High degree vertices and eigenvalues in the preferential attachment graph

Abraham Flaxman [*]
Department of Mathematical Sciences
Carnegie Mellon University
Pittsburgh, PA, 15213, USA
email abie@cmu.edu

Alan Frieze [†]
Department of Mathematical Sciences
Carnegie Mellon University
Pittsburgh, PA, 15213, USA
email alan@random.math.cmu.edu

Trevor Fenner
School of Computer Science
Birkbeck College, University of London
Malet Street, London WC1E 7HX
email trevor@dcs.bbk.ac.uk

October 2, 2003

Abstract

The preferential attachment graph is a random graph formed by adding a new vertex at each time step, with a single edge which points to a vertex selected at random with probability proportional to its degree. Every m steps the most recently added m vertices are contracted into a single vertex, so at time t there are roughly t/m vertices and exactly t edges. This process yields a graph which has been proposed as a simple model of the world wide web [BA99]. For any constant k, let $\Delta_1 \geq \Delta_2 \geq \cdots \geq \Delta_k$ be the degrees of the k highest degree vertices. We show that at time t, for any function f with $f(t) \to \infty$ as $t \to \infty$, $\frac{t^{1/2}}{f(t)} \leq \Delta_1 \leq t^{1/2} f(t)$, and for $i = 2, \ldots, k$, $\frac{t^{1/2}}{f(t)} \leq \Delta_i \leq \Delta_{i-1} - \frac{t^{1/2}}{f(t)}$, with high probability (**whp**). We use this to show that at time t the largest k eigenvalues of the adjacency matrix of this graph have $\lambda_k = (1 \pm o(1))\Delta_k^{1/2}$ **whp**.

1 Introduction

Recently there has been much interest in understanding the properties of real-world large-scale networks such as the structure of the Internet and the World Wide Web. For a general introduction to this topic, see Bollobás and Riordan [BR02], Hayes [Hay00], or Watts [Wat99]. One approach is to model these networks by random graphs. Experimental studies by Albert, Barabási, and

[*]Supported in part by NSF VIGRE Grant DMS-9819950

[†]Supported in part by NSF Grant CCR-0200945

Jeong [ABJ99], Broder et al [BKM$^+$00], and Faloutsos, Faloutsos, and Faloutsos [FFF99] have demonstrated that in the World Wide Web/Internet the proportion of vertices of a given degree follows an approximate inverse power law i.e. the proportion of vertices of degree k is approximately $Ck^{-\alpha}$ for some constants C, α. The classical models of random graphs introduced by Erdős and Renyi [ER59] do not have power law degree sequences, so they are not suitable for modeling these networks. This has driven the development of various alternative models for random graphs.

One approach to remedy this situation is to study graphs with a prescribed degree sequence (or prescribed expected degree sequence). This is proposed as a model for the web graph by Aiello, Chung, and Lu in [ACL00]. Mihail and Papadimitriou also use this model [MP02] in their study of large eigenvalues, as do Chung, Lu, and Vu in [CLV03a, CLV03b].

An alternative approach, which we will follow in this paper, is to sample graphs via some generative procedure which yields a power law distribution. There is a long history of such models, outlined in the survey by Mitzenmacher [Mit]. We will use the preferential attachment model to generate our random graph. The preferential attachment random graph has been the subject of recently revived interest. It dates back to Yule [Yul25] and Simon [Sim55]. It was proposed as a model for the web by Barabási and Albert [BA99], and their description was elaborated by Bollobás and Riordan in [BR]. It was used by Bollobás, Riordan, Spencer, and Tusnády [BRST01] who proved that the degree sequence does follow a power law distribution. Bollobás and Riordan obtained several additional results regarding the diameter and connectivity of such graphs [BR]. We use the generative model of [BR] (see also [BRST01]) and build a graph sequentially as follows:

- At each time step t, we add a vertex v_t, and we add an edge from v_t to some other vertex u, where u is chosen at random according to the distribution:

$$\Pr[u = v_i] = \begin{cases} \frac{d_t(v_i)}{2t-1}, & \text{if } v_i \neq v_t; \\ \frac{1}{2t-1}, & \text{if } v_i = v_t; \end{cases}$$

  where $d_t(v)$ denotes the degree of vertex v at time t. This means that each vertex receives an additional edge with probability proportional to its current degree. The probability of choosing v_t (and forming a loop) is consistent with this, since we've already committed "half" an edge to v_t and are deciding where to put the other half.

- For some constant m, every m steps we contract the most recently added m vertices to form a supervertex.

Let G_t^m denote the random graph at time step t with contractions of size m. Note that contracting each set of vertices $\{im+1, im+2, \ldots, (i+1)m\}$ of G_t^1 yields a graph identically distributed with G_t^m.

It is worth mentioning that there are several alternative simple models for the World Wide Web and for general power law graphs. A generalization of the preferential attachment model is described by Drinea, Enachescu, and Mitzenmacher in [DEM01], and degree sequence results analogous to [BRST01] are proved for this model by Buckley and Osthus in [BO01]. A completely different generative model, based on the idea that new webpages are often consciously or unconsciously copies of existing pages, is developed by Kleinberg et al and Kumar et al in [KKR$^+$99], [KRRT99], [KRR$^+$00b], [KRR$^+$00a]. Cooper and Frieze analyze a model combining these approaches in [CF01].

Several previous results have studied the structure of low degree vertices in the preferential attachment graph. For example, the results in [BRST01] concern degrees up to $t^{1/15}$. The maximum degree vertex of the preferential attachment graph is the subject of Theorem 17 of [BR02], where an elegant static description of the preferential attachment graph is used to show that $\Delta_1/\sqrt{t}$ converges in distribution to a certain non-negative distribution. The technique used there extends to give the asymptotic distribution of $\Delta_i/\sqrt{t}$ for any constant i. Our first theorem also deals with the highest degree vertices:

Theorem 1 *Let m and k be fixed positive integers, and let $f(t)$ be a function with $f(t) \to \infty$ as $t \to \infty$. Let $\Delta_1 \geq \Delta_2 \geq \cdots \geq \Delta_k$ denote the degrees of the k highest degree vertices of G_t^m. Then*

$$\frac{t^{1/2}}{f(t)} \leq \Delta_1 \leq t^{1/2} f(t)$$

and for $i = 2, \ldots, k$,

$$\frac{t^{1/2}}{f(t)} \leq \Delta_i \leq \Delta_{i-1} - \frac{t^{1/2}}{f(t)},$$

whp[1].

