Improved Degree Bounds and Full Spectrum Power Laws in Preferential Attachment Networks

Chen Avin
Ben Gurion University of the Negev
Be'er-Sheva, Israel
avin@cse.bgu.ac.il

Zvi Lotker
Ben Gurion University of the Negev
Be'er-Sheva, Israel
zvilo@cse.bgu.ac.il

Yinon Nahum
Weizmann Institute of Science
Rehovot, Israel
yinon.nahum@weizmann.ac.il

David Peleg
Weizmann Institute of Science
Rehovot, Israel
david.peleg@weizmann.ac.il

ABSTRACT

Consider a random preferential attachment model $G(p)$ for network evolution that allows both node and edge arrivals. Starting with an arbitrary nonempty graph G_0, at each time step, there are two possible events: with probability $p > 0$ a new node arrives and a new edge is added between the new node and an existing node, and with probability $1 - p$ a new edge is added between two existing nodes. In both cases, the involved existing nodes are chosen at random according to preferential attachment, i.e., with probability proportional to their degree. $G(p)$ is known to generate *power law networks*, i.e., the fraction of nodes with degree k is proportional to $k^{-\beta}$. Here $\beta = (4 - p)/(2 - p)$ is in the range $(2, 3]$.

Denoting the number of nodes of degree k at time t by $m_{k,t}$, we significantly improve some long-standing results. In particular, we show that $m_{k,t}$ is concentrated around its mean with a deviation of $O(\sqrt{t})$, which is independent of k. We also tightly bound the expectation $\mathbb{E}\left[m_{k,t}\right]$ with an additive error of $O(1/k)$, which is independent of t. These new bounds allow us to tightly estimate $m_{k,t}$ for a considerably larger k values than before. This, in turn, enables us to estimate other important quantities, e.g., the size of the k-rich club, namely, the set of all nodes with a degree at least k.

Finally, we introduce a new generalized model, $G(p_t, r_t, q_t)$, which extends $G(p)$ by allowing also *time-varying* probabilities for node and edge arrivals, as well as the formation of new components. We show that the extended model can produce power law networks with any exponent β in the range $(1, \infty)$. Furthermore, the concentration bounds established for $m_{k,t}$ in $G(p)$ also apply in $G(p_t, r_t, q_t)$.

KEYWORDS

preferential attachment; degree bounds; power law

KDD '17, August 13-17, 2017, Halifax, NS, Canada

DOI: 10.1145/3097983.3098012

1 INTRODUCTION

1.1 Background

Preferential attachment is one of the prevalent mechanisms for network evolution, in the context of both social networks and other complex systems, and considerable efforts have been invested in studying its properties and behavior. The basic idea in preferential attachment is that "the rich get richer", which translates in a network evolution context to the concept that nodes with high degrees have a better chance to gain new edges, i.e., the probability that a node gains a new edge is proportional to its degree.

A key feature of preferential attachment models is that they are known to generate networks with a power law degree distribution [18], i.e., a distribution where the fraction of nodes with degree k is proportional to $k^{-\beta}$ for some fixed β. The power law degree distribution has been observed in many real life networks (and complex systems), as well as non-network scenarios, and is thought to be a universal property. For example, the preferential attachment model itself was proposed by Price [19] to explain observed power law degree distributions in citation networks and later by Barabasi et al. [4] in order to explain the observed heavy-tailed degree distributions in various other networks.

Specifically, consider the following model $G(p)$ defined in [6]. Starting at time 0 with a graph consisting of a single node with a single self-loop, at time t, with probability $p > 0$ a new node arrives and a new edge is added between the new node and an existing node, and with probability $1 - p$ a new edge is added between two existing nodes. In both cases, existing nodes are chosen at random according to preferential attachment, i.e., with probability proportional to their degree. Letting $m_{k,t}$ denote the number of nodes with degree k at time t, and setting

$$\beta = \frac{4-p}{2-p} \in (2,3] \qquad \text{and} \qquad M_k = \frac{2}{4-p}\prod_{j=1}^{k-1}\frac{j}{j+\beta}\ ,$$

the following is shown in [6].

Lemma 1.1. *[6] In the $G(p)$ model the following hold:*

(i) M_k is proportional to $k^{-\beta}$.
*(ii) For every **fixed** k we have*

$$\lim_{t\to\infty}\frac{\mathbb{E}\left[m_{k,t}\right]}{p\cdot t} = M_k\ .$$

(iii) For every integers k, t and real $c \geq 0$ we have

$$\mathbb{P}\left[|m_{k,t} - p \cdot M_k \cdot (t+1)| \geq 2k\sqrt{t} \cdot \sqrt{c + \ln 2 + (k-1)\ln(t+1)}\right] \leq e^{-c} .$$

Note that while (ii) suggests that the fraction of nodes with degree k obeys a power law with exponent β, it only holds for fixed k and only in the limit ($t \to \infty$). Therefore, it is insufficient in order to estimate $\mathbb{E}\left[m_{k,t}\right]$ for large k, e.g., $k = \sqrt{t}$. Also note that (iii) implies a deviation of $O(\sqrt{k^3 \cdot t \cdot \ln t})$ for $m_{k,t}$ from $p \cdot M_k \cdot (t+1)$. This deviation bound is somewhat dissatisfactory since it increases with k, contrary to the intuition that $m_{k,t}$ should become more concentrated as k increases.

1.2 Paper Contributions

The contribution of our paper is threefold. First, we improve Lemma 1.1(ii) by providing bounds on $\mathbb{E}\left[m_{k,t}\right]$ for every k and t, rather than only for fixed k's and only in the limit $t \to \infty$. This, in turn, allows us to estimate various quantities in $G(p)$, for example, the number of nodes with degree $k = \sqrt{t}$, which was not known before. Our second contribution lies in significantly improving the best previously known concentration bounds for $m_{k,t}$ given by Lemma 1.1(iii), established over a decade ago. Specifically, we bound the deviation by $O(\sqrt{t})$, independent of k, Our third contribution is introducing an extended preferential attachment model that can produce a power law network with any exponent in the range $(1, \infty)$.

Let us now describe our results in more detail. Letting $e_t = e_0 + t$ denote the number of edges at time t, we first prove the following theorem.

Theorem 1.2. *In the $G(p)$ model, for every integers $t \geq 0$ and $1 \leq k \leq 2e_t$ we have*

$$\mathbb{E}\left[m_{k,t}\right] \leq p \cdot M_k \cdot e_t - \frac{p^2(1-p)}{2(2-p)} \cdot M_k + \frac{c^*}{k} , \quad (1)$$

$$\mathbb{E}\left[m_{k,t}\right] \geq p \cdot M_k \cdot e_t + p(1-p)M_k - \frac{2p(e_0+1-p)}{k(4-p)} , \quad (2)$$

where $c^* = \max\left\{2e_0 \ , \ 2e_1(1-p) \ , \ 2e_1^2 p \ , \ -\frac{8}{\ln(1-p)}\right\}$.

Clearly, Thm. 1.2 implies Lemma 1.1(ii). Note, however, that it also implies a considerably tighter estimate for $\mathbb{E}\left[m_{k,t}\right]$, namely, $|\mathbb{E}\left[m_{k,t}\right] - p \cdot M_k \cdot e_t| \leq O(1/k)$, and thus can be applied to estimate $\mathbb{E}\left[m_{k,t}\right]$ for any $k = o(t^{1-p/2})$.

Also note that Thm. 1.2 may be used to bound the expected size of the "rich club". Formally, defining the *k-rich club* of the network at time t to be the set of all nodes with degree at least k, and denoting its size by

$$RC_{k,t} = \sum_{j \geq k} m_{j,t} ,$$

we prove the following corollary.

Corollary 1.3. *In the $G(p)$ model, for every integers $t \geq 0$ and $1 \leq k \leq 2e_t$ we have*

$$\mathbb{E}\left[RC_{k,t}\right] \leq p \cdot e_t \cdot (\varphi_k - \varphi_{2e_t+1}) + c^* \cdot \ln\left(\frac{2e_t}{k-1}\right), \quad (3)$$

$$\mathbb{E}\left[RC_{k,t}\right] \geq p \cdot e_t \cdot (\varphi_k - \varphi_{2e_t+1}) - \frac{2e_0}{3} \cdot \ln\left(\frac{2e_t+1}{k}\right), \quad (4)$$

where $\varphi_z = (1 + (1-p/2)z)M_z = \Theta(z^{1-\beta})$.

Note that Cor. 1.3 implies $|\mathbb{E}\left[m_{k,t}\right] - pe_t(\varphi_k - \varphi_{2e_t+1})| = O(\ln(t/k))$, and thus can be applied to estimate $RC_{k,t}$ for any $k = o((t/\ln t)^{1-p/2})$.

For our second contribution, we prove the following rather general theorem, which extends Lemma 1.1(iii) by bounding any weighted combination of $m_{k,t}$ values. For integer $t \geq 0$ and reals $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_{2e_t}) \in \mathbb{R}^{2e_t}$, let $\Delta_{\boldsymbol{\varepsilon}} = \max_{i,j} |\varepsilon_i - \varepsilon_j|$ be the maximum distance between any two coordinates of $\boldsymbol{\varepsilon}$, and $\|\boldsymbol{\varepsilon}\|_\infty = \max_i |\varepsilon_i|$ be the maximum distance between a coordinate and 0. Also let $S_t(\boldsymbol{\varepsilon}) = \sum_{k=1}^{2e_t} \varepsilon_k \cdot m_{k,t}$.

