Vertices of high degree in the preferential attachment tree[*]

Graham Brightwell[†]   Malwina Luczak[‡]

Abstract

We study the basic preferential attachment process, which generates a sequence of random trees, each obtained from the previous one by introducing a new vertex and joining it to one existing vertex, chosen with probability proportional to its degree. We investigate the number $D_t(\ell)$ of vertices of each degree ℓ at each time t, focussing particularly on the case where ℓ is a growing function of t. We show that $D_t(\ell)$ is concentrated around its mean, which is approximately $4t/\ell^3$, for all $\ell \leq (t/\log t)^{-1/3}$; this is best possible up to a logarithmic factor.

Keywords: random graphs; web graphs; concentration of measure; martingales; preferential attachment.
AMS MSC 2010: 05C80; 60J10; 60G42.
Submitted to EJP on December 27, 2010, final version accepted on January 18, 2012.
Supersedes arXiv:1012.5550.

1 Introduction

In this paper, we study the basic *preferential attachment* process, which is defined as follows. We start with a (small) tree on $\tau_0 \geq 1$ vertices. At each integer time $t > \tau_0$, a new vertex arrives, and is joined to one existing vertex; a vertex is chosen as the other endpoint of the new edge with probability proportional to its current degree. Thus, at each time t, we have a tree with t vertices. The random tree obtained at any time t is called the *preferential attachment tree.*

The first appearance of this process can be traced back at least to Yule [25] in 1925, and in probability theory the model is sometimes referred to as a *Yule process.* Subsequently, Szymański [22] studied the preferential attachment process in the guise of *plane-oriented recursive trees.* He gave a formula for the expected number $d_t(\ell)$ of vertices of degree ℓ at time t, namely

$$d_t(\ell) = \frac{4t}{\ell(\ell+1)(\ell+2)} + O(1).$$

[†]London School of Economics, UK. E-mail: g.r.brightwell@lse.ac.uk
[‡]University of Sheffield, UK. E-mail: m.luczak@sheffield.ac.uk

The structure of such trees was further analysed by Mahmood, Smythe and Szymański [17], and by Mahmood and Smythe [16]. Lu and Feng [12] proved a concentration result for the random number $D_t(\ell)$ of vertices of degree ℓ, for fixed ℓ.

Interest in the model surged after a paper of Barabási and Albert [2] in 1999, who proposed preferential attachment as a model of the growth of "web graphs", i.e., graphs possessing many of the same properties as "real-world networks" such as the worldwide web. Barabási and Albert studied not just the preferential attachment process as defined above, but also the variant where each new vertex chooses some fixed number $m \geq 1$ of neighbours. For $m > 1$, the *preferential attachment graphs* produced are of course not trees, but for properties such as the degree sequence, the overall pattern of behaviour is the same for any fixed m.

Preferential attachment graphs were studied formally by Bollobás, Riordan, Spencer and Tusnády [4], who proved that the degree sequence follows a power law with exponent 3, i.e., the expected number $d_t^m(\ell)$ of vertices of degree ℓ at time t is of order t/ℓ^3, more precisely

$$d_t^m(\ell) \simeq \frac{2m(m+1)}{(\ell+m)(\ell+m+1)(\ell+m+2)} t$$

for all $\ell \leq t^{1/15}$. They also showed that the random number $D_t^m(\ell)$ of vertices of degree ℓ at time t is concentrated within $O(\sqrt{t \log t})$ of its expectation $d_t^m(\ell)$. They further indicated how the results could be extended to somewhat larger values of ℓ.

Szymański [23] gave a more precise estimate for $d_t^1(\ell) = d_t(\ell)$. Combining this with the concentration result of Bollobás, Riordan, Spencer and Tusnády [4] shows that $D_t(\ell)$ is concentrated within a factor $(1 + o(1))$ of its mean for ℓ up to nearly $t^{1/6}$. Szymański [23] also gave a precise estimate for the expected number $u_t(\ell)$ of vertices of degree *at least* ℓ, namely $u_t(\ell) = \dfrac{2t}{\ell(\ell+1)} + O(1)$. Janson [11], extending a result of Mahmoud, Smythe and Szymański [17], proved a central limit theorem for $D_t(\ell)$ as $t \to \infty$, jointly for all $\ell \geq 1$.

A much more general model was introduced and studied by Cooper and Frieze [6], and Cooper [5]: in the latter paper, Cooper proved a general result that implies (weak) concentration for $D_t(\ell)$ whenever $\ell \leq t^{1/6}/\log^2 t$.

The maximum degree Δ_t of the preferential attachment tree is known to behave as $t^{1/2}$ as $t \to \infty$. Móri [19] proved a law of large numbers and a central limit theorem for the Δ_t: in particular, he showed that $\Delta_t t^{-1/2}$ converges almost surely to some positive (non-constant) random variable, as $t \to \infty$. Further, he showed that the fluctuations of $\Delta_t t^{-1/2}$ around the limit, scaled by $t^{-1/4}$, converge in distribution to a normal law.

Many variants of the preferential attachment process have been studied. The limiting proportion of vertices of each *fixed* degree ℓ has been investigated in many different models extending and generalising that of preferential attachment trees. See for instance, Rudas, Toth and Valko [21], Athreya, Ghosh and Sethuraman [1], Deijfen, van den Esker, van der Hofstad and Hooghiemstra [8], and Dereich and Mörters [9]. See also the survey of Bollobás and Riordan [3] for a number of other results on related models.

Our principal aim in this paper is to prove concentration of measure results for $D_t(\ell)$ for all values of ℓ up to the expected maximum degree. For values of ℓ above $\varepsilon t^{1/3}$, the expectation of $D_t(\ell)$ is of order at most 1, and all we show is that $D_t(\ell)$ is, with high probability, at most about $\log t$. For values of ℓ at most $(t/\log t)^{1/3}$, we shall prove that $D_t(\ell)$ is concentrated within about $\sqrt{t \log t/\ell^3}$ of its mean.

We can write $D_t(\ell) = \sum_{s=1}^{t} I(s,t,\ell)$, where $I(s,t,\ell)$ is the indicator of the event that the vertex arriving at time s (or, for $s \leq \tau_0$, the initial vertex labelled s) has degree exactly ℓ at time t. One would expect that, for large t and ℓ in an appropriate range, most

of the variables $I(s, t, \ell)$ are approximately independent of each other, each with mean bounded away from 1. This would suggest that the variance of $D_t(\ell)$ is of the same order as its mean, and that $D_t(\ell)$ should be concentrated within about $\sqrt{\mathbb{E}\, D_t(\ell)} \simeq \sqrt{t/\ell^3}$ of its mean. This is indeed the case for constant ℓ: see Janson [11] for asymptotic formulae for the covariances $\mathrm{Cov}(D_t(\ell)/t, D_t(j)/t)$. So our concentration result is likely to be best possible up to logarithmic factors.

Our methods can also be used to prove similar results for the random variable $U_t(\ell)$, the number of vertices of degree *at least* ℓ at time t. The expectation of $U_t(\ell)$ is approximately $2t/\ell^2$ for large t: at the end of this paper, we indicate briefly how to adapt our proof to show that $U_t(\ell)$ is concentrated within about $\sqrt{t \log t/\ell^2}$ of its mean as long as $\ell \leq t^{1/2}/\log^{13/2} t$.

Before stating our results, we specify our model precisely. We start at some time $\tau_0 \geq 1$, with an initial graph $G(\tau_0) = (V(\tau_0), E(\tau_0))$, with $|V(\tau_0)| = \tau_0$, $|E(\tau_0)| = \tau_0 - 1$; we think of $G(\tau_0)$ as a tree, although it need not be. At each step $t > \tau_0$, a new vertex v_t is created, and is joined to existing vertices by one new edge, whose other endpoint is chosen by *preferential attachment*, that is, a vertex v is chosen as an endpoint with probability proportional to its degree at time $t - 1$. Note that, if $G(\tau_0)$ is a tree, then the graph at all later stages is also a tree.

Our main theorem concerns the number $D_t(\ell)$ of vertices of degree exactly ℓ at time t, for all $\ell \geq 1$ and $t \geq \tau_0$.

Theorem 1.1. *Let $\tau_0 \geq 4$ and $\psi \geq 10^5 \sqrt{\tau_0 - 1} \log^3 \tau_0$ be constants. Let $G(\tau_0)$ be any graph with τ_0 vertices and $\tau_0 - 1$ edges, and consider the preferential attachment process with initial graph $G(\tau_0)$ at time τ_0, and the associated Markov chain $D = (D_t(\ell) : t \geq \tau_0, \ell \in \mathbb{N})$.*

With probability at least $1 - \frac{4}{\psi}$, we have

$$\left| D_t(\ell) - \frac{4t}{\ell(\ell+1)(\ell+2)} \right| \leq 120\sqrt{\frac{t \log(\psi t)}{\ell^3}} + 301\psi^2 \log(\psi t),$$

for all $\ell \geq 1$, and all $t \geq \tau_0$.

The parameter ψ is a constant which may be chosen arbitrarily large in order to make the probability of failure arbitrarily small; the results are only of interest when t is larger than some $t_0(\psi)$. All our results are stated in terms of such a parameter (denoted ψ or ω).

We state here an analogous result about the number $U_t(\ell)$ of vertices of degree *at least* ℓ at time t, for all $\ell \geq 2$ and all $t \geq \tau_0$. (Note that $U_t(1)$ is equal to t for each $t \geq \tau_0$.)

Theorem 1.2. *Let $\tau_0 \geq 4$ and $\psi \geq \max\left(\tau_0, 10^5 \sqrt{\tau_0 - 1} \log^3 \tau_0\right)$ be constants. Let $G(\tau_0)$ be any graph with τ_0 vertices and $\tau_0 - 1$ edges, and consider the preferential attachment process with initial graph $G(\tau_0)$ at time τ_0, and the associated Markov chain $U = (U_t(\ell) : t \geq \tau_0, \ell \in \mathbb{N})$.*

With probability at least $1 - \frac{4}{\psi}$, we have

$$\left| U_t(\ell) - \frac{2t}{\ell(\ell+1)} \right| \leq 45\frac{\sqrt{t \log(\psi t)}}{\ell} + 4 \times 10^9 \psi \log^7(\psi t),$$

for all $\ell \geq 2$, and all $t \geq \tau_0$.

We do not give a detailed proof of Theorem 1.2 in this paper, but we do give an indication of how to adapt our proof of Theorem 1.1 to give this result. The term $\log^7(\psi t)$ appearing above could certainly be improved with more work.

Theorem 1.2 seems to be the first explicit result concerning concentration of measure for $U_t(\ell)$, although some weak concentration can be deduced from concentration results for $D_t(\ell)$. Moreover, Talagrand's inequality [24] can be applied readily: to demonstrate that $U_t(\ell) \geq x$, a certificate of length at most $O(x\ell)$ suffices (see for instance [18] for details of the method). This method gives concentration for $U_t(\ell)$ up to about $\ell = t^{1/3}$, and indeed concentration for $D_t(\ell)$ up to about $\ell = t^{1/5}$.

For constant values of ℓ, Bollobás, Riordan, Spencer and Tusnády [4] showed that $D_t(\ell)$ is concentrated within about $t^{1/2}$ of its mean, which is best possible; a similar result for $U_t(\ell)$ follows. For larger values of ℓ, in particular where ℓ is growing as a small power of t, earlier methods (including the method based on Talagrand's inequality that we mentioned above) do not give the "optimal" concentration of $D_t(\ell)$ or $U_t(\ell)$ about their respective means. Our results above do give what should be optimal concentration, up to logarithmic factors, for $D_t(\ell)$ and $U_t(\ell)$, whenever the expectations of these random variables tend to infinity, again up to logarithmic factors.

In Section 2, we give an exposition of a method based on exponential supermartingales, that is widely used in the analysis of continuous time Markov processes. We transfer the method to the discrete time setting, and state two theorems that we shall use, and that can be applied in other similar contexts.

In Section 3, we apply our method to describe the evolution of the degree of a fixed vertex in the preferential attachment model. We do this partly to illustrate the method, but mostly so that we can use the results in later sections. We prove a result on the maximum degree Δ_t that is weaker than Móri's [19], but simple to prove, in the interests of keeping the paper self-contained.

Sections 4 to 6 are devoted to the proof of Theorem 1.1. Section 4 contains the main thread of the proof, and we defer some calculations to Sections 5 and 6. One difficulty we face is that we cannot get sharp results by working directly with the natural martingale associated to the Markov chain $D = (D_t(\ell) : t \geq \tau_0, \ell \in \mathbb{N})$, so we work instead with a suitable transform of that martingale. Proving concentration of measure for the transform is not straightforward, so we introduce another Markov process derived from D, and apply our methods from Section 2 to that process.

Section 7 contains a brief sketch of the proof of Theorem 1.2.

In this paper, we deal only with the preferential attachment tree. However, our methods will extend to more general settings, and indeed we believe we can prove results similar to those above for the general Cooper-Frieze model. We intend to address this elsewhere; in a very brief final section, we make a few remarks on the difficulties involved in extending our proof to other preferential attachment models.

2 Our method: exponential supermartingales

The following technique is adapted from a fairly standard method used in the analysis of continuous-time random processes; see for instance [7], [13] and [15]. We have not been able to find a suitable account in the literature of a discrete-time, and time-dependent, version of the method for us to quote, so we develop the theory here. We provide results that we hope may prove useful in other settings.

Let $X = (X_t : t \in \mathbb{Z}^+)$ be a discrete-time Markov chain, possibly time non-homogeneous, with countable state space E and transition matrix $P_t = (P_t(x, x') : x, x' \in E)$ at time t. (Here and in what follows, our matrices – which will normally be infinite – have rows and columns indexed by the countable set E.) Let $(\mathcal{F}_t)$ be a filtration, and suppose that (X_t) is adapted.

Let I denote the identity matrix. Further, let us write, for a matrix A,

$$(Af)(x) = \sum_{x' \in E} A(x, x')f(x').$$

Then we see that

$$[(P_t - I)f](x) = \sum_{x' \in E} [P_t(x, x') - I(x, x')]f(x') = \sum_{x' \in E} P_t(x, x')(f(x') - f(x)).$$

Further, note that

$$[(P_t - I)f](x) = \mathbb{E}[f(X_{t+1}) - f(x) \mid X_t = x], \tag{2.1}$$

that is, $[(P_t - I)f](x)$ is the expected change in f at the t-th step given that $X_t = x$.

Lemma 2.1. *Suppose $X_0 = x_0$ a.s. Let $f : E \to \mathbb{R}$ be a function such that $\mathbb{E}[|f(X_t)| \mid X_0 = x_0]$ is finite for each t. Then*

$$
\begin{aligned}
M_t^f &= f(X_t) - f(X_0) - \sum_{s=0}^{t-1} [(P_s - I)f](X_s) \\
&= f(X_t) - f(X_0) - \sum_{s=0}^{t-1} \sum_{x'} P_s(X_s, x')(f(x') - f(X_s))
\end{aligned}
$$

is an $(\mathcal{F}_t)$-martingale.

Proof. The proof for the time homogeneous case can be found in Norris [20]. Checking that M_t^f is a martingale in the time non-homogeneous case is just as easy. Consider

$$
\begin{aligned}
\mathbb{E}[M_{t+1}^f \mid \mathcal{F}_t] &= \mathbb{E}\left[f(X_{t+1}) - f(X_0) - \sum_{s=0}^{t} [(P_s - I)f](X_s) \mid \mathcal{F}_t \right] \\
&= \mathbb{E}[f(X_{t+1}) \mid X_t] - f(X_0) - [(P_t - I)f](X_t) \\
&\quad - \sum_{s=0}^{t-1} [(P_s - I)f](X_s) \\
&= f(X_t) - f(X_0) - \sum_{s=0}^{t-1} [(P_s - I)f](X_s) \\
&= M_t^f,
\end{aligned}
$$

where we used (2.1). Also, for each $t \geq 0$,

$$
\begin{aligned}
\mathbb{E}|M_t^f| &\leq \mathbb{E}|f(X_t)| + \mathbb{E}|f(X_0)| + \sum_{s=0}^{t-1} \mathbb{E}|[(P_s - I)]f(X_s)| \\
&\leq \mathbb{E}|f(X_t)| + \mathbb{E}|f(X_0)| + \sum_{s=0}^{t-1} (\mathbb{E}|f(X_s)| + \mathbb{E}|f(X_{s+1})|) \\
&< \infty.
\end{aligned}
$$

$\square$

Lemma 2.2. *Suppose $X_0 = x_0$ a.s. Let $f : E \to \mathbb{R}^+$ be a function. Then*

$$Z_t^f = \frac{f(X_t)}{f(X_0)} \exp\left(-\sum_{s=0}^{t-1} \frac{[(P_s - I)f](X_s)}{f(X_s)} \right)$$

is an $\mathcal{F}_t$-supermartingale, as long as $\mathbb{E}\, Z_t^f < \infty$ for all t.

Proof. Consider

$$
\begin{aligned}
\mathbb{E}[Z_{t+1}^f \mid \mathcal{F}_t] &= \mathbb{E}[f(X_{t+1}) \mid X_t] \frac{1}{f(X_0)} \exp\left(-\sum_{s=0}^{t} \frac{[(P_s - I)f](X_s)}{f(X_s)}\right) \\
&= \frac{f(X_t)}{f(X_0)} \frac{(P_t f)(X_t)}{f(X_t)} \exp\left(1 - \frac{(P_t f)(X_t)}{f(X_t)}\right) \\
&\quad \times \exp\left(-\sum_{s=0}^{t-1} \frac{[(P_s - I)f](X_s)}{f(X_s)}\right) \\
&\leq \frac{f(X_t)}{f(X_0)} \exp\left(-1 + \frac{(P_t f)(X_t)}{f(X_t)}\right) \exp\left(1 - \frac{(P_t f)(X_t)}{f(X_t)}\right) \\
&\quad \times \exp\left(-\sum_{s=0}^{t-1} \frac{[(P_s - I)f](X_s)}{f(X_s)}\right) \\
&= Z_t^f,
\end{aligned}
$$

where we have used the fact that $\mathbb{E}[f(X_{t+1}) \mid X_t] = P_t f$, and the fact that $x \leq \exp(-1+x)$ for all x. $\qquad\square$

Note that, for a continuous-time Markov chain, the analogue of Z_t^f in Lemma 2.2 is in fact a martingale; see for example Lemma 3.2 in Chapter 4 of [10]. In the time-continuous case, the matrix $(P_t - I)$ is replaced by the generator matrix A_t of the Markov chain, which is the derivative at time t of its transition semigroup P_t.

We shall show how, under certain conditions, Lemma 2.1 and Lemma 2.2 can be used to prove a law of large numbers for a Markov chain.

