

# DIAMETER OF P.A. RANDOM GRAPHS WITH EDGE-STEP FUNCTIONS

CAIO ALVES<sup>1</sup>, RODRIGO RIBEIRO<sup>2</sup>, AND RÉMY SANCHIS<sup>3</sup>

**ABSTRACT.** In this work we prove general bounds for the diameter of random graphs generated by a preferential attachment model whose parameter is a function  $f : \mathbb{N} \rightarrow [0, 1]$  that drives the asymptotic proportion between the numbers of vertices and edges. These results are sharp when  $f$  is a *regularly varying function at infinity* with strictly negative index of regular variation  $-\gamma$ . For this particular class, we prove a characterization for the diameter that depends only on  $-\gamma$ . More specifically, we prove that the diameter of such graphs is of order  $1/\gamma$  with high probability, although its vertex set order goes to infinity polynomially. Sharp results for the diameter for a wide class of *slowly varying functions* are also obtained.

**Keywords:** complex networks; cliques; preferential attachment; concentration bounds; diameter; scale-free; small-world

MSC 2010 subject classifications. Primary 05C82; Secondary 60K40, 68R10

## 1. INTRODUCTION

P. Erdős and A. Rényi in their seminal paper [15] introduced the random graph model that now carries their name in order to solve combinatorial problems. However, the theory of Random Graphs as a whole has proven to be a useful tool for treating concrete problems as well. Any discrete set of entities whose elements interact in a pairwise fashion may be seen as a graph: the vertices represent the entities, and the edges, the possible interactions. This approach is nowadays intuitive and very fruitful. In the scenario where there exists some randomness on the interactions among the entities, random graphs became the natural tool to represent abstract or real phenomena.

From a mathematical/statistical point of view, the Erdős-Rényi model – and many others related to it – is *homogeneous*, in the sense that its vertices are *statistically indistinguishable*. However, the empirical findings of the seminal work of A. Bárábasi and R. Albert [6] suggested that many real-world networks are *non-homogeneous*. They observed that such graphs were *scale-free*, i.e., their degree sequence had a power-law distribution. The authors

---

*Date:* July 4, 2023

<sup>1</sup> Institute of Mathematics, University of Leipzig

<sup>2</sup> PUC Chile, Pontificia Universidad Católica de Chile.

<sup>3</sup> Departamento de Matemática, Universidade Federal de Minas Gerais.

proposed a mechanism – known as *preferential attachment* – that could explain the emergence of such highly skewed distributions. Roughly speaking, the idea is that some sort of popularity drives the interaction among the entities.

Motivated mainly by these empirical findings, nowadays Preferential Attachment models (PA-models for short) constitute a well known class of random graph models investigated from both theoretical and applied perspectives. Recently, the preferential attachment mechanism has been generalized in many ways and combined with other rules of attachment, such as *spatial proximity* [17] and *fitness of vertices* [12]. It also arises naturally even in models where it is not entirely explicit such as the deletion-duplication models [5, 22], in which vertices' degree still evolve according to the PA-rule. Furthermore, the PA-models provide an interesting and natural environment for other random processes, such as *bootstrap percolation*, *contact process* and *random walks*, see [4, 9, 18] for recent examples of random processes whose random media is sampled from some PA-model.

When dealing with PA-models, there exists a set of natural questions that arises, such as the degree distribution and the order of the diameter. Their interest relies on modeling purposes and on the implications for the graph's combinatorial structures.

In this paper we address the latter topic on a PA-model which is a modification of the Barabási-Albert model introduced in [1] (BA-model for short). The kind of result we pursue is to show that some graph properties hold *asymptotically almost surely* (a.a.s). Given a sequence of random graphs  $\{G_t\}_{t \in \mathbb{N}}$ , we say that a graph property  $\mathcal{P}$  holds a.a.s, if

$$\mathbb{P}(G_t \in \mathcal{P}) = 1 - o(1)$$

i.e., the probability of observing such property increases to 1 as  $t$  goes to infinity. For instance,  $\mathcal{P}$  may be the set of graphs having diameter less than the logarithm of the total number of vertices.

In order to offer a clearer discussion of our results, we introduce the model in the next subsection, then we discuss separately the properties which we want the graph to satisfy a.a.s, as well as the associated motivation.

**1.1. Preferential attachment model with an edge-step function.** The model we investigate here in its generality was proposed in [3] and combines the traditional preferential attachment rule with a function called *edge-step function* that drives the growth rate of the vertex set.

The model has one parameter  $f$  which is a real non-negative function with domain given by  $\mathbb{N}$  and bounded by one on the  $L_\infty$ -norm, we will see  $f$  as a sequence of probabilities indexed by a time parameter  $t \in \mathbb{N}$ . Without loss of generality and to simplify the expressions we deal with, we start the process from an initial graph  $G_1$  consisting in one vertex and one loop. The model evolves inductively and at each step the next graph is obtained by performing one of the two stochastic operations defined below on the previous one:

- *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

$$P(v \rightarrow u|G) = \frac{\text{degree}(u)}{\sum_{w \in G} \text{degree}(w)}.$$

- *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.

The model alternates between the two types of operations according to a sequence  $\{Z_t\}_{t \geq 1}$  of independent random variables such that  $Z_t \stackrel{d}{=} \text{Ber}(f(t))$ . 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  $Z_t = 0$ . Notice that  $f(t)$  is the probability of adding a new vertex at time  $t$ , thus the reader may think of  $f$  as a *vertex-step probability*.

Given  $f$ , its partial sum is an important quantity for us and we reserve the letter  $F$  to denote it, i.e.,  $F$  is a function defined as

$$(1.1) \quad F(t) := 1 + \sum_{s=2}^t f(s).$$

Observe that the edge-step function  $f$  is intimately related to the growth of the vertex set. If we let  $V_t$  denote the number of vertices added up to time  $t$ , then, if  $F(t) \xrightarrow{t \rightarrow \infty} \infty$ , then a concentration of measures result will imply

$$(1.2) \quad V_t = 1 + \sum_{s=2}^t Z_s \approx F(t),$$

since  $\{Z_s\}_{s \geq 1}$  is a sequence of independent random variables. Thus, abusing the notation for a brief moment, we may write

$$\frac{dV_t}{dt} = f(t).$$

When the proper machinery has been settled, we will discuss in Section 6 that some regularity should be imposed on  $f$  in order to avoid some pathological behaviors. For now, we define a list of conditions we may impose on  $f$  at different points of the paper in order to get the proper results. For instance, we say  $f$  satisfies condition (D) if it is *non-increasing*. We define the further conditions:

$$(D_0) \quad f \text{ is non-increasing and } \lim_{t \rightarrow \infty} f(t) = 0;$$

$$(S) \quad \sum_{s=1}^{\infty} \frac{f(s)}{s} < \infty;$$

$$(L_\kappa) \quad \sum_{s=t^{1/13}}^t \frac{f(s)}{s} < (\log t)^\kappa, \text{ for all } t \in \mathbb{N} \text{ and some } \kappa \in (0, 1);$$

$$(RV_\gamma) \quad \exists \gamma, \text{ such that } \forall a > 0, \lim_{t \rightarrow \infty} \frac{f(at)}{f(t)} = \frac{1}{a^\gamma}.$$

We must point out that for modeling purposes, conditions (D) and (D<sub>0</sub>) may be desirable. For instance, in the context of social networks, these conditions assure that the rate at which new individuals join the network is decreasing as the size of the network increases. Whereas, conditions (S) and (L<sub>κ</sub>) are related to the order of the maximum degree of  $G_t(f)$ . In [3], the authors point out that the maximum degree at time  $t$  should be of order

$$(1.3) \quad t \cdot \exp \left\{ -\frac{1}{2} \sum_{s=2}^t \frac{f(s)}{s-1} \right\}.$$

A function satisfying condition (RV<sub>γ</sub>) is called *regularly varying* at infinity and the exponent  $\gamma$  is called the *index of regular variation*. Functions in this class are well-studied in mathematics in many contexts and a variety of asymptotic results for them and their integrals is known due mainly to the theory developed by Karamata, see [7] for a complete reference.

In general, we may say that this paper investigates how sensitive is the diameter to changes of  $f$  and aims at a general characterization of such observable for a class of functions  $f$  that is as wide as possible.

**1.2. Shaping the diameter.** An important property of graphs which is also related to spread of rumors and connectivity of networks is the diameter, that is, the maximal graph distance between two vertices of said graph. Originally, investigating the diameter of real-world networks, the authors in [21] observed that, although coming from different contexts, those networks usually have diameter of order less than the logarithm of the number of vertices, the so-called *small-world* phenomena.

In this paper we address the issue of determining the order of the diameter of  $G_t(f)$ . Our main goals in this subject are to obtain a characterization for the diameter imposing conditions on  $f$  as weak as possible and also to obtain regimes for the diameter arbitrarily small but still preserving the scalefreeness of the graph.

In order to slow the growth of the diameter of PA-models, two observable play important roles: the maximum degree and the proportion of vertices with low degree. The former tends to concentrate connections on vertices with very high degree which acts in the way of shortening the diameter, since they attract connections to them; whereas the latter acts in the opposite way. In [24] and [16], the authors have shown that in the configuration model with power-law distribution the diameter order is extremely sensitive to the proportion of vertices with degree 1 and 2.

One way to reduce the effect of low degree vertices on the diameter is via *affine* preferential attachment rules, i.e., introducing a parameter  $\delta$  and choosing vertices with probability proportional to their degree plus  $\delta$ . In symbols, conditionally on  $G_t$ , we connect a new vertex  $v_{t+1}$  to an existing one  $u$  with probability

$$\mathbb{P}(v_{t+1} \rightarrow u | G_t) = \frac{\text{degree}(u) + \delta}{\sum_{w \in G_t} (\text{degree}(w) + \delta)}.$$

By taking a negative  $\delta$ , the above rule increases the influence of high degree vertices and indeed decreases drastically the diameter's order. For instance, for positive  $\delta$  the diameter of  $G_t$  is at least  $\log(t)$ , whereas for  $\delta < 0$  the diameter of  $G_t$  is at most  $\log \log t$ . See [13] for several results on the diameter of different combinations for the affine preferential attachment rule.

Diminishing the effect of low degree vertices is not enough to break the growth of the diameter completely. The reason for that is, despite their low degree, these vertices exist in large amount. Even the existence of a vertex with degree close to  $t$  at time  $t$  may not be enough to freeze the diameter's growth. In [20] the authors have proven that the maximum degree of a modification of the BA-model is of order  $t$  at time  $t$ . However, the authors believe that this is not enough to obtain a diameter of order  $\log \log t$ , the reason being that this large hub still has to compete with a large number of low degree vertices.

**1.2.1. General bounds for the diameter.** As said before, our goal is to develop bounds for the diameter of  $G_t(f)$  with  $f$  as general as possible. Under the condition of monotonicity, we prove the following lower bound

**Theorem 1** (Lower bound on the diameter). *Let  $f$  be an edge-step function satisfying condition (D). Then*

$$(1.4) \quad \mathbb{P} \left( \text{diam}(G_t(f)) \geq \frac{1}{3} \left( \frac{\log t}{\log \log t} \wedge \frac{\log t}{-\log f(t)} \right) \right) = 1 - o(1).$$

Requiring more information on  $f$ , we prove upper bounds that, for a broad class of functions, are of the same order of the lower bounds given by the previous theorem. This is all summarized in the Theorem below.