Theorem 1.4. *In the $G(p)$ model, for every integer $t \geq 0$, real $c \geq 0$, and reals $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{2e_t}) \in \mathbb{R}^{2e_t}$, we have*

$$\mathbb{P}\left[S_t(\boldsymbol{\varepsilon}) - \mathbb{E}\left[S_t(\boldsymbol{\varepsilon})\right] \geq \sqrt{2ct} \cdot (2\Delta_{\boldsymbol{\varepsilon}} + 2\max(\Delta_{\boldsymbol{\varepsilon}}, \|\boldsymbol{\varepsilon}\|_\infty))\right] \leq e^{-c} ,$$

$$\mathbb{P}\left[S_t(\boldsymbol{\varepsilon}) - \mathbb{E}\left[S_t(\boldsymbol{\varepsilon})\right] \leq -\sqrt{2ct} \cdot (2\Delta_{\boldsymbol{\varepsilon}} + 2\max(\Delta_{\boldsymbol{\varepsilon}}, \|\boldsymbol{\varepsilon}\|_\infty))\right] \leq e^{-c} .$$

Note that Thm. 1.4 suggests that $m_{k,t}$ is concentrated around its mean with a deviation of $O(\sqrt{t})$. This is given by the following direct corollary.

Corollary 1.5. *In the $G(p)$ model, for every integers $k \geq 0, t \geq 0$ and real $c \geq 0$, we have*

$$\mathbb{P}\left[m_{k,t} - \mathbb{E}\left[m_{k,t}\right] \geq \sqrt{32ct}\right] \leq e^{-c},$$

$$\mathbb{P}\left[m_{k,t} - \mathbb{E}\left[m_{k,t}\right] \leq -\sqrt{32ct}\right] \leq e^{-c}.$$

Note that Thm. 1.4 also suggests that $RC_{k,t}$ is concentrated around its mean with a deviation of $O(\sqrt{t})$. This is given by the following direct corollary.

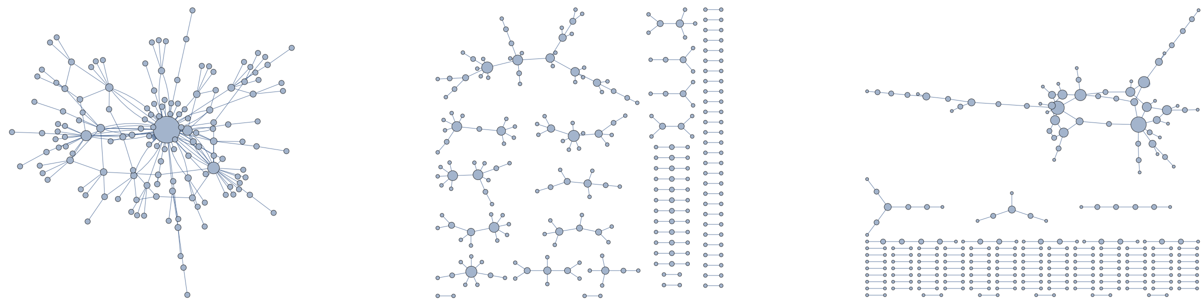
Corollary 1.6. *In the model $G(p)$, for every integers $k \geq 0, t \geq 0$ and real $c \geq 0$, we have*

$$\mathbb{P}\left[RC_{k,t} - \mathbb{E}\left[RC_{k,t}\right] \geq \sqrt{32ct}\right] \leq e^{-c},$$

$$\mathbb{P}\left[RC_{k,t} - \mathbb{E}\left[RC_{k,t}\right] \leq -\sqrt{32ct}\right] \leq e^{-c},$$

For our third contribution, we consider a rather general preferential attachment model $G(p_t, r_t, q_t)$ for network evolution that allows three types of events. Starting with an arbitrary nonempty graph G_0, at time t, the following three types of events may occur. i) **Node arrival event:** with probability p_t a new node arrives and a new edge is added between the new node and an existing node. ii) **Edge arrival event:** with probability r_t a new edge is added between two existing nodes, and iii) **Component arrival event:** with probability $q_t = 1 - p_t - r_t$ two new nodes arrive and a new edge is added between them. In the first two cases, existing nodes that participate in the event are chosen at random according to preferential attachment.

The model $G(p_t, r_t, q_t)$ was previously studied with null q_t and constant p_t, i.e., $q_t = 0$ and $p_t = p > 0$ for all t. It is known



(a) (b) (c)

Figure 1: Three networks generated by the model $G(p_t, r_t, q_t)$. Each with 200 edges. Node sizes are proportional to their degrees. (a) No component events: $p_t = 0.75, r_t = 0.25, q_t = 0$. This is the standard model with one connected component. Here $\beta = 2.6$. (b) No edge events: $p_t = 0.75, r_t = 0, q_t = 0.25$. We obtain many connected components. Here $\beta = 3\frac{2}{3}$. (c) No node events: $p_t = 0, r_t = 0.25, q_t = 0.75$. One giant component emerges. Here $\beta = 5$.

that for the special case where q_t is null and p_t is constant, the resulting graph consists of a single connected component and obeys a power law with exponent $\beta = (4-p)/(2-p) \in (2,3]$. In this paper, we extend the model by allowing p_t to vary over time, and even diminish as t tends to infinity. The rationale for this is that it allows us to model situations where the nodes already existing in the network continue to create new connections among themselves, hence as their number grows, the fraction of new edge events grows as well. This also allows the model both to account for changes in the growth rate of large networks, and to explain phenomena which were observed empirically in many networks, but could not be captured by the original model. Such examples are *densification* [17] (i.e., unbounded average degree as the network grows), power law distribution with exponent smaller than 2 or larger than 3 [7], multiple connected components, etc. We show that in this model it is possible to obtain networks obeying a power law with any exponent in the range $(1, \infty)$. In particular, letting $y_t = p_t + 2q_t$, we show that whenever $\lim_{t\to\infty} y_t = y < 2$, $\sum_{i=1}^{\infty} y_i = \infty$ and $\lim_{t\to\infty} t \cdot y_{t+1} / \sum_{j=1}^{t} y_j = \Gamma > 0$, the resulting network obeys a power law with exponent $\beta = 1 + \frac{2\Gamma}{2-y}$. Thus selecting the parameters so as to satisfy $y_t = t^{\alpha-2}$ $(1 < \alpha < 2)$ yields networks obeying a power law with exponent $\beta = \alpha$. Likewise, setting the parameters to satisfy $y_t = 1/\ln(t+2)$ yields networks obeying a power law with exponent $\beta = 2$, and aiming for $y_t = 2 - 2/(\alpha - 1)$ $(\alpha > 2)$ yields networks obeying a power law with exponent $\beta = \alpha$. In addition, we show that Thm. 1.4 applies also in the extended model $G(p_t, r_t, q_t)$, see Thm. 3.1. Figure 1 presents networks generated by the model $G(p_t, r_t, q_t)$ for selected parameters p_t, r_t and q_t.

1.3 Related Work

The first to consider a *preferential attachment* based network evolution process was Price [19] who studied citation networks. Motivated by early work of Simon [20] who showed power laws in non networked settings like wealth, city sizes and word distributions, Price was interested in a network model that will generate a power law degree distribution. He proposed a directed graph model where a new paper cites an existing paper i with probability $\frac{q_i+a}{\sum_i q_i+a} = \frac{q_i+a}{n(c+a)}$, where q_i is the number of papers that already cite i (i.e., the in-degree of i), a is a positive constant and c is the average (in- and out-)degree. So a new paper generates on average c citations. This model leads to a power law distribution of the in-degrees with exponent $\beta = 2 + \frac{a}{c} > 2$.

Barabási and Albert [4] independently proposed their model for growing undirected networks and coined the term "preferential attachment". In their model each new node adds exactly c new edges, each of which connects to (existing) node i with probability $d(i)/(\sum_v d(v))$, where $d(v)$ is the degree of node v. It can be shown directly, or by a careful reduction from Price's model (setting $a = c$), that this leads to a power law degree distribution with $\beta = 3$.

The model proposed here is a generalization of the one introduced in [1, 8] and we follow the formulation of Chung and Lu [6]. They analyzed an undirected network evolution with two types of preferential attachment events: a node event with probability p and an *edge event* with probability $1-p$. As discussed later on, this leads to a power law with $\beta = \frac{4-p}{2-p} \in (2,3]$.

A different type of models are the *copying* based models [15, 16]. In the simple copy model, every new node that joins the network adds a single directed edge. It selects an existing node u uniformly at random and then connects to u with probability $1-p$ or to the single node that u points to with probability p. It can then be shown that the in-degree distribution follows a power law with exponent $\beta = 1 + \frac{1}{p} \in (2, \infty)$.

While models with time varying probabilities were considered before [9, 11–14], they have not captured the full range $(1, \infty)$.

The exponent β of the power law has a significant impact on the networks structure. In [2] we studied the core size in preferential attachment models. It can be shown that for $\beta = 2$, the core size becomes *sublinear* in the network size, which may have implications on the size of elites in social networks [3].

