The Concentration of the Maximum Degree in the Duplication-Divergence Models*

Alan Frieze[1[0000−0002−8481−5615]], Krzysztof Turowski[2[0000−0003−3871−1038]], and Wojciech Szpankowski[2,3[0000−0001−9062−0067]]

[1] Department of Mathematical Sciences,
Carnegie Mellon University, Pittsburgh, PA, USA
alan@random.math.cmu.edu
[2] Department of Theoretical Computer Science, Jagiellonian University, Poland
{krzysztof.turowski, wojciech.szpankowski}@uj.edu.pl
[3] Center for Science of Information, Department of Computer Science,
Purdue University, West Lafayette, IN, USA
szpan@purdue.edu

Abstract. We pursue the analysis of the maximum degree in a dynamic duplication-divergence graph model defined by Solé et al. in which a new node arriving at time t first randomly selects an existing node and connects to its neighbors with probability p, and then connects to the other nodes with probability r/t. This model is often said to capture the growth of some real-world processes e.g. biological or social networks. However, there are only a handful of rigorous results concerning this model. Here we study the distribution of the maximum degree of a vertex in graphs generated by this model.
In this paper we prove that for $\frac{1}{2} < p < 1$ with high probability the maximum degree is asymptotically quite surely concentrated around t^p, i.e. it deviates from this value by at most a polylogarithmic factor. Our findings are a step towards a better understanding of the overall structure of graphs generated by this model, especially the degree distribution, compression, and symmetry.

Keywords: Random graphs · Duplication-divergence model · Degree distribution · Maximum degree · Large deviation.

1 Introduction

Studying structural properties of graphs (e.g., symmetry, compressibility, vertex degree) is a popular topic of research in computer science and discrete mathematics ever since the seminal work of Paul Erdős and Alfréd Rényi [8] . Recently attention has turned to dynamic graphs such as preferential attachment (Barabási-Albert) graphs [1], Watts-Strogatz small world graphs [25]

* This work was supported by NSF Center for Science of Information (CSoI) Grant CCF-0939370, by NSF Grants CCF-1524312, CCF-2006440, CCF-2007238, DMS1952285 and, in addition, by the National Science Center, Poland, Grant 2020/39/D/ST6/00419.

or duplication-divergence graphs. Dynamic graphs, in which the edge- and/or vertex-sets are functions of time, are ubiquitous in diverse application domains ranging from biology to finance to social science. Deriving novel insights and knowledge from dynamic structures is a key challenge and understanding the structural properties of such dynamic graphs is critical for new characterizations and insights of the underlying dynamic processes.

Numerous networks in the real world change over time, in the sense that nodes and edges enter and leave the networks. To explain their macroscopic properties (e.g., subgraph frequencies, diameter, degree distribution, symmetry) and to make predictions and other inferences (such as community detection, graph compression, order of node arrivals), several generative models have been proposed [19, 24]. Typically, one tries to capture the behavior of well-known graph parameters under probability distributions induced by the models, e.g. the distribution of the number of vertices with a given degree, the number of connected components, the existence of Hamiltonian paths or other parameters like clique number and chromatic number (see [3, 9, 13] for overviews of the main results in the area).

In this paper we make further progress on structural properties of the *duplication-divergence* graph models, in which vertices arrive one by one, select an existing node as a parent, connect to the some neighbors of its parent and other vertices according to some pre-defined rule. More precisely, a newly arriving node at time t first selects randomly an existing node and connects to its neighbors with probability p; and then connects to other nodes with probability r/t. The particular model which we bring under consideration is a duplication-divergence model, first defined by Solé, Pastor-Satorras et al. [21]. It has been a popular object of study because it has been shown empirically that its degree distribution, small subgraph (graphlets) counts and number of symmetries fit very well with the structure of some real-world biological and social networks, e.g. protein-protein and citation networks [5, 20, 22]. This suggests a possible real-world significance for the duplication-divergence model, which further motivates the studies of its structural properties. However, it is also one of the least understood models, much less so than the Erdős-Rényi or preferential attachment models. At the moment there exist only a handful of results related to the behavior of the degree distribution of the graphs generated by this model. Unlike other dynamic graphs such as the preferential attachment model, the graphs generated by the duplication-divergence model can be very symmetric or quite asymmetric. In Figure 1 from [22] it is shown that there exist certain ranges of the model parameters p and r such that the graphs generated from the model are highly symmetric, and certain ranges such that the graphs are asymmetric. Here the symmetry is measured by the size of the automorphism group $|\mathrm{Aut}(G)|$, i.e. the number of distinct mapping of vertices onto themselves preserving the adjacency matrix. Still the basic question about the conditions under which the generated graph is symmetric or not remains unanswered. We believe that proving results about the range of the maximum degree can be a stepping stone for

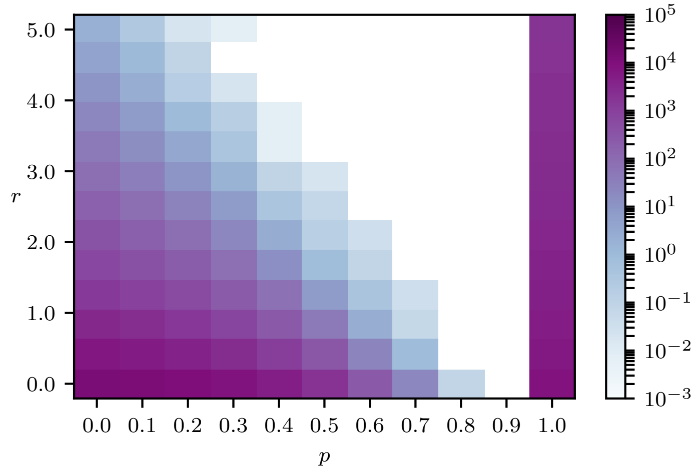

Fig. 1. Symmetry of graphs ($\log|\mathrm{Aut}(G)|$) generated by the Solé-Pastor-Satorras duplication-divergence model, based on simulations from [22].

rigorous general results regarding symmetries and compression, just as it has been in the case for other random graph models.

In particular, the parameters such as the maximum degree of a random graph and the degree of a given vertex are parameters that are studied not only for their own sake, but it turns out that their analysis opens the way to further results. Let us recall here two examples of these insights related to the questions of graph asymmetry and incompressibility.

First, Łuczak et al. [17] used the estimation of these parameters to prove that the preferential attachment model with $m \geq 3$ (where m is the number of edges added when a new node arrives) generates asymmetric graphs (i.e. graphs with only one automorphism) with high probability. This was achieved by proving two properties: (A) for any pair of early vertices t_1 and t_2 the degrees of both nodes t_1 and t_2 are distinct, and (B) for any pair of late vertices their corresponding neighbors are not the same, in particular, they have different sets of early neighbors (and therefore, a permutation of t_1 and t_2 does not produce symmetry). We believe that this approach to asymmetry analysis can be extended to the duplication-divergence model and it requires knowledge of the maximum degree which is exactly the topic of this paper.