Lemma 2.3. *Let $g : E \to \mathbb{R}$ be a function, and suppose that $X_0 = x_0$ a.s., for some $x_0 \in E$. For $\theta \in \mathbb{R}$, let*

$$
\varphi_s^g(x, \theta) = \sum_{x' \in E} P_s(x, x') \left(e^{\theta(g(x') - g(x))} - 1 - \theta(g(x') - g(x))\right).
$$

Then

$$
Z_t^g(\theta) = \exp\left(\theta M_t^g - \sum_{s=0}^{t-1} \varphi_s^g(X_s, \theta)\right)
$$

is an $\mathcal{F}_t$-supermartingale, as long as $\mathbb{E}\, Z_t^g(\theta) < \infty$ for each t.

Proof. The result is a consequence of Lemma 2.2, with $f(x) = e^{\theta(g(x) - g(x_0))}$. That lemma tells us that Z_t^f is a supermartingale, and we need only verify that $Z_t^f = Z_t^g(\theta)$ for this choice of f.

The calculation goes as follows:

$$
\begin{aligned}
Z_t^f &= \frac{f(X_t)}{f(X_0)} \exp\Big(-\sum_{s=0}^{t-1} \frac{[(P_s - I)f](X_s)}{f(X_s)} \Big) \\
&= \exp(\theta(g(X_t) - g(X_0))) \\
&\quad \times \exp\Big(-\sum_{s=0}^{t-1} \frac{\sum_{x'} P_s(X_s, x')[e^{\theta(g(x')-g(X_0))} - e^{\theta(g(X_s)-g(X_0))}]}{e^{\theta(g(X_s)-g(X_0))}} \Big) \\
&= \exp\Big(\theta(g(X_t) - g(X_0)) - \sum_{s=0}^{t-1} \sum_{x'} P_s(X_s, x')[e^{\theta(g(x')-g(X_s))} - 1] \Big) \\
&= \exp\Big(\theta(g(X_t) - g(X_0)) - \theta \sum_{s=0}^{t-1} \sum_{x'} P_s(X_s, x')(g(x') - g(X_s)) \\
&\quad - \sum_{s=0}^{t-1} \varphi_s^g(X_s, \theta) \Big) \\
&= \exp\Big(\theta M_t^g - \sum_{s=0}^{t-1} \varphi_s^g(X_s, \theta) \Big).
\end{aligned}
$$

$\square$

Note that, while X_t remains in a 'good' set S_t of states x where $e^{\theta(g(x')-g(x))}$ is bounded by some constant (possibly depending on t) over all x' such that $P_t(x, x') > 0$ and all $x \in S_t$, (i.e., the size of changes in g stays uniformly bounded), then the finiteness assumption of Lemma 2.3 holds. Furthermore, we can approximate $e^{\theta(g(x')-g(X_t))}$ using a Taylor expansion.

In many applications, in particular those in this paper, $|g(x') - g(x)|$ will be uniformly bounded over the entire state space E and over all transition matrices P_t: if we work up to some fixed time τ, then it suffices to have the bound valid for $t < \tau$. We assume from now on that, for every $\tau \geq 0$, there is some real number $J = J(\tau)$ such that g satisfies:

$$
\sup_{s<\tau, x} \sup_{x':P_s(x,x')\neq 0} |g(x') - g(x)| \leq J < \infty. \tag{2.2}
$$

Now we fix some real number $\alpha > 0$, and restrict attention to values of θ such that $|\theta| \leq \alpha$. We use the identity

$$
e^z - 1 - z = z^2 \int_{r=0}^1 e^{rz}(1 - r)\, dr
$$

to deduce that

$$
\begin{aligned}
\varphi_s^g(x, \theta) &= \sum_{x'} P_s(x, x')\theta^2(g(x') - g(x))^2 \int_0^1 e^{r\theta(g(x')-g(x))}(1 - r)\, dr \\
&\leq \theta^2 \sum_{x'} P_s(x, x')(g(x') - g(x))^2 e^{\alpha J} \int_0^1 (1 - r)\, dr \\
&= \frac{1}{2}\theta^2 e^{\alpha J} \sum_{x'} P_s(x, x')(g(x') - g(x))^2.
\end{aligned}
$$

Suppose that $X_0 = x_0$ a.s., for some $x_0 \in E$, and that we study the chain up to some time $\tau > 0$. Our aim is to show that M_t^g remains small over the period $0 \leq t \leq \tau$. For a precise statement, we need a few more definitions.

We set

$$\Phi_t^g(X) = \sum_{s=0}^{t} \sum_{x'} P_s(X_s, x')(g(x') - g(X_s))^2,$$

so that

$$Z_t^g(\theta) \geq \exp\left(\theta M_t^g - \frac{1}{2}\theta^2 e^{\alpha J} \Phi_{t-1}^g(X)\right),$$

for all θ with $|\theta| \leq \alpha$.

Now let R be a positive real number, and set

$$T_R = \inf\{t \geq 0 : \Phi_t^g(X) > R\}.$$

Thus, for $t \leq T_R$, we have $\Phi_{t-1}^g(X) \leq R$, and therefore

$$Z_t^g(\theta) \geq \exp\left(\theta M_t^g - \frac{1}{2}\theta^2 e^{\alpha J} R\right),$$

provided $|\theta| \leq \alpha$.

Also, for $\delta > 0$, we define

$$T_g^+(\delta) = \inf\{t : M_t^g > \delta\}, \quad T_g^-(\delta) = \inf\{t : M_t^g < -\delta\},$$

and

$$T_g(\delta) = T_g^+(\delta) \wedge T_g^-(\delta) = \inf\{t : |M_t^g| > \delta\}.$$

Lemma 2.4. *Fix $\tau > 0$ and $R > 0$, and let $g : E \to \mathbb{R}$ be a function satisfying (2.2) for some $J \in \mathbb{R}$. Also, let $\alpha > 0$ and $\delta > 0$ be any constants such that $\delta \leq e^{\alpha J} \alpha R$. Then*

$$\mathbb{P}\left(T_g(\delta) \leq T_R \wedge \tau\right) \leq 2e^{-\delta^2/(2Re^{\alpha J})},$$

and hence

$$\mathbb{P}\left((\sup_{0 \leq t \leq \tau} |M_t^g| > \delta) \wedge (T_R \geq \tau)\right) \leq 2e^{-\delta^2/(2Re^{\alpha J})}.$$

In particular:

(i) *for any $\omega \leq R/J^2$, we obtain the following by choosing $\alpha = \log 2/J$ and $\delta = \sqrt{\omega R}$:*

$$\mathbb{P}\left(\left(\sup_{0 \leq t \leq \tau} |M_t^g| > \sqrt{\omega R}\right) \wedge (T_R \geq \tau)\right)$$
$$\leq \quad \mathbb{P}\left(T_g(\sqrt{\omega R}) \leq T_R \wedge \tau\right)$$
$$\leq \quad 2e^{-\omega/4};$$

(ii) *for any $\omega \geq R/J^2$, we set $\alpha = \frac{1}{J} \log\left(2\omega J^2/R\right)$ and $\delta = \omega J$, and obtain*

$$\mathbb{P}\left((\sup_{0 \leq t \leq \tau} |M_t^g| > \omega J) \wedge (T_R \geq \tau)\right) \leq \mathbb{P}\left(T_g(\omega J) \leq T_R \wedge \tau\right) \leq 2e^{-\omega/4}.$$

Proof. Fix any real θ with $|\theta| \leq \alpha$. For ease of notation, we write T_g^+ for $T_g^+(\delta)$ and T_g^- for $T_g^-(\delta)$. By Lemma 2.3, $(Z_t^g(\theta))_{t \geq 0}$ is a supermartingale.

On the event $\{T_g^+ \leq T_R \wedge \tau\}$, we have $M_{T_g^+}^g > \delta$ and

$$Z_{T_g^+}^g(\theta) > \exp\left(\theta\delta - \frac{1}{2}\theta^2 e^{\alpha J} R\right).$$

By optional stopping,

$$\mathbb{E}[Z^g_{T^+_g}(\theta)] \leq \mathbb{E}[Z^g_0(\theta)] = 1.$$

Hence, using the Markov inequality,

$$
\begin{aligned}
\mathbb{P}(T^+_g \leq T_R \wedge \tau) \quad &\leq \quad \mathbb{P}\left(Z^g_{T^+_g}(\theta) > \exp\left(\delta\theta - \frac{1}{2}\theta^2 e^{\alpha J}R\right)\right) \\
&\leq \quad \exp\left(-\delta\theta + \frac{1}{2}\theta^2 e^{\alpha J}R\right).
\end{aligned}
$$

Optimising in θ, we find that $\theta = \delta/e^{\alpha J}R$ is the best choice, and note that $|\theta| \leq \alpha$. This yields

$$\mathbb{P}(T^+_g \leq T_R \wedge \tau) \leq \exp\left(-\delta^2 \Big/ 2e^{\alpha J}R\right).$$

An almost identical calculation gives

$$\mathbb{P}(T^-_g \leq T_R \wedge \tau) \leq \exp\left(-\delta^2 \Big/ 2e^{\alpha J}R\right),$$

and the first part of the result follows.

The two special cases are obtained by choosing the given values of α and δ, and verifying that

$$\delta \leq e^{\alpha J}\alpha R \quad \text{and} \quad \frac{\delta^2}{2e^{\alpha J}R} \geq \frac{\omega}{4}$$

in each case. $\qquad\square$

We summarise what we have proved in a theorem.

Theorem 2.5. *Let $X = (X_t)_{t \in \mathbb{Z}^+}$ be a discrete-time Markov chain, with countable state space E and transition matrix P_t at time t, and suppose that (X_t) is adapted to a filtration $(\mathcal{F}_t)$. Let $g : E \to \mathbb{R}$ be any function, τ any natural number, and J any real number, satisfying*

$$\sup_{s<\tau, x \in E} \sup_{x' \in E : P_s(x, x')>0} |g(x') - g(x)| \leq J.$$

Set

$$\Phi^g_t(X) = \sum_{s=0}^{t} \sum_{x' \in E} P_s(X_s, x')\big(g(x') - g(X_s)\big)^2.$$

Let $R > 0$ be a real number, and set

$$T_R = \inf\{t \geq 0 : \Phi^g_t(X) > R\}.$$

Then

$$M^g_t = g(X_t) - g(X_0) - \sum_{s=0}^{t-1} \sum_{x' \in E} P_s(X_s, x')(g(x') - g(X_s))$$

is an $(\mathcal{F}_t)$-martingale and, for any $\omega > 0$,

$$\mathbb{P}\left(\left(\sup_{0 \leq t \leq \tau} |M^g_t| > \max\left(\sqrt{\omega R}, \omega J\right)\right) \wedge (T_R \geq \tau)\right) \leq 2e^{-\omega/4}.$$

We have demanded that the state space be countable, so that we can express our results in terms of sums over the state space. It suffices to assume instead that, for any state $x \in E$, and any time s, there is a countable set $E(x, s) \subseteq E$ such that $\sum_{x' \in E} P_s(x, x') = 1$. Indeed, under this assumption, if $X_0 = x_0$ a.s. for some state x_0, then there is a countable set $E' \subseteq E$ such that, a.s., $X_t \in E'$ for all $t \geq 0$.

We also remark that, in the statement above, we begin our consideration of the chain at time 0. When applying Theorem 2.5 in the analysis of the preferential attachment tree, we shall instead start at some fixed time τ_0: of course this makes no substantive difference.

In some instances, for example in Section 3, we will want to bound the probability that $|M_t^g| \leq \delta(t)$ for all $t \leq \tau$, where $\delta(t)$ is a suitable function growing with t. One easy way to do this is to apply the above theorem for each value $t \leq \tau$, choosing an appropriate value $R(t)$ of R at each time. This approach has the drawback that it is necessary to sum the probabilities of failure over $t \leq \tau$. Better bounds may be obtained by applying the lemma only for a sparse sequence of values t, as we illustrate in the proof of the following result.

The notation here is essentially as for Theorem 2.5. We again have a real-valued function g defined on the state space E of a Markov chain X, and the change in g is uniformly bounded by J over all possible transitions of the chain. The function $\Phi_t^g(X)$ is as in Theorem 2.5. Now we have a non-decreasing function $R : \mathbb{Z}^+ \to \mathbb{R}^+$, and we set

$$T_R = \inf\{t \geq 0 : \Phi_t^g(X) > R(t)\}.$$

Also, for any non-decreasing function $\delta : \mathbb{Z}^+ \to \mathbb{R}^+$, we define an associated stopping time

$$T_g(\delta) = \inf\{t \geq 0 : |M_t^g| > \delta(t)\}.$$

With the notation as above, we have the following result.

Theorem 2.6. (a) Fix $\omega > 4$, and let $\delta(t) = \max(\omega J, 2\sqrt{\omega R(t-1)})$ for $t \geq 1$. Then, for any $\tau > 0$ such that $R(\tau - 1) \geq \omega J^2$,

$$\mathbb{P}(T_g(\delta) \leq T_R \wedge \tau) \leq 2 \log\left(\frac{8R(\tau-1)}{\omega J^2}\right) e^{-\omega/4}.$$

(b) Fix $\psi \geq 4/J^2$, and suppose that $R(t)$ tends to infinity as $t \to \infty$. For $t \geq 1$, let $\widetilde{\delta}(t) = 2\max\left(\psi J \log(\psi J^2), \sqrt{\psi R(t-1) \log R(t-1)}\right)$. Then

$$\mathbb{P}(T_g(\widetilde{\delta}) \leq T_R) \leq 5e^{-\psi/4}.$$

Proof. For (a), we define a finite sequence of times $\tau_1, \tau_2, \ldots$ as follows. Let τ_1 be the first t for which $R(t) > \omega J^2$: by assumption $\tau_1 \leq \tau$. For $j > 1$, if $\tau_j < \tau$ then we set

$$\tau_{j+1} = \inf\{t > \tau_j : R(t) > 4R(\tau_j)\} \wedge \tau.$$

The final term τ_N in the sequence is the first τ_j with $\tau_j = \tau$, and the number N of terms in the sequence is then no greater than $2 + \log_4\left(\frac{R(\tau-1)}{\omega J^2}\right) \leq \log\left(\frac{8R(\tau-1)}{\omega J^2}\right)$.

We first apply Lemma 2.4(ii) with $\tau = \tau_1$ and $R = R(\tau_1 - 1)$, noting that $\omega \geq R(\tau_1 - 1)/J^2$ by definition of τ_1. We obtain that

$$\mathbb{P}(T_g(\omega J) \leq T_R \wedge \tau_1) \leq \mathbb{P}(T_g(\omega J) \leq T_{R(\tau_1-1)} \wedge \tau_1) \leq 2e^{-\omega/4}.$$

As $\delta(t) \geq \omega J$ for all $t \leq \tau_1$, this means that, with probability at least $1 - 2e^{-\omega/4}$, $|M_t^g| \leq \delta(t)$ for all times $t \leq T_R \wedge \tau_1$.

For each of the times $\tau = \tau_j$ ($j \geq 2$), we apply Lemma 2.4(i) with $R = R(\tau_j - 1)$, noting now that $\omega \leq R(\tau_j - 1)/J^2$ by choice of τ_1. We obtain that

$$\mathbb{P}\left(T_g\left(\sqrt{\omega R(\tau_j - 1)}\right) \leq T_R \wedge \tau_j\right) \leq 2e^{-\omega/4}.$$

For each t with $\tau_{j-1} < t \leq \tau_j$, we have

$$\delta(t) \geq \delta(\tau_{j-1} + 1) \geq 2\sqrt{\omega R(\tau_{j-1})} \geq \sqrt{\omega R(\tau_j - 1)},$$

since $R(\tau_j - 1) \leq 4R(\tau_{j-1})$ by definition of τ_j. We conclude that, for each $j \geq 2$, with probability at least $1 - 2e^{-\omega/4}$, $|M_t^g| \leq \delta(t)$ for all times t with $\tau_{j-1} < t \leq T_R \wedge \tau_j$.

It now follows that, with probability at least $1 - 2Ne^{-\omega/4}$, we have $|M_t^g| \leq \delta(t)$ for all times $t \leq T_R \wedge \tau$, and part (a) follows.

The proof of part (b) is very similar in style. This time we let τ_1 be the first t for which $R(t) > 2\psi J^2 \log(\psi J^2)$. Given τ_j, we let τ_{j+1} be the minimum t such that $R(t) > 4R(\tau_j)$. The assumption that $R(t)$ tends to infinity ensures that we obtain an infinite sequence $(\tau_j)_{j \geq 1}$ of times.

We apply Lemma 2.4(ii) with $\omega = 2\psi \log(\psi J^2)$, $\tau = \tau_1$, and $R = R(\tau_1 - 1)$, noting that $\omega \geq R(\tau_1 - 1)/J^2$ by choice of τ_1. We obtain that

$$\mathbb{P}\left(T_g(2\psi J \log(\psi J^2)) \leq T_R \wedge \tau_1\right) \leq 2e^{-\psi/4}.$$

Since $\widetilde{\delta}(t) \geq 2\psi J \log(\psi J^2)$ for all t, this implies that, with probability at least $1 - 2e^{-\psi/4}$, $|M_t^g| \leq \widetilde{\delta}(t)$ for all times $t \leq T_R \wedge \tau_1$.

Now, for each $\tau = \tau_j$ ($j \geq 2$), we apply Lemma 2.4(i) with $R = R(\tau_j - 1)$ and $\omega = \psi \log R(\tau_{j-1})$ for each j. We need to check that $\omega \leq R(\tau_j - 1)/J^2$, i.e., that $R(\tau_j - 1)/\log R(\tau_{j-1}) \geq \psi J^2$: we have

$$\frac{R(\tau_j - 1)}{\log R(\tau_{j-1})} \geq \frac{R(\tau_{j-1})}{\log R(\tau_{j-1})} \geq \frac{2\psi J^2 \log(\psi J^2)}{\log(2\psi J^2 \log(\psi J^2))} > \frac{2\psi J^2 \log(\psi J^2)}{2\log(\psi J^2)} = \psi J^2,$$

as required; here we used the facts that (i) $R(\tau_j - 1) \geq R(\tau_{j-1}) \geq 2\psi J^2 \log(\psi J^2)$, by the definition of the τ_j, and $R(\cdot)$ is increasing, (ii) $\psi J^2 \geq 4$ and $x/\log x$ is an increasing function, with minimum value e, for $x \geq e$ (so in particular $2\log(\psi J^2) < \psi J^2$).

We obtain:

$$\mathbb{P}\left(T_g\left(\sqrt{\psi R(\tau_j - 1)\log R(\tau_{j-1})}\right) \leq T_R \wedge \tau_j\right) \leq 2e^{-\psi \log R(\tau_{j-1})/4}.$$

We have, for $\tau_{j-1} < t \leq \tau_j$,

$$\widetilde{\delta}(t) \geq \widetilde{\delta}(\tau_{j-1}) \geq 2\sqrt{\psi R(\tau_{j-1})\log R(\tau_{j-1})} \geq \sqrt{\psi R(\tau_j - 1)\log R(\tau_{j-1})}$$

and so, with probability at least $1 - 2e^{-\psi \log R(\tau_{j-1})/4}$, $|M_t^g| \leq \widetilde{\delta}(t)$ for all times t with $\tau_{j-1} < t \leq T_R \wedge \tau_j$.