Additional study of large scale social networks has revealed several other universal properties, for example the "small-world" phenomena, short average path lengths, navigability and high clustering coefficients.

The rest of the paper is organized as follows. Section 2 presents the model formally. Section 3 establishes improved concentration bounds for $m_{k,t}$, Section 4 establishes bounds for $\mathbb{E}\left[m_{k,t}\right]$, and section 5 deals with power laws for degree distributions, proves their existence, and analyzes their exponent.

2 THE GENERALIZED $G(p_t, r_t, q_t)$ MODEL

In this section we describe a generalization of the preferential attachment model $G(p)$ defined in Ch. 3.1 of [6]. $G(p)$ has one parameter, p, assumed therein to be constant with t. Here we allow p to vary with t. Let us now introduce the model $G(p_t, r_t, q_t)$ formally.

Consider a sequence $(G_t)_{t=1}^{\infty}$ of graphs, $G_t = (V_t, E_t)$, where V_t (respectively, E_t) denotes the set of nodes (resp. edges) in G_t and $n_t = |V_t|$ (resp. $e_t = |E_t|$). Let $d_t(w)$ denote the degree of node w in G_t. Starting from an arbitrary initial nonempty graph G_0, for $t \geq 1$ the graph G_t is constructed from G_{t-1} at time t by performing either a *node event* with probability $p_t \in [0,1]$, an *edge event* with probability r_t, or a *component event* with probability $q_t = 1 - p_t - r_t$.

In a node event, a new vertex v is added to the graph, along with a new edge selected using *preferential attachment*. Formally, the new edge is (v, u), where $u \in V_{t-1}$ is chosen with probability proportional to its degree, i.e., $\gamma_t(u) = d_{t-1}(u)/(\sum_{w \in V_{t-1}} d_{t-1}(w))$. In an edge event, a new edge (u, w) is added, where both $u, w \in V_{t-1}$ are chosen by preferential attachment, independently of each other, i.e., (u, w) is chosen with probability $\gamma_t(u) \cdot \gamma_t(w)$. Finally, in a component event, two new nodes v and u are added, along with a new edge (v, u). Note that exactly one edge is added at each time step, so the number of edges in the graph G_t is $e_t = e_0 + t$ and the sum of degrees is $2e_t = 2(e_0 + t)$.[1] We also have

$$\mathbb{E}[n_t] = n_0 + \sum_{j=1}^{t} (p_j + 2q_j).$$

3 CONCENTRATION BOUNDS FOR $m_{k,t}$

We now establish Thm. 1.4. To do that, we prove the following more general theorem.

For integer $t \geq 0$ and reals $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_{2e_t}) \in \mathbb{R}^{2e_t}$, let $\Delta_{\boldsymbol{\varepsilon}} = \max_{i,j} |\varepsilon_i - \varepsilon_j|$ be the maximum distance between any two coordinates of $\boldsymbol{\varepsilon}$, and $\|\boldsymbol{\varepsilon}\|_\infty = \max_i |\varepsilon_i|$ be the maximum distance between a coordinate and 0. Also let $S_t(\boldsymbol{\varepsilon}) = \sum_{k=1}^{2e_t} \varepsilon_k \cdot m_{k,t}$.

Theorem 3.1. *In the $G(p_t, r_t, q_t)$ preferential attachment model, for every integer $t \geq 0$, real $c \geq 0$, and reals $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{2e_t}) \in \mathbb{R}^{2e_t}$, we have*

$$\mathbb{P}\left[S_t(\boldsymbol{\varepsilon}) - \mathbb{E}[S_t(\boldsymbol{\varepsilon})] \geq \sqrt{2ct} \cdot (2\Delta_{\boldsymbol{\varepsilon}} + 2\max(\Delta_{\boldsymbol{\varepsilon}}, \|\boldsymbol{\varepsilon}\|_\infty))\right] \leq e^{-c},$$
$$\mathbb{P}\left[S_t(\boldsymbol{\varepsilon}) - \mathbb{E}[S_t(\boldsymbol{\varepsilon})] \leq -\sqrt{2ct} \cdot (2\Delta_{\boldsymbol{\varepsilon}} + 2\max(\Delta_{\boldsymbol{\varepsilon}}, \|\boldsymbol{\varepsilon}\|_\infty))\right] \leq e^{-c}.$$

Note that Thm. 3.1 suggests a deviation of $O(\sqrt{t})$ from the mean. This can be obtained by setting all the coordinates in $\boldsymbol{\varepsilon}$ to zero except for a single ε_k which is set to 1, yielding the following.

Corollary 3.2. *In the model $G(p_t, r_t, q_t)$, for every integers $k \geq 0, t \geq 0$ and real $c \geq 0$, we have*

$$\begin{aligned}
\mathbb{P}\left[m_{k,t} - \mathbb{E}\left[m_{k,t}\right] \geq \sqrt{32ct}\right] &\leq e^{-c}, \\
\mathbb{P}\left[m_{k,t} - \mathbb{E}\left[m_{k,t}\right] \leq -\sqrt{32ct}\right] &\leq e^{-c},
\end{aligned}$$

[1] Here a self loop increments the degree of a node by two.

Note that for $k > 2e_t$, we have $m_{k,t} = 0$ and Cor. 3.2 follows trivially. Thm. 3.1 also implies concentration bounds for $RC_{k,t}$. This can be obtained by using a vector $\boldsymbol{\varepsilon}$ in which $\varepsilon_i = 0$ for $i < k$ and $\varepsilon_i = 1$ for $i \geq k$, yielding the following.

Corollary 3.3. *In the model $G(p_t, r_t, q_t)$, for every integers $k \geq 0, t \geq 0$ and real $c \geq 0$, we have*

$$\begin{aligned}
\mathbb{P}\left[RC_{k,t} - \mathbb{E}\left[RC_{k,t}\right] \geq \sqrt{32ct}\right] &\leq e^{-c}, \\
\mathbb{P}\left[RC_{k,t} - \mathbb{E}\left[RC_{k,t}\right] \leq -\sqrt{32ct}\right] &\leq e^{-c},
\end{aligned}$$

Finally, note that Thm. 3.1 implies Thm. 1.4, Cor. 3.2 implies Cor. 1.5, and Cor. 3.3 implies Cor. 1.6 as a special case since the model $G(p)$ can be obtained from the model $G(p_t, r_t, q_t)$ by setting $p_t = p$ and $q_t = 0$.

The rest of this section is dedicated to proving Thm. 3.1.

3.1 Recursive Relation for $m_{k,t}$

As a first step, we find a recursive representation for $m_{k,t+1}$ in terms of $m_{k',t}$ for smaller $k' \leq k$. Note that $m_{0,t} = m_{0,0}$ for every t. Also note that the number of edges at time t is $e_t = e_0 + t$. Hence, since the sum of the degrees at time t equals $2e_t$, we have $m_{k,t} = 0$ for $k > 2e_t$. Therefore, we focus on $m_{k,t}$ for $1 \leq k \leq 2e_t$. Let $\mathcal{F}_t$ denote the σ-algebra generated by the graphs $G_0, \ldots, G_t$. Intuitively, $\mathcal{F}_t$ encodes all the history until time t. Fix $k \geq 2$. Since $0 \leq d_{t+1}(v) - d_t(v) \leq 2$ for every node v and time t, we have

$$\mathbb{E}\left[m_{k,t+1} \mid \mathcal{F}_t\right] = \sum_{\{v \,:\, k-2 \leq d_t(v) \leq k\}} \mathbb{P}[d_{t+1}(v) = k]\,. \tag{5}$$

Recalling that $\gamma_t(v) = d_{t-1}(v)/\sum_{w \in V_{t-1}} d_{t-1}(w)$, for every node v such that $d_{t+1}(v) = k$, there are at most three possible values for $d_t(v)$:

(i) $d_t(v) = k - 2$ and at time $t + 1$ there was an edge event where the edge (v, v) was added (this happens with probability $r_{t+1} \cdot (\gamma_{t+1}(v))^2$).
(ii) $d_t(v) = k - 1$ and at time $t + 1$ there was either a node event involving v (this happens with probability $p_{t+1} \cdot \gamma_{t+1}(v)$) or an edge event where the edge (v, w) $w \neq v$ was added (this happens with probability $r_{t+1} \cdot 2\gamma_{t+1}(v) \cdot (1 - \gamma_{t+1}(v))$).
(iii) $d_t(v) = k$ and at time $t+1$ there was either a node event where v was not involved (this happens with probability $p_{t+1} \cdot (1 - \gamma_{t+1}(v))$), an edge event where the edge (u, w) $v \notin \{u, w\}$ was added (this happens with probability $r_{t+1} \cdot (1 - \gamma_{t+1}(v))^2$) or a component event (this happens with probability q_{t+1}).