A second usage of these parameters was presented by Chierichetti et al. in [4]. For example, for the preferential attachment model they used an upper bound on the maximum degree and the degree of a vertex arriving at time s to show that the entropy over all graphs on t vertices generated by this model is bounded by $\Omega(t \log t)$. They also used their bound on vertex degrees to provide lower bounds on graph entropy for several other random graph models known in the literature, e.g. the copying model or ACL model (see also [18] for the preferential attachment graph compression algorithm).

Therefore, we turn our attention to the asymptotic behavior of the distribution of degrees of vertices in random graphs generated by the duplication-divergence model. Let us recall that, for example, for Erdős-Rényi model $\mathtt{ER}(t,p)$ it is known that the degree distribution approximately follows the Poisson distribution with a tail decreasing exponentially [2]. Clearly, the degree of each vertex is a random variable with the binomial distribution, so it is highly concentrated around its mean $(t-1)p$. Moreover, the maximum degree is also highly concentrated around $(t-1)p+\sqrt{2p(1-p)(t-1)\log t}$ [9, Theorem 3.5]. For the preferential attachment model $\mathtt{PA}(t,m)$ it was proved that the degree distribution exhibits scale-free behaviour, i.e. the number of vertices with degree k is proportional to k^{-3} [3]. In addition, if we consider a vertex arriving at time s, its degree in graph on t vertices is proportional to $\sqrt{t/s}$ on average and with high probability it does not exceed $\sqrt{t/s}\log^3 t$ [6]. In the next section we discuss in some details recent results regarding the degree distribution of the duplication-divergence graph model.

Here we provide analogous results for the duplication-divergence model. The paper is organized as follows: in Section 2 we present a formal definition of the duplication-divergence model, recall previous results related to the properties of the degree distribution and introduce our main results. In Section 3.1 and Section 3.2 we prove upper bounds for the degrees for earlier and later vertices arriving in the graph, respectively. Finally, in Section 3.3 we give a proof of the lower bound for the maximum degree in the graph.

2 Model definition and main results

We formally define the duplication-divergence model $\mathtt{DD}(t,p,r)$, introduced by Solé et al. [21]. Then we summarize our main results about high-probability bounds on the the maximum degree.

Throughout the paper we use standard graph notation from [7], e.g. $V(G)$ denotes the vertex set of a graph G, $\deg_G(s)$ – the degree of node s in G and $\Delta(G)$ – the maximum degree of a vertex in G. All graphs considered in the paper are simple.

G_t denotes a graph on t vertices. Because in the paper we deal with graphs that are dynamically generated, we assume that the vertices are identified with the natural numbers according to their arrival time. We use the notation $\deg_t(s)$ for the random variable denoting the degree of vertex s at time t i.e. after t vertices have been added in total.

Let us now formally define the model $\mathtt{DD}(t,p,r)$ as follows: let G_T be a fixed graph on $T \le t$ vertices, with vertices having distinct labels from 1 to T. Let also $0 \le p \le 1$ and $0 \le r \le T$ be the parameters of the model. Now, for every $n = T, T+1, \ldots, t-1$ we create G_{n+1} from G_n according to the following rules:

1. we add a new vertex with label $n+1$ to the graph,
2. we choose a vertex u from G_n uniformly at random – and we denote u as $\mathrm{parent}(n+1)$,

3. for every vertex v:
   (a) if v is adjacent to u in G_n, then add an edge between v and $n+1$ with probability p,
   (b) if v is not adjacent to u in G_n, then add an edge between v and $n+1$ with probability $\frac{r}{n}$.

All edge additions are independent random Bernoulli variables.

We now review in some detail, recent results on the degree distribution. For example, for $p < 1$ and $r = 0$, it is shown in [11] that even for large p the limiting distribution of degree frequencies indicates that almost all vertices are isolated as $t \to \infty$. Moreover, from [16] we know that the number of vertices of degree one is $\Omega(\log t)$ but again the precise rate of growth of the number of vertices with any fixed degree $k > 0$ is currently unknown. Recently, also for $r = 0$, in [14, 12] the authors showed that for $0 < p < e^{-1}$ the non-trivial connected component has a degree distribution that has a power-law behavior with the exponent is equal to γ satisfying $3 = \gamma + p^{\gamma-2}$.

Now let us turn to results directly related to the question of maximum degree. For example, in [23] it was shown that for any fixed s asymptotically as $t \to \infty$ it holds that

$$\mathbb{E}[\deg_t(s)] = \begin{cases} \Theta(\ln t) & \text{if } p = 0 \text{ and } r > 0, \\ \Theta(t^p) & \text{otherwise.} \end{cases}$$

Note that by the close relation between parameters $\Delta(G_t)$ and $\deg_t(s)$ we can establish easily that $\mathbb{E}[\Delta(G_t)] = \Omega(t^p)$ when $p > 0$ or $r = 0$, and $\mathbb{E}[\Delta(G_t)] = \Omega(\ln t)$ otherwise.

It turns out that a lower bound on maximum degree is easily established as a byproduct of existing results by Frieze et al. [10]: for $\frac{1}{2} < p < 1$ and $G_t \sim \mathtt{DD}(t, p, r)$ with $p > 0$ and $s = O(1)$ it holds that

$$\Pr\left[\deg_t(s) \leq \frac{C}{A} t^p \log^{-3-\varepsilon}(t)\right] = O(t^{-A})$$

for some fixed constant $C > 0$ and any $A > 0$. This lower bound holds for the maximum degree because for any s it holds that $\deg_t(s) \leq \Delta(G_t)$. In the same paper, Frieze et al. also proved that for $\frac{1}{2} < p < 1$, $G_t \sim \mathtt{DD}(t, p, r)$ and $s = O(1)$ it holds that

$$\Pr[\deg_t(s) \geq A\, C\, t^p \log^2(t)] = O(t^{-A})$$

for some fixed constant $C > 0$ and any $A > 0$. They also left as an open problem the question of the behavior of the right tail of the maximum degree distribution or, equivalently, of the upper bound on $\deg_t(s)$ for larger s that holds with high probability.

In this paper, we solve this problem. More precisely, we obtain two major results: first, we provide a bound $\deg_t(s) \leq t^p \text{polylog}(t)$ which holds quite surely (i.e. at least $1 - O(t^{-A})$ for any given $A > 0$ [15]). We prove that this bound is valid for all vertices in G_t, not only for $s = O(1)$ as before, leading to the estimate $\Delta(G_t) \leq t^p \text{polylog}(t)$ for any $\varepsilon > 0$ with high probability. Next, we provide a precise lower bound and we show that there exists an early vertex s such that $\deg_t(s) \geq (1-\varepsilon)t^p$ for any $\varepsilon > 0$ quite surely. Putting everything together we obtain the main result of this paper, that is:

Theorem 1. *Let $\frac{1}{2} < p < 1$. Asymptotically for $G_t \sim \mathtt{DD}(t,p,r)$*

$$\Pr[(1-\varepsilon)t^p \leq \Delta(G_t) \leq (1+\varepsilon)t^p \log^{5-4p}(t)] = O(t^{-A})$$

for any constants $\varepsilon > 0$ and $A > 0$,

In other words, we are now certain that the maximum degree of the graph is concentrated in the sense that by moving only by some polylogarithmic factor from the mean to both left and right we observe the tail decay which is greater than any polynomial.