**Theorem 2** (Upper bound on the diameter). *Let  $f$  be an edge-step function. Then*

(a) *if  $f$  also satisfies conditions (S) and (D<sub>0</sub>) then*

$$\mathbb{P} \left( \text{diam}(G_t(f)) \leq 2 + 6 \left( \frac{\log t}{-\log \left( \sum_{s=t^{\frac{1}{13}}}^t \frac{f(s)}{s-1} \right)} \wedge \frac{\log t}{\log \log t} \right) \right) = 1 - o(1);$$

(b) *if  $f$  satisfies condition (L <sub>$\kappa$</sub> ) then*

$$\mathbb{P} \left( \text{diam}(G_t(f)) \leq 2 + \frac{6}{1 - \kappa} \frac{\log t}{\log \log t} \right) = 1 - o(1).$$

1.2.2. *The class regularly of varying functions.* In [3], the authors prove a characterization of the empirical degree distribution of graphs generated by  $f$  satisfying condition  $(RV_\gamma)$ , for  $\gamma \in [0, 1)$ . More specifically, they prove that the degree distribution of such graphs obeys a power law distribution whose exponent depends only on the index of regular variation  $-\gamma$ .

A byproduct of our general bounds is a similar characterization for the diameter. For edge-step functions satisfying conditions  $(D_0)$  and  $(RV_\gamma)$  for  $\gamma \in (0, \infty)$  the graphs generated by such functions have constant diameter and its order depends only on the index of regular variation  $-\gamma$ . We state this result in the theorem below

**Theorem 3** (Diameter of regularly varying functions). *Let  $f$  be an edge-step function satisfying conditions  $(D_0)$  and  $(RV_\gamma)$ , for  $\gamma \in (0, \infty)$ . Then,*

$$\mathbb{P} \left( \frac{1}{4^\gamma} \leq \text{diam}(G_t(f)) \leq \frac{100}{\gamma} + 2 \right) = 1 - o(1).$$

1.2.3. *The class of slowly varying functions.* The case when  $\gamma = 0$  is richer in terms of possible orders of the diameter and does not admit a nice characterization as the one we obtain for positive  $\gamma$ . In this settings, we present another consequence of our bounds for particular subclasses of the class of slowly varying functions. Let us first define the subclass of functions and later state how our results fit these specific classes.

$$(1.5) \quad L := \left\{ f \text{ is an edge-step function such that } f(t) = \frac{1}{\log^\alpha(t)}, \text{ for some } \alpha > 0 \right\};$$

$$(1.6) \quad E := \left\{ f \text{ is an edge-step function such that } f(t) = e^{-\log^\alpha(t)}, \text{ for some } \alpha \in (0, 1) \right\}.$$

It is straightforward to verify that functions belonging to the set above defined are slowly varying. For functions belonging to the two subclasses  $L$  and  $E$ , our results have the following consequences, verifiable through elementary calculus,

**Corollary 1.** *Let  $f$  be an edge-step function.*

(a) *if  $f$  belongs to  $L$ , with  $\alpha \leq 1$ , then*

$$\mathbb{P} \left( \frac{1}{3} \frac{\log t}{\log \log t} \leq \text{diam}(G_t(f)) \leq \frac{8}{\alpha} \frac{\log t}{\log \log t} \right) = 1 - o(1);$$

(b) *if  $f$  belongs to  $L$ , with  $\alpha > 1$ , then*

$$\mathbb{P} \left( \frac{1}{3\alpha} \frac{\log t}{\log \log t} \leq \text{diam}(G_t(f)) \leq \frac{7}{\alpha - 1} \frac{\log t}{\log \log t} \right) = 1 - o(1);$$

(c) *if  $f$  belongs to  $E$ , then*

$$\mathbb{P} \left( C_\alpha^{-1} (\log t)^{1-\alpha} \leq \text{diam}(G_t(f)) \leq C_\alpha (\log t)^{1-\alpha} \right) = 1 - o(1),$$

*for some  $C_\alpha \geq 1$ .*

**1.3. Main technical ideas.** In order to prove the existence of some given subgraph in the (affine) BA-random graphs a key ingredient is usually to use the fact that two given vertices  $v_i$  and  $v_j$  – born at times  $i, j \in \mathbb{N}$  respectively – may be connected only at one specific time-step, since (assuming  $i < j$ ) the model’s dynamic only allows  $v_j$  to connect to  $v_i$  at the moment in which  $v_j$  is created. This property facilitates the computation of the probability of the occurrence of a given subgraph and decreases the combinatorial complexity of the arguments. In [8, 14] the authors estimate the number of triangles and cherries (paths of length 3) of the (affine) BA-model and their argument relies heavily on this feature of the model. In our case, however, the *edge-step* prevents an application of such arguments, since a specific subgraph may appear at any time after the vertices have been added.

Another difficulty in our setup is the degree of generality we work with. Our case replaces the parameter  $p \in (0, 1]$  in the models investigated in [2, 10, 11] by any non-negative real function  $f$  with  $\|f\|_\infty \leq 1$ . The introduction of such function naturally increases the complexity of any analytical argument one may expect to rely on, as will be clear during our estimates for the vertices’ degree, and makes it harder to discover *threshold* phenomena. This is the reason why in our work the Karamata’s Theory of *regularly varying functions* is crucial in order to prove sharper results.

In order to prove Theorem 1, we apply the second moment method on the number of *isolated paths*. This approach demands correlation estimations for the existence of two such paths in  $G_t(f)$ , which we do under the assumption of  $f$  being monotonic only.

For the proof of Theorem 2, a key step is to obtain a lower bound for the maximum degree, since high degree vertices tend to attract vertices to themselves. However, as said above, the degree of generality of  $f$  makes all analytical arguments more involved. Then, in order to overcome part of the heavy computation we have to deal with in our setup, in Lemma 1 we obtain lower bounds for the maximum degree by constructing a monotonic coupling with the case in which  $f \equiv p \in (0, 1)$ . This coupling allow us to transpose known results about the maximum degree of  $G_t(p)$  to the general case of  $G_t(f)$ .

We then combine Lemma 1 with a lower bound for the degree of earlier vertices, which is obtained by estimation of negative moments of a given vertex’s degree, in order to show that, under conditions (S) or  $(L_\kappa)$ , long paths of younger vertices are unlikely and older vertices are all very close in graph distance. Finally, using results from the Karamata’s theory of regularly varying functions, we verify that this broad class of functions satisfies our assumptions, proving Theorem 3.

**1.4. Organization.** In Section 2, we prove technical estimates for the degree of a given vertex, which are needed for the upper bound on the diameter. Section 3 is devoted to the general lower bound for the diameter, i.e., for the proof of Theorem 1. We prove the upper bound for the diameter, Theorem 2, in Section 4. Finally, in Section 5 we show what our results say for the class of regularly varying functions. We end the paper at Section 6



with some comments on the affine version of our model and a brief discussion on what may happen to the model if some regularity conditions are dropped.

**1.5. Notation.** We let  $V(G_t(f))$  and  $E(G_t(f))$  denote the set of *vertices* and *edges* of  $G_t(f)$ , respectively. Given a vertex  $v \in V(G_t(f))$ , 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,

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

When necessary in the context, we may use  $D_G(v)$  to denote the degree of  $v$  in the graph  $G$ .

Given two sets  $A, B \subseteq V(G_t(f))$ , we let  $\{A \leftrightarrow B\}$  denote the event where there exists an edge connecting a vertex from  $A$  to a vertex from  $B$ . We denote the complement of this event by  $\{A \nleftrightarrow B\}$ . We let  $\text{dist}(A, B)$  denote the graph distance between  $A$  and  $B$ , i.e. the minimum number of edges that a path that connects  $A$  to  $B$  must have. When one of these subsets consists of a single vertex, i.e.  $A = \{v\}$ , we drop the brackets from the definition and use  $\{v \leftrightarrow B\}$  and  $\text{dist}(v, B)$ , respectively.

Regarding constants, we let  $C, C_1, C_2, \dots$  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  $t$ . In this way we obtain the natural filtration  $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \dots$  associated to the process.

## 2. TECHNICAL ESTIMATES FOR THE DEGREE

In this section we develop technical estimates related to the degree of a given vertex. We begin by stating one of the most fundamental identities in the study of preferential attachment models: the conditional distribution of the increment of the degree of a given vertex. Given  $v \in V(G_t(f))$ , we have

$$\begin{aligned} \mathbb{P}(\Delta D_t(v) = 0 | \mathcal{F}_t) &= f(t+1) \left(1 - \frac{D_t(v)}{2t}\right) + (1 - f(t+1)) \left(1 - \frac{D_t(v)}{2t}\right)^2 \\ (2.1) \quad \mathbb{P}(\Delta D_t(v) = 1 | \mathcal{F}_t) &= f(t+1) \frac{D_t(v)}{2t} + 2(1 - f(t+1)) \frac{D_t(v)}{2t} \left(1 - \frac{D_t(v)}{2t}\right), \\ \mathbb{P}(\Delta D_t(v) = 2 | \mathcal{F}_t) &= (1 - f(t+1)) \frac{D_t(v)^2}{4t^2}. \end{aligned}$$

To see why the above identities hold true, observe for example that in order for  $\Delta D_t(v) = 0$ , either a vertex step was taken, and the vertex did not connect to  $v$ , or an edge step was taken and neither of the endpoints of the new edge connected to  $v$ . The other equations



follow from analogous reasonings. As a direct consequence, we obtain

$$\begin{aligned}
 \mathbb{E} [\Delta D_t(v) | \mathcal{F}_t] &= 1 \cdot f(t+1) \cdot \frac{D_t(v)}{2t} + 1 \cdot 2(1-f(t+1)) \frac{D_t(v)}{2t} \left(1 - \frac{D_t(v)}{2t}\right) \\
 (2.2) \quad &+ 2 \cdot (1-f(t+1)) \frac{D_t^2(v)}{4t^2} \\
 &= \left(1 - \frac{f(t+1)}{2}\right) \frac{D_t(v)}{t}.
 \end{aligned}$$

Using the above equation repeatedly one obtains, conditioned on the event where the vertex  $v$  is born at time  $t_0$ ,

$$\begin{aligned}
 \mathbb{E} [D_t(v)] &= \mathbb{E} [\mathbb{E} [D_t(v) | \mathcal{F}_{t-1}]] \\
 (2.3) \quad &= \left(1 + \frac{1}{t-1} - \frac{f(t)}{2(t-1)}\right) \mathbb{E} [D_{t-1}(v)] \\
 &= \prod_{s=t_0}^{t-1} \left(1 + \frac{1}{s} - \frac{f(s+1)}{2s}\right).
 \end{aligned}$$

The next lemma gives a lower bound for the maximum degree of  $G_t(f)$  when  $f$  decreases to zero. This result will be crucial in the proof of Theorem 2, which gives general upper bounds for the diameter. Its proof involves the construction of a coupling between  $\{G_t(f)\}_{t \geq 1}$  and  $\{G_t(p)\}_{t \geq 1}$ , where  $G_t(p)$  is constructed from the constant function equal to  $p \in (0, 1]$  for every  $t \in \mathbb{N}$ . This coupling will be monotonic in the sense that, if  $G_t(p)$  has a vertex with high degree, then so has  $G_t(f)$ .