Unfortunately, the slowly growing function $f(t)$ in the result above cannot be removed. Indeed, Theorem 17 of [BR02] and its extension to the k largest degrees shows that for any constants $a < b$ we have

$$\lim_{t\to\infty} \Pr\left[\Delta_1 \in (at^{1/2}, bt^{1/2})\right] > 0 \quad \text{and} \quad \lim_{t\to\infty} \Pr\left[\Delta_i - \Delta_{i-1} \in (at^{1/2}, bt^{1/2})\right] > 0.$$

The next theorem relates maximum eigenvalues and maximum degrees. It mirrors results of Mihail and Papadimitriou [MP02] and Chung, Liu and Vu [CLV03a, CLV03b] for fixed degree expectation models and at a high level, the proof follows the same lines as these two papers. Experimentally, a power law distribution for eigenvalues was observed in "real-world" graphs in [FFF99].

Theorem 2 *Let m and k be fixed positive integers, and let $f(t)$ be a function with $f(t) \to \infty$ as $t \to \infty$. Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k$ be the k largest eigenvalues of the adjacency matrix of G_t^m. Then for $i = 1, \ldots, k$ we have $\lambda_i = (1 \pm o(1))\Delta_i^{1/2}$* **whp**.

Our proofs of these theorems require two lemmas.

Lemma 1 *Let $d_t^m(s)$ denote the degree of vertex s in G_t^m, and let $a^{(k)} = a(a+1)(a+2)\cdots(a+k-1)$ denote the rising factorial function. Then for any positive integer k,*

$$\mathrm{E}\left[(d_t^m(s))^{(k)}\right] \leq (2m)^{(k)} 2^{k/2} \left(\frac{t}{s}\right)^{k/2}.$$

[1] In this paper an event $\mathcal{E}$ is said to hold *with high probability* (**whp**) if $\Pr[\mathcal{E}] \to 1$ as $t \to \infty$.

To simplify the exposition, we speak of a *supernode*, which is simply a collection of vertices viewed as one vertex. So the degree of a supernode is the sum of the degrees of the vertices in the supernode, and an edge is incident to a supernode if it is incident to some vertex in the supernode.

Lemma 2 *Let* $\mathbf{S} = (S_1, S_2, \ldots, S_\ell)$ *be a collection of disjoint supernodes, and let* $p_{\mathbf{S}}(\mathbf{r}; \mathbf{d}, t_0, t)$ *denote the probability that each supernode* S_i *has degree* $r_i + d_i$ *at time* t *conditioned on* $d_{t_0}(S_i) = d_i$*. Let* $d = \sum_{i=1}^{\ell} d_i$ *and* $r = \sum_{i=1}^{\ell} r_i$*. If* $d = o(t^{1/2})$ *and* $r = o(t^{2/3})$*, then*

$$p_{\mathbf{S}}(\mathbf{r}; \mathbf{d}, t_0, t) \leq \left(\prod_{i=1}^{\ell} \binom{r_i + d_i - 1}{d_i - 1} \right) \left(\frac{t_0 + 1}{t} \right)^{d/2} \exp\left\{ 2 + t_0 - \frac{d}{2} + \frac{2r}{t^{1/2}} \right\}.$$

In the next section we prove Lemma 1 and Theorems 1 and 2. We defer the proof of Lemma 2 (which consists of carefully bounding a sum) to the appendix.

2 Proofs

2.1 Proof of Lemma 1

An earlier version of the paper bounded $\mathrm{E}\left[(d_t^m(s))^k\right]$. This was a quite involved calculation. One of the referees suggested that we bound $\mathrm{E}\left[(d_t^m(s))^{(k)}\right]$ because this would be simpler using an idea from [BR02]. This is indeed the case, as the reader can see next.

Let $Z_t = d_t^m(s)$ denote the degree of vertex s at time t (when the graph contains t edges), and let Y_t be an indicator variable for the event that the edge added at time t is incident to s.

Then we have

$$\begin{aligned}
\mathrm{E}\left[Z_t^{(k)}\right] &= \mathrm{E}\left[\mathrm{E}\left[(Z_{t-1} + Y_t)^{(k)} \right] \,\middle|\, Z_{t-1} \right] \\
&= \mathrm{E}\left[Z_{t-1}^{(k)} \left(1 - \frac{Z_{t-1}}{2t-1}\right) + (Z_{t-1}+1)^{(k)} \left(\frac{Z_{t-1}}{2t-1}\right) \right] \\
&= \left(1 + \frac{k}{2t-1}\right) \mathrm{E}\left[Z_{t-1}^{(k)}\right].
\end{aligned}$$

Since $Z_s^{(k)} \leq (2m)^{(k)}$, we have

$$\mathrm{E}\left[Z_t^{(k)}\right] \leq (2m)^{(k)} \prod_{t'=s+1}^{t} \left(1 + \frac{k}{2t'-1}\right) \leq (2m)^{(k)} \exp\left\{ \frac{k}{2} \sum_{t'=s+1}^{t} \frac{1}{t' - \frac{1}{2}} \right\}.$$

We upper bound the sum with an integral,

$$\sum_{t'=s+1}^{t} \frac{1}{t' - \frac{1}{2}} \leq \int_{x=s}^{t} \frac{1}{t' - \frac{1}{2}} dx = \log \frac{t - \frac{1}{2}}{s - \frac{1}{2}},$$

and the bound on the expectation becomes

$$\mathrm{E}\left[Z_t^{(k)}\right] \leq (2m)^{(k)} \left(\frac{t-\frac{1}{2}}{s-\frac{1}{2}}\right)^{k/2} = (2m)^{(k)} \left(\frac{t}{s}\right)^{k/2} \left(\frac{2-1/t}{2-1/s}\right)^{k/2}.$$

Since $\frac{2-1/t}{2-1/s} \leq 2$, we may conclude that

$$\mathrm{E}\left[Z_t^{(k)}\right] \leq (2m)^{(k)} 2^{k/2} \left(\frac{t}{s}\right)^{k/2}.$$

□

2.2 Proof of Theorem 1

We partition the vertices into those added before time t_0, before t_1, and after t_1 and argue about the maximum degree of vertices in each set. Here

$$t_0 = \log\log\log f(t) \text{ and } t_1 = \log\log f(t).$$

We break the proof of Theorem 1 into 5 Claims.

Claim 1 *In G_t^m the degree of the supernode of vertices added before time t_0 is at least $t_0^{1/3} t^{1/2}$* **whp**.

