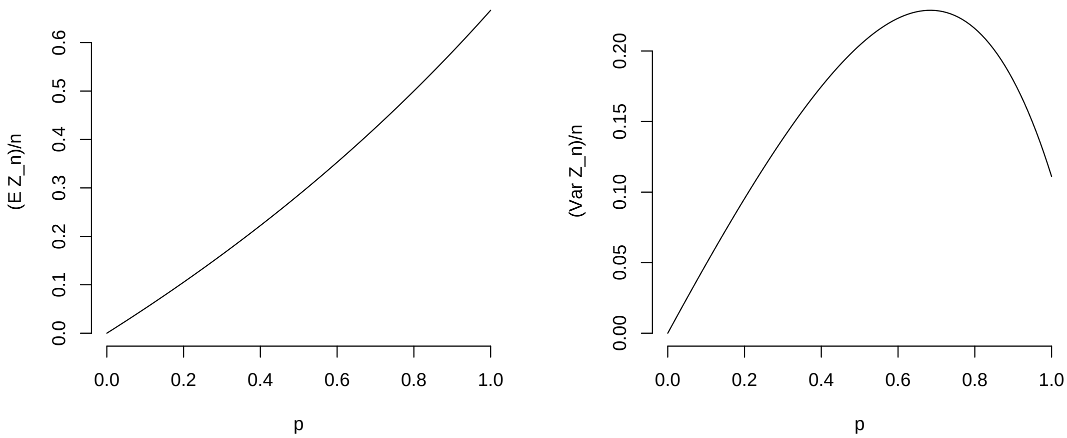

FIGURE 1. The asymptotic behavior of $\frac{\mathbb{E}Z_n}{n}$ and $\frac{\operatorname{Var} Z_n}{n}$ depending on p

For later reference, we finally want to compute the asymptotic behavior of the 2nd central moment $\mathbb{E}\overline{Z_n^2} = \mathbb{E}(Z_n - \mathbb{E}Z_n)^2$ and the 4th central moment $\mathbb{E}\overline{Z_n^4} = \mathbb{E}(Z_n - \mathbb{E}Z_n)^4$. Clearly,

$$\mathbb{E}\overline{Z_n^2} = \operatorname{Var} Z_n \quad \sim n.$$

In order to calculate the leading term of $\mathbb{E}\overline{Z_n^4}$ exactly, we would eventually need more terms of the raw moments than we calculated. However, we can derive a bound for the asymptotic behavior of $\mathbb{E}\overline{Z_n^4}$. First notice that

$$\mathbb{E}\overline{Z_n^4} = -3(\mathbb{E}Z_n)^4 + 6(\mathbb{E}Z_n)^2\mathbb{E}Z_n^2 - 4\mathbb{E}Z_n\mathbb{E}Z_n^3 + \mathbb{E}Z_n^4.$$

Therefore, when we are using our approximations for $\mathbb{E}Z_n^2$, $\mathbb{E}Z_n^3$ and $\mathbb{E}Z_n^4$, we can expect all terms of order 3 or higher of $\mathbb{E}\overline{Z_n^4}$ to be exact. However, these terms cancel out. Hence we have

$$\mathbb{E}\overline{Z_n^4} \quad \sim n^i \quad \text{for some } i < 3.$$

For the case $p = 1$, we can solve the recurrence relations above exactly. This can be found in appendix B.

4. Law of Large Numbers for Z_n

Having already computed the 4th central moment of Z_n, we can give a simple proof of a Law of Large Numbers for Z_n:

Theorem 4.1 (Law of Large Numbers)**.** *We have*

$$\frac{Z_n}{n} \xrightarrow[n\to\infty]{} \frac{2p}{4-p} \quad a.s.$$

Proof. For all $k > 0$, Chebyshev's inequality yields,

$$\mathbb{P}\left(\left|\frac{Z_n - \mathbb{E}Z_n}{n}\right| > \frac{1}{k}\right) \le \frac{\mathbb{E}(Z_n - \mathbb{E}Z_n)^4}{n^4} \cdot k^4 = \frac{\mathbb{E}\overline{Z_n^4}}{n^4} \cdot k^4 \sim \frac{1}{n^i} \cdot k^4 \xrightarrow[n\to\infty]{} 0,$$

with $i > 1$. Therefore we have

$$\sum_{j=1}^{\infty} \mathbb{P}\left(\left|\frac{Z_j - \mathbb{E}Z_j}{j}\right| > \frac{1}{k}\right) \leq k^4 \sum_{j=1}^{\infty} \frac{\mathbb{E}\overline{Z_j^4}}{j^4} \sim k^4 \sum_{j=1}^{\infty} \frac{1}{j^i} < \infty.$$

Thus, by the Borel-Cantelli Lemma,

$$\mathbb{P}\left(\left|\frac{Z_n - \mathbb{E}Z_n}{n}\right| > \frac{1}{k} \text{ i.o.}\right) = 0.$$

Now let

$$A_k = \left\{\omega : \left|\frac{Z_n(\omega)}{n} - \mu\right| > \frac{1}{k} \text{ i.o.}\right\},$$

so

$$A_k^C = \left\{\omega : \left|\frac{Z_n(\omega)}{n} - \mu\right| \leq \frac{1}{k} \text{ for all } n \geq n_0(\omega) \text{ for some } n_0(\omega)\right\}.$$

Moreover, let

$$A = \bigcup_{k \geq 1} A_k.$$

Since $\mathbb{P}(A_k) = 0$ for all $k \geq 1$, we have $\mathbb{P}(A) = 0$. Now consider $A^C = \bigcap_{k \geq 1} A_k^C$. Clearly, $\mathbb{P}(A^C) = 1$ and $A^C \subseteq \{\omega : \lim_{n \to \infty} \frac{Z_n(\omega)}{n} = \mu\}$. □

5. Central Limit Theorem for Z_n

We want to find values α_n and β_n such that

$$M_n = \alpha_n Z_n + \beta_n$$

is a martingale. Taking expectation yields

$$\begin{aligned}
\mathbb{E}[M_{n+1} \mid \mathcal{F}_n] &= \alpha_{n+1}\left(\left(1 + \tfrac{p-2}{2n}\right) Z_n + p\right) + \beta_{n+1} \\
&= \alpha_{n+1}\left(1 + \tfrac{p-2}{2n}\right) Z_n + \alpha_{n+1} p + \beta_{n+1} \overset{!}{=} M_n.
\end{aligned}$$