Therefore, summing over j, and noting that $R(\tau_{j-1}) \geq 2\psi J^2 \log(\psi J^2)4^{j-2} > e^{j-1}$ for each $j \geq 2$,

$$\mathbb{P}\left(|M_t^g| > \widetilde{\delta}(t) \text{ for some } t \leq T_R\right) \leq 2e^{-\psi/4} + 2\sum_{j=2}^{\infty} e^{-\psi(j-1)/4}$$

$$= 2e^{-\psi/4}\left(1 + \frac{1}{1 - e^{-\psi/4}}\right) \leq 5e^{-\psi/4},$$

as desired. $\qquad\square$

3 Evolution of the degree of a vertex

For the remainder of the paper, we will use the results of the previous section to analyse various aspects of the preferential attachment process.

Our first, relatively simple, application is to the evolving degree of a single vertex; loosely, we prove that, if a vertex has degree k at time s, its degree at a later time t is unlikely to be far from $k\sqrt{t/s}$. Results of a similar flavour are in the literature already (see for instance Cooper [5], Athreya, Ghosh and Sethuraman [1] and Dereich and Mörters [9]); we give them here partly to illustrate our methods and partly because we shall have need of the results from this section later on.

We assume as always that our process starts at some time τ_0, with τ_0 vertices and $\tau_0 - 1$ edges. At each stage, one new vertex and one new incident edge are created, so that at each time $s \geq \tau_0$ there are s vertices and $s - 1$ edges. We identify the vertex set at time τ_0 with the set $[\tau_0] = \{1, \ldots, \tau_0\}$, and label the new vertex arriving at each later time s with s, so that the set of vertices present at time $t \geq \tau_0$ is exactly $[t]$.

For a vertex v, and $t \in \mathbb{N}$ with $t \geq \max(\tau_0, v)$, let $X_t(v)$ be the degree of vertex v at time t. For $\tau_0 \leq t < v$, we set $X_t(v) = 0$.

Set $(X_t) = (X_t(v) : v = 1, 2, \ldots)$; it is easily seen that (X_t) is a Markov process. Indeed, each component $X_t(v)$ is separately a Markov process. For $v \in \mathbb{N}$, we take $f_v(x) = x(v)$, the degree of vertex v in state x; the function f_v is the projection of the state onto its v-th component. In what follows, we shall assume that $v \leq \tau_0$, so that vertex v is present in the graph at time τ_0, and we let $m_0 = X_{\tau_0}(v)$ be its degree at the initial time.

We now want to calculate the corresponding martingale from Lemma 2.1. First we note that

$$
\begin{aligned}
\mathbb{E}[f_v(X_{t+1}) - f_v(X_t) \mid X_t = x] &= [(P_t - I)f_v](x) \\
&= \frac{x(v)}{2(t-1)}.
\end{aligned}
$$

This is because the sum of all vertex degrees at time t is $2(t-1)$, and the probability that a vertex w is chosen as the endpoint of the new edge from vertex t at time t, conditional on $X_t = x$, is proportional to its degree $x(w)$, and therefore the conditional probability that vertex v is chosen is $x(v)/2(t-1)$.

By Lemma 2.1, we know that the process $M(v)$ given by

$$
\begin{aligned}
M_t(v) &= f_v(X_t) - f_v(X_{\tau_0}) - \sum_{s=\tau_0}^{t-1} [(P_s - I)f_v](X_s) \\
&= X_t(v) - m_0 - \sum_{s=\tau_0}^{t-1} \frac{X_s(v)}{2(s-1)}
\end{aligned}
$$

is a martingale. We re-write the above as

$$
X_t(v) = M_t(v) + m_0 + \sum_{s=\tau_0}^{t-1} \frac{X_s(v)}{2(s-1)}. \tag{3.1}
$$

Let x_t solve the recurrence relation

$$
x_{t+1} = x_t \left(1 + \frac{1}{2(t-1)} \right),
$$

for $t \geq \tau_0$, with $x_{\tau_0} = m_0$. A simple induction argument shows that, for all $t \geq \tau_0$,

$$
m_0 \sqrt{\frac{t-1}{\tau_0 - 1}} \leq x_t \leq m_0 \sqrt{\frac{t-2}{\tau_0 - 2}}.
$$

Provided $\tau_0 \geq 4$, we have

$$\sqrt{\frac{t-2}{\tau_0 - 2}} \leq \sqrt{\frac{t-1}{\tau_0 - 1}}\left(1 + \frac{1}{2(\tau_0 - 2)}\right) \leq \frac{5}{4}\sqrt{\frac{t-1}{\tau_0 - 1}},$$

and so

$$x_t \leq \frac{5}{4}m_0\sqrt{\frac{t-1}{\tau_0 - 1}}. \tag{3.2}$$

Now fix any $\omega \geq 4$. For a vertex v, we define the time

$$T_v = \inf\left\{t \geq \tau_0 : X_t(v) > 60\omega^3 m_0\sqrt{\frac{t-1}{\tau_0 - 1}}\right\}.$$

Then for $\tau_0 \leq t < T_v$, we have $X_t(v) \leq 60\omega^3 m_0\sqrt{\frac{t-1}{\tau_0-1}}$.

Let $E_t = X_t(v) - x_t$, with $E_{\tau_0} = 0$; we want to bound $|E_t|$. Substituting $X_t(v) = x_t + E_t$ in (3.1), and using the recurrence

$$x_t = x_{\tau_0} + \sum_{s=\tau_0}^{t-1}\frac{x_s}{2(s-1)},$$

we obtain that, for $\tau_0 \leq t$,

$$|E_t| \leq |M_t(v)| + \sum_{s=\tau_0}^{t-1}\frac{|E_s|}{2(s-1)}.$$

For $\tau_0 \leq s < T_v$, $X_{s+1}(v) - X_s(v)$ is either 0 or 1, and the probability that it is equal to 1, conditional on X_s, is

$$\frac{X_s(v)}{2(s-1)} \leq 30\omega^3 m_0\sqrt{\frac{1}{(s-1)(\tau_0 - 1)}}.$$

So, for $t < T_v$, we have

$$\begin{aligned}
\Phi_t^{f_v}(X) &= \sum_{s=\tau_0}^{t}\sum_{x'} P_s(X_s, x')\left(f_v(x') - f_v(X_s)\right)^2 \\
&\leq \sum_{s=\tau_0}^{t} 30\omega^3 m_0\sqrt{\frac{1}{(s-1)(\tau_0-1)}} \\
&\leq 60\omega^3 m_0\sqrt{\frac{t}{\tau_0 - 1}}.
\end{aligned}$$

We now apply Theorem 2.6(b) to the function f_v, with $R(s) = 60\omega^3 m_0\sqrt{\frac{s}{\tau_0 - 1}}$, $J = 1$, and with $\psi = \omega$. Thus we have

$$\begin{aligned}
\widetilde{\delta}(t) &= 2\max\left(\omega\log\omega, \sqrt{60\omega^4 m_0 \log\left(60\omega^3 m_0\sqrt{(t-1)/(\tau_0-1)}\right)}\left(\frac{t-1}{\tau_0-1}\right)^{1/4}\right) \\
&\leq 8\omega^2\sqrt{m_0}\left(\frac{t-1}{\tau_0-1}\right)^{1/4}\sqrt{\log\left(60\omega^3 m_0\sqrt{(t-1)/(\tau_0-1)}\right)},
\end{aligned}$$

for $\tau_0 \leq t$. The result implies that, with probability at least $1 - 5e^{-\omega/4}$, we have $|M_t(v)| \leq \widetilde{\delta}(t)$ for all t with $\tau_0 \leq t \leq T_v$.

We thus obtain that, with probability at least $1 - 5e^{-\omega/4}$, for $\tau_0 \leq t \leq T_v$:

$$|E_t| \leq \sum_{s=\tau_0}^{t-1} \frac{|E_s|}{2(s-1)} + 8\omega^2 \sqrt{m_0} \left(\frac{t-1}{\tau_0-1}\right)^{1/4} \sqrt{\log\left(60\omega^3 m_0 \sqrt{(t-1)/(\tau_0-1)}\right)}.$$

Now $\log(xy) - y^{1/3}\log x$ is decreasing in y for $y \geq 1$ when $x \geq 20$, and is zero at $y = 1$: we apply this with $x = 60\omega^3 m_0$ and $y = \sqrt{(t-1)/(\tau_0-1)}$, to obtain that

$$\log\left(60\omega^3 m_0 \sqrt{(t-1)/(\tau_0-1)}\right) \leq \left(\frac{t-1}{\tau_0-1}\right)^{1/6} \log(60\omega^3 m_0)$$

$$\leq \omega^2 \left(\frac{t-1}{\tau_0-1}\right)^{1/6} \log(2m_0),$$

for any $m_0 \geq 1$, and so

$$|E_t| \leq \sum_{s=\tau_0}^{t-1} \frac{|E_s|}{2(s-1)} + 8\omega^3 \sqrt{m_0 \log(2m_0)} \left(\frac{t-1}{\tau_0-1}\right)^{1/3}.$$

We analyse the recurrence above using the following simple lemma.

Lemma 3.1. *Let A be a positive constant, and τ_0 a positive integer. Suppose the sequence e_t, for $t \geq \tau_0$, satisfies $e_{\tau_0} = 0$ and*

$$e_t \leq \sum_{s=\tau_0}^{t-1} \frac{e_s}{2(s-1)} + A\left(\frac{t-1}{\tau_0-1}\right)^{1/3},$$

for all $t > \tau_0$. Then

$$e_t \leq e_t^* = 6A\left[\sqrt{\frac{t-1}{\tau_0-1}} - \frac{2}{3}\left(\frac{t-1}{\tau_0-1}\right)^{1/3}\right],$$

for all $t \geq \tau_0$.

Of course, the conclusion that we shall use is that $e_t < 6A\sqrt{\frac{t-1}{\tau_0-1}}$, but the bound above is easier to establish by induction.

Proof. The proof is by induction on t, the result being true with something to spare for $t = \tau_0$.

Suppose the result is true for all s with $\tau_0 \leq s < t$. Then, by the induction hypothesis and the recursive bound, we have:

$$e_t \leq \sum_{s=\tau_0}^{t-1} \frac{e_s^*}{2(s-1)} + A\left(\frac{t-1}{\tau_0-1}\right)^{1/3}$$

$$\leq 3A \sum_{s=\tau_0}^{t-1} \frac{1}{s-1}\left[\sqrt{\frac{s-1}{\tau_0-1}} - \frac{2}{3}\left(\frac{s-1}{\tau_0-1}\right)^{1/3}\right] + A\left(\frac{t-1}{\tau_0-1}\right)^{1/3}.$$

Now the function $g(s) = \frac{1}{s-1}\left[\sqrt{\frac{s-1}{\tau_0-1}} - \frac{2}{3}\left(\frac{s-1}{\tau_0-1}\right)^{1/3}\right]$ is decreasing for all $s > \tau_0$, so $g(s) \leq 1/3\tau_0$ for all $s \geq \tau_0$, and we have

$$\sum_{s=\tau_0}^{t-1} g(s) \leq \int_{s=\tau_0}^{t} g(s)\,ds + \frac{1}{3(\tau_0-1)}$$

$$= \frac{1}{3(\tau_0-1)} + \left[2\left(\frac{s-1}{\tau_0-1}\right)^{1/2} - 2\left(\frac{s-1}{\tau_0-1}\right)^{1/3}\right]_{\tau_0}^{t}$$

$$= \frac{1}{3(\tau_0-1)} + 2\left[\left(\frac{t-1}{\tau_0-1}\right)^{1/2} - \left(\frac{t-1}{\tau_0-1}\right)^{1/3}\right].$$

This gives

$$
\begin{aligned}
e_t &\leq \frac{A}{(\tau_0 - 1)} + 6A \left(\frac{t-1}{\tau_0 - 1} \right)^{1/2} - 5A \left(\frac{t-1}{\tau_0 - 1} \right)^{1/3} \\
&\leq 6A \left(\frac{t-1}{\tau_0 - 1} \right)^{1/2} - 4A \left(\frac{t-1}{\tau_0 - 1} \right)^{1/3}.
\end{aligned}
$$

This is the desired inequality for e_t. $\hfill\square$

We can now deduce that, with probability at least $1 - 5e^{-\omega/4}$, for $\tau_0 \leq t \leq T_v$,

$$
|E_t| < 48\omega^3 \sqrt{m_0 \log(2m_0)} \sqrt{\frac{t-1}{\tau_0 - 1}}, \tag{3.3}
$$

and this bound is at most $48\omega^3 m_0 \sqrt{\frac{t-1}{\tau_0 - 1}}$. Combined with (3.2), this implies that $X_t(v) \leq 50\omega^3 m_0 \sqrt{\frac{t-1}{\tau_0 - 1}}$ for all times $t \leq T_v$. We deduce that $T_v = \infty$, since otherwise this would contradict the definition of T_v. This means that, with probability at least $1 - 5e^{-\omega/4}$, the bound (3.3) is valid for all times $t \geq \tau_0$.

We thus have the following theorem.

Theorem 3.2. *For all $\omega \geq 4$, $\tau_0 \geq 4$, and $m_0 \geq 1$, we have*

$$
\mathbb{P} \left(|X_t(v) - x_t| < 48\omega^3 \sqrt{m_0 \log(2m_0)} \sqrt{\frac{t-1}{\tau_0 - 1}} \text{ for all } t \geq \tau_0 \,\Big|\, X_{\tau_0}(v) = m_0 \right)
$$

$$
\geq 1 - e^{-5\omega/4},
$$

and therefore

$$
\mathbb{P} \left(X_t(v) \leq 50\omega^3 m_0 \sqrt{\frac{t-1}{\tau_0 - 1}} \text{ for all } t \geq \tau_0 \,\Big|\, X_{\tau_0}(v) = m_0 \right)
$$

$$
\geq 1 - e^{-5\omega/4}.
$$

We note two consequences of the result above that we shall use later.

Corollary 3.3. *For $\tau_0 \geq 4$, $\omega \geq 4$ and $m_0 \geq 10^5 \omega^7$, we have*

$$
\mathbb{P} \left(X_t(v) \leq 2m_0 \sqrt{\frac{t-1}{\tau_0 - 1}} \text{ for all } t \geq \tau_0 \,\Big|\, X_{\tau_0}(v) = m_0 \right) \geq 1 - e^{-5\omega/4}.
$$

Proof. By (3.2), we have $x_t \leq \frac{5}{4} m_0 \sqrt{\frac{t-1}{\tau_0 - 1}}$ for all $t \geq \tau_0$, given the initial condition $x_{\tau_0} = m_0$.

The result will then follow from Theorem 3.2 as long as

$$
48\omega^3 \sqrt{\frac{\log(2m_0)}{m_0}} \leq \frac{3}{4}.
$$

We see that

$$
\frac{m_0}{\log(2m_0)} \geq \frac{10^5 \omega^7}{\log(2 \times 10^5 \omega^7)} \geq \omega^6 \frac{10^5 \times 4}{\log(2 \times 10^5 \times 4^7)} \geq \omega^6 2^{12},
$$

which implies the desired inequality. $\hfill\square$

We shall also use the following result, stating that the maximum degree at time t is unlikely to be larger than $\psi\sqrt{t-1}$, where ψ is a large constant.

Theorem 3.4. *Let $\tau_0 \geq 4$ and $\psi \geq 10^5\sqrt{\tau_0-1}\log^3\tau_0$ be constants. For the preferential attachment model, with any initial graph on τ_0 vertices and τ_0-1 edges,*

$$\mathbb{P}\left(X_t(v) > \psi\sqrt{t-1} \text{ for some vertex } v \text{ and some } t \geq \tau_0\right)$$

$$\leq 2\tau_0\exp\left(-\frac{\psi^{1/3}}{3(\tau_0-1)^{1/6}}\right) \leq \frac{1}{\psi}.$$

Proof. Let P_1 be the probability that $X_t(v) \geq \psi\sqrt{t-1}$ for some $t \geq \tau_0$ and some vertex v already present at time τ_0, and P_2 be the probability that $X_t(v) \geq \psi\sqrt{t-1}$ for some $t \geq \tau_0$ and some vertex v arriving at a time later than τ_0.

We begin by bounding P_1. For a fixed vertex v present at time τ_0, its degree at that time is certainly at most τ_0-1. We apply Theorem 3.2 with $m_0 = \tau_0-1$, and $\omega = (\psi/50)^{1/3}(\tau_0-1)^{-1/6} \geq 4$, so that

$$50\omega^3 m_0\sqrt{\frac{t-1}{\tau_0-1}} \leq 50\omega^3\sqrt{\tau_0-1}\sqrt{t-1} = \psi\sqrt{t-1}.$$

We obtain that

$$\mathbb{P}\left(X_t(v) > \psi\sqrt{t-1} \text{ for some } t \geq \tau_0\right) \leq e^{-5\omega/4} \leq \exp\left(-\frac{\psi^{1/3}}{3(\tau_0-1)^{1/6}}\right).$$

We therefore have

$$P_1 \leq \tau_0\exp\left(-\frac{\psi^{1/3}}{3(\tau_0-1)^{1/6}}\right)$$

We now bound P_2. For each time $s > \tau_0$, consider the new vertex v of degree 1 born at time s. We apply Theorem 3.2 to this vertex, with $m_0 = 1$, τ_0 replaced by s, and $\omega = \left(\psi\sqrt{s-1}/50\right)^{1/3}$, so $50\omega^3 m_0\sqrt{\frac{t-1}{s-1}} \leq \psi\sqrt{t-1}$, and we have

$$\mathbb{P}\left(X_t(v) > \psi\sqrt{t-1} \text{ for some } t \geq s\right) \leq e^{-5\omega/4} \leq \exp\left(-\frac{\psi^{1/3}(s-1)^{1/6}}{3}\right).$$

Summing over s, we have

$$P_2 \leq \sum_{s=2}^{\infty}\exp\left(-\frac{\psi^{1/3}(s-1)^{1/6}}{3}\right) \leq 2\exp\left(-\psi^{1/3}/3\right).$$

The overall probability that there is, at any time t, a vertex of degree at least $\psi\sqrt{t-1}$, is thus at most $P_1 + P_2 \leq 2P_1$, as claimed.

