ON THE CONCENTRATION OF THE MAXIMUM DEGREE
IN THE DUPLICATION-DIVERGENCE MODELS

ALAN FRIEZE†, KRZYSZTOF TUROWSKI‡, AND WOJCIECH SZPANKOWSKI§

Abstract. We present a rigorous and precise analysis of the maximum degree and the average degree in a dynamic duplication-divergence graph model introduced by Solé, Pastor-Satorras et al. in which the graph grows according to a duplication-divergence mechanism, i.e. by iteratively creating a copy of some node and then randomly alternating the neighborhood of a new node with probability $p$. This model captures the growth of some real-world processes e.g. biological or social networks. In this paper, we prove that for some $0 < p < 1$ the maximum degree and the average degree of a duplication-divergence graph on $t$ vertices are asymptotically concentrated with high probability around $t^p$ and $\max\{t^{2p-1}, 1\}$, respectively, i.e. they are within at most a polylogarithmic factor from these values with probability at least $1 - t^{-A}$ for any constant $A > 0$.

Key words. random graphs, dynamic graphs, duplication-divergence model, degree distribution, maximum degree, average degree, large deviation

AMS subject classifications. 05C07, 05C80, 68R10

1. Introduction. Studying properties of random graphs is a popular topic of research in computer science and discrete mathematics since the seminal work of Paul Erdős and Alfréd Rényi [8]. This model was studied extensively using various probabilistic and analytic methods. The research mostly concentrated on a few broad topics: distribution of structural properties of graphs (e.g. the number of edges, degrees of fixed vertex, maximum degree, diameter), the existence of special subgraphs (e.g. motif counting, longest paths, maximum matching, Hamilton cycles), values of well-known combinatorial parameters (e.g. largest independent set, chromatic number), or extremal properties (Ramsey- and Turán-type) – see e.g. surveys of results in [2, 10, 17, 31]. The widening array of application domains ranging from biology to finance to social science inspired further directions of research: first, there appeared an idea to bring the models to the real-world data and to study important aspects, such as centrality, degree correlation, community detection, or graph compression [19, 20, 24]. Second, more models of random networks were developed e.g. for inhomogeneous random graphs, geometric random graphs, preferential attachment graphs, or duplication graphs [4, 10, 31]. Often, these models were inspired by some generation mechanisms (e.g. rich-get-richer), or properties (e.g. scale-free/power-law property) that were
claimed at work for the real-world networks [9].

In particular, since the late 1990s attention turned toward dynamics graphs in which the behavior of networks evolves in time, e.g. when sets of vertices and/or edges are functions of time, which is definitely the case for certain biological (e.g. protein-protein networks) and social networks (e.g. graph of citations).

One of the family of such networks is the so-called duplication models [5, 4]. It was observed that the evolutionary dynamics of protein interaction networks can be described by simple duplication and mutation rules [25, 32]. For example, the main mechanisms in such models are duplication and divergence: when vertices arrive one by one, they are created as copies of some already existing node, chosen uniformly at random (duplication), and then the neighborhood is typically altered randomly according to some predefined rules (divergence).

In this paper we study a particular duplication-divergence model, first introduced by Solé, Pastor-Satorras et al. [28]. This model is a promising object of inquiry since it has been shown empirically that its degree distribution, small subgraph (graphlets) counts, and the number of symmetries fit very well the structure of some real-world biological and social networks, e.g. protein-protein and citation networks. More precisely, there exist heuristics to infer the underlying parameters of the model from various biological networks, which enable us to generate similar graphs in terms of degree distributions, k-hop reachability, closeness, betweenness, and graphlet frequency [15, 22] (see also an alternative method of parameter estimation in [29]). It also turns out that this model often outperformed alternative ones in terms of systematically replicating the degree distribution, small subgraph (graphlets) counts, and symmetries of the input networks [6, 27, 29]. 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 precise results related to the behavior of the degree distribution of the graphs generated by this model. Our contribution is a step towards closing this gap. In short, we prove an asymptotic tight concentration of two parameters in duplication-divergence graphs: maximum degree, and average degree (or, equivalently, the number of edges) around their mean values.

The paper is organized as follows: in Section 2 we define formally the duplication-divergence model, and we present an overview of the previous results related to the properties of the degree distribution. Then, in Section 3 we introduce our result for the maximum degree, with proof split into three parts: in Subsection 3.1 and Subsection 3.2 we prove upper bounds for the degrees of the earliest and later vertices arriving in the graph, respectively, and in Subsection 3.3 we give a proof of the lower bound for the degree of the first vertices, which is effectively also the lower degree of the maximum degree. Next, we proceed with Section 4, containing the proofs of the upper and the lower bounds for the average degree (or, equivalently, the total number of edges in the graph), respectively. Finally, we offer some further problems and hypotheses that stem from our current research.

This work is a substantial extension of two conference papers: the one presented at COCOON 2021 [12] which contained the weaker concentration results for maximum degree only for the case $\frac{1}{2} < p < 1$, and the one presented at WG 2020 [11] which contained the weaker claims (proved using different methods) for average degree and for degrees only of the earliest vertices.
2. Model definition and earlier work. Throughout the paper we use standard graph notation from [7], e.g. $V(G)$ denotes the vertex set of a graph $G$, $\deg_G(s)$ is the degree of node $s$ in $G$, and we write $\Delta(G)$ and $D(G)$ for the maximum degree and the average degree in $G$. Let also $N_G(s)$ denote the open neighborhood of $s$ in $G$.

All graphs considered in the paper are simple, i.e. without loops or multiple edges. Additionally, since we are eventually dealing with a probability space over graphs on $t$ vertices, let $G_t$ denote a random variable representing a graph on $t$ vertices. Finally, since we are dealing with graphs growing sequentially, we assume that the vertices are identified with the natural numbers according to their arrival time. For simplicity, we introduce the notation $\deg_t(s)$ for the random variable denoting the degree of vertex $s$ in $G_t$. Clearly, $\Delta(G_t)$ and $D(G_t)$ are random variables denoting the maximum degree and the average degree in $G_t$.

Let us now formally define the duplication-divergence model, denoted $\text{DD}(t, p, r)$, introduced by Solé et al. [28, 26]. Let $G_{t_0}$ be some graph on $t_0 \leq t$ vertices, with vertices having distinct labels from 1 to $t_0$. Now, for every $i = t_0, t_0 + 1, \ldots, t - 1$ we create $G_{i+1}$ from $G_i$ according to the following rules:

1. we add a new vertex with label $i + 1$ to the graph,
2. we choose a vertex $u$ from $G_i$ uniformly at random – and we denote $u$ as parent$(i + 1)$,
3. for every vertex $v$:
   (a) if $v$ is adjacent to $u$ ($v \in N_{G_i}(u)$) in $G_i$, then add an edge between $v$ and $i + 1$ with probability $p$,
   (b) if $v$ is not adjacent to $u$ in $G_i$ ($v \notin N_{G_i}(u)$), then add an edge between $v$ and $i + 1$ with probability $r_i$. Note that this case also occurs when $v = u$, since $u \notin N_{G_i}(u)$.

All edge additions are independent Bernoulli random variables.
Since both $p$ and $\frac{r}{t}$ for $i = t_0, \ldots, t - 1$ are probabilities, we allow the parameter space to be $p \in [0, 1]$ and $r \in [0, t_0]$.

There is indeed a duplication-divergence mechanism at work since we can think of the equivalent set of rules in the form “copy a vertex from $G_t$ uniformly at random”, “remove its neighbors independently at random with probability $1 - p^r$”, and “add edges to all other vertices independently at random with probability $\frac{r}{t}$”.

Throughout the paper we will refer to the standard Big-O Landau notation, as popularized e.g. in [13]. Let us recall its basic notion: $f(n) = O(g(n))$ for some functions $f$ and $g$ such that $\exists k > 0 \exists n_0 \forall n > n_0 \ |f(n)| \leq |k \cdot g(n)|$. Additionally, we will use

- $f(n) = \Omega(g(n))$ when $g(n) = O(f(n))$,
- $f(n) = \Theta(g(n))$ when both $f(n) = O(g(n))$ and $g(n) = O(f(n))$,
- $f(n) = o(g(n))$ when $f(n) = O(g(n))$ but not $f(n) = O(g(n))$.

Intuitively, $f(n) = \Omega(g(n))$ when $\lim_{n \to \infty} \frac{f(n)}{g(n)} \in [k_1, k_2]$ for some $0 < k_1 < k_2$. Since in the model both $p$ and $r$ (and the order of initial graph $G_{t_0}$) are constants, the asymptotic results are given exclusively in terms of $t$.