This yields the recurrence relations $\alpha_n = \alpha_{n+1}\left(1 + \frac{p-2}{2n}\right)$ and $\beta_n = \alpha_{n+1} p + \beta_{n+1}$, which can be solved using Lemma A.5 and Lemma A.6 with initial conditions $\alpha_1 = 1$ and $\beta_1 = 1$ as

$$\begin{aligned}
\alpha_n &= \frac{\Gamma(\frac{1}{2}p)\Gamma(n)}{\Gamma(n + \frac{1}{2}p - 1)} \quad \sim n^{1-\frac{1}{2}p} \\
\beta_n &= 1 - \frac{2p}{p-4} + \frac{4\Gamma(\frac{1}{2}p + 1)\Gamma(n+1)}{(p-4)\Gamma(n + \frac{1}{2}p - 1)} \quad \sim n^{2-\frac{1}{2}p}.
\end{aligned}$$

Hence the asymptotic behavior of the martingale is

$$M_n \quad \sim n^{1-\frac{1}{2}p}\left(Z_n - \tfrac{2p}{4-p} n\right).$$

Now for $k \geq 1$ we define ε_k by

$$\varepsilon_k = M_k - M_{k-1} = \alpha_k Z_k - \alpha_{k-1} Z_{k-1} + \beta_k - \beta_{k-1} = \alpha_k Z_k - \alpha_{k-1} Z_{k-1} - \alpha_k p,$$

where $M_0 = 0$. Since

$$\mathbb{E}[\varepsilon_k \mid \mathcal{F}_{k-1}] = \alpha_k \left(\left(1 + \tfrac{p-2}{2k-2}\right) Z_{k-1} + p\right) - \left(1 + \tfrac{p-2}{2k-2}\right) \alpha_k Z_{k-1} - \alpha_k p = 0,$$

ε_k is a martingale difference array. Now, for $k \le n$, let

$$\xi_k^{(n)} = \frac{\varepsilon_k}{\alpha_n \sqrt{\operatorname{Var} Z_n}}.$$

Clearly, $\xi_k^{(n)}$ is also a martingale difference array since for all $k = 1, 2, \ldots, n$, we have

$$\mathbb{E}\left[\xi_k^{(n)} \;\middle|\; \mathcal{F}_{k-1}\right] = \frac{1}{\alpha_n \sqrt{\operatorname{Var} Z_n}} \mathbb{E}[\varepsilon_k \mid \mathcal{F}_{k-1}] = 0.$$

Lemma 5.1. *We have*

$$\max_{1 \le k \le n} |\xi_k^{(n)}| \xrightarrow[n\to\infty]{} 0 \quad \textit{uniformly.}$$

Proof. By the recurrence relation of α_k, $|\xi_k^{(n)}|$ can be rewritten as

$$\begin{aligned} |\xi_k^{(n)}| &= \frac{\alpha_k}{\alpha_n \sqrt{\operatorname{Var} Z_n}} \left| Z_k - \left(1 + \tfrac{p-2}{2k-2}\right) Z_{k-1} - p \right| \\ &\le \frac{1}{\sqrt{\operatorname{Var} Z_n}} \cdot \underbrace{\frac{\alpha_k}{\alpha_n}}_{\le 1} \cdot \Bigg(\underbrace{|Z_k - Z_{k-1}|}_{\le 1} + \underbrace{\left| \frac{(2-p)Z_{k-1}}{2(k-1)} \right|}_{\le 1} + \underbrace{|p|}_{\le 1} \Bigg) \le \frac{3}{\sqrt{\operatorname{Var} Z_n}}. \end{aligned}$$

So, $|\xi_k^{(n)}|$ is uniformly bounded by $\frac{3}{\sqrt{\operatorname{Var} Z_n}} \xrightarrow[n\to\infty]{} 0.$ ∎

Corollary 5.2. *For every $\varepsilon > 0$, the conditional Lindeberg condition*

$$\sum_{k=1}^{n} \mathbb{E}\left[(\xi_k^{(n)})^2 \; I(|\xi_k^{(n)}| > \varepsilon) \;\middle|\; \mathcal{F}_{k-1} \right] \xrightarrow[n\to\infty]{\mathrm{p}} 0$$

holds.

Rewriting $\xi_k^{(n)}$ as

$$\xi_k^{(n)} = \frac{\alpha_k}{\alpha_n \sqrt{\operatorname{Var} Z_n}} (Z_k - \mathbb{E}[Z_k \mid \mathcal{F}_{k-1}]),$$

we can compute the conditional expectation

$$\begin{aligned} \mathbb{E}\left[(\xi_k^{(n)})^2 \;\middle|\; \mathcal{F}_{k-1} \right] &= \mathbb{E}\left[\tfrac{\alpha_k^2}{\alpha_n^2 \operatorname{Var} Z_n} \left(Z_k - \mathbb{E}[Z_k \mid \mathcal{F}_{k-1}]\right)^2 \;\middle|\; \mathcal{F}_{k-1} \right] \\ &= \frac{\alpha_k^2}{\alpha_n^2 \operatorname{Var} Z_n} \left(- \frac{(p^2 - 2p + 2) Z_{k-1}^2}{4(k-1)^2} - \frac{(2p^2 - p - 2) Z_{k-1}}{2(k-1)} + p(1-p) \right). \end{aligned}$$

Lemma 5.3. *For all $k \le n$, we have*

$$\sum_{k=1}^{n} \mathbb{E}\left[(\xi_k^{(n)})^2 \;\middle|\; \mathcal{F}_{k-1} \right] \xrightarrow[n\to\infty]{\mathrm{p}} 1.$$

Proof. First define

$$\eta_k = \mathbb{E}\left[(\xi_k^{(n)})^2 \;\middle|\; \mathcal{F}_{k-1} \right] - \mathbb{E}\left(\xi_k^{(n)} \right)^2.$$