**Lemma 1.** *Let  $f$  be an edge-step function such that  $f(t) \searrow 0$  as  $t$  goes to infinity. Then, for any fixed  $\varepsilon > 0$ ,*

$$\mathbb{P} (D_{\max}(G_t(f)) < t^{1-\varepsilon}) \xrightarrow{t \rightarrow \infty} 0.$$

Let us describe the general idea behind the proof of the above result. Exclusively in the proof of this Lemma we will denote as  $i \in \mathbb{N}$  the  $i$ -th vertex to be added to the graph. We couple the degree sequence of the first  $m$  vertices added by the process  $\{G_s(f)\}_{s \geq 1}$ , for some large constant  $m$ , to the degree sequence of vertices  $\{t^\varepsilon + 1, \dots, t^\varepsilon + m\}$  in  $\{G_s(p)\}_{s \geq 1}$  for large  $t$  and small  $\varepsilon$ . The coupling is constructed in such way that the  $i$ -th vertex of  $\{G_s(f)\}_{s \geq 1}$  always has larger degree than the  $(t^\varepsilon + i)$ -th vertex in  $\{G_s(p)\}_{s \geq 1}$ . This is possible for two reasons, the first being due to the fact that, since vertices are chosen with probability proportional to their degrees and both graphs have the same number of edges, if vertex  $i$  in  $G_t(f)$  has larger degree than vertex  $t^\varepsilon + i$  in  $G_t(p)$ , then  $i$  is more likely to increase its degree in the next step. Moreover, when using this fact to show stochastic domination by induction, the first step follows because, if at the moment in which process  $\{G_s(p)\}_{s \geq 1}$  adds its  $(t^\varepsilon + i)$ -th vertex process  $\{G_s(f)\}_{s \geq 1}$  has already added vertex  $i$ , then  $i$  has degree at least one. The second reason is the fact that  $f$  decreases to zero. This property guarantees that  $f(t)$  is always smaller than  $p$  after some time, which allows us to couple both process in a way that process

$\{G_s(p)\}_{s \geq 1}$  takes less edge-steps than  $\{G_s(f)\}_{s \geq 1}$ , and therefore vertices in  $\{G_s(p)\}_{s \geq 1}$  have less chances to increase their degrees. We finally combine the coupling with previous results about  $G_t(p)$  in [2], which imply that, *w.h.p.*, some vertex  $j \in G_t(p)$  added at some time between  $t^\varepsilon + 1$  and  $t^\varepsilon + m$  has degree at least  $t^{(1-\varepsilon)(1-p/2)}$ .

Although the coupling above described might be believable, its formalization may be a little involved since we have to deal with times in which the two graph processes perform different graph operations (vertex or edge-step), as well as the case when  $\{G_s(f)\}_{s \geq 1}$  has less than  $m$  vertices at the time when  $\{G_s(p)\}_{s \geq 1}$  adds its  $t^\varepsilon$ -th vertex.

*Proof.* The case in which  $f$  decreases fast enough to zero so that  $\sum_{s=1}^{\infty} f(s) < \infty$  is simpler. By the hypothesis on  $f$  and the fact that in this case  $|V_t|$  is the sum of independent Bernoulli random variables, it follows that  $\{\exp\{|V_t|\}\}_{t \geq 1}$  is limited in  $L_1$ . Thus, by Markov's inequality

$$\mathbb{P}(|V_t| > 2 \log t) \leq \frac{\sup_t \mathbb{E} \exp\{|V_t|\}}{t^2}.$$

And since the sum of all degrees at time  $t$  is  $2t$ , we have *w.h.p.* that  $G_t(f)$  has at most  $2 \log t$  vertices, and by the pigeonhole principle it follows in this case that there exists a vertex of degree at least  $t/\log t$ .

We now consider the case in which the average number of vertices goes to infinity. We begin by applying Theorem 2 of [2], choosing  $m, R$  large enough,  $p$  small enough, and letting  $j = t^\varepsilon$  in order to obtain, for large enough  $t$ , the bound

$$(2.4) \quad \mathbb{P} \left( \sum_{i=1}^m D_{G_t(p)}(j+i) < t^{1-\varepsilon} \right) \leq t^{-C},$$

for some  $C$  depending on  $\varepsilon, m, p$  and  $R$ . That is, with probability at least  $1 - t^{-C}$ , for small  $p$  and large  $m$ , at least one of the vertices  $\{j+1, \dots, j+m\}$  has degree at least  $t^{1-\varepsilon}/m$  in  $G_t(p)$ .

Let  $t_0 := \min\{t \geq 0 ; f(t) \leq p\}$ , which is well-defined since  $f$  decreases to zero, and consider  $t$  in the definition of  $j$  be larger than  $t_0$ . We will couple  $\{G_t(f)\}_{t \in \mathbb{N}}$  and  $\{G_t(p)\}_{t \in \mathbb{N}}$  for all  $i \in \{1, \dots, m\}$  in such way that

$$(2.5) \quad D_{G_s(p)}(j+i) \leq D_{G_s(f)}(i), \forall s \geq 0, \text{ w.h.p.}$$

In order to formalize the coupling, we assume that  $|V(G_{t^\varepsilon}(f))| \geq m$  and that our probability space is large enough so we have at our disposal two independent sequences of r.v.s  $\{(U_s^{(1)}, U_s^{(2)})\}_{s \in \mathbb{N}}$  and  $\{U'_s\}_{s \in \mathbb{N}}$  such that  $U^{(1)} \stackrel{d}{=} U_s^{(2)} \stackrel{d}{=} U' \stackrel{d}{=} \text{Uni}[0, 1]$ , and that all these variables are mutually independent. We will use the sequence  $\{U'_s\}_{s \in \mathbb{N}}$  to control simultaneously the sequence of edge and vertex-steps we take in each graph process. The sequence  $\{(U_s^{(1)}, U_s^{(2)})\}_{s \in \mathbb{N}}$  on the other hand will be used to select vertices in both graphs.

We proceed by induction. Define the degree of a yet-unborn vertex as 0. Then in the event where  $|V(G_{t^\varepsilon}(f))| \geq m$ , (2.5) holds up to time  $s = t^\varepsilon$ . Suppose we have succeeded in coupling  $\{G_r(f)\}_{r \leq s}$  and  $\{G_r(p)\}_{r \leq s}$  in such way that (2.5) is satisfied for  $s \geq t^\varepsilon$ . Now we have in the

same probability space both graphs  $G_s(f)$  and  $G_s(p)$  with the property that (2.5) holds. To advance the induction, we need to know how to generate  $G_{s+1}(f)$  and  $G_{s+1}(p)$ . To do so, let us first introduce some notation. For  $i \in \{1, \dots, m\}$ , define inductively the random intervals

$$I_s^{(1)} := \left[0, \frac{D_{G_s(f)}(1)}{2s}\right] =: [b_0(s), b_1(s)], \quad I_s^{(i)} \equiv [b_{i-1}(s), b_i(s)] := \left[b_{i-1}(s), b_{i-1}(s) + \frac{D_{G_s(f)}(i)}{2s}\right],$$

and for each  $I_s^{(i)}$  we let  $I_{s,p}^{(j+i)}$  be

$$(2.6) \quad I_{s,p}^{(j+i)} := \left[b_{i-1}(s), b_{i-1}(s) + \frac{D_{G_s(p)}(j+i)}{2s}\right].$$

Notice that, since (2.5) holds,  $I_{s,p}^{(j+i)} \subset I_s^{(i)}$ , and that, since the total degree of  $G_t(f)$  is always  $2s$ , all the above intervals are contained in  $[0, 1]$ .

Now we need to know which kind of step we take on each graph and we do this according to  $U'_{s+1}$  (recall that this variable is independent of the whole past).

$U'_{s+1} \leq f(s+1)$ . In this case we perform a vertex-step in *both graphs*. We use  $U_{s+1}^{(1)}$  in order to select two vertices (one in each graph) to connect the new vertex that was born in each graph. We increase the degree of vertex  $i \in G_s(f)$  (resp.  $j+i \in G_s(p)$ ) if  $U_{s+1}^{(1)} \in I_s^{(i)}$  (resp.  $U_{s+1}^{(1)} \in I_{s,p}^{(j+i)}$ ). If  $U_{s+1}^{(1)}$  does not belong to any interval  $I_s^{(i)}$  for  $i = 1, \dots, m$  (resp. any interval  $I_{s,p}^{(j+i)}$ ), we let  $G_{s+1}(f)$  (resp.  $G_{s+1}(p)$ ) be constructed in some arbitrary manner that preserves its law. By the construction of the intervals it is clear that (2.5) is preserved in time  $s+1$ .

$f(s+1) \leq U'_{s+1} \leq p$ . In this case, we perform an edge-step on  $G_{s+1}(f)$  and a vertex-step on  $G_{s+1}(p)$ . As before, we sample  $U_{s+1}^{(1)}$  and if  $U_{s+1}^{(1)} \in I_{s,p}^{(j+i)}$  we connect the new vertex of  $G_{s+1}(p)$  to  $j+i$ . As for  $G_{s+1}(f)$ , we interpret  $U_{s+1}^{(1)}$  as indicating the endvertex of one of the halves of the new edge. Then we sample  $U_{s+1}^{(2)}$  and perform a similar procedure as before, but only considering the intervals with respect to  $f$ , in order to choose the second half of the new edge added to  $G_{s+1}(f)$ . We again construct  $G_{s+1}(f)$  and  $G_{s+1}(p)$  in some arbitrary manner in the case the respective uniform variables do not fall in any of the associated intervals.

$p \leq U'_{s+1}$ . In this case we perform an edge-step in both graphs. To do so, it is enough to repeat of procedure of selecting vertices using both  $U_{s+1}^{(1)}$  and  $U_{s+1}^{(2)}$ .

Finally, with Markov's inequality we bound the probability that  $|V(G_{t^\varepsilon}(f))| \leq m$ . Recalling that  $|V(G_{t^\varepsilon}(f))|$  is the sum of independent Bernoulli random variables, we obtain

$$(2.7) \quad \mathbb{P}(|V(G_{t^\varepsilon}(f))| \leq m) \leq e^m \mathbb{E}[\exp\{-|V(G_{t^\varepsilon}(f))|\}] \leq \exp\{m - (1 - e^{-1})F(t^\varepsilon)\},$$

which completes the proof since  $\varepsilon$  was chosen arbitrarily.  $\square$

Our objective now is to obtain a polynomial lower bound for the degree of older vertices, which will be important in the proof of the upper bound for the diameter in Theorem 2. We begin with an upper bound for the expectation of the multiplicative inverse of the degree. Recall the definition of the process  $(Z_t)_{t \geq 1}$ , consisting of independent Bernoulli variables that dictate whether a vertex-step or an edge-step is performed at time  $t$ .

