Preferential Attachment Random Graphs with Edge-Step Functions

Caio Alves1 · Rodrigo Ribeiro2 · Rémy Sanchis3

Received: 13 August 2019 / Revised: 12 October 2019 / Published online: 7 November 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

We analyze a random graph model with preferential attachment rule and edge-step functions that govern the growth rate of the vertex set, and study the effect of these functions on the empirical degree distribution of these random graphs. More specifically, we prove that when the edge-step function $f$ is a monotone regularly varying function at infinity, the degree sequence of graphs associated with it obeys a (generalized) power-law distribution whose exponent belongs to $(1, 2]$ and is related to the index of regular variation of $f$ at infinity whenever said index is greater than $-1$. When the regular variation index is less than or equal to $-1$, we show that the empirical degree distribution vanishes for any fixed degree.

Keywords Complex networks · Preferential attachment · Concentration bounds · Power-law · Scale-free · Karamata’s theory · Regularly varying functions

Mathematics Subject Classification (2010) Primary 05C82; Secondary 60K40 · 68R10
1 Introduction

In the late 1990s, the seminal works of Strogatz and Watts [17] and of Álbert and Barabási [4] brought to light two common features shared by real-life networks: small diameter and power-law degree distribution. In the first work, the authors observed that large-scale networks of biological, social and technological origins presented diameters of much smaller order than the order of the entire network, a phenomenon they called small world. In the second paper, the authors noted that the fraction of nodes having degree $d$ decays roughly as $d^{-\beta}$ for some $\beta > 1$, a feature known as scale-freeness.

Many models proposed by the community over the years in order to understand these phenomena are Markovian, in the sense that at each step $t$ one obtains the random graph $G_t$ by performing some stochastic operation on $G_{t-1}$. In the well-known Barabási-Álbert (BA-) model [4], the stochastic operation consists of at each step a new vertex being added and a neighbor to it chosen among the previous vertices with probability proportional to its degree. This simple attachment rule, which is known as preferential attachment, or PA rule for short, is capable of producing graphs whose empirical degree distribution is well approximated by a power-law distribution with exponent $\beta = 3$ (see [6] for a rigorous proof). Many variants of the original B-A preferential attachment [7,9,11,14,15] have been introduced. These models also are capable of exhibiting power law with different values of $\beta$ and small-world phenomenon.

In what follows, we will make use of asymptotic notation $o$, $\Theta$ and $O$, which will presuppose asymptotic in the time parameter $t$, except when another parameter is explicitly indicated. We also use the notation $O_{b_1, b_2}$, indicating that the implied constant depends only on the quantities $b_1$ and $b_2$.

In the remainder of this introduction, we define our model in the next subsection and discuss the questions we have addressed in this paper. We end this section settling down some conventions and notations and explaining the paper’s structure.

1.1 The Preferential Attachment Scheme with Edge-Step Functions

The model we investigate here generalize the edge-step rule introduced in [8]. It has one parameter: a real nonnegative function $f$ with domain given by the semi-line $[1, \infty)$ such that $\|f\|_\infty \leq 1$. For the sake of simplicity, we start the process from an initial graph $G_1$ which is taken to be the graph with one vertex and one loop. We consider the two stochastic operations below that can be performed on any finite graph $G$:

- **Vertex step** Add a new vertex $v$ and add an edge $\{u, v\}$ by choosing $u \in G$ with probability proportional to its degree. More formally, conditionally on $G$, the probability of attaching $v$ to $u \in G$ is given by

$$P(v \to u | G) = \frac{\text{degree}(u)}{\sum_{w \in G} \text{degree}(w)}. \quad (1.1)$$
• **Edge step** Add a new edge \( \{u_1, u_2\} \) by independently choosing vertices \( u_1, u_2 \in G \) according to the same rule described in the vertex step. We note that both loops and parallel edges are allowed.

We consider a sequence \( \{Z_t\}_{t \geq 1} \) of independent random variables such that \( Z_t \overset{d}{=} \text{Ber}(f(t)) \). We then define inductively a random graph process \( \{G_t(f)\}_{t \geq 1} \) as follows: start with \( G_1 \). Given \( G_t(f) \), obtain \( G_{t+1}(f) \) by either performing a **vertex step** on \( G_t(f) \) when \( Z_t = 1 \) or performing an **edge step** on \( G_t(f) \) when \( Z_t = 0 \).

We will call the function \( f \) by **edge-step function**, though we follow an edge step at time \( t \) with probability \( 1 - f(t) \).

### 1.2 The Edge-Step Rule

The introduction of the edge step has some advantages over some generalizations of the traditional preferential attachment rule proposed in \([4] \). Some of them come directly from the definition of the edge-step rule.

The first one is that this mechanism is easy to describe and is even natural to expect. Using the social-network terminology, it is expected that users already in the network may be connected over the time. How often old users may interact among themselves is controlled by the edge-step rule and the edge-step function. Moreover, the usefulness of the edge step has been verified empirically, in \([20] \), a statistical analysis has been made comparing prediction capabilities between a class of models with edge step (called in the paper GLP) and other influential network models, such as the Erdős-Rényi, Álbert-Barabási and Tel Aviv Network Generator. Their results suggest that the particular case \( f \equiv p \in (0, 1) \) outperforms these popular models when the task is either to predict or mimic real-world networks, pointing out that the edge-step rule does make the model realistic.

The second advantage, from the modeling perspective, is that the edge step naturally offers control over the growth rate of the vertex set. Fix an edge-step function \( f \) and let \( V_t(f) \) be the random number of vertices in \( G_t(f) \). Then, by Chernoff bounds, \( V_t(f) \) concentrates around its expected value, which is \( \sum_{s=1}^{t} f(s) \).

Regarding the order of \( V_t(f) \), in many preferential attachment random graph models it grows linearly with \( t \), meaning that \( V_t = \Theta(t) \), w.h.p or deterministically depending on the model, see \([4,10,11,13] \) for a few important examples. For modeling purposes, a sub-linear growth and some control over the growth rate of the vertex set may be desirable, since in many real-world networks the rate of newborn nodes decreases with time. This tendency has been verified empirically. In \([16] \), it is verified that the user growth of Facebook slows down in time. In our setup, this may be achieved by choosing \( f \) such that \( f(t) \downarrow 0 \) as \( t \) goes to infinity. Since \( V_t(f) \approx \sum_{s=1}^{t} f(s) \) and, for a wide class of functions, it holds that \( \sum_{s=1}^{t} f(s) \approx \int_{1}^{t} f(s) \, ds \), we may abuse the notation for a moment and see \( f \) as the growth rate of the vertex set, i.e., we may write

\[
\frac{dV_t(f)}{dt} = f(t).
\]
Furthermore, the edge step also produces many interesting topological properties on the graph sequence \( \{G_t(f)\}_{t \in \mathbb{N}} \) such as existence of complete subgraph of polynomial order \([1,2]\) and distances of order much smaller than the more traditional models [3].

### 1.3 The Empirical Degree Distribution

Given a vertex \( v \) in \( G_t(f) \), we let \( D_t(v) \) be its degree in \( G_t(f) \). In this work, we focus on the empirical degree distribution

\[
\hat{P}_t(d, f) := \frac{1}{V_t(f)} \sum_{v \in G_t(f)} \mathbbm{1}\{D_t(v) = d\},
\]

i.e., the random proportion of vertices having degree \( d \) in \( G_t(f) \), for any \( d \in \mathbb{N} \).

In many works, a combination of the preferential attachment rule (1.1) with other attachment rules [11,14,18] proved itself to be an efficient mechanism for generating graphs where \( \hat{P}_t(d) \) is essentially a power-law distribution, meaning that \( \text{w.h.p.} \) for large \( t \),

\[
\hat{P}_t(d) \approx d^{-\beta},
\]

for some exponent \( \beta \) generally lying in \((2, 3]\). In [8], the authors investigated a very general model whose growth rule involves the case \( f(t) \equiv p \) and the possibility of choosing vertices uniformly instead of preferentially. Their model produces graphs whose empirical degree distribution follows a power-law distribution whose exponent lies in the range \((2, 3]\). More specifically, in the particular case of \( f(t) \equiv p \), with \( p \in (0,1) \), studied in [7,8], the edge-step functions provided a control over the tail of the power-law distribution producing graphs obeying such laws with a tunable exponent \( \beta = 2 + \frac{p}{2-p} \), i.e., they have shown that

\[
\hat{P}_t(d) = C d^{-2 - \frac{p}{2-p}} + O \left( \frac{d}{\sqrt{t}} \right),
\]

\( \text{w.h.p.} \). As pointed out in [10], it may be interesting to investigate models capable of generating graphs with \( \beta \) lying in the range \((1, 2]\). In this paper, the authors propose a model in which the number of edges added at each step is given by a sequence of independent random variables. This new rule is capable of reducing \( \beta \), but the vertex set still grows linearly in time.

In [13], the authors have introduced a Markovian model that combines the PA rule with spatial proximity, i.e., the vertices are added on some metric space and the closer the vertices are the more likely they are to become connected. In the paper, the authors have addressed the characterization of the empirical degree distribution, proving that it is also well approximated by a power law.

In our case, one of our results (Theorem 1) states that, for a broad class of edge-step functions, \( \hat{P}_t(d) \) is approximated by a power-law distribution whose exponent lies in \((1, 2]\). More formally, we prove that
\[
\hat{P}_t(d, f) = C d^{-2+\gamma} + O \left( \frac{d}{\sqrt{\int_1^t f(s)ds}} \right)
\]

where \( \gamma \in [0, 1) \) depends only on the class \( f \) belongs to.

### 1.4 Main Results

Our main goal in this paper is to characterize \( \hat{P}_t(\cdot, f) \) for a class of edge-step functions as general as possible. More precisely, we would like to obtain a very broad family \( \mathcal{F} \) of functions and a (generalized) distribution over the positive integers \( (p(d, f))_{d \in \mathbb{N}} \) such that, for every fixed \( d \in \mathbb{N} \) and for all \( f \in \mathcal{F} \)

\[
\left| \hat{P}_t(d, f) - p(d, f) \right| \leq o_f(1), \quad (1.5)
\]

with high probability.

The class we investigate here is the class of regularly varying functions. A positive function \( f \) is said to be a regularly varying function at infinity with index of regular variation \( \gamma \) if, for all \( a \in \mathbb{R}_+ \), the identity below is satisfied

\[
\lim_{t \to \infty} \frac{f(at)}{f(t)} = a^{\gamma}. \quad (1.6)
\]

In the particular case where \( \gamma = 0 \), \( f \) is said to be a slowly varying function. It will be useful to our purposes to recall that if \( f \) is a regularly varying function with index \( \gamma \), the representation theorem (Theorem A.2) assures the existence of a slowly varying function \( \ell \) such that, for all \( t \) in the domain of \( f \), \( f(t) = \ell(t)t^{\gamma} \).

For each \( \gamma \in [0, \infty) \), we take the family \( \mathcal{F} \) to be the subclass of all regularly varying function of index \( \gamma \), bounded by one and converging monotonically to zero. In notation, we will focus on functions belonging to the family defined below

\[
\text{RES}(-\gamma) := \{ f : [1, \infty] \to [0, 1] | f \text{ is continuous, decreases to zero and has index } -\gamma \}. \quad (1.7)
\]

The goal is to characterize \( \hat{P}_t(\cdot, f) \) for all functions in \( \text{RES}(-\gamma) \), for all \( \gamma \in [0, \infty) \). Our results establish a characterization for the empirical distribution depending only on the index \( -\gamma \) and show a phase transition at \( \gamma = 1 \), meaning that for all \( \gamma \) below this value, \( \hat{P}_t(d, f) \) is well approximated by a power law whose exponent depends on \( \gamma \) only, whereas for \( \gamma \geq 1 \) the empirical distribution vanishes for all fixed \( d \).

Specifically, if we let \( (p_\gamma(d))_{d \in \mathbb{N}} \) be the (generalized) distribution on \( \mathbb{N} \) given by

\[
p_\gamma(d) := \frac{(1-\gamma)\Gamma(2-\gamma)\Gamma(d)}{\Gamma(d+2-\gamma)}, \quad (1.8)
\]
for $\gamma \in [0, 1)$, a consequence of our results is that, for fixed $d \in \mathbb{N}$, w.h.p,

$$\left| \hat{P}_t(d, f) - p\gamma(d) \right| \leq o(1) \quad (1.9)$$

for any $f \in \text{RES}(-\gamma)$, with $\gamma \in [0, 1)$. The error $o(1)$ on (1.9) may depend on $f$ in an involved way, and it is specified combining the estimates given by the two theorems below. We note that the distribution defined in (1.8) is not a probability distribution since some mass escapes to infinity when $\gamma > 0$, see the discussion in Sect. 2.3.

The power-law degree distribution of the random graph when $f \in \text{RES}(-\gamma)$ is provided by Propositions 3.1 and 4.2 and is formally stated in the theorem below.

**Theorem 1** (Power-law degree distribution) Let $f \in \text{RES}(-\gamma)$ with $\gamma \in [0, 1)$. Then, for all $d \in \mathbb{N}$, $\alpha \in (0, 1)$, and $A > 0$ satisfying (4.27),

$$\left| \hat{P}_t(d, f) - \left(1 - \gamma\right)\Gamma(2 - \gamma)\Gamma(d) / \Gamma(d + 2 - \gamma) \right| \leq A \sqrt{40d^2 / E\nu(f)} + o_{\alpha, f}(1), \quad (1.10)$$

with probability at least $1 - 3e^{-A^2/3}$.

We must point out that Proposition 3.1 provides an explicit expression for $o_{\alpha, f}(1)$; however, since it is too involved, we decided to omit it in the above statement for the sake of simplicity. In Sect. 2, we discuss in more detail the above statement as well as where the $o_{\alpha, f}(1)$ term comes from.