Since $\xi_k^{(n)}$ is a martingale difference array, $\sum_{k=1}^{n} \mathbb{E}(\xi_k^{(n)})^2 = 1$ and it suffices to show that

$$\sum_{k=1}^{n} \eta_k \xrightarrow[n\to\infty]{\mathrm{p}} 0.$$

In fact, we will show that

$$\mathbb{E}\left(\sum_{k=1}^{n}\eta_k\right)^2 = \sum_{k=1}^{n}\mathbb{E}\eta_k^2 + 2\sum_{i=1}^{n}\sum_{j=i+1}^{n}\mathbb{E}\eta_i\eta_j \xrightarrow[n\to\infty]{} 0. \tag{5.1}$$

Let

$$w_k = \frac{-\alpha_k^2(p^2-2p+2)\overline{Z_{k-1}^2}}{4(k-1)^2\alpha_n^2 \operatorname{Var} Z_n} \quad \text{and} \quad y_k = \frac{-\alpha_k^2(2p^2-p-2)\overline{Z_{k-1}}}{2(k-1)\alpha_n^2 \operatorname{Var} Z_n},$$

so $\eta_k = w_k + y_k$. Notice that

$$\mathbb{E}\left(\sum_{k=1}^{n}\eta_k\right)^2 = \mathbb{E}\left(\sum_{k=1}^{n} w_k + y_k\right)^2 \le 2\mathbb{E}\left(\sum_{k=1}^{n} w_k\right)^2 + 2\mathbb{E}\left(\sum_{k=1}^{n} y_k\right)^2.$$

Since

$$\mathbb{E}\left(\sum_{k=1}^{n} w_k\right)^2 = \sum_{k=1}^{n}\mathbb{E}w_k^2 + 2\sum_{i=1}^{n}\sum_{j=i+1}^{n}\mathbb{E}w_iw_j$$

and

$$\mathbb{E}\left(\sum_{k=1}^{n} y_k\right)^2 = \sum_{k=1}^{n}\mathbb{E}y_k^2 + 2\sum_{i=1}^{n}\sum_{j=i+1}^{n}\mathbb{E}y_iy_j,$$

we actually have to calculate four limits to prove equation (5.1):

(a) $\displaystyle\sum_{k=1}^{n}\mathbb{E}y_k^2 \xrightarrow[n\to\infty]{} 0$

(b) $\displaystyle\sum_{i=1}^{n}\sum_{j=i+1}^{n}\mathbb{E}y_iy_j \xrightarrow[n\to\infty]{} 0$

(c) $\displaystyle\sum_{k=1}^{n}\mathbb{E}w_k^2 \xrightarrow[n\to\infty]{} 0$

(d) $\displaystyle\sum_{i=1}^{n}\sum_{j=i+1}^{n}\mathbb{E}w_iw_j \xrightarrow[n\to\infty]{} 0$

(a) Since

$$|y_k| = \frac{\alpha_k^2}{\alpha_n^2}\cdot\frac{1}{\operatorname{Var} Z_n}\cdot|2p^2-p-2|\cdot\left|\frac{\overline{Z_{k-1}}}{2(k-1)}\right| \le 1\cdot\frac{1}{n}\cdot 1\cdot 1 = \frac{1}{n},$$

we have

$$\sum_{k=1}^{n}\mathbb{E}y_k^2 \le n\cdot\frac{1}{n^2} = \frac{1}{n} \xrightarrow[n\to\infty]{} 0.$$

(b) Note that since $M_{k-1} = \alpha_{k-1}\overline{Z_{k-1}}$, we can rewrite y_k as

$$y_k = \frac{-\alpha_k^2(2p^2-p-2)M_{k-1}}{2(k-1)\alpha_{k-1}\alpha_n^2 \operatorname{Var} Z_n}.$$

Therefore,

$$\begin{aligned}\sum_{i=1}^{n}\sum_{j=i+1}^{n}\mathbb{E}y_iy_j &= \sum_{i=1}^{n}\mathbb{E}\left(y_i\sum_{j=i+1}^{n}y_j\right)\\ &= \sum_{i=1}^{n}\mathbb{E}\left(y_i\mathbb{E}\left[\sum_{j=i+1}^{n}y_j \;\middle|\; \mathcal{F}_{i-1}\right]\right)\\ &= \sum_{i=1}^{n}\mathbb{E}\left(y_iM_{i-1}\sum_{j=i+1}^{n}\frac{-\alpha_j^2(2p^2-p-2)}{2(j-1)\alpha_{j-1}\alpha_n^2\operatorname{Var}Z_n}\right).\end{aligned}$$

Now,

$$-\frac{2p^2-p-2}{\alpha_n^2\operatorname{Var}Z_n}\sum_{j=i+1}^{n}\frac{\alpha_j^2}{2(j-1)\alpha_{j-1}}\sim\frac{1}{n^{3-p}}\sum_{j=i+1}^{n}\frac{1}{j^{p/2}}\sim\frac{1}{n^{3-p}}$$

and since $\mathbb{E}M_i^2\sim i^{2-p}\mathbb{E}\overline{Z_i^2}\sim i^{3-p}$,

$$\begin{aligned}y_iM_{i-1} &= -\frac{1}{\alpha_{i-1}}\cdot\frac{\alpha_i^2}{\alpha_n^2}\cdot\frac{2p^2-p-2}{\operatorname{Var}Z_n}\cdot\frac{M_{i-1}^2}{2(i-1)}\\ &\sim \frac{1}{i^{1-\frac{1}{2}p}}\cdot\frac{i^{2-p}}{n^{2-p}}\cdot\frac{1}{n}\cdot i^{2-p}=\frac{1}{n^{3-p}}i^{3-\frac{3}{2}p}.\end{aligned}$$

So,

$$\sum_{i=1}^{n}\sum_{j=i+1}^{n}\mathbb{E}y_iy_j\sim\sum_{i=1}^{n}\frac{1}{n^{6-2p}}i^{3-\frac{3}{2}p}\sim\frac{1}{n^{6-2p}}\cdot n^{4-\frac{3}{2}p}\xrightarrow[n\to\infty]{}0.$$