As it was mentioned earlier, there are only a few rigorous results for the $\mathcal{DD}(t, p, r)$ model and its special cases. For $0 < p < 1$ and $r = 0$, it was proved in [14] that asymptotically there exists a phase transition for the limiting distribution of degree frequencies: if $p \leq p^*$, then almost all vertices are isolated, i.e. the number of non-isolated vertices in $G_t$ is $o(t)$, and if $p > p^*$, then only a constant fraction of vertices (with an explicit constant) are isolated. Moreover, it was proved that for any $k$ the fraction of vertices of degree $k$ in $G_t$ converges to 0, and therefore there is no limiting degree distribution for $p > p^*$. From [21] it is known that the number of vertices of degree one in $G_t$ 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.

However, also for the same case in [18, 16] it was shown for $p < \exp(-1)$ that the (only) connected component in $G_t$ exhibits a power-law property with the scale parameter $\gamma$ which is the solution of $3 = \gamma + p^{\gamma - 2}$.

For the general case, the two main parameters under consideration were the degree of fixed vertices $\deg_t(s)$ and the average degree of $G_t$ defined as

$$D(G_t) = \frac{1}{t} \sum_{s=1}^{t} \deg_t(s).$$

It was shown in [30] that we can solve the recurrence equation for the expected average degree and obtain

**Theorem 2.1.** For $t \to \infty$ it holds that

$$\mathbb{E}[D(G_t)] = \begin{cases} \Theta(1) & \text{if } p < \frac{1}{2} \text{ and } r > 0, \\
\Theta(\ln t) & \text{if } p = \frac{1}{2} \text{ and } r > 0, \\
\Theta(t^{2p-1}) & \text{otherwise}. \end{cases}$$

In a similar fashion it was shown that the expected degree of a vertex $s$ is given by the following theorem:
Theorem 2.2. For $t \to \infty$, it holds that

\[ E[\deg_t(s)] = \begin{cases} 
\Theta\left(\log\left(\frac{1}{p}\right)\right) & \text{if } p = 0 \text{ and } r > 0, \\
\Theta\left(\left(\frac{1}{p}\right)^p\right) & \text{if } 0 < p < \frac{1}{2} \text{ and } r > 0, \\
\Theta\left(\sqrt{\frac{1}{2p}\log s}\right) & \text{if } p = \frac{1}{2} \text{ and } r > 0, \\
\Theta\left(\left(\frac{1}{r}\right)^{2p^2}\right) & \text{otherwise.}
\end{cases} \]

Clearly, the latter result for the earliest vertices implies that the expected maximum degree is $\Omega(t^p)$ for all $0 < p < 1$.

In fact, in [30] the authors obtained more than just Theorem 2.1 and Theorem 2.2, because they derived the exact formulae for both $E[D(G_t)]$ and $E[\deg_t(s)]$ with their very convoluted leading coefficients (depending on $s$, $p$, $r$) together with the asymptotics for $\text{Var}[D(G_t)]$ and $\text{Var}[\deg_t(s)]$.

The natural question then is to show that these random variables are concentrated, i.e. whether by moving only some small (e.g. polylogarithmic) factor from the mean we could observe the polynomial tail decay. Intuitively, for the later vertices we should not expect such a phenomenon: since the parent of a new vertex is drawn uniformly, and there are two binomial processes on top of it, we expect the degree distribution of $\deg_t(t)$ rather reflect the whole degree distribution, which for some cases we know (and for all other we stipulate, based on simulations) is not concentrated. However, as we will see in the next sections for the maximum degree and the average degree we can answer this question in the affirmative.

3. Maximum degree. In this section we present our main result concerning the concentration of the maximum degree $\Delta(G_t)$. We formulate it in the next theorem.

Theorem 3.1. Let $0 < p < 1$. Asymptotically for $G_t \sim DD(t, p, r)$

\[ \Pr[(1 - \alpha)t^p \leq \Delta(G_t) \leq (1 + \alpha)t^p \log^{2-p^2}(t)] = 1 - O(t^{-A}) \]

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

We prove separately a lower bound and a matching (within a polylogarithmic factor) upper bound. The main idea of the upper bound proof, presented in the next subsection, is as follows: we first in Definition 3.2 introduce auxiliary deterministic sequences $(t_i)_{i=0}^{k}$ and $(X_i)_{i=0}^{k}$ such that $t_0 < \ldots < t_{k-1} < t \leq t_k$. Although at first glance the dependency between $(t_i)_{i=0}^{k}$ and $(X_i)_{i=0}^{k}$ given in this definition could seem very convoluted, the intuition behind it is very simple: by doing this we can prove with little effort that $X_t$ grows close to $t^p$, provided that we choose the right parameters. Indeed, we show that $X_t \leq (1 + \alpha)t^p \log^{2-p^2}(t)$ for any constant $\alpha > 0$.

This way, we want $(X_i)_{i=0}^{k}$ to be a good (i.e. holding with high probability) upper bound for $\deg_t(s)$ for all $i = 0, \ldots, k$ and all $s \leq t_0$ (denoted as early vertices), which in turn should give us a similar lower bound $\deg_t(s)$ in terms of $X_t$ whp. We proceed in two major steps: first, by construction, we have $\deg_{t_0}(s) \leq t_0 = X_{t_0}$, and second, we prove a bound on $\deg_{t_{k+1}}(s) - \deg_{t_k}(s)$ that ensures it does not exceed $X_{t_{k+1}} - X_{t_k}$ with high probability. The latter part is achieved by providing an adequate upper bounding of $\deg_{t_{k+1}}(s) - \deg_{t_k}(s)$ by a sum of independent Bernoulli variables, so the Chernoff bound can be employed – and by applying a telescoping sum we establish that $\deg_{s}(s) \leq X_t$ with high probability for all $s \leq t_0$. Therefore, we find for early vertices $s$ (i.e. $s \leq t_0$) a Chernoff-type bound on the growth of $\deg_{s}(s)$ over an interval of certain length $h$.

This manuscript is for review purposes only.
The second part of the proof of our upper bound on the maximum degree is inductive: we prove that with high probability for any vertex \( s \in (t_i, t_{i+1}] \) it holds that \( \deg_t(s) \leq \max_{\tau \leq t_i} \{ \deg_t(\tau) \} \), that is, the later vertices (that is, for any \( s > t_0 \)) can have maximum degree only with a negligible probability. This proof can also be decomposed into three steps: first, we show that a vertex \( s \) on its arrival cannot have a degree greater than \((1 + \varepsilon)(pX_t + \tau)\) with high probability, and then it cannot increase between time \( s \) and \( t_{i+1} \) to exceed \( X_{t_{i+1}} \). Finally, to proceed from \( \deg_{t_i+1}(s) \leq X_{t_{i+1}} \) whp to \( \deg_t(s) \leq X_t \) whp we use exactly the same Chernoff bound as for early vertices.

To prove the lower bound we follow the steps from the upper bound for the early vertices: we show a respective lower Chernoff-type bound on the growth of \( \deg_t(s) \) over an interval of certain length \( h \) and we combine it with different (but very similar) sequences \( t_i \) and \( X_{t_i} \), thus proving that in this case \( \deg_t(s) \geq X_\tau - \ln^{1+\varepsilon}(\tau) + 1 \) with high probability for all early vertices (that is, \( s \leq t_0 \)), and that \( X_i \geq (1 - \alpha)t_i^p \) for any \( \alpha > 0 \).

Note that the asymmetry between the proofs of both bounds stems from the fact that for the lower bound we only needed to find an inequality that holds with high probability for a single vertex, whereas for the upper bound we had to prove an inequality that holds with high probability for all vertices \( s = 1, \ldots, t \).

### 3.1. Upper bound, early vertices \( (s \leq t_0) \)

We begin with the definitions for two auxiliary sequences that we mentioned earlier:

**Definition 3.2.** For any \( t \) and the given coefficients \( \phi(t) \), \((\beta_i(t))_{i=0}^{k-1}\) and the sequence of positive jumps \((w_i(t))_{i=0}^{k-1}\) we define the sequences \((t_i)_{i=0}^{k}\) and \((X_{t_i})_{i=0}^{k}\) and a number \( k(t) \in \mathbb{N} \) also implicitly dependent on \( t \) as follows:

\[
\begin{align*}
t_0 &= \phi(t), \\
t_{i+1} &= t_i + w_i(t), \\
X_{t_0} &= t_0, \\
X_{t_{i+1}} &= X_{t_i} + \beta_i(t) \frac{w_i(t)X_{t_i}}{t_i}, \\
k &\text{ is such that } t_{k-1} < t \leq t_k.
\end{align*}
\]

Moreover, to prove the desired bounds it would be ultimately necessary that \( \phi(t) \) and all \( w_i(t) \) tend to infinity with \( t \). For brevity, from now on we assume the dependency on \( t \) as implicit and write \( \phi \), \( \beta_i \), and \( w_i \) instead of \( \phi(t) \), \( \beta_i(t) \), \( w_i(t) \), respectively.