Proof Let $\mathcal{A}_1$ denote the event that the supernode consisting of the first t_0 vertices has degree less than $t_0^{1/3} t^{1/2}$. We bound the probability of this event using Lemma 2 with $\ell = 1$. Since at time t_0 the supernode of all vertices added by this time has all of the edges, we take $\mathbf{d} = d_1 = 2t_0$. Then

$$\begin{aligned}
\Pr[\mathcal{A}_1] &\leq \sum_{r_1=0}^{t_0^{1/3} t^{1/2} - 2t_0} \binom{r_1 + 2t_0 - 1}{2t_0 - 1} \left(\frac{t_0+1}{t}\right)^{d/2} e^{2 + t_0 - d/2 + 2r/t^{1/2}} \\
&\leq (t_0^{1/3} t^{1/2}) \frac{(t_0^{1/3} t^{1/2})^{2t_0 - 1}}{(2t_0 - 1)!} \left(\frac{t_0+1}{t}\right)^{t_0} e^{2 + t_0 + 2t_0^{1/3}} \\
&\leq t_0^{2t_0/3} \frac{e^{2t_0 - 1}}{(2t_0 - 1)^{2t_0 - 1}} (t_0 + 1)^{t_0} e^{2 + t_0 + 2t_0^{1/3}} \\
&\leq \frac{e^{3t_0 + 2t_0^{1/3} + 2}}{(2t_0 - 1)^{t_0/3 - 1}} \\
&= o(1).
\end{aligned}$$

□

Claim 2 *In G_t^m no vertex added after time t_1 has degree exceeding $t_0^{-2} t^{1/2}$* **whp**.

Proof Let $\mathcal{A}_2$ denote the event that some vertex added after time t_1 has degree exceeding $t_0^{-2}t^{1/2}$. Then we have

$$\Pr[\mathcal{A}_2] \le \sum_{s=t_1}^{t} \Pr[d_t(s) \ge t_0^{-2}t^{1/2}] = \sum_{s=t_1}^{t} \Pr\left[(d_t(s))^{(3)} \ge \left(t_0^{-2}t^{1/2}\right)^{(3)}\right] \le \sum_{s=t_1}^{t} t_0^6 t^{-3/2}\,\mathrm{E}\left[d_t(s)^{(3)}\right]$$

Using Lemma 1 this bound becomes

$$\Pr[\mathcal{A}_2] \le \sum_{s=t_1}^{t} t_0^6 t^{-3/2}(2m)^{(3)}2^{3/2}\left(\frac{t}{s}\right)^{3/2} \le (2m)^{(3)}2^{3/2}t_0^6 \sum_{s=t_1}^{t} s^{-3/2} \le (2m)^{(3)}2^{3/2}t_0^6 t_1^{-1/2} = o(1).$$

□

Claim 3 *In G_t^m no vertex added before time t_1 has degree exceeding $t_0^{1/6}t^{1/2}$* **whp***.*

Proof Let $\mathcal{A}_3$ denote the event that some vertex added before t_1 has degree exceeding $t_0^{1/6}t^{1/2}$. We use Lemma 1 for a third moment argument as above.

$$\Pr[\mathcal{A}_3] \le \sum_{s=1}^{t_1} \left(t_0^{1/6}t^{1/2}\right)^{-3}(2m)^{(3)}2^{3/2}\left(\frac{t}{s}\right)^{3/2} = (2m)^{(3)}2^{3/2}t_0^{-1/2}\sum_{s=1}^{t_1} s^{-3/2} \le (2m)^{(3)}2^{3/2}t_0^{-1/2}$$
$$= o(1).$$

□

Claim 4 *The k highest degree vertices of G_t^m are added before time t_1 and have degree Δ_i bounded by $t_0^{-1}t^{1/2} \le \Delta_i \le t_0^{1/6}t^{1/2}$* **whp***.*

Proof

(Upper bound on Δ_i) By Claim 2, all vertices added after time t_1 have degree at most $t_0^{-2}t^{1/2}$ **whp**. Combining this with Claim 3 we have $\Delta_1 \le t_0^{1/6}t^{1/2}$ **whp**.

(Lower bound on Δ_i) The conditions from Claims 1,2, and 3 imply the lower bound. To see this, suppose the conditions of these claims are satisfied, but assume for contradiction that at most $k-1$ vertices added before t_1 have degree exceeding $t_0^{-1}t^{1/2}$. Then the total degree of vertices added before t_0 is less than $k(t_0^{1/6}t^{1/2}) + t_0(t_0^{-1}t^{1/2}) \le 2kt_0^{1/6}t^{1/2}$. But this contradicts the condition of Claim 1, which says the total degree of vertices added before t_0 at least $t_0^{1/3}t^{1/2}$.

(Added before t_1) By Claim 2 all vertices added after time t_1 have degree at most $t_0^{-2}t^{1/2}$ **whp**. So the lower bound on Δ_i shows the k highest degree vertices are added before time t_1 **whp**.

□

Claim 5 *The k highest degree vertices of G_t^m have $\Delta_i \leq \Delta_{i-1} - t^{1/2}/f(t)$* **whp**.

Proof Let $\mathcal{A}_4$ denote the event that there are 2 vertices among the first t_1 with degrees exceeding $t_0^{-1}t^{1/2}$ and within $t^{1/2}/f(t)$ of each other.

Let $p_{\ell,s_1,s_2} = \Pr[d_t(s_1) - d_t(s_2) = \ell \mid \overline{\mathcal{A}_3}]$, for $|\ell| \leq \sqrt{t}/f(t)$. Then

$$\Pr[\mathcal{A}_4 \mid \overline{\mathcal{A}_3}] \leq \sum_{1\leq s_1<s_2\leq t_1} \sum_{\ell=-t^{1/2}/f(t)}^{t^{1/2}/f(t)} p_{\ell,s_1,s_2}.$$