(c) Since

$$|w_k|=\frac{\alpha_k^2}{\alpha_n^2}\cdot\frac{1}{\operatorname{Var}Z_n}\cdot|p^2-2p+2|\cdot\frac{\overline{Z_{k-1}^2}}{4(k-1)^2}\le 1\cdot\frac{1}{n}\cdot 2\cdot 1=\frac{2}{n},$$

we have

$$\sum_{k=1}^{n}\mathbb{E}w_k^2\le n\cdot\frac{4}{n^2}=\frac{4}{n}\xrightarrow[n\to\infty]{}0.$$

(d) We first need to find a martingale N_n of the form

$$N_n=\gamma_nZ_n^2+\delta_nZ_n+\epsilon_n.$$

Taking expectation yields

$$\begin{aligned}\mathbb{E}[N_{n+1}\mid\mathcal{F}_n] &= \gamma_{n+1}\left(1+\tfrac{p-2}{n}+\tfrac{1-p}{2n^2}\right)Z_n^2\\ &\quad+\left(\gamma_{n+1}\left(2p+\tfrac{2-3p}{2n}\right)+\delta_{n+1}\left(1+\tfrac{p-2}{2n}\right)\right)Z_n\\ &\quad+\gamma_{n+1}p+\delta_{n+1}p+\epsilon_{n+1}\quad\stackrel{!}{=}N_n\end{aligned}$$

We will use the initial conditions $\gamma_2=1$, $\delta_1=1$ and $\epsilon_1=1$. The solution of the recurrence relation for γ can be written as

$$\gamma_n=\prod_{k=2}^{n-1}\frac{2k^2}{2k^2+(2p-4)k+1-p}.$$

Splitting up the quadratic denominator into its roots, we can use formula (A.3) to compute

$$\gamma_n = \frac{\Gamma(\frac{1}{2}p - \frac{1}{2}\sqrt{p^2-2p+2})\Gamma(\frac{1}{2}p + \frac{1}{2}\sqrt{p^2-2p+2})\Gamma(n)^2}{\Gamma(n+\frac{1}{2}p - 1 - \frac{1}{2}\sqrt{p^2-2p+2})\Gamma(n+\frac{1}{2}p-1+\frac{1}{2}\sqrt{p^2-2p+2})}.$$

The leading term of γ_n is

$$\gamma_n = \Gamma(\tfrac{1}{2}p - \tfrac{1}{2}\sqrt{p^2-2p+2})\Gamma(\tfrac{1}{2}p + \tfrac{1}{2}\sqrt{p^2-2p+2})n^{2-p} + \text{l. o. t.}$$

Using Corollary A.4, we get

$$\delta_n = \frac{4p\Gamma(\frac{1}{2}p - \frac{1}{2}\sqrt{p^2-2p+2})\Gamma(\frac{1}{2}p + \frac{1}{2}\sqrt{p^2-2p+2})}{4-p} n^{3-p} + \text{l. o. t.}$$

Since $\epsilon_n = -p\sum_{k=2}^n (\gamma_k + \delta_k)$, we have

$$\epsilon_n \sim n^{4-p} \quad \text{at most.}$$

We can rewrite the martingale N_n as

$$N_n = \gamma_n (Z_n - \tfrac{\delta_n}{2\gamma_n})^2 + \nu_n \quad \sim \gamma_n \overline{Z_n^2} + \nu_n.$$

Now, $\overline{Z_{k-1}^2} \sim (N_{k-1} - \nu_{k-1})/\gamma_{k-1}$, so

$$w_k \sim \frac{-\alpha_k^2(p^2-2p+2)(N_{k-1}-\nu_{k-1})}{4(k-1)^2\alpha_n^2\gamma_{k-1}\operatorname{Var} Z_n}.$$

Similarly as in (b),

$$\sum_{i=1}^n \sum_{j=i+1}^n \mathbb{E} w_i w_j \sim \sum_{i=1}^n \mathbb{E}\left(w_i (N_{i-1} - \nu_{i-1}) \sum_{j=i+1}^n \frac{-\alpha_j^2(p^2-2p+2)}{4(j-1)^2\gamma_{j-1}\alpha_n^2 \operatorname{Var} Z_n}\right).$$

Now,

$$\frac{p^2-2p+2}{\operatorname{Var} Z_n} \sum_{j=i+1}^n \left(\frac{1}{\alpha_n^2}\cdot\frac{1}{4(j-1)^2}\cdot\frac{-\alpha_j^2}{\gamma_{j-1}}\right) \sim \frac{1}{n^{3-p}} \sum_{j=i+1}^n \frac{1}{j^2}.$$

Since $\mathbb{E}(N_i - \nu_i)^2 \sim \gamma_i^2 \mathbb{E}\overline{Z_i^4} \sim i^{4-2p}\cdot i^3 = i^{7-2p}$ at most,

$$\begin{aligned} w_i(N_{i-1}-\nu_{i-1}) \quad &\sim \quad \frac{-\alpha_i^2}{\gamma_{i-1}}\cdot\frac{1}{\alpha_n^2}\cdot\frac{p^2-2p+2}{\operatorname{Var} Z_n}\cdot\frac{(N_{i-1}-\nu_{i-1})^2}{4(i-1)^2} \\ &\sim \quad 1\cdot\frac{1}{n^{2-p}}\cdot\frac{1}{n}\cdot i^{5-2p} = \frac{1}{n^{3-p}} i^{5-2p}. \end{aligned}$$

Hence,

$$\sum_{i=1}^n \sum_{j=i+1}^n \mathbb{E} w_i w_j \sim \frac{1}{n^{6-2p}} \sum_{i=1}^n i^{5-2p} \sum_{j=i+1}^n \frac{1}{j^2} \xrightarrow[n\to\infty]{} 0.$$

This completes our proof of Lemma 5.3. □

Now let us recall the Martingale Central Limit Theorem in the version of Corollary 3.2 in [HH80].