Note that inductively from the definition it follows that if \( \beta_i \leq 1 \), then \( X_{t_i} \leq t_i \) for all \( i = 0, 1, \ldots, k \).

Moreover, observe that we do not need to specify the values of \( X_\tau \) for \( \tau \) other than \( \{t_0, t_1, \ldots, t_k\} \). 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 reader’s convenience we shall 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:

**Lemma 3.3.** Assume \( \beta_i \geq p > \frac{p(1-p)}{4\ln(1/\varepsilon)} \) and \( w_i \leq \frac{t_i}{\ln(1/\varepsilon)} \). For \( t \to \infty \) we have \( X_{t_i} \geq t_i^p \) for all \( i = 0, 1, \ldots, k \).

**Proof.** Let us define \( X_\tau = \tau^p \). By definition we know that \( X_{t_0} = t_0 \geq Y_{t_0} \). Now,
ON THE CONCENTRATION OF THE MAXIMUM DEGREE...

233 let us assume that $X_i \geq Y_i$ holds for some $i \geq 0$. Then we have

$$Y_{i+1} - Y_i = ((t_i + w_i)^p - t_i^p) = t_i^p \left( \left(1 + \frac{w_i}{t_i}\right)^p - 1 \right)$$

$$\leq t_i^p \left( \frac{pw_i}{t_i} - \frac{p(1-p)w_i^2}{4t_i^2} \right) \leq t_i^p \frac{w_i}{t_i} \left( p - \frac{p(1-p)}{4 \ln t_i} \right),$$

since from Taylor expansion it follows that $(1 + x)^p \leq 1 + px - \frac{p(1-p)x^2}{2}$ for any $p \in [0, 1]$ and any $x \in (0, 1)$. Therefore,

$$Y_{i+1} - Y_i \leq Y_i \frac{w_i}{t_i} \left( p - \frac{p(1-p)}{4 \ln t_i} \right) \leq X_i \frac{w_i}{t_i} = X_{i+1} - X_i,$$

so clearly $X_{i+1} \geq Y_{i+1}$ holds as well, which completes the inductive step. \(\Box\)

Now we prove an upper bound on $X_i$.

**Lemma 3.4.** Assume that $\phi \geq \ln t$, $\beta_i \leq p + \frac{1}{2 \ln t_i^k}$ and $w_i \leq \frac{1}{\ln t_i^k}$. It holds asymptotically as $t \to \infty$ that $X_i \leq \phi^{1-p}\tau_i^p \ln t_i$ for all $i = 0, 1, \ldots, k$.

**Proof.** We again proceed by induction with $Y_\tau = \phi^{1-p}\tau^p \ln \tau$. Clearly, $X_0 = t_0 \leq Y_0 = t_0 \ln t_0$. Directly from the definition we get

$$Y_{i+1} - X_{i+1} = Y_{i+1} - X_i \left( 1 + \frac{\beta_i w_i}{t_i} \right)$$

$$\geq \phi^{1-p} \tau_i^p \ln t_{i+1} - \phi^{1-p} \tau_i^p \ln t_i \left( 1 + \frac{\beta_i w_i}{t_i} \right)$$

$$\geq \phi^{1-p} \tau_i^p \ln t_i \left( \frac{t_i+1}{t_i} \right) \left( \ln t_{i+1} \right) - \frac{1}{t_i} \left( \beta_i w_i \right)$$

$$= \phi^{1-p} \tau_i^p \ln t_i \left( 1 + \frac{w_i}{t_i} \right) \left( 1 + \frac{\ln(1+w_i/t_i)}{\ln t_i} \right) - 1 - \frac{\beta_i w_i}{t_i}.$$

Now we use the inequalities derived from the respective Taylor expansions: $(1 + x)^p \geq 1 + px - \frac{p(1-p)x^2}{2}$ and $\ln(1+x) \geq x - \frac{x^2}{2} \geq 0$, true for any $p \in [0, 1]$ and any $x \in (0, 1)$. In particular, in our case $x = \frac{w_i}{t_i} \leq \frac{1}{\ln t_i} \leq \frac{1}{\ln \tau_i} = o(1)$. Therefore

$$Y_{i+1} - X_{i+1} \geq \phi^{1-p} \tau_i^p \ln t_i \left( \frac{p - \beta_i w_i}{t_i} \right) + \left( 1 + \frac{pw_i}{t_i} \right) \left( \frac{w_i}{t_i} \ln t_i - \frac{w_i^2}{2t_i^2 \ln t_i} \right)$$

$$- \frac{p(1-p)w_i^2}{2t_i^2} \left( 1 + \frac{w_i}{t_i} \ln t_i - \frac{w_i^2}{2t_i^2 \ln t_i} \right)$$

$$\geq \phi^{1-p} \tau_i^p \ln t_i \cdot w_i \left( \frac{1}{2 \ln t_i} + \frac{1}{w_i} \ln t_i - \frac{w_i^2}{8t_i} \left( 1 + \frac{w_i}{t_i} \ln t_i - \frac{w_i^2}{2t_i^2 \ln t_i} \right) \right)$$

$$\geq \phi^{1-p} \tau_i^p \ln t_i \cdot w_i \left( \frac{1}{8} - \frac{1}{8 \ln t_i} \left( \frac{w_i}{t_i} - \frac{w_i^2}{2t_i^2} \right) \right),$$

and for sufficiently large $t$ the last expression is clearly non-negative since $\frac{w_i}{t_i} \leq \frac{1}{\ln t_i} \leq \frac{1}{\ln \tau} \to 0$, which completes the proof. \(\Box\)

Next, we need some bounds on $\deg_r(s)$ holding with high probability to match with the sequence $X_r$. Let us begin with the following estimate:
Lemma 3.5. For any $\phi \leq \tau \leq t$ and any $0 \leq d \leq h$ it is true that

$$\Pr[\deg_{\tau+h}(s) - \deg_{\tau}(s) \geq d \mid \deg_{\tau}(s)] \leq \exp\left(d \ln \frac{\exp(1) \cdot h(p \deg_{\tau}(s) + pd + r)}{d^2} \right).$$

Proof. First, it follows from the definition of the model that $\deg_{\tau+i+1}(s) = \deg_{\tau+i}(s) + I_{\tau+i}$ for $i = 0, 1, \ldots, h - 1$ where $I_{\tau+i} \sim \text{Be}(q_{\tau+i})$ for some $q_{\tau+i} \in [0, 1]$. The probability $q_{\tau+i}$ of adding an edge between $s$ and $\tau + i + 1$ is just a sum of probabilities of two events:

1. when $\text{parent}(\tau + i + 1) \in N_{G_{\tau+i}}(s)$ holds, i.e. with probability $\frac{\deg_{\tau+i}(s)}{r_{\tau+i}}$ (since we draw the parent uniformly), we add an edge with probability $p$ so the whole event has probability $\frac{p \deg_{\tau+i}(s)}{r_{\tau+i}}$,

2. when $\text{parent}(\tau + i + 1) \notin N_{G_{\tau+i}}(s)$ holds, i.e. with probability $1 - \frac{\deg_{\tau+i}(s)}{r_{\tau+i}}$, we add an edge with probability $\frac{r_{\tau+i}}{r_{\tau+i}}$ so the whole event has probability $\frac{r_{\tau+i}}{r_{\tau+i}} - \frac{\deg_{\tau+i}(s)}{r_{\tau+i}}$.

Both events are disjoint, so we obtain $q_{\tau+i} = \frac{p \deg_{\tau+i}(s)}{r_{\tau+i}} + \frac{r_{\tau+i}}{r_{\tau+i}} - \frac{\deg_{\tau+i}(s)}{r_{\tau+i}} \leq \frac{p \deg_{\tau+i}(s) + r}{r_{\tau+i}}$.