3 Analysis and proofs

3.1 Upper bound, early vertices

The main idea of the proof of the upper bound of the maximum degree is as follows: we first find for small s (i.e. $s \leq t_0$) a Chernoff-type bound on the growth of $\deg_\tau(s)$ over an interval of certain length h.

Then, we introduce auxiliary deterministic sequences t_i and X_{t_i} such that $t_0 < \ldots < t_{k-1} < t \leq t_k$. The definition of these sequences stems from the bound mentioned above, in particular from the relation between h and the growth of the degree, guaranteed with high probability. Ultimately, we prove $\deg_\tau(s) \leq X_\tau$ with high probability for all $s \leq t_0$.

Let us start with providing a Chernoff-type bound on the growth of the degree of a given early vertex (with proof in Appendix A):

Lemma 1. *Let $1 \leq s \leq \tau \leq t$. Let X_τ be any value such that $\deg_\tau(s) \leq X_\tau$. Then for any $h \leq \varepsilon X_\tau$ with $\varepsilon \in (0,1)$ it is true that*

$$\Pr\left[\deg_{\tau+h}(s) \geq \deg_\tau(s) + (1+3\varepsilon)\,\tfrac{h(pX_\tau+r)}{\tau}\right] \leq \exp\left(-\tfrac{h\varepsilon^2(1+\varepsilon)(pX_\tau+r)}{3\tau}\right).$$

We can immediately deduce how large h has to be to get a polynomial tail:

Corollary 1. *Let $1 \leq s \leq \tau \leq t$. Let $X_\tau \geq 0$, $\varepsilon \in (0,1)$ be values such that asymptotically for any $A > 0$, it holds that $\deg_\tau(s) \leq X_\tau$ and $3A\tau\log t \leq \varepsilon^3 X_\tau(pX_\tau + r)$. Then for any $h \in \left[\frac{3A\tau\log t}{\varepsilon^2(pX_\tau+r)}, \varepsilon X_\tau\right]$ it is true that*

$$\Pr\left[\deg_{\tau+h}(s) > \deg_\tau(s) + (1+3\varepsilon)\,\tfrac{h(pX_\tau+r)}{\tau}\right] = O(t^{-A}).$$

Now we provide the definitions for two auxiliary sequences that we mentioned earlier:

Definition 1. *Let $0 < p < 1$ be fixed with certain α, β_i and ϕ. We define the increasing sequences $(t_i)_{i=0}^{k}$ and $(X_{t_i})_{i=0}^{k}$ and an integer k in the following way:*

$$t_0 = \phi, \qquad t_{i+1} = t_i + \frac{\alpha\, t_i \log t_i}{X_{t_i}}, \qquad t_{k-1} < t \leq t_k,$$
$$X_{t_0} = t_0, \qquad X_{t_{i+1}} = X_{t_i} + \beta_i \log t_i.$$

Note that α, β_i and ϕ can be, and indeed we will specify them, as dependent on t. However, for brevity, we assume the possible dependency on t as implicit.

Observe that inductively from the definition it follows that if $\alpha \geq \beta_i$, then $X_{t_i} \leq t_i$ for all $i = 0, 1, \ldots, k$.

Moreover, note that we do not specify the values of X_τ for τ other than $\{t_0, t_1, \ldots, t_k, \ldots\}$. In the rest of the paper we will be using precisely these values in the proofs, so such a definition is sufficient for our purposes. For convenience, we only assume that for any $\tau \in (t_l, t_{l+1})$ for some $l = 0, 1, \ldots, k-1$ the sequence is completed in any way such that $X_{t_l} \leq X_\tau \leq X_{t_{l+1}}$.

Now we analyze the asymptotic properties of these sequences. We start with a simple lower bound (see Appendix B for proof):

Lemma 2. *Assume that $\phi \geq \log^2 t$, $\alpha \leq \sqrt{\phi}$ and $\beta_i \geq \alpha(p-\delta)$ for some $\delta \in [0, p)$. Asymptotically as $t \to \infty$ for any $i = 0, 1, \ldots, k$ we have $X_{t_i} \geq t_i^{p-\delta}$.*

It enables us to we prove (in Appendix C) the upper bound:

Lemma 3. *Assume that $\phi \geq \log^3 t$, $\alpha(p-\delta) \leq \beta_i \leq \alpha p + \frac{\alpha}{2\log t_i}$ for some $\delta \in [0, p)$. It holds asymptotically as $t \to \infty$ that $X_{t_i} \leq \phi^{1-p} t_i^p \log t_i$ for all $i = 0, 1, \ldots, k$.*

Corollary 2. *If $\alpha \leq \phi$, then for the value of k such that $t_{k-1} < t \leq t_k$ it is true that $\alpha k < t$.*

Proof. We know from the definition of t_i and Lemma 3 that

$$t > t_{k-1} \geq t_0 + \sum_{i=0}^{k-2} \frac{\alpha t_i \log t_i}{\phi^{1-p} t_i^p \log t_i} \geq t_0 + \sum_{i=0}^{k-2} \alpha \geq \phi + (k-1)\alpha > \alpha\, k$$

as needed.

Here let us note (and prove in Appendix D) the relation between the last elements of the sequences $(t_i)_{i=0}^{k}$, $(X_{t_i})_{i=0}^{k}$ and the final values themselves:

Lemma 4. *Let ε be any positive constant. Assume that $\phi \geq \log^3 t$, $\alpha \leq \sqrt{\phi}$, $\alpha(p-\delta) < \beta_i \leq \alpha p + \frac{\alpha}{2\log t_i}$ for some $\delta \in [0, p)$. It holds asymptotically as $t \to \infty$ that $(1-\varepsilon)t_k \leq t \leq (1+\varepsilon)t_{k-1}$ and $(1-\varepsilon)X_{t_k} \leq X_{t_k} \leq (1+\varepsilon)X_{t_{k-1}}$.*

Observe that since we will use $\phi < t$, it holds that $k \geq 1$.