Theorem 5.4 (MCLT)**.** *Let* $\{S_k^{(n)}, \mathcal{F}_k^{(n)}, 1 \le k \le n, n \ge 1\}$ *be a zero-mean, square-integrable martingale array with differences* $X_k^{(n)}$ *and let* η^2 *be an a.s. finite r.v. Suppose that*

(a) $\sum_{k=1}^n \mathbb{E}\left[(X_k^{(n)})^2 \ I(X_k^{(n)} > \varepsilon) \middle| \ \mathcal{F}_{k-1}^{(n)} \right] \xrightarrow[n\to\infty]{\mathrm{p}} 0$ *for all* $\varepsilon > 0$,

(b) $V_n^2 = \sum_{k=1}^n \mathbb{E}\left[(X_k^{(n)})^2 \ \middle| \ \mathcal{F}_{k-1}^{(n)} \right] \xrightarrow[n\to\infty]{\mathrm{p}} \eta^2$,

(c) *the* σ*-fields are nested:* $\mathcal{F}_k^{(n)} \subseteq \mathcal{F}_k^{(n+1)}$ *for* $1 \le i \le n$, $n \ge 1$, *and*

(d) $\mathbb{P}(\eta^2 > 0) = 1$.

Then,

$$S_n^{(n)}/V_n^2 \Rightarrow N(0,1) \text{ in distribution.}$$

Therewith, the Central Limit Theorem for Z_n follows immediately.

Theorem 5.5 (Central Limit Theorem)**.** *We have*

$$Z_n \Rightarrow N(\mu, \sigma),$$

with

$$\mu = \tfrac{2p}{4-p} n \quad \text{and} \quad \sigma^2 = \tfrac{2p(p^2-12p+12)}{(4-p)^2(3-p)} n.$$

Proof. Note that $\overline{Z_n}/\sqrt{\operatorname{Var} Z_n} = \sum_{k=1}^n \xi_k^{(n)}$. Corollary 5.2 is part (a) of Theorem 5.4, Lemma 5.3 is part (b). Parts (c) and (d) are obviously true. Therefore,

$$\sum_{k=1}^n \xi_k^{(n)} \Rightarrow N(0,1).$$

□

6. Simulating the Distribution of Z_n

We wrote a C program, cf. appendix C, to simulate graphs in our model. For different values of p, we were letting 100,000 graphs grow with 10,000 edges each and we obtained the following distributions for Z_{10000}:

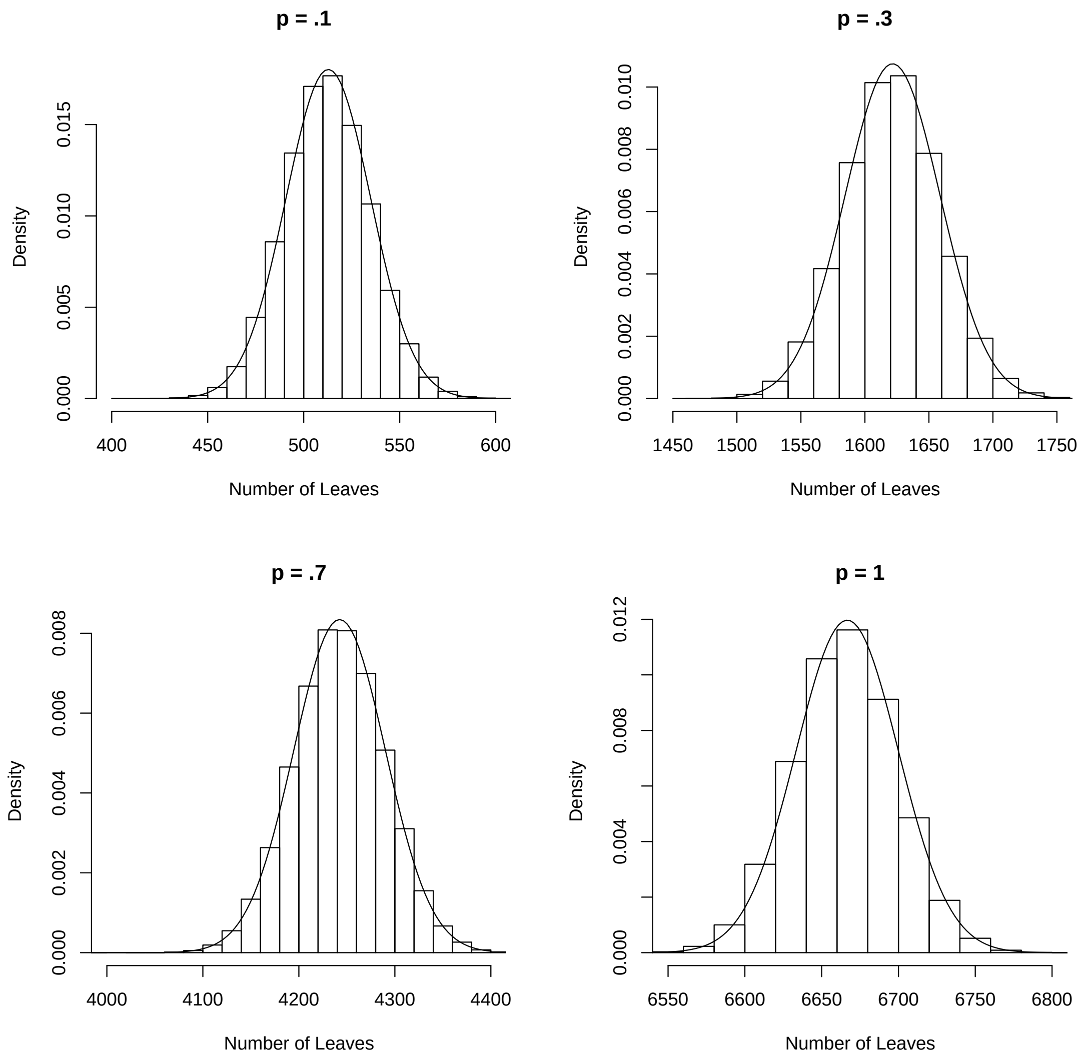

Figure 2. Histograms of Z_{10000} for different values of p

The histograms show the actual simulated data, whereas the overlying line corresponds to a normal distribution with mean μ and variance σ^2 as defined in Theorem 5.5.