Next, we note that the degree grows by at least $d$ if there is a subsequence of $d$ successes $i_1, i_2, \ldots, i_d$ with only failures between them:

$$\Pr[\deg_{\tau+h}(s) - \deg_{\tau}(s) \geq d \mid \deg_{\tau}(s)]$$

$$= \sum_{0 \leq i_1 < \ldots < i_d < h} \Pr\left[ \bigcup_{j \in \{i_1, \ldots, i_d\}} I_{\tau+j} \cup \bigcup_{j \in [0, i_d] \setminus \{i_1, \ldots, i_d\}} -I_{\tau+j} \right]$$

$$= \sum_{0 \leq i_1 < \ldots < i_d < h} \prod_{j \in \{i_1, \ldots, i_d\}} \Pr[\deg_{\tau+j}(s)] \prod_{j \in [0, i_d] \setminus \{i_1, \ldots, i_d\}} \Pr[-\deg_{\tau+j}(s)].$$

Now observe that $\Pr[-\deg_{\tau+j}(s)] \leq 1$ for any $j$ and $\Pr[I_{\tau+i_1}\deg_{\tau+i_j}(s)] \leq \frac{p(\deg_{\tau}(s)+j-1)+r}{r_{\tau+i_j}}$ for $j = 1, 2, \ldots, d$ since $j$-th success occurs after exactly $j - 1$ successes, i.e. when the degree of the vertex $s$ is exactly equal to $\deg_{\tau}(s) + j - 1$. Thus

$$\Pr[\deg_{\tau+h}(s) - \deg_{\tau}(s) \geq d \mid \deg_{\tau}(s)] \leq \sum_{0 \leq i_1 < \ldots < i_d < h} \prod_{j=1}^{d} \frac{p(\deg_{\tau}(s)+j-1)+r}{\tau+i_j}$$

$$\leq \left(\frac{h}{d}\right)^{d \max_{0 \leq i_1 < \ldots < i_d < h} \left\{ \prod_{j=1}^{d} \frac{p(\deg_{\tau}(s)+j-1)+r}{\tau+i_j} \right\}}.$$
This lemma gives a far better bound than the simple estimation \(\deg_{\tau+h}(s) \leq \deg_{\tau}(s) + h\) (e.g. used in [12]). However, it is still too coarse to obtain a desired upper bound that could be coupled with the sequence \(X_{\tau}\). But we can still use it to kickstart the Chernoff bound by bounding the probabilities of all Bernoulli variables:

**Lemma 3.6.** For \(\ln^{1+p} t \leq \tau \leq t\) and \(\varepsilon = \frac{1}{5 \ln \tau}\) with \(h = \frac{\varepsilon^2 \tau}{\ln(1+2\varepsilon)\exp(2)}\) it holds for any constant \(A > 0\) that

\[
\Pr \left[ \max_{j=0,\ldots,h-1} \left\{ \frac{p \deg_{\tau+j}(s) + r}{\tau + j} \right\} \geq \left( 1 + \varepsilon \right) \frac{pX_{\tau} + r}{\tau} \left| \deg_{\tau}(s) \leq X_{\tau} \right. \right] = O(t^{-A}).
\]

**Proof.** Substituting \(d = \varepsilon X_{\tau}\) in Lemma 3.5 we get asymptotically as \(t \to \infty\) that

\[
\Pr \left[ \frac{p \deg_{\tau+h}(s) + r}{\tau + h} \geq \left( 1 + \varepsilon \right) \frac{pX_{\tau} + r}{\tau} \left| \deg_{\tau}(s) \leq X_{\tau} \right. \right] 
\leq \Pr \left[ \deg_{\tau+h}(s) \geq \left( 1 + \varepsilon \right) X_{\tau} \left| \deg_{\tau}(s) \leq X_{\tau} \right. \right] 
\leq \exp \varepsilon X_{\tau} \ln \frac{\exp(1) \cdot h(pX_{\tau} + \varepsilon X_{\tau} + r))}{\varepsilon X_{\tau} \cdot \tau}
\leq \exp \varepsilon X_{\tau} \left( \frac{\exp(1) \cdot h(pX_{\tau} + \varepsilon X_{\tau} + r))}{\varepsilon X_{\tau} \cdot \tau} \right) \leq \exp(-\varepsilon X_{\tau})
\leq \exp \left( - \frac{\max \{\ln^{1+p} t, \tau \}}{5 \ln \tau} \right) \leq \exp \left( - \frac{\ln t \cdot \tau^p / (1 + p)}{5 \ln \tau} \right) \leq t^{-A-1},
\]

for any constant \(A > 0\). In the fourth line we applied inequality \(r \leq pX_{\tau}\). Moreover, in the last line we used the facts that \(X_{\tau} \geq \max\{\phi, \tau^p\}\) and \(\max\{a, b\} \geq a^\gamma b^{1-\gamma}\) for any \(a, b > 0\) and \(\gamma \in [0, 1]\).

To complete the proof it is sufficient to use a union bound over all values up to \(h = O(t)\).

Let us now proceed with providing a Chernoff-type bound on the growth of the degree of a given early vertex:

**Lemma 3.7.** Let \(1 \leq s \leq \tau \leq t\) such that \(\tau \geq \phi = \ln^{1+p} t\). Then for any \(A > 0\) it is true that

\[
\Pr \left[ \deg_{\tau+h}(s) - \deg_{\tau}(s) \geq \frac{3A(1+d)}{\delta^2} \ln t \left| \deg_{\tau}(s) \leq X_{\tau} \right. \right] = O(t^{-A}),
\]

with \(\varepsilon = \delta = \frac{1}{5 \ln \tau}\), and \(h = \frac{3A + \ln t}{\tau^p (1 + \tau^p) (pX_{\tau} + r)}\).

**Proof.** Let us first define an event

\[
D_c(\tau, h) = \left[ \max_{j=0,\ldots,h-1} \left\{ \frac{p \deg_{\tau+j}(s) + r}{\tau + j} \right\} \geq \left( 1 + \varepsilon \right) \frac{pX_{\tau} + r}{\tau} \left| \deg_{\tau}(s) \leq X_{\tau} \right. \right].
\]

Clearly,

\[
\Pr \left[ \deg_{\tau+h}(s) - \deg_{\tau}(s) \geq d \left| \deg_{\tau}(s) \leq X_{\tau} \right. \right] 
\leq \Pr \left[ \deg_{\tau+h}(s) - \deg_{\tau}(s) \geq d \left\{ \deg_{\tau}(s) \leq X_{\tau}, \neg D_c(\tau, h) \right\} + \Pr[D_c(\tau, h)],
\]

This manuscript is for review purposes only.
Let us estimate the probability of the second event. If \( h = \frac{3A\tau\ln t}{\sigma^2(1+\tau)(pX_\tau+r)} \) and \( \varepsilon = \frac{1}{5\ln \tau} \), then the condition \( h \leq \frac{\varepsilon \tau}{p(1+\varepsilon)\exp(2)\ln \tau} \) is met since for some constant \( C > 0 \)
we have
\[
h \leq \frac{C\tau \ln t}{\delta^2 X_\tau} = \frac{C\tau \ln t \cdot 25 \ln^2 \tau}{\max\{\ln^{1+p} t, \tau^p\}} \leq \frac{C\tau \ln t \cdot 25 \ln^2 \tau}{\ln^2 \tau} \leq \frac{\tau}{\varepsilon \tau}
\]
and from Lemma 3.6 we obtain that \( \Pr[D_\varepsilon(\tau, h)] = O(t^{-A}) \). Here we again used the facts that \( X_\tau \geq \max\{\phi, \tau^p\} \) and \( \max\{a, b\} \geq a^\gamma b^{1-\gamma} \) for any \( a, b > 0 \) and \( \gamma \in [0, 1] \).
Thus, it is sufficient to bound \( \deg_{\tau+h}(s) - \deg_\tau(s) \) with high probability when
\( D_\varepsilon(\tau, h) \) does not hold, that is, when for all \( i = 1, \ldots, h \) it is true that
\[
\frac{\deg_{\tau+i}(s)}{\tau + i} < (1 + \varepsilon) \frac{X_\tau}{\tau}.
\]
It follows that \( I_{\tau+i} = \deg_{\tau+i+1}(s) - \deg_{\tau+i}(s) \) is stochastically dominated by independent
random variables \( I_{\tau+i} \sim Be\left(\left(1 + \varepsilon\right)\frac{pX_\tau+r}{\tau}\right) \) for any \( i = 0, 1, \ldots, h - 1 \).
In the case of Bernoulli variables \( Be(p_1) \) is stochastically dominated by \( Be(p_2) \)
whenever \( p_1 \leq p_2 \). This way we can eliminate dependencies – the outcome of each
\( I_\tau \) influences the distributions for \( I_{\tau'}, \tau' > \tau \) – and work with independent variables
\( I_{\tau+i} \).
Now, since the new variables are both Bernoulli and independent, we can use the
well-known left tail Chernoff bound for binomial setting from [10] (see Corollary 21.7)
which states that for any \( \delta \in (0, 1) \)
\[
\Pr\left[\sum_{i=0}^{h-1} I_{\tau+i} \geq (1 + \delta) E\left[\sum_{i=0}^{h-1} I_{\tau+i}\right]\right] \leq \exp\left(-\frac{\delta^2}{3} E\left[\sum_{i=0}^{h-1} I_{\tau+i}\right]\right)
\]
and therefore
\[
\Pr\left[\deg_{\tau+h}(s) - \deg_\tau(s) \geq (1 + \delta)(1 + \varepsilon) \frac{h(pX_\tau+r)}{\tau} \left| \deg_\tau(s) \leq X_\tau, \neg D_\varepsilon(\tau, h)\right.\right] \leq \exp\left(-\frac{h\delta^2(1 + \varepsilon)(pX_\tau+r)}{3\tau}\right).
\]
To finish the proof it is sufficient to see that \( h = \frac{3A\tau\ln t}{\sigma^2(1+\tau)(pX_\tau+r)} \) gives the required
\( O(t^{-A}) \) bound in the last equation.