Let us denote by $\mathcal{A}_i(s)$ the event that $\deg_{t_i}(s) \leq X_{t_i}$ for a fixed $s \leq t_i$. Now we proceed with the main theorem:

Theorem 2. *For $G_t \sim \mathtt{DD}(t,p,r)$ with $\frac{1}{2} < p < 1$ and $s \in [1, 74529(A+1)^2 \log^4 t]$ it holds asymptotically that*

$$\Pr\left[\deg_t(s) > (1+\varepsilon)t^p \log^{5-4p} t\right] = O(t^{-A})$$

for any constants $\varepsilon > 0$ and $A > 0$.

Proof. Throughout the proof we will use sequences $(t_i)_{i=0}^k$ and $(X_{t_i})_{i=0}^k$ with $\alpha = 273p^3(A+1)\log^2 t$, $\beta_i = \alpha p + \frac{\alpha}{2\log t_i}$ and $\phi = 74529(A+1)^2 \log^4 t$ and $t_{k-1} < t \leq t_k$.

Observe that all the assumptions of Lemma 2, Lemma 3 and Corollary 2 are met so we know that $\max\{74529(A+1)^2 \log^4 t, t_i^p\} \leq X_{t_i} \leq t_i^p \log^{5-4p} t$ for all $i = 0, 1, \ldots, k$ and also $k < \frac{t}{\log^2 t}$. Moreover, if $\mathcal{A}_i(s)$ holds, then the assumptions of Corollary 1 also are true for $\tau = t_i$ and $h = \frac{\alpha t_i \log t_i}{X_{t_i}}$ as $t_i \to \infty$ since for any constant $A > 0$ and $\varepsilon = \frac{1}{9p \log t_i}$ it holds that

$$\frac{3At_i \log t}{\varepsilon^2(pX_{t_i} + r)} < h = \frac{\alpha t_i \log t_i}{X_{t_i}} < \varepsilon X_{t_i}.$$

The left inequality is easy to verify as the left element is $\Theta\left(\frac{t_i \log^2 t_i \log t}{X_{t_i}}\right)$ and h grows like $\Theta\left(\frac{t_i \log t_i \log^2 t}{X_{t_i}}\right)$. The right inequality follows directly from Lemma 2, provided we choose some $\delta \in \left[0, p - \frac{1}{2}\right)$ so that X_{t_i} grows sufficiently fast.

Moreover, since $\beta_i > \alpha p$, we know that for $\varepsilon = \frac{1}{9p\log t_i}$ asymptotically

$$X_{t_{i+1}} - X_{t_i} = \beta_i \log t_i \geq (1+3\varepsilon)\frac{h(pX_{t_i} + r)}{t_i}.$$

where $h = \frac{1}{1+\frac{1}{2p\log t_i}} \frac{\beta_i t_i \log t_i}{pX_{t_i}} \leq \varepsilon X_{t_i}$.

Therefore, Corollary 1 implies that for any constant $A > 0$ and $\varepsilon = \frac{1}{9p\log t_i}$ it is true that $\Pr[\neg\mathcal{A}_{i+1}(s)|\mathcal{A}_i(s)] = O(t^{-A})$.

Clearly, for any $1 \leq s \leq t_0$ we know that $\mathcal{A}_0(s)$ always holds so $\Pr[\neg\mathcal{A}_0(s)] = 0$. Finally, we obtain using Lemma 4 and Corollary 1 that

$$\begin{aligned} \Pr[\deg_t(s) > X_{t_k}] &\leq \Pr[\deg_{t_k}(s) > X_{t_k}] = \Pr[\neg\mathcal{A}_k(s)] \\ &\leq \sum_{i=0}^{k-1} \Pr[\neg\mathcal{A}_{i+1}(s)|\mathcal{A}_i(s)] + \Pr[\neg\mathcal{A}_0(s)] = \sum_{i=0}^{k-1} O(t^{-A}) = O(t^{-A+1}). \end{aligned}$$

3.2 Upper bound, late vertices

In the second part of the proof we also use the sequences $(t_i)_{i=0}^k$ and $(X_{t_i})_{i=0}^k$ as defined in Definition 1. Moreover, in their definition throughout this section we use the same constants as in the proof of Theorem 2: $\alpha = 273p^3(A+1)\log^2 t$, $\beta_i = \alpha p + \frac{\alpha}{2\log t_i}$ and $\phi = 74529(A+1)^2 \log^4 t$.

The proof consists of showing that for $s \in [t_i, t_{i+1})$ for some $i = 0, 1, \ldots, k-1$ the degree of the vertex when it appears in the graph (i.e. $\deg_s(s)$) is with high probability significantly smaller than its respective $X_{t_{i+1}}$. Furthermore, we show that the increase in the degree between $\deg_s(s)$ and $\deg_{t_{i+1}}(s)$ with high probability cannot compensate for this difference. Thus, X_t (or, to be more precise, X_{t_k}) gives us a good upper bound on $\deg_t(s)$ for all s – and therefore also we obtain an upper bound for $\Delta(G_t)$.

Let us introduce an auxiliary event $\mathcal{B}_l(s) = \bigcup_{\tau=1}^{s} \mathcal{A}_l(\tau) = [\deg_{t_l}(s) \leq X_{t_l}$ for any s and l such that $s \leq t_l]$.

Lemma 5. *Let $s \in (t_l, t_{l+1}]$ for some $l = 0, 1, \ldots, k-1$. Then, for any $\varepsilon \in (0, 1)$*

$$\Pr\left[\deg_s(s) \geq (1+\varepsilon)(pX_{t_{l+1}} + r) | \mathcal{B}_l(t_l) \wedge \mathcal{B}_{l+1}(s-1)\right] \leq \exp\left(-\frac{\varepsilon^2(pX_{t_{l+1}}+r)}{3}\right).$$

Proof. First, we notice the fact that $\max\{\deg_{t_{l+1}}(\tau) \colon 1 \leq \tau \leq s-1\} \leq X_{t_{l+1}}$ guarantees that $\max\{\deg_s(\tau) \colon 1 \leq \tau \leq s-1\} \leq X_{t_{l+1}}$. Therefore, $\deg_s(s)$ is stochastically dominated by $A_s \sim Bin\left(s, \frac{pX_{t_{l+1}}+r}{s}\right)$ so for any $\varepsilon \in (0, 1)$ we obtain the result directly using the Chernoff bound with $\mathbb{E}[A_s] = pX_{t_{l+1}} + r$.

Note that the result implies that with high probability at most slightly more than a p fraction of the maximum allowed degree is already used at time s. Therefore, we are interested in bounding the remaining part of the degree, i.e. $\deg_{t_{l+1}}(s) - \deg_s(s)$, by something smaller than the remaining $(1-p)$ fraction of the maximum allowed degree.