We point out that a by-product of our theorems is that for all $d \in \mathbb{N}$, we have that

$$\lim_{t \to \infty} \hat{P}_t(d, f) = p\gamma(d), \ a.s$$

for any $f \in \text{RES}(-\gamma)$, with $\gamma \in [0, 1)$.

For functions whose index $-\gamma$ lies on $(-\infty, -1]$, all the mass of the empirical degree distribution escapes to infinity in the sense that the fraction of vertices having degree $d$ goes to zero for any value of $d$. We note that the case $-\gamma < -1$ is not interesting since in this case $V_t(f)$ is finite almost surely and all vertices are selected infinitely many times with probability one. Therefore, the only remaining case is the one considered in the theorem below.

**Theorem 2** (All mass escapes to infinity) Let $f \in \text{RES}(-\gamma)$ with $\gamma = 1$. Then, for all fixed $d \in \mathbb{N},$

$$\lim_{t \to \infty} \hat{P}_t(d, f) = 0, \ a.s. \quad (1.11)$$
1.5 Notation and Conventions

1.5.1 General

Regarding constants, we let $C, C_1, C_2, \ldots$ and $c, c_1, c_2, \ldots$ be positive real numbers that do not depend on $t$ whose values may vary in different parts of the paper. The dependence on other parameters will be highlighted throughout the text.

Since our model is inductive, we use the notation $\mathcal{F}_t$ to denote the $\sigma$-algebra generated by all the random choices made up to time $t$. We then have the natural filtration $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots$ associated with the process.

1.5.2 Graph Theory

We abuse the notation, and let $V_t(f)$ denote the set and the number of vertices in $G_t(f)$. Given a vertex $v \in V_t(f)$, we will denote by $D_t(v)$ its degree in $G_t(f)$. We will also denote by $\Delta D_t(v)$ the increment of the discrete function $D_t(v)$ between times $t$ and $t+1$, that is,

$$\Delta D_t(v) = D_{t+1}(v) - D_t(v).$$ (1.12)

For every $d \in \mathbb{N}$ and edge-step function $f$, we let $N_t(d, f)$ be the number of vertices of degree $d$. Naturally, $N_t(\leq d, f)$ stands for the number of vertices having degree at most $d$. Our empirical degree distribution is written as

$$\hat{P}_t(d, f) = \frac{N_t(d, f)}{V_t(f)}.$$ (1.13)

Since the expected number of vertices appear repeatedly throughout the paper, we reserve a special notation for it

$$F(t) := \mathbb{E}V_t(f).$$ (1.14)

We also drop the dependency on $f$ on all the above notations when the function $f$ is clear from the context or when we are talking about these observables in a very general way, including in other preferential attachment models.

1.6 Organization

In Sect. 2, we discuss in more detail the statement of our main results as well as the main technical ideas behind their proof. Section 3 is devoted to the analysis of expected number of vertices having degree $d$ for $f \in \text{RES}(-\gamma)$ with $\gamma \in [0, 1)$. In Sect. 4, we prove a general concentration result for $\hat{P}_t(d, f)$ which holds for any edge-step function $f$. Then, we use this general result exploiting our knowledge about $f$ when it belongs to the class $\text{RES}(-\gamma)$, for $\gamma \in [0, 1)$, to prove Theorem 1. The case $\gamma \geq 1$ is treated separately in Sect. 5, where we prove Theorem 2 and all the
results needed. We end this paper presenting in Sect. 6 some brief discussion about
the affine case of this model and the maximum degree. In the first topic, we show that
the presence of edge-step functions inhibits the effect of constant terms added to rule
(1.1). In the second topic, we provide some computations that indicate that the order
of the maximum degree also varies according to how fast the edge step goes to zero.
For $f$ whose index of regular variation $-\gamma$ lies on $(-1, 0)$ a maximum degree of order
$t$ seems to be achieved, whereas the case where $f$ is slowly varying seems to be richer
in the sense that the order of the maximum degree at time $t$ may depend on $f$.

2 Discussion on the Main Results and Technical Ideas

2.1 Expected Value Analysis and Convergence Rate

The first step to prove the power-law degree distribution stated by Theorem 1 is to
prove a version of it in average, i.e., we prove that for $t$ large enough the sequence
of real numbers $(\mathbb{E}[\hat{P}_t(d, f)])_{d \in \mathbb{N}}$ is close to $(p_\gamma(d))_{d \in \mathbb{N}}$ for any $f \in \text{RES}(-)$ in the
$L_\infty(\mathbb{N})$-norm. How close these sequences depend on $f$ and an auxiliary parameter
$\alpha \in (0, 1)$.

The proof of this first step is inspired by [19] Section 8.6.2, though the degree of
generality we work on prevents a straightforward application. The essential idea is
that $\mathbb{E}[\hat{N}_t(d, f)]$ and $p_\gamma(F(t))$ satisfy very similar recurrence relations in $d$ when $t$
is large, which implies in turn that they, as sequences in $d$, are fixed points of similar
operators in $L_\infty(\mathbb{N}) \to L_\infty(\mathbb{N})$. By quantifying this similarity and using a contraction
argument, we can prove that the sequences are indeed close in the $L_\infty(\mathbb{N})$ sense. This is
the statement of Proposition 3.1, which gives the term $o_{\alpha, f}(1)$ present in the statement
of Theorem 1.

The distance between $(\mathbb{E}[\hat{P}_t(d, f)])_{d \in \mathbb{N}}$ and $(p_\gamma(d))_{d \in \mathbb{N}}$ is the main bottleneck in
the approximation of the empirical degree distribution by a power-law distribution.
Whereas Proposition 4.2 states essentially that w.h.p we have

$$\hat{P}_t(d, f) \approx \mathbb{E}[\hat{P}_t(d, f)] \pm C \frac{d^2}{\int_1^t f(s)ds},$$

in Sect. 3, we discuss some examples which illustrate that, by choosing $f$ properly,
our bounds on the convergence rate of $\mathbb{E}[\hat{P}_t(d, f)]$ to $p_\gamma(d)$ may be arbitrarily slow.

2.2 Concentration of Measure Results

The second step toward the proof of Theorem 1 is to prove concentration results
for $N_t(d, f)$. However, in our setup, the presence of the edge-step function requires
concentration results sharper than those found in the present literature. For general
concentration results for $N_t(d)$, the usual approach is to obtain a (sub, super)martingale
involving $N_t(d)$ and then to prove that it has bounded increments and finally to apply
Azuma’s inequality (Theorem B1). These (sub, super)martingales are usually $N_t(d)$
properly normalized or the Doob martingale, see [7,19] for the two distinct approaches. This sort of argument leads to concentration results for \( N_t(d) \) with a deviation from its mean typically of order \( \sqrt{t} \). More precisely, it is proven that

\[
N_t(d) \sim \mathbb{E}[N_t(d)] \pm A\sqrt{t},
\]

with high probability, and from the analysis of the expected value it comes that

\[
\mathbb{E}[N_t(d)] \sim \frac{t}{d^\beta},
\]

where \( \beta \) is the power-law exponent. Since the edge-step function controls the growth rate of the vertex set, in the presence of a regular varying edge-step function the expected value of \( N_t(d, f) \) analysis leads to

\[
\mathbb{E}[N_t(d, f)] \sim \int_1^t \frac{f(s)ds}{d^\beta},
\]

see Proposition 3.1. On the other hand, a straightforward application of the usual approach would give us

\[
\int_1^t \frac{f(s)ds}{d^\beta} - A\sqrt{t} \leq N_t(d, f) \leq \int_1^t \frac{f(s)ds}{d^\beta} + A\sqrt{t},
\]

with high probability. However, this is trivially true for some choices of \( f \), e.g., if \( f \in \text{RES}(-\gamma) \) with \( \gamma > 1/2 \). This issue demands a result finer than those found in the literature, at least for a particular class of functions. We overcome it by applying Freedman’s inequality (Theorem B2) instead of Azuma’s. Freedman’s inequality takes into account our knowledge about the past of the martingale to estimate its increments instead of simply bounding them deterministically as it is done in Azuma’s. However, Freedman’s inequality requires upper bounds on the conditional quadratic variation of the martingale (see (B.1)), which may be more involved than obtaining deterministic bounds for the increments.

### 2.3 The Phase Transition on \( \gamma = 1 \)

Theorems 1 and 2 show a sharp difference between cases \( \gamma \in [0, 1) \) and \( \gamma \geq 1 \). In the former case, the limit degree sequence distribution gives positive mass to any \( d \in \mathbb{N} \), indicating that in the limit we do have a positive fraction of vertices whose degree is exactly \( d \). However, on the latter case, all mass escapes to infinity, indicating that in the limit all vertices have infinite degree. The explanation for the loss of mass is due to the growth rate of \( V_t(f) \). In the regime \( \gamma \geq 1 \), we have considerably less vertices than we have in the case \( \gamma \in [0, 1) \). Consequently, there exist less competition for the edges coming at each time. This allows all vertices in the graph to increase their degrees “simultaneously.”
The phenomena of loss of mass to infinity do not occur in the case $\gamma = 0$. Indeed, observe that in this case $p_\gamma (d)$ assumes a simpler expression

$$p_\gamma (d) = \frac{\Gamma(2) \Gamma(d)}{\Gamma(d + 2)} = \frac{(d - 1)!}{(d + 1)!} = \frac{1}{d(d + 1)} = \frac{1}{d - \frac{1}{d + 1}}.$$  

Then, summing over $d$ we obtain a telescoping sum which converges to 1. However, in the regime $\gamma \in (0, 1)$ we do have an amount of mass escaping to infinity. If we denote by $\hat{P}_\infty$ the limit distribution and recall the following upper bound for ratio of gamma functions

$$\frac{\Gamma(d)}{\Gamma(d + 2 - \gamma)} \leq \frac{1}{d^{2 - \gamma}},$$

we have for any $\gamma \in [0, 1)$

$$\hat{P}_\infty ([1, \infty)) = \sum_{d=1}^{\infty} p_\gamma (d) = \sum_{d=1}^{\infty} \frac{(1 - \gamma) \Gamma(2 - \gamma) \Gamma(d)}{\Gamma(d + 2 - \gamma)} \leq \Gamma(2 - \gamma),$$

which suggests that $\hat{P}_\infty (\infty) = 1 - \Gamma(2 - \gamma)$. This is due, possibly, to the existence of some positive fraction of vertices with very high degree (c.f. Sect. 6 for a discussion about the maximum degree). On the other hand, which may be surprising in this case, is the fact that $G_t(f)$ has mean degree of order $t^{\gamma}$ w.h.p, but it still has a positive proportion of vertices of constant degree.

### 2.4 The Vanishing Distribution Regime

Theorem 2 states that when $\gamma \geq 1$ the limiting distribution $\hat{P}_t(d, f)$ vanishes a.s. as $t$ tends to infinity. In order to prove this result, we apply a second moment estimation method.

### 3 Expected Value Analysis

In this section, we prove Proposition 3.1, which gives us estimates on the expected number of vertices having degree exactly $d$ for $f \in \text{RES}(-\gamma)$, with $\gamma \in [0, 1)$. Our first result in this direction is the following recurrence relation for $\mathbb{E}N_t(d, f)$ which holds for any edge-step function $f$.

**Lemma 1** Let $\mathbb{E}N_t(d)$ denote $\mathbb{E}N_t(d, f)$ for a fixed edge-step function $f$. Then, $\mathbb{E}N_t(d)$ satisfies

$$\mathbb{E}N_{t+1}(1) = \left(1 - \frac{2 - f(t + 1)}{2t} + \frac{(1 - f(t + 1))}{4t^2}\right) \mathbb{E}N_t(1) + f(t + 1), \quad (3.1)$$
and for a fixed integer $d \geq 2$,

$$
\mathbb{E}N_{t+1}(d) = \left(1 - \frac{(2 - f(t + 1))d}{2t} + \frac{(1 - f(t + 1))d^2}{4t^2}\right) \mathbb{E}N_t(d)
+ \left(\frac{(2 - f(t + 1))(d - 1)}{2t} - \frac{2(1 - f(t + 1))(d - 1)^2}{4t^2}\right) \mathbb{E}N_t(d - 1)
+ \left(\frac{1 - f(t + 1)(d - 2)^2}{4t^2}\right) \mathbb{E}N_t(d - 2).
$$

(3.2)

**Proof** There are two possible ways in which a vertex $v$ increases its degree by 1 at time $t + 1$: either a vertex is created at time $t + 1$ and connects to $v$, or an edge is created instead and exactly one of its endpoints connects to $v$. This implies

$$
\mathbb{P}(\Delta D_t(v) = 1|\mathcal{F}_t) = f(t + 1) \frac{D_t(v)}{2t} + 2(1 - f(t + 1)) \frac{D_t(v)}{2t} \left(1 - \frac{D_t(v)}{2t}\right)
= \left(1 - \frac{f(t + 1)}{2}\right) \frac{D_t(v)}{t} - 2(1 - f(t + 1)) \frac{D_t^2(v)}{4t^2}.
$$

(3.3)

In order for the degree of $v$ to increase by 2 at time $t + 1$, the only possibility is that an edge step occurs and both endpoints of the new edge are attached to $v$, creating a loop. This implies

$$
\mathbb{P}(\Delta D_t(v) = 2|\mathcal{F}_t) = (1 - f(t + 1)) \frac{D_t^2(v)}{4t^2}.
$$

(3.4)

We may write $N_{t+1}(d)$ as

$$
N_{t+1}(d) = \sum_{v \in V_t(f) \atop D_t(v) = d} \mathbb{1}\{\Delta D_t(v) = 0\} + \sum_{v \in V_t(f) \atop D_t(v) = d-1} \mathbb{1}\{\Delta D_t(v) = 1\}
+ \sum_{v \in V_t(f) \atop D_t(v) = d-2} \mathbb{1}\{\Delta D_t(v) = 2\}.
$$

(3.5)

Combining the three above equations and taking the expected value on (3.5), we obtain (3.2). For the case $d = 1$, just observe that

$$
N_{t+1}(1) = \sum_{v \in V_t(f) \atop D_t(v) = 1} \mathbb{1}\{\Delta D_t(v) = 0\} + \mathbb{1}\{a \text{ vertex born at time } t + 1\}.
$$

\[\square\]

From now on, we restrict our edge-step functions to the class $\text{RES}(-\gamma)$ with $\gamma$ always in the range $[0, 1)$. Note that in this case, by the representation theorem (Theorem A.2), there exists a *slowly varying* function $\ell$ such that

\(\mathcal{G}\) Springer
\( f(t) = t^{-\gamma} \ell(t), \quad (3.6) \)

for all \( t \). Before proving Proposition 3.1, we introduce notation and state a crucial lemma about regularly varying functions and their sums.