Finally, we proceed with the proof of the main result of this section.

**Theorem 3.8.** For \( G_t \sim DD(t, p, r) \) with \( 0 < p < 1 \) and \( s \in [1, \ln^{1+p} t] \) it holds
asymptotically that
\[
\Pr\left[\deg_\tau(s) \geq (1 + \alpha) t^p \ln^{2-p^2} t \right] = O(t^{-A})
\]
for any constants \( \alpha > 0 \) and \( A > 0 \).

**Proof.** Throughout the proof we will use sequences \((t_i)_{i=0}^k\) and \((X_i)_{i=0}^k\) with
\( \phi = \ln^{1+p} t, \beta_i = p + \frac{1}{2 \ln t_i}, w_i = \frac{1}{\sigma^2(1+\tau)(pX_i+r)} \), and \( \varepsilon = \delta = \frac{1}{5 \ln t_i} \).

This manuscript is for review purposes only.
Observe that all the assumptions of Lemma 3.3 and Lemma 3.4 are met, i.e. $p \leq \beta_i \leq p + \frac{1}{2 \ln t_i}$ and $w_i \leq \frac{t_i}{2 \ln t_i}$, so we know that $\max \{ \ln^{1+p} t_i^p \} \leq X_i \leq t^p \ln^{2-p^2} t$

for all $i = 0, 1, \ldots, k$.

Now let us define events $A_i(s) = [\deg_{t_i}(s) < X_i]$ for $i = 0, \ldots, k$. Clearly, $A_0(s)$ holds since by definition of $X_0 = X_0$ we have $\deg_{t_0}(s) < t_0 = X_0$.

Suppose that $A_i(s)$ holds. Then we can apply Lemma 3.7 with $\tau = t_i$ and $h = w_i$:

$$Pr[\neg A_{i+1}(s) | A_i(s)] = Pr[\deg_{t_{i+1}}(s) \geq X_{t_{i+1}} \mid \deg_{t_i}(s) < X_i]$$

$$\leq Pr[\deg_{t_{i+1}}(s) - \deg_{t_i}(s) \geq X_{t_{i+1}} - X_t \mid \deg_{t_i}(s) < X_i]$$

$$= Pr \left[ \deg_{t_{i+1}}(s) - \deg_{t_i}(s) \geq \frac{\beta_i X_i}{w_i} \mid \deg_{t_i}(s) < X_i \right]$$

$$= Pr \left[ \deg_{t_{i+1}}(s) - \deg_{t_i}(s) \geq \frac{\beta_i X_i}{(1 + \varepsilon)(p X_i + r)} \frac{3(A + 1)}{\delta^2} \ln t \mid \deg_{t_i}(s) < X_i \right]$$

$$\leq Pr \left[ \deg_{t_{i+1}}(s) - \deg_{t_i}(s) \geq \frac{3(A + 1)(1 + \delta)}{\delta^2} \ln t \mid \deg_{t_i}(s) < X_i \right] = O(t^{-A-1}),$$

where we used the fact that asymptotically as $t \to \infty$

$$\frac{\beta_i X_i}{(1 + \delta)(1 + \varepsilon)(p X_i + r)} = \frac{\beta_i X_i}{\delta^2} \frac{3(A + 1)}{\delta^2} \ln t \to \frac{p + \delta}{p + \delta + \varepsilon + \frac{r}{X_i}} \geq 1,$$

where in the denominator of the first inequality we used the facts that $p + \frac{1}{X_i} \leq p + \varepsilon + \frac{r(1+\varepsilon)}{X_i} = p + o(1) \leq 1$ for any constants $0 < p < 1$, $0 \leq r \leq t_0$ when $t \to \infty$.

Next, we get

$$Pr[\deg_{t_i}(s) \geq X_i] \leq Pr[\deg_{t_k}(s) \geq X_k] = Pr[\neg A_k(s)]$$

$$\leq \sum_{i=0}^{k-1} Pr[\neg A_{i+1}(s) | A_i(s)] + Pr[\neg A_0(s)] = \sum_{i=0}^{k-1} O(t^{-A-1}) = O(t^{-A}),$$

since asymptotically it is true that $w_i \geq 1$ for all $i = 0, \ldots, k$, and therefore $k \leq t$.

To complete the proof it is sufficient to note that $t_k = t_{k-1} (1 + \alpha) \leq (1 + \alpha) t$ and thus $X_k \leq (1 + \alpha) t^p \ln^{2-p^2} t$ for any constant $\alpha > 0$.

### 3.2. Upper bound, late vertices ($s > t_0$)

In the second part of the proof we also use the sequences $(t_i)_{i=0}^k$ and $(X_i)_{i=0}^k$ as defined in Definition 3.2. Moreover, throughout this section we use the same constants as in the proof of Theorem 3.8:

$$\phi = \ln^{1+p} t_i, \beta_i = p + \frac{1}{2 \ln t_i}, \text{ and } w_i = \frac{\beta_i X_i}{(1 + \varepsilon)(p X_i + r)} \ln t$$

The proof consists of showing that for $s \in [t_i, t_{i+1})$ for some $i = 0, 1, \ldots, k - 1$ the degree graph (i.e. $\deg_{s}(s)$) is with high probability significantly smaller than its corresponding $X_{t_{i+1}}$. 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, $X_{tm}$) gives us a good upper bound on $\deg_s(s)$ for all $s$ – and therefore we also obtain an upper bound for $\Delta(G_t)$.

Let us introduce auxiliary events $E_t(s) = \bigcup_{\tau = 1}^s A_t(\tau) \leq X_t$ for all $\tau \leq s \leq t_1$ where $A_t(s)$ is, as before, the event that $\deg_t(s) \leq X_t$ for a fixed $s \leq t_1$.

**Lemma 3.9.** Let $s \in \{t_l, t_{l+1}\}$ for some $l = 0, 1, \ldots, k-1$. Then, for any constants $\epsilon > 0$ and $A > 0$

$$\Pr \left[ \deg_s(s) \geq (1 + \epsilon)(pX_{t_{l+1}} + r) \mid E_t(t_l) \land E_{t+1}(s-1) \right] = O(t^{-A}).$$

**Proof.** First, we notice the fact that $\max\{\deg_{s+1}(\tau): 1 \leq \tau \leq s - 1\} \leq X_{t_{l+1}}$ guarantees that $\max\{\deg_s(s): 1 \leq \tau \leq s - 1\} \leq X_{t_{l+1}}$. Therefore, $\deg_s(s)$ is stochastically dominated by $A_s \sim \text{Bin}(X_{t_{l+1}}, p) + \text{Bin}(s-1, \frac{r}{s-1})$ and we directly obtain the result using the Chernoff bound with $\mathbb{E}[A_s] = pX_{t_{l+1}} + r$:

$$\Pr \left[ \deg_s(s) \geq (1 + \epsilon)(pX_{t_{l+1}} + r) \mid E_t(t_l) \land E_{t+1}(s-1) \right] \leq \exp \left( -\frac{\epsilon^2}{\epsilon + 2}(pX_{t_{l+1}} + r) \right) \leq t^{-A},$$

asymptotically for any constants $\epsilon, A > 0$ since $X_{t_{l+1}} \geq \ln^{1+p} t$.