Lemma 6. *Let $\frac{1}{2} < p < 1$ and $s \in (t_l, t_{l+1}]$ for some $l = 0, 1, \ldots, k-1$. Then asymptotically as $t \to \infty$, for any constant $A > 0$ it holds that*

$$\Pr\left[\deg_{t_{l+1}}(s) \geq X_{t_{l+1}} | \mathcal{B}_l(t_l) \wedge \mathcal{B}_{l+1}(s)\right] = O(t^{-A}).$$

Lemma 7. *Let $\frac{1}{2} < p < 1$ and $s \in (t_l, t_{l+1}]$ for some $l = 0, 1, \ldots, k-1$. Then asymptotically as $t \to \infty$, for any constant $A > 0$ it holds that*

$$\Pr\left[\neg\mathcal{B}_{l+1}(t_{l+1}) | \mathcal{B}_l(t_l)\right] = O(t^{-A}).$$

The proofs of both lemmas above are presented in Appendices E and F.

Theorem 3. *Let $\frac{1}{2} < p < 1$. Then asymptotically as $t \to \infty$, for any constant $A > 0$ it holds that*

$$\Pr\left[\Delta(G_t) \geq (1+\varepsilon)t^p \log^{5-4p} t\right] = O(t^{-A}).$$

Proof. From Lemma 3 we know that $X_{t_k} \leq (1+\varepsilon)t^p \log^{5-4p} t$ holds quite surely. It follows that

$$\Pr\left[\Delta(G_t) \geq (1+\varepsilon)t^p \log^{5-4p} t\right] \leq \Pr\left[\Delta(G_t) \geq X_{t_k}\right] \leq \Pr\left[\neg\mathcal{B}_k(t_k)\right]$$

$$\leq \sum_{l=0}^{k-1} \Pr\left[\neg \mathcal{B}_{l+1}(t_{l+1}) | \mathcal{B}_l(t_l)\right] + \Pr\left[\neg \mathcal{B}_0(t_0)\right].$$

Now, from Theorem 2 and Lemma 7 we know that both $\Pr\left[\neg \mathcal{B}_0(t_0)\right] = O(t^{-A})$ and $\Pr[\neg \mathcal{B}_{l+1}(t_l) | \mathcal{B}_l(t_l)] = O(t^{-A})$ for any $A > 0$, respectively. Putting this all together with Lemma 4 we obtain the result.

3.3 Lower bound

Here we proceed analogously to the case of the upper bound for early vertices. First, we provide an appropriate Chernoff-type bound for the degree of a given vertex with respect to some deterministic sequence. Then we again use a special sequence, which has the desired rate of growth and serves as a lower bound on $\deg_t(s)$. Note that we don't need to extend our analysis for the late vertices since a lower bound for the degree of any vertex s at time t is also a lower bound for the minimum degree of G_t.

First, we note that if we start the whole process from a non-empty graph, then there exists $s \in [1, t_0]$ such that $\deg_{t_0}(s) \geq 1$. Moreover, even if the starting graph is empty, but $r > 0$, then with high probability there exists a vertex with positive degree, as the probability of adding another isolated vertex to an empty graph on t vertices is at most $(1 - \frac{r}{t})^t \leq \exp(-r)$, so within first $\frac{A}{r} \log t$ vertices for any $A > 0$ we have a non-isolated vertex with probability at least $1 - O(t^{-A})$. Of course, if we start from an empty graph and $r = 0$, then for any p there cannot arise any edge in the duplication process. However, in this case it trivially follows that $\Delta(G_t) = 0$, so we omit this case in further analysis.

Let us now return to the aforementioned Chernoff-type lower bound:

Lemma 8. *Let $1 \leq s \leq \tau \leq t$. Let X_τ be any value such that $\deg_\tau(s) \geq X_\tau$. Then for any $h \leq \varepsilon\tau$ with $\varepsilon \in \left(0, \frac{1}{3}\right)$ it is true that*

$$\Pr\left[\deg_{\tau+h}(s) \leq \deg_\tau(s) + (1 - 2\varepsilon)\,\frac{hpX_\tau}{\tau}\right] \leq \exp\left(-\frac{h\varepsilon^2 pX_\tau}{3\tau}\right).$$

Corollary 3. *Let $1 \leq s \leq \tau \leq t$. Let $X_\tau \geq 0$, $A > 0$, $\varepsilon \in \left(0, \frac{1}{3}\right)$ be values such that $\deg_\tau(s) \leq \tau$ and $3A \log t \leq \varepsilon^3 pX_\tau$. Then for any $h \in \left[\frac{3A\log t}{\varepsilon^2 pX_\tau}, \varepsilon\tau\right]$ it is true that*

$$\Pr\left[\deg_{\tau+h}(s) \leq \deg_\tau(s) + (1 - 2\varepsilon)\,\frac{hpX_\tau}{\tau}\right] = O(t^{-A}).$$

In the following, we again use sequences $(t_i)_{i=1}^k$ and $(X_{t_i})_{i=1}^k$ from Definition 1. Let us also define $\mathcal{C}_i(s)$ as the event that $\deg_{t_i}(s) \geq X_{t_i} - \phi + 1$ for a fixed $s \leq t_i$. This allows us to proceed with the main theorem of this section:

Theorem 4. *For $G_t \sim \mathtt{DD}(t, p, r)$ with $\frac{1}{2} < p < 1$ there exists s such that for any constants $\varepsilon > 0$ and $A > 0$ it holds asymptotically that*

$$\Pr\left[\deg_t(s) < (1 - \varepsilon)t^p\right] = O(t^{-A}).$$

Proof. Again let us use sequences $(t_i)_{i=0}^k$ and $(X_{t_i})_{i=0}^k$ with $\alpha = 12p^3(A+1)\log^2 t$, $\beta_i = \alpha p - \frac{\alpha}{\log t_i}$ and $\phi = 144(A+1)^2 \log^4 t$. These parameters satisfy the assumptions of Lemma 3 and Corollary 2.

Moreover, if $\mathcal{C}_i(s)$ holds, then the assumptions of Corollary 3 are also true for $\tau = t_i$ and $h = \frac{\alpha t_i \log t_i}{X_{t_i}}$ as $t_i \to \infty$, since for any constant $A > 0$ and $\varepsilon = \frac{1}{2p \log t_i}$

$$\frac{3A\tau \log t}{\varepsilon^2 p X_{t_i}} < h = \frac{\alpha t_i \log t_i}{X_{t_i}} < \varepsilon t_i.$$

The left inequality is easy to verify as the left hand side is $\Theta\left(\frac{t_i \log^2 t_i \log t}{X_{t_i}}\right)$ and h grows like $\Theta\left(\frac{t_i \log t_i \log^2 t}{X_{t_i}}\right)$. The right inequality follows directly from Lemma 2.