For the final inequality, we need to show that our bounds imply that

$$f(\psi) = \psi^{1/3} - 3(\tau_0-1)^{1/6}\log(2\tau_0\psi) \geq 0.$$

We note that $f'(\psi) = \frac{1}{3\psi}\left[\psi^{1/3} - 9(\tau_0-1)^{1/6}\right] > 0$, so it is enough to verify that the inequality holds at $\psi = 10^5\sqrt{\tau_0-1}\log^3\tau_0$, at which point $2\psi \leq \tau_0^{12}$ for all $\tau_0 \geq 4$. The desired inequality is equivalent to $10^{5/3}\log(\tau_0) \geq 3\log(2\tau_0\psi)$: this holds since $10^{5/3} > 39$. □

4 Concentration for $D_t(\ell)$

In this section, we again consider the basic preferential attachment model, but now we are concerned with the number of vertices of degree exactly ℓ at time t.

Recall that, for $t \geq \tau_0$ and $\ell \in \mathbb{N}$, $D_t(\ell)$ denotes the number of vertices of degree exactly ℓ at time t. It is easy to see that $D = (D_t(\ell) : t \geq \tau_0, \ell \in \mathbb{N})$ is a Markov chain.

We recall our main theorem.

Theorem 1.1. *Let $\tau_0 \geq 4$ and $\psi \geq 10^5 \sqrt{\tau_0 - 1} \log^3 \tau_0$ be constants. Let $G(\tau_0)$ be any graph with τ_0 vertices and $\tau_0 - 1$ edges, and consider the preferential attachment process with initial graph $G(\tau_0)$ at time τ_0, and the associated Markov chain $D = (D_t(\ell) : t \geq \tau_0, \ell \in \mathbb{N})$.*

With probability at least $1 - \frac{4}{\psi}$, we have

$$\left| D_t(\ell) - \frac{4t}{\ell(\ell+1)(\ell+2)} \right| \leq 120 \sqrt{\frac{t \log(\psi t)}{\ell^3}} + 301 \psi^2 \log(\psi t),$$

for all $\ell \geq 1$, and all $t \geq \tau_0$.

As is well-known (and as we shall show shortly), the expectation of $D_t(\ell)$ is very close to $4t/\ell(\ell+1)(\ell+2)$, for all $\ell \geq 1$ and all $t \geq \tau_0$, so the theorem shows concentration of measure of these random variables about their means.

For $\ell \leq (t/\log(\psi t))^{1/3}/\psi^2$, the bound on the deviation of $D_t(\ell)$ is at most $125 \sqrt{\frac{t \log(\psi t)}{\ell^3}}$, which is, up to the log factor, on the order of $\sqrt{\mathbb{E}\, D_t(\ell)}$. We get concentration within a factor $(1 + o(1))$ of the mean as long as $\ell = o(t/\log t)^{1/3}$.

For all values of ℓ larger than $(t/\log t)^{1/3}$, the bound on the deviation that we obtain is of order $\log t$. This result might conceivably be of interest for values of ℓ between about $(t/\log t)^{1/3}$ and $t^{1/2}$, but for larger values of ℓ we already have a stronger result: Theorem 3.4 tells us that $D_t(\ell) = 0$ when ℓ is larger than $\psi\sqrt{t}$, with probability at least $1 - 1/\psi$.

The proof of Theorem 1.1 takes up the rest of this section, although we defer the bulk of the calculations until later sections.

Proof. It will shortly turn out to be convenient to truncate the range of ℓ, so that we consider only values of ℓ with $1 \leq \ell \leq \ell_0$, for some fixed ℓ_0. We remark now that we may freely do this, as we are proving an explicit bound on the probability of failure that is independent of ℓ_0.

For the moment though, we consider all values $\ell \in \mathbb{N}$ simultaneously, and consider the evolution of the entire process $D = (D_t(\ell))$ for $t \geq \tau_0$.

We have, for $t \geq \tau_0 + 1$,

$$\mathbb{E}[D_t(1) - D_{t-1}(1) \mid D_{t-1}] = 1 - \frac{D_{t-1}(1)}{2(t-1)}.$$

Also, for $\ell \geq 2$,

$$\mathbb{E}[D_t(\ell) - D_{t-1}(\ell) \mid D_{t-1}] = \frac{(\ell-1)D_{t-1}(\ell-1)}{2(t-1)} - \frac{\ell D_{t-1}(\ell)}{2(t-1)}.$$

Then, by Lemma 2.1,

$$D_t(1) = D_{\tau_0}(1) + \sum_{s=\tau_0}^{t-1} \left(1 - \frac{D_s(1)}{2s} \right) + M_t(1)$$

$$D_t(\ell) = D_{\tau_0}(\ell) + \sum_{s=\tau_0}^{t-1} \left(\frac{(\ell-1)D_s(\ell-1)}{2s} - \frac{\ell D_s(\ell)}{2s} \right) + M_t(\ell), \quad (\ell \geq 2),$$

where $M_t(\ell)$ is a martingale for each $\ell \geq 1$.

We want to show that, for $\ell \geq 1$, $D_t(\ell)$ is close to $d_t(\ell)$, where the $d_t(\ell)$ satisfy $d_{\tau_0}(\ell) = D_{\tau_0}(\ell)$ for all ℓ, and:

$$
\begin{aligned}
d_t(1) &= d_{\tau_0}(1) + \sum_{s=\tau_0}^{t-1} \left(1 - \frac{d_s(1)}{2s}\right) \\
d_t(\ell) &= d_{\tau_0}(\ell) + \sum_{s=\tau_0}^{t-1} \left(\frac{(\ell-1)d_s(\ell-1)}{2s} - \frac{\ell d_s(\ell)}{2s}\right), \quad (\ell \geq 2).
\end{aligned}
$$

Given the initial values $d_{\tau_0}(\ell) = D_{\tau_0}(\ell)$, for $\ell \geq 1$, the equations above are equivalent to:

$$
\begin{aligned}
d_t(1) &= 1 + d_{t-1}(1)\left(1 - \frac{1}{2(t-1)}\right) \\
d_t(\ell) &= d_{t-1}(\ell)\left(1 - \frac{\ell}{2(t-1)}\right) + \frac{\ell-1}{2(t-1)}d_{t-1}(\ell-1), \quad (\ell \geq 2).
\end{aligned}
$$

These equations are known to admit the explicit solution

$$
d_t(\ell) = \frac{4t}{\ell(\ell+1)(\ell+2)},
$$

if the initial conditions correspond (which of course cannot happen for a concrete graph at time τ_0, since then all the $D_{\tau_0}(\ell)$ are natural numbers). More generally, we have the following result, which is very similar to results of Szymański [22, 23] and Bollobás, Riordan, Spencer and Tusnády [4].

Lemma 4.1. *Take any $\tau_0 \geq 1$, and any sequence $(D_{\tau_0}(\ell))_{\ell \geq 1}$ of non-negative integers with $\sum_{\ell \geq 1} D_{\tau_0}(\ell) = \tau_0$. Then the solution $d_t(\ell)$ of the equations above with $d_{\tau_0}(\ell) = D_{\tau_0}(\ell)$ for all ℓ satisfies*

$$
\left|d_t(\ell) - \frac{4t}{\ell(\ell+1)(\ell+2)}\right| \leq \frac{\tau_0^{3/2}}{t^{1/2}} \leq \tau_0,
$$

for all $\ell \geq 1$ and all $t \geq \tau_0$.

Proof. Set

$$
z_t(\ell) = d_t(\ell) - \frac{4t}{\ell(\ell+1)(\ell+2)},
$$

for all $t \geq \tau_0$ and $\ell \geq 1$.

Note first that $D_{\tau_0}(\ell) \leq \tau_0$, for all ℓ, and so also $|z_{\tau_0}(\ell)| \leq \tau_0$. Thus the lemma holds for $t = \tau_0$.

For each $t > \tau_0$, it is straightforward to verify that

$$
z_t(1) = z_{t-1}(1)\left(1 - \frac{1}{2(t-1)}\right),
$$

and, for $\ell > 1$,

$$
z_t(\ell) = z_{t-1}(\ell)\left(1 - \frac{\ell}{2(t-1)}\right) + z_{t-1}(\ell-1)\frac{\ell-1}{2(t-1)}.
$$

If Z is a common upper bound on $|z_{t-1}(\ell)|$ and $|z_{t-1}(\ell-1)|$, we deduce that

$$
|z_t(\ell)| \leq Z\left(1 - \frac{\ell}{2(t-1)} + \frac{\ell-1}{2(t-1)}\right) = Z\left(1 - \frac{1}{2(t-1)}\right).
$$

By induction, it now follows that

$$|z_t(\ell)| \leq \tau_0 \prod_{u=\tau_0}^{t-1} \left(1 - \frac{1}{2u}\right) \leq \tau_0 \sqrt{\frac{\tau_0}{t}} = \frac{\tau_0^{3/2}}{t^{1/2}},$$

for all $t \geq \tau_0$ and every $\ell \geq 1$, as claimed. $\qquad\square$

Define $E_t(\ell) = D_t(\ell) - d_t(\ell)$, where, as above, we set $d_{\tau_0}(\ell) = D_{\tau_0}(\ell)$ for all $\ell \geq 1$. Note that $D_{\tau_0}(\ell)$ is an integer-valued random variable, determined by the graph at the initial time τ_0. Note also that $E_{\tau_0}(\ell) = 0$ for all $\ell \geq 1$. For the moment, we shall keep the term E_{τ_0} in our expressions, to show how the calculation would be affected in a setting where E_{τ_0} is not necessarily zero.

For $t \geq \tau_0 + 1$, we have

$$E_t(1) \;=\; E_{\tau_0}(1) - \sum_{s=\tau_0}^{t-1} \frac{E_s(1)}{2s} + M_t(1)$$

$$E_t(\ell) \;=\; E_{\tau_0}(\ell) + \sum_{s=\tau_0}^{t-1} \left(\frac{(\ell-1)E_s(\ell-1)}{2s} - \frac{\ell E_s(\ell)}{2s}\right) + M_t(\ell), \quad (\ell \geq 2).$$

This means that, for $t \geq \tau_0 + 1$,

$$E_t(1) \;=\; E_{t-1}(1)\left(1 - \frac{1}{2(t-1)}\right) + M_t(1) - M_{t-1}(1)$$

and, for $\ell \geq 2$,

$$E_t(\ell) \;=\; E_{t-1}(\ell)\left(1 - \frac{\ell}{2(t-1)}\right) + \frac{(\ell-1)E_{t-1}(\ell-1)}{2(t-1)} + M_t(\ell) - M_{t-1}(\ell).$$

At this point, we truncate the process D: we fix some $\ell_0 \geq 1$, and set $D^{\ell_0} = (D_t^{\ell_0}(\ell) : t \in \mathbb{Z}^+, \ell = 1, \ldots, \ell_0)$ – in other words, we restrict attention to the numbers of vertices with degrees at most ℓ_0. The truncated process D^{ℓ_0} remains Markov, since the distribution of $D_t(\ell)$ conditioned on D_{t-1} depends only on $D_{t-1}(\ell)$ and $D_{t-1}(\ell-1)$, for each $\ell \leq \ell_0$.

Once we have fixed ℓ_0, we may restate the previous system of equations as a matrix equation, giving a recurrence for

$$E_t = \begin{pmatrix} E_t(1) \\ \vdots \\ E_t(\ell_0) \end{pmatrix}.$$

We have, for $t \geq \tau_0 + 1$,

$$E_t = A_{t-1}E_{t-1} + \Delta M_t,$$

where

$$\Delta M_t = \begin{pmatrix} \Delta M_t(1) \\ \vdots \\ \Delta M_t(\ell_0) \end{pmatrix},$$

with $\Delta M_t(\ell) = M_t(\ell) - M_{t-1}(\ell)$ for each ℓ, and, for $s \geq \tau_0$, the matrix A_s is given by

$$\begin{pmatrix} 1 - \frac{1}{2s} & 0 & \cdots & 0 \\ \frac{1}{2s} & 1 - \frac{2}{2s} & \cdots & 0 \\ \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \frac{\ell_0-1}{2s} & 1 - \frac{\ell_0}{2s} \end{pmatrix}.$$

Hence it follows that, for $t \geq \tau_0 + 1$,

$$E_t = \left(\prod_{s=\tau_0}^{t-1} A_s\right) E_{\tau_0} + \sum_{s=\tau_0+1}^{t} \left(\prod_{u=s}^{t-1} A_u\right) \Delta M_s.$$

Here and subsequently, the notation $\prod_{s=\tau_0}^{t-1} A_s$ indicates the matrix product $A_{t-1} \cdots A_{\tau_0}$, with the indices taken in decreasing order. At this point, we recall that $E_{\tau_0} = 0$, so that

$$E_t = \sum_{s=\tau_0+1}^{t} \left(\prod_{u=s}^{t-1} A_u\right) \Delta M_s.$$

We shall control the deviations of E_t, although this process is not itself a martingale, and so we cannot directly apply our martingale deviation inequalities. The process (E_t) is a transform of the martingale (M_t), in that it is a sum of its differences, multiplied by the appropriate $\prod_{u=s}^{t-1} A_u$, which depend on t. In order to get around this difficulty, we now introduce, for each $\tau > \tau_0$, a martingale $\widetilde{M}^\tau$ stopped at τ, whose value at τ is the quantity E_τ of interest.

We fix $\tau > \tau_0$ and define, for $t \leq \tau$,

$$\widetilde{M}_t^\tau = \sum_{s=\tau_0+1}^{t} \left(\prod_{u=s}^{\tau-1} A_u\right) \Delta M_s,$$

and $\widetilde{M}_t^\tau = \widetilde{M}_\tau^\tau$ for $t > \tau$, then it is easily checked that $\widetilde{M}^\tau = (\widetilde{M}_t^\tau)$ is a martingale, and that $\widetilde{M}_\tau^\tau = E_\tau$. Thus we can obtain bounds on E_τ by studying the martingale $\widetilde{M}^\tau$.

From Lemma 4.1, we have that, for every $\tau \geq \tau_0$ and every $\ell = 1, \ldots, \ell_0$,

$$D_\tau(\ell) = d_\tau(\ell) + \widetilde{M}_\tau^\tau(\ell) \leq \frac{4\tau}{\ell^3} + \tau_0 + \widetilde{M}_\tau^\tau(\ell).$$

We now consider the transitions of the truncated process D^{ℓ_0}, with state space $(\mathbb{Z}^+)^{\ell_0}$. Recall that each transition involves the creation of one new vertex of degree 1, and the increase of a degree of an existing vertex by 1. This means that a transition of the truncated process involves an increase of 1 in $D^{\ell_0}(1)$, and either: (i) a decrease of 1 in $D^{\ell_0}(k)$ and an increase of 1 in $D^{\ell_0}(k+1)$, for some $k \in \{1, \ldots, \ell_0 - 1\}$, (ii) a decrease of 1 in $D^{\ell_0}(\ell_0)$, or (iii) no further change. In other words, the vector $D_{s+1}^{\ell_0}$ is obtained from $D_s^{\ell_0}$ by adding one of the following vectors:
(i) $y_k = e_1 - e_k + e_{k+1}$, for some $k \in \{1, \ldots, \ell_0 - 1\}$,
(ii) $y_{\ell_0} = e_1 - e_{\ell_0}$,
(iii) $y_0 = e_1$.
Here e_j denotes the standard basis vector in $\mathbb{Z}^{\ell_0}$ with a 1 in the jth coordinate and 0s elsewhere: here and in what follows, we abuse notation by suppressing the dependence on ℓ_0. The transition probabilities are then given by

$$P_s(D_s, D_s + y_k) = \frac{kD_s(k)}{2s-2} \qquad (k = 1, \ldots, \ell_0)$$

$$P_s(D_s, D_s + y_0) = 1 - \sum_{k=1}^{\ell_0} \frac{kD_s(k)}{2s-2}.$$

Here too we have removed the superscripts ℓ_0 for clarity.

We can write, for $s \geq \tau_0$,

$$
\begin{aligned}
\Delta M_{s+1} &= D_{s+1} - D_s - \sum_{k=0}^{\ell_0} P_s(D_s, D_s + y_k) y_k \\
&= \sum_{k=0}^{\ell_0} y_k [\mathbb{1}_{D_{s+1} - D_s = y_k} - P_s(D_s, D_s + y_k)].
\end{aligned}
$$

We consider running the process up to some fixed $\tau > \tau_0$: all our notation should specify the dependence on τ, but again where possible we shall suppress this.

For $\tau_0 \leq s < \tau$, we define $B_s = B_s^\tau = \prod_{u=s+1}^{\tau-1} A_u$, so that

$$
\widetilde{M}_s^\tau = \sum_{w=\tau_0}^{s-1} B_w \Delta M_{w+1}.
$$

We then have, for $\tau_0 \leq s < \tau$,

$$
\Delta \widetilde{M}_{s+1}^\tau = \widetilde{M}_{s+1}^\tau - \widetilde{M}_s^\tau = B_s \Delta M_{s+1} = B_s \sum_{k=0}^{\ell_0} y_k [\mathbb{1}_{D_{s+1} - D_s = y_k} - P_s(D_s, D_s + y_k)].
$$

Now we define $\widetilde{D} = \widetilde{D}^{\ell_0}$ by

$$
\begin{aligned}
\widetilde{D}_t &= \sum_{s=\tau_0}^{t-1} B_s \sum_{k=0}^{\ell_0} y_k P_s(D_s, D_s + y_k) \\
&\quad + \sum_{s=\tau_0}^{t-1} B_s \sum_{k=0}^{\ell_0} y_k [\mathbb{1}_{D_{s+1} - D_s = y_k} - P_s(D_s, D_s + y_k)], \quad (t \leq \tau), \\
&= \sum_{s=\tau_0}^{t-1} \sum_{k=0}^{\ell_0} P_s(D_s, D_s + y_k)[B_s y_k] + \widetilde{M}_t^\tau.
\end{aligned}
$$

so that, for $t \leq \tau$,

$$
\begin{aligned}
\widetilde{D}_t &= \sum_{s=\tau_0}^{t-1} B_s \sum_{k=0}^{\ell_0} y_k \mathbb{1}_{D_{s+1} - D_s = y_k} \\
&= \sum_{s=\tau_0}^{t-1} B_s (D_{s+1} - D_s)
\end{aligned}
$$

and so

$$
\widetilde{D}_{t+1} - \widetilde{D}_t = B_t(D_{t+1} - D_t).
$$

The process $\widetilde{D}$ is not in general a Markov process. However, we may define a process $Y = Y^\tau$ by setting $Y_t = (D_t, \widetilde{D}_t)$ for $t \leq \tau$, and $Y_t = Y_\tau$ for $t \geq \tau$. This extended process Y is Markovian, with state space $E = (\mathbb{Z}^+)^{\ell_0} \times (\mathbb{R}^+)^{\ell_0}$. At each time t with $\tau_0 \leq t < \tau$, the one-step transition matrix $\widetilde{P}_t$ for Y is derived from that of D. Specifically, if $D_{t+1} - D_t = y_k$, then $\widetilde{D}_{t+1} - \widetilde{D}_t = B_t y_k$, and $Y_{t+1} - Y_t = (y_k, B_t y_k)$.