Letting $\alpha_{k,t} = k/(2e_t)$, for every v such that $d_t(v) = i$ we have $\gamma_{t+1}(v) = \alpha_{i,t}$. Denoting

$$\begin{aligned}
A_{k,t} &= p_{t+1}(1 - \alpha_{k,t}) + r_{t+1} \cdot (1 - \alpha_{k,t})^2 + q_{t+1}\,, \\
B_{k,t} &= p_{t+1} \cdot \alpha_{k,t} + 2 \cdot r_{t+1} \cdot \alpha_{k,t} \cdot (1 - \alpha_{k,t})\,, \\
C_{k,t} &= r_{t+1} \cdot \alpha_{k,t}^2\,,
\end{aligned}$$

we have

$$A_{k,t} + B_{k,t} + C_{k,t} = 1 \tag{6}$$

and $A_{k,t}, B_{k,t}, C_{k,t} \geq 0$ for every $0 \leq k \leq 2e_t$. Furthermore, by Eq. (5), for every $k \geq 2$ we have

$$\mathbb{E}\left[m_{k,t+1} \mid \mathcal{F}_t\right] = m_{k,t} \cdot A_{k,t} + m_{k-1,t} \cdot B_{k-1,t} + C_{k-2,t} \cdot m_{k-2,t}\,. \tag{7}$$

Similarly, for $k = 1$ we have

$$\mathbb{E}\left[m_{1,t+1} \mid \mathcal{F}_t\right] = A_{1,t} \cdot m_{1,t} + p_{t+1} + 2q_{t+1} \,. \tag{8}$$

Letting

$$X_{k,t} = \begin{cases} m_{k-1,t} \cdot B_{k-1,t} + m_{k-2,t} \cdot C_{k-2,t} & k \geq 2\,, \\ p_{t+1} + 2q_{t+1} & k = 1\,, \end{cases}$$

Eq. (7) and (8) can be rewritten as

$$\mathbb{E}\left[m_{k,t+1} \mid \mathcal{F}_t\right] = m_{k,t} \cdot A_{k,t} + X_{k,t} \,. \tag{9}$$

Taking expectation on both sides, we get

$$\mathbb{E}\left[m_{k,t+1}\right] = \mathbb{E}\left[m_{k,t}\right] \cdot A_{k,t} + \mathbb{E}\left[X_{k,t}\right] \,. \tag{10}$$

3.2 Analysis of the Martingale $\mathbb{E}\left[m_{k,t} \mid \mathcal{F}_\ell\right]$

For integers $0 \leq \ell \leq t$ and $0 \leq k \leq 2e_t$, denote $Y_\ell^{k,t} = \mathbb{E}\left[m_{k,t} \mid \mathcal{F}_\ell\right]$, where the expectation is taken over all that occurred from time $\ell + 1$ till time t. By the law of total expectation, we have

$$\begin{aligned} \mathbb{E}\left[Y_\ell^{k,t} \mid \mathcal{F}_{\ell-1}\right] &= \mathbb{E}\left[\mathbb{E}\left[m_{k,t} \mid \mathcal{F}_\ell\right] \mid \mathcal{F}_{\ell-1}\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[m_{k,t} \mid \mathcal{F}_\ell, \mathcal{F}_{\ell-1}\right] \mid \mathcal{F}_{\ell-1}\right] \\ &= \mathbb{E}\left[m_{k,t} \mid \mathcal{F}_{\ell-1}\right] \;=\; Y_{\ell-1}^{k,t} \,. \end{aligned}$$

Therefore, $Y_0^{k,t}, Y_1^{k,t}, \ldots, Y_t^{k,t}$ is a martingale (with respect to the filtration $\mathcal{F}_0 \subseteq \cdots \subseteq \mathcal{F}_t$). This general construction is commonly referred to as a Doob Martingale [10]. Therefore, for a vector $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_{2e_t}) \in \mathbb{R}^{2e_t}$, letting

$$Y_\ell^t(\boldsymbol{\varepsilon}) \;=\; \sum_{k=1}^{2e_t} \varepsilon_k \cdot Y_\ell^{k,t} \;=\; \sum_{k=1}^{2e_t} \varepsilon_k \cdot \mathbb{E}\left[m_{k,t} \mid \mathcal{F}_\ell\right],$$

we have that $Y_0^t(\boldsymbol{\varepsilon}), Y_1^t(\boldsymbol{\varepsilon}) \ldots, Y_t^t(\boldsymbol{\varepsilon})$ is also a Doob martingale (with respect to the same filtration $\mathcal{F}_0 \subseteq \cdots \subseteq \mathcal{F}_t$). We next bound the difference between consecutive elements in the sequence. For $1 \leq \ell \leq t$ and reals $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_{2e_t}) \in \mathbb{R}^{2e_t}$, let

$$\Delta_\ell^t(\boldsymbol{\varepsilon}) \;=\; |Y_\ell^t(\boldsymbol{\varepsilon}) - Y_{\ell-1}^t(\boldsymbol{\varepsilon})|.$$

Lemma 3.4. *For every $1 \leq \ell \leq t$ and $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_{2e_t})$, we have*

$$\Delta_\ell^t(\boldsymbol{\varepsilon}) \leq 2\Delta_{\boldsymbol{\varepsilon}} + 2\max(\Delta_{\boldsymbol{\varepsilon}}, \|\boldsymbol{\varepsilon}\|_\infty).$$

Proof. Fixing ℓ, the proof is by induction on t. For $t = \ell$, letting

$$\begin{aligned} \Psi_1 &= \sum_{k=1}^{2e_\ell} \varepsilon_k \cdot (m_{k,\ell} - m_{k,\ell-1})\,, \\ \Psi_2 &= \varepsilon_1 \cdot (p_\ell + 2q_\ell) + \sum_{k=1}^{2e_\ell} \varepsilon_k \cdot ((1 - A_{k,\ell-1})m_{k,\ell-1} - X_{k,\ell-1})\,, \end{aligned}$$

by Eq. (9) we have

$$\begin{aligned} \Delta_\ell^\ell(\boldsymbol{\varepsilon}) &= \left|\sum_{k=1}^{2e_\ell} \varepsilon_k \cdot \left(\mathbb{E}\left[m_{k,\ell} \mid \mathcal{F}_\ell\right] - \mathbb{E}\left[m_{k,\ell} \mid \mathcal{F}_{\ell-1}\right]\right)\right| \\ &= \left|\sum_{k=1}^{2e_\ell} \varepsilon_k \cdot (m_{k,\ell} - A_{k,\ell-1} \cdot m_{k,\ell-1} - X_{k,\ell-1})\right| \\ &= |-\varepsilon_1 \cdot (p_\ell + 2q_\ell) + \Psi_1 + \Psi_2|\,. \end{aligned} \tag{11}$$

Note that

$$\Psi_1 = \begin{cases} 2\varepsilon_1, & \text{At time } \ell \text{ a compenent event occurred,} \\ \varepsilon_1 + \varepsilon_{i+1} - \varepsilon_i, & \text{At time } \ell \text{ a new node was connected to an existing node of degree } i, \\ \varepsilon_{i+2} - \varepsilon_i, & \text{At time } \ell \text{ a self loop was added to a node of degree } i, \\ \varepsilon_{i+1} - \varepsilon_i + \varepsilon_{j+1} - \varepsilon_j, & \text{At time } \ell \text{ a new edge was added between two distinct existing nodes with degrees } i \text{ and } j. \end{cases}$$

Therefore, we have

$$\begin{aligned} |\Psi_1 - \varepsilon_1 \cdot (p_\ell + 2q_\ell)| &\leq \max(2|\varepsilon_1|, |\varepsilon_1| + \Delta_{\boldsymbol{\varepsilon}}, (p_\ell + 2q_\ell)|\varepsilon_1| + 2\Delta_{\boldsymbol{\varepsilon}}) \\ &\leq (p_\ell + 2q_\ell)\Delta_{\boldsymbol{\varepsilon}} + 2\max(|\varepsilon_1|, \Delta_{\boldsymbol{\varepsilon}})\,. \end{aligned}$$

For Ψ_2, recalling that $m_{k,\ell-1} = 0$ for every $k > 2e_{\ell-1}$, and by Eq. (6) we obtain

$$\begin{aligned} |\Psi_2| &= \left|\sum_{k=1}^{2e_\ell} m_{k,\ell-1}\left((1 - A_{k,\ell-1})\varepsilon_k - B_{k,\ell-1}\varepsilon_{k+1} - C_{k,\ell-1}\varepsilon_{k+2}\right)\right| \\ &= \left|\sum_{k=1}^{2e_\ell} m_{k,\ell-1}\left(B_{k,\ell-1}(\varepsilon_k - \varepsilon_{k+1}) + C_{k,\ell-1}(\varepsilon_k - \varepsilon_{k+2})\right)\right| \\ &\leq \sum_{k=1}^{2e_\ell} m_{k,\ell-1}\left(B_{k,\ell-1} \cdot \Delta_{\boldsymbol{\varepsilon}} + C_{k,\ell-1} \cdot 2\Delta_{\boldsymbol{\varepsilon}}\right) \\ &= (p_\ell + 2r_\ell)\Delta_{\boldsymbol{\varepsilon}} \sum_{k=1}^{2e_\ell} \alpha_{k,\ell-1} m_{k,\ell-1} \;=\; (p_\ell + 2r_\ell)\Delta_{\boldsymbol{\varepsilon}}\,. \end{aligned}$$