Appendix A. Recurrence relations

Lemma A.1. *The recurrence relation*

$$a_{n+1} = \left(1 + \tfrac{\alpha}{n}\right) n + \beta,$$

with initial condition a_1, has the solution

$$a_n = \frac{\beta}{1-\alpha} n + \frac{((\alpha-1)a_1 + \beta)\Gamma(\alpha+n)}{(\alpha-1)\Gamma(1+\alpha)\Gamma(n)}.$$

Proof. The solution of the recurrence relation can be written as

$$a_1 \prod_{j=1}^{n-1} \left(1 + \frac{\alpha}{j}\right) + \beta \sum_{j=1}^{n-1} \prod_{k=j+1}^{n-1} \left(1 + \frac{\alpha}{k}\right). \tag{A.2}$$

Now,

$$\prod_{j=1}^{n-1} \left(1 + \frac{\alpha}{j}\right) = \frac{1}{(n-1)!} \prod_{j=1}^{n-1} (j+\alpha) = \frac{\Gamma(\alpha+n)}{\Gamma(n)\Gamma(1+\alpha)}, \tag{A.3}$$

and

$$\prod_{k=j+1}^{n-1} \left(1 + \frac{\alpha}{k}\right) = \frac{\Gamma(1+j)\Gamma(\alpha+n)}{\Gamma(1+\alpha+j)\Gamma(n)}. \tag{A.4}$$

Finally, we use the identity

$$\sum_{j=2}^{n} \frac{\Gamma(j)}{\Gamma(j+\alpha)} = \frac{\Gamma(1+\alpha)\Gamma(n+1) - \Gamma(\alpha+n)}{(1-\alpha)\Gamma(1+\alpha)\Gamma(\alpha+n)}. \tag{A.5}$$

The latter equality can be proven easily by induction. Plugging in equations (A.3), (A.4) and (A.5) into equation (A.2) yields the result. □

Lemma A.2. *We have the following asymptotic expansions:*

$$\begin{aligned} \tfrac{\Gamma(n)}{\Gamma(n+\alpha)} &= n^{-\alpha} - \tfrac{\alpha(\alpha-1)}{2} n^{-\alpha-1} + \mathrm{l.\,o.\,t.} \\ \tfrac{\Gamma(n+\alpha)}{\Gamma(n)} &= n^{\alpha} + \tfrac{\alpha(\alpha-1)}{2} n^{\alpha-1} + \mathrm{l.\,o.\,t.} \end{aligned}$$

Proof. Using the Euler-Maclaurin formula, we can expand

$$\Gamma(n+\alpha) = e^{-n}\sqrt{2\pi} n^{n+\alpha-\frac{1}{2}} + e^{-n}\tfrac{1}{6}(1 - 6\alpha + 6\alpha^2)\sqrt{\tfrac{\pi}{2}} n^{n+\alpha-\frac{3}{2}} + \mathrm{l.\,o.\,t.}$$

Now the claim follows by simple algebra. □

Lemma A.3. *The recurrence relation*

$$a_{n+1} = \left(1 - \tfrac{\theta}{n} + \tfrac{\alpha}{n^2}\right) a_n + (\gamma n^{\nu} + \eta n^{\nu-1} + \mathrm{l.\,o.\,t.}),$$

with $\nu \geq 1$ and initial condition a_1, has the solution

$$\frac{\gamma}{\theta+\nu+1} n^{\nu+1} + \frac{2\alpha\gamma - \gamma\nu(1+\nu) + 2\eta(\theta+\nu+1)}{2(\theta+\nu)(\theta+\nu+1)} n^{\nu} + \mathrm{l.\,o.\,t.}$$

Proof. The solution can be written as

$$a_n = a_1 \prod_{j=1}^{n-1} \left(1 - \tfrac{\theta}{j} + \tfrac{\alpha}{j^2}\right) + \sum_{j=1}^{n-1} \prod_{k=j+1}^{n-1} \left(1 - \tfrac{\theta}{k} + \tfrac{\alpha}{k^2}\right) \left(\gamma j^\nu + \eta j^{\nu-1} + \mathrm{l.\,o.\,t.}\right).$$

First we want to approximate

$$\prod_{k=j+1}^{n-1} \left(1 - \tfrac{\theta}{k} + \tfrac{\alpha}{k^2}\right) = A_1 j^\theta + A_2 j^{\theta-1} + \mathrm{l.\,o.\,t.}$$

Factoring $1 - \frac{\theta}{k} + \frac{\alpha}{k^2} = \left(1 - \frac{\theta+\sqrt{\theta^2-4\alpha}}{2k}\right)\left(1 - \frac{\theta-\sqrt{\theta^2-4\alpha}}{2k}\right)$, we can use equation (A.4) and the asymptotic expansion from Lemma A.2 compute

$$\begin{aligned} A_1 &= \frac{\Gamma\left(n-\frac{\theta}{2}-\frac{1}{2}\sqrt{\theta^2-4\alpha}\right)\Gamma\left(n-\frac{\theta}{2}+\frac{1}{2}\sqrt{\theta^2-4\alpha}\right)}{\Gamma(n)^2} \\ A_2 &= \frac{\left(2\alpha+\theta-\theta^2\right)\Gamma\left(n-\frac{\theta}{2}-\frac{1}{2}\sqrt{\theta^2-4\alpha}\right)\Gamma\left(n-\frac{\theta}{2}+\frac{1}{2}\sqrt{\theta^2-4\alpha}\right)}{2\Gamma(n)^2} \end{aligned}$$

Using Lemma A.2 again, we can expand A_1 and A_2 as

$$\begin{aligned} A_1 &= n^{-\theta} + \tfrac{\theta^2+\theta-2\alpha}{2} n^{-\theta-1} + \mathrm{l.\,o.\,t.} \\ A_2 &= \tfrac{2\alpha+\theta-\theta^2}{2} n^{-\theta} + \tfrac{\theta^2-(\theta^2-2\alpha)^2}{4} n^{-\theta-1} + \mathrm{l.\,o.\,t.} \end{aligned}$$