Our plan is to apply Theorem 2.5 to the Markov process Y, and, for $\ell = 1, \ldots, \ell_0$, to the projection function $g = g^\ell$ taking $(x, \widetilde{x}) \in E$ to $\widetilde{x}(\ell)$. For each $k = 0, \ldots, \ell_0$, if $D_{t+1} - D_t = y_k$, then $g(Y_{t+1}) - g(Y_t) = [B_t y_k](\ell)$, the ℓ-th entry of the vector $B_t y_k$. Since B_t is a product of non-negative sub-stochastic matrices, it too is non-negative and sub-stochastic. The vector y_k has all its entries in $\{0, +1, -1\}$, with at most two positive and one negative entries, so each co-ordinate of the vector $B_t y_k$ is a sum of at most two

entries of B_t, minus at most one other entry. Therefore $|[B_t y_k](\ell)| \leq 1$ for all t, k and ℓ. So, in applying Theorem 2.5, we may take $J = 1$.

For $\ell = 1, \ldots, \ell_0$, and $\tau_0 \leq t < \tau$, we therefore have

$$
\begin{aligned}
\Phi_t^{g^\ell}(Y) &= \sum_{s=\tau_0}^{t} \sum_{x'} P_s(D_s, x') \left(g^\ell(x') - g^\ell(D_s)\right)^2 \\
&= \sum_{s=\tau_0}^{t} \sum_{k=0}^{\ell_0} P_s(D_s, D_s + y_k) \left([B_s y_k](\ell)\right)^2.
\end{aligned}
$$

For brevity, we set $\Phi_t^\ell(Y) = \Phi_t^{g^\ell}(Y)$ from now on.

For each $\ell = 1, \ldots, \ell_0$, we set

$$
R^\ell = 1600 \frac{\tau - 1}{\ell^3} + (10\psi)^4 \log(\psi\tau),
$$

and, for $\ell = 1, \ldots, \ell_0$, we set

$$
T_R^\ell = \inf\{t \geq \tau_0 : \Phi_t^\ell(Y) > R^\ell\}.
$$

We now apply Theorem 2.5, with $J = 1$, $R = R^\ell$, and $\omega = 9 \log(\psi\tau)$, noting that $\omega J^2 \leq R^\ell$; we obtain that, for each $\tau \geq \tau_0$ and $\ell = 1, \ldots, \ell_0$,

$$
\mathbb{P}\left(\left(\sup_{\tau_0 \leq t \leq \tau} |\widetilde{M_t^\tau}(\ell)| > 3\sqrt{\log(\psi\tau) R^\ell}\right) \wedge (T_R^\ell \geq \tau)\right) \leq \frac{2}{\psi^2 \tau^2}.
$$

Let $\widehat{T}_\Delta = \inf\{s \geq \tau_0 : D_s(k) > 0 \text{ for some } k > \psi\sqrt{s-1}\}$, the first time s such that there is a vertex of degree greater than $\psi\sqrt{s-1}$: by Theorem 3.4, $\mathbb{P}(\widehat{T}_\Delta < \infty) \leq 1/\psi$. Also, for each $k = 1, \ldots, \ell_0$, let

$$
\widehat{T}_k = \inf\left\{s \geq \tau_0 : D_s(k) > 5\frac{s}{k^3} + 400\psi^2 \log(\psi s)\right\}.
$$

Note that, for $\tau_0 \leq s < \min\left((k/\psi)^2 + 1, \widehat{T}_\Delta\right)$, $D_s(k) = 0$, and so $\widehat{T}_k \geq \min\left((k/\psi)^2, \widehat{T}_\Delta\right)$ for each $k = 1, \ldots, \ell_0$. Finally, let $\widehat{T}$ be the minimum of $\widehat{T}_1, \ldots, \widehat{T}_{\ell_0}$.

In the next section, we shall prove the following result.

Lemma 4.2. *For all $t \leq \widehat{T} \wedge \widehat{T}_\Delta$, and all $\ell = 1, \ldots, \ell_0$, $\Phi_{t-1}^\ell(Y) \leq R^\ell$.*

This result can be restated as saying that $T_R^\ell \geq \widehat{T} \wedge \widehat{T}_\Delta$ for each $\ell = 1, \ldots, \ell_0$.

Now, for each $\tau \geq \tau_0$ and each $\ell = 1, \ldots, \ell_0$, set

$$
\delta_\tau(\ell) = 120\sqrt{\frac{\tau \log(\psi\tau)}{\ell^3}} + 300\psi^2 \log(\psi\tau),
$$

which is slightly less than the bound on the deviation appearing in the statement of the theorem. Observe that, for $\tau \geq \tau_0$ and $\ell = 1, \ldots, \ell_0$,

$$
\delta_\tau(\ell)^2 \geq 14400 \frac{\tau \log(\psi\tau)}{\ell^3} + 9 \times (10\psi)^4 \log^2(\psi\tau) = 9 \log(\psi\tau) R^\ell.
$$

Thus we have

$$
\delta_\tau(\ell) \geq 3\sqrt{\log(\psi\tau) R^\ell},
$$

and so, using Lemma 4.2,

$$\mathbb{P}\left(\left(\sup_{\tau_0 \leq t \leq \tau} |\widetilde{M}_t^\tau(\ell)| > \delta_\tau(\ell)\right) \wedge (\widehat{T} \wedge \widehat{T}_\Delta \geq \tau)\right)$$

$$\leq \quad \mathbb{P}\left(\left(\sup_{\tau_0 \leq t \leq \tau} |\widetilde{M}_t^\tau(\ell)| > \delta_\tau(\ell)\right) \wedge (T_R^\ell \geq \tau)\right)$$

$$\leq \quad \frac{2}{\psi^2 \tau^2}.$$

Recall that $\widetilde{M}_\tau^\tau = E_\tau$ for each $\tau \geq \tau_0$, so we now deduce that

$$\mathbb{P}\left((|E_\tau(\ell)| > \delta_\tau(\ell)) \wedge (\widehat{T} \wedge \widehat{T}_\Delta \geq \tau)\right) \leq \frac{2}{\psi^2 \tau^2}, \tag{4.1}$$

for all $\tau \geq \tau_0$ and $\ell = 1, \ldots, \ell_0$.

We now wish to bound the total probability that there is some pair (τ, ℓ), with $\tau \geq \tau_0$ and $\ell = 1, \ldots, \ell_0$, such that $|E_\tau(\ell)| > \delta_\tau(\ell)$ and $\tau < \widehat{T} \wedge \widehat{T}_\Delta$: recall that we want a bound independent of ℓ_0.

For those pairs (τ, ℓ) with $\ell \leq \psi\sqrt{\tau - 1}$, we sum the bounds from (4.1), and obtain that

$$\mathbb{P}\left((|E_\tau(\ell)| > \delta_\tau(\ell)) \wedge (\widehat{T} \wedge \widehat{T}_\Delta > \tau)\right.$$

$$\left. \text{for some } \tau \geq \tau_0, 1 \leq \ell \leq \min(\ell_0, \psi\sqrt{\tau - 1})\right)$$

$$\leq \quad \frac{2}{\psi^2} \sum_{\tau = \tau_0}^\infty \frac{\psi\sqrt{\tau - 1}}{\tau^2}$$

$$\leq \quad \frac{2}{\psi} \int_3^\infty \tau^{-3/2} \, d\tau$$

$$\leq \quad \frac{3}{\psi}.$$

For those pairs (τ, ℓ) with $\ell > \psi\sqrt{\tau - 1}$, we have either $\widehat{T}_\Delta \leq \tau$ or $D_\tau(\ell) = 0$, and in the latter case we have

$$|E_\tau(\ell)| = d_\tau(\ell) \leq \frac{4\tau}{\ell^3} + \tau_0 < \frac{4}{\psi^2 \ell} + \frac{1}{\ell^3} + 10^4 \psi^2 < \delta_\tau(\ell).$$

Therefore, with probability at least $1 - 3/\psi$, we have, for all $\ell = 1, \ldots, \ell_0$ and all $\tau \geq \tau_0$, that either $|E_\tau(\ell)| \leq \delta_\tau(\ell)$ or $\tau \geq \widehat{T} \wedge \widehat{T}_\Delta$.

We now set $T_\ell^* = \inf\{s \geq \tau_0 : |E_s(\ell)| > \delta_s(\ell)\}$ for each $\ell = 1, \ldots, \ell_0$, and $T^* = \min(T_\ell^*, \ell = 1, \ldots, \ell_0)$; we obtain that

$$\mathbb{P}(T^* < \infty \text{ and } T^* \leq \widehat{T} \wedge \widehat{T}_\Delta) \leq 3/\psi. \tag{4.2}$$

On the other hand, if $\widehat{T}_\ell < T_\ell^*$ for some $\ell = 1, \ldots, \ell_0$, there is an $s \geq \tau_0$ with $D_s(\ell) > 5s/\ell^3 + 400\psi^2 \log(\psi s)$ and

$$\begin{aligned}
D_s(\ell) &\leq \frac{4s}{\ell^3} + \tau_0 + \delta_s(\ell) \\
&= \frac{4s}{\ell^3} + \tau_0 + 120\sqrt{\frac{s \log(\psi s)}{\ell^3}} + 300\psi^2 \log(\psi s) \\
&\leq \frac{4s}{\ell^3} + \psi^2 + \left(\frac{s}{\ell^3} + 3600 \log(\psi s)\right) + 300\psi^2 \log(\psi s) \\
&\leq 5\frac{s}{\ell^3} + 350\psi^2 \log(\psi s),
\end{aligned}$$

which is a contradiction.

We conclude that $T_\ell^* \leq \widehat{T}_\ell$ for all $\ell = 1, \ldots, \ell_0$. Recalling that $\widehat{T}$ is the minimum of $\widehat{T}_1, \ldots, \widehat{T}_{\ell_0}$ and T^* is the minimum of $T_1^*, \ldots, T_{\ell_0}^*$, this implies that $T^* \leq \widehat{T}$. Equation (4.2) now implies that $\mathbb{P}(T^* < \infty$ and $T^* \leq \widehat{T}_\Delta) \leq 3/\psi$. However, we also have that $\mathbb{P}(\widehat{T}_\Delta < \infty) \leq 1/\psi$, so $\mathbb{P}(T^* < \infty) \leq 4/\psi$.

This conclusion is equivalent to the statement that

$$\mathbb{P}\left(|E_s(\ell)| \leq \delta_s(\ell) \text{ for all } \ell \geq 1 \text{ and } s \geq \tau_0\right) \geq 1 - \frac{4}{\psi}.$$

This implies the result stated, since

$$\left| D_s(\ell) - \frac{4s}{\ell(\ell+1)(\ell+2)} \right| \leq |E_s(\ell)| + \tau_0 \leq |E_s(\ell)| + \psi^2.$$

$\square$

5 Bounds for $\Phi_{\tau-1}^\ell(Y)$

Our aim in this section is to prove Lemma 4.2, which states that

$$\Phi_{\tau-1}^\ell(Y) = \sum_{s=\tau_0}^{\tau-1} \sum_{k=0}^{\ell_0} P_s(D_s, D_s + y_k) \left([B_s y_k](\ell)\right)^2$$

is at most

$$R^\ell = 1600 \frac{\tau-1}{\ell^3} + (10\psi)^4 \log(\psi\tau),$$

whenever $\tau_0 \leq \tau \leq \widehat{T} \wedge \widehat{T}_\Delta$ and $1 \leq \ell \leq \ell_0$.

Recall that, for $s < \tau \leq \widehat{T} \wedge \widehat{T}_\Delta$, and $1 \leq k \leq \ell_0$, we have

$$D_s(k) \leq \begin{cases} 0 & k > \psi\sqrt{s-1} \\ 5\frac{s}{k^3} + 400\psi^2 \log(\psi s) & k \leq \psi\sqrt{s-1} \end{cases}. \tag{5.1}$$

For this section, we may and shall assume that we do indeed have these bounds on the values of $D_s(k)$.

Recall also that, for $s = \tau_0, \ldots, \tau-1$, B_s is the matrix product $A_{\tau-1} \cdots A_{s+1}$, and that

$$y_k = \begin{cases} e_1 & k = 0 \\ e_1 - e_k + e_{k+1} & 1 \leq k < \ell_0 \\ e_1 - e_{\ell_0} & k = \ell_0. \end{cases}$$

We may now write, for $1 \leq \ell \leq \ell_0$, and $\tau_0 \leq s < \tau$,

$$\begin{aligned}
[B_s y_0](\ell) &= [B_s e_1](\ell) \\
&= B_s(\ell, 1) \\
[B_s y_k](\ell) &= [B_s e_1](\ell) - [B_s e_k](\ell) + [B_s e_{k+1}](\ell) \\
&= B_s(\ell, 1) - B_s(\ell, k) + B_s(\ell, k+1) \quad (1 \leq k < \ell_0) \\
[B_s y_{\ell_0}](\ell) &= [B_s e_1](\ell) - [B_s e_{\ell_0}](\ell) \\
&= B_s(\ell, 1) - B_s(\ell, \ell_0).
\end{aligned}$$

where $[B_s](i, j)$ denotes the (i, j)-entry of the matrix B_s. We then have

$$\begin{aligned}
[B_s y_0](\ell)^2 &= B_s(\ell, 1)^2 \\
[B_s y_k](\ell)^2 &\leq 2B_s(\ell, 1)^2 + 2(B_s(\ell, k) - B_s(\ell, k+1))^2 \quad (1 \leq k < \ell_0) \\
[B_s y_{\ell_0}](\ell)^2 &\leq 2B_s(\ell, 1)^2 + 2B_s(\ell, \ell_0)^2.
\end{aligned}$$

Provided we interpret $B_s(\ell_0, \ell_0 + 1)$ as equal to zero, we can now bound the sum over k, for any s and any $\ell \leq \ell_0$, as

$$\sum_{k=0}^{\ell_0} P_s(D_s, D_s + y_k)([B_s y_k](\ell))^2$$

$$\leq P_s(D_s, D_s + y_0)B_s(\ell, 1)^2$$

$$+ 2\sum_{k=1}^{\ell_0} P_s(D_s, D_s + y_k) \left[B_s(\ell, 1)^2 + (B_s(\ell, k) - B_s(\ell, k+1))^2 \right]$$

$$\leq 2B_s(\ell, 1)^2 + 2\sum_{k=1}^{\ell_0} P_s(D_s, D_s + y_k)(B_s(\ell, k) - B_s(\ell, k+1))^2.$$

For $1 \leq \ell < \ell_0$, all terms in the sum with $k > \ell$ are zero, since the matrix B_s is lower-triangular, and therefore we have

$$\Phi_{\tau-1}^{\ell}(Y) \leq 2\sum_{s=\tau_0}^{\tau-1} B_s(\ell, 1)^2 + 2\sum_{s=\tau_0}^{\tau-1} \sum_{k=1}^{\ell} P_s(D_s, D_s + y_k)(B_s(\ell, k) - B_s(\ell, k+1))^2.$$

The key task is thus to estimate the entries $B_s(\ell, k)$ of the matrix product $B_s = A_{\tau-1} \cdots A_{s+1}$, and in particular the differences $|B_s(\ell, k) - B_s(\ell, k+1)|$. The recurrence satisfied by these matrix entries is that, for $0 \leq j < \ell$:

$$B_{s-1}(\ell, \ell - j) = [B_s A_s](\ell, \ell - j)$$

$$= B_s(\ell, \ell - j)A_s(\ell - j, \ell - j) + B_s(\ell, \ell - j + 1)A_s(\ell - j + 1, \ell - j),$$

since the only non-zero entries of A_s in column $(\ell - j)$ are those in rows $(\ell - j)$ and $(\ell - j + 1)$. Substituting for the values of these entries yields

$$B_{s-1}(\ell, \ell - j) = B_s(\ell, \ell - j)\left(1 - \frac{\ell - j}{2s}\right) + B_s(\ell, \ell - j + 1)\frac{\ell - j}{2s}.$$

For notational convenience, we fix $\ell \geq 1$ and write

$$a_j(s) = a_j^{(\ell)}(s) = B_s(\ell, \ell - j),$$

for $s = \tau_0, \ldots, \tau - 1$ and $j = -1, 0, \ldots, \ell - 1$.

Rewriting in terms of the $a_j(s)$ gives:

$$\Phi_{\tau-1}^{\ell}(Y) \leq 2\sum_{s=\tau_0}^{\tau-1} a_{\ell-1}(s)^2 + 2\sum_{s=\tau_0}^{\tau-1} \sum_{k=1}^{\ell} P_s(D_s, D_s + y_k)(a_{\ell-k}(s) - a_{\ell-k-1}(s))^2. \qquad (5.2)$$

The transition probabilities $P_s(D_s, D_s + y_k)$ can be expressed explicitly as $\dfrac{kD_s(k)}{2(s-1)}$ for each s and k.

The recurrence satisfied by the $a_j(s)$ is then:

$$a_j(s-1) = \frac{\ell - j}{2s}a_{j-1}(s) + \left(1 - \frac{\ell - j}{2s}\right)a_j(s),$$

for $0 \leq j \leq \ell - 1$, and $\tau_0 \leq s \leq \tau - 1$. We also have $B_{\tau-1} = I$, the identity matrix, so that $a_0(\tau - 1) = 1$, and $a_j(\tau - 1) = 0$ for $j > 0$. Note also that $a_{-1}(s) = 0$ for all s, since the matrix B_s is lower-triangular. These boundary conditions, together with the recurrence relation, suffice to determine all the values $a_j(s)$.

There is a natural interpretation of the term $a_j(s)$: it is the probability that a fixed vertex v with degree $\ell - j$ at time s will have degree ℓ at time $\tau - 1$. This can most easily be seen by checking that this system of probabilities satisfies the boundary conditions and the recurrence relation. In the notation of Section 3,

$$a_j(s) = \mathbb{P}(X_{\tau-1}(v) = \ell \mid X_s(v) = \ell - j).$$

One immediate consequence is that $0 \leq a_j(s) \leq 1$ for all j and s.

It may be of interest to note that there is a formula for the $a_j(s)$ as an alternating sum:

$$a_j(s) = \binom{\ell - 1}{j} \sum_{i=\ell-j}^{\ell} \binom{j}{\ell - i} (-1)^{i-\ell+j} \prod_{u=s+1}^{\tau-1} \left(1 - \frac{i}{2u}\right).$$

One may verify that this formula satisfies the recurrence. It can also be obtained by observing that the matrices A_s can be simultaneously diagonalised, leading to a formula for the matrix B_s. We also obtain

$$a_{\ell-k}(s) - a_{\ell-k-1}(s) = \binom{\ell - 1}{k} \frac{1}{\ell - k} \sum_{i=k}^{\ell} i \binom{\ell - k}{i - k} (-1)^{i-k} \prod_{u=s+1}^{\tau-1} \left(1 - \frac{i}{2u}\right).$$

Although these formulae are quite appealing, we have been unable to extract useful bounds from them.