**Lemma 2** (Proof in Appendix A) Let \( \gamma \in [0, 1) \) and let \( \ell : \mathbb{R} \to \mathbb{R} \) be a continuous slowly varying function such that \( s \mapsto \ell(s)s^{-\gamma} \) is non-increasing. Define

\[
\mathcal{H}_{\ell, \gamma}(t) := \int_0^1 \left| \frac{\ell(ut)}{\ell(t)} - 1 \right| u^{-\gamma} du. \quad (3.7)
\]

Then \( \mathcal{H}_{\ell, \gamma}(t) \) is well defined and the following holds

(i) \( \mathcal{H}_{\ell, \gamma}(t) \xrightarrow{t \to \infty} 0 \);

(ii) \( G_{\ell, \gamma}(t) := \sum_{k=1}^t \ell(k)k^{-\gamma} - t^{1-\gamma} \ell(t) \left( t^{1-\gamma} \ell(t) \right)^{-1} \leq \mathcal{H}_{\ell, \gamma}(t) + \left( t^{1-\gamma} \ell(t) \right)^{-1} \).

Given \( c \in \mathbb{R}, \delta \in (0, 1) \), we consider the functions that to each \( t > 1 \) associate, respectively

\[
c, \quad (\log t)^c, \quad \log \log t, \quad \exp((\log t)^\delta). \quad (3.8)
\]

**Remark 1** Here we provide some examples of the kind of rate of decay that is associated with the above lemma. Consider the functions defined in (3.8). Elementary calculations then show that their associated error terms are, respectively,

\[
\mathcal{H}_c(t) = 0, \quad \mathcal{H}_{(\log t)^c}(t) = O((\log t)^{-1}), \\
\mathcal{H}_{\log \log t}(t) = O((\log t \log \log t)^{-1}), \quad \mathcal{H}_{\exp((\log t)^\delta)}(t) = O((\log t)^{-(1-\delta)}). \quad (3.9)
\]

Now we have all the tools needed for the proof of the main result of this section.

**Proposition 3.1** Let \( f \in \text{RES}(-\gamma) \) with \( \gamma \in [0, 1) \) be such that \( f(t) = t^{-\gamma} \ell(t) \), where \( \ell \) is a slowly varying function. Then there exists a positive constant \( \tilde{C} = \tilde{C}(f) \), such that for every \( \alpha \in (0, 1) \),

\[
\sup_{d \leq t} \left\{ \mathbb{E} \left[ V_t(f) \hat{P}_t(d, f) \right] - \mathbb{E} [V_t(f)] \cdot p_\gamma(d) \right\} \leq \tilde{C} \cdot \text{err}_t(\alpha, f), \quad (3.10)
\]

where \( \text{err}_t(\alpha, f) \) is defined as

\[
\text{err}_t(\alpha, f) := 1 + \log t + \ell(t^\alpha)t^{\alpha(1-\gamma)} + \frac{\ell(t)}{\ell(t^\alpha)}t^{1-\alpha(1-\gamma)} \sup_{s \geq t^\alpha} \mathcal{H}_{\ell, \gamma}(s) t^{1-\gamma} \ell(t). \quad (3.11)
\]
Before we go to the proof of the above proposition, let us say some words about its statement. First, we stress the fact that the LHS of (3.10) is $o(\ell(t)t^{1-\gamma}) = o(\mathbb{E}V_t(f))$, a fact implied by Lemma 2. This allows one to employ Proposition 3.1 to extract results about the behavior of the expected number of vertices having degree $d$ even when $d$ is a function of $t$ such that $d = d(t) \to \infty$, though the rate of growth of $d(t)$ cannot be taken arbitrarily, but dependent on $\mathcal{H}_{\ell,\gamma}$ and $\gamma$. In fact, $d(t) = d_{\ell,\gamma}(t)$ should be chosen in such a way that

$$
\ell(t)^{-1}t^{-(1-\gamma)} \left( 1 + \log t + \ell(t)^{\alpha(1-\gamma)} + \frac{\ell(t)}{\ell(t^\alpha)}t^{(1-\alpha)(1-\gamma)} + \sup_{s \geq t^\alpha} \mathcal{H}_{\ell,\gamma}(s)t^{1-\gamma} \ell(t) \right) = o(d_{\ell,\gamma}(t)^{(2-\gamma)}). \tag{3.12}
$$

For the specific slowly varying functions in (3.8), one can see by Remark 1, (3.9) and elementary asymptotic analysis,

- $d_{c,\gamma}(t)$ can be chosen in $o\left( t^{\frac{1-\gamma}{2(2-\gamma)}} \right)$;
- $d_{\log c,\gamma}(t)$ can be chosen in $o\left( (\log t)^{\frac{1}{2-\gamma}} \right)$;
- $d_{\log \log c,\gamma}(t)$ can be chosen in $o\left( (\log t \cdot \log \log t)^{\frac{1}{2-\gamma}} \right)$;
- $d_{\exp (\log c),\gamma}(t)$ can be chosen in $o\left( (\log t)^{\frac{1-\delta}{2-\gamma}} \right)$.

Now we prove the proposition.

**Proof of Proposition 3.1** For each $t \geq 2$, we define the linear operator

$$
T_t : L_\infty(\mathbb{N}) \to L_\infty(\mathbb{N})
$$

that maps each bounded sequence $(a_j)_{j \geq 1}$ to a sequence defined by

$$(T_t((a_j)_{j \geq 1}))_k := \left( 1 - \frac{2 - f(t)}{2(t-1)} k + \frac{1 - f(t)}{4(t-1)^2} k^2 \right) a_k$$

$$+ \left( \frac{2 - f(t)}{2(t-1)} (k-1) - \frac{2(1 - f(t))}{4(t-1)^2} (k-1)^2 \right) a_{k-1} 1\{k > 1\}$$

$$+ \frac{1 - f(t)}{4(t-1)^2} (k-2)^2 a_{k-2} 1\{k > 2\}. \tag{3.13}$$

Since the coefficients of $a_k$, $a_{k-1}$, and $a_{k-2}$ above are all nonnegative, we get

$$
\|T_t((a_j)_{j \geq 1})\|_\infty \leq \sup_k \left( \left( 1 - \frac{2 - f(t)}{2(t-1)} k + \frac{1 - f(t)}{4(t-1)^2} k^2 \right) \|a_j\|_\infty \right).
$$
which implies $T_t$ is a contraction on $L_\infty(\mathbb{N})$. Furthermore, by Lemma 1, we have
\[
\mathbb{E}[N_t(d)] = (T_t((\mathbb{E}[N_{t-1}(k)])_{k \geq 1}))_{d} + f(t) \cdot 1[d = 1].
\]

Our goal is to use $T_t$ to bound the distance between the sequence of expectations above and the sequence $(F(t) \cdot p_\gamma(d))_{d \geq 1}$. We will do so by showing that $(F(t) \cdot p_\gamma(d))_{d \geq 1}$ is very close to being a fixed point of another operator defined below in (3.17), this operator being itself very close to $T_t$ for large $t$.

By elementary properties of the Gamma function, we see that $(p_\gamma(d))_{d \geq 1}$ is defined recursively by
\[
 p_\gamma(d) = \frac{d - 1}{d + 1 - \gamma} p_\gamma(d - 1); \quad p_\gamma(1) = \frac{1 - \gamma}{2 - \gamma}.
\] (3.15)

By Lemma 2, we have
\[
\frac{F(t - 1)}{F(t)} = 1 - \frac{f(t)}{F(t)} = 1 - \frac{t^{-\gamma} \ell(t)}{(1 + O(\mathcal{G}_t, \gamma(t)))^{\frac{t^{-\gamma} \ell(t)}{1 - \gamma}}}
= 1 - \frac{1 - \gamma}{t} (1 + O(\mathcal{G}_t, \gamma(t))).
\] (3.16)

Observe that the sequence $(\mathbb{E}N_t(d))_{d \geq 1}$ has all its coordinates, for $d > 2t$, equal zero. Therefore, we must truncate the sequence $(p_\gamma(d))_{d \geq 1}$ for $d > 2t$ obtaining the sequence $(m_{d,t})_{d \geq 1}$ defined by
\[
m_{d,t} := p_\gamma(d) 1[d \leq 2t].
\]

Now consider the operator $S_t : \mathbb{R}^N \to \mathbb{R}^N$ defined by
\[
(S_t(a_j)_{j \geq 1})_d := \left(\frac{d - 1}{1 - \gamma} a_{d-1} - \frac{d}{1 - \gamma} a_d\right) 1[d \leq t],
\] (3.17)
and note that, by an application of (3.15), the sequence $(m_{d,t})_{d \geq 1}$ satisfies
\[
m_{d,t} = (S_t(m_{j,t})_{j \geq 1})_d + 1[d = 1].
\] (3.18)

Defining then
\[
\mathcal{E}_d(t) := ((f(t) S_t + F(t - 1)(I - T_t))(m_{j,t})_{j \geq 1})_d,
\] (3.19)
where $I$ denotes the identity operator in $\mathbb{R}^N$, we get

$$
F(t)m_{d,t} = F(t - 1)m_{d,t} + f(t)m_{d,t}
$$

$$
= F(t - 1)m_{d,t} + f(t)(S_t(m_{j,t})_{j \geq 1}d) + f(t)\mathbb{1}[d = 1]
$$

$$
= (T_t(F(t - 1)m_{j,t})_{j \geq 1}d) + f(t)\mathbb{1}[d = 1] + \mathcal{E}_d(t).
$$

(3.20)

We will now bound from above the terms in $(\mathcal{E}_d(t))_{d \geq 1}$, which will be the main error terms associated with the approximation of $\mathbb{E}[N_t(d)]$ by $F(t)m_{d,t}$. Note that $\|m_{j,t}\|_\infty \leq 1$ and $\sup_d d^{2-\gamma} p_\gamma (d) < \infty$, which together with (3.13) imply

$$
((T_t - I)((m_{j,t})_{j \geq 1}))d = - \frac{d}{t - 1}m_{d,t} + \frac{d - 1}{t - 1}(d - 1)m_{d-1,t}
$$

$$
+ O(f(t)t^{-1} + d^\gamma t^{-2}\mathbb{1}[d \leq 2t]).
$$

(3.21)

Note that the function represented by the $O$ notation above is actually $o(t^{-1})$, since $f$ decreases to zero. We then obtain, by (3.6,3.15,3.16), for $d \leq 2t$,

$$
\mathcal{E}_d(t) = F(t - 1)\left(\frac{d}{t - 1}m_{d,t} - \frac{d - 1}{t - 1}m_{d-1,t}\right)
$$

$$
+ f(t)\left(- \frac{d}{1 - \gamma}m_{d,t} + \frac{d - 1}{1 - \gamma}m_{d-1,t}\right) + o(t^{-1})
$$

$$
= \frac{d}{1 - \gamma}p_\gamma (d)
$$

$$
\left((1 + O(\mathcal{G}_t,\gamma (t - 1)))\ell(t - 1)(t - 1)^{1-\gamma} - \ell(t)t^{-\gamma}\right)
$$

$$
+ \frac{(d - 1) d + 1 - \gamma}{1 - \gamma}p_\gamma (d)
$$

$$
\left(\ell(t)t^{-\gamma} - \frac{(1 + O(\mathcal{G}_t,\gamma (t - 1)))\ell(t - 1)(t - 1)^{1-\gamma}}{(t - 1)}\right)
$$

$$
+ o(t^{-1})
$$

$$
= t^{-\gamma} p_\gamma (d)\left((1 + O(\mathcal{G}_t,\gamma (t - 1)))\ell(t) - \ell(t - 1)\left(1 - \frac{\gamma}{t} + O(t^{-2})\right)\right)
$$

$$
+ o(t^{-1})
$$

$$
= t^{-\gamma} p_\gamma (d)(\ell(t) - \ell(t - 1)) + t^{-\gamma} \ell(t - 1) O(\mathcal{G}_t,\gamma (t - 1)) + o(t^{-1}).
$$

(3.22)