Plugging this into Eq. (11), we obtain

$$\begin{aligned} \Delta_\ell^\ell(\boldsymbol{\varepsilon}) &\leq |\Psi_1 - \varepsilon_1 \cdot (p_\ell + 2q_\ell)| + |\Psi_2| \\ &\leq (2p_\ell + 2r_\ell + 2q_\ell)\Delta_{\boldsymbol{\varepsilon}} + 2\max\{|\varepsilon_1|, \Delta_{\boldsymbol{\varepsilon}}\} \\ &= 2\Delta_{\boldsymbol{\varepsilon}} + 2\max\{|\varepsilon_1|, \Delta_{\boldsymbol{\varepsilon}}\} \;\leq\; 2\Delta_{\boldsymbol{\varepsilon}} + 2\max\{\|\boldsymbol{\varepsilon}\|_\infty, \Delta_{\boldsymbol{\varepsilon}}\}\,. \end{aligned}$$

Hence the induction basis holds. Next, assuming the claim for t, we prove it for $t + 1$. By the law of total expectation and Eq. (9),

$$\begin{aligned} Y_\ell^{t+1}(\boldsymbol{\varepsilon}) &= \sum_{k=1}^{2e_{t+1}} \varepsilon_k \cdot \mathbb{E}\left[m_{k,t+1} \mid \mathcal{F}_\ell\right] \\ &= \sum_{k=1}^{2e_{t+1}} \varepsilon_k \cdot \mathbb{E}\left[\mathbb{E}\left[m_{k,t+1} \mid \mathcal{F}_t\right] \mid \mathcal{F}_\ell\right] \\ &= \sum_{k=1}^{2e_{t+1}} \varepsilon_k \cdot \mathbb{E}\left[A_{k,t} \cdot m_{k,t} + X_{k,t} \mid \mathcal{F}_\ell\right] \\ &= \varepsilon_1 \cdot (p_t + 2q_t) + \sum_{k=1}^{2e_t} \tilde{\varepsilon}_k \cdot \mathbb{E}\left[m_{k,t} \mid \mathcal{F}_\ell\right] \\ &= \varepsilon_1 \cdot (p_t + 2q_t) + Y_\ell^t(\tilde{\boldsymbol{\varepsilon}})\,, \end{aligned}$$

where $\tilde{\varepsilon}_k = \varepsilon_k \cdot A_{k,t} + \varepsilon_{k+1} \cdot B_{k,t} + \varepsilon_{k+2} \cdot C_{k,t}$ and $\tilde{\boldsymbol{\varepsilon}} = (\tilde{\varepsilon}_1, \ldots, \tilde{\varepsilon}_{2e_t})$.[2] Therefore, by the induction hypothesis we have

$$\begin{aligned}\Delta_\ell^{t+1}(\boldsymbol{\varepsilon}) &= \left|Y_\ell^{t+1}(\boldsymbol{\varepsilon}) - Y_{\ell-1}^{t+1}(\boldsymbol{\varepsilon})\right| = \left|Y_\ell^t(\tilde{\boldsymbol{\varepsilon}}) - Y_{\ell-1}^t(\tilde{\boldsymbol{\varepsilon}})\right| \\ &= \Delta_\ell^t(\tilde{\boldsymbol{\varepsilon}}) \leq 2\Delta_{\tilde{\boldsymbol{\varepsilon}}} + 2\max(\|\tilde{\varepsilon}\|_\infty, \Delta_{\tilde{\boldsymbol{\varepsilon}}}) \\ &\leq 2\Delta_{\boldsymbol{\varepsilon}} + 2\max(\|\boldsymbol{\varepsilon}\|_\infty, \Delta_{\boldsymbol{\varepsilon}}) ,\end{aligned}$$

noting that every coordinate of $\tilde{\boldsymbol{\varepsilon}}$ is a convex combination of three coordinates of $\boldsymbol{\varepsilon}$, so $\|\tilde{\boldsymbol{\varepsilon}}\|_\infty \leq \|\boldsymbol{\varepsilon}\|_\infty$ and $\Delta_{\tilde{\boldsymbol{\varepsilon}}} \leq \Delta_{\boldsymbol{\varepsilon}}$. The lemma follows. □

Proof of Thm. 3.1: By Lemma 3.4, we may apply Azuma's inequality to the martingale $Y_0^t(\boldsymbol{\varepsilon}), Y_1^t(\boldsymbol{\varepsilon}), \ldots, Y_t^t(\boldsymbol{\varepsilon})$, and obtain for every $\lambda \geq 0$

$$\mathbb{P}[Y_t^t(\boldsymbol{\varepsilon}) - Y_0^t(\boldsymbol{\varepsilon}) \geq \lambda] \leq \exp\left(-\frac{\lambda^2}{2t \cdot (2\Delta_{\boldsymbol{\varepsilon}} + 2\max(\Delta_{\boldsymbol{\varepsilon}}, \|\boldsymbol{\varepsilon}\|_\infty))^2}\right) ,$$

$$\mathbb{P}[Y_t^t(\boldsymbol{\varepsilon}) - Y_0^t(\boldsymbol{\varepsilon}) \leq -\lambda] \leq \exp\left(-\frac{\lambda^2}{2t \cdot (2\Delta_{\boldsymbol{\varepsilon}} + 2\max(\Delta_{\boldsymbol{\varepsilon}}, \|\boldsymbol{\varepsilon}\|_\infty))^2}\right) .$$

Noting that $Y_t^t(\boldsymbol{\varepsilon}) = \sum_{k=1}^{2e_t} \varepsilon_k \cdot m_{k,t}$ and $Y_0^t = \sum_{k=1}^{2e_t} \varepsilon_k \cdot \mathbb{E}\left[m_{k,t}\right]$ and letting $\lambda = \sqrt{2ct} \cdot (2\Delta_{\boldsymbol{\varepsilon}} + 2\max(\Delta_{\boldsymbol{\varepsilon}}, \|\boldsymbol{\varepsilon}\|_\infty))$, Thm. 3.1 follows. □

4 IMPROVED BOUNDS FOR $\mathbb{E}\left[m_{k,t}\right]$

We now prove Thm. 1.2 for $G(p)$. Consider the model $G(p)$, which can be derived from the model $G(p_t, r_t, q_t)$ by setting $p_t = p$ and $q_t = 0$ for every t. Thus we obtain the simplified expressions

$$\begin{aligned}A_{k,t} &= 1 - \frac{(2-p)k}{2e_t} + (1-p) \cdot \frac{k^2}{4e_t^2} , \\ B_{k,t} &= \frac{(2-p)k}{2e_t} - 2(1-p) \cdot \frac{k^2}{4e_t^2} , \\ C_{k,t} &= (1-p) \cdot \frac{k^2}{4e_t^2} .\end{aligned}$$

Recall that $\beta = \frac{4-p}{2-p}$ and $M_k = \frac{2}{4-p} \prod_{j=1}^{k-1} \frac{j}{j+\beta}$.

Proof of Thm. 1.2 : To prove Ineq. (1), letting

$$\ell_{k,t} = \mathbb{E}\left[m_{k,t}\right] - \left(pe_t - \frac{p^2(1-p)}{2(2-p)^2}\right) \cdot M_k - \frac{c^*}{k} ,$$

it suffices to show that $\ell_{k,t} \leq 0$ for every t and $1 \leq k \leq 2e_t$. We show this by induction on t. For $t = 0$, since $e_t \geq 1$, we have

$$k \cdot \ell_{k,0} \leq k \cdot \mathbb{E}\left[m_{k,0}\right] - c^* \leq k \cdot \mathbb{E}\left[m_{k,0}\right] - 2e_0 \leq 0 ,$$

where the last inequality follows since $\sum_{k=1}^{2e_0} k \cdot m_{k,0} = 2e_0$, hence the basis of the induction holds. Assuming the claim for t, we prove it for $t + 1$. For $k = 1$, by Eq. (10), it can be shown that

$$\begin{aligned}\ell_{1,t+1} &= A_{1,t} \cdot \ell_{1,t} + \frac{p(1-p)M_1}{4e_t} \\ &- \left(c^* - \frac{p^2(1-p)M_1}{2(2-p)^2}\right)(1 - A_{1,t}) \\ &\leq 0 + \frac{p(1-p)M_1}{4e_t} - \left(c^* - \frac{p(1-p)M_1}{2}\right) \cdot \frac{1}{2e_t} \\ &= \frac{p(1-p)M_1 - c^*}{2e_t} \leq \frac{1-2e_0}{2e_t} \leq 0 .\end{aligned}$$

For $2 \leq k \leq 2e_t$, by Eq. (10), it can be shown that

$$\begin{aligned}\ell_{k,t+1} &= A_{k,t} \cdot \ell_{k,t} + B_{k-1,t} \cdot \ell_{k-1,t} + C_{k-2,t} \cdot \ell_{k-2,t} \\ &\quad - \frac{p(1-p)M_k}{2(2-p)^2 e_t}\left(\frac{p^2(1-p)}{2(2-p)^2 e_t} + \frac{4-p}{k-1}\left(1 - \frac{p(1-p)}{2(2-p)^2 e_t}\right)\right) \\ &\leq A_{k,t} \cdot \ell_{k,t} + B_{k-1,t} \cdot \ell_{k-1,t} + C_{k-2,t} \cdot \ell_{k-2,t} \leq 0 .\end{aligned}$$