**Lemma 2.** *Given any edge-step function  $f$ , consider the process  $\{G_t(f)\}_{t \geq 1}$ . Denote by  $v_i$  the vertex born at time  $i \in \mathbb{N}$ . We have*

$$(2.8) \quad \mathbb{E} [(D_t(v_i))^{-1} Z_i] \leq f(i) \left( \frac{t-1}{i} \right)^{-\frac{1}{6}}.$$

And consequently

$$(2.9) \quad \mathbb{P} \left( D_t(v_i) \leq \left( \frac{t-1}{i} \right)^{\frac{1}{12}} \middle| Z_i = 1 \right) \leq \left( \frac{t-1}{i} \right)^{-\frac{1}{12}}.$$

*Proof.* If  $Z_i = 1$ , then for every  $s \geq i$  we have that

$$\Delta D_s(v_i) \geq 0, \quad D_{s+1}(v_i) \leq D_s(v_i) + 2 \leq 3D_s(v_i),$$

and that  $D_s(v_i)$  is  $\mathcal{F}_s$  measurable. Together with (2.2), these facts imply, on the event where  $Z_i = 1$ ,

$$(2.10) \quad \begin{aligned} \mathbb{E} \left[ \frac{1}{D_{s+1}(v_i)} - \frac{1}{D_s(v_i)} \middle| \mathcal{F}_s \right] &= \mathbb{E} \left[ -\frac{\Delta D_s(v_i)}{D_{s+1}(v_i) D_s(v_i)} \middle| \mathcal{F}_s \right] \\ &\leq -\frac{1}{3(D_s(v_i))^2} \left( 1 - \frac{f(s+1)}{2} \right) \frac{D_s(v_i)}{s} \\ &\leq -\frac{1}{6s} \frac{1}{D_s(v_i)}, \end{aligned}$$

since  $f(k) \leq 1$  for very  $k \in \mathbb{N}$ . Therefore,

$$(2.11) \quad \mathbb{E} [Z_i \cdot (D_{s+1}(v_i))^{-1} | \mathcal{F}_s] \leq Z_i (1 - (6s)^{-1}) (D_s(v_i))^{-1}.$$

Iterating the above argument from  $i$  until  $t$ , we obtain

$$(2.12) \quad \mathbb{E} [Z_i \cdot (D_t(v_i))^{-1}] \leq f(i) \prod_{s=i}^{t-1} \left( 1 - \frac{1}{6s} \right) \leq f(i) \exp \left\{ -\frac{1}{6} \sum_{s=i}^{t-1} \frac{1}{s} \right\} \leq f(i) \left( \frac{t-1}{i} \right)^{-\frac{1}{6}},$$

proving (2.8). We then obtain (2.9) by an elementary application of the Markov inequality:

$$\begin{aligned}
 \mathbb{P} \left( D_t(v_i) \leq \left( \frac{t-1}{i} \right)^{\frac{1}{12}} \middle| Z_i = 1 \right) &= \mathbb{P} \left( (D_t(v_i))^{-1} \geq \left( \frac{t-1}{i} \right)^{-\frac{1}{12}} \middle| Z_i = 1 \right) \\
 (2.13) \qquad \qquad \qquad &\leq \left( \frac{t-1}{i} \right)^{\frac{1}{12}} \mathbb{E} [(D_t(v_i))^{-1} | Z_i = 1] \\
 &\leq \left( \frac{t-1}{i} \right)^{-\frac{1}{12}}.
 \end{aligned}$$

□

We now provide an elementary consequence of the above result, which uses the union bound in order to show that, with high probability, every vertex born before time  $t^{\frac{1}{12}}$  has degree at least  $t^{\frac{1}{15}}$  by time  $t$ .

**Lemma 3.** *Using the same notation as in Lemma 2 we have, for every edge-step function  $f$  and for sufficiently large  $t \in \mathbb{N}$ ,*

$$(2.14) \qquad \mathbb{P} \left( \exists i \in \mathbb{N}, 1 \leq i \leq t^{\frac{1}{12}}, \text{ such that } Z_i = 1 \text{ and } D_t(v_i) \leq t^{\frac{1}{15}} \right) \leq Ct^{-\frac{1}{144}}.$$

*Proof.* By the union bound and equation (2.9), we have that the probability in the left hand side of (2.14) is smaller than or equal to

$$\begin{aligned}
 \sum_{i=1}^{t^{\frac{1}{12}}} \mathbb{P} \left( Z_i = 1, D_t(v_i) \leq t^{\frac{1}{15}} \right) &\leq \sum_{i=1}^{t^{\frac{1}{12}}} f(i) \mathbb{P} \left( D_t(v_i) \leq \left( \frac{t-1}{i} \right)^{\frac{1}{12}} \middle| Z_i = 1 \right) \\
 (2.15) \qquad \qquad \qquad &\leq \sum_{i=1}^{t^{\frac{1}{12}}} \left( \frac{t-1}{i} \right)^{-\frac{1}{12}} \\
 &\leq Ct^{-\frac{1}{12} + \frac{1}{12}(1 - \frac{1}{12})} \\
 &\leq Ct^{-\frac{1}{144}},
 \end{aligned}$$

finishing the proof of the Lemma. □

### 3. GENERAL LOWER BOUND FOR THE DIAMETER: PROOF OF THEOREM 1

The proof of Theorem 1 follows a second moment argument. The idea is to count the number of “long” (the specific size depending on  $f$  and  $t$ ) *isolated paths* in  $G_t(f)$ . We begin by stating two key lemmas for the proof. The former, Lemma 4, states that the expected number of isolated paths goes to infinity with  $t$ . The latter, Lemma 5, guarantees that the presence of a specific isolated path is almost independent from the presence of some other given isolated path whenever said paths are disjoint. This “almost independence” makes the

second moment of the number of such paths very close to the first moment squared, which allows us to apply Paley-Zygmund inequality.

We start by defining precisely what we mean by an isolated path.

**Definition 1** (*t*-isolated path). *Let  $l$  be a positive integer. Let  $\vec{t} = (t_1, \dots, t_l)$  be a vector of distinct positive integers. We say that this vector corresponds to an isolated path  $\{v_{t_1}, \dots, v_{t_l}\}$  in  $G_t(f)$  if and only if:*

- $t_l \leq t$ ;
- $t_i < t_j$  whenever  $1 \leq i < j \leq l$ ;
- during each time  $t_i$ ,  $i = 1, \dots, l$ , a vertex-step is performed;
- for every integer  $k \leq l$ , the subgraph induced by the vertices  $\{v_{t_i}\}_{1 \leq i \leq k}$  is connected in  $G_{t_k}(f)$ ;
- for  $i = 1, \dots, l-1$ , the degree of  $v_{t_i}$  in  $G_t(f)$  is 2. The degree of  $v_{t_l}$  in  $G_t(f)$  is 1.

In other words, an *t*-isolated path  $\{v_{t_i}\}_{1 \leq i \leq l}$  is a path where each vertex  $v_{t_i}$ , for  $i = 2, \dots, l$ , is born at time  $t_i$  and makes its first connection to its predecessor  $v_{t_{i-1}}$ . Other than that, no other vertex or edge gets attached to  $\{v_{t_i}\}_{1 \leq i \leq l}$ . We will denote  $\{v_{t_i}\}_{1 \leq i \leq l}$  by  $v_{\vec{t}}$ .

Given  $\xi \in (0, 1)$ , we will denote by  $\mathcal{S}_{l,\xi}(t)$  the set of all *t*-isolated paths in  $G_t(f)$  of size  $l$  whose vertices were created between times  $\xi t$  and  $t$ . Having all the necessary notation, we now state the two technical lemmas needed for the proof of Theorem 1.

**Lemma 4** (Average number of *t*-isolated paths). *Let  $f$  be a non-increasing edge-step function, then, for any  $0 < \xi < 1$  and any integer  $l$ , the following lower bound holds:*

$$(3.1) \quad \mathbb{E}[|\mathcal{S}_{l,\xi}(t)|] \geq \binom{(1-\xi)t}{l} \frac{f(t)^l}{(2t)^{l-1}} \left(1 - \frac{2l}{\xi t}\right)^t.$$

Furthermore, for

$$(3.2) \quad l \leq \frac{1}{3} \left( \frac{\log t}{\log \log t} \wedge \frac{\log t}{-\log f(t)} \right),$$

we have that, for sufficiently large  $t$ ,

$$(3.3) \quad \mathbb{E}[|\mathcal{S}_{l,\xi}(t)|] \geq t^{\frac{1}{4}}.$$

**Lemma 5** (Correlation between *t*-isolated paths). *Let  $l$  be such that (3.2) is satisfied. For two *t*-isolated paths with disjoint time vectors  $\vec{t}$  and  $\vec{r}$ , we have*

$$(3.4) \quad \frac{\mathbb{P}(v_{\vec{t}} \in \mathcal{S}_{l,\xi}(t)) \mathbb{P}(v_{\vec{r}} \in \mathcal{S}_{l,\xi}(t))}{\mathbb{P}(v_{\vec{t}}, v_{\vec{r}} \in \mathcal{S}_{l,\xi}(t))} = 1 + o(1).$$

We postpone the proof of these two lemmas, since their proofs are technical. Below, we show how the main result of this section follows from them.

*Proof of Theorem 1.* Recall that Paley-Zygmund's inequality (see e.g. section 5.5 of [19]) assures us that, for any  $0 \leq \theta \leq 1$ ,

$$(3.5) \quad \mathbb{P}(|\mathcal{S}_{l,\xi}(t)| > \theta \mathbb{E}[|\mathcal{S}_{l,\xi}(t)|]) \geq (1 - \theta)^2 \frac{(\mathbb{E}[|\mathcal{S}_{l,\xi}(t)|])^2}{\mathbb{E}[|\mathcal{S}_{l,\xi}(t)|^2]}.$$

On the other hand, let  $l$  be such that (3.2) is satisfied, and consider  $\xi \in (0, 1)$ . Since it is impossible for two non disjoint and non equal isolated paths to exist at the same time, we have that

$$(3.6) \quad \begin{aligned} \mathbb{E}[|\mathcal{S}_{l,\xi}(t)|^2] &= \mathbb{E}\left[\left(\sum_{\vec{t}} \mathbb{1}\{v_{\vec{t}} \in \mathcal{S}_{l,\xi}(t)\}\right) \left(\sum_{\vec{r}} \mathbb{1}\{v_{\vec{r}} \in \mathcal{S}_{l,\xi}(t)\}\right)\right] \\ &= \sum_{\substack{\vec{t}, \vec{r} \\ \text{disjoint}}} \mathbb{P}(v_{\vec{t}}, v_{\vec{r}} \in \mathcal{S}_{l,\xi}(t)) + \mathbb{E}[|\mathcal{S}_{l,\xi}(t)|], \end{aligned}$$

and that

$$(\mathbb{E}[|\mathcal{S}_{l,\xi}(t)|])^2 = \sum_{\substack{\vec{t}, \vec{r} \\ \text{disjoint}}} \mathbb{P}(v_{\vec{t}} \in \mathcal{S}_{l,\xi}(t)) \mathbb{P}(v_{\vec{r}} \in \mathcal{S}_{l,\xi}(t)) + \sum_{\substack{\vec{t}, \vec{r} \\ \vec{r} \cap \vec{t} \neq \emptyset}} \mathbb{P}(v_{\vec{t}} \in \mathcal{S}_{l,\xi}(t)) \mathbb{P}(v_{\vec{r}} \in \mathcal{S}_{l,\xi}(t)).$$