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

**Lemma 3.10.** Let $s \in \{t_l, t_{l+1}\}$ for some $l = 0, 1, \ldots, k-1$. Then, for any constant $\alpha > 0$ and $A > 0$

$$\Pr \left[ \deg_{s+1}(s) - \deg_s(s) \geq \alpha X_{t_{l+1}} \mid E_t(t_l) \land E_{t+1}(s-1) \right] = O(t^{-A}).$$

**Proof.** We use Lemma 3.5 with $d = \alpha X_{t_{l+1}}$ to obtain asymptotically as $t \to \infty$ that for any $A > 0$ it holds that

$$\Pr \left[ \deg_{s+1}(s) - \deg_s(s) \geq \alpha X_{t_{l+1}} \mid E_t(t_l) \land E_{t+1}(s-1) \right] = \Pr \left[ \deg_{s+1}(s) - \deg_s(s) \geq \alpha X_{t_{l+1}} \right] \leq \Pr \left[ \deg_{s+1}(s) - \deg_s(s) \geq \alpha X_{t_{l+1}} \right] \leq \exp \left( \alpha X_{t_{l+1}} \ln \left( \exp(1) \cdot \frac{wp(1 + 2\alpha)}{X_{t_{l+1}} \cdot s} \right) \right) \leq \exp \left( \alpha X_{t_{l+1}} \left( \exp(1) \cdot \frac{1 + 2\alpha}{\alpha(1 + \alpha)} \cdot 3(A + 1) + \ln \frac{\ln t}{\delta^2(X_{t_{l+1}} + r/p)} \right) \right) \leq \exp \left( \alpha X_{t_{l+1}} \left( \Theta(1) + \ln \frac{25 \ln t \cdot \ln^2 t_l}{\max\{1, t_{l_{l}}^2(1 + p)\}} \right) \right) \leq \exp \left( \alpha \ln^{1+p} t \left( \Theta(1) + \ln \frac{25 \ln^2 t_l}{t_{l_{l}}^2(1 + p)} \right) \right) \leq \exp(-A \ln t) \leq t^{-A}$$

as needed.
Lemma 3.11. Let \( s \in (t_i, t_{i+1}] \) for some \( l = 0, 1, \ldots, k - 1 \). Then asymptotically as \( t \to \infty \), for any constant \( A > 0 \) it holds that
\[
\Pr \left[ \deg_{t_{i+1}}(s) \geq X_{t_{i+1}} | \mathcal{B}_l(t_i) \land \mathcal{B}_{l+1}(s-1) \right] = O(t^{-A}).
\]

Proof. We combine Lemma 3.9 with \( \varepsilon = \frac{1-p}{4p} \) and Lemma 3.10 with \( \alpha = \frac{1-p}{2} \) to obtain
\[
\Pr \left[ \deg_{t_{i+1}}(s) \geq X_{t_{i+1}} | \mathcal{B}_l(t_i) \land \mathcal{B}_{l+1}(s-1) \right] 
\leq \Pr \left[ \deg_s(s) \geq \left( 1 + \frac{1-p}{4p} \right) (pX_{t_{i+1}} + r) | \mathcal{B}_l(t_i) \land \mathcal{B}_{l+1}(s-1) \right] 
+ \Pr \left[ \deg_{t_{i+1}}(s) - \deg_s(s) \geq \frac{1-p}{2} X_{t_{i+1}} | \mathcal{B}_l(t_i) \land \mathcal{B}_{l+1}(s-1) \right] = O(t^{-A}). \]

Lemma 3.12. Let \( s \in (t_i, t_{i+1}] \) for some \( l = 0, 1, \ldots, k - 1 \). Then asymptotically as \( t \to \infty \), for any constant \( A > 0 \) it holds that
\[
\Pr \left[ \neg \mathcal{B}_{l+1}(t_{i+1}) | \mathcal{B}_l(t_i) \right] = O(t^{-A}).
\]

Proof. Let \( l \) be the first value for which the lemma does not hold. Then, from Lemma 3.11 we get that for any constant \( A > 0 \) it holds that
\[
\Pr \left[ \neg \mathcal{B}_{l+1}(t_{i+1}) | \mathcal{B}_l(t_i) \right] = \sum_{s=t_i}^{t_{i+1}-1} \Pr \left[ \neg \mathcal{B}_{l+1}(s+1) | \mathcal{B}_l(t_i) \land \mathcal{B}_{l+1}(s) \right]
= \sum_{s=t_i}^{t_{i+1}-1} \Pr \left[ \neg \mathcal{A}_{l+1}(s+1) | \mathcal{B}_l(t_i) \land \mathcal{B}_{l+1}(s) \right] = O(t^{-A}).
\]

From Theorem 3.8 we know that \( \Pr[\mathcal{B}_l(t_0)] = 1 - O(t^{-A}) \). Recall that by our assumption \( \Pr[\neg \mathcal{B}_{l+1}(t_{i+1}) | \mathcal{B}_l(t_i)] = 1 - O(t^{-A}) \) for all \( i = 0, 1, \ldots, l - 1 \), so it follows that \( \Pr[\mathcal{B}_l(t_i)] = 1 - O(t^{-A}) \) for all \( i = 0, 1, \ldots, l \). We use this fact, combined with the observation that \( \mathcal{B}_l(t_i) \subseteq \mathcal{A}_i(s) \) and Theorem 3.8 to get
\[
\Pr \left[ \neg \mathcal{B}_{l+1}(t_i) | \mathcal{B}_l(t_i) \right] \leq \sum_{s=1}^{t_i} \Pr \left[ \neg \mathcal{A}_{l+1}(s) | \mathcal{B}_l(t_i) \right]
\leq \sum_{s=1}^{t_i} \frac{\Pr[\neg \mathcal{A}_{l+1}(s) \land \mathcal{A}_l(s)]}{\Pr[\mathcal{B}_l(t_i)]}
\leq \sum_{s=1}^{t_i} \frac{\Pr[\neg \mathcal{A}_{l+1}(s) \land \mathcal{A}_l(s)]}{\Pr[\mathcal{B}_l(t_i)]}
= \sum_{s=1}^{t_i} \frac{O(t^{-A})}{1 - O(t^{-A})} = O(t^{-A}).
\]

Finally, for any events \( E_1, E_2, E_3 \) we have
\[
\Pr[\neg E_1|E_2] = \Pr[\neg E_1 \land E_3|E_2] + \Pr[\neg E_1 \land \neg E_3|E_2]
\leq \Pr[\neg E_1|E_3 \land E_2] + \Pr[\neg E_3|E_2].
\]

We substitute \( E_1 = \mathcal{B}_{l+1}(t_{i+1}), E_2 = \mathcal{B}_l(t_i) \) and \( E_3 = \mathcal{B}_{l+1}(t_i) \) to obtain the final result. \qed
Finally, we present the main result of this section.

**Theorem 3.13.** For $G_t \sim \mathcal{D}(t, p, r)$ with $0 < p < 1$ and any constants $\alpha, A > 0$

it holds asymptotically that

$$
\Pr \left[ \Delta(G_t) \geq (1 + \alpha) t^p \ln^{2-p^2} t \right] = O(t^{-A}).
$$

**Proof.** From Lemma 3.4 we know that $X_{t_k} \leq (1 + \alpha) t^p \ln^{2-p^2} t$ holds asymptotically. It follows that in this case

$$
\Pr \left[ \Delta(G_t) \geq (1 + \alpha) t^p \ln^{2-p^2} t \right] \leq \Pr[\Delta(G_t) \geq X_{t_k}] \leq \Pr[\neg B_k(t_k)]
$$

$$
\leq \sum_{l=0}^{k-1} \Pr[\neg B_{l+1}(t_{l+1})|B_l(t_l)] + \Pr[\neg B_0(t_0)].
$$

Now, from Theorem 3.8 and Lemma 3.12 we know that both $\Pr[\neg B_0(t_0)] = O(t^{-A})$ and $\Pr[\neg B_{l+1}(t_{l+1})|B_l(t_l)] = O(t^{-A})$ for any $A > 0$, respectively. Putting this all together with the fact that asymptotically as $t \to \infty$ it holds that $k \leq t$ we obtain the final result. \qed

### 3.3. Lower bound.

Here we proceed analogously to the case of the upper bound for early vertices. 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 $\deg_s(s)$. 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 $G_t$.

Now, 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 $(1 - \frac{1}{2} \tau)^t \leq \exp(-\tau)$, so within first $\frac{3}{2} \ln t$ 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 is no 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.