Since

$$\begin{aligned}
p_{\ell,s_1,s_2} &\leq \sum_{r_1=t_0^{-1}t^{1/2}}^{t_0^{1/6}t^{1/2}} \sum_{d_1,d_2=1}^{2t_1} p_{(s_1,s_2)}((r_1, r_1-\ell); (d_1,d_2), t_1, t) \\
&\leq t_0^{1/6}t^{1/2} \sum_{d_1,d_2=1}^{2t_1} \binom{2t_0^{1/6}t^{1/2}}{d_1-1}\binom{2t_0^{1/6}t^{1/2}}{d_2-1}\left(\frac{t_1+1}{t}\right)^{(d_1+d_2)/2} e^{t_0+2+4t_0^{1/6}} \\
&\leq t_0^{1/6}t^{1/2} \sum_{d_1,d_2=1}^{2t_1} \left(2t_0^{1/6}t^{1/2}\right)^{d_1+d_2-2} (t_1+1)^{2t_1} t^{-(d_1+d_2)/2} e^{2t_0} \\
&\leq t_0^{1/6}(2t_1)^2 2^{4t_1} t_0^{2t_1/3}(t_1+1)^{2t_1} e^{2t_0} t^{-1/2} \\
&= o(t_1^{-2}t^{-1/2}f(t)),
\end{aligned}$$

we have

$$\Pr[\mathcal{A}_4 \mid \overline{\mathcal{A}_3}] \leq \sum_{1\leq s_1<s_2\leq t_1} \sum_{\ell=-t^{1/2}/f(t)}^{t^{1/2}/f(t)} p_{\ell,s_1,s_2} = o(1).$$

So

$$\Pr[\mathcal{A}_4] = \Pr[\mathcal{A}_4 \mid \mathcal{A}_3]\Pr[\mathcal{A}_3] + \Pr[\mathcal{A}_4 \mid \overline{\mathcal{A}_3}]\Pr[\overline{\mathcal{A}_3}] \leq \Pr[\mathcal{A}_3] + \Pr[\mathcal{A}_4 \mid \overline{\mathcal{A}_3}] = o(1).$$

□

2.3 Proof of Theorem 2

We partition the vertices into 3 sets; let S_i be the vertices added after time t_{i-1} and at or before time t_i, for

$$t_0 = 0, \quad t_1 = t^{1/8}, \quad t_2 = t^{9/16}, \quad t_3 = t.$$

To reduce the number of subscripts necessary, we use G to denote the graph G_t.

For any graph H, we let M_H denote the adjacency matrix of H, and we let $\lambda_i(H)$ denote the i-th largest eigenvalue of M_H. We will use the identity (*Rayleigh's Principle*)

$$\lambda_i(H) = \min_L \max_{\mathbf{x}\in L, \mathbf{x}\neq 0} \frac{\mathbf{x}^T M_H \mathbf{x}}{\mathbf{x}^T\mathbf{x}} \tag{1}$$

where L ranges over all $(n-i+1)$-dimensional subspaces of $\mathbb{R}^n$. (See, for example, [Str88]).

Our approach, as in [MP02], [CLV03a, CLV03b], is to show that **whp** G contains a star forest F with stars of degree asymptotic to the maximum degree vertices of G. Then we will show $G \setminus F$ has small eigenvalues. Then Rayleigh's Principle is sufficient to conclude that the large eigenvalues of G cannot be too different from the large eigenvalues of F.

To do this, we need reasonable bounds on the degrees and codegrees in G. Recall that $d_s^m(r)$ is the degree at time s of the vertex added at time r with contractions of size m.

Claim 6 *For any $\epsilon > 0$ and any $f(t)$ with $f(t) \to \infty$ as $t \to \infty$ the following holds* **whp**: *for all s with $f(t) \le s \le t$, for all vertices $v \in G_s^m$, if v was added at time r, then $d_s^m(v) \le s^{1/2+\epsilon} r^{-1/2}$.*

Proof We use Lemma 1 and the union bound. Let $\ell = \lceil 3/\epsilon \rceil$.

$$\begin{aligned}
\Pr\left[\bigcup_{s=f(t)}^{t}\bigcup_{r=1}^{s}\{d_s^m(r) \ge s^{1/2+\epsilon}r^{-1/2}\}\right] &\le \sum_{s=f(t)}^{t}\sum_{r=1}^{s}\Pr\left[d_s^m(r) \ge s^{1/2+\epsilon}r^{-1/2}\right] \\
&= \sum_{s=f(t)}^{t}\sum_{r=1}^{s}\Pr\left[(d_s^m(r))^{(\ell)} \ge \left(s^{1/2+\epsilon}r^{-1/2}\right)^{(\ell)}\right] \\
&\le \sum_{s=f(t)}^{t}\sum_{r=1}^{s} s^{-\ell(1/2+\epsilon)} r^{\ell/2}\, \mathrm{E}\left[(d_s^m(r))^{\ell}\right] \\
&\le \sum_{s=f(t)}^{t}\sum_{r=1}^{s} s^{-\ell(1/2+\epsilon)} r^{\ell/2} (2m)^{(\ell)} 2^{\ell/2} (s/r)^{\ell/2} \\
&= (2m)^{(\ell)} 2^{\ell/2} \sum_{s=f(t)}^{t} s^{1-\epsilon\ell}.
\end{aligned}$$

Since $\ell \ge 3/\epsilon$,

$$\sum_{s=f(t)}^{t} s^{1-\epsilon\ell} \le \int_{f(t)-1}^{\infty} x^{1-\epsilon\ell} dx = \frac{1}{\epsilon\ell - 2}(f(t)-1)^{2-\epsilon\ell} = o(1).$$

□

Claim 7 *Let S_3' be the set of vertices in S_3 which are adjacent to more than 1 vertex of S_1 in G. Then $|S_3'| \le t^{7/16}$* **whp**.

Proof Let $\mathcal{B}_1$ be the event that the conditions of Claim 6 hold with $f(t) = t_2$ and $\epsilon = 1/16$. Then for a vertex $v \in S_3$ added at time s,

$$\Pr[|N(v) \cap S_1| \ge 2 \mid \mathcal{B}_1] \le \binom{m}{2}\left(\frac{s^{1/2+\epsilon}t_1}{2s-1}\right)^2 \le m^2 s^{-7/8} t^{1/4}.$$

Let X denote the number of $v \in S_3$ adjacent to more than 1 vertex of S_1. Then

$$\mathrm{E}\left[X \mid \mathcal{B}_1\right] \leq \sum_{s=t_2+1}^{t} m^2 s^{-7/8} t^{1/4} \leq m^2 t^{1/4} \int_{t_2}^{t} x^{-7/8} dx \leq 8m^2 t^{3/8}.$$

We finish the claim with Markov's inequality,

$$\Pr[X \geq t^{7/16} \mid \mathcal{B}_1] \leq E[X \mid \mathcal{B}_1]/t^{7/16} = o(1).$$

$\square$

Now, let $F \subseteq G$ be the star forest consisting of edges between S_1 and $S_3 \setminus S_3'$.