Therefore, by lemmas 4 and 5, we obtain

$$(3.7) \quad \begin{aligned} \frac{\mathbb{E}[|\mathcal{S}_{l,\xi}(t)|^2]}{(\mathbb{E}[|\mathcal{S}_{l,\xi}(t)|])^2} &= \frac{\sum_{\text{disjoint}} \mathbb{P}(v_{\vec{t}}, v_{\vec{r}} \in \mathcal{S}_{l,\xi}(t))}{(\mathbb{E}[|\mathcal{S}_{l,\xi}(t)|])^2} + \frac{\mathbb{E}[|\mathcal{S}_{l,\xi}(t)|]}{(\mathbb{E}[|\mathcal{S}_{l,\xi}(t)|])^2} \\ &\leq \frac{\sum_{\text{disjoint}} \mathbb{P}(v_{\vec{t}}, v_{\vec{r}} \in \mathcal{S}_{l,\xi}(t))}{\sum_{\text{disjoint}} \mathbb{P}(v_{\vec{t}} \in \mathcal{S}_{l,\xi}(t)) \mathbb{P}(v_{\vec{r}} \in \mathcal{S}_{l,\xi}(t))} + \frac{1}{\mathbb{E}[|\mathcal{S}_{l,\xi}(t)|]} \\ &\leq \frac{\sum_{\text{disjoint}} (1 + o(1)) \mathbb{P}(v_{\vec{t}} \in \mathcal{S}_{l,\xi}(t)) \mathbb{P}(v_{\vec{r}} \in \mathcal{S}_{l,\xi}(t))}{\sum_{\text{disjoint}} \mathbb{P}(v_{\vec{t}} \in \mathcal{S}_{l,\xi}(t)) \mathbb{P}(v_{\vec{r}} \in \mathcal{S}_{l,\xi}(t))} + \frac{1}{\mathbb{E}[|\mathcal{S}_{l,\xi}(t)|]} \\ &\leq 1 + o(1), \end{aligned}$$

then by choosing  $\theta = \theta(t) = \mathbb{E}[|\mathcal{S}_{l,\xi}(t)|]^{-1/2}$  and combining (3.7) and (3.5) we conclude the proof.  $\square$

We now prove the previously stated lemmas.

*Proof of Lemma 4.* The random variable  $|\mathcal{S}_{l,\xi}(t)|$  can be written as

$$(3.8) \quad |\mathcal{S}_{l,\xi}(t)| = \sum_{t_1 < t_2 < \dots < t_l} \mathbb{1}\{v_{\vec{t}} \in \mathcal{S}_{l,\xi}(t)\} \implies \mathbb{E}[|\mathcal{S}_{l,\xi}(t)|] = \sum_{t_1 < t_2 < \dots < t_l} \mathbb{P}(v_{\vec{t}} \in \mathcal{S}_{l,\xi}(t)).$$



So it will be important to obtain a proper lower bound for  $\mathbb{P}(v_{\vec{t}} \in \mathcal{S}_{l,\xi}(t))$ . Given a time vector of an isolated path  $\vec{t} = (t_1, \dots, t_l)$  such that  $\xi t \leq t_1 < t_2 < \dots < t_l \leq t$ , it follows that

$$(3.9) \quad \mathbb{P}(v_{\vec{t}} \in \mathcal{S}_{l,\xi}(t)) \geq \frac{f(t)^l}{(2t)^{l-1}} \left(1 - \frac{2l}{\xi t}\right)^t,$$

since in order for  $v_{\vec{t}}$  to be in  $\mathcal{S}_{l,\xi}(t)$ , we need to assure that  $l$  vertices are born exactly at times  $t_1, \dots, t_l$  (which happens with probability greater than  $f(t)^l$ , by the monotonicity of  $f$ ), that  $v_{t_i}$  connects to  $v_{t_{i-1}}$  for every  $i = 2, \dots, l$  (which happens with probability greater than  $(2t)^{-(l-1)}$ ), and that no other vertex or edge connects to  $v_{\vec{t}}$  until time  $t$  (which happens with probability greater than  $\left(1 - \frac{2l}{\xi t}\right)^t$ ).

Finally, by counting the number of possible ways to choose  $t_1 < t_2 < \dots < t_l$  so that  $t_i \in [\xi t, t]$  for  $1 \leq i \leq l$ , we obtain 3.1.

We now assume  $l$  to be such that (3.2) holds. Stirling's formula gives us

$$\begin{aligned} \log \left( \binom{(1-\xi)t}{l} \right) &\geq C + (1-\xi)t \log((1-\xi)t) - (1-\xi)t + \frac{\log((1-\xi)t)}{2} \\ &\quad - l \log l + l - \frac{\log l}{2} \\ &\quad - ((1-\xi)t - l) \log((1-\xi)t - l) + (1-\xi)t - l - \frac{\log((1-\xi)t - l)}{2}, \end{aligned}$$

Since  $l \ll t$ , we obtain

$$(3.10) \quad \log \left( \binom{(1-\xi)t}{l} \right) \geq l \log t - l \log l - Cl.$$

Equation (3.9) then implies, again for sufficiently large  $t$ ,

$$(3.11) \quad \log(\mathbb{P}(v_{\vec{t}} \in \mathcal{S}_{l,\xi}(t))) \geq l \log f(t) - (l-1) \log t - Cl.$$

Combining the above inequality with (3.1), (3.2), and (3.10) gives us that, for large enough  $t$ ,

$$\mathbb{E}[|\mathcal{S}_{l,\xi}(t)|] \geq \exp\{l \log t - l \log l + l \log f(t) - (l-1) \log t - Cl\} \geq t^{\frac{1}{4}},$$

since

$$l \log l \leq \frac{1}{3} \frac{\log t}{\log \log t} \log \log t \left(1 - \frac{\log \log \log t + \log 3}{\log \log t}\right) \leq \frac{\log t}{3},$$

which finishes the proof of the lemma.  $\square$

*Proof of Lemma 5.* We begin the proof by making some remarks regarding the existence of a specific  $t$ -isolated path and by introducing a new notation. Given a  $t$ -isolated path  $v_{\vec{t}} = \{v_{t_j}\}_{1 \leq j \leq l}$ , we denote by  $D_r(v_{\vec{t}})$  the sum of the degrees of each of its vertices at time  $r$ , i.e. :

$$(3.12) \quad D_r(v_{\vec{t}}) = \sum_{t_i \in \vec{t}} D_r(v_{t_i}),$$

where we assumed that  $D_r(v_{t_i}) = 0$  if  $t_i > r$ . Note that if  $t_l \leq r \leq t$  and  $v_{\vec{t}}$  has size  $l$ , then  $D_r(v_{\vec{t}}) = 2l - 1$ . Furthermore, by the same reasoning as in (2.1),

$$(3.13) \quad \begin{aligned} \mathbb{P}(\Delta D_s(v_{\vec{t}}) = 0 \mid \mathcal{F}_s) &= f(s+1) \left(1 - \frac{D_s(v_{\vec{t}})}{2s}\right) + (1 - f(s+1)) \left(1 - \frac{D_s(v_{\vec{t}})}{2s}\right)^2 \\ &= 1 - \left(1 - \frac{f(s+1)}{2} - (1 - f(s+1)) \frac{D_s(v_{\vec{t}})}{4s}\right) \frac{D_s(v_{\vec{t}})}{s}. \end{aligned}$$

Regarding  $t$ -isolated paths, observe that, in order for a  $t$ -isolated path  $v_{\vec{t}}$  to appear the following must happen:

- a vertex  $v_{t_1}$  must be created at time  $t_1$ , which happens with probability  $f(t_1)$ ;
- between times  $t_1 + 1$  and  $t_2 - 1$  there can be no new connection to  $v_{t_1}$ , which, by (3.13), happens with probability

$$\prod_{r_1=t_1+1}^{t_2-1} \left(1 - \left(1 - \frac{f(r_1)}{2} - (1 - f(r_1)) \frac{1}{4(r_1-1)}\right) \frac{1}{r_1-1}\right);$$

- In general, at time  $t_k$  a vertex  $v_{t_k}$  is created and makes its first connection to  $v_{t_{k-1}}$ , no new connection is then made to  $\{v_{t_j}\}_{1 \leq j \leq k}$  between times  $t_k + 1$  and  $t_{k+1} - 1$  for every  $k = 2, \dots, l - 1$ , all this happens with probability equal to

$$\begin{aligned} &f(t_k) \frac{1}{2(t_k-1)} \prod_{r_k=t_k+1}^{t_{k+1}-1} \left(1 - \left(1 - \frac{f(r_k)}{2} - (1 - f(r_k)) \frac{D_{r_k}(v_{\vec{t}})}{4(r_k-1)}\right) \frac{D_{r_k}(v_{\vec{t}})}{r_k-1}\right) \\ &= f(t_k) \frac{1}{2(t_k-1)} \prod_{r_k=t_k+1}^{t_{k+1}-1} \left(1 - \left(1 - \frac{f(r_k)}{2} - (1 - f(r_k)) \frac{2k-1}{4(r_k-1)}\right) \frac{2k-1}{r_k-1}\right); \end{aligned}$$

- finally, a vertex  $v_{t_l}$  is born at time  $t_l$ , connects to  $v_{t_{l-1}}$  and no new connection is made to  $\{v_{t_j}\}_{1 \leq j \leq l}$  between times  $t_l + 1$  and  $t$ .

This implies

$$(3.14) \quad \begin{aligned} &\mathbb{P}(v_{\vec{t}} \in \mathcal{S}_{l,\xi}(t)) \\ &= f(t_1) \prod_{r_1=t_1+1}^{t_2-1} \left(1 - \left(1 - \frac{f(r_1)}{2} - (1 - f(r_1)) \frac{1}{4(r_1-1)}\right) \frac{1}{r_1-1}\right) \\ &\quad \times \dots \times f(t_k) \frac{1}{2(t_k-1)} \prod_{r_k=t_k+1}^{t_{k+1}-1} \left(1 - \left(1 - \frac{f(r_k)}{2} - (1 - f(r_k)) \frac{2k-1}{4(r_k-1)}\right) \frac{2k-1}{r_k-1}\right) \\ &\quad \times \dots \times f(t_l) \frac{1}{2(t_l-1)} \prod_{r_l=t_l+1}^t \left(1 - \left(1 - \frac{f(r_l)}{2} - (1 - f(r_l)) \frac{2l-1}{4(r_l-1)}\right) \frac{2l-1}{r_l-1}\right). \end{aligned}$$

Moreover, given two time vectors  $\vec{r}$  and  $\vec{t}$ , we note that  $\mathbb{P}(v_{\vec{t}}, v_{\vec{r}} \in \mathcal{S}_{l,\xi}(t))$  is only nonzero if  $\vec{r}$  and  $\vec{t}$  have either disjoint or identical sets of entries.

Then, to prove the lemma we will make a comparison between the two probabilities terms  $\mathbb{P}(v_{\vec{t}}, v_{\vec{r}} \in \mathcal{S}_{l,\xi}(t))$  and  $\mathbb{P}(v_{\vec{t}} \in \mathcal{S}_{l,\xi}(t)) \mathbb{P}(v_{\vec{r}} \in \mathcal{S}_{l,\xi}(t))$ . We can write both these terms as products in the manner of (3.14). We can then compare the terms from these products associated to each time  $s \in [\xi t, t]$ . There are two cases we must study.