That said, let us now proceed with the aforementioned Chernoff-type lower bound for the degree of a given early vertex:

**Lemma 3.14.** Let $1 \leq s \leq \tau \leq t$ such that $\tau \geq \phi = \ln^{1+p} t$. Then for any $A > 0$

it is true that

$$
\Pr \left[ \deg_{\tau+h}(s) - \deg_{\tau}(s) \leq \frac{2A(1-\delta)}{\delta^2} \ln t \right| \deg_{\tau}(s) \geq X_{\tau}] = O(t^{-A}),
$$

with $\varepsilon = \delta = \frac{\nu(1-\rho)}{8 \ln \tau}$ and $h = \frac{2A \tau \ln t}{8 \varepsilon^2 (1-\varepsilon)(p X_{\tau}+\varepsilon)}$.

**Proof.** Let us recall (as in the proof of Lemma 3.5) that for $i = 0, 1, \ldots, h-1$ we have $\deg_{\tau+i}(s) = \deg_{\tau+i}(s) + I_{\tau+i}$ where $I_{\tau+i} \sim \text{Be}(q_{\tau+i})$ for $q_{\tau+i} = \frac{p \deg_{\tau+i}(s)+r}{\tau+i}$. 

This manuscript is for review purposes only.
for any \( k \) as clearly \( \Pr_{\delta} \) is independent. So this time we can use the right tail Chernoff bound for binomial settings from [10] (see Corollary 21.7) which states that for any \( \delta \in (0,1) \)

\[
\Pr\left[ \sum_{i=0}^{h-1} I^*_{r+i} \leq (1 - \delta) E\left[ \sum_{i=0}^{h-1} I^*_{r+i} \right] \right] \leq \exp\left( -\frac{\delta^2}{2} E\left[ \sum_{i=0}^{h-1} I^*_{r+i} \right] \right)
\]

and therefore

\[
\Pr\left[ \deg_{r+h}(s) \leq \deg_r(s) + (1 - \delta)(1 - \varepsilon) \frac{h(pX_r + r)}{\tau} \right] \leq \exp\left( -\frac{h\delta^2(1 - \varepsilon)(pX_r + r)}{2\tau} \right)
\]

as clearly \( \Pr[\deg_{r+h}(s) - \deg_r(s) \leq k] = \Pr[\sum_{i=0}^{h-1} I_{r+i} \leq k] \leq \Pr[\sum_{i=0}^{h-1} I^*_{r+i} \leq k] \)

for any \( k \), due to the stochastic dominance.

To finish the proof it is sufficient to see that \( h = \frac{2A_{r-1} \ln t}{\delta^2(1 - \varepsilon)(pX_r + r)} \) gives the required \( O(t^{-A}) \) bound in the last equation.

In the following, we again use sequences \((t_i)_{i=1}^k\) and \((X_{t_i})_{i=1}^k\) from Definition 3.2.

Let us also define \( C_{t_i}(s) = [deg_{t_i}(s) > X_{t_i} - \phi + 1] \) for a fixed \( s \leq t_i \). Now we are in the position to proceed with the main theorem of this section:

**Theorem 3.15.** For \( G_t \sim DD(t, p, r) \) with \( 0 < p < 1 \) there exists \( s \) such that it holds asymptotically that

\[
\Pr[deg_{t_i}(s) < (1 - \alpha)t^p] = O(t^{-A})
\]

for any constants \( \alpha, A > 0 \).

**Proof.** Let us use \( \phi = \ln^{1+p} t \), \( \beta_i = p - \frac{p(1-p)}{4\ln t_i} \) and \( u_i = \frac{2(A+1)t_i \ln t}{\delta^2(1 - \varepsilon)(pX_{t_i} + r)} \) with \( \delta = \varepsilon = \frac{p(1-p)}{8\ln t_i} \).

Suppose that \( C_{t_i}(s) \) holds. Then we can apply Lemma 3.14 with \( \tau = t_i \) and \( h = u_i \):

\[
\Pr[-C_{t_{i+1}}(s)|C_{t_i}(s)] = \Pr[deg_{t_{i+1}}(s) \leq X_{t_{i+1}} | deg_{t_{i}}(s) > X_{t_i} - \phi + 1] \leq \Pr[deg_{t_{i+1}}(s) - deg_{t_{i}}(s) \leq X_{t_{i+1}} - X_{t_i} | deg_{t_{i}}(s) > X_{t_i} - \phi + 1] = \Pr\left[ deg_{t_{i+1}}(s) - deg_{t_{i}}(s) \leq \beta_i \frac{u_i X_{t_i}}{t_i} \left| deg_{t_{i}}(s) > X_{t_i} - \phi + 1 \right. \right] \leq \Pr\left[ deg_{t_{i+1}}(s) - deg_{t_{i}}(s) \leq \frac{2(A+1)(1 - \delta)}{\delta^2} \ln t \left| deg_{t_{i}}(s) > X_{t_i} - \phi + 1 \right. \right] = O(t^{-A-1}),
\]
where we used the fact that asymptotically as $t \to \infty$ it holds that
\[
\frac{\beta X_t}{(1-\delta)(1-\varepsilon)(pX_t+r)} \leq \frac{p-p(1-p)}{p(1-\delta-\varepsilon)} = 1.
\]

Next, we get
\[
\Pr[\deg_t(s) \leq X_{tk} - \phi + 1] \leq \Pr[\deg_{tk}(s) \leq X_{tk} - \phi + 1] = \Pr[\neg C_k(s)]
\]
\[
\leq \sum_{i=0}^{k-1} \Pr[\neg C_{k+1}(s)|C_i(s)] + \Pr[\neg C_0(s)] = \sum_{i=0}^{k-1} O(t^{-A-1}) = O(t^{-A}),
\]
since asymptotically it is true that $w_i \geq 1$ for all $i = 0, \ldots, k$, and therefore $k \leq t$.

To complete the proof it is sufficient to note that $t \leq tk \leq (1+\alpha)tk-1 \leq (1+\alpha)t$ for any constant $\alpha > 0$ and thus $X_{tk} \leq (1+\alpha)t^p$.

**4. Average degree.** Now let us proceed to the results on the average degree of $G_t$ defined as
\[
D(G_t) = \frac{1}{t} \sum_{s=1}^{t} \deg_t(s).
\]

First, we recall from [30, Theorem 9(iii)] that for any $\tau = t_0, \ldots, t-1$ it holds
asymptotically (i.e. when $t_0 \to \infty$) that
\[
E[\deg_t(\tau)] = \begin{cases} 
D(G_{t_0}) \frac{p^{f(t_0)}\Gamma(t_0+1)}{(t_0+c)^2} 2^{2p-1}(1+o(1)) & \text{if } p \leq \frac{1}{2}, \tau = 0, \\
r(1+o(1)) & \text{if } p = 0, \tau > 0, \\
\tau(1-\frac{p}{2}) & \text{if } 0 < p < \frac{1}{2}, \tau > 0, \\
r \log \tau \left( 1 + o(1) \right) & \text{if } p = \frac{1}{2}, \tau > 0, \\
\left( D(G_{t_0}) + \frac{2p_{t_0}}{\Gamma(t_0+c+1)} \right) 2^{2p-1} \frac{\gamma_{t_0}^{t_0+1,t_0+1,1}}{\Gamma(t_0+c+1)} + 1 & \text{if } p > \frac{1}{2},
\end{cases}
\]
where $D(G_{t_0})$ is the average degree of the initial graph $G_{t_0}$ and
\[
\gamma_{t_0}^{t_0+1,t_0+1,1} = \sum_{l=0}^{\infty} \left( \frac{a_1 a_2 a_3}{b_1 b_2 b_3} \right) \frac{z^l}{l!}
\]
is the generalized hypergeometric function with $(a)_l = a(a+1) \ldots (a+l-1)$, $(a)_0 = 1$
the rising factorial (see [1] for details).

In short, if we omit constant factors, there are three regimes of growth: constant, \ln $t$, and $t^{2p-1}$. We need to find the proper high probability bound for each case separately, however it turns out that the proofs are very similar.