For $2e_{t+1} - 1 \leq k \leq 2e_{t+1}$, we have

$$k \cdot \ell_{k,t+1} \leq k \cdot \mathbb{E}\left[m_{k,t+1}\right] - c^* .$$

Note that for $k = 2e_{t+1}$ and $m_{k,t+1} = 1$, necessarily the initial graph G_0 is composed of a single node (of degree $2e_0$) and only edge events occur until time $t + 1$, thus

$$\begin{aligned}k \cdot \ell_{k,t+1} &\leq 2e_{t+1} \cdot (1-p)^{t+1} - c^* \\ &\leq \max\left\{2e_1(1-p) , \frac{2}{-\ln(1-p)}\right\} - c^* \leq 0 ,\end{aligned}$$

where the second inequality follows by maximizing over the range $t \geq 0$. Finally, for $k = 2e_{t+1} - 1$ and $m_{k,t+1} > 0$, we analyze the two possible cases. In the first case, the initial graph G_0 is composed of two nodes, one of degree 1 and the other of degree $2e_0 - 1$. Therefore, necessarily only edge events occur until time $t + 1$, thus

$$k \cdot \ell_{k,t+1} \leq (2e_{t+1} - 1) \cdot (1-p)^{t+1} - c^* \leq 0 ,$$

as before. In the second case, the initial graph G_0 is composed of a single node v_0 (of degree $2e_0$). Therefore, exactly t edge events (involving the edge (v_0, v_0)) and one node event occur until time $t + 1$. Therefore,

$$\begin{aligned}k \cdot \ell_{k,t+1} &\leq (2e_{t+1} - 1) \cdot (1-p)^t \cdot p \cdot t - c^* \\ &\leq 2e_{t+1}^2 \cdot (1-p)^t \cdot p - c^* \\ &\leq \max\left\{2e_1^2 \cdot p , \frac{8p}{(\ln(1-p))^2}\right\} - c^* \leq 0 ,\end{aligned}$$

where the third inequality follows by maximizing over the range $t \geq 0$, thus concluding the proof of Ineq. (1). To prove Ineq. (2), letting

$$Z_{k,t} = p(e_t + 1 - p)M_k - \frac{p(e_0 + 1 - p)M_1}{k} - \mathbb{E}\left[m_{k,t}\right] ,$$

it suffices to show that $Z_{k,t} \leq 0$ for every $t \geq 0$ and $1 \leq k \leq 2e_t$. We show this by induction on t. For $t = 0$, since $\beta > 1$, we have

$$\frac{Z_{k,0}}{p(e_0 + 1 - p)} \leq M_k - \frac{M_1}{k} \leq -\frac{M_1}{k} + M_1 \prod_{j=1}^{k-1} \frac{j}{j+1} = 0 .$$

[2] Note that $\boldsymbol{\varepsilon} \in \mathbb{R}^{2e_{t+1}}$ and $\tilde{\boldsymbol{\varepsilon}} \in \mathbb{R}^{2e_t}$, thus $Y_\ell^{t+1}(\boldsymbol{\varepsilon})$ and $Y_\ell^t(\tilde{\boldsymbol{\varepsilon}})$ are well defined.

Assuming the claim for t, we prove for $t+1$. For $k=1$, by Eq. (10), it can be shown that

$$\begin{aligned} Z_{1,t+1} &= A_{1,t}\cdot Z_{1,t} - \frac{p(1-p)}{2e_t(4-p)} - p\cdot e_0\cdot M_1\left(1-A_{1,t}\right) \\ &\le A_{1,t}\cdot Z_{1,t} - 0 - 0 \le 0 . \end{aligned}$$

For $2\le k\le 2e_t$, by Eq. (10), it can be shown that

$$\begin{aligned} Z_{k,t+1} &= A_{k,t}\cdot Z_{k,t} + B_{k-1,t}\cdot Z_{k-1,t} + C_{k-2,t}\cdot Z_{k-2,t} \\ &+ \frac{p(1-p)M_k}{e_t}\left(-1+\frac{1+(1-p)/e_t}{2(2-p)^2}\left(\frac{4-p}{k-1}-p\right)\right) \\ &\le 0+0+0+\frac{p(1-p)M_k}{e_t}\left(-1+\frac{1+(1-p)/e_t}{2-p}\right) \\ &\le \frac{p(1-p)M_k}{e_t}\left(-1+\frac{1+(1-p)}{2-p}\right) = 0 , \end{aligned}$$

and for $2e_{t+1}-1\le k\le 2e_{t+1}$, since $\beta\ge 2$, we have

$$\begin{aligned} Z_{k,t+1} &\le p(e_{t+1}+1-p)M_k - \frac{p(e_0+1-p)M_1}{k} \\ &\le p(e_{t+1}+1-p)M_1\prod_{j=1}^{k-1}\frac{j}{j+2} - \frac{p(e_0+1-p)M_1}{k} \\ &= \frac{pM_1}{k}\left(\frac{2e_{t+1}}{k+1} - e_0 + (1-p)\left(\frac{2}{k+1}-1\right)\right) \\ &\le \frac{pM_1}{k}\left(1-1+(1-p)(1-1)\right) = 0 , \end{aligned}$$

thus concluding the proof of Ineq. (2). Theorem 1.2 follows. □

We now prove Cor. 1.3 using Thm. 1.2.

Proof of Cor. 1.3 : By Ineq. (1),

$$\begin{aligned} \mathbb{E}\left[RC_{k,t}\right] &= \sum_{j=k}^{2e_t}\mathbb{E}\left[m_{j,t}\right] \le p\cdot e_t\sum_{j=k}^{2e_t}M_j + c^*\sum_{j=k}^{2e_t}1/j \\ &\le c^*\cdot\ln\left(\frac{2e_t}{k-1}\right) + p\cdot e_t\sum_{j=k}^{2e_t}M_j . \end{aligned}$$

Noting that $M_j=\varphi_j-\varphi_{j+1}$, Ineq. (3) follows. The proof of Ineq. (4) is similar. □

5 POWER LAW IN $G(p_t,r_t,q_t)$

In this section, we show that the in the $G(p_t,r_t,q_t)$ model, it is possible to obtain networks obeying a power law degree distribution with any exponent in the range $(1,\infty)$, thus expanding the range of $(2,3]$ obtainable in the standard $G(p)$ model (see Lemma 1.1).

5.1 Existence of Power Law

The following lemma can be found in [6].

Lemma 5.1. *[6] Suppose that a sequence $\{a_t\}$ satisfies the recurrence relation*

$$a_{t+1} = \left(1-\frac{b_t}{t+t_1}\right)a_t + c_t \text{ for } t\ge t_0 .$$

Furthermore, suppose $\lim_{t\to\infty} b_t = b > 0$ and $\lim_{t\to\infty} c_t = c$. Then $\lim_{t\to\infty} a_t/t$ exists and $\lim_{t\to\infty} a_t/t = c/(1+b)$.

We use Lemma 5.1 to prove the following generalized lemma.

Lemma 5.2. *Suppose that a sequence $\{a_t\}$ satisfies the recurrence relation*

$$a_{t+1} = \left(1-\frac{b_t}{t+t_1}\right)a_t + c_t \text{ for } t\ge t_0 . \tag{12}$$

Furthermore, letting $\{s_t\}$ be a sequence of reals satisfying $\lim_{t\to\infty} s_t/s_{t+1}=1$, and $d_t = t(1-s_t/s_{t+1})$, suppose that $\lim_{t\to\infty} b_t = b$, $\lim_{t\to\infty} c_t\cdot t/s_t = c$, $\lim_{t\to\infty} d_t = d$ and $b+d>1$. Then $\lim_{t\to\infty} a_t/s_t$ exists and $\lim_{t\to\infty} a_t/s_t = c/(b+d)$.

Note that Lemma 5.1 is a special case of Lemma 5.2 obtained by setting $s_t=t$.

Proof. Denote

$$\begin{aligned} a_t^* &= \frac{a_t\cdot t}{s_t} , \\ b_t^* &= \frac{t+t_1}{t}\cdot\left(d_t - \frac{s_t}{s_{t+1}}\right) + \frac{t+1}{t}\cdot\frac{s_t}{s_{t+1}}\cdot b_t , \\ c_t^* &= c_t\cdot\frac{t+1}{s_{t+1}} . \end{aligned}$$

By Eq. (12),

$$\begin{aligned} a_{t+1}^* &= \frac{a_{t+1}\cdot(t+1)}{s_{t+1}} = \left(1-\frac{b_t}{t+t_1}\right)a_t\cdot\frac{t+1}{s_{t+1}} + c_t\cdot\frac{t+1}{s_{t+1}} \\ &= \left(1-\frac{b_t^*}{t+t_1}\right)a_t^* + c_t^* . \end{aligned}$$

Since $\lim_{t\to\infty} b_t^* = b+d-1>0$, and $\lim_{t\to\infty} c_t^* = c$, we may apply Lemma 5.1, and obtain $\lim_{t\to\infty}\frac{a_t}{s_t} = \lim_{t\to\infty}\frac{a_t^*}{t} = \frac{c}{b+d}$. □

The following theorem and corollary complete the proof of a power law in $G(p_t,r_t,q_t)$.