Furthermore, $\mathcal{E}_d(t) = 0$ for $d > 2t$. The above equation together with (3.14) and (3.20) implies

$$
\|(\mathbb{E}[N_t(d)] - F(t)m_{d,t})_{d \geq 1}\|_\infty \leq \|T_t((\mathbb{E}[N_{t-1}(d)] - F(t - 1)m_{d,t})_{d \geq 1})\|_\infty
$$

$$
+ \|(\mathcal{E}_d(t))_{d \geq 1}\|_\infty
$$

$$
\leq \|(\mathbb{E}[N_{t-1}(d)] - F(t - 1)m_{d,t-1})_{d \geq 1}\|_\infty
$$

$$
+ \|(\mathcal{E}_d(t))_{d \geq 1}\|_\infty + F(t - 1) p_\gamma (t).
$$
\[
\leq C + \sum_{s=1}^{t} (\| (E_d(s))_{d \geq 1} \|_\infty + F(s-1)p_\gamma(s)),
\]
(3.23)
since
\[
\| (E[N_1(d)] - F(1)m_{d,1})_{d \geq 1} \|_\infty < C.
\]
for some constant \( C > 0 \). Since \( p_\gamma(s) = O(s^{-2+\gamma}) \), Lemma 2 implies
\[
\sum_{s=1}^{t} F(s-1)p_\gamma(s) \leq C \sum_{s=1}^{t} s^{-1} \leq C \log t,
\]
and the proof will be finished once we show an upper bound for \( \sum_{s=1}^{t} \| (E_d(s))_{d \geq 1} \|_\infty \) of the desired order. Since \( \ell(s)s^{-\gamma} \) is decreasing, we get
\[
\sum_{s=1}^{t} s^{-\gamma} |\ell(s) - \ell(s-1)| \leq C + \int_{1}^{t} s^{-\gamma} |\ell(s) - \ell(t)| ds
\]
\[
+ \int_{1}^{t} s^{-\gamma} |\ell(s-1) - \ell(t)| ds
\]
\[
\leq C + \ell(t)t^{1-\gamma}H_{\ell,\gamma}(t) + \ell(t)t^{1-\gamma}
\]
\[
\int_{0}^{t-1} \frac{|\ell(y)|}{\ell(t)} - 1 \left| \frac{(y+1)^{-\gamma}}{t^{1-\gamma}} \right| dy
\]
\[
\leq C + 2\ell(t)t^{1-\gamma}H_{\ell,\gamma}(t).
\]
(3.24)

By Lemma 2, we have
\[
\sum_{s=1}^{t} s^{-\gamma} \ell(s-1)G_{\ell,\gamma}(s-1)
\]
\[
= \sum_{s=1}^{t^\alpha} s^{-\gamma} \ell(s-1)G_{\ell,\gamma}(s-1) + \sum_{s=t^\alpha+1}^{t} s^{-\gamma} \ell(s-1)G_{\ell,\gamma}(s-1)
\]
\[
\leq C\ell(t^\alpha)t^{\alpha(1-\gamma)} + C \sup_{s \geq t^\alpha} G_{\ell,\gamma}(s)t^{1-\gamma} \ell(t)
\]
\[
\leq C \left( \ell(t^\alpha)t^{\alpha(1-\gamma)} + \sup_{s \geq t^\alpha} H_{\ell,\gamma}(s)t^{1-\gamma} \ell(t) + \frac{\ell(t)}{\ell(t^\alpha)}t^{(1-\alpha)(1-\gamma)} \right)
\]
(3.25)

Together with (3.23,3.24), this implies
\[
\| (E[N_t(d)] - F(t)m_{d,t})_{d \geq 1} \|_\infty
\]
\[
\leq C \left( 1 + \log t + \ell(t)\ell'(1-\gamma) + \frac{\ell(t)}{\ell(t^\alpha)} t^{(1-\alpha)(1-\gamma)} + \sup_{s \geq t^\alpha} \mathcal{H}_{t,\gamma}(s) t^{1-\gamma} \ell(t) \right),
\]

finishing the proof of the result. \hfill \Box

### 4 Concentration Results for \( \hat{P}_t(d, f) \)

In this section, we prove a concentration measure result for the empirical degree distribution. In essence, we will show that \( \hat{P}_t(d) \) is close to the estimates for its expected value we gave in the previous section.

For a fixed time \( t \geq 1 \), \( d \in \mathbb{N} \) and any edge-step function \( f \), we define the following sequence of random variables

\[
M_s(d, f) := E \left[ N_t(d, f) \mid \mathcal{F}_s \right].
\]  

Since the degree \( d \) and the edge-step function \( f \) will be fixed for the remainder of this section, we will omit the dependency on them, denoting simply \( \{M_s\}_{s \geq 1} \) when there is no risk of confusion. Observe that by the tower property of the conditional expected value, it follows that \( \{M_s\}_{s \geq 1} \) is a martingale.

We will obtain our concentration result applying Freedman’s inequality (Theorem B2) to \( M_t \). It requires estimates on the increments of \( \{M_s\}_{s \geq 1} \) as well as on its conditional quadratic variation, see B.1. We begin by showing that \( \{M_s\}_{s \geq 1} \) is actually a bounded increment martingale, which is done in the next lemma. Since it is almost in line with proof of Lemma 8.6 in [19], we skip some details throughout the proof.

**Lemma 3** (Bounded increments) Let \( \{M_s\}_{s \geq 1} \) be as in (4.1). Then, it satisfies

\[
|M_{s+1} - M_s| \leq 4,
\]  

for all values of \( s \).

**Proof** For a fixed \( s \), consider in the same probability space the process

\[
\{G_r'(f)\}_{r \geq 1} \overset{d}{=} \{G_r(f)\}_{r \geq 1},
\]

which evolves following exactly the steps of \( \{G_r(f)\}_{r \geq 1} \) for all \( r \leq s \) and then evolves independently for \( r \geq s + 1 \). Let \( \{F'_r\}_{r \geq 1} \) be the natural filtration associated with the prime process.

Denote by \( v_r \) the vertex born at time \( r \geq 1 \) and recall the the definition of \( (Z_r)_{r \geq 1} \), the Bernoulli variables that control whether a vertex or edge step was taken at each time. Observe that we may write \( N_t(d, f) \) as

\[
N_t(d, f) = \sum_{r=1}^{t} 1(D_t(v_r) = d) Z_r,
\]  

\( \square \) Springer
consequently, we may express $\Delta M_s$ as

$$M_{s+1} - M_s = \sum_{r=1}^t \left[ \Pr(D_t(v_\tau) = d, Z_r = 1|\mathcal{F}_{s+1}) - \Pr(D_t(v_\tau) = d, Z_r = 1|\mathcal{F}_s) \right]. \quad (4.4)$$

Let $D'_t(v_\tau)$ and $Z'_r$ denote the counterpart to $D_t(v_\tau)$ and $Z_r$ in the prime process, respectively, and note that

$$\Pr(D'_t(v_\tau) = d, Z'_r = 1|\mathcal{F}_s) = \Pr(D'_t(v_\tau) = d, Z'_r = 1|\mathcal{F}_{s+1}), \quad (4.5)$$

since $\mathcal{F}_{s+1}$ is $\mathcal{F}_s$ (which is equal to $\mathcal{F}'_s$) with information independent of $D'_t(i)$ and $Z'_r$ added. Moreover, since the evolution of each vertex’s degree only depends on itself, we also have

$$\Pr(D_t(v_\tau) = d, Z_r = 1|\mathcal{F}_{s+1}) = \Pr(D_t(v_\tau) = d, Z_r = 1|D_{s+1}(v_\tau)) \quad (4.6)$$

and

$$\Pr(D'_t(v_\tau) = d, Z'_r = 1|\mathcal{F}_{s+1}) = \mathbb{E}\left[ \Pr(D'_t(v_\tau) = d, Z'_r = 1|\mathcal{F}'_{s+1}) | \mathcal{F}_{s+1} \right]$$

$$= \mathbb{E}\left[ \Pr(D'_t(v_\tau) = d, Z'_r = 1 | D'_{s+1}(v_\tau)) | \mathcal{F}_{s+1} \right]. \quad (4.7)$$

Now, observe that if $D_{s+1}(v_\tau) = D'_{s+1}(v_\tau)$, then

$$\Pr(D_t(v_\tau) = d, Z_r = 1 | D_{s+1}(v_\tau)) = \Pr(D'_t(v_\tau) = d, Z'_r = 1 | D'_{s+1}(v_\tau)), \quad (4.8)$$

since both processes evolve with the same distribution. Furthermore, at time $s$, we have that $D_s(v_\tau) = D'_s(v_\tau)$, for all $r \leq s$; thus, the number of vertices which have $D_{s+1} \neq D'_{s+1}$ is at most 4. By the definition of $M_s$ and the above observations, the increment $|\Delta M_s|$ is equal to the sum below

$$\left| \sum_{i=1}^{t} \mathbb{E}\left[ \Pr(D_t(v_\tau) = d, Z_r = 1 | D_{s+1}(v_\tau)) - \Pr(D'_t(v_\tau) = d, Z_r = 1 | D'_{s+1}(v_\tau)) \right] \mathcal{F}_{s+1} \right| \quad (4.8)$$

and all we have concluded so far leads to the following bound from above

$$|M_{s+1} - M_s| \leq \mathbb{E}\left[ \sum_{i=1}^{t} \mathbb{1}\{D_{s+1}(i) \neq D'_{s+1}(i)\} | \mathcal{F}_{s+1} \right] \leq 4, \quad (4.9)$$

which concludes the proof. \qed

The next step is to bound the conditional quadratic variation of $\{M_s\}_{s \geq 1}$ in order to apply Freedman’s inequality, which is done in the lemma below.
Lemma 4 (Upper bound for the quadratic variation) Let \( f \) be any edge-step function and \( \{M_s\}_{s \geq 1} \) be as in (4.1). Then, the following bound holds

\[
\mathbb{E} \left[ (M_{s+1} - M_s)^2 \g F_s \right] \leq \frac{10d^2 N_s(\leq d, f)}{s},
\]

(4.10)

for all time \( s \) and degree \( d \).

Proof By (4.8) and Jensen’s inequality, we have that \((M_{s+1} - M_s)^2\) is bounded from above by

\[
\mathbb{E} \left[ \left( \sum_{r=1}^{t} \mathbb{P}(D_t(v_r) = d, Z_r = 1 \mid \mathcal{F}_{s+1}) - \mathbb{P}(D_t'(v_r) = d, Z_r' = 1 \mid \mathcal{F}_{s+1}') \right)^2 \g \mathcal{F}_{s+1} \right].
\]

(4.11)

Taking the conditional expectation w.r.t \( \mathcal{F}_s \), using the tower property and recalling that we must have \( D_s(v_r) \leq d \) yields

\[
\mathbb{E} \left[ (\Delta M_s)^2 \g \mathcal{F}_s \right] \leq \mathbb{E} \left[ \left( \sum_{r=1}^{t} \mathbb{I}(D_{s+1}(v_r) \neq D'_{s+1}(v_r)) \mathbb{I}(D_s(v_r) \leq d) \right)^2 \g \mathcal{F}_s \right].
\]

(4.12)

Now, observe that the following upper bound holds deterministically

\[
\mathbb{I}(D_{s+1}(v_r) \neq D'_{s+1}(v_r)) \leq \Delta D_s(v_r) + \Delta D'_s(v_r)
\]

(4.13)

and identities (3.3) and (3.4) give us

\[
\mathbb{E} \left[ \Delta D'_s(v_r) \g \mathcal{F}_s \right] = \mathbb{E} \left[ \Delta D_s(v_r) \g \mathcal{F}_s \right] = \left( 1 - \frac{f(s + 1)}{2} \right) \frac{D_s(v_r)}{s},
\]

(4.14)

which, in turn, leads to

\[
\mathbb{P}(D_{s+1}(v_r) \neq D'_{s+1}(v_r) \g \mathcal{F}_s) \leq \frac{2D_s(v_r)}{s},
\]

(4.15)

for all \( r \in \{1, \ldots, t\} \). For \( u \geq 1 \), using that the product \( \Delta D_s(v_r) \Delta D_s(v_u) \) is nonzero if and only if both vertices are selected at the same time and that \( \Delta D_s(v_r) \) and \( \Delta D'_s(v_u) \) are independent given \( \mathcal{F}_s \), we also derive

\[
\mathbb{P}(D_{s+1}(v_r) \neq D'_{s+1}(v_r), D_{s+1}(v_u) \neq D'_{s+1}(v_u) \g \mathcal{F}_s) \leq \frac{4D_s(v_r)D_s(v_u)}{s^2}.
\]

(4.16)
Expanding the summand on the RHS of (4.12) and substituting (4.15) and (4.16) in it, we obtain

\[
\mathbb{E}\left[ (\Delta M_s)^2 \bigg| \mathcal{F}_s \right] \leq 2 \sum_{r=1}^{t} D_s(v_r) \mathbbm{1}[D_s(v_r) \leq d] \]

\[
+ 8 \sum_{1 \leq r < u \leq t} \frac{D_s(v_r)D_s(v_u) \mathbbm{1}[D_s(v_r) \leq d, D_s(v_u) \leq d]}{s^2}
\]

\[
\leq 2dN_s(\leq d, f) + 8d^2 \sum_{1 \leq r < u \leq t} \frac{\mathbbm{1}[D_s(v_r) \leq d, D_s(v_u) \leq d]}{s^2}
\]

\[
\leq \frac{10d^2 N_s(\leq d, f)}{s}
\]

since

\[
\sum_{1 \leq r < u \leq t} \mathbbm{1}[D_s(v_r) \leq d, D_s(v_u) \leq d]
\]

\[
\leq \left( \sum_{r=1}^{t} \mathbbm{1}[D_s(v_r) \leq d] \right) \left( \sum_{u=1}^{t} \mathbbm{1}[D_s(v_u) \leq d] \right) = N_s^2(\leq d, f)
\]

and \(N_s(d, f)\) is less than \(s\) deterministically. This finishes the proof. \(\square\)

Now we are able to prove a general concentration result for \(N_t(d, f)\), which holds for any edge-step function \(f\). Then, we obtain Theorem 1 as a consequence of exploiting additional information about \(f\).

### 4.1 The General Case

For the general picture, our estimates of the deviation of \(N_t(d, f)\) from its expected value depend on

\[
\sum_{s=1}^{t} \frac{1}{s} \sum_{r=1}^{s} f(r)
\]

which cannot be well estimated in this degree of generality. In this section, we will prove a general concentration result, which holds for any \(f\), but later we will see that this result can be very sharp if more information on the asymptotic behavior of \(f\) is provided. For now, our goal is to prove the proposition below

**Proposition 4.1** Let \(f\) be any edge-step function. Then, for all \(\lambda > 0\) and \(d \in \mathbb{N}\) it follows that

\[
\mathbb{P}\left( |N_t(d, f) - \mathbb{E}[N_t(d, f)]| \geq \lambda \right) \leq \exp\left\{-\frac{\lambda^2}{2\sigma_{d,t}^2 + 8\lambda/3}\right\}
\]
\[ + \exp \left\{ -\frac{\lambda^2}{2F(t) + 4\lambda/3} \right\} \]  

(4.18)

where

\[ \sigma_{d,t}^2 := 10d^2 \sum_{s=1}^{t-1} \frac{F(s) + \lambda}{s} \].  