*Case 1:  $s \in \vec{t}$  but  $s \notin \vec{r}$  ( $s \notin \vec{t}$  but  $s \in \vec{r}$ ).*

The product term related to time  $s$  in  $\mathbb{P}(v_{\vec{t}}, v_{\vec{r}} \in \mathcal{S}_{l,\xi}(t))$  is

$$(3.15) \quad \frac{f(s)}{2(s-1)},$$

since a new vertex is created and then makes its first connection specifically to the latest vertex of  $\vec{t}$ . On the other hand, the term related to time  $s$  in  $\mathbb{P}(v_{\vec{t}} \in \mathcal{S}_{l,\xi}(t)) \mathbb{P}(v_{\vec{r}} \in \mathcal{S}_{l,\xi}(t))$  is

$$\frac{f(s)}{2(s-1)} \left( 1 - \left( 1 - \frac{f(s)}{2} - (1-f(s)) \frac{D_{s-1}(v_{\vec{r}})}{4(s-1)} \right) \frac{D_{s-1}(v_{\vec{r}})}{s-1} \right),$$

since the term related to  $s$  in the product form of  $\mathbb{P}(v_{\vec{t}} \in \mathcal{S}_{l,\xi}(t))$  continues to be equal to (3.15), but the related term in  $\mathbb{P}(v_{\vec{r}} \in \mathcal{S}_{l,\xi}(t))$  is

$$(3.16) \quad \left( 1 - \left( 1 - \frac{f(s)}{2} - (1-f(s)) \frac{D_{s-1}(v_{\vec{r}})}{4(s-1)} \right) \frac{D_{s-1}(v_{\vec{r}})}{s-1} \right).$$

The above expression is the term that will appear regarding the time  $s$  in the fraction in the left hand side of (3.4). This case occurs  $2l$  times since the isolated paths are disjoint. Thus, recalling that  $s \in [\xi t, t]$ ,  $l \leq 3^{-1} \log(t)/\log(\log(t))$ , and that the degree of each isolated path is at most  $2l-1$ , we obtain that there exist constants  $C_1, C_2 > 0$  such that we can bound the product of all the terms of the form (3.16) from above by

$$\left( 1 - \frac{C_1}{t} \right)^{2l},$$

and from below by

$$\left( 1 - \frac{C_2 l}{t} \right)^{2l}.$$

Observe that both products go to 1 as  $t$  goes to infinity.

*Case 2:  $s \notin \vec{t}$  and  $s \notin \vec{r}$ .*

In  $\mathbb{P}(v_{\vec{r}} \in \mathcal{S}_{l,\xi}(t))$  as well as in  $\mathbb{P}(v_{\vec{t}} \in \mathcal{S}_{l,\xi}(t))$  we see terms of the form (3.16), since we must avoid the isolated paths in both events. But in the term related to  $\mathbb{P}(v_{\vec{t}}, v_{\vec{r}} \in \mathcal{S}_{l,\xi}(t))$  we actually observe

$$\left( 1 - \left( 1 - \frac{f(s)}{2} - (1-f(s)) \frac{(D_{s-1}(v_{\vec{t}}) + D_{s-1}(v_{\vec{r}}))}{4(s-1)} \right) \frac{(D_{s-1}(v_{\vec{t}}) + D_{s-1}(v_{\vec{r}}))}{s-1} \right),$$

since we must guarantee that neither isolated path receives a connection. We note however that

$$\begin{aligned}
& \left( 1 - \left( 1 - \frac{f(s)}{2} - (1 - f(s)) \frac{D_{s-1}(v_{\vec{r}})}{4(s-1)} \right) \frac{D_{s-1}(v_{\vec{r}})}{s-1} \right) \\
& \quad \times \left( 1 - \left( 1 - \frac{f(s)}{2} - (1 - f(s)) \frac{D_{s-1}(v_{\vec{t}})}{4(s-1)} \right) \frac{D_{s-1}(v_{\vec{t}})}{s-1} \right) \\
& = \left( 1 - \left( 1 - \frac{f(s)}{2} - (1 - f(s)) \frac{(D_{s-1}(v_{\vec{t}}) + D_{s-1}(v_{\vec{r}}))}{4(s-1)} \right) \frac{(D_{s-1}(v_{\vec{t}}) + D_{s-1}(v_{\vec{r}}))}{s-1} \right) \\
& \quad \times \left( 1 + O\left(\frac{l^2}{t^2}\right) \right),
\end{aligned}$$

since  $D_{s-1}(v_{\vec{t}}), D_{s-1}(v_{\vec{r}}) \leq 2l - 1$  and  $s \geq \xi t$ . In the fraction in the left hand side of (3.4), we will then have  $\Theta(t)$  terms of the form

$$\left( 1 + O\left(\frac{l^2}{t^2}\right) \right),$$

But again, as in Case 1, their product goes to 1 as  $t \rightarrow \infty$  since  $l^2 = o(t)$ . This finishes the proof of the lemma.  $\square$

#### 4. GENERAL UPPER BOUND FOR THE DIAMETER: PROOF OF THEOREM 2

In this section we provide a proof for Theorem 2. The main idea is to use Lemmas 1 and 3 to show that, with high probability, all vertices born up to time  $t^{\frac{1}{12}}$  are in a connected component with diameter 2. We then use a first moment estimate (Lemma 6 below) to show that the lengths of the paths formed by newer vertices have the desired upper bound. As in the previous section, we will state only the essential lemma needed for the proof of the main result and postpone its proof to the end of this section.

Given  $k \in \mathbb{N}$  and  $k$  time steps  $s_1, \dots, s_k \in \mathbb{N}$  such that  $s_1 < \dots < s_k$ , we say that the vector of time steps  $\vec{s} = (s_1, \dots, s_k)$  is a *vertex path* if in the process  $\{G_t(f)\}_{t \geq 1}$ , at each time  $s_j$ , for  $j = 2, \dots, k$ , a vertex is born and makes its first (vertex-step) connection to the vertex born at time  $s_{j-1}$ . We denote by

$$\{s_1 \leftarrow s_2 \leftarrow \dots \leftarrow s_k\}$$

the event where  $\vec{s}$  is a vertex path. Moreover, given  $k, t_0, t \in \mathbb{N}$ , denote by  $\mathcal{V}_{k,t_0}(t)$  the set of all vertex-paths of length  $k$  whose vertices were born between times  $t_0$  and  $t$ .

The lemma below gives an upper bound to the average cardinality of  $\mathcal{V}_{k,t_0}(t)$  in terms of  $k, t_0$  and  $t$ .

**Lemma 6.** *Using the notation above defined, we have, for  $k, t_0, t \in \mathbb{N}$ ,*

$$(4.1) \quad \mathbb{E} [|\mathcal{V}_{k,t_0}(t)|] \leq C_1 \exp \left\{ 2 \log t - (k-2) \left( \log(k-2) + C_2 + \log \left( \sum_{j=t_0}^t \frac{f(j)}{j-1} \right) \right) \right\}.$$

Next we show how the above lemma and estimates on vertices' degree implies Theorem 2.

*Proof of Theorem 2.* We prove part (a) first.

Proof of part (a): Recall that, in this part of the theorem,  $f$  is under condition (S), which holds if  $\sum_{s=2}^{\infty} f(s)/s$  is finite.

For  $t_0, t \in \mathbb{N}$  and  $\delta \in (0, 1)$ , let  $A_\delta(t_0, t)$  be the event where every vertex born before time  $t_0$  has degree at least  $t^\delta$  in  $G_t(f)$ . For  $\varepsilon \in (0, 1)$ , let  $B_\varepsilon(t)$  be the event where there exists a vertex  $v$  in  $V(G_t(f))$  such that  $D_t(v) \geq t^{1-\varepsilon}$ . Then, by Lemmas 1 and 3, we have

$$(4.2) \quad \mathbb{P} \left( A_{\frac{1}{15}}(t^{\frac{1}{12}}, t) \right) = 1 - o(1), \quad \mathbb{P} \left( B_{\frac{1}{30}}(t) \right) = 1 - o(1).$$

We have, on the event where two vertices  $u_1, u_2 \in V(G_t(f))$  are such that  $D_t(u_1) \geq t^{15^{-1}}$  and  $D_t(u_2) \geq t^{1-30^{-1}}$ ,

$$(4.3) \quad \begin{aligned} \mathbb{P}(u_1 \leftrightarrow u_2 \text{ in } G_{2t}(f)) &\leq \prod_{s=t+1}^{2t} \left( 1 - \frac{(1-f(s))t^{15^{-1}}t^{1-30^{-1}}}{2(s-1)^2} \right) \\ &\leq \exp \left\{ -\frac{t^{15^{-1}}t^{1-30^{-1}}}{16t^2} \cdot t \right\} \\ &= \exp \left\{ -\frac{t^{-30^{-1}}}{16} \right\}. \end{aligned}$$

Recall the notation  $v_i$  symbolizing the vertex born at time  $i \in \mathbb{N}$ . Together with (4.2) and the union bound, the above equation implies

$$\begin{aligned} &\mathbb{P} \left( \exists i, j \in \mathbb{N}, \text{ such that } 1 \leq i < j \leq t^{12^{-1}}, Z_i = Z_j = 1, \text{ and } \text{dist}(v_i, v_j) > 2 \text{ in } G_{2t}(f) \right) \\ &\leq o(1) + \sum_{1 \leq i < j \leq t^{\frac{1}{12}}} \mathbb{P} \left( A_{\frac{1}{15}}(t^{\frac{1}{12}}, t), B_{\frac{1}{30}}(t), \text{dist}(v_i, v_j) > 2 \text{ in } G_{2t}(f) \right) \\ &\leq o(1) + t^{-6^{-1}} \exp \left\{ -\frac{t^{-30^{-1}}}{16} \right\} \\ &= o(1). \end{aligned}$$

This implies the existence of a constant  $C_1 > 0$  such that the probability of there existing two vertices born before time  $C_1 t^{12^{-1}}$  such that the distance between said vertices is larger than 2 in  $G_t(f)$  is polynomially small in  $t$ . We now turn our attention to vertices born after  $t^{\frac{1}{13}}$ .

We will use Lemma 6 in order to bound the probability of there existing long vertex-paths formed by vertices born after  $t^{\frac{1}{13}}$ , the notion of a “long” path being  $f$ -dependent. Let  $V_t(t^{\frac{1}{13}})$  denote the set of all vertices of  $V(G_t(f))$  born before time  $t^{\frac{1}{13}}$ , let  $\mathbf{d}_{\max}(t^{\frac{1}{13}}, t)$  be the length of a maximal vertex-path of vertices born between times  $t^{\frac{1}{13}}$  and  $t$ . Let  $u_1, u_2 \in V(G_t(f))$ . Since  $G_t(f)$  is connected,

$$(4.4) \quad \begin{aligned} \text{dist}(u_1, u_2) &\leq \text{dist}(u_1, V_t(t^{\frac{1}{13}})) + \text{diam}(V_t(t^{\frac{1}{13}})) + \text{dist}(u_2, V_t(t^{\frac{1}{13}})) \\ &\leq 2\mathbf{d}_{\max}(t^{\frac{1}{13}}, t) + \text{diam}(V_t(t^{\frac{1}{13}})). \end{aligned}$$

But we know that  $\text{diam}(V_t(t^{\frac{1}{13}})) \leq 2$  with high probability. Bounding  $\mathbf{d}_{\max}(t^{\frac{1}{13}}, t)$  then gives us an a.a.s. upper bound for the diameter of  $G_t(f)$ .