Theorem 5.3. *Consider $G(p_t,r_t,q_t)$ and let $y_t=p_t+2q_t$. Assume that $\lim_{t\to\infty} y_t = y < 2$, $\sum_{t=1}^{\infty} y_t=\infty$ and $\lim_{t\to\infty} t\cdot y_{t+1}/\sum_{j=1}^{t} y_j = \Gamma > 0$. Then letting $\beta = 1+\frac{2\Gamma}{2-y}$, the limit $M_k = \lim_{t\to\infty}\frac{\mathbb{E}[m_{k,t}]}{\mathbb{E}[n_t]}$ exists for every $k\ge 1$ and*

$$M_k = \frac{\Gamma}{\Gamma+1-y/2}\cdot\prod_{j=1}^{k-1}\frac{j}{j+\beta} . \tag{13}$$

Proof. By induction on k. For $k=1$, consider Eq. (10). By applying Lemma 5.2 with (t_1,s_t,a_t,b_t,c_t) set to $(e_0,\mathbb{E}[n_t],\mathbb{E}[m_{1,t}],e_t(1-A_{1,t}),y_{t+1})$, yielding the limits $b=1-y/2$, and $c=d=\Gamma$, Eq. (13) follows. Assuming the claim for $k-1$, we prove it for k. Consider Eq. (10). Applying Lemma 5.2 with (t_1,s_t,a_t,b_t,c_t) set to $(e_0,\mathbb{E}[n_t],\mathbb{E}\left[m_{k,t}\right],e_t(1-A_{k,t}),B_{k-1,t}\cdot\mathbb{E}\left[m_{k-1,t}\right]+C_{k-2,t}\cdot\mathbb{E}\left[m_{k-2,t}\right])$, we obtain the limits $d=\Gamma$, $b=k\cdot(1-y/2)$ and by the induction hypothesis,

$$c = \lim_{t\to\infty}\frac{c_t\cdot t}{s_t} = (k-1)\left(1-\frac{y}{2}\right)M_{k-1} .$$

Therefore, M_k exists and

$$M_k = \frac{(k-1)(1-y/2)M_{k-1}}{k\cdot(1-y/2)+\Gamma} = \frac{k-1}{k-1+\beta}\cdot M_{k-1} .$$

□

Corollary 5.4. *Under the assumptions stated in Thm. 5.3, M_k is proportional to $k^{-\beta}$.*

Proof. Consider Eq. (13). By the First Order Condition [5], a differentiable function f is convex if and only if $f(x_2) \geq f(x_1) + f'(x_1)(x_2 - x_1)$ for every x_1 and x_2. Applying this with $(f(z), x_1, x_2)$ set to $(z^\beta, 1, 1 + 1/j)$, we obtain

$$\prod_{j=1}^{k-1} \frac{j}{j+\beta} = \prod_{j=1}^{k-1} \left(1 + \frac{\beta}{j}\right)^{-1} \geq \prod_{j=1}^{k-1} \left(1 + \frac{1}{j}\right)^{-\beta} = k^{-\beta} .$$

Applying the First Order Condition with $(f(z), x_1, x_2)$ set to $(z^\beta, 1, 1 - 1/(j+\beta))$, we obtain

$$\begin{aligned}\prod_{j=1}^{k-1} \frac{j}{j+\beta} &= \prod_{j=1}^{k-1} \left(1 - \frac{\beta}{j+\beta}\right) \leq \prod_{j=1}^{k-1} \left(1 - \frac{1}{j+\beta}\right)^{\beta} \\ &= \left(\frac{\beta}{k-1+\beta}\right)^{\beta} \leq \beta^\beta \cdot k^{-\beta} .\end{aligned}$$

Therefore, $c_1 \cdot k^{-\beta} \leq M_k \leq c_2 \cdot k^{-\beta}$ for some positive constants c_1, c_2 and every $k \geq 1$. □

Note that $\lim_{t\to\infty} \mathbb{E}[n_t] = \infty$. Thus n_t is concentrated around $\mathbb{E}[n_t]$, yielding an expected power law with exponent β. Furthermore, for $y_t = \omega(1/\sqrt{t})$, we have $\mathbb{E}[n_t] = \omega(\sqrt{t})$. By Thm. 5.3, $\mathbb{E}\left[m_{k,t}\right] = \Omega(\mathbb{E}[n_t]) = \omega(\sqrt{t})$ for fixed k. Therefore, by Thm. 3.1, $m_{k,t}$ is concentrated around $\mathbb{E}\left[m_{k,t}\right]$, thus the power law is obtained almost surely.

5.2 Power Law Exponent Characterization

In Sect. 5.1 we showed that letting $y_t = p_t + 2q_t$, and whenever $\lim_{t\to\infty} y_t = y < 2$, $\sum_{t=1}^{\infty} y_t = \infty$ and $\lim_{t\to\infty} t \cdot y_{t+1} / \sum_{j=1}^{t} y_j = \Gamma > 0$, the resulting network obeys a power law with exponent $\beta = 1 + \frac{2\Gamma}{2-y}$. Under these assumptions, we have $\beta > 1$. The following lemma shows that any value of β in the range $(1, \infty)$ is attainable for some possible choice of the probabilities p_t, r_t and q_t.

Lemma 5.5. *For any $x \in (1, \infty)$, there exists a choice of p_t, r_t, q_t such that in $G(p_t, r_t, q_t)$ the resulting network follows a power law degree distribution with exponent $\beta = x$.*

Proof. Consider the following three cases.

(i) For $x \in (2, \infty)$, setting $y_t = 2 - 2/(x-1) \in (0, 2)$ for every t and applying Cor. 5.4, we obtain a power law with exponent $\beta = 1 + \frac{2}{2-(2-2/(x-1))} = x$.

(ii) For $x \in (1, 2)$, setting $y_t = t^{x-2}$ for every t, we have

$$\frac{t^{x-1}}{\int_{j=0}^{t} j^{x-2} dj} = \frac{(x-1) \cdot t^{x-1}}{[j^{x-1}]_{j=0}^{t}} = x - 1 .$$

Therefore,

$$\begin{aligned}\Gamma &= \lim_{t\to\infty} \frac{y_{t+1} \cdot t}{\sum_{j=1}^{t} y_j} = \lim_{t\to\infty} \frac{(t+1)^{x-2} \cdot t}{\sum_{j=1}^{t} j^{x-2}} = \lim_{t\to\infty} \frac{t^{x-1}}{\int_{j=0}^{t} j^{x-2} dj} \\ &= x - 1 .\end{aligned}$$

By Cor. 5.4, $G(p_t, r_t, q_t)$ follows a power law degree distribution with exponent
$\beta = 1 + 2\Gamma/(2 - y) = 1 + 2(x-1)/(2-0) = x$.

(iii) For $x = 2$, setting $y_t = 1/\ln(t+2)$ for every t, we have

$$\begin{aligned}\Gamma &= \lim_{t\to\infty} \frac{y_{t+1} \cdot t}{\sum_{j=1}^{t} y_j} = \lim_{t\to\infty} \frac{t/\ln(t+3)}{\sum_{j=1}^{t} 1/\ln(j+2)} \\ &= \lim_{t\to\infty} \frac{t/\ln(t+3)}{t/\ln t} = 1 .\end{aligned}$$

By Cor. 5.4, $G(p_t, r_t, q_t)$ follows a power law degree distribution with exponent
$\beta = 1 + 2\Gamma/(2 - y) = 1 + 2/(2 - 0) = 2$.

□

We now show that component events are needed in $G(p_t, r_t, q_t)$ in order to obtain a power law with exponent larger than 3, namely, in a model with no component events, the exponent β is restricted to the range $(1, 3]$.

Lemma 5.6. *Assume $\lim_{t\to\infty} y_t = y$ and $\lim_{t\to\infty} y_{t+1} \cdot t / \sum_{j=1}^{t} y_j = \Gamma$. Then:*

(i) For $y > 0$, we have $\Gamma = 1$.
(ii) For $y = 0$ we have $\Gamma \leq 1$. Here $0/0$ is interpreted as 0.

Proof. Let $\overline{y}_t = \frac{1}{t} \sum_{j=1}^{t} y_j$. For $y > 0$, we have also $\lim_{t\to\infty} \overline{y}_t = y$, so $\lim_{t\to\infty} y_{t+1} \cdot t / \sum_{j=1}^{t} y_j = y/y = 1$. Part (i) follows. Turning to (ii), for $y = 0$, if $y_t = 0$ for infinitely many t's, then clearly $\Gamma = 0$. Otherwise, there exists t_0 such that $y_t > 0$ for every $t \geq t_0$. Consider the following positive descending subsequence. For every $i \geq 1$ we let t_i be the minimum integer such that $t_i > t_{i-1}$ and $y_{t_i} \leq y_{t_{i-1}}$. The minimum exists since $y_{t_{i-1}}$ is positive and $\lim_{t\to\infty} y_t = 0$. By construction, $y_{t_i} \leq y_t$ for every $t \in [t_{i-1}, t_i]$. Since $(y_{t_i})_{i=0}^{\infty}$ is descending, we have $y_{t_i} \leq y_t$ for every $t \in [t_0, t_i]$, thus letting

$$Q(i) = \frac{y_{t_i+1} \cdot t_i}{\sum_{j=1}^{t_i} y_j} \leq \frac{y_{t_i+1} \cdot t_i}{\sum_{j=t_0}^{t_i} y_j} \leq \frac{y_{t_i+1} \cdot t_i}{(t_i - t_0 + 1) \cdot y_{t_i+1}} = \frac{t_i}{t_i - t_0 + 1} ,$$

it follows that $\Gamma = \lim_{i\to\infty} Q(i) \leq \lim_{i\to\infty} t_i/(t_i - t_0 + 1) = 1$. □