(4.19)

**Proof** We apply Freedman’s inequality (Theorem B2) to the Doob martingale \( \{M_s\}_{s \geq 1} \) defined on (4.1). Before, however, it will be important to control the number of vertices at time \( t \), \( V_t \). Recall that \( V_t \) is 1 plus the sum of the independent random variables \( Z_2, \ldots, Z_t \), and that \( Z_s \sim \text{Ber}(f(s)) \). Thus, \( V_t - F(t) \) is a mean zero martingale whose increments are bounded by 2. And since the \( Z_s \)'s are independent, it follows that

\[ t - 1 \sum_{s=1}^{t-1} \mathbb{E} \left[ (V_{s+1} - F(s + 1) - V_s + F(s))^2 \bigg| \mathcal{F}_s \right] \]

(4.20)

\[ = \sum_{s=1}^{t-1} \mathbb{E} \left[ (Z_{s+1} - f(s + 1))^2 \bigg| \mathcal{F}_s \right] \leq F(t). \]

Then, applying Freedman’s inequality on the martingale \( V_t - F(t) \), with \( \sigma^2 = F(t) \), we obtain that

\[ \mathbb{P} \left( \max_{s \leq t} \{ V_s - F(s) \} \geq \lambda \right) \leq \exp \left\{ -\frac{\lambda^2}{2F(t) + 4\lambda/3} \right\}. \]  

(4.21)

Now, for a fixed \( \lambda > 0 \), define the stopping time

\[ \tau := \inf \{ s \geq 1 \mid V_s - F(s) \geq \lambda \}. \]  

(4.22)

Observe that (4.21) gives us

\[ \mathbb{P} \left( \tau \leq t \right) = \mathbb{P} \left( \max_{s \leq t} \{ V_s - F(s) \} \geq \lambda \right) \leq \exp \left\{ -\frac{\lambda^2}{2F(t) + 4\lambda/3} \right\}. \]  

(4.23)

Now consider the stopped martingale \( \{M_{s \wedge \tau}\}_{s \geq 1} \), whose conditional quadratic variation is bounded in the following way

\[ \sum_{s=1}^{t-1} \mathbb{E} \left[ (\Delta M_{s \wedge \tau})^2 \bigg| \mathcal{F}_s \right] \leq 10d^2 \sum_{s=1}^{t-1} \frac{10d^2N_s(\leq d, f) \mathbb{I}\{s \leq \tau\}}{s} \]

(4.24)
deterministically, since the number of vertices having degree at most \( d \) is less than the total number of vertices, and \( V_s \leq F(s) + \lambda \) whenever \( s < \tau \). Recalling (4.19), we have that the LHS above is smaller than or equal to \( \sigma_{d,t}^2 \). Defining then

\[
W_t := \sum_{k=1}^{t-1} \mathbb{E} \left[ (M_{k+1} - M_k)^2 \bigg| \mathcal{F}_k \right],
\]

we have by Freedman’s inequality,

\[
\mathbb{P} \left( |M_t - \mathbb{E} N_t(d, f)| \geq \lambda, W_t \leq \sigma_{d,t}^2 \right) \leq \exp \left\{ -\frac{\lambda^2}{2\sigma_{d,t}^2 + 8\lambda/3} \right\}. (4.25)
\]

Finally, we obtain

\[
\mathbb{P} \left( |M_t - \mathbb{E} N_t(d, f)| \geq \lambda \right) \leq \mathbb{P} \left( |M_t - \mathbb{E} N_{t\wedge \tau}(d, f)| \geq \lambda, \tau > t \right) + \mathbb{P} \left( \tau \leq t \right)
\leq \exp \left\{ -\frac{\lambda^2}{2\sigma_{d,t}^2 + 8\lambda/3} \right\} + \exp \left\{ -\frac{\lambda^2}{2F(t) + 4\lambda/3} \right\}, (4.26)
\]

finishing the proof. \( \square \)

### 4.2 Index of Regular Variation in \((-1, 0]\)

Now, we will explore Proposition 4.1 when more properties of \( f \) are available in order to prove a general concentration measure result for \( \hat{P}_t(d, f) \). As we will see, information about the asymptotic behavior of \( f \) is enough to derive useful concentration results. Our goal is to prove that the fluctuations around the mean of \( N_t(d, f) \) are of order \( \sqrt{F(t)} \), which can be of order much smaller than \( \sqrt{t} \), as discussed in Sect. 2.

Recall that Proposition 3.1 assures us that \( \mathbb{E} [V_t(f) \hat{P}_t(d, f)] / \mathbb{E} V_t(f) \) is close to \( p_{\gamma}(d) \). The next proposition assures that \( \hat{P}_t(d, f) \) is concentrated around \( \mathbb{E} [V_t(f) \hat{P}_t(d, f)] / \mathbb{E} V_t(f) \).

**Proposition 4.2** Let \( f \in \text{RES}(-\gamma) \) with \( \gamma \in [0, 1) \). Then, for all \( d \in \mathbb{N} \) and \( A > 0 \) such that

\[
A < \frac{1}{4d \log(t)} \sqrt{\frac{\mathbb{E} V_t(f)}{1 - \gamma}}, (4.27)
\]

we have

\[
\left| \hat{P}_t(d, f) - \frac{\mathbb{E} \left[ V_t(f) \hat{P}_t(d, f) \right]}{\mathbb{E} V_t(f)} \right| \leq A \cdot \frac{10d}{\sqrt{(1 - \gamma) \mathbb{E} V_t(f)}}, (4.28)
\]

\( \square \) Springer
with probability at least \(1 - 3e^{-A^2/3}\).

**Proof** We apply Proposition 4.1 combined with the fact that we are now considering edge-step functions which are regularly varying, which gives us extra knowledge about the quantities involved in the statement of Proposition 4.1.

We begin observing that by Lemma 2 we have

\[
F(t) \sim \int_1^t f(s)ds \sim (1 - \gamma)^{-1} \ell(t)t^{1-\gamma},
\]

(4.29)

for \(\gamma \in [0, 1]\). Consequently, we have that

\[
\sum_{s=1}^t \frac{F(s) + \lambda}{s} \leq (1 - \gamma)^{-1} F(t) + \lambda \log(t).
\]

(4.30)

We set \(\lambda = A\sqrt{40d^2(1 - \gamma)^{-1}}F(t)\) with \(A < \sqrt{F(t)/(1 - \gamma)^{-1}(4d \log(t))^{-1}}\). Using Proposition 4.1, we obtain that for large enough \(t\)

\[
\mathbb{P} \left( |N_t(d, f) - \mathbb{E}[N_t(d, f)]| \geq A\sqrt{40d^2 F(t)(1 - \gamma)^{-1}} \right)
\]

\[
\leq \exp \left\{ - \frac{\lambda^2}{20d^2(1 - \gamma)^{-1} F(t) + \lambda (20d^2 \log(t) + 8/3)} \right\} + \exp \left\{ - \frac{\lambda^2}{2F(t) + 4\lambda/3} \right\}
\]

\[
\leq \exp \left\{ - \frac{A^2 \cdot 40d^2(1 - \gamma)^{-1} F(t)}{20d^2(1 - \gamma)^{-1} F(t) + A\sqrt{40d^2(1 - \gamma)^{-1} F(t) \cdot (20d^2 \log(t) + 8/3)}} \right\}
\]

\[
+ \exp \left\{ - \frac{A^2 \cdot 40d^2(1 - \gamma)^{-1} F(t)}{2F(t) + 4\lambda/3 \cdot A\sqrt{40d^2(1 - \gamma)^{-1} F(t)}} \right\}
\]

\[
\leq 2 \exp \left\{ -A^2 \right\}.
\]

To prove the proposition from the above result, note that by triangle inequality and the fact that \(N_t(d, f) \leq V_t(f)\) deterministically, we may obtain

\[
\left| \hat{P}_t(d) - \frac{\mathbb{E}N_t(d, f)}{F(t)} \right| \leq \left| \frac{N_t(d, f)(F(t) - V_t(f))}{V_t(f)F(t)} \right| + \left| \frac{N_t(d, f) - \mathbb{E}N_t(d, f)}{F(t)} \right|
\]

\[
\leq \left| \frac{V_t(f)}{F(t)} - 1 \right| + \left| \frac{N_t(d, f) - \mathbb{E}N_t(d, f)}{F(t)} \right|
\]

(4.32)

By the multiplicative form of the Chernoff bound, we have

\[
\mathbb{P} \left( \left| \frac{V_t}{F(t)} - 1 \right| > \frac{A}{\sqrt{F(t)}} \right) \leq \exp \left\{ - \frac{A^2}{3} \right\}.
\]

(4.33)

The second term is then bounded by (4.31), giving

\[
\mathbb{P} \left( \left| \hat{P}_t(d) - \frac{\mathbb{E}N_t(d, f)}{F(t)} \right| > 10d \frac{A}{\sqrt{(1 - \gamma)F(t)}} \right) \leq \exp \left\{ - \frac{A^2}{3} \right\} + 2e^{-A^2},
\]

(4.34)
finishing the proof of the proposition.

\[ \Box \]

5 The Case \( \gamma \in [1, \infty) \)

In this section, we prove Theorem 2, which states that when the index of regular variation is less than \(-1\) the empirical distribution \( \{ \hat{P}_t(d, f) \}_{t \in \mathbb{N}} \) converges to zero almost surely for any fixed \( d \). The mass on finite degrees is completely lost in this regime. We start by showing that this phenomenon happens in expectation.

**Proposition 5.1** Let \( f \in \text{RES}(-\gamma) \), with \( \gamma \in [1, \infty) \). Then, for all \( d \in \mathbb{N} \), we have

\[
\lim_{t \to \infty} \frac{\mathbb{E}N_t(d, f)}{F(t)} = 0.
\]

**Proof** We proceed by induction on \( d \). Again, by the representation theorem (Theorem A.2), there exists a slowly varying function \( \ell \) such that \( f(t) = t^{-\gamma} \ell(t) \) for all \( t \geq 1 \). In order to simplify our writing, we let \( a_t(d) \) be \( \mathbb{E}N_t(d, f) \).

**Base case of the induction** According to Lemma 1, we have, for \( d = 1 \)

\[
a_{t+1}(1) \leq \left( 1 - \frac{1}{t} + \frac{\ell(t)}{2t^{1+\gamma}} \right) a_t(1) + \frac{\ell(t + 1)}{(t + 1)^\gamma} + O \left( \frac{F(t)}{t^2} \right). \tag{5.1}
\]

Expanding the above recurrence relation yields

\[
a_{t+1}(1) \leq \frac{\ell(t + 1)}{(t + 1)^\gamma} + O \left( \frac{F(t)}{t^2} \right) + \sum_{s=1}^t \left[ \left( \frac{\ell(s)}{s^\gamma} + O \left( \frac{F(s)}{s^2} \right) \right) \prod_{r=s}^t \left( 1 - \frac{1}{r} + \frac{\ell(r)}{2r^{1+\gamma}} \right) \right] \leq \exp \left\{ \sum_{r=1}^\infty \frac{\ell(r)}{2r^{1+\gamma}} \right\} \sum_{s=1}^t \left[ \left( \frac{\ell(s)}{s^\gamma} + O \left( \frac{F(s)}{s^2} \right) \right) \exp \left\{ - \sum_{r=s}^t \frac{1}{r} \right\} \right] + o(1). \tag{5.2}
\]

When \( \gamma > 1 \) it is straightforward to verify that \( a_t(d) < C_d \) for all \( t \geq 1 \). Thus, from now on, we assume \( \gamma = 1 \), which is the hardest case. Observe that, by Karamata’s theorem (Theorem A1), it follows that

\[
\exp \left\{ \sum_{r=1}^\infty \frac{\ell(r)}{2r^{1+\gamma}} \right\} \leq c_1 \tag{5.3}
\]
and by Corollary A.3, we also have

$$\lim_{s \to \infty} \frac{F(s)}{s} = 0 \implies \lim_{t \to \infty} \frac{1}{t} \sum_{s=1}^{t} \frac{F(s)}{s} = 0. \quad (5.4)$$

By Lemma 2, for large enough $t$, we have that

$$\frac{1}{t} \sum_{s=1}^{t} \ell(s) \leq \frac{2t \ell(t)}{t} = 2\ell(t). \quad (5.5)$$

Therefore, we have that, for some positive constant $C$,

$$a_t(1) \leq C \ell(t) \quad (5.6)$$

and by Corollary A.1 (whose proof we postpone to “Appendix”) it follows that

$$\lim_{t \to \infty} \frac{a_t(1)}{F(t)} = 0. \quad (5.7)$$

concluding the base step.