Claim 8 *Let $\Delta_1 \geq \Delta_2 \geq \cdots \geq \Delta_k$ denote the degrees of the k highest degree vertices of G. Then $\lambda_i(F) = (1 - o(1))\Delta_i^{1/2}$* **whp**.

Proof Let H be the star forest $H = K_{1,d_1} \cup K_{1,d_2} \cup \cdots \cup K_{1,d_k}$, with $d_1 \geq d_2 \geq \cdots \geq d_k$. Then for $i = 1, \ldots, k$, $\lambda_i(H) = d_i^{1/2}$. So it is sufficient to show that $\Delta_i(F) = (1 - o(1))\Delta_i(G)$ for $i = 1, \ldots, k$.

Claim 4 shows that the k highest degree vertices of G are added before time t_1, so these vertices are all in F. The only edges to these vertices that are not in F are those added before time t_2 and those incident to S_3'. By Theorem 1 we have $\Delta_1(G_{t_2}^m) \leq t_2^{7/9} = t^{7/16}$ and, also by Theorem 1, $\Delta_i(G) \geq t^{1/2}/\log t$ for $i = 1, \ldots, k$, **whp**. Claim 7 says that **whp** $|S_3'| \leq t^{7/16}$, and so **whp**

$$\Delta_i(F) \geq \Delta_i(G) - t^{7/16} - mt^{7/16} = (1 - o(1))\Delta_i(G).$$

$\square$

Let $H = G \setminus F$. We now show that $\lambda_1(H)$ is $o(\lambda_k(F))$. This completes the proof of Theorem 2 because, for any subspace L we have

$$\max_{\mathbf{x} \in L, \mathbf{x} \neq 0} \frac{\mathbf{x}^T M_G \mathbf{x}}{\mathbf{x}^T \mathbf{x}} = \max_{\mathbf{x} \in L, \mathbf{x} \neq 0} \frac{\mathbf{x}^T M_F \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \pm O\left(\max_{\mathbf{x} \neq 0} \frac{\mathbf{x}^T M_H \mathbf{x}}{\mathbf{x}^T \mathbf{x}}\right),$$

and so, for $i \leq k$, Rayleigh's Principle gives $\lambda_i(G) = \lambda_i(F)(1 \pm o(1))$.

Claim 9 $\lambda_1(H) \leq 6mt^{15/64}$ **whp**.

Proof We bound the eigenvalues of H in 6 parts. Let

$$H_i = H[S_i], \qquad H_{ij} = H(S_i, S_j),$$

where $H[S]$ is the subgraph of H induced by the vertex set S, and $H(S,T)$ is the subgraph containing only edges with one vertex in S and the other in T.

To bound $\lambda_1(H_i)$ we use the fact that the maximum eigenvalue of a graph is at most the maximum degree of the graph. This is easily verified from (1).

We use Claim 6 with $f(t) = t_1$ and $\epsilon = 1/64$ to conclude that **whp**

$$\begin{array}{lllll} \lambda_1(H_1) & \le \Delta_1(H_1) & = \max\limits_{v \le t_1}\{d_{t_1}^m(v)\} & \le t_1^{1/2+\epsilon} & = t^{33/512}, \\ \lambda_1(H_2) & \le \Delta_1(H_2) & \le \max\limits_{t_1 \le v \le t_2}\{d_{t_2}^m(v)\} & \le t_2^{1/2+\epsilon} t_1^{-1/2} & = t^{233/1024}, \\ \lambda_1(H_3) & \le \Delta_1(H_3) & \le \max\limits_{t_2 \le v \le t_3}\{d_{t_3}^m(v)\} & \le t_3^{1/2+\epsilon} t_2^{-1/2} & = t^{15/64}. \end{array}$$

To bound $\lambda_1(H_{ij})$, we begin by considering the case $m = 1$. Then, for $i < j$, each vertex in S_j has at most 1 edge in H_{ij}, so H_{ij} is a star forest. As observed in Claim 8, the eigenvalues of a star forest are directly related to the degrees of the stars.

When $m > 1$, we let G' denote a preferential attachment graph with t edges and $m = 1$. Recall that by contracting vertices $\{(i-1)m+1, \ldots, im\}$ into a single vertex i, we obtain a graph identically distributed with G. There is a simple representation of this observation in terms of linear algebra: we can write the adjacency matrix of G in terms of the adjacency matrix of the graph G':

$$M_G = C_m^T M_{G'} C_m,$$

where C_m is the $t \times t/m$ matrix with i-th column

$$[\ \underbrace{0 \ \cdots \ 0}_{(i-1)m} \ \ \underbrace{1 \ \cdots \ 1}_{m} \ \ \underbrace{0 \ \cdots \ 0}_{(t/m-i)m}\ ]^T.$$

Similarly, we can write the adjacency matrix of H_{ij} in terms of the adjacency matrix of H'_{ij} using this "contraction matrix" C_m.

Note that for $\mathbf{y} = C_m\mathbf{x}$ we have $\mathbf{y}^T\mathbf{y} = m(\mathbf{x}^T\mathbf{x})$. So

$$\begin{aligned} \lambda_1(H_{ij}) = \max_{\mathbf{x}\neq 0} \frac{\mathbf{x}^T M_{H_{ij}} \mathbf{x}}{\mathbf{x}^T\mathbf{x}} = \max_{\mathbf{x}\neq 0} \frac{\mathbf{x}^T C_m^T M_{H'_{ij}} C_m \mathbf{x}}{\mathbf{x}^T\mathbf{x}} &= \max_{\mathbf{y}:\ \mathbf{y}=C_m\mathbf{x}\neq 0} m \frac{\mathbf{y}^T M_{H'_{ij}} \mathbf{y}}{\mathbf{y}^T\mathbf{y}} \\ &\le m \max_{\mathbf{y}\neq 0} \frac{\mathbf{y}^T M_{H'_{ij}} \mathbf{y}}{\mathbf{y}^T\mathbf{y}} = m\lambda_1(H'_{ij}). \end{aligned}$$

We use Claim 6 with $f(t) = t_1$ and $\epsilon = 1/64$ as above to conclude that **whp**

$$\begin{array}{llll} \Delta_1(H'_{12}) & = \max\limits_{v \le t_2}\{d_{t_2}^1(v)\} & \le t_2^{1/2+\epsilon} & = t^{297/1024} \\ \Delta_1(H'_{23}) & = \max\limits_{t_1 \le v \le t_3}\{d_{t_3}^1(v)\} & \le t_3^{1/2+\epsilon} t_1^{-1/2} & = t^{29/64} \end{array}$$

Finally, all edges in H'_{13} are between S_1 and S'_3, so Claim 7 shows that $\Delta_1(H'_{13}) \le t^{7/16}$ **whp**.