Corollary 5.7. *Consider $G(p_t, r_t, q_t)$. Assume $\lim_{t\to\infty} q_t = 0$, $\lim_{t\to\infty} y_t = y$, $\sum_{j=1}^{t} y_j = \infty$, and $\lim_{t\to\infty} y_{t+1} \cdot t / \sum_{j=1}^{t} y_j = \Gamma > 0$. Then the resulting graph follows a power law degree distribution with exponent $\beta \in (1, 3]$.*

Proof. By Cor. 5.4, $G(p_t, r_t, q_t)$ follows a power law degree distribution with exponent $\beta = 1 + 2\Gamma/(2 - y) > 1$. By Lemma 5.6, for $0 < y \leq 1$ we have $\beta = 1 + \frac{2}{2-y} \in (2, 3]$ and for $y = 0$ we have $\beta = 1 + \Gamma \in (1, 2]$. □

6 CONCLUSIONS AND DISCUSSION

Consider the preferential attachment model $G(p)$ described in [6]. The contribution of this paper is threefold. First, we improve the concentration bounds on the number of nodes of degree k at time t, denoted $m_{k,t}$. Specifically, we show that $m_{k,t}$ is concentrated around its mean with a deviation of $O(\sqrt{t})$, an improvement of the previously known deviation of $O(\sqrt{k^3 t \ln t})$. These bounds also allow us to bound the concentration of the k-rich club, namely the number of nodes with degree at least k.

Second, we bound the expectation of $m_{k,t}$, showing that
$\mathbb{E}\left[m_{k,t}\right] = pM_k e_t \pm O(1/k)$, where $\beta = 1 + \frac{2}{2-p}$ and $M_k =$

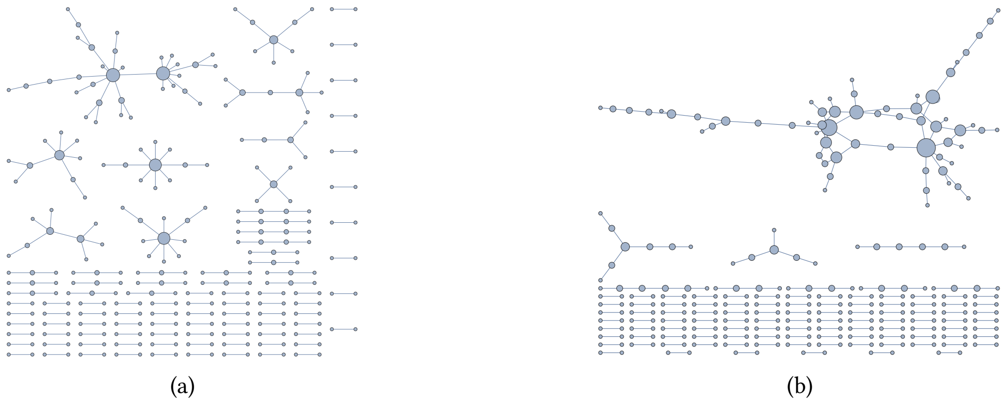


Figure 2: An example of two different networks with the same power law exponent $\beta = 5$. Each with 200 edges. (a) No edge events: $p_t = 0.5, r_t = 0, q_t = 0.5$. Many connected components. (b) No node events: $p_t = 0, r_t = 0.25, q_t = 0.75$. A giant connected component.

$\frac{2}{4-p}\prod_{j=1}^{k-1}\frac{j}{j+\beta}$ is proportional to $k^{-\beta}$, and $e_t = e_0 + t$ is the number of edges at time t. This allows us to estimate $\mathbb{E}\left[m_{k,t}\right]$ for any $k = o(t^{1-p/2})$, and the k-rich club, namely $\mathbb{E}\left[RC_{k,t}\right] = \sum_{j\geq k}\mathbb{E}\left[m_{j,t}\right]$, for any $k = o((t/\ln t)^{1-p/2})$.

Third, we extend the $G(p)$ model in a natural way to a model $G(p_t, r_t, q_t)$, by allowing the creation of new connected components, and time varying p. This results in a model for networks obeying a power law degree distribution with any exponent β in the full spectrum $1 < \beta < \infty$.

The new model allows us to point out some delicate issues that come up in the analysis of social networks. Let us mention two such examples. The first deals with the difference between the degree distribution in the network as a whole and that of the largest component in the same network. Consider the case where $p_t = q_t = 1/2$ for every t, thus there are no edge events. An illustration of the generated network appears in Figure 2(a). By Thm. 5.3, we obtain $y = 1.5$, $\Gamma = 1$ and so the whole network follows a power law degree distribution with exponent $\beta = 5$. However, any single component of this network is created using node events only, and hence can be described within the classical preferential attachment model of Barabási and Albert [4], or in our model $G(p_t, r_t, q_t)$ with $p_t = 1$. By Thm. 5.3, the degree distribution in this component follows a power law with exponent $\hat{\beta} = 3$, which is considerably different from $\beta = 5$, the exponent of the entire network. The surprising implication is that, focusing on studying the largest component alone might be insufficient for determining the correct power law exponent of the network as a whole.

The second example highlights yet another important feature of our model, namely, that it can generate many networks with the same power law distribution, but strikingly different structure. To illustrate this, consider the case where $r_t = 0.25$ and $q_t = 0.75$ for every t, thus there are no node events. An illustration of the generated network appears in Figure 2(b). By Thm. 5.3, the generated network follows a power law with exponent $\beta = 5$. While the above two networks display the same power law distribution, they also differ in some of their properties. For example, it can be shown that only the latter will have a single giant component, which follows a power law degree distribution with exponent $\beta = 5$. These examples emphasize some of the advantages of our extended preferential attachment model. We leave the formal proof of the above claims to the complete version of the paper.

REFERENCES

[1] William Aiello, Fan Chung, and Linyuan Lu. Random evolution in massive graphs. In *Handbook of massive data sets*, pages 97–122. Springer, 2002.
[2] Chen Avin, Zvi Lotker, Yinon Nahum, and David Peleg. Core size and densification in preferential attachment networks. In *Automata, Languages, and Programming*, pages 492–503. Springer, 2015.
[3] Chen Avin, Zvi Lotker, David Peleg, Yvonne Anne Pignolet, and Itzik Turkel. Core-periphery in networks: an axiomatic approach. *arXiv preprint arXiv:1411.2242*, 2014.
[4] A-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286:509–512, 1999.
[5] S. Boyd and L. Vandenberghe. *Convex optimization.* Cambridge Univ. Press, 2004.
[6] F.R.K. Chung and L. Lu. *Complex graphs and networks.* AMS, 2006.
[7] Aaron Clauset, Cosma Rohilla Shalizi, and Mark EJ Newman. Power-law distributions in empirical data. *SIAM review*, 51(4):661–703, 2009.
[8] Colin Cooper and Alan M Frieze. A general model of undirected web graphs. In *Algorithms—ESA 2001*, pages 500–511. Springer, 2001.
[9] Colin Cooper and Pawe l Pra lat. Scale-free graphs of increasing degree. *Random Structures & Algorithms*, 38(4):396–421, 2011.
[10] Joseph L Doob. Regularity properties of certain families of chance variables. *Transactions of the American Mathematical Society*, 47(3):455–486, 1940.
[11] S. N. Dorogovtsev and J. F. F. Mendes. Effect of the accelerating growth of communications networks on their structure. *Phys. Rev. E*, 63:025101, Jan 2001.
[12] Sergey N Dorogovtsev and JFF Mendes. Accelerated growth of networks. *arXiv preprint cond-mat/0204102*, 2002.
[13] Sergey N Dorogovtsev and Jose FF Mendes. Evolution of networks. *Advances in physics*, 51(4):1079–1187, 2002.
[14] M. J. Gagen and J. S. Mattick. Accelerating, hyperaccelerating, and decelerating networks. *Phys. Rev. E*, 72:016123, Jul 2005.
[15] Jon M Kleinberg, Ravi Kumar, Prabhakar Raghavan, Sridhar Rajagopalan, and Andrew S Tomkins. The web as a graph: measurements, models, and methods. In *Computing and combinatorics*, pages 1–17. Springer, 1999.
[16] Ravi Kumar, Prabhakar Raghavan, Sridhar Rajagopalan, Dandapani Sivakumar, Andrew Tompkins, and Eli Upfal. The web as a graph. In *Proceedings of the nineteenth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*, pages 1–10. ACM, 2000.
[17] J. Leskovec, J. Kleinberg, and C. Faloutsos. Graph evolution: Densification and shrinking diameters. *Trans. Knowledge Discovery from Data*, 1:2, 2007.
[18] M. Newman. *Networks: An Introduction.* Oxford Univ. Press, 2010.
[19] D. de Solla Price. A general theory of bibliometric and other cumulative advantage processes. *J. Amer. Soc. Inform. Sci.*, 27(5):292–306, 1976.
[20] H.A. Simon. On a class of skew distribution functions. *Biometrika*, pages 425–440, 1955.