Inductive step Assume that for all $k \leq d - 1$ there exists $C_k$ such that

$$a_t(k) \leq C_k \ell(t). \quad (5.8)$$

Recall the recurrence relation given by (3.2), which gives us

$$a_{t+1}(d) \leq \left(1 - \frac{d}{t} + \frac{d \ell(t)}{2t^{1+\gamma}}\right) a_t(d) + \left(\frac{d - 1}{t} - \frac{(d - 1)\ell(t)}{2t^{1+\gamma}}\right) a_t(d - 1) + O_d \left(\frac{F(t)}{t^2}\right).$$

Expanding the above equality and recalling that $\gamma = 1$, we obtain

$$a_{t+1}(d) = \sum_{s=1}^{t} \left[\left(\frac{d - 1}{s} - \frac{(d - 1)\ell(s)}{2s^2}\right) a_s(d - 1) + O_d \left(\frac{F(s)}{s^2}\right)\right] \times \prod_{r=s+1}^{t} \left(1 - \frac{d}{r} + \frac{d \ell(r)}{2r^2}\right) \quad (5.9)$$

$$\leq \frac{c_d}{t^d} \sum_{s=1}^{t} \left[s^{d-1} a_s(d - 1) + O_d \left(F(s)s^{d-2}\right)\right].$$
From Corollary A.3, it follows that, for some \( \varepsilon > 0 \),

\[
\lim_{t \to \infty} \frac{c_d}{t^d} \sum_{s=1}^{t} O_d \left( \frac{F(s)}{s} s^{d-1} \right) \leq \frac{c_d}{t^d} \sum_{s=1}^{t} O_d \left( \frac{s^\varepsilon}{s} s^{d-1} \right) \leq c_d t^{-(1-\varepsilon)}
\]

Finally, the inductive hypothesis and Karamata’s theorem lead to

\[
\frac{1}{t^d} \sum_{s=1}^{t} s^{d-1} a_s (d - 1) \leq \frac{C_{d-1}}{t^{d'}} \sum_{s=1}^{t} s^{d-1} \ell(s) \leq c_d' \ell(t),
\]

proving the inductive step, since \( \ell(t) \geq t^{-(1-\varepsilon)} \) for sufficiently large \( t \).

Combining (5.8) with Corollary A.1, it is proved that

\[
\lim_{t \to \infty} \frac{a_t(d)}{F(t)} = 0
\]

for all \( d \in \mathbb{N} \), finishing the proof. \( \square \)

From Proposition 5.1, we will prove the a.s. convergence employing a second moment estimate. For this, we will need a new definition and a few lemmas.

**Definition 1 (d-admissible vectors)** Given \( d, t, r, s \in \mathbb{N} \), with \( r < s < t \) and two vertices \( v_s \) and \( v_r \) born at time \( r \) and \( s \), respectively, we say that two vectors \( \mathbf{x}_{s,t} \) and \( \mathbf{y}_{r,t} \) are \( d \)-admissible for \( v_s \) and \( v_r \) if \( x_u, y_u \in \{0, 1, 2\} \) for all \( u \), the sum of their coordinates is at most \( d \), \( y_s \neq 2 \) and the vectors do not have a 2 in the same coordinate.

Observe that given a vertex \( v_s \), the vector \( \mathbf{x}_{s,t} \in \{0, 1, 2\}^{t-s} \) induces an event in which the trajectory of the degree of \( v_s \) up to time \( t \) is completely characterized by said vector. More specifically, \( \mathbf{x}_{s,t} = (x_u)_{u=s+1}^{t} \) characterizes the event

\[
\{ \Delta D_t(v_s) = x_t \} \cap \cdots \cap \{ \Delta D_{s+1}(v_s) = x_{s+1} \} \cap \{ Z_s = 1 \}.
\]

Thus, two vectors are \( d \)-admissible if the events induced by them imply that both \( D_t(v_r) \) and \( D_t(v_s) \) are at most \( d \) and that their intersection is not empty. Moreover, given two \( d \)-admissible vectors \( \mathbf{x}_{s,t} \) and \( \mathbf{y}_{r,t} \), we denote by \( \mathbb{P}_{\mathbf{x}_{s,t}, \mathbf{y}_{r,t}} \) the distribution \( \mathbb{P} \) conditioned on the intersection of the events induced by the vectors. Also, to simplify our writing, fixed the vertices \( v_s \) and \( v_r \) and two \( d \)-admissible vectors, we write for all \( u \)

\[
\Delta_u := \Delta D_u(v_s); \quad \Delta_u' := \Delta D_u(v_r).
\]

The following lemma is the first step in obtaining a decorrelation estimate that will allow us to estimate the variance of \( N_t(\leq d, f) \), the number of vertices at time \( t \) with degree lesser than or equal to \( d \).
Lemma 5 Let $\vec{x}_{s,t+1} = (x_{u})_{u=s+1}^{t+1}$ and $\vec{y}_{r,t+1} = (y_{u})_{u=r+1}^{t+1}$ be two $d$-admissible vectors for some $d \in \mathbb{N}$ and vertices $v_{s}$ and $v_{r}$. Then,

$$
\mathbb{P}_{\vec{x}_{s,t},\vec{y}_{r,t}} (\Delta_t = x_{t+1}, \Delta'_r = y_{t+1}) \\
\leq \left( 1 + O \left( \frac{\ell(t) + d}{t} \right) \right) \mathbb{P}_{\vec{x}_{s,t},\vec{y}_{r,t}} (\Delta_t = x_{t+1}) \mathbb{P}_{\vec{x}_{s,t},\vec{y}_{r,t}} (\Delta'_r = y_{t+1}),
$$

(5.11)

for all $t > s$. Furthermore, for the special case where $x_{t+1} = y_{t+1} = 0$, we have, also for all $t > s$,

$$
\mathbb{P}_{\vec{x}_{s,t},\vec{y}_{r,t}} (\Delta_t = 0, \Delta'_r = 0) \leq \mathbb{P}_{\vec{x}_{s,t},\vec{y}_{r,t}} (\Delta_t = 0) \mathbb{P}_{\vec{x}_{s,t},\vec{y}_{r,t}} (\Delta'_r = 0).
$$

(5.12)

Proof The proof is done by direct computation. We compute the probabilities of all possible combinations for $x_{t+1}$ and $y_{t+1}$ in $\{0, 1, 2\}$ and compare them. We will write the degree $d_{t}(v_{s})$ in lower case meaning the degree of $v_{s}$ at time $t$ according to the event induced by the vector $\vec{x}_{s,t}$, analogously defining $d_{t}(v_{r})$ for $v_{r}$. Note that since the two vectors are $d$-admissible, $d_{t}(v_{s})$ and $d_{t}(v_{r})$ are both less than $d$. We have

$$
\mathbb{P}_{\vec{x}_{s,t},\vec{y}_{r,t}} (\Delta_t = 0) = \left( 1 - \frac{d_{t}(v_{s})}{2t} \right) \left[ 1 - (1 - f(t+1)) \frac{d_{t}(v_{s})}{2t} \right],
$$

(5.13)

$$
\mathbb{P}_{\vec{x}_{s,t},\vec{y}_{r,t}} (\Delta_t = 1) = \frac{d_{t}(v_{s})}{2t} \times \left( f(t+1) + 2(1 - f(t+1)) - 2(1 - f(t+1)) \frac{d_{t}(v_{s})}{2t} \right)
\times \left[ 1 + O(f(t+1) + dt^{-1}) \right],
$$

(5.14)

And finally

$$
\mathbb{P}_{\vec{x}_{s,t},\vec{y}_{r,t}} (\Delta_t = 2) = (1 - f(t+1)) \frac{d_{t}^{2}(v_{s})}{4t^{2}}.
$$

(5.15)

Now, we consider the cases in which $\Delta_{u}$ and $\Delta'_{u}$ change simultaneously.

$$
\mathbb{P}_{\vec{x}_{s,t},\vec{y}_{r,t}} (\Delta_t = 0, \Delta'_r = 0) \\
= \left( 1 - \frac{d_{t}(v_{s})}{2t} - \frac{d_{t}(v_{r})}{2t} \right) \left[ 1 - (1 - f(t+1)) \left( \frac{d_{t}(v_{s})}{2t} + \frac{d_{t}(v_{r})}{2t} \right) \right],
$$

(5.16)

$$
\mathbb{P}_{\vec{x}_{s,t},\vec{y}_{r,t}} (\Delta_t = 1, \Delta'_r = 0) \\
= \frac{d_{t}(v_{s})}{2t} f(t+1) + 2(1 - f(t+1)) \frac{d_{t}(v_{s})}{2t} \left( 1 - \frac{d_{t}(v_{s})}{2t} - \frac{d_{t}(v_{r})}{2t} \right) \\
= \frac{d_{t}(v_{s})}{t} (1 + O(f(t+1) + dt^{-1})).
$$

(5.17)

$$
\mathbb{P}_{\vec{x}_{s,t},\vec{y}_{r,t}} (\Delta_t = 1, \Delta'_r = 1) = 2(1 - f(t+1)) \frac{d_{t}(v_{s})d_{t}(v_{r})}{4t^{2}}.
$$

(5.18)
Finally,

\[ P_{\bar{x}_{s,t},\bar{y}_{r,t}}(\Delta_t = 2, \Delta'_t = 0) = (1 - f(t + 1)) \frac{d_l^2(v_s)}{4t^2}. \] (5.19)

These cases are enough to cover all possible combinations. The result then follows by a direct comparison between product of probabilities given by (5.13,5.14, 5.15) with those obtained in (5.16,5.17, 5.18, 5.19). In particular, we note that from (5.13) and (5.16) we obtain

\[
P_{\bar{x}_{s,t},\bar{y}_{r,t}}(\Delta_t = 0) P_{\bar{x}_{s,t},\bar{y}_{r,t}}(\Delta'_t = 0)
= \left(1 - \frac{d_l(v_s)}{2t} - \frac{d_l(v_r)}{2t} + \frac{d_l(v_s) d_l(v_r)}{4t^2}\right)
\times \left[1 - (1 - f(t + 1)) \left(\frac{d_l(v_s)}{2t} + \frac{d_l(v_r)}{2t}\right) + (1 - f(t + 1))^2 \frac{d_l(v_s) d_l(v_r)}{4t^2}\right]
\geq P_{\bar{x}_{s,t},\bar{y}_{r,t}}(\Delta_t = 0, \Delta'_t = 0),
\] (5.20)

finishing the proof of the lemma.$\square$

For a fixed vertex $v_s$ and $d \in \mathbb{N}$, define the event

\[ E_{t,d}(v_s) := \{D_t(v_s) \leq d, Z_s = 1\}. \] (5.21)

In the next lemma, we prove that the events $E_{t,d}(v_r)$ and $E_{t,d}(v_s)$ are almost uncorrelated.

**Lemma 6** For $d, r, s, t \in \mathbb{N}$ with $r < s \leq t$, we have

\[ \mathbb{P}(E_{t,d}(v_s), E_{t,d}(v_r)) \leq \left(1 + dO(\frac{\ell(s) + d}{s})\right) \mathbb{P}(E_{t,d}(v_s)) \mathbb{P}(E_{t,d}(v_r)). \] (5.22)

**Proof** Fix two $d$-admissible vectors $\bar{x}_{s,t} = (x_u)_{u=s+1}^t$ and $\bar{y}_{r,t} = (y_u)_{u=r+1}^t$ and denote by $\Xi(\bar{x}_{s,t})$ the event

\[ \Xi(\bar{x}_{s,t}) := \{\Delta_{t-1} = x_t \} \cap \cdots \cap \{\Delta_s = x_{s+1} \} \cap \{Z_s = 1\}, \] (5.23)

and analogously define $\Xi(\bar{y}_{r,t})$. Observe that for each $m \in s + 1, \ldots, t$ we have

\[ P_{\bar{x}_{s,m},\bar{y}_{r,m}}(\Delta_m = x_{m+1}) = \mathbb{P}(\Delta_m = x_{m+1} \mid \Delta_{m-1} = x_m, \ldots, \Delta_s = x_{s+1}, Z_s = 1), \] (5.24)

where $P_{\bar{x}_{s,m},\bar{y}_{r,m}}$ is defined for the restrictions of the vectors $\bar{x}_{s,t}$ and $\bar{y}_{r,t}$ up to time $m$. We apply Lemma 5 iteratively and then use (5.24) to regroup the terms in a convenient
way. For the first step, we note that
\[
\mathbb{P}(\Xi(\vec{x}_{s,t}), \Xi(\vec{y}_{r,t})) \\
\leq (1 + O(\frac{\ell(t) + d}{t})) \mathbb{P}_{\vec{x}_{s,t-1}, \vec{y}_{r,t-1}}(\Delta_{t-1} = x_t) \mathbb{P}_{\vec{x}_{s,t-1}, \vec{y}_{r,t-1}}(\Delta'_{t-1} = y_t) \\
\times \mathbb{P}(\Xi(\vec{x}_{s,t-1}), \Xi(\vec{y}_{r,t-1}))
\]

We iterate this procedure until \( u = s + 1 \). The case \( u = s \) we must handle in a slightly different way. Note that, when \( u = s \) we have to deal with the term
\[
\mathbb{P}_{\vec{x}_{s,s-1}, \vec{y}_{r,s-1}}(\Delta'_s = y_s, Z_s = 1) = \mathbb{P}_{\vec{x}_{s,s-1}, \vec{y}_{r,s-1}}(\Delta'_s = y_s | Z_s = 1) f(s), \quad (5.25)
\]

since \( Z_s \) is independent of \( \mathcal{F}_{s-1} \). Now, for \( y_s = 1 \) we have
\[
\frac{\mathbb{P}_{\vec{x}_{s,s-1}, \vec{y}_{r,s-1}}(\Delta'_s = 1 | Z_s = 1)}{\mathbb{P}_{\vec{x}_{s,s-1}, \vec{y}_{r,s-1}}(\Delta'_s = 1)} = \frac{d_{s-1}(v_r)}{f(s) \frac{d_{s-1}(v_r)}{2(s-1)} + 2(1 - f(s)) \frac{d_{s-1}(v_r)}{2(s-1)} \left(1 - \frac{d_{s-1}(v_r)}{2(s-1)}\right)} = \frac{1}{2 - f(s) - 2(1 - f(s)) \frac{d_{s-1}(v_r)}{2(s-1)}} = \frac{1}{2}(1 + O(f(s) + ds^{-1})). \quad (5.26)
\]