Next, $X_{t_{i+1}} - X_{t_i} = \beta_i \log t_i = (1-2\varepsilon)\frac{hpX_{t_i}}{t_i}$, where $h = \frac{1}{1-\frac{1}{p\log t_i}} \frac{\beta_i t_i \log t_i}{pX_{t_i}}$. Therefore, Corollary 3 implies that for any constant $A > 0$ and $\varepsilon = \frac{1}{2p\log t_i}$ it is true that $\Pr[\neg\mathcal{C}_{i+1}(s)|\mathcal{C}_i(s)] = O(t^{-A})$. Note that we apply this with a sequence $X_{t_i} - \phi + 1$, not with X_{t_i} itself this time. This is so because to use Corollary 3 we need $\deg_{t_0}(s) \geq X_{t_0} - \phi + 1 = 1$, which holds with high probability – as e.g. $\deg_{t_0}(s) \geq X_{t_0}$ is false with high probability.

Since $X_{t_0} = 144(A+1)^2 \log^4 t$ we know that $\mathcal{C}_0(s)$ holds with high probability: either the starting graph is nonempty, or $r > 0$ and some edges appear before t_0. Using Lemma 4 and Corollary 3 for any $\varepsilon > 0$ and $A > 0$ we get

$$\begin{aligned}\Pr[\deg_t(s) <&(1-\varepsilon)t^p] \leq \Pr[\deg_t(s) < X_{t_{k-1}} - \phi + 1] \leq \Pr[\neg\mathcal{C}_{k-1}(s)] \\ &\leq \sum_{i=0}^{k-2} \Pr[\neg\mathcal{C}_{i+1}(s)|\mathcal{C}_i(s)] + \Pr[\neg\mathcal{C}_0(s)] = \sum_{i=0}^{k-1} O(t^{-A}) = O(t^{-A+1}).\end{aligned}$$

We conclude our analysis with the following corollary.

Corollary 4. *For $G_t \sim \mathtt{DD}(t,p,r)$ with $\frac{1}{2} < p < 1$ for any constants $\varepsilon > 0$ and $A > 0$ it holds asymptotically that*

$$\Pr\left[\Delta(G_t) \leq (1-\varepsilon)t^p\right] = O(t^{-A}).$$

References

1. Barabási, A.L., Albert, R.: Emergence of scaling in random networks. Science **286**(5439), 509–512 (1999)
2. Bollobás, B.: Random graphs. Cambridge University Press (2001)
3. Bollobás, B., Riordan, O., Spencer, J., Tusnády, G.: The degree sequence of a scale-free random graph process. In: The Structure and Dynamics of Networks, pp. 384–395. Princeton University Press (2011)
4. Chierichetti, F., Kumar, R., Lattanzi, S., Panconesi, A., Raghavan, P.: Models for the compressible web. SIAM Journal on Computing **42**(5), 1777–1802 (2013)
5. Colak, R., Hormozdiari, F., Moser, F., Schönhuth, A., Holman, J., Ester, M., Sahinalp, S.C.: Dense graphlet statistics of protein interaction and random networks. In: Biocomputing 2009, pp. 178–189. World Scientific Publishing, Singapore (2009)

6. Cooper, C., Frieze, A.: The cover time of the preferential attachment graph. Journal of Combinatorial Theory, Series B **97**(2), 269–290 (2007)
7. Diestel, R.: Graph Theory. Springer (2005)
8. Erdős, P., Rényi, A.: On random graphs I. Publicationes Mathematicae **6**, 290–297 (1959)
9. Frieze, A., Karoński, M.: Introduction to Random Graphs. Cambridge University Press (2016)
10. Frieze, A., Turowski, K., Szpankowski, W.: Degree distribution for duplication-divergence graphs: Large deviations. In: Adler, I., Müller, H. (eds.) Graph-Theoretic Concepts in Computer Science - 46th International Workshop, WG 2020, Leeds, UK, June 24-26, 2020, Revised Selected Papers. Lecture Notes in Computer Science, vol. 12301, pp. 226–237. Springer (2020)
11. Hermann, F., Pfaffelhuber, P.: Large-scale behavior of the partial duplication random graph. ALEA **13**, 687–710 (2016)
12. Jacquet, P., Turowski, K., Szpankowski, W.: Power-law degree distribution in the connected component of a duplication graph. In: Drmota, M., Heuberger, C. (eds.) 31st International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms, AofA 2020, June 15-19, 2020, Klagenfurt, Austria (Virtual Conference). LIPIcs, vol. 159, pp. 16:1–16:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik (2020)
13. Janson, S., Łuczak, T., Ruciński, A.: Random graphs. John Wiley & Sons (2011)
14. Jordan, J.: The connected component of the partial duplication graph. ALEA – Latin American Journal of Probability and Mathematical Statistics **15**, 1431–1445 (2018)
15. Knuth, D., Motwani, R., Pittel, B.: Stable husbands. Random Structures & Algorithms **1**(1), 1–14 (1990)
16. Li, S., Choi, K.P., Wu, T.: Degree distribution of large networks generated by the partial duplication model. Theoretical Computer Science **476**, 94–108 (2013)
17. Łuczak, T., Magner, A., Szpankowski, W.: Asymmetry and structural information in preferential attachment graphs. Random Structures & Algorithms **55**(3), 696–718 (2019)
18. Łuczak, T., Magner, A., Szpankowski, W.: Compression of preferential attachment graphs. In: IEEE International Symposium on Information Theory, ISIT 2019, Paris, France, July 7-12, 2019. pp. 1697–1701. IEEE (2019)
19. Newman, M.: Networks: An Introduction. Oxford University Press (2010)
20. Shao, M., Yang, Y., Guan, J., Zhou, S.: Choosing appropriate models for protein–protein interaction networks: a comparison study. Briefings in Bioinformatics **15**(5), 823–838 (2013)
21. Solé, R., Pastor-Satorras, R., Smith, E., Kepler, T.: A model of large-scale proteome evolution. Advances in Complex Systems **5**(01), 43–54 (2002)
22. Sreedharan, J., Turowski, K., Szpankowski, W.: Revisiting parameter estimation in biological networks: Influence of symmetries. IEEE/ACM Transactions on Computational Biology and Bioinformatics (2020)
23. Turowski, K., Szpankowski, W.: Towards degree distribution of a duplication-divergence graph model. Electronic Journal of Combinatorics **28**(1), P1.18 (2021)
24. Van Der Hofstad, R.: Random graphs and complex networks. Cambridge University Press (2016)
25. Watts, D., Strogatz, S.: Collective dynamics of "small-world" networks. Nature **393**(6684), 440–442 (1998)

A Proof of Lemma 1

First, recall that for $i = 0, 1, \dots, h-1$ we have $\deg_{\tau+i+1}(s) = \deg_{\tau+i}(s) + I_{\tau+i}$ where $I_{\tau+i} \sim Be\left(\frac{p \deg_{\tau+i}(s)+r}{\tau+i}\right)$. Also clearly $\deg_{\tau+i}(s) \le \deg_{\tau}(s) + i$ for any $i = 0, 1, \dots, h$, so we have

$$\frac{\deg_{\tau+i}(s)}{\tau+i} \le \frac{\deg_{\tau}(s)+i}{\tau} \le \left(1 + \frac{i}{X_\tau}\right)\frac{X_\tau}{\tau} \le \left(1 + \frac{h}{X_\tau}\right)\frac{X_\tau}{\tau} \le (1+\varepsilon)\frac{X_\tau}{\tau}.$$

Therefore for any $i = 0, 1, \dots, h-1$ we know that $I_{\tau+i}$ is stochastically dominated by $I^*_{\tau+i} \sim Be\left((1+\varepsilon)\frac{pX_\tau+r}{\tau}\right)$.