We now conclude that **whp**

$$\lambda_1(H_{ij}) \le m\lambda_1(H'_{ij}) \le m\Delta_1(H'_{ij})^{1/2} \le mt^{15/64},$$

and so **whp**

$$\lambda_1(H) \le \sum_{i=1}^{3} \lambda_1(H_i) + \sum_{i<j} \lambda_1(H_{ij}) \le 6mt^{15/64}.$$

□

Acknowledgement: We would like to thank the referee whose suggestion concerning Lemma 1 has considerably simplified part of the paper.

References

[ABJ99] R. Albert, A. Barabási, and H. Jeong. Diameter of the world wide web. *Nature*, 401:103–131, 1999.

[ACL00] W. Aiello, F. R. K. Chung, and L. Lu. A random graph model for massive graphs. In *Proc. of the 32nd Annual ACM Symposium on the Theory of Computing*, pages 171–180, 2000.

[BA99] A. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286:509–512, 1999.

[BKM+00] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener. Graph structure in the web. In *Proc. of the 9th Intl. World Wide Web Conference*, pages 309–320, 2000.

[BO01] G. Buckley and D. Osthus. Popularity based random graph models leading to a scale-free degree distribution, 2001.

[BR] B. Bollobás and O. Riordan. The diameter of a scale-free random graph.

[BR02] B. Bollobás and O. Riordan. Mathematical results on scale-free random graphs. In *Handbook of Graphs and Networks.* Wiley-VCH, Berlin, 2002.

[BRST01] B. Bollobás, O. Riordan, J. Spencer, and G. Tusanády. The degree sequence of a scale-free random graph process. *Random Structures and Algorithms*, 18:279–290, 2001.

[CF01] C. Cooper and A. M. Frieze. A general model of undirected Web graphs. In *Proc. of ESA*, pages 500–511, 2001.

[CLV03a] F.R.K. Chung, L. Lu, and V. Vu. Eigenvalues of random power law graphs. *Annals of Combinatorics*, 7:21–33, 2003.

[CLV03b] F.R.K. Chung, L. Lu, and V. Vu. The spectra of random graphs with expected degrees. *Proceedings of national Academy of Sciences*, 100(11):6313–6318, 2003.

[DEM01] E. Drinea, M. Enachescu, and M. Mitzenmacher. Variations on random graph models for the web. Technical report, Harvard University, 2001.

[ER59] P. Erdös and A. Rényi. On random graphs I. *Publicationes Mathematicae Debrecen*, 6:290–297, 1959.

[FFF99] M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the internet topology. In *SIGCOMM*, pages 251–262, 1999.

[Hay00] B. Hayes. Graph theory in practice: Part II. *American Scientist*, 88:104–109, 2000.

[KKR+99] J. M. Kleinberg, R. Kumar, P. Raghavan, S. Rajagopalan, and A. S. Tomkins. The Web as a graph: Measurements, models and methods. *Lecture Notes in Computer Science*, 1627:1–??, 1999.

[KRR+00a] R. Kumar, P. Raghavan, S. Rajagopalan, D. Sivakumar, A. Tomkins, and E. Upfal. Stochastic models for the web graph. In *FOCS: IEEE Symposium on Foundations of Computer Science (FOCS)*, 2000.

[KRR+00b] R. Kumar, P. Raghavan, S. Rajagopalan, D. Sivakumar, A. Tomkins, and E. Upfal. The Web as a graph. In *Proc. 19th ACM SIGACT-SIGMOD-AIGART Symp. Principles of Database Systems, PODS*, pages 1–10. ACM Press, 15–17 2000.

[KRRT99] R. Kumar, P. Raghavan, S. Rajagopalan, and A. Tomkins. Trawling the Web for emerging cyber-communities. *Computer Networks (Amsterdam, Netherlands: 1999)*, 31(11–16):1481–1493, 1999.

[Mit] M. Mitzenmacher. A brief history of generative models for power law and lognormal distributions.

[MP02] M. Mihail and C. H. Papadimitriou. On the eigenvalue power law. In *Randomization and Approximation Techniques, 6th International Workshop*, pages 254–262, 2002.

[Sim55] H. A. Simon. On a class of skew distribution functions. *Biometrika*, 42(3/4):425–440, 1955.

[Str88] G. Strang. *Linear algebra and its applications.* Hardcourt Brace Jovanovich Publishing, 1988.

[Wat99] D. J. Watts. *Small Worlds: The Dynamics of Networks between Order and Randomness.* Princeton: Princeton University Press, 1999.

[Yul25] G. Yule. A mathematical theory of evolution based on the conclusions of Dr. J.C. Willis. *Philosophical Transactions of the Royal Society of London (Series B)*, 213:21–87, 1925.

A Proof of Lemma 2

We calculate the probability as the union of disjoint events by fixing the times when the degrees of the S_i change. Let $\tau^{(i)} = (\tau_1^{(i)}, \ldots, \tau_{r_i}^{(i)})$ where $\tau_j^{(i)}$ is the time when we add an edge incident to S_i and increase the degree of S_i from $d_i + j - 1$ to $d_i + j$. We will see that in the calculation it doesn't matter much which S_i increases in degree, so we let $d = \sum_{i=1}^{\ell} d_i$ and $r = \sum_{i=1}^{\ell} r_i$ and define $\tau = (\tau_0, \tau_1, \ldots, \tau_{r+1})$ to be the ordered union of the $\tau^{(i)}$, with $\tau_0 = t_0$ and $\tau_{r+1} = t$.