Now, given  $t \in \mathbb{N}$ , if

$$(4.5) \quad k \geq 3 \left( \frac{\log t}{-\log \left( \sum_{s=t^{\frac{1}{13}}}^t \frac{f(s)}{s-1} \right)} \wedge \frac{\log t}{\log \log t} \right),$$

then, by Lemma 6, and since  $\log \left( \sum_{j=t^{\frac{1}{13}}}^t \frac{f(j)}{j-1} \right)$  is eventually negative for large  $t$  (recall again that  $f$  satisfies (S)), we have,

$$(4.6) \quad \begin{aligned} &\mathbb{P} \left( \mathbf{d}_{\max}(t^{\frac{1}{13}}, t) > k \right) \\ &\leq \mathbb{E} [|\mathcal{V}_{k, t^{1/13}}(t)|] \\ &\leq C_1 \exp \left\{ 2 \log t - (k-2) \left( \log(k-2) + C_2 - \log \left( \sum_{j=t^{\frac{1}{13}}}^t \frac{f(j)}{j-1} \right) \right) \right\} \\ &\leq C t^{-\frac{1}{2}}. \end{aligned}$$

The above upper bound and (4.4) proves part (a).

Proof of part (b): Recall that in this part, we are under condition  $(L_\kappa)$ , which holds if, for  $\kappa \in (0, 1)$  and every  $t \in \mathbb{N}$  sufficiently large, one has

$$(4.7) \quad \sum_{s=t^{1/13}}^t \frac{f(s)}{s} < (\log t)^\kappa.$$

Then, let  $k$  be so that

$$k \geq \frac{3}{1 - \kappa} \frac{\log t}{\log \log t}.$$

We then obtain, again by Lemma 6, for sufficiently large  $t$ ,

$$\begin{aligned}
(4.8) \quad & \mathbb{P} \left( \mathbf{d}_{\max}(t^{\frac{1}{13}}, t) > k \right) \\
& \leq \mathbb{E} \left[ |\mathcal{V}_{k, t^{1/13}}(t)| \right] \\
& \leq C_1 \exp \left\{ 2 \log t - (k-2) \left( \log(k-2) + C_2 - \log \left( \sum_{j=t^{\frac{1}{13}}}^t \frac{f(j)}{j-1} \right) \right) \right\} \\
& \leq C_1 \exp \left\{ 2 \log t - \frac{3}{1-\kappa} \frac{\log t}{\log \log t} ((1-\kappa) \log \log t - C - \log \log \log t) \right\} \\
& \leq Ct^{-\frac{1}{2}}.
\end{aligned}$$

Finally, the above upper bound together with (4.4) finishes the proof of the Theorem.  $\square$

Before proving Lemma 6, we will need an intermediate result.

**Lemma 7.** *Using the notation above, we have, for each vector  $\vec{s} = (s_1, \dots, s_k)$  and step function  $f$ ,*

$$(4.9) \quad \mathbb{P}(s_1 \leftarrow s_2 \leftarrow \dots \leftarrow s_k) \leq f(s_1) \frac{s_k - 1}{s_1 + 1} \prod_{m=2}^k \frac{f(s_m)}{2(s_m - 1)}.$$

*Proof.* Consider the events  $\{s_1 \leftarrow s_2 \leftarrow \dots \leftarrow s_{k-1}\}$  and  $\{s_{k-1} \leftarrow s_k\}$ , defined analogously as the event in the above equation, but for the vectors  $(s_1, \dots, s_{k-1})$  and  $(s_{k-1}, s_k)$  respectively. We have

$$\begin{aligned}
(4.10) \quad & \mathbb{P}(s_1 \leftarrow s_2 \leftarrow \dots \leftarrow s_k) = \mathbb{E} \left[ \mathbb{1}\{s_1 \leftarrow s_2 \leftarrow \dots \leftarrow s_{k-1}\} \mathbb{P}(s_{k-1} \leftarrow s_k | \mathcal{F}_{s_{k-1}}) \right] \\
& = \mathbb{E} \left[ \mathbb{1}\{s_1 \leftarrow s_2 \leftarrow \dots \leftarrow s_{k-1}\} f(s_k) \cdot \frac{D_{s_{k-1}}(v_{s_{k-1}})}{2(s_k - 1)} \right].
\end{aligned}$$

But, crucially, conditioned on the event where a vertex is born at time  $s_{k-1}$ , the degree of said vertex at time  $s_k - 1$  depends only on the connections made after time  $s_{k-1}$ , and is



therefore independent of the indicator function above. We then obtain, by (2.3),

$$\begin{aligned}
(4.11) \quad & \mathbb{P}(s_1 \leftarrow s_2 \leftarrow \cdots \leftarrow s_k) \\
&= \mathbb{E} \left[ \mathbb{1}\{s_1 \leftarrow s_2 \leftarrow \cdots \leftarrow s_{k-1}\} f(s_k) \cdot \frac{D_{s_{k-1}}(v_{s_{k-1}})}{2(s_k - 1)} \middle| Z_{s_{k-1}} = 1 \right] f(s_{k-1}) \\
&= f(s_{k-1}) f(s_k) \mathbb{P}(s_1 \leftarrow s_2 \leftarrow \cdots \leftarrow s_{k-1} | Z_{s_{k-1}} = 1) \mathbb{E} \left[ \frac{D_{s_{k-1}}(v_{s_{k-1}})}{2(s_k - 1)} \middle| Z_{s_{k-1}} = 1 \right] \\
&= \mathbb{P}(s_1 \leftarrow s_2 \leftarrow \cdots \leftarrow s_{k-1}) \frac{f(s_k)}{2(s_k - 1)} \prod_{m=s_{k-1}}^{s_k-2} \left( 1 + \frac{1}{m} - \frac{f(m+1)}{2m} \right) \\
&\leq \mathbb{P}(s_1 \leftarrow s_2 \leftarrow \cdots \leftarrow s_{k-1}) \frac{f(s_k)}{2(s_k - 1)} \exp \left\{ \sum_{m=s_{k-1}}^{s_k-2} \left( \frac{1}{m} - \frac{f(m+1)}{2m} \right) \right\},
\end{aligned}$$

by elementary properties of the exponential function. Iterating the above argument and recalling that the vertex  $s_1$  is born with probability  $f(s_1)$ , we obtain

$$\begin{aligned}
(4.12) \quad & \mathbb{P}(s_1 \leftarrow s_2 \leftarrow \cdots \leftarrow s_k) \leq f(s_1) \exp \left\{ \sum_{m=s_1}^{s_k-2} \left( \frac{1}{m} - \frac{f(m+1)}{2m} \right) \right\} \prod_{m=2}^k \frac{f(s_m)}{2(s_m - 1)} \\
&\leq f(s_1) \frac{s_k - 1}{s_1 + 1} \prod_{m=2}^k \frac{f(s_m)}{2(s_m - 1)},
\end{aligned}$$

finishing the proof of the lemma. □

We are finally able to prove Lemma 6.

*Proof of Lemma 6.* We will use Lemma 7 and then an application of the union bound. First, fix  $s_1, s_k \in \mathbb{N}$  such that  $t_0 \leq s_1 < s_k \leq t$ . We have, by Stirling's approximation formula and the positivity of the terms involved,

$$\begin{aligned}
(4.13) \quad & \sum_{\substack{s_2, s_3, \dots, s_{k-1} \\ s_1 < s_2 < \cdots < s_{k-1} < s_k}} \prod_{m=2}^{k-1} \frac{f(s_m)}{s_m - 1} \leq \frac{1}{(k-2)!} \left( \sum_{j=t_0}^t \frac{f(j)}{j-1} \right)^{k-2} \\
&\leq C \exp \left\{ -(k-2) \left( \log(k-2) - 1 - \log \left( \sum_{j=t_0}^t \frac{f(j)}{j-1} \right) \right) \right\}.
\end{aligned}$$

We can then show, by the above equation, Lemma 7, and the union bound,

$$\begin{aligned}
& \mathbb{E} [|\mathcal{V}_{k,t_0}(t)|] \\
& \leq \sum_{\substack{s_1, \dots, s_k \\ t_0 \leq s_1 < \dots < s_k \leq t}} \mathbb{P}(s_1 \leftarrow s_2 \leftarrow \dots \leftarrow s_k) \\
& \leq \sum_{\substack{s_1, \dots, s_k \\ t_0 \leq s_1 < \dots < s_k \leq t}} f(s_1) \frac{s_k - 1}{s_1 + 1} \prod_{m=2}^k \frac{f(s_m)}{2(s_m - 1)} \\
(4.14) \quad & = \frac{1}{2^{k-1}} \sum_{\substack{s_1, s_k \\ t_0 \leq s_1 < s_k \leq t}} \frac{f(s_1)f(s_k)}{s_1 + 1} \sum_{\substack{s_2, \dots, s_{k-1} \\ s_1 < s_2 < \dots < s_{k-1} < s_k}} \prod_{m=2}^{k-1} \frac{f(s_m)}{s_m - 1} \\
& \leq C \exp \left\{ -(k-2) \left( \log(k-2) + C_2 - \log \left( \sum_{j=t_0}^t \frac{f(j)}{j-1} \right) \right) \right\} \sum_{\substack{s_1, s_k \\ t_0 \leq s_1 < s_k \leq t}} \frac{f(s_1)f(s_k)}{s_1 + 1} \\
& \leq C \exp \left\{ 2 \log t - (k-2) \left( \log(k-2) + C_2 - \log \left( \sum_{j=t_0}^t \frac{f(j)}{j-1} \right) \right) \right\},
\end{aligned}$$

thus concluding the proof of the result.  $\square$

## 5. THE FAMILY OF REGULARLY VARYING FUNCTIONS: SHARP BOUNDS FOR THE DIAMETER

In this section we explore our results when more information on  $f$  is provided. In particular, we assume that  $f$  satisfies condition  $(RV_\gamma)$  for  $\gamma \in (0, 1)$  and prove sharp bounds for the diameter. Recall that this condition holds whenever

$$\exists \gamma, \text{ such that } \forall a > 0, \lim_{t \rightarrow \infty} \frac{f(at)}{f(t)} = \frac{1}{a^\gamma}.$$

Also recall that when  $f$  satisfies the above condition for  $\gamma > 0$  we say it is a regular varying function with index of regular variation  $-\gamma$ . The case  $\gamma = 0$  is said to be slowly varying.