Now, from the well known Chernoff bounds we know that for any $\varepsilon \in (0,1)$

$$\Pr\left[\deg_{\tau+h}(s) - \deg_{\tau}(s) \ge (1+\varepsilon)\,\mathbb{E}\left[\sum_{i=0}^{h-1} I^*_{\tau+i}\right]\right] \le \exp\left(-\frac{\varepsilon^2}{3}\mathbb{E}\left[\sum_{i=0}^{h-1} I^*_{\tau+i}\right]\right)$$

and therefore

$$\begin{aligned}
&\Pr\left[\deg_{\tau+h}(s) \ge \deg_{\tau}(s) + (1+3\varepsilon)\,\frac{h(pX_\tau+r)}{\tau}\right] \\
&\le \Pr\left[\deg_{\tau+h}(s) \ge \deg_{\tau}(s) + (1+\varepsilon)^2\,\frac{h(pX_\tau+r)}{\tau}\right] \\
&\le \exp\left(-\frac{h\varepsilon^2(1+\varepsilon)(pX_\tau+r)}{3\tau}\right).
\end{aligned}$$

This completes the proof.

B Proof of Lemma 2

Let us define $Y_\tau = \tau^{p-\delta}$. By definition we know that $X_{t_0} = \phi \ge Y_{t_0}$.

Now, let us assume that $X_{t_i} \ge Y_{t_i}$ holds for some $i \ge 0$. Let us also denote by $h = t_{i+1} - t_i = \frac{\alpha\, t_i \log t_i}{X_{t_i}}$. Then we have asymptotically

$$Y_{t_{i+1}} - Y_{t_i} = (t_i+h)^{p-\delta} - t_i^{p-\delta} = t_i^{p-\delta}\left(\left(1+\frac{h}{t_i}\right)^{p-\delta} - 1\right) \le t_i^{p-\delta}\frac{(p-\delta)h}{t_i},$$

for any $\delta \in [0,p)$, because $X_{t_i} \ge \phi \ge \log^2 t$, so $\frac{h}{t_i} = \frac{\alpha \log t_i}{X_{t_i}} \le \frac{\alpha \log t_i}{\phi} \le \frac{\log t}{\sqrt{\phi}} \le 1$. Thus,

$$Y_{t_i}\frac{(p-\delta)\,h}{t_i} \le X_{t_i}\frac{(p-\delta)\,h}{t_i} = \alpha(p-\delta)\log t_i \le \beta_i \log t_i = X_{t_{i+1}} - X_{t_i},$$

so clearly $X_{t_{i+1}} \ge Y_{t_{i+1}}$ holds as well, which completes the inductive step.

C Proof of Lemma 3

We again proceed by induction. Clearly, $X_{t_0} = t_0 \le t_0 \log t_0$.

Directly from the definition we get

$$\begin{aligned}
&\phi^{1-p} t_{i+1}^p \log t_{i+1} - X_{t_{i+1}} = \phi^{1-p} t_{i+1}^p \log t_{i+1} - X_{t_i} - \beta_i \log t_i \\
&\quad \ge \phi^{1-p} t_{i+1}^p \log t_{i+1} - \phi^{1-p} t_i^p \log t_i - \beta_i \log t_i \\
&\quad = \phi^{1-p} t_i^p \log t_i \left(\left(1 + \frac{\alpha \log t_i}{X_{t_i}}\right)^p \left(1 + \frac{\log(1 + \alpha \log t_i / X_{t_i})}{\log t_i}\right) - 1 \right) - \beta_i \log t_i.
\end{aligned}$$

Now we use the inequalities $(1+x)^p \ge 1 + px - \frac{p(1-p)x^2}{2} + O(x^3)$ and $\log(1+x) \ge x - O(x^2)$, true for any $p \in [0,1]$ and any $x \to 0$. In particular, in our case $x = \frac{\alpha \log t_i}{X_{t_i}} \le \frac{1}{\sqrt{\log t}}$ since $\alpha \le \sqrt{\phi}$ and $X_{t_i} \ge \phi \ge \log^3 t$. Therefore

$$\begin{aligned}
&\phi^{1-p} t_{i+1}^p \log t_{i+1} - X_{t_{i+1}} \\
&\quad \ge \phi^{1-p} t_i^p \log t_i \left(\frac{\alpha\, p \log t_i}{X_{t_i}} + \frac{\alpha}{X_{t_i}}(1 - o(1)) - \frac{\alpha^2\, p(1-p) \log^2 t_i}{2X_{t_i}^2}(1 - o(1)) \right) \\
&\qquad - \beta_i \log t_i \\
&\quad \ge \alpha \log t_i \left(p + \frac{1}{\log t_i}(1 - o(1)) - \frac{p(1-p)\log t_i}{2\sqrt{t_i^{p-\delta}}}(1 - o(1)) \right) - \beta_i \log t_i,
\end{aligned}$$

where in the last line we used $X_{t_i} \ge \sqrt{\phi t_i^{p-\delta}} \ge \alpha \sqrt{t_i^{p-\delta}}$ – derived as the geometric mean between the bounds from Definition 1 and Lemma 2.

Finally, we note the assumption $\beta_i \le \alpha p + \frac{\alpha}{2 \log t_i}$ ensures that for sufficiently large t last expression is clearly non-negative, which completes the proof.

D Proof of Lemma 4

Clearly from the previous lemmas we know that for any constant $\varepsilon > 0$ it is true that

$$\frac{t_k}{t_{k-1}} = 1 + \frac{\alpha \log t_{k-1}}{X_{t_{k-1}}} \le 1 + \frac{\alpha \log t_{k-1}}{\sqrt{\phi t_{k-1}^{p-\delta}}} \in (1, 1+\varepsilon).$$

The first claim follows from this and from the fact that $t_{k-1} < t \le t_k$.

Similarly, for any constant $\varepsilon > 0$ the second claim follows from the fact that $X_{t_{k-1}} < X_t \le X_{t_k}$ and that

$$\frac{X_{t_k}}{X_{t_{k-1}}} = 1 + \frac{\beta_k \log t_k}{X_{t_{k-1}}} \le 1 + \frac{\alpha \log t_{k-1}(p+\varepsilon)}{\sqrt{\phi t_{k-1}^{p-\delta}}} \in (1, 1+\varepsilon)$$

which completes the proof.