Now since

$$\sum_{j=1}^{n-1} j^\theta = \tfrac{1}{\theta+1} n^{\theta+1} - \tfrac{1}{2} n^\theta + \mathrm{l.\,o.\,t.},$$

we have

$$\begin{aligned} &\sum_{j=1}^{n-1} \prod_{k=j+1}^{n-1} \left(1 - \tfrac{\theta}{k} + \tfrac{\alpha}{k^2}\right) \left(\gamma j^\nu + \eta j^{\nu-1} + \mathrm{l.\,o.\,t.}\right) \\ &\quad = \sum_{j=1}^{n-1} (A_1 j^\theta + A_2 j^{\theta-1} + \mathrm{l.\,o.\,t.})(\gamma j^\nu + \eta j^{\nu-1} + \mathrm{l.\,o.\,t.}) \\ &\quad = \tfrac{\gamma A_1}{\theta+\nu+1} n^{\theta+\nu+1} + \left(\tfrac{\gamma A_2 + \eta A_1}{\theta+\nu} - \tfrac{\gamma A_1}{2}\right) n^{\theta+\nu} + \mathrm{l.\,o.\,t.} \\ &\quad = \tfrac{\gamma}{\theta+\nu+1} n^{\nu+1} + \tfrac{2\alpha\gamma - \gamma\nu(1+\nu) + 2\eta(\theta+\nu+1)}{2(\theta+\nu)(\theta+\nu+1)} n^\nu + \mathrm{l.\,o.\,t.} \end{aligned}$$

Clearly, these will be the leading terms of a_n, since

$$a_1 \prod_{j=1}^{n-1} \left(1 - \tfrac{\theta}{j} + \tfrac{\alpha}{j^2}\right) \sim A_1 \sim n^{-\theta}. \tag{A.6}$$

□

Corollary A.4. *The recurrence relation*

$$a_n = a_{n+1}\left(1 + \tfrac{\theta}{n}\right) + (\gamma n^\nu + \mathrm{l.\,o.\,t.}),$$

with $\nu \geq 1$, has the solution

$$a_n = \frac{\gamma}{\theta+\nu+1} n^{\theta+1} + \mathrm{l.\,o.\,t.}$$

Proof. Rewrite

$$a_{n+1} = (1 - \tfrac{\theta}{n} + \mathrm{l.\,o.\,t.})a_n + (\gamma n^\nu + \mathrm{l.\,o.\,t.})$$

and use Lemma A.3 (notice that the first l. o. t.-term doesn't affect the proof). □

Lemma A.5. *The recurrence relation*

$$a_{n+1} = \tfrac{n}{n+\beta} a_n,$$

with initial condition a_1, has the solution

$$a_n = \frac{\Gamma(1+\beta)\Gamma(n)}{\Gamma(\beta+n)} a_1.$$

Proof. The solution can be written as

$$a_n = a_1 \prod_{k=1}^{n-1} \frac{k}{k+\beta} = a_1 \frac{\Gamma(k)}{\frac{\Gamma(\beta+n)}{\Gamma(1+\beta)}}.$$

□

Lemma A.6. *The recurrence relation*

$$a_{n+1} = a_n - \tfrac{\beta\Gamma(n)}{\Gamma(n+\gamma)},$$

with initial condition a_1, has the solution

$$a_n = a_1 - \frac{\beta}{(\gamma-1)\Gamma(\gamma)} + \frac{\beta\Gamma(\gamma)\Gamma(n)}{(\gamma-1)\Gamma(\gamma-1+n)}.$$

Proof. The solution can be written as

$$a_n = a_1 - \beta \sum_{k=1}^{n-1} \frac{\Gamma(k)}{\Gamma(k+\gamma)}.$$

Using the identity (A.5) we receive the result. □

Appendix B. Expectation and Higher Moments in the BA Model

In the Barabási-Albert model, i.e. if $p = 1$, we can calculate the moments similarly as in section 3. However, this time the recurrence relations are simple enough to be solved exactly.

First some notation: Define

$$g(n) = \frac{\Gamma(n)}{\Gamma(n-\frac{1}{2})} \quad \text{and} \quad g_\pi^{-1}(n) = \frac{\Gamma(n-\frac{1}{2})}{\sqrt{\pi}\,\Gamma(n)}.$$

Note that $g(n) \sim \sqrt{n}$ and $g(n+1) = \frac{2n}{2n-1} g(n)$.

We have

$$Z_{n+1} = \begin{cases} Z_n & \text{with probability } \frac{Z_n}{2n} \\ Z_n + 1 & \text{with probability } 1 - \frac{Z_n}{2n} \end{cases}.$$

Therefore we can compute

$$\begin{aligned} \mathbb{E}[Z_{n+1} \mid \mathcal{F}_n] &= Z_n \tfrac{Z_n}{2n} + (Z_n + 1)\left(1 - \tfrac{Z_n}{2n}\right) \\ &= \left(1 - \tfrac{1}{2n}\right) Z_n + 1. \end{aligned}$$

Taking expectations on both sides we obtain the recurrence relation

$$\mathbb{E}Z_{n+1} = \left(1 - \tfrac{1}{2n}\right) \mathbb{E}Z_n + 1,$$

which, with initial condition $Z_1 = 2$, can be solved as

$$\mathbb{E}Z_n = \tfrac{2}{3}n + \tfrac{4}{3} g_\pi^{-1}(n) \quad \sim \tfrac{2}{3}n.$$

Similarly, we can compute the higher moments

$$\begin{aligned} \mathbb{E}Z_n^2 &= \tfrac{1}{9}\left(4n^2 + n\right) + \tfrac{4}{9}(4n+1) g_\pi^{-1}(n) \quad \sim \tfrac{4}{9}n^2 \\ \mathbb{E}Z_n^3 &= \tfrac{1}{135}\left(40n^3 + 30n^2 - n\right) + \tfrac{1}{90}\left(160n^3 - 120n^2 - 160n - 217\right) \tfrac{2}{2n-3} g_\pi^{-1}(n) \quad \sim \tfrac{8}{27}n^3 \\ \mathbb{E}Z_n^4 &= \tfrac{1}{405}\left(80n^4 + 120n^3 + 7n^2 - 6n\right) \\ &\quad + \tfrac{1}{405}\left(640n^4 - 1096n^2 - 2712n - 1755\right) \tfrac{2}{2n-3} g_\pi^{-1}(n) \quad \sim \tfrac{16}{81}n^4. \end{aligned}$$