At this point, we break into three cases. The main case of interest is when $8 \leq \ell \leq 2\psi\sqrt{\tau - 1}$, but we also need to deal with values of ℓ outside this range, and we do this first.

For $\ell \leq 7$, all we have to do is note that, from (5.2),

$$\Phi_{\tau-1}^{\ell}(Y) \leq 4(\tau - 1) \leq 1600 \frac{\tau - 1}{\ell^3}.$$

Now suppose that $\ell > 2\psi\sqrt{\tau - 1}$. By assumption, whenever $k > \psi\sqrt{s - 1}$, we have $D_s(k) = 0$ and so $P_s(D_s, D_s + y_k) = 0$, and such terms contribute nothing to the double sum. We now need to bound the contribution of terms where $k \leq \psi\sqrt{s - 1}$ and $\ell > 2\psi\sqrt{\tau - 1}$. To do this, we use the inequality $(a_{\ell-k}(s) - a_{\ell-k-1}(s))^2 \leq a_{\ell-k}(s)^2 + a_{\ell-k-1}(s)^2$, and bound the size of any term $a_{\ell-k}(s)$ subject to the given conditions. For this, we observe that

$$\begin{aligned}
a_{\ell-k}(s) &\leq \mathbb{P}(X_{\tau-1}(v) \geq \ell \mid X_s(v) = k) \\
&\leq \mathbb{P}\left(X_{\tau-1}(v) \geq 2\psi\sqrt{\tau - 1} \mid X_s(v) = \lfloor \psi\sqrt{s - 1} \rfloor\right).
\end{aligned}$$

We now apply Corollary 3.3, with τ_0 replaced by s, m_0 replaced by $\lfloor \psi\sqrt{s - 1} \rfloor$, and ω replaced by $(s - 1)^{1/14}$. We have $\lfloor \psi\sqrt{s - 1} \rfloor \geq 10^5(s - 1)^{7/14}$, since $\psi > 2 \times 10^5$. We also have $(s - 1)^{1/14} \geq 4$ provided $s > 2^{28}$. So, for $s > 2^{28}$, we have

$$a_{\ell-k}(s) \leq e^{-\frac{5}{4}(s-1)^{1/14}}.$$

Thus, for each $s > 2^{28}$,

$$\sum_{k=1}^{\lfloor \psi\sqrt{s-1} \rfloor} P_s(D_s, D_s + y_k) a_{\ell-k}(s)^2 \leq e^{-\frac{5}{2}(s-1)^{1/14}}.$$

For $s < 2^{28}$, we have

$$\sum_{k=1}^{\lfloor \psi\sqrt{s-1} \rfloor} P_s(D_s, D_s + y_k) a_{\ell-k}(s)^2 \leq 1.$$

Therefore

$$
\begin{aligned}
\Phi_{\tau-1}^{\ell}(Y) &\leq 2\sum_{s=\tau_0}^{\tau-1}\left[a_{\ell-1}(s)^2 + 2\sum_{k=1}^{\lfloor\psi\sqrt{s-1}\rfloor}P_s(D_s, D_s + y_k)a_{\ell-k}(s)^2\right]\\
&\leq 6\left[2^{28} + \sum_{s=2^{28}+1}^{\infty}e^{-\frac{5}{2}(s-1)^{1/14}}\right]\\
&\leq 2\times 10^9,
\end{aligned}
$$

which, since $\psi \geq 2 \times 10^5$, is at most ψ^2. This comfortably gives the required result in the case where $\ell > 2\psi\sqrt{\tau-1}$.

For the remainder of this section, we assume that $8 \leq \ell \leq 2\psi\sqrt{\tau-1}$.

Although our exact expression for the $a_j(s)$ proved difficult to work with, we now give a function $f_j(s)$ which has a simple form, and which satisfies the boundary conditions and an approximate version of the recurrence; our plan is to show that $a_j(s)$ is close to $f_j(s)$ for all values of j and s.

For $0 \leq j \leq \ell-1$ and $0 \leq s \leq \tau-1$, set

$$
f_j(s) = \binom{\ell-1}{j}\left(1-\sqrt{\frac{s}{\tau-1}}\right)^j\sqrt{\frac{s}{\tau-1}}^{\,\ell-j}.
$$

Throughout what follows, we shall set $v = v_s = \sqrt{s/(\tau-1)}$, so

$$
f_j(s) = \binom{\ell-1}{j}(1-v)^j v^{\ell-j}.
$$

We note that $v_{\tau-1} = 1$, and so $f_j(\tau-1) = 0$ for $j \neq 0$, while $f_0(\tau-1) = 1$. We could formally define the function f_{-1} to be identically 0: the key identity we use for the binomial coefficients is $\binom{\ell-1}{j-1} = \binom{\ell-1}{j}\frac{j}{\ell-j}$, which indeed entails $\binom{\ell-1}{-1} = 0$. However, we find ourselves having to deal with the case $j = 0$ as a boundary case separately anyway, and so we need make no (further) explicit mention of the case $j = -1$.

We claim that

$$
f_j(s-1) = \frac{\ell-j}{2s}f_{j-1}(s) + \left(1-\frac{\ell-j}{2s}\right)f_j(s) + \left[f_j(s-1) - f_j(s) + f_j'(s)\right],
$$

for all $j \geq 1$ and $1 \leq s \leq \tau-1$. Our aim will then be to show that the term in square brackets is usually small, and that this is thus a good approximation to the recurrence satisfied by the $a_j(s)$. Rearranging the claimed identity, we see that it is equivalent to

$$
f_j'(s) = \frac{\ell-j}{2s}(f_j(s) - f_{j-1}(s)). \tag{5.3}
$$

To verify this identity, we write

$$
\begin{aligned}
f_j'(s) &= \binom{\ell-1}{j}v'(s)\frac{d}{dv}\left((1-v)^j v^{\ell-j}\right)\Big|_{v=v_s}\\
&= \binom{\ell-1}{j}\frac{v}{2s}(1-v)^{j-1}v^{\ell-j-1}\left((\ell-j)(1-v)-jv\right) \tag{5.4}\\
&= \frac{\ell-j}{2s}(1-v)^{j-1}v^{\ell-j}\left(\binom{\ell-1}{j}(1-v)-\frac{j}{\ell-j}\binom{\ell-1}{j}v\right)\\
&= \frac{\ell-j}{2s}\left(\binom{\ell-1}{j}(1-v)^j v^{\ell-j}-\binom{\ell-1}{j-1}(1-v)^{j-1}v^{\ell-j+1}\right)\\
&= \frac{\ell-j}{2s}(f_j(s) - f_{j-1}(s)).
\end{aligned}
$$

Equation (5.3) demonstrates that the $f_j(s)$ are the analogues to the $a_j(s)$ for a continuous time version of the preferential attachment process. In this continuous time version, at time s, each vertex of degree k attracts a new edge (whose other endpoint is a new vertex of degree 1) at rate $k/2s$, independent of the degrees of other vertices. The degree of a given vertex is then a pure birth process with this transition rate. The probability that a vertex with degree $\ell - j$ at time s has degree ℓ at time $\tau - 1$ satisfies the differential equation (5.3), as well as the boundary condition $f_j(\tau - 1) = \delta_{j0}$.

It seems intuitively plausible that the difference $e_j(s) = f_j(s) - a_j(s)$ between the continuous and the discrete "solutions" will always be small. Indeed we shall prove the following lemma, which is very crude in most ranges.

Lemma 5.1. *For all $\ell \geq 8$ and $0 \leq j \leq \ell - 1$, we have:*

$$|e_j(s)| \leq \begin{cases} \frac{2200\ell}{\tau - 1} & (\tau - 1)/\ell^2 < s \leq \tau - 1 \text{ or } j \leq \ell - 2 \\ \frac{800\ell^{3/2}}{\tau - 1} & (\tau - 1)/\ell^3 < s \leq (\tau - 1)/\ell^2 \\ 1 & \tau_0 \leq s \leq (\tau - 1)/\ell^3. \end{cases}$$

We shall defer the proof of Lemma 5.1 to the next section.

We set

$$\Psi_{\tau-1}^\ell(Y) = 4 \sum_{s=\tau_0}^{\tau-1} f_{\ell-1}(s)^2 + 4 \sum_{s=\tau_0}^{\tau-1} \sum_{k=1}^{\ell} P_s(D_s, D_s + y_k)(f_{\ell-k}(s) - f_{\ell-k-1}(s))^2.$$

We now show that the bound in Lemma 5.1 suffices to show that $\Phi_{\tau-1}^\ell(Y)$ is not much larger than $\Psi_{\tau-1}^\ell(Y)$.

Lemma 5.2. *For any ℓ and τ, with $8 \leq \ell \leq 2\psi\sqrt{\tau - 1}$,*

$$\Phi_{\tau-1}^\ell(Y) \leq \Psi_{\tau-1}^\ell(Y) + 5 \times 10^8 \psi^2 + 20\frac{\tau - 1}{\ell^3}.$$

Proof. Equation (5.2) tells us that $\Phi_{\tau-1}^\ell(Y)$ is at most

$$2 \sum_{s=\tau_0}^{\tau-1} a_{\ell-1}(s)^2 + 2 \sum_{s=\tau_0}^{\tau-1} \sum_{k=1}^{\ell} P_s(D_s, D_s + y_k)(a_{\ell-k}(s) - a_{\ell-k-1}(s))^2.$$

Using the inequalities $a_j(s)^2 \leq 2f_j(s)^2 + 2e_j(s)^2$ and

$$(a_j(s) - a_{j-1}(s))^2 \leq 2(f_j(s) - f_{j-1}(s))^2 + 4e_j(s)^2 + 4e_{j-1}(s)^2,$$

we deduce that

$$\begin{aligned} \Phi_{\tau-1}^\ell(Y) &\leq \Psi_{\tau-1}^\ell(Y) + 4 \sum_{s=\tau_0}^{\tau-1} e_{\ell-1}(s)^2 \\ &\quad + 8 \sum_{s=\tau_0}^{\tau-1} \sum_{k=1}^{\ell} P_s(D_s, D_s + y_k) \left(e_{\ell-k}(s)^2 + e_{\ell-k-1}(s)^2 \right). \end{aligned}$$

Now we apply the bounds from Lemma 5.1:

$$\begin{aligned} \sum_{s=\tau_0}^{\tau-1} e_{\ell-1}(s)^2 &\leq (\tau - 1)\left(\frac{2200\ell}{\tau - 1}\right)^2 + \frac{\tau - 1}{\ell^2}\left(\frac{800\ell^{3/2}}{\tau - 1}\right)^2 + \frac{\tau - 1}{\ell^3} \\ &\leq \frac{5 \times 10^6 \ell^2}{\tau - 1} + \frac{10^6 \ell}{\tau - 1} + \frac{\tau - 1}{\ell^3} \\ &\leq \frac{6 \times 10^6 \ell^2}{\tau - 1} + \frac{\tau - 1}{\ell^3}, \end{aligned}$$

and similarly

$$\sum_{s=\tau_0}^{\tau-1} \sum_{k=1}^{\ell} P_s(D_s, D_s + y_k)\left(e_{\ell-k}(s)^2 + e_{\ell-k-1}(s)^2\right)$$

$$\leq 2\left((\tau-1)\left(\frac{2200\ell}{\tau-1}\right)^2 + \frac{\tau-1}{\ell^2}\left(\frac{800\ell^{3/2}}{\tau-1}\right)^2 + \frac{\tau-1}{\ell^3}\right)$$

$$\leq 2\left(\frac{6\times10^6\ell^2}{\tau-1} + \frac{\tau-1}{\ell^3}\right).$$

Therefore

$$\Phi_{\tau-1}^{\ell}(Y) \leq \Psi_{\tau-1}^{\ell}(Y) + 20\left(\frac{6\times10^6\ell^2}{\tau-1} + \frac{\tau-1}{\ell^3}\right)$$

$$\leq \Psi_{\tau-1}^{\ell}(Y) + \frac{120\times10^6\ell^2}{\tau-1} + 20\frac{\tau-1}{\ell^3}$$

$$\leq \Psi_{\tau-1}^{\ell}(Y) + 5\times10^8\psi^2 + 20\frac{\tau-1}{\ell^3},$$

as claimed. □

For $k = 1, \ldots, \ell$, we have that $P_s(D_s, D_s + y_k) = kD_s(k)/2(s-1)$, since each of the $D_s(k)$ vertices of degree k has probability $k/2(s-1)$ of receiving an extra edge at time $s+1$. Therefore

$$\Psi_{\tau-1}^{\ell}(Y) = 4\sum_{s=\tau_0}^{\tau-1} f_{\ell-1}(s)^2 + 2\sum_{s=\tau_0}^{\tau-1}\sum_{k=1}^{\ell}\frac{kD_s(k)}{s-1}(f_{\ell-k}(s) - f_{\ell-k-1}(s))^2.$$

The double sum is the main term here, and we mainly concentrate on this; we will obtain adequate bounds on $\sum_s f_{\ell-1}(s)^2$ as a byproduct of our estimates.

Recall our assumptions (5.1) that $D_s(k) = 0$ for all $k > \psi\sqrt{s-1}$, and that $D_s(k) \leq 5\frac{s}{k^3} + 400\psi^2\log(\psi s)$ for all $k = 1, \ldots, \ell_0$ with $k \leq \psi\sqrt{s-1}$. Using these bounds, we find that, for all $k = 1, \ldots, \ell_0$, and all $s \geq 4$,

$$\frac{kD_s(k)}{s-1} \leq \frac{7}{k^2} + 550\psi^2\log(\psi s)\frac{k}{s}.$$

Thus we have

$$\Psi_{\tau-1}^{\ell}(Y) \leq 4\sum_{s=\tau_0}^{\tau-1} f_{\ell-1}(s)^2$$

$$+ 14\sum_{s=\tau_0}^{\tau-1}\sum_{k=1}^{\ell}\frac{1}{k^2}(f_{\ell-k}(s) - f_{\ell-k-1}(s))^2$$

$$+ 1100\psi^2\log(\psi\tau)\sum_{s=\tau_0}^{\tau-1}\sum_{k=1}^{\ell}\frac{k}{s}(f_{\ell-k}(s) - f_{\ell-k-1}(s))^2.$$

We define

$$Q_1(\tau, \ell) = \sum_{s=\tau_0}^{\tau-1}\sum_{k=1}^{\ell}\frac{1}{k^2}(f_{\ell-k}(s) - f_{\ell-k-1}(s))^2$$

$$Q_2(\tau, \ell) = \sum_{s=\tau_0}^{\tau-1}\sum_{k=1}^{\ell}\frac{k}{s}(f_{\ell-k}(s) - f_{\ell-k-1}(s))^2$$

so that

$$\Psi_{\tau-1}^{\ell}(Y) \leq 4 \sum_{s=\tau_0}^{\tau-1} f_{\ell-1}(s)^2 + 14Q_1(\tau,\ell) + 1100\psi^2 \log(\psi\tau)Q_2(\tau,\ell). \tag{5.5}$$

To estimate Q_1, we exchange the order of summation and substitute $j = \ell - k$:

$$Q_1(\tau,\ell) = \sum_{j=0}^{\ell-1} \frac{1}{(\ell-j)^2} \sum_{s=\tau_0}^{\tau-1} (f_j(s) - f_{j-1}(s))^2.$$

From (5.3) and (5.4), we have

$$f_j(s) - f_{j-1}(s) = \binom{\ell-1}{j} \frac{1}{\ell-j} (1-v)^{j-1} v^{\ell-j} (\ell(1-v)-j),$$

where $v = v_s = \sqrt{s/(\tau-1)}$, as before. We estimate the sum over s by approximating it by the integral

$$\int_{s=\tau_0}^{\tau-1} \binom{\ell-1}{j}^2 \frac{1}{(\ell-j)^2} (1-v)^{2j-2} v^{2\ell-2j} (\ell(1-v)-j)^2 \, ds.$$

The integrand here is bounded above by 1, since each $f_j(s)$ is at most 1. The function $(1-v)^{j-1} v^{\ell-j} (\ell(1-v)-j)$ has derivative which is a positive multiple of a quadratic function of v, so the function has just two stationary points, one either side of the zero $v = (\ell-j)/\ell$. Therefore the integrand, which is a positive multiple of the square of this function, has two local maxima. The sum is then at most the value of the integral plus the values of the integrand at the two local maxima, and so

$$Q_1(\tau,\ell) \leq 2 \sum_{j=0}^{\ell-1} \frac{1}{(\ell-j)^2}$$

$$+ \sum_{j=0}^{\ell-1} \frac{1}{(\ell-j)^4} \binom{\ell-1}{j}^2 \int_{s=\tau_0}^{\tau-1} (1-v)^{2j-2} v^{2\ell-2j} (\ell(1-v)-j)^2 \, ds$$

$$\leq 4 + \sum_{j=0}^{\ell-1} \frac{2(\tau-1)}{(\ell-j)^4} \binom{\ell-1}{j}^2 \int_{v=0}^{1} (1-v)^{2j-2} v^{2\ell-2j+1} (\ell(1-v)-j)^2 \, dv.$$

In the last line, we changed variable: recall that $s = v^2(\tau-1)$.

We write

$$Q_1(\tau,\ell) \leq 4 + \sum_{j=0}^{\ell-1} \frac{2(\tau-1)}{(\ell-j)^4} \binom{\ell-1}{j}^2 I(\ell,j,1),$$

where

$$I(\ell,j,\alpha) = \int_{v=0}^{1} (1-v)^{2j-2} v^{2\ell-2j+\alpha} (\ell(1-v)-j)^2 \, dv,$$

for positive integers ℓ and j, and integer α, where $\ell > j$ and $\alpha \geq -1$.

The integral above can be evaluated as a sum of Beta functions. We will be confronted by a very similar integral when estimating Q_2, and it is convenient to prove a lemma covering both cases (here we need $\alpha = 1$ and later we shall take $\alpha = -1$).

Lemma 5.3. *For integers ℓ and j with $\ell > j \geq 0$, and integer $\alpha \geq -1$,*

$$I(\ell,j,\alpha) \leq \frac{(2\ell-2j+\alpha)!(2j-2)!}{(2\ell+\alpha+1)!} j\ell \{2(\ell-j+1) + \alpha(3+\alpha)\} \quad (j \geq 1)$$

and

$$I(\ell,0,\alpha) \leq \frac{\ell}{2}.$$

Proof. For non-negative integers a and b, we have the identity

$$\int_{v=0}^{1} (1-v)^a v^b \, dv = B(a+1, b+1) = \frac{a! \, b!}{(a+b+1)!},$$

where $B(\cdot, \cdot)$ denotes the Beta function.