E Proof of Lemma 6

Let us denote $d = \frac{1-p}{2} X_{t_{l+1}} - \frac{(1+p)r}{2p}$. If $s \in [t_{l+1} - d, t_{l+1}]$, then the result is directly implied by Lemma 5 with $\varepsilon = \frac{1-p}{2p}$, as the degree of the vertex during an interval of length d cannot grow more than d.

Otherwise $s \in (t_l, t_{l+1} - d)$. But if such an s exists, then it is the case that $d \le t_{l+1} - t_l = \frac{C t_l \log t_l \log^2 t}{X_{t_l}}$ for some constant $C > 0$ so from Lemma 2 with $\delta = 0$ and by the fact that $X_{t_i} \ge \phi$ we get that asymptotically $X_{t_l} \ge t_l^{\gamma p} \log^{4(1-\gamma)} t$ for any $\gamma \in [0, 1]$ and therefore

$$
\begin{aligned}
C t_l \log t_l \log^2 t &\ge \left(\frac{1-p}{2} X_{t_{l+1}} - \frac{(1+p)r}{2p} \right) X_{t_l} \\
&\ge \frac{1-p}{4} X_{t_l}^2 \ge \frac{1-p}{4} t_l^{2\gamma p} \log^{8(1-\gamma)} t.
\end{aligned}
$$

However, if we set e.g. $\gamma = \frac{3}{5}$, then we can bound the right side from below by $\frac{1-p}{4} t_l^{6/5} \log^{16/5} t$ – and for sufficiently large t we obtain a contradiction, as each term on the right side asymptotically dominates the respective one on the left side.

F Proof of Lemma 7

Let l be the first value for which the theorem does not hold. Then, from Lemma 6 we get that for any constant $A > 0$ it holds that

$$
\begin{aligned}
\Pr\left[\neg \mathcal{B}_{l+1}(t_{l+1}) | \mathcal{B}_l(t_l) \wedge \mathcal{B}_{l+1}(t_l)\right] &= \sum_{s=t_l}^{t_{l+1}-1} \Pr\left[\neg \mathcal{B}_{l+1}(s+1) | \mathcal{B}_l(t_l) \wedge \mathcal{B}_{l+1}(s)\right] \\
= \sum_{s=t_l}^{t_{l+1}-1} \Pr\left[\neg \mathcal{A}_{l+1}(s+1) | \mathcal{B}_l(t_l) \wedge \mathcal{B}_{l+1}(s)\right] &= O(t^{-A}).
\end{aligned}
$$

From Theorem 2 we know that $\Pr[\mathcal{B}_0(t_0)] = 1 - O(t^{-A})$. Recall that by our assumption $\Pr[\neg \mathcal{B}_{i+1}(t_{i+1}) | \mathcal{B}_i(t_i)] = 1 - O(t^{-A})$ for all $i = 0, 1, \ldots, l-1$, so it follows that $\Pr[\mathcal{B}_i(t_i)] = 1 - O(t^{-A})$ for all $i = 0, 1, \ldots, l$. We use this fact, the observation that $\mathcal{B}_l(t_l) \subseteq \mathcal{A}_l(s)$ and Theorem 2 to get

$$
\begin{aligned}
\Pr\left[\neg \mathcal{B}_{l+1}(t_l) | \mathcal{B}_l(t_l)\right] &\le \sum_{s=1}^{t_l} \Pr\left[\neg \mathcal{A}_{l+1}(s) | \mathcal{B}_l(t_l)\right] \le \sum_{s=1}^{t_l} \frac{\Pr\left[\neg \mathcal{A}_{l+1}(s) \wedge \mathcal{B}_l(t_l)\right]}{\Pr\left[\mathcal{B}_l(t_l)\right]} \\
\le \sum_{s=1}^{t_l} \frac{\Pr\left[\neg \mathcal{A}_{l+1}(s) \wedge \mathcal{A}_l(s)\right]}{\Pr\left[\mathcal{B}_l(t_l)\right]} &\le \sum_{s=1}^{t_l} \frac{\Pr\left[\neg \mathcal{A}_{l+1}(s) | \mathcal{A}_l(s)\right]}{\Pr\left[\mathcal{B}_l(t_l)\right]} = O(t^{-A}).
\end{aligned}
$$

Finally, for any events E_1, E_2, E_3 we have $\Pr[\neg E_1 | E_2] \le \Pr[\neg E_1 | E_3 \wedge E_2] + \Pr[\neg E_3 | E_2]$. We substitute $E_1 = \mathcal{B}_{l+1}(t_{l+1})$, $E_2 = \mathcal{B}_l(t_l)$ and $E_3 = \mathcal{B}_{l+1}(t_l)$ to obtain the final result.

G Proof of Lemma 8

Let us recall that for $i = 0, 1, \ldots, h-1$ we have $\deg_{\tau+i+1}(s) = \deg_{\tau+i}(s) + I_{\tau+i}$ where $I_{\tau+i} \sim Be\left(\frac{p\deg_{\tau+i}(s)+r}{\tau+i}\right)$. Also clearly $\deg_{\tau+i}(s) \geq \deg_{\tau}(s)$ for any $i = 0, 1, \ldots, h$, so we have

$$\frac{\deg_{\tau+i}(s)}{\tau+i} \geq \frac{\deg_{\tau}(s)}{\tau+h} \geq \frac{X_\tau}{\tau(1+\varepsilon)} \geq (1-\varepsilon)\frac{X_\tau}{\tau}.$$

Therefore for any $i = 0, 1, \ldots, h-1$ we know that $I_{\tau+i}$ stochastically dominates $I^*_{\tau+i} \sim Be\left((1-\varepsilon)\frac{pX_\tau}{\tau}\right)$.

Now, from the Chernoff bounds we know that for any $\varepsilon \in (0,1)$

$$\Pr\left[\deg_{\tau+h}(s) - \deg_{\tau}(s) \leq (1-\varepsilon)\,\mathbb{E}\left[\sum_{i=0}^{h-1} I^*_{\tau+i}\right]\right] \leq \exp\left(-\frac{\varepsilon^2}{2}\mathbb{E}\left[\sum_{i=0}^{h-1} I^*_{\tau+i}\right]\right)$$

and therefore

$$\begin{aligned}
&\Pr\left[\deg_{\tau+h}(s) \leq \deg_{\tau}(s) + (1-2\varepsilon)\,\frac{hpX_\tau}{\tau}\right] \\
&\quad \leq \Pr\left[\deg_{\tau+h}(s) \leq \deg_{\tau}(s) + (1-\varepsilon)^2\,\frac{hpX_\tau}{\tau}\right] \leq \exp\left(-\frac{h\varepsilon^2(1-\varepsilon)pX_\tau}{2\tau}\right).
\end{aligned}$$

Finally, is is sufficient to see that if $\varepsilon < \frac{1}{3}$, then we can replace $\frac{1-\varepsilon}{2}$ by $\frac{1}{3}$ in the last formula, which completes the proof.