To prove the results under the assumption of regular variation our arguments rely on the theorems from Karamata's theory. In particular, the *Representation Theorem* (Theorem 1.4.1 of [7]) and *Karamata's theorem* (Proposition 1.5.8 of [7]). The former states that if  $f$  is a regularly varying function with index  $-\gamma$ , then there exists a slowly varying function  $\ell$  such that  $f$  is of the form

$$(5.1) \quad f(t) = \frac{\ell(t)}{t^\gamma}.$$

Whereas, the latter states that if  $\ell$  is a slowly varying function, then, for any  $a < 1$

$$(5.2) \quad \int_1^t \frac{\ell(x)}{x^a} dx \sim \frac{\ell(t)t^{1-a}}{1-a},$$

and for  $a > 1$ , we have

$$(5.3) \quad \int_t^\infty \frac{\ell(x)}{x^a} dx \sim \frac{\ell(t)}{(a-1)t^{a-1}}.$$

Our proofs will also rely on another result from Karamata's Theory (Corollary A.3 of [3]) which assures that for all  $\varepsilon > 0$ ,

$$(5.4) \quad \frac{f(t)}{t^\varepsilon} \xrightarrow{t \rightarrow \infty} 0.$$

The next step is to prove the constant order of the diameter of the graphs generated by regularly varying functions and how it depends on the index of regular variation on infinity. This result is stated on Theorem 3 and we provide a proof for it below.

*Proof of Theorem 3.* We begin proving the lower bound.

Lower bound: By (5.1) we have that there exists a slowly varying function  $\ell$  such that  $f(t) = \ell(t)/t^\gamma$ . Moreover, (5.4) implies that

$$(5.5) \quad \frac{\log \ell(t)}{\log t} \rightarrow 0.$$

Thus, we have that

$$(5.6) \quad -\frac{\log t}{\log f(t)} = \frac{\log t}{\gamma \log(t) + \log \ell(t)} = \frac{1 - o(1)}{\gamma}.$$

Applying Theorem 1 gives us the desired lower bound.

Upper bound: By the Representation Theorem and (5.4), we have that  $f$  also satisfies condition (S). Just notice that

$$(5.7) \quad \sum_{s=1}^\infty \frac{f(s)}{s} = \sum_{s=1}^\infty \frac{\ell(s)}{s^{1+\gamma}} < \infty.$$

And by Karamata's Theorem (Equation (5.3) in particular) we have that

$$(5.8) \quad \sum_{s=t^{\frac{1}{13}}}^t \frac{\ell(s)}{s^{1+\gamma}} \sim \frac{\ell(t)}{t^{\gamma \frac{12}{13}}} \implies -\frac{\log t}{\log \left( \sum_{s=t^{\frac{1}{13}}}^t \frac{\ell(s)}{s^{1+\gamma}} \right)} \sim \frac{\log t}{\frac{\gamma}{13} \log t + \log \ell(t)} = \frac{13(1 - o(1))}{\gamma}$$

And finally, applying Theorem 2 we prove the result. □

## 6. FINAL COMMENTS

We end this paper with a brief discussion on the affine version of our model and how dropping some regularity conditions on  $f$  may produce a sequence of graphs  $\{G_t\}_{t \in \mathbb{N}}$  that has a subsequence which is essentially of complete graphs and another one which is close to the BA-model.

**Affine version.** At the introduction, we have discussed the affine version of the PA-rule, which we recall below.

$$\mathbb{P}(v_{t+1} \rightarrow u | G_t) = \frac{\text{degree}(u) + \delta}{\sum_{w \in G_t} (\text{degree}(w) + \delta)}.$$

In [3], the authors showed that the effect of the affine term  $\delta$  vanishes in the long run when one is dealing with the empirical degree distribution. I.e., their results show that  $\delta$  has no effect on the degree sequence of the graphs, which is not observed in the affine version of the BA-model, for which the exponent of the power-law distribution depends on  $\delta$ , see [11]. However, we believe  $\delta$  may have an increasing/decreasing effect on the diameter's order. One also may find interesting to consider  $\delta = \delta(t)$  and investigate which influence takes over: is it the edge-step function or the affine term?

**Dropping regularity conditions.** In this part we discuss an example which hints that dropping some assumptions on  $f$  may produce a somewhat pathological sequence of graphs. For instance, if we drop the assumption of  $f$  being non-increasing, we may obtain a sequence of graphs whose diameter sequence oscillates between 1 and  $\log t$ . More generally, the sequence of graphs oscillates between graphs similar to the BA-random tree,  $\{\text{BA}_t\}_{t \in \mathbb{N}}$ , and graphs close to complete graphs.

Let  $(t_k)_{k \in \mathbb{N}}$  be the following sequence:  $t_0 = 1$  and  $t_{k+1} := \exp\{t_k\}$ , for  $k > 1$ . Now, let  $h$  be the edge-step function defined as follows

$$(6.1) \quad h(t) = \begin{cases} 1 & \text{if } t \in [t_{2k}, t_{2k+1}], \\ 0 & \text{if } t \in (t_{2k+1}, t_{2k+2}). \end{cases}$$

The idea behind such  $h$  is that between times  $[t_{2k}, t_{2k+1}]$  the process behaves essentially as the traditional BA-model, whereas at interval  $(t_{2k+1}, t_{2k+2})$  the process “messes things up” connecting almost all vertices by only adding new edges. Moreover, in both regimes the process has time enough to “forget about what was built in the past”.

Using Lemma 3 and reasoning as in (4.3), one may prove that

$$\text{diam}_{G_{t_{2k+1}}^{13}}(h)G_{t_{2k+1}}(h) \leq 2, \text{ a.a.s.}$$

However, the process does not add any new vertex in the interval  $(t_{2k+1}, t_{2k+2})$ . Therefore,  $\text{diam}_{G_{t_{2k+2}}}(h) \leq 2$ , a.a.s.

On the other hand, we may collapse all vertices of  $G_{t_{2k+2}}(h)$  into a single super vertex with  $t_{2k+2}$  loops losing just one unit in its diameter. This way  $\{G_s(h)\}_{t_{2k+2}}^{t_{2k+3}}$  is distributed as the BA-random tree started from the super vertex. Thus (see Theorem 7.1 of [23]),

$$\text{diam}G_{t_{2k+3}}(h) \approx \log t_{2k+3} = t_{2k+2}.$$

Roughly speaking,  $\{\text{diam}G_t(h)\}_{t \geq 1}$  has a subsequence that is bounded by 2 and another one that grows logarithmic in time. The conclusion is that if we drop some monotonicity assumption on  $f$ , we may obtain a sequence of graphs having at least two subsequences that are completely different as graphs.

**Acknowledgements** We are thankful to Roberto Imbuzeiro Oliveira for the many inspiring and pleasant conversations. *C.A.* was partially supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), grants 2013/24928-2 and 2015/18930-0, and by the DFG grant SA 3465/1-1. *R.R.* was partially supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and by the project Stochastic Models of Disordered and Complex Systems. The Stochastic Models of Disordered and Complex Systems is a Millennium Nucleus (NC120062) supported by the Millenium Scientific Initiative of the Ministry of Science and Technology (Chile). *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.

## REFERENCES

- [1] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Rev. Modern Phys.*, 2002.
- [2] C. Alves, R. Ribeiro, and R. Sanchis. Large communities in a scale-free network. *Journal of Statistical Physics*, 166(1):137–149, 2017.
- [3] C. Alves, R. Ribeiro, and R. Sanchis. Preferential attachment random graphs with edge-step functions. *Journal of Theoretical Probability*, 11 2019.
- [4] M. Amin Abdullah and N. Fountoulakis. A phase transition in the evolution of bootstrap percolation processes on preferential attachment graphs. *Random Structures & Algorithms*, 52(3):379–418, 2018.
- [5] A. Backhausz and T. F. Móri. Asymptotic properties of a random graph with duplications. *J. Appl. Probab.*, 52(2):375–390, 06 2015.
- [6] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 1999.
- [7] N. H. Bingham, C. M. Goldie, and J. L. Teugels. Regular variation, volume 27 of encyclopedia of mathematics and its applications, 1989.
- [8] B. Bollobás and O. Riordan. Mathematical results on scale-free random graphs. in handbook of graphs and networks: From the genome to the internet. pages 1–34, 2003.
- [9] S. Chatterjee and R. Durrett. Contact processes on random graphs with power law degree distributions have critical value 0. *Ann. Probab.*, 37(6):2332–2356, 11 2009.
- [10] F. Chung and L. Lu. *Complex Graphs and Networks (Cbms Regional Conference Series in Mathematics)*. American Mathematical Society, Boston, MA, USA, 2006.
- [11] C. Cooper and A. Frieze. A general model of web graphs. *Random Structures & Algorithms*, 22(3):311–335, 2003.
- [12] S. Dereich and M. Ortgiese. Robust analysis of preferential attachment models with fitness. *Combinatorics, Probability and Computing*, 23(3):386–411, 2014.

- [13] S. Dommers, R. van der Hofstad, and G. Hooghiemstra. Diameters in preferential attachment models. *Journal of Statistical Physics*, 139(1):72–107, 2010.
- [14] N. Eggemann and S.D. Noble. The clustering coefficient of a scale-free random graph. *Discrete Applied Mathematics*, 159(10):953 – 965, 2011.
- [15] P. Erdős and A. Rényi. On random graphs i. *Publicationes Mathematicae Debrecen*, 6:290, 1959.
- [16] D. Fernholz and V. Ramachandran. The diameter of sparse random graphs. *Random Structures & Algorithms*, 30(4):482–516, 2007.
- [17] E. Jacob and P. Mörters. Spatial preferential attachment networks: Power laws and clustering coefficients. *Ann. Appl. Probab.*, 25(2):632–662, 04 2015.
- [18] E. Jacob and P. Mörters. The contact process on scale-free networks evolving by vertex updating. *Royal Society Open Science*, 4(5):170081, 2017.
- [19] Russell Lyons and Yuval Peres. *Probability on Trees and Networks*, volume 42 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, New York, 2016. Available at <http://pages.iu.edu/~rdlyons/>.
- [20] Y. Malyshkin and E. Paquette. The power of choice combined with preferential attachment. *Electronic Communications in Probability*, 19(44):1–13, 2014.
- [21] S.H Strogatz and D. J. Watts. Collective dynamics of ‘small-world’ networks. *Nature*, 393(6684):440–442, June 1998.
- [22] E. Thörnblad. Asymptotic degree distribution of a duplication–deletion random graph model. *Inter. Math.*, 11, 03 2014.
- [23] R. Van Der Hofstad. Random graphs and complex networks volume 2. Available on [http://www.win.tue.nl/~hofstad/ NotesRGCNII. pdf](http://www.win.tue.nl/~hofstad/NotesRGCNII.pdf), 2018.
- [24] D. Znamenski, G. Hooghiemstra, and R. van der Hofstad. A phase transition for the diameter of the configuration model. *Internet Mathematics*, 4(1):113–128, 2007.

<sup>1</sup> INSTITUTE OF MATHEMATICS, , UNIVERSITY OF LEIPZIG – AUGUSTUSPLATZ 10, 04109 LEIPZIG  
E-MAIL: [caio.alves@math.uni-leipzig.de](mailto:caio.alves@math.uni-leipzig.de)

<sup>2</sup> PUC CHILE, AV. VICUÑA MACKENNA 4860, MACUL, LA FLORIDA, REGIÓN METROPOLITANA, CHILE.  
E-MAIL: [rribeiro@impa.br](mailto:rribeiro@impa.br)

<sup>3</sup>DEPARTAMENTO DE MATEMÁTICA, UNIVERSIDADE FEDERAL DE MINAS GERAIS, AV. ANTÔNIO CARLOS 6627 C.P. 702 CEP 30123-970 BELO HORIZONTE-MG, BRAZIL  
E-MAIL: [rsanchis@mat.ufmg.br](mailto:rsanchis@mat.ufmg.br)