**4.1. Upper bound.** Now we may proceed to the main result of this section: the upper bound for the average degree of $G_t$. It turns out that there are exactly two regimes with somewhat different behavior:

**Theorem 4.1.** Asymptotically for $G_t \sim DD(t,p,r)$ it holds that
\[
\Pr[D(G_t) \geq AC \ln t] = O(t^{-A}) \quad \text{for } p \leq \frac{1}{2},
\]
\[
\Pr[D(G_t) \geq Ct^{2p-1}] = O(t^{-A}) \quad \text{for } p > \frac{1}{2},
\]
for some fixed constant $C > 0$ and any $A > 0$.
Proof. For simplicity, we will work with the total number of edges $\tau D(G_\tau)$ instead of $D(G_\tau)$. Clearly, for any $\tau = t_0, \ldots, t-1$ it holds that

\[(\tau + 1)D(G_{\tau+1}) - \tau D(G_\tau) = 2\deg_{\tau+1}(\tau + 1),\]

so therefore

\[\deg_{\tau+1}(\tau + 1) \sim \text{Bin}(\deg_\tau(\text{parent}(\tau+1)), p) + \text{Bin}(\tau - \deg_\tau(\text{parent}(\tau+1)), r/\tau).\]

Therefore, we can use Chernoff bound to obtain for any $\delta \geq 0$

\[\Pr \left[ (\tau + 1)D(G_{\tau+1}) - \tau D(G_\tau) \geq 2(1 + \delta)\mathbb{E}[\deg_{\tau+1}(\tau + 1)] \right] \leq \exp\left( -2\delta^2\frac{\mathbb{E}[\deg_{\tau+1}(\tau + 1)]}{2 + \delta} \right).\]

Now, for $p > \frac{1}{2}$ we know that $\mathbb{E}[\deg_\tau(\tau)] \leq C^* \tau^{2p-1}$ for some constant $C^* > 0$. Thus, it is sufficient to set $t_0 = t^{p/3}$ and $\delta = \sqrt{\frac{3(A+1)\ln t}{2C^*\tau^{2p-1}}} = o(1)$ for all $\tau = t_0, \ldots, t-1$ to get

\[\Pr \left[ (\tau + 1)D(G_{\tau+1}) - \tau D(G_\tau) \geq 2(1 + \delta)C^* \tau^{2p-1} \right] = O(t^{-A-1}),\]

and by summing over all $\tau$ that no event from polynomial tails happens we obtain

\[\Pr \left[ tD(G_t) \geq Ct^{2p} \right] \leq \Pr \left[ tD(G_t) - t_0 D(G_{t_0}) \geq \sum_{i=t_0}^{t-1} 2(1 + \delta)C^* \tau^{2p-1} \right] = O(t^{-A}),\]

for any constant $C \geq t^{-2p} \sum_{i=t_0}^{t-1} 2(1 + \delta)C^* \tau^{2p-1} + t^{-2p} t_0 D(G_{t_0})$ and such constant $C$ indeed exists since it is not hard to verify that the latter sum is finite.

In all cases $0 < p \leq \frac{1}{2}$ it turns out that $\sqrt{\frac{3(A+1)\ln t}{2C^*\tau^{2p-1}}} \to \infty$. However, for $0 < p \leq \frac{1}{2}$, $r > 0$ we have $\mathbb{E}[\deg_\tau(\tau)] \leq C^* \ln \tau$ for some constant $C^* > 0$, and we can assume $\delta \to \infty$ such that

\[\frac{1 + \delta}{2} \leq \frac{\delta^2}{2 + \delta} = \frac{(A + 1) \ln t}{2C^* \ln \tau},\]

so therefore

\[\Pr \left[ (\tau + 1)D(G_{\tau+1}) - \tau D(G_\tau) \geq 2(A + 1) \ln t \right] = O(t^{-A-1}),\]

\[\Pr \left[ tD(G_t) \geq ACt \ln t \right] \leq \Pr \left[ tD(G_t) - t_0 D(G_{t_0}) \geq \sum_{i=t_0}^{t-1} 2(A + 1) \ln i \right] = O(t^{-A}),\]

for some constant $C \geq 2 + \frac{t_0}{\ln t} D(G_{t_0})$ when $t_0 = t^{1/3}$.

Finally, let us study the case $0 < p < \frac{1}{2}$, $r = 0$. Again we know that $\mathbb{E}[\deg_\tau(\tau)] \leq C^* \tau^{2p-1}$ for some constant $C^* > 0$. Again, we can assume

\[\frac{1 + \delta}{2} \leq \frac{\delta^2}{2 + \delta} = \frac{(A + 1) \ln t}{2C^* \tau^{2p-1}},\]

so by a similar reasoning as before we get

\[\Pr \left[ (\tau + 1)D(G_{\tau+1}) - \tau D(G_\tau) \geq 2(A + 1) \ln t \right] = O(t^{-A-1}),\]

\[\Pr \left[ tD(G_t) \geq ACt \ln t \right] \leq \Pr \left[ tD(G_t) - t_0 D(G_{t_0}) \geq \sum_{i=t_0}^{t-1} 2(A + 1) \ln i \right] = O(t^{-A}),\]

for sufficiently large constant $C$ when $t_0 = t^{1/3}$. \qed
4.2. Lower bound. We now turn our attention to establishing the corresponding lower bound. Note that since $E[D(G_1)] = O(\log t)$ for $p \leq \frac{1}{2}$, the lower polynomial tail is trivial in this range since all smaller values are within the polylogarithmic distance from the mean. However, we can investigate the case $p > \frac{1}{2}$.

**Theorem 4.2.** For $G_t \sim DD(t, p, r)$ with $p > \frac{1}{2}$ asymptotically it holds that

$$\Pr[D(G_t) \leq C t^{2p-1}] = O(t^{-A}).$$

for some fixed constant $C > 0$ and any $A > 0$.

**Proof.** Similarly as before, we invoke the appropriate Chernoff bound for $\delta \in (0, 1)$

$$\Pr \left[(\tau + 1)D(G_{\tau+1}) - \tau D(G_{\tau}) \leq 2(1 - \delta) E[\deg_{\tau+1}(\tau + 1)] \right] \leq \exp(-\delta^2 E[\deg_{\tau+1}(\tau + 1)]).$$

For $p > \frac{1}{2}$ it is true that $E[\deg_{\tau}(\tau)] \geq C^* \tau^{2p-1}$ for some constant $C^* > 0$. Thus, it is sufficient to set $t_0 = t^{p/3}$ and $\delta = \sqrt{(4+1)\ln t}{C^* \tau^{2p-1}} \leq \frac{1}{2}$ for all $\tau = t_0, \ldots, t-1$ to get

$$\Pr \left[(\tau + 1)D(G_{\tau+1}) - \tau D(G_{\tau}) \leq 2(1 - \delta) C^* \tau^{2p-1} \right] = O(t^{-A-1}),$$

which leads us to

$$\Pr \left[t(D(G_t) - t_0 D(G_{t_0}) \leq C t^{2p} \right] \leq \Pr \left[t D(G_t) - t_0 D(G_{t_0}) \leq \sum_{i=t_0}^{t-1} 2(1 - \delta) C^* \tau^{2p-1} \right] = O(t^{-A}).$$

for any constant $0 < C \leq t^{-2p} \sum_{i=t_0}^{t-1} 2(1 - \delta) C^* \tau^{2p-1} + t^{-2p} t_0 D(G_{t_0})$ and such constant indeed exists since it is not hard to verify that the latter sum is non-zero and finite when $t_0 = t^{1/3}$.

5. Further challenges. In this paper we focus on deriving large deviations for the average and the maximum degree in the duplication-divergence networks. By a simple martingale argument one can show that $\Delta(G_t)/t^p$ converges to some random variable $\Delta$. However, it is still worth asking whether $\Delta$ has finite support (e.g. dependent only on $p$ and $r$, but not on $t$).

A natural next challenge would be to obtain the exact asymptotic formula for the whole degree distribution. For example, there is an open question whether $DD(t, p, r)$ graphs are scale-free, i.e. they have $\Theta(k^{-\gamma})$ fraction of vertices with degree $k$. A first step towards this goal was already done for $r = 0$ in [18, 16], where it was proved that this property indeed holds for the (only) giant component $p < e^{-1}$. However, it was noticed in [14] that for $r = 0$ and all $0 < p < 1$ such phenomenon does not appear in the whole graph, since almost all vertices are isolated, thus for any $k > 0$ the fraction of vertices of degree $k$ tends to 0 as $t \to \infty$.

Finally, finding good bounds on the concentration of both $D(G_t)$ and $\Delta(G_t)$ is only the step towards the full understanding of this model, as we still do not know for example how symmetric such networks are. This, in turn, we believe could help find good compression algorithms for these types of networks, as was the case with other graph models [3, 23].

REFERENCES
ON THE CONCENTRATION OF THE MAXIMUM DEGREE... 19