For $j \geq 1$, the required integral can be written as a sum of three integrals of the form above, and we obtain

$$
\begin{aligned}
I(\ell, &j, \alpha) \\
&= (2\ell - 2j + \alpha)! \left(\ell^2 \frac{(2j)!}{(2\ell + \alpha + 1)!} - 2\ell j \frac{(2j-1)!}{(2\ell + \alpha)!} + j^2 \frac{(2j-2)!}{(2\ell + \alpha - 1)!} \right) \\
&= (2\ell - 2j + \alpha)! \frac{(2j-2)!}{(2\ell + \alpha + 1)!} \\
&\quad \times \left\{ \ell^2 (2j-1)(2j) - 2j\ell(2j-1)(2\ell + \alpha + 1) + j^2 (2\ell + \alpha + 1)(2\ell + \alpha) \right\} \\
&= \frac{(2\ell - 2j + \alpha)!(2j-2)!}{(2\ell + \alpha + 1)!} j \left\{ 2\ell(\ell - j + 1) + \alpha(j(1+\alpha) + 2\ell) \right\} \\
&\leq \frac{(2\ell - 2j + \alpha)!(2j-2)!}{(2\ell + \alpha + 1)!} j\ell \left\{ 2(\ell - j + 1) + \alpha(3 + \alpha) \right\},
\end{aligned}
$$

as claimed.

For $j = 0$, we have

$$I(\ell, 0, \alpha) = \ell^2 \int_{v=0}^{1} v^{2\ell + \alpha} \, dv = \frac{\ell^2}{2\ell + \alpha + 1} \leq \frac{\ell}{2},$$

for all $\alpha \geq -1$, also as claimed. $\qquad \square$

Lemma 5.3, with $\alpha = 1$, tells us that

$$
\begin{aligned}
Q_1(\tau, \ell) &\leq 4 + 2\frac{\tau - 1}{\ell^4} I(\ell, 0, 1) \\
&\quad + 2(\tau - 1) \sum_{j=1}^{\ell-1} \frac{j\ell\{2(\ell - j + 1) + 4\}}{(\ell - j)^4} \binom{\ell - 1}{j}^2 \frac{(2\ell - 2j + 1)! \, (2j - 2)!}{(2\ell + 2)!} \\
&= 4 + \frac{\tau - 1}{\ell^3} + 4(\tau - 1) \sum_{j=1}^{\ell-1} \frac{j\ell(\ell - j + 3)}{(\ell - j)^4} \frac{(\ell - 1)!^2}{(2\ell + 2)!} \frac{(2\ell - 2j + 1)!}{(\ell - j - 1)!^2} \frac{(2j - 2)!}{j!^2} \\
&= 4 + \frac{\tau - 1}{\ell^3} + 4(\tau - 1) \sum_{j=1}^{\ell-1} \binom{2\ell}{\ell}^{-1} \frac{1}{\ell(2\ell + 2)(2\ell + 1)} \\
&\qquad \binom{2\ell - 2j}{\ell - j} \frac{(\ell - j + 3)(2\ell - 2j + 1)}{(\ell - j)^2} \binom{2j}{j} \frac{j}{2j(2j - 1)}.
\end{aligned}
$$

This is the first of several occasions in the paper where we use the inequalities

$$\frac{2^{2x}}{2\sqrt{x}} \leq \binom{2x}{x} \leq \frac{2^{2x}}{\sqrt{x+1}};$$

the first is valid for all integers $x \geq 1$, and the second for all integers $x \geq 0$. Sometimes, as below, we use simply that $\binom{2x}{x} \leq \frac{2^{2x}}{\sqrt{x}}$.

We obtain

$$
\begin{aligned}
Q_1(\tau, \ell) &\leq 4 + \frac{\tau - 1}{\ell^3} + 4(\tau - 1) \sum_{j=1}^{\ell-1} \frac{2\sqrt{\ell}}{2^{2\ell}} \frac{1}{4\ell^3} \frac{2^{2\ell-2j}}{\sqrt{\ell-j}} 12 \frac{2^{2j}}{\sqrt{j}} \frac{1}{2j} \\
&= 4 + \frac{\tau - 1}{\ell^3} + \frac{12(\tau - 1)}{\ell^3} \sum_{j=1}^{\ell-1} \frac{1}{j^{3/2}} \sqrt{\frac{\ell}{\ell-j}}.
\end{aligned}
$$

To estimate the sum appearing above, we use the numerical value $\sum_{j=1}^{\infty} j^{-3/2} \leq 2.61238$, and the crude bound

$$
\sum_{j=1}^{\ell-1} \frac{1}{j^{3/2}} \sqrt{\frac{\ell}{\ell-j}} \leq 2 \sum_{j=1}^{\ell/2} \frac{1}{j^{3/2}} \sqrt{\frac{\ell}{\ell-j}} \leq 2\sqrt{2} \sum_{j=1}^{\infty} \frac{1}{j^{3/2}} \leq 8,
$$

and obtain

$$
Q_1(\tau, \ell) \leq 4 + \frac{100(\tau - 1)}{\ell^3}. \tag{5.6}
$$

The next step is to estimate

$$
\begin{aligned}
Q_2(\tau, \ell) &= \sum_{s=\tau_0}^{\tau-1} \sum_{k=1}^{\ell} \frac{k}{s} \left(f_{\ell-k}(s) - f_{\ell-k-1}(s) \right)^2 \\
&\leq \sum_{j=0}^{\ell-1} (\ell - j) \sum_{s=\tau_0}^{\tau-1} \frac{1}{s} \left(f_j(s) - f_{j-1}(s) \right)^2.
\end{aligned}
$$

As before, we shall start by fixing j, and estimating the sum over s by the integral

$$
\begin{aligned}
&\int_{s=\tau_0}^{\tau-1} \frac{1}{s} \left(f_j(s) - f_{j-1}(s) \right)^2 ds \\
&\leq \binom{\ell-1}{j}^2 \frac{1}{(\ell-j)^2} \int_{v=0}^{1} \frac{(1-v)^{2j-2} v^{2\ell-2j} (\ell(1-v) - j)^2}{v^2(\tau-1)} 2(\tau-1) v \, dv \\
&= 2 \binom{\ell-1}{j}^2 \frac{1}{(\ell-j)^2} I(\ell, j, -1).
\end{aligned}
$$

We used the expression for $f_j(s) - f_{j-1}(s)$ derived earlier, and made the substitution $s = v^2(\tau - 1)$.

To bound the difference between the sum $\sum_{s=\tau_0}^{\tau-1} \frac{1}{s} \left(f_j(s) - f_{j-1}(s) \right)^2$ and the corresponding integral is not completely straightforward. The integrand $\frac{1}{s} \left(f_j(s) - f_{j-1}(s) \right)^2$ can be written as

$$
H_j(v) = \frac{h_j(v)^2}{\tau - 1}, \text{ where } h_j(v) = \frac{1}{\ell-j} \binom{\ell-1}{j} (1-v)^{j-1} v^{\ell-j-1} (\ell(1-v) - j).
$$

The function $h_j(v)$ has stationary points at

$$
v_* = \frac{\ell - j}{\ell} \pm \frac{1}{\ell} \sqrt{\frac{j(\ell - j)}{\ell - 1}}.
$$

Therefore $H_j(v)$ has a global minimum at $v = (\ell - j)/\ell$, and local maxima at the two points v_*, and so the global maximum of $H_j(v)$ is attained at one of the v_*. For $j \geq 1$,

we can write

$$
\begin{aligned}
|h_j(v_*)| &= \frac{\ell-1}{j(\ell-j)}\left[\binom{\ell-2}{j-1}(1-v_*)^{j-1}v_*^{\ell-j-1}\right]|\ell(1-v_*)-j| \\
&\leq \frac{\ell-1}{j(\ell-j)}\sqrt{\frac{j(\ell-j)}{\ell-1}} \\
&= \sqrt{\frac{\ell-1}{j(\ell-j)}},
\end{aligned}
$$

where we used the fact that the term in square brackets is the probability that a Binomial random variable with parameters $(\ell-2, v_*)$ takes the value $\ell-j-1$, and is therefore at most 1. Hence we have

$$
H_j(v) \leq \frac{\ell-1}{j(\ell-j)}\frac{1}{\tau-1},
$$

for all $j \geq 1$ and all v. For $j=0$, the maximum value of $h_j(v)$ is 1, and thus $H_0(v)$ is at most $\frac{1}{\tau-1}$ for all v.

We have that

$$
\begin{aligned}
Q_2(\tau,\ell) &\leq \sum_{j=0}^{\ell-1}(\ell-j)\left[2\binom{\ell-1}{j}^2\frac{1}{(\ell-j)^2}I(\ell,j,-1)+2\max_v H_j(v)\right] \\
&\leq 2\sum_{j=0}^{\ell-1}\binom{\ell-1}{j}^2\frac{1}{\ell-j}I(\ell,j,-1)+\frac{2}{\tau-1}\left\{\ell+\sum_{j=1}^{\ell-1}\frac{\ell-1}{j}\right\} \\
&\leq 2\sum_{j=0}^{\ell-1}\binom{\ell-1}{j}^2\frac{1}{\ell-j}I(\ell,j,-1)+\frac{2\ell^2}{\tau-1}.
\end{aligned}
$$

Now we use the bounds for $I(\ell,j,-1)$ from Lemma 5.3. We also use that $\ell \leq 2\psi\sqrt{\tau-1}$, and $\psi \geq 3$, to obtain:

$$
\begin{aligned}
Q_2(\tau,\ell) &\leq \frac{2\ell^2}{\tau-1}+\frac{1}{\ell}\frac{\ell}{2} \\
&\quad +2\sum_{j=1}^{\ell-1}\binom{\ell-1}{j}^2\frac{1}{\ell-j}\frac{(2\ell-2j-1)!\,(2j-2)!}{(2\ell)!}2j\ell(\ell-j) \\
&= 8\psi^2+\frac{1}{2}+4\sum_{j=1}^{\ell-1}\frac{(\ell-1)!^2}{j!^2(\ell-j-1)!^2}\frac{(2\ell-2j-1)!\,(2j-2)!}{(2\ell)!}j\ell \\
&= 8\psi^2+\frac{1}{2}+4\sum_{j=1}^{\ell-1}\binom{2\ell}{\ell}^{-1}\frac{1}{\ell}\binom{2\ell-2j-2}{\ell-j-1}(2\ell-2j-1)\binom{2j-2}{j-1}\frac{1}{j} \\
&\leq 8\psi^2+\frac{1}{2}+4\sum_{j=1}^{\ell-1}\left(\frac{2\ell-2j-1}{j\ell}\right)\left(\frac{2\sqrt{\ell}}{2^{2\ell}}\right)\left(\frac{2^{2\ell-2j-2}}{\sqrt{\ell-j}}\right)\left(\frac{2^{2j-2}}{\sqrt{j}}\right) \\
&\leq 8\psi^2+\frac{1}{2}+\sum_{j=1}^{\ell-1}\frac{1}{j^{3/2}}\sqrt{\frac{\ell-j}{\ell}} \\
&\leq 8\psi^2+5 \leq 9\psi^2. \qquad\qquad (5.7)
\end{aligned}
$$

The next task is to bound the sum

$$
\sum_{s=\tau_0}^{\tau-1}f_{\ell-1}(s)^2 = \sum_{s=\tau_0}^{\tau-1}(1-v)^{2\ell-2}v^2,
$$

where, as before, $v = \sqrt{s/\tau - 1}$. This sum is bounded above by the integral $\int_{s=0}^{\tau-1}(1 - v)^{2\ell-2}v^2\,dv$, plus the maximum value of the integrand. The integral is equal to

$$2(\tau - 1)\int_{v=0}^{1}(1 - v)^{2\ell-2}v^3\,dv = 2(\tau - 1)\frac{(2\ell - 2)!3!}{(2\ell + 2)!} \leq \frac{12(\tau - 1)}{(2\ell)^4},$$

which is more than small enough for our purposes, and the integrand is certainly at most 1, so

$$\sum_{s=\tau_0}^{\tau-1} f_{\ell-1}(s)^2 \leq 1 + \frac{\tau - 1}{\ell^4}. \tag{5.8}$$

Finally, we combine all our estimates. For any $\ell = 1,\dots,\ell_0$ we have, by Lemma 5.2, (5.5), (5.6), (5.7) and (5.8), that

$$
\begin{aligned}
\Phi_{\tau-1}^{\ell}(Y) &\leq& \Psi_{\tau-1}^{\ell}(Y) + 5 \times 10^8 \psi^2 + 20\frac{\tau - 1}{\ell^3} \\
&\leq& 4\sum_{s=\tau_0}^{\tau-1} f_{\ell-1}(s)^2 + 14Q_1(\tau,\ell) + 1100\psi^2\log(\psi\tau)Q_2(\tau,\ell) \\
&& + 5 \times 10^8 \psi^2 + 20\frac{\tau - 1}{\ell^3} \\
&\leq& 4 + \frac{4(\tau - 1)}{\ell^4} + 56 + 1400\frac{\tau - 1}{\ell^3} + 9900\psi^4\log(\psi\tau) \\
&& + 5 \times 10^8 \psi^2 + 20\frac{\tau - 1}{\ell^3} \\
&<& 1600\frac{\tau - 1}{\ell^3} + 10^4\psi^4\log(\psi\tau).
\end{aligned}
$$

Thus $\Phi_{\tau-1}^{\ell}(Y) < R^{\ell}$, as required. This completes the proof of Lemma 4.2, except for the proof of Lemma 5.1, to which the next section is devoted.

6 Proof of Lemma 5.1

Our aim in this section is to prove the following upper bound on $e_j(s) = f_j(s) - a_j(s)$, to be valid whenever $\ell \geq 8$, $0 \leq j \leq \ell - 1$ and $\tau_0 \leq s \leq \tau - 1$:

$$
|e_j(s)| \leq \begin{cases} \frac{2200\ell}{\tau-1} & (\tau - 1)/\ell^2 < s \leq \tau - 1 \text{ or } j \leq \ell - 2 \\ \frac{800\ell^{3/2}}{\tau-1} & (\tau - 1)/\ell^3 < s \leq (\tau - 1)/\ell^2 \\ 1 & \tau_0 \leq s \leq (\tau - 1)/\ell^3. \end{cases}
$$

The final case is straightforward, since both $a_j(s)$ and $f_j(s)$ lie between 0 and 1 for all $0 \leq j \leq \ell - 1$ and $\tau_0 \leq s \leq \tau - 1$. So from now on we assume that, if $j = \ell - 1$, then $s > (\tau - 1)/\ell^3$.

Recall that $a_j(s)$ and $f_j(s)$ satisfy the recurrences:

$$
\begin{aligned}
a_j(s - 1) &=& \frac{\ell - j}{2s}a_{j-1}(s) + \left(1 - \frac{\ell - j}{2s}\right)a_j(s), \\
f_j(s - 1) &=& \frac{\ell - j}{2s}f_{j-1}(s) + \left(1 - \frac{\ell - j}{2s}\right)f_j(s) + \left[f_j(s - 1) - f_j(s) + f_j'(s)\right],
\end{aligned}
$$

for all $j \geq 1$ and all s with $\tau_0 < s \leq \tau$.

The term in square brackets is, by Taylor's Theorem, equal to $\frac{1}{2}f_j''(w)$ for some $w \in (s-1, s)$. We will thus estimate it by bounding the absolute value of the second derivative of f_j.

Lemma 6.1. *For all* $1 \leq j \leq \ell - 2$ *and* $\tau_0 \leq s \leq \tau - 1$,

$$|f_j''(s)| \leq \frac{140}{s(\tau-1)} \frac{(\ell-1)^{5/2}}{j^{3/2}(\ell-j)^{1/2}}. \tag{6.1}$$

This bound also holds if $j = \ell - 1$ *and* $s > (\tau-1)/\ell^2$.
  For $(\tau-1)/\ell^3 \leq s \leq (\tau-1)/\ell^2$, *we have*

$$|f_{\ell-1}''(s)| \leq \ell^{3/2}/s(\tau-1).$$

Proof. Using the expression for $f_j'(s)$ in (5.4), as well as the identities $s = v^2(\tau-1)$ and $(\ell-j)(1-v) - jv = \ell(1-v) - j$, we can write

$$f_j'(s) = \binom{\ell-1}{j} \frac{1}{2(\tau-1)} (1-v)^{j-1} v^{\ell-j-2} (\ell(1-v) - j),$$

and then we have:

$$
\begin{aligned}
f_j''(s) &= \binom{\ell-1}{j} \frac{1}{2(\tau-1)} \frac{v}{2s} (1-v)^{j-2} v^{\ell-j-3} \\
&\quad \times \left[(\ell(1-v) - j) \left\{ (\ell-j-2)(1-v) - (j-1)v \right\} - \ell v(1-v) \right] \\
&= \binom{\ell-1}{j} \frac{1}{4s(\tau-1)} (1-v)^{j-2} v^{\ell-j-2} \\
&\quad \times \left[\{(\ell-1)(1-v) - j\}^2 - (1-v)^2 - vj \right].
\end{aligned}
$$

Let us first verify the result for $j = 1$, when we can write

$$4s(\tau-1) f_1''(s) = (\ell-1) v^{\ell-3} [(\ell-1)(\ell-3) - v\ell(\ell-2)].$$

The right-hand side is increasing from $v = 0$ to $v = (\ell-1)(\ell-3)^2/\ell(\ell-2)^2$, and decreasing thereafter. It is thus always at least its value at $v = 1$, which is $-(\ell-1)(2\ell-3)$, and at most its value at the stationary point, which is at most

$$(\ell-1)1^{\ell-3}(\ell-1)(\ell-3) \left(1 - \frac{\ell-3}{\ell-2} \right) = \frac{(\ell-1)^2(\ell-3)}{(\ell-2)} \leq (\ell-1)^2,$$

and thus

$$|f_1''(s)| \leq \frac{1}{2s(\tau-1)} (\ell-1)^2,$$

which is as required.

We now embark on the calculation for $2 \leq j \leq \ell - 2$. We define a parameter $\varphi = \varphi(v)$ by $v = (\ell - j - 2 - \varphi)/(\ell - 4)$, so $1 - v = (j - 2 + \varphi)/(\ell - 4)$, and $-(j-2) \leq \varphi \leq \ell - j - 2$. The point is that the "main term" $(1-v)^{j-2} v^{\ell-j-2}$ in our expression for the second derivative of f is maximised at $\varphi = 0$, whereas the other term $\left[\{(\ell-1)(1-v) - j\}^2 - (1-v)^2 - vj \right]$ is small for small φ. We write

$$4s(\tau-1) f_j''(s) = k_1 k_2 k_3,$$

where

$$
\begin{aligned}
k_1 &= \binom{\ell-1}{j} \left(\frac{j-2}{\ell-4} \right)^{j-2} \left(\frac{\ell-j-2}{\ell-4} \right)^{\ell-j-2}, \\
k_2 &= \left(1 + \frac{\varphi}{j-2} \right)^{j-2} \left(1 - \frac{\varphi}{\ell-j-2} \right)^{\ell-j-2}, \\
k_3 &= \left((\ell-1) \frac{j-2+\varphi}{\ell-4} - j \right)^2 - \left(\frac{j-2+\varphi}{\ell-4} \right)^2 - j \frac{\ell-j-2-\varphi}{\ell-4}.
\end{aligned}
$$

If $j = 2$ or $j = \ell - 2$, the terms in k_1 and k_2 with a power of $j - 2$ or $\ell - j - 2$ respectively are treated as equal to 1, and therefore absent from the products. We shall estimate k_1, k_2 and k_3 separately to start with, and then consider $k_2 k_3$.