Let $p(\tau; \mathbf{d}, t_0, t)$ denote the probability that (super)nodes S_i increase in degree at exactly the times specified by τ between time t_0 and t given $d_{t_0}(s_i) = d_i$. Then

$$
\begin{aligned}
p(\tau; \mathbf{d}, t_0, t) &= \left(\prod_{i=1}^{\ell} \prod_{k=1}^{r_i} \frac{d_i + k - 1}{2\tau_k^{(i)} - 1} \right) \left(\prod_{k=0}^{r} \prod_{j=\tau_k+1}^{\tau_{k+1}-1} \left(1 - \frac{d+k}{2j-1} \right) \right) \\
&= \left(\prod_{i=1}^{\ell} \frac{(r_i + d_i - 1)!}{(d_i - 1)!} \right) \left(\prod_{k=1}^{r} \frac{1}{2\tau_k - 1} \right) \\
&\qquad \times \exp\left\{ \sum_{k=0}^{r} \sum_{j=\tau_k+1}^{\tau_{k+1}-1} \log\left(1 - \frac{d+k}{2j-1} \right) \right\}.
\end{aligned}
$$

We bound the inner sum by an integral

$$
\sum_{j=\tau_k+1}^{\tau_{k+1}-1} \log\left(1 - \frac{d+k}{2j-1}\right) \leq \sum_{j=\tau_k+1}^{\tau_{k+1}-1} \log\left(1 - \frac{d+k}{2j}\right) \leq \int_{\tau_k+1}^{\tau_{k+1}} \log\left(1 - \frac{d+k}{2x}\right) dx.
$$

Then, since

$$
\int \log\left(1 - \frac{d+k}{2x}\right) dx = -x\log(2x) + \frac{2x - (d+k)}{2} \log(2x - (d+k)),
$$

we have

$$
\begin{aligned}
&\int_{\tau_k+1}^{\tau_{k+1}} \log\left(1 - \frac{d+k}{2x}\right) dx \\
&\quad = -\tau_{k+1}\log(2\tau_{k+1}) + \frac{2\tau_{k+1} - (d+k)}{2} \log(2\tau_{k+1} - (d+k)) \\
&\qquad + (\tau_k + 1)\log(2\tau_k + 2) - \frac{2\tau_k + 2 - (d+k)}{2} \log(2\tau_k + 2 - (d+k)).
\end{aligned}
$$

By grouping like terms and noting that $\tau_0 = t_0$ and $\tau_{r+1} = t$, we have

$$
\begin{aligned}
&\sum_{k=0}^{r} \int_{\tau_k+1}^{\tau_{k+1}} \log\left(1 - \frac{d+k}{2x}\right) dx \\
&\quad = (t_0 + 1)\log(2t_0 + 2) - \frac{2t_0 + 2 - d}{2} \log(2t_0 + 2 - d) \\
&\qquad - t\log(2t) + \frac{2t - (d+r)}{2} \log(2t - (d+r)) \\
&\qquad + \sum_{k=1}^{r} \Bigg((\tau_k + 1)\log(2\tau_k + 2) - \frac{2\tau_k + 2 - (d+k)}{2} \log(2\tau_k + 2 - (d+k)) \\
&\qquad\qquad - \tau_k \log(2\tau_k) + \frac{2\tau_k - (d+k-1)}{2} \log(2\tau_k - (d+k-1)) \Bigg) \\
&\quad = A + \sum_{k=1}^{r} B_k,
\end{aligned}
$$

where A is the term outside the summation and B_k is the k-th term of the sum.

We concentrate first on the term B_k. Rearranging terms yields

$$B_k = \tau_k \log(1+1/\tau_k) + \log(2\tau_k+2) + \frac{2\tau_k+2-(d+k)}{2}\log\left(1-\frac{1}{2\tau_k+2-(d+k)}\right) - \frac{1}{2}\log(2\tau_k+1-(d+k)).$$

Since $1+x \le e^x$, this is bounded as

$$B_k \le \frac{1}{2}\log(2\tau_k+2) - \frac{1}{2}\log\left(1-\frac{d+k+1}{2\tau_k+2}\right) + \frac{1}{2}.$$

Now we turn our attention to A. Rearranging terms, we have

$$A = -(t_0+1)\log\left(1-\frac{d}{2t_0+2}\right) + \frac{d}{2}\log(2t_0+2-d) + t\log\left(1-\frac{d+r}{2t}\right) - \frac{d+r}{2}\log(2t-(d+r))$$

So

$$\begin{aligned} e^A &= \left(1-\frac{d}{2t_0+2}\right)^{-(t_0+1)} (2t_0+2-d)^{d/2} \left(1-\frac{d+r}{2t}\right)^{t} (2t-(d+r))^{-(d+r)/2} \\ &= \left(1-\frac{d}{2(t_0+1)}\right)^{-\left(1-\frac{d}{2(t_0+1)}\right)(t_0+1)} \left(1-\frac{d+r}{2t}\right)^{t-(d+r)/2} \left(\frac{t_0+1}{t}\right)^{d/2} (2t)^{-r/2}. \end{aligned}$$

Since $1-x \le e^{-x-x^2/2}$ for $0<x<1$ we have

$$\left(1-\frac{d+r}{2t}\right)^{t-(d+r)/2} \le \exp\left\{-\frac{d+r}{2} + \frac{(d+r)^2}{8t} + \frac{(d+r)^3}{16t^2}\right\}.$$

So

$$\begin{aligned} e^{A+\sum B_k} &\le \left(1-\frac{d}{2(t_0+1)}\right)^{-\left(1-\frac{d}{2(t_0+1)}\right)(t_0+1)} \exp\left\{-\frac{d+r}{2} + \frac{(d+r)^2}{8t} + \frac{(d+r)^3}{16t^2}\right\} \\ &\quad \times \left(\frac{t_0+1}{t}\right)^{d/2} (2t)^{-r/2} \left(\prod_{k=1}^{r}\left(\left(1-\frac{d+k+1}{2\tau_k+2}\right)^{-1/2}(2\tau_k+2)^{1/2}\right)\right) e^{r/2} \\ &= err(r,d,t_0,t)\left(\frac{t_0+1}{t}\right)^{d/2} (2t)^{-r/2} \left(\prod_{k=1}^{r}\left(\left(1-\frac{d+k+1}{2\tau_k+2}\right)^{-1/2}(2\tau_k+2)^{1/2}\right)\right), \end{aligned}$$

where

$$err(r,d,t_0,t) = \left(1-\frac{d}{2(t_0+1)}\right)^{-\left(1-\frac{d}{2(t_0+1)}\right)(t_0+1)} \exp\left\{-\frac{d}{2} + \frac{(d+r)^2}{8t} + \frac{(d+r)^3}{16t^2}\right\}.$$

Inserting the bounds for $A+\sum B_k$ into the bound on $p(\tau;\mathbf{d},t_0,t)$, we have

$$\begin{aligned} p(\tau;\mathbf{d},t_0,t) &\le \left(\prod_{i=1}^{\ell}\frac{(r_i+d_i-1)!}{(d_i-1)!}\right) err(r,d,t_0,t)\left(\frac{t_0+1}{t}\right)^{d/2}(2t)^{-r/2} \\ &\quad \times \left(\prod_{k=1}^{r}\left(1-\frac{d+k+1}{2\tau_k+2}\right)^{-1/2}(2\tau_k+2)^{1/2}(2\tau_k-1)^{-1}\right). \end{aligned}$$