And for \( y_s = 0 \), we get
\[
\frac{\mathbb{P}_{\vec{x}_{s,s-1}, \vec{y}_{r,s-1}}(\Delta'_s = 0 | Z_s = 1)}{\mathbb{P}_{\vec{x}_{s,s-1}, \vec{y}_{r,s-1}}(\Delta'_s = 0)} = \frac{1 - \frac{d_{s-1}(v_r)}{2(s-1)}}{f(s) \left(1 - \frac{d_{s-1}(v_r)}{2(s-1)}\right) + (1 - f(s)) \left(1 - \frac{d_{s-1}(v_r)}{2(s-1)}\right)^2} = (1 + O(f(s) + ds^{-1})). \quad (5.27)
\]

Iterating the procedure, we obtain
\[
\mathbb{P}(\Xi(\vec{x}_{s,t}), \Xi(\vec{y}_{r,t})) \leq \prod_{u=3}^t \left(1 + O\left(\frac{\ell(u) + d}{u}\right)\right) \times \mathbb{P}_{\vec{x}_{u,u-1}, \vec{y}_{r,u-1}}(\Delta_{u-1} = x_u) \mathbb{P}_{\vec{x}_{u,u-1}, \vec{y}_{r,u-1}}(\Delta'_{u-1} = y_u) \mathbb{P}(\Xi(\vec{y}_{r,s-1})).
\]

Note that, since \( \vec{x}_{s,t} \) and \( \vec{y}_{r,t} \) are \( d \)-admissible, in all but at most \( 2d \) steps the increments are both 0. Therefore, by (5.20) and the fact that \( f(u) = \ell(u)/u \) is non-increasing, we can use (5.24) to regroup separately all terms involving \( v_s \) and \( v_r \) to obtain

\( \Box \) Springer
\[ P(\Xi(\tilde{x}_r,t), \Xi(\tilde{y}_r,t)) \leq \left( 1 + O\left( \frac{\ell(s) + d}{s} \right) \right)^{2d} P(\Xi(\tilde{x}_s,t)) P(\Xi(\tilde{y}_s,t)) \]

\[
(5.28)
\]

We can then use the above equation to get

\[
P(\mathbb{E}_t, d(v_s), \mathbb{E}_t, d(v_r)) = \sum_{\tilde{x}_s,t, \tilde{y}_r,t \text{ d-admissible}} P(\Xi(\tilde{x}_s,t)) P(\Xi(\tilde{y}_r,t)) \]

\[
\leq \left( 1 + dO\left( \frac{\ell(s) + d}{s} \right) \right) \sum_{\tilde{x}_s,t, \tilde{y}_r,t \text{ d-admissible}} P(\Xi(\tilde{x}_s,t)) P(\Xi(\tilde{y}_r,t))
\]

\[
\leq \left( 1 + dO\left( \frac{\ell(s) + d}{s} \right) \right) P(\mathbb{E}_t, d(v_s)) P(\mathbb{E}_t, d(v_r)),
\]

finishing the proof of the lemma. \(\square\)

The next lemma shows that it is hard for earlier vertices to have small degrees.

**Lemma 7** Let \( \delta \in (0, 1) \) and \( d \in \mathbb{N} \). Then, for \( r \leq t^{1-\delta} \) we have that

\[
P(D_t(v_r) \leq d \mid Z_r = 1) \leq e^d t^{-\delta/4}.
\]

**Proof** Let \( \tilde{G}_r \) be a possible realization of the process \( (G_t)_{t \geq 1} \) at time \( r \) such that the vertex \( v_r \) belongs to \( V(\tilde{G}_r) \), and let \( \tilde{P}_{\tilde{G}_r} \) be the distribution \( \tilde{P} \) conditioned on the event where \( G_r = \tilde{G}_r \). By the simple Markov property, \( \tilde{P}_{\tilde{G}_r} \) has the same distribution as our model started from \( \tilde{G}_r \). Now from (5.13) and (5.14), we obtain, for any step \( u \geq r \),

\[
P_{\tilde{G}_r}(\Delta D_u(v_r) \geq 1 | \mathcal{F}_u) = (2 - f(u + 1)) \frac{D_u(v_r)}{2u} - (1 - f(u + 1)) \frac{D_u^2(v_r)}{4u^2} \geq \frac{D_u(v_r)}{2u},
\]

since \( D_u(v_r) \leq 2u \) deterministically. Using the fact that the degree is at least one, we obtain that \( D_r(v_r) \) dominates a sum of independent random variables \( \{Y_u\}_{u = t^{1-\delta}} \) where \( Y_u \overset{d}{=} \text{Ber}(1/2u) \). Observe that, bounding the sum by the integral, we obtain the following lower bound for the expectation of the degree of \( v_r \) under \( \tilde{P}_{\tilde{G}_r} \)

\[
\mu_r := \mathbb{E}_{\tilde{G}_r} \left[ \sum_{u = t^{1-\delta}}^t Y_u \right] \geq \frac{\delta}{2} \log t.
\]

\(\square\) Springer
Consequently, taking $\varepsilon$ as
\[ \varepsilon = 1 - \frac{d}{\mu_t} \quad (5.32) \]
and applying the Chernoff bound leads to
\[
\mathbb{P}_{\tilde{G}_r} (D_t(v_r) \leq d) \leq \mathbb{P} \left( \sum_{u=t^{1-\delta}}^{t} Y_u \leq (1 - \varepsilon) \mu_t \right) \leq \exp \left\{ -\frac{\varepsilon^2 \mu_t}{2} \right\} \leq \frac{e^d}{t^{\delta/4}}. \quad (5.33)
\]

Integrating over all possible graphs $\tilde{G}_r$ gives the desired result. \hspace{1cm} \square

We have now the ingredients needed in order to bound $\text{Var} (N_t(\leq d, f))$, which in turn will let us finish the argument using the Chebyshev inequality, the Borel–Cantelli lemma and an elementary subsequence argument.

Lemma 8 For any $d \in \mathbb{N}$ and $f \in \text{RES}(-\gamma)$, with $\gamma \in [1, \infty)$, we have
\[
\text{Var} (N_t(\leq d, f)) \leq \mathbb{E} N_t(\leq d, f) (1 + o(1))
\]

**Proof** By definition, we may write $N_t(\leq d, f)$ as
\[
N_t(\leq d, f) = \sum_{s \leq t} \mathbb{I} \{ D_t(v_s) \leq d \} Z_s = \sum_{s \leq t} \mathbb{I} \{ E_{t,d}(v_s) \}. \quad (5.34)
\]

To control $N_t^2(\leq d, f)$, we split the sum over $s$ into two sets: the vertices added before $t^{1-\delta}$ and those added after, for some small $\delta$. By Lemma 6, for $s > t^{1-\delta}$ and $r < s$, we have that
\[
\mathbb{P} \left( E_{t,d}(v_s), E_{t,d}(v_r) \right) - \mathbb{P} \left( E_{t,d}(v_s) \right) \mathbb{P} \left( E_{t,d}(v_r) \right) \leq \frac{C (\ell(t^{1-\delta}) + d)}{t^{1-\delta}} \mathbb{P} \left( E_{t,d}(v_s) \right) \mathbb{P} \left( E_{t,d}(v_r) \right). \quad (5.35)
\]

Thus,
\[
\sum_{s=t^{1-\delta}}^{t} \sum_{r=1}^{s-1} \mathbb{P} \left( E_{t,d}(v_s), E_{t,d}(v_r) \right) - \mathbb{P} \left( E_{t,d}(v_s) \right) \mathbb{P} \left( E_{t,d}(v_r) \right) \leq C \frac{\mathbb{E} \left[ N_t(\leq d, f) \right]^2 (\ell(t^{1-\delta}) + d)}{t^{1-\delta}} \quad (5.36)
\]

\[
\text{Lemma } 5.1 \quad \leq C \frac{\mathbb{E} \left[ N_t(\leq d, f) \right] (\ell(t^{1-\delta}) + d)}{t^{1-\delta}} \leq o(\mathbb{E} \left[ N_t(\leq d, f) \right]).
\]
Using Lemmas 6 and 7, and assuming \( r < s \leq t^{1-\delta} \), we get
\[
\mathbb{P} \left( E_{t,d}(v_r), E_{t,d}(v_s) \right) \leq C \mathbb{P} \left( E_{t,d}(v_r) \right) \mathbb{P} \left( E_{t,d}(v_s) \right) \\
\leq C \mathbb{P} \left( E_{t,d}(v_r) \right) \mathbb{P} \left( D_t(v_r) \leq d | Z_r = 1 \right) f(r) \leq C \mathbb{P} \left( E_{t,d}(v_s) \right) t^{-\delta/4} f(r). \tag{5.37}
\]

We therefore have
\[
\sum_{s=1}^{t^{1-\delta}} \sum_{r=1}^{s-1} \mathbb{P} \left( E_{t,d}(v_r), E_{t,d}(v_s) \right) \leq C \mathbb{E} \left[ N_t(\leq d, f) \right] \frac{F(t^{1-\delta})}{t^{\delta/4}} \\
= o \left( \mathbb{E} \left[ N_t(\leq d, f) \right] \right), \tag{5.38}
\]
which implies, together with (5.36),
\[
\text{Var} \left( N_t(\leq d, f) \right) = \sum_{s=1}^{t} \mathbb{P} \left( E_{t,d}(v_s) \right) + 2 \sum_{s=1}^{t} \sum_{r=1}^{s-1} \mathbb{P} \left( E_{t,d}(v_r), E_{t,d}(v_s) \right) \leq \mathbb{E} N_t(\leq d, f) + o(\mathbb{E} N_t(\leq d, f)) \tag{5.39}
\]
finishing the proof. \( \square \)

Now, we have all the tools needed to prove Theorem 2

**Proof of Theorem 2** We will be only interested in the case where \( F(t) \uparrow \infty \), since otherwise, by Borel–Cantelli lemma, there are only finitely many vertices in the graph. Given \( \varepsilon > 0 \), let \( t_k \) be the following deterministic index:
\[
t_k = \inf \{ s > 0 \; ; \; F(s) \geq (1 + \varepsilon)^k \}, \tag{5.40}
\]
which exists because we are assuming \( F(t) \) goes to infinity. Chebyshev inequality implies then, for every \( \delta > 0 \),
\[
\mathbb{P} \left( \left( \frac{V_{t_k}(f)}{F(t_k)} - 1 \right) > \delta \right) \leq \frac{\sum_{s=1}^{t_k} f(s)(1 - f(s))}{\delta^2 F(t_k)^2} \leq \delta^{-2} (1 + \varepsilon)^{-k}, \tag{5.41}
\]
implying that
\[
\frac{V_{t_k}(f)}{F(t_k)} \to 1 \text{ a.s. as } k \to \infty.
\]

Combining Lemma 8 with the Chebyshev inequality, we also get, for any \( \delta > 0 \),
\[
\mathbb{P} \left( \left( \frac{N_{t_k}(\leq d, f)}{F(t_k)} - \mathbb{E}(N_{t_k}(\leq d, f)) \right) > \delta \right) \leq \frac{\text{Var}(N_{t_k}(\leq d, f))}{\delta^2 (1 + \varepsilon)^{2k}} \leq C \delta^{-2} (1 + \varepsilon)^{-2k}, \tag{5.42}
\]
and the Borel–Cantelli lemma together with Proposition 5.1 then implies
\[
\lim_{t \to \infty} \frac{N_{t_k}(\leq d, f)}{F(t_k)} = \lim_{t \to \infty} \frac{\mathbb{E}(N_{t_k}(\leq d, f))}{F(t_k)} = 0 \quad (5.43)
\]
almost surely. Now, for \( s \in (t_{k-1}, t_k) \) the fact that \( V_t(f) \) and \( F(t) \) are non-decreasing leads to
\[
\frac{V_s(f)}{F(s)} \geq \frac{V_{t_{k-1}}(f)}{(1 + \varepsilon)^k} \geq \frac{V_{t_{k-1}}(f)}{F(t_{k-1})(1 + \varepsilon)} > 1 - 2\varepsilon, \quad (5.44)
\]
for sufficiently small \( \varepsilon \). Therefore, since \( \varepsilon \) was chosen arbitrarily,
\[
\frac{V_t(f)}{F(t)} \to 1 \quad \text{a.s. as } t \to \infty, \quad (5.45)
\]
implying
\[
\lim_{t \to \infty} \frac{N_{t_k}(> d, f)}{F(t_k)} = 1, \quad (5.46)
\]
a.s. since \( N_t(> d, f) = V_t(f) - N_t(\leq d, f) \). But observe that \( N_t(> d, f) \) is also non-decreasing in \( t \). Then, for \( s \in (t_{k-1}, t_k) \) we have that
\[
\frac{N_s(> d, f)}{F(s)} \geq \frac{N_{t_{k-1}}(> d, f)}{(1 + \varepsilon)^k} \geq \frac{N_{t_{k-1}}(> d, f)}{F(t_{k-1})(1 + \varepsilon)} > 1 - 2\varepsilon \quad (5.47)
\]
for sufficiently small \( \varepsilon \), which is enough to conclude that the whole sequence \( N_s(> d, f)/F(s) \) converges to 1 a.s, and consequently \( N_s(\leq d, f)/F(s) \) converges to zero a.s. This, together with (5.45), concludes the proof.

\( \square \)

6 Final Remarks

6.1 Edge-Step Functions VS Affine PA Rule

One of the natural generalizations of the preferential attachment rule proposed in [4] is known as the affine preferential attachment rule. One may introduce a constant \( \delta > -1 \) on the PA rule (1.1) so that the probability of a new vertex \( v \) connecting to a previous one \( u \) is now given by
\[
P(v \to u | G) = \frac{\text{degree}(u) + \delta}{\sum_{w \in G} (\text{degree}(w) + \delta)}. \quad (6.1)
\]

This slight modification is capable of producing graphs obeying a power-law degree distribution with a tunable exponent lying on \( (2, \infty) \). It would be natural to ask the effects on the degree distribution of an affine version of our model, since the addition
of \( \delta \) may lower the degree distribution’s tail, whereas the edge-step function may lift it. However, the effect of the edge-step function overcomes the presence of \( \delta \) in (1.1) in the long term. We give here some indications of why this is true for \( f \in \text{RES}(-\gamma) \), with \( \gamma \in [0, 1] \).