Now we can write down the variance of Z_n as

$$\operatorname{Var} Z_n = \mathbb{E}Z_n^2 - (\mathbb{E}Z_n)^2 = \tfrac{1}{9}n + \tfrac{4}{9}(4n+1) g_\pi^{-1}(n) - \tfrac{16}{9}\left(g_\pi^{-1}(n)\right)^2 \quad \sim \tfrac{1}{9}n.$$

Clearly,

$$\mathbb{E}\overline{Z_n^2} = \operatorname{Var} Z_n \quad \sim \tfrac{1}{9}n,$$

and we can compute

$$\mathbb{E}\overline{Z_n^4} = \mathbb{E}\left(Z_n - \mathbb{E}Z_n\right)^4 \quad \sim \tfrac{1}{27}n^2.$$

Appendix C. C Code

```
#include <stdio.h>
#include <stdlib.h>
#include <math.h>

int main (int argc, const char * argv[]) {
    double p = 1;
    int n = 10000;
    int numSweeps = 100000;
    int Z1;

    FILE *pFile;
    pFile = fopen("myfile.txt","w");

    int sweeps;
    int i;

    int *Deg = malloc(sizeof (int) * n+1);
    if (Deg == NULL) exit(EXIT_FAILURE);              // No memory.

    int *Edges = malloc(sizeof (int) * 2*n+1);
    if (Edges == NULL) exit(EXIT_FAILURE);            // No memory.

    for (sweeps=1; sweeps<=numSweeps; sweeps++) {
        for (i=1; i<=n; i++) Deg[i] = (int)0;
        Deg[1] = 1; Deg[2] =  1;

        for (i=1; i<=2*n; i++) Edges[i] = (int)0;
        Edges[1] = 1; Edges[2] = 2;

        int ii = 2;
        for (i=2; i<=n; i++) {
            double rr = rand()/((double)RAND_MAX + 1);
            if (rr < p) {
                int r = (int)((2*(double)i-2) * rand()/((double)RAND_MAX ...
                    + 1))+1;
                Deg[Edges[r]]++;
                Deg[++ii]++;
                Edges[2*i-1] = ii;
                Edges[2*i] = Edges[r];
            } else {
                int r = (int)((2*(double)i-2) * rand()/((double)RAND_MAX ...
                    + 1))+1;
                int s = (int)((2*(double)i-2) * rand()/((double)RAND_MAX ...
                    + 1))+1;
                Deg[Edges[r]]++;
                Deg[Edges[s]]++;
                Edges[2*i-1] = Edges[r];
                Edges[2*i] = Edges[s];
            }
        }
        Z1 = 0;                         // creating histogram
        for (i=1; i<=ii; i++)
            if (Deg[i] == 1) Z1++;
        fprintf(pFile, "%i\n", Z1);
    }

    fclose (pFile);
    return 0;
}
```

Acknowledgments

I would like to take this opportunity to express my thanks to those who helped me with various aspects of conducting research and the writing of this paper.

First and foremost, Dr. Sunder Sethuraman for his guidance, patience and support throughout this research and the writing of this paper. His insights and words of encouragement have often inspired me and renewed my hopes for completing my creative component.

I would also like to thank my committee members for their efforts and contributions to this work: Dr. Wolfgang Kliemann and Dr. Michael Smiley.

Finally, I would like to thank my officemate and friend Doug Ray for checking and correcting my English.

References

[AB02] Réka Albert and Albert-László Barabási, *Statistical mechanics of complex networks*, Rev. Modern Phys. **74** (2002), no. 1, 47–97.

[BA99] Albert-László Barabási and Réka Albert, *Emergence of scaling in random networks*, Science **286** (1999), no. 5439, 509–512.

[Bol01] Béla Bollobás, *Random graphs*, second ed., Cambridge Studies in Advanced Mathematics, vol. 73, Cambridge University Press, Cambridge, 2001.

[BRST01] Béla Bollobás, Oliver Riordan, Joel Spencer, and Gábor Tusnády, *The degree sequence of a scale-free random graph process*, Random Structures Algorithms **18** (2001), no. 3, 279–290.

[CF03] Colin Cooper and Alan Frieze, *A general model of web graphs*, Random Structures Algorithms **22** (2003), no. 3, 311–335.

[CHJ03] Fan Chung, Shirin Handjani, and Doug Jungreis, *Generalizations of Polya's urn problem*, Ann. Comb. **7** (2003), no. 2, 141–153.

[CL06] Fan Chung and Linyuan Lu, *Concentration inequalities and martingale inequalities — a survey*, Internet Math. (to appear 2006).

[ER59] Paul Erdős and Alfred Rényi, *On random graphs. I*, Publ. Math. Debrecen **6** (1959), 290–297.

[ER60] ______, *On the evolution of random graphs*, Magyar Tud. Akad. Mat. Kutató Int. Közl. **5** (1960), 17–61.

[HH80] Peter Hall and Christopher C. Heyde, *Martingale limit theory and its application*, Academic Press, New York, NY, 1980.

[KRL00] Paul. L. Krapivsky, Sidney Redner, and Francois Leyvraz, *Connectivity of growing random networks*, Phys. Rev. Lett. **85** (2000), 4629–4632.

[Mit04] Michael Mitzenmacher, *A brief history of generative models for power law and lognormal distributions*, Internet Math. **1** (2004), no. 2, 226–251.

[Mór02] Tamás F. Móri, *On random trees*, Studia Sci. Math. Hungar. **39** (2002), no. 1-2, 143–155.

[OS05] Roberto Oliveira and Joel Spencer, *Connectivity transitions in networks with super-linear preferential attachment*, Internet Math. **2** (2005), no. 2, 121–163.

[Szy05] Jerzy Szymański, *Concentration of vertex degrees in a scale-free random graph process*, Random Structures Algorithms **26** (2005), no. 1-2, 224–236.

Department of Mathematics, Iowa State University, Ames, IA 50011

E-mail address: `wimmer@iastate.edu`