Now observe that

$$\left(1-\frac{d+k+1}{2\tau_k+2}\right)^{-1/2}(2\tau_k+2)^{1/2}(2\tau_k-1)^{-1}=(2\tau_k+1-(d+k))^{-1/2}\left(1+\frac{3}{2\tau_k-1}\right).$$

In order to bound the probability of interest, we sum $p(\tau;\mathbf{d},t_0,t)$ over all ordered choices of τ.

$$\begin{aligned}
p_{\mathbf{S}}(\mathbf{r};\mathbf{d},t_0,t) &= \sum_{\tau^{(1)},\ldots,\tau^{(\ell)}} p(\tau;\mathbf{d},t_0,t)\\
&\le \binom{r}{r_1,\ldots,r_\ell} \sum_{t_0+1\le\tau_1<\cdots<\tau_r\le t}\left(\prod_{i=1}^{\ell}\frac{(r_i+d_i-1)!}{(d_i-1)!}\right) err(r,d,t_0,t)\\
&\quad\times\left(\frac{t_0+1}{t}\right)^{d/2}(2t)^{-r/2}\left(\prod_{k=1}^{r}(2\tau_k+1-(d+k))^{-1/2}\left(1+\frac{3}{2\tau_k-1}\right)\right)\\
&= r!\left(\prod_{i=1}^{\ell}\binom{r_i+d_i-1}{d_i-1}\right) err(r,d,t_0,t)\left(\frac{t_0+1}{t}\right)^{d/2}(2t)^{-r/2}\\
&\qquad\times\sum_{t_0+1\le\tau_1<\tau_2\cdots<\tau_r\le t}\left(\prod_{k=1}^{r}(2\tau_k+1-(d+k))^{-1/2}\left(1+\frac{3}{2\tau_k-1}\right)\right).
\end{aligned}$$

Now we make a change of variables, introducing $\tau_k'=\tau_k-\lceil(d+k)/2\rceil$. For some τ_k,τ_{k+1}, this can result in $\tau_k'=\tau_{k+1}'$, so we relax the strict inequalities to less-than-or-equals. Also, since d and k are both at least 1, we have $2\lceil(d+k)/2\rceil\ge 2$. So

$$\begin{aligned}
&\sum_{t_0<\tau_1<\tau_2<\cdots<\tau_r\le t}\left(\prod_{k=1}^{r}(2\tau_k+1-(d+k))^{-1/2}\left(1+\frac{3}{2\tau_k-1}\right)\right)\\
&\qquad\le\sum_{(t_0-\lceil d/2\rceil+1)\le\tau_1'\le\tau_2'\le\cdots\le\tau_r'\le(t-\lceil(d+r)/2\rceil)}\left(\prod_{k=1}^{r}(2\tau_k'+1)^{-1/2}\left(1+\frac{3}{2\tau_k'+1}\right)\right)
\end{aligned}$$

We simplify this sum by unordering the variables,

$$\begin{aligned}
&\sum_{(t_0-\lceil d/2\rceil+1)\le\tau_1'\le\tau_2'\le\cdots\le\tau_r'\le(t-\lceil(d+r)/2\rceil)}\left(\prod_{k=1}^{r}\left((2\tau_k'+1)^{-1/2}+3\,(2\tau_k+1)^{-3/2}\right)\right)\\
&\qquad=\frac{1}{r!}\left(\sum_{\tau'=t_0-\lceil d/2\rceil+1}^{t-\lceil(d+r)/2\rceil}\left((2\tau'+1)^{-1/2}+3\left(2\tau'+1\right)^{-3/2}\right)\right)^{r},
\end{aligned}$$

and then using an integral, which we start from $x=0$, since $t_0 - \lceil d/2 \rceil + 1 \geq 1$,

$$\begin{aligned}
&\sum_{\tau'=t_0-\lceil d/2\rceil+1}^{t-\lceil (d+r)/2\rceil} (2\tau'+1)^{-1/2} + 3\,(2\tau'+1)^{-3/2} \\
&\qquad \leq \int_{x=0}^{t-(d+r)/2} \left((2x+1)^{-1/2} + 3\,(2x+1)^{-3/2}\right) dx \\
&\qquad \leq (2t-(d+r)+1)^{1/2} - 1 - 3(2t-(d+r)+1)^{-1/2} + 3) \\
&\qquad \leq (2t-(d+r)+1)^{1/2} + 2 \\
&\qquad = (2t)^{1/2}\left(1 - \frac{d+r-1}{2t}\right)^{1/2}\left(1 + \frac{2}{(2t-(d+r)+1)^{1/2}}\right).
\end{aligned}$$

Again using $1+x \leq e^x$ we have

$$\left(1 - \frac{d+r-1}{2t}\right)^{r/2} \leq \exp\left\{-\frac{r(d+r-1)}{4t}\right\}$$

and

$$\left(1 + \frac{2}{(2t-(d+r)+1)^{1/2}}\right)^{r} \leq \exp\left\{\frac{2r}{(2t-(d+r)+1)^{1/2}}\right\}$$

So

$$\begin{aligned}
p_{\mathbf{S}}(\mathbf{r};\mathbf{d},t_0,t) \leq \left(\prod_{i=1}^{\ell}\binom{r_i+d_i-1}{d_i-1}\right) err(r,d,t_0,t)\left(\frac{t_0+1}{t}\right)^{d/2} \\
\exp\left\{-\frac{r(d+r-1)}{4t} + \frac{2r}{(2t-(d+r)+1)^{1/2}}\right\}.
\end{aligned}$$

For $d = o(t^{1/2})$ and $r = o(t^{2/3})$, we have

$$\begin{aligned}
&err(r,d,t_0,t)\exp\left\{-\frac{r(d+r-1)}{4t} + \frac{2r}{(2t-(d+r)+1)^{1/2}}\right\} \\
&\qquad \leq \left(1 - \frac{d}{2(t_0+1)}\right)^{-\left(1-\frac{d}{2(t_0+1)}\right)(t_0+1)} \exp\left\{1 - \frac{d}{2} - \frac{r^2}{8t} + \frac{2r}{t^{1/2}}\right\}.
\end{aligned}$$

Since $x^{-x} \leq e$, this completes the proof of the lemma. $\square$