For an affine version of our model, one may start evaluating identity (3.3), which is crucial for proof of Proposition 3.1, to obtain

\[
P(\Delta D_t(v) = 1 | \mathcal{F}_t) = \left( 1 - \frac{f(t+1)}{2} \right) \frac{D_t(v) + \delta}{t + V_t(f)\delta} - 2 (1 - f(t+1)) \frac{(D_t(v) + \delta)^2}{(2t + V_t(f)\delta)^2},
\]

However, for \( f \in \text{RES}(-\gamma) \), with \( \gamma \in [0, 1] \), we have by (5.45) that \( V_t(f) = o(t) \). Thus, the above equation is also equal to

\[
\left( 1 - \frac{f(t+1)}{2} \right) \frac{D_t(v) + \delta}{t} (1 + o(1)) - 2 (1 - f(t+1)) \frac{(D_t(v) + \delta)^2}{4t^2} (1 + o(1)).
\]

The same goes for (3.4). The same sort of computation also leads to

\[
\mathbb{E}[\Delta D_t(v)|\mathcal{F}_t] = \left( 1 - \frac{f(t+1)}{2} \right) \frac{D_t(v) + \delta}{t} (1 + o(1)),
\]

which is not the case on the usual affine models. The above identities imply that many of the recursive computations one usually makes regarding these models are not altered by the introduction of the term \( \delta \). In particular, one can use the above recursion and elementary analysis to show that a process with this affine rule produces graphs with the same power-law exponents as the non affine rule \( (\delta = 0) \), though an analogous result to Proposition 3.1 would be significantly more involved. We therefore opted to focus the results in this paper on the case \( \delta = 0 \).

### 6.2 Maximum Degree

Since all the choices are made following the preferential attachment rule, the first vertices on the graph are good candidates for being the ones of highest degree. In this sense, estimates on their degree usually give the exact order of the maximum degree. In this subsection, we estimate the expected degree for the first vertex in order to argue that the presence of edge-step functions may also shape other graph observables, as the maximum degree.

From Eqs. (3.3) and (3.4), one may deduce the recurrence relation below for the expected degree of the very first vertex in the graph

\[
\mathbb{E}[D_{t+1}(1)] = \left( 1 + \frac{1}{t} - \frac{f(t+1)}{2t} \right) \mathbb{E}[D_t(1)],
\]

(6.2)
which implies

$$
\mathbb{E} \left[ D_{t+1}(1) \right] = 2 \prod_{s=2}^{t} \left( 1 + \frac{1}{s} - \frac{f(s+1)}{2s} \right) \\
\sim 2 \exp \left\{ \sum_{s=2}^{t} \left( \frac{1}{s} - \frac{f(s+1)}{2s} \right) \right\}. \tag{6.3}
$$

If $f$ is taken to be a regularly varying function with index of regular variation $\gamma \in (-\infty, 0)$, then, by the representation theorem and Karamata’s theorem (Corollary A.2 and Theorem A1, respectively), it follows that $\sum_{s=2}^{\infty} f(s)s^{-1}$ is finite and consequently

$$
\mathbb{E} \left[ D_{t+1}(1) \right] \approx t.
$$

For a slowly varying $f$, the order of $\mathbb{E} \left[ D_{t+1}(1) \right]$ depends on

$$
\int_{2}^{t} \frac{f(s)}{s} \, ds.
$$

If $f(s) = \log^{-1}(s)$, we have that

$$
\mathbb{E} \left[ D_{t+1}(1) \right] \approx tf(t).
$$

Whereas, for $f(s) = \log^{-2}(s)$, order $t$ is again achieved. The above discussion suggests that even when $f$ is slowly varying, which produces graphs with power-law exponent equal to 2, the maximum degree may still be $f$ dependent. One interesting question would be to determine the precise order of the maximum degree in terms of $f$ when it is taken to be a slowly varying function.

**Acknowledgements** C. A. was supported by the Deutsche Forschungsgemeinschaft (DFG). R. R. was partially supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). R.S. has been partially supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and by FAPEMIG (Programa Pesquisador Mineiro), Grant PPM 00600/16.

**Appendix A. Important Results on Regularly Varying Functions**

In this “Appendix,” we collect some results regarding regularly varying functions that will be useful throughout the paper, as well as providing a proof for an earlier lemma.

**Corollary A.1** Let $\ell$ be a slowly varying function. Then,

$$
\lim_{t \to \infty} \frac{\ell(t)}{\sum_{s=1}^{t} \ell(s)s^{-1}} = 0 \tag{A.1}
$$
Proof Since \( s^{-1} \ell(s) \) is a RV function which is eventually monotone, we may bound the sum by the integral. Now, by Theorem 1.5.2 of [5], for a fixed small \( \varepsilon \), we know that

\[
\lim_{t \to \infty} \frac{\ell(t)}{\ell(t)} = 1
\]

uniformly for \( x \in [\varepsilon, 1] \). Therefore, for large enough \( t \)

\[
\ell^{-1}(t) \int_{1}^{t} \frac{\ell(s)}{s} ds \geq \int_{\varepsilon}^{1} \frac{\ell(tx)}{\ell(t)} dx \geq -(1 - \delta) \log \varepsilon,
\]

for some small \( \delta \). This proves the desired result. \( \square \)

The three following results are used throughout the paper.

**Corollary A.2** (Representation theorem Theorem 1.4.1 of [5]) Let \( f \) be a continuous regularly varying function with index of regular variation \( \gamma \). Then, there exists a slowly varying function \( \ell \) such that

\[
f(t) = t^\gamma \ell(t),
\]

for all \( t \) in the domain of \( f \).

**Corollary A.3** Let \( f \) be a continuous regularly varying function with index of regular variation \( \gamma < 0 \). Then,

\[
f(x) \to 0,
\]

as \( x \) tends to infinity. Moreover, if \( \ell \) is a slowly varying function, then for every \( \varepsilon > 0 \)

\[
x^{-\varepsilon} \ell(x) \to 0 \text{ and } x^\varepsilon \ell(x) \to \infty
\]

**Proof** Comes as a straightforward application of Theorem 1.3.1 of [5] and Corollary A.2. \( \square \)

**Theorem A1** (Karamata’s theorem Proposition 1.5.8 of [5]) Let \( \ell \) be a continuous slowly varying function and locally bounded in \([x_0, \infty)\) for some \( x_0 \geq 0 \). Then

(a) for \( \alpha > -1 \)

\[
\int_{x_0}^{x} t^\alpha \ell(t) dt \sim \frac{x^{1+\alpha} \ell(x)}{1 + \alpha}.
\]

(b) for \( \alpha < -1 \)

\[
\int_{x}^{\infty} t^\alpha \ell(t) dt \sim \frac{x^{1+\alpha} \ell(x)}{1 + \alpha}.
\]
We finish this section with the proof of an earlier lemma.

**Proof of Lemma 2** (i) By Potter’s theorem (Theorem 1.5 of [5]), if \( \ell \) is slowly varying, then for every \( \delta > 0 \) there exists \( M > 0 \) such that

\[
\frac{\ell(x)}{\ell(y)} \leq 2 \max \left\{ \frac{x^\delta}{y^\delta}, \frac{y^\delta}{x^\delta} \right\}
\]  

(A.8)

for every \( x, y > M \). We have

\[
\int_0^1 \left| \frac{\ell(ut)}{\ell(t)} - 1 \right| u^{-\gamma} du = \int_0^M \left| \frac{\ell(ut)}{\ell(t)} - 1 \right| u^{-\gamma} du + \int_M^1 \left| \frac{\ell(ut)}{\ell(t)} - 1 \right| u^{-\gamma} du.
\]

We then obtain

\[
\int_0^M \left| \frac{\ell(ut)}{\ell(t)} - 1 \right| u^{-\gamma} du \leq \left( \sup_{y \in [0,M]} \frac{\ell(y)}{\ell(t)} - 1 \right) \frac{M^{1-\gamma}}{t^{1-\gamma} (1-\gamma)} \xrightarrow{t \to \infty} 0,
\]

by Corollary A.3. Choosing \( \delta < 1 - \gamma \) in (A.8), we see that

\[
\int_0^1 \left| \frac{\ell(ut)}{\ell(t)} - 1 \right| u^{-\gamma} 1\{u \geq M/t\} du \leq \int_0^1 (2 \max\{u^{-\delta}, u^\delta\} - 1) u^{-\gamma} du < \infty,
\]

and therefore the LHS of the above equation tends to 0 by the dominated convergence theorem. This and another elementary application of Corollary A.3 finish the proof of item (i).

(ii) We have

\[
\left| \sum_{k=1}^t \ell(k)k^{-\gamma} - \frac{t^{1-\gamma}\ell(t)}{1-\gamma} \right| \leq \left| \sum_{k=1}^t \ell(k)k^{-\gamma} - \int_0^t \ell(s)s^{-\gamma} ds \right| + \left| \int_0^t \ell(s)s^{-\gamma} ds - \ell(t) \cdot \int_0^t s^{-\gamma} ds \right| \\
\leq C + \left| \int_0^t s^{-\gamma} (\ell(s) - \ell(t)) ds \right|,
\]

(A.9)

since \( \ell(s)s^{-\gamma} \) is eventually monotone decreasing. Dividing both sides by \( t^{1-\gamma}\ell(t) \) and making the substitution \( u = st^{-1} \) in the integral gives the result. \( \square \)

**Appendix B. Martingales Concentration Inequalities**

For the sake of completeness, we state here two useful concentration inequalities for martingales which are used throughout the paper.
Theorem B1 (Azuma–Höffeding Inequality [7]) Let \((M_n, \mathcal{F})_{n \geq 1}\) be a (super)martingale satisfying
\[
|M_{i+1} - M_i| \leq a_i
\]
Then, for all \(\lambda > 0\) we have
\[
\mathbb{P}(M_n - M_0 > \lambda) \leq \exp\left(-\frac{\lambda^2}{\sum_{i=1}^{n} a_i^2}\right).
\]

Theorem B2 (Freedman’s inequality [12]) Let \((M_n, \mathcal{F}_n)_{n \geq 1}\) be a (super)martingale. Write
\[
W_n := \sum_{k=1}^{n-1} \mathbb{E}\left[\left((M_{k+1} - M_k)^2\right) \mid \mathcal{F}_k\right] \tag{B.1}
\]
and suppose that \(M_0 = 0\) and
\[
|M_{k+1} - M_k| \leq R, \quad \text{for all } k.
\]
Then, for all \(\lambda > 0\) we have
\[
\mathbb{P}\left(M_n \geq \lambda, W_n \leq \sigma^2, \text{ for some } n\right) \leq \exp\left(-\frac{\lambda^2}{2\sigma^2 + 2R\lambda/3}\right).
\]

References
1. Alves, C., Ribeiro, R., Sanchis, R.: Large communities in a scale-free network. J. Stat. Phys. 166(1), 137–149 (2017)
2. Alves, C., Ribeiro, R., Sanchis, R.: Agglomeration in a preferential attachment random graph with edge-steps. (2019). Preprint. arxiv:1901.02486
3. Alves, C., Ribeiro, R., Sanchis, R.: Topological properties of p.a. random graphs with edge-step functions (2019). Preprint. arxiv:1902.10165
4. Barabási, A.-L., Albert, R.: Emergence of scaling in random networks. Science (1999)
5. Bingham, N.H., Goldie, C.M., Teugels, J.L.: Regular Variation. Encyclopedia of Mathematics and its Applications, vol. 27. CUP, Cambridge (1989)
6. Bollobás, B., Riordan, O., Spencer, J., Tusnády, G.: The degree sequence of a scale-free random graph process. Random Struct. Algorithms 18, 279–290 (2001)
7. Chung, F., Lu, L.: Complex Graphs and Networks (Cbms Regional Conference Series in Mathematics). American Mathematical Society, Boston (2006)
8. Cooper, C., Frieze, A.: A general model of web graphs. Random Struct. Algorithms 22(3), 311–335 (2003)
9. Crane, Harry, Dempsey, Walter: Edge exchangeable models for interaction networks. J. Am. Stat. Assoc. 113(523), 1311–1326 (2018)
10. Deijfen, M., van den Esker, H., van der Hofstad, R., Hooghiemstra, G.: A preferential attachment model with random initial degrees. Arkiv för Matematik 47(1), 41–72 (2009)
11. Dereich, S., Ortgiese, M.: Robust analysis of preferential attachment models with fitness. Comb. Probab. Comput. 23(3), 386–411 (2014)
12. Freedman, D.A.: On tail probabilities for martingales. Ann. Probab. 3(1), 100–118 (1975)
13. Jacob, E., Mörters, P.: Spatial preferential attachment networks: power laws and clustering coefficients. Ann. Appl. Probab. 25(2), 632–662, 04 (2015)
14. Kim, B., Holme, P.: Growing scale-free networks with tunable clustering. Phys. Rev. E (2002)
15. Malyshevkin, Y., Paquette, E.: The power of choice combined with preferential attachment. Electron. Commun. Probab. 19(44), 1–13 (2014)
16. Pew Research Center. Social media update 2014 (2014)
17. Strogatz, S.H., Watts, D.J.: Tcollective dynamics of `small-world’ networks. Nature (1998)
18. Thörnblad, E.: Asymptotic degree distribution of a duplication-deletion random graph model. Int. Math. 11, 03 (2014)
19. Van Der Hofstad, R.: Random graphs and complex networks. Available on http://www.win.tue.nl/rhofstad/NotesRGCN.pdf (2009)
20. Wang, W.-Q., Zhang, Q.-M., Zhou, T.: Evaluating network models: a likelihood analysis. EPL (Europhysics Letters) 98(2), 28004 (2012)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.