We have

$$
\begin{aligned}
k_1 &= \binom{\ell - 1}{j} \left(\frac{j - 2}{\ell - 4} \right)^{j-2} \left(\frac{\ell - j - 2}{\ell - 4} \right)^{\ell - j - 2} \\
&= \frac{(\ell - 1)(\ell - 2)(\ell - 3)}{j(j - 1)(\ell - j - 1)} \left(\frac{(\ell - 4)!\, e^{\ell - 4}}{(\ell - 4)^{\ell - 4}} \right) \left(\frac{(j - 2)^{j-2}}{(j - 2)!\, e^{j-2}} \right) \\
&\quad \times \left(\frac{(\ell - j - 2)^{\ell - j - 2}}{(\ell - j - 2)!\, e^{\ell - j - 2}} \right).
\end{aligned}
$$

Again, if $j - 2$ or $\ell - j - 2$ is zero, the related term is absent (i.e., the ratio is equal to 1). We now use the inequalities

$$
\sqrt{x + 1} \left(\frac{x}{e} \right)^x \leq x! \leq 3\sqrt{x} \left(\frac{x}{e} \right)^x,
$$

to obtain that

$$
k_1 \leq 3 \frac{(\ell - 1)(\ell - 2)(\ell - 3)}{j(j - 1)(\ell - j - 1)} \sqrt{\frac{\ell - 4}{(j - 1)(\ell - j - 1)}}.
$$

(Note that this remains valid if $j = 2$ or $j = \ell - 2$.)

We next consider k_2. We assume for the moment that $j \leq \ell/2$ (the other case is symmetric) and distinguish two ranges. First, we consider the case where $|\varphi| < j - 2$. In this case, we use the bound $\log(1 + x) \leq x - \frac{x^2}{4}$, valid for all $|x| < 1$, and obtain:

$$
\begin{aligned}
\log k_2 &= (j - 2) \log \left(1 + \frac{\varphi}{j - 2} \right) + (\ell - j - 2) \log \left(1 - \frac{\varphi}{\ell - j - 2} \right) \\
&\leq (j - 2) \left(\frac{\varphi}{j - 2} - \frac{\varphi^2}{4(j - 2)^2} \right) \\
&\quad + (\ell - j - 2) \left(\frac{-\varphi}{\ell - j - 2} - \frac{\varphi^2}{4(\ell - j - 2)^2} \right) \\
&= -\frac{\varphi^2}{4} \left(\frac{1}{j - 2} + \frac{1}{\ell - j - 2} \right) \\
&= -\frac{\varphi^2}{4} \frac{\ell - 4}{(j - 2)(\ell - j - 2)}.
\end{aligned}
$$

In the case where $\varphi = \alpha(j - 2)$ with $\alpha \geq 1$ and $j > 2$, we estimate

$$
\begin{aligned}
\left(1 + \frac{\varphi}{j - 2} \right)^{j-2} \left(1 - \frac{\varphi}{\ell - j - 2} \right)^{\ell - j - 2} &\leq (1 + \alpha)^{\varphi/\alpha} e^{-\varphi} \\
&= \left((1 + \alpha)^{1/\alpha} e^{-1} \right)^{\varphi} \\
&\leq (2/e)^{\varphi}.
\end{aligned}
$$

If $j = 2$, then k_2 is just $(1 - \varphi/(\ell - 4))^{\ell - 4}$, which is at most $e^{-\varphi} \leq (2/e)^{\varphi}$. There is a final case where $\varphi = -(j - 2)$, i.e., $v = 1$, and we may dispose of this immediately since the second derivative is zero unless $j = 2$. In summary, $k_2 \leq \exp \left(-\frac{\varphi^2}{4} \frac{\ell - 4}{(j - 2)(\ell - j - 2)} \right)$ if $|\varphi| < \min(j - 2, \ell - j - 2)$, and $k_2 \leq (2/e)^{|\varphi|}$ otherwise.

Let us organise k_3 as a quadratic in φ:

$$
\begin{aligned}
k_3 \;=\; & \frac{1}{(\ell-4)^2}\Big[\varphi^2\ell(\ell-2)-\varphi(4\ell^2-7\ell j-8\ell+12j) \\
& + \big((2\ell-3j-2)^2-(j-2)^2-j(\ell-4)(\ell-j-2)\big)\Big].
\end{aligned}
$$

Using the inequality

$$
|a\varphi^2+b\varphi+c|\le 2a\varphi^2+b^2/4a+|c|, \tag{6.2}
$$

valid for all positive a, and all real b,c,φ, we have, for all $\ell\ge 8$,

$$
\begin{aligned}
k_3 \;\le\;& \frac{2\ell(\ell-2)}{(\ell-4)^2}\varphi^2+\frac{(4\ell^2-7\ell j-8\ell+12j)^2}{4\ell(\ell-2)(\ell-4)^2} \\
& +\frac{\big|(2\ell-3j-2)^2-(j-2)^2-j(\ell-4)(\ell-j-2)\big|}{(\ell-4)^2} \\
\le\;& 6\varphi^2+\frac{\big(2(\ell-4)(2\ell-3)-(j-2)(7\ell-12)\big)^2}{4\ell(\ell-2)(\ell-4)^2} \\
& +\left(\frac{|2\ell-3j-2|}{\ell-4}\right)^2+\left(\frac{j-2}{\ell-4}\right)^2+\frac{j(\ell-j-2)}{\ell-4} \\
\le\;& 6\varphi^2+4+4+1+\frac{j(\ell-j-2)}{\ell-4} \\
=\;& 6\varphi^2+9+\frac{j(\ell-j-2)}{\ell-4}.
\end{aligned}
$$

Now we combine our bounds to produce a single bound on k_2k_3. The product of k_2 with $9+\dfrac{j(\ell-j-2)}{\ell-4}$ is certainly at most $9+\dfrac{j(\ell-j-2)}{\ell-4}\le 10\dfrac{j(\ell-j-1)}{\ell-4}$, while the product of k_2 with $6\varphi^2$ is at most the maximum of $6\varphi^2\left(\dfrac{2}{e}\right)^{|\varphi|}\le 35\le 35\dfrac{j(\ell-j-1)}{\ell-4}$ and

$$
6\varphi^2\exp\left(-\frac{\varphi^2}{4}\frac{\ell-4}{(j-2)(\ell-j-2)}\right)\le\frac{24}{e}\frac{(j-2)(\ell-j-2)}{\ell-4}<35\frac{j(\ell-j-1)}{\ell-4}.
$$

We can summarise by saying that, provided $\ell\ge 8$, for all values of φ,

$$
k_2k_3\le 45\frac{j(\ell-j-1)}{\ell-4}.
$$

Therefore, overall, we have, for $2\le j\le\ell-2$,

$$
\begin{aligned}
4s(\tau-1)|f_j''(s)| \;\le\;& 135\frac{(\ell-1)(\ell-2)(\ell-3)}{j(j-1)(\ell-j-1)}\sqrt{\frac{\ell-4}{(j-1)(\ell-j-1)}\frac{j(\ell-j-1)}{\ell-4}} \\
=\;& 135\frac{(\ell-1)(\ell-2)(\ell-3)}{\sqrt{\ell-4}}\frac{1}{(j-1)^{3/2}}\frac{1}{\sqrt{\ell-j-1}} \\
\le\;& 560\frac{(\ell-1)^{5/2}}{j^{3/2}(\ell-j)^{1/2}},
\end{aligned}
$$

provided $\ell\ge 8$.

Consider now the special case $j=\ell-1$, when the bound (6.1) translates to $4s(\tau-$

$1)|f_j''(s)| \le 560(\ell - 1)$. Using (6.2), we have

$$
\begin{aligned}
4s(\tau - 1)|f_{\ell-1}''(s)| &= \frac{(1-v)^{\ell-3}}{v}\left|(\ell-1)^2 v^2 - (\ell-1)v - (1-v)^2\right| \\
&\le \frac{(1-v)^{\ell-3}}{v}\left(2(\ell-1)^2 v^2 + \frac{5}{4}\right) \\
&\le 2(\ell-1)^2 v(1-v)^{\ell-3} + \frac{2(1-v)^{\ell-3}}{v}.
\end{aligned}
$$

The first term above is maximised at $v = 1/(\ell - 2)$, so we have

$$
2(\ell-1)^2 v(1-v)^{\ell-3} \le 2\frac{(\ell-1)^2}{\ell-2}\left(1 - \frac{1}{\ell-2}\right)^{\ell-3} \le 2(\ell-1).
$$

If $s > (\tau-1)/\ell^2$, then $v > 1/\ell$, and the second term above is at most $2\ell(1-v)^{\ell-3} \le 3(\ell-1)$, so we do have $4s(\tau - 1)|f_j''(s)| \le 5(\ell - 1)$, as desired.

For the range $(\tau - 1)/\ell^3 < s \le (\tau - 1)/\ell^2$, we have $v > \ell^{-3/2}$. This gives $4s(\tau - 1)|f_{\ell-1}''(s)| \le 2(\ell-1) + 2\ell^{3/2}$, and thence $|f_{\ell-1}''(s)| \le \ell^{3/2}/s(\tau - 1)$, as claimed.

This completes the proof. $\qquad\square$

We are now ready to bound the difference $e_j(s) = a_j(s) - f_j(s)$. Recall that we have, from comparing the recurrences satisfied by the two systems:

$$
e_j(s-1) = \left(1 - \frac{\ell - j}{2s}\right)e_j(s) + \frac{\ell - j}{2s}e_{j-1}(s) - \left[f_j(s-1) - f_j(s) + f_j'(s)\right].
$$

For $0 \le j \le \ell - 1$ and $\tau_0 \le s \le \tau - 1$, set

$$
C_j = 1 + 140\sum_{i=1}^{j}\left(\frac{\ell}{i(\ell - i)}\right)^{3/2}.
$$

We now use induction on j and $\tau - s$ to show that $|e_j(s)| \le C_j\frac{\ell}{\tau-1}$, for all $0 \le j \le \ell - 2$ and $\tau_0 \le s \le \tau - 1$. (We shall return to the case $j = \ell - 1$ afterwards.)

We first check the inequality $|e_0(s)| \le \dfrac{\ell}{\tau - 1}$ for $j = 0$. We have that $a_0(\tau - 1) = 1$, and, for $\tau_0 < s \le \tau - 1$,

$$
a_0(s-1) = \left(1 - \frac{\ell}{2s}\right)a_0(s),
$$

so

$$
a_0(s) = \prod_{w=s+1}^{\tau-1}\left(1 - \frac{\ell}{2w}\right),
$$

while

$$
f_0(s) = \left(\frac{s}{\tau - 1}\right)^{\ell/2} = \prod_{w=s+1}^{\tau-1}\left(1 - \frac{1}{w}\right)^{\ell/2}.
$$

Thus we have

$$
\frac{f_0(s)}{a_0(s)} = \prod_{w=s+1}^{\tau-1}\frac{(1 - 1/w)^{\ell/2}}{(1 - \ell/2w)}.
$$

Each term in the product is clearly at least 1, so $f_0(s) \ge a_0(s)$ for all s. If $s \le \ell$, then we certainly have $|e_0(s)| \le f_0(s) \le s/(\tau - 1) \le \ell/(\tau - 1)$, so we may assume that $s \ge \ell$. Now we have, for all $w \ge s \ge 2\ell$,

$$
\frac{(1 - 1/w)^{\ell/2}}{1 - \ell/2w} \le \frac{1 - \frac{\ell}{2w} + \frac{\ell^2}{8w^2}}{1 - \frac{\ell}{2w}} \le 1 + \frac{1}{1 - 1/2}\frac{\ell^2}{8w^2} = 1 + \frac{\ell^2}{4w^2} \le \exp\left(\frac{\ell^2}{4w^2}\right).
$$

This means that

$$\frac{f_0(s)}{a_0(s)} \le \exp\left(\sum_{w=s+1}^{\tau-1} \frac{\ell^2}{4w^2}\right) \le \exp\left(\frac{\ell^2}{4}\left(\frac{1}{s} - \frac{1}{\tau-1}\right)\right) = \exp\left(\frac{\ell^2(\tau-1-s)}{4s(\tau-1)}\right).$$

Now we write

$$
\begin{aligned}
|e_0(s)| &= f_0(s)(1 - a_0(s)/f_0(s)) \\
&\le \left(\frac{s}{\tau-1}\right)^{\ell/2}\left(1 - \exp(-\ell^2(\tau-1-s)/4s(\tau-1))\right) \\
&\le \left(\frac{s}{\tau-1}\right)^{\ell/2}\frac{\ell^2(\tau-1-s)}{4s(\tau-1)}.
\end{aligned}
$$

This function is maximised at $s = (1 - 2/\ell)(\tau-1)$, and its value there is equal to

$$\frac{(1-2/\ell)^{\ell/2-1}\ell^2(2/\ell)}{4(\tau-1)} \le \frac{\ell}{2(\tau-1)},$$

as required for the case $j = 0$.

For $j > 0$, we have $e_j(\tau-1) = a_j(\tau-1) - f_j(\tau-1) = 0 - 0 = 0$. Now, for the induction step, suppose that $0 < j \le \ell-2$, $\tau_0 < s \le \tau-1$, and that we have verified our inequality for both $|e_{j-1}(s)|$ and $|e_j(s)|$. Hence we have

$$
\begin{aligned}
&|e_j(s-1)| \\
&\le \left(1 - \frac{\ell-j}{2s}\right)|e_j(s)| + \frac{\ell-j}{2s}|e_{j-1}(s)| + |f_j(s-1) - f_j(s) + f_j'(s)| \\
&\le \frac{\ell}{\tau-1}\left[\left(1 - \frac{\ell-j}{2s}\right)C_j + \frac{\ell-j}{2s}C_{j-1}\right] + |f_j(s-1) - f_j(s) + f_j'(s)|.
\end{aligned}
$$

We now note that $f_j(s-1) - f_j(s) + f_j'(s) = f_j''(w)/2$ for some w in $(s-1, s)$, and so its absolute value is at most $\dfrac{70}{s(\tau-1)}\dfrac{\ell^{5/2}}{j^{3/2}(\ell-j)^{1/2}}$, by Lemma 6.1, Therefore

$$
\begin{aligned}
|e_j(s-1)| &\le \frac{\ell}{\tau-1}\left[\left(1 - \frac{\ell-j}{2s}\right)C_j + \frac{\ell-j}{2s}C_{j-1} + \frac{70}{s}\frac{\ell^{3/2}}{j^{3/2}(\ell-j)^{1/2}}\right] \\
&= \frac{\ell}{\tau-1}\left[C_j + \frac{\ell-j}{2s}\left(-C_j + C_{j-1} + 140\frac{\ell^{3/2}}{j^{3/2}(\ell-j)^{3/2}}\right)\right] \\
&= C_j\frac{\ell}{\tau-1},
\end{aligned}
$$

where the last line is by the definition of the C_j's.

For $j = \ell-1$ and $s-1 > (\tau-1)/\ell^2$, the same calculation still gives us that $|e_{\ell-1}(s-1)| \le C_{\ell-1}\ell/(\tau-1)$, since the bound (6.1) is valid for $|f_{\ell-1}''(w)|$ as long as $w > (\tau-1)/\ell^2$. For values of $s-1$ with $(\tau-1)/\ell^3 < s-1 \le (\tau-1)/\ell^2$, we replace the bound on the second derivative by $|f''(w)| \le \dfrac{\ell^{3/2}}{s(\tau-1)}$, and the same calculation gives

$$|e_{\ell-1}(s)| \le C_{\ell-2}\frac{\ell}{\tau-1} + \frac{\ell^{3/2}}{\tau-1}.$$

We now observe that the sum $\sum_{i=1}^{\ell-1} \dfrac{\ell^{3/2}}{i^{3/2}(\ell-i)^{3/2}}$ is uniformly bounded, being at most

$2\sum_{i=1}^{\ell/2}\dfrac{2^{3/2}}{i^{3/2}}\le 15$. Therefore we have

$$|e_j(s)|\le \frac{2200\ell}{\tau-1},$$

for all $0\le j\le \ell-2$ and $\tau_0\le s\le \tau-1$, and also for $j=\ell-1$ and $(\tau-1)/\ell^2<s\le\tau-1$.
For $(\tau-1)/\ell^3<s\le(\tau-1)/\ell^2$, we have

$$|e_{\ell-1}(s)|\le \frac{2200\ell}{\tau-1}+\frac{\ell^{3/2}}{\tau-1}\le \frac{800\ell^{3/2}}{\tau-1},$$

since $\ell\ge 8.0000$

This completes the proof of Lemma 5.1, and hence in turn the proofs of Lemma 4.2 and Theorem 1.1.

7 Concentration for $U_t(\ell)$

In this section, we give a very brief sketch of the proof of Theorem 1.2. The proof proceeds on very similar lines to that of Theorem 1.1.

Recall that $U_t(\ell)$ is the number of vertices of degree *at least* ℓ at time t. It is easy to show that the expected value of $U_t(\ell)$ is close to $u_t(\ell)=2t/\ell(\ell+1)$, uniformly over t and ℓ.

The difference $F_t(\ell)=U_t(\ell)-u_t(\ell)$ satisfies the matrix equation

$$F_t=W_{t-1}F_{t-1}+\Delta M_t.$$

Here ΔM_t is a vector of martingale differences, and

$$W_u=\begin{pmatrix}1-\frac{1}{2u} & 0 & \cdots & 0\\ \frac{2}{2u} & 1-\frac{2}{2u} & \cdots & 0\\ \vdots & \ddots & \ddots & 0\\ 0 & 0 & \frac{\ell_0-1}{2u} & 1-\frac{\ell_0-1}{2u}\end{pmatrix}.$$

The indexing of vectors and matrices runs from $\ell=2$ to $\ell=\ell_0$; we do not need to track the component $\ell=1$ since $U_t(1)=t$ for all t. As for D_t, the plan is to apply Theorem 2.5 to the martingale

$$\widetilde{M_t^\tau}=\sum_{s=\tau_0+1}^{t}C_{s-1}\Delta M_s,$$

where $C_s=\prod_{u=s+1}^{\tau-1}W_u$, for fixed $\tau>\tau_0$.

A transition of U_s involves an increase of 1 in at most one component $U(\ell)$; in other words U_{s+1} is obtained from U_s by adding some unit vector e_ℓ, or leaving the vector unchanged.

For $\ell=2,\dots,\ell_0$, and $\tau_0\le t<\tau$, we set

$$\Phi_t^\ell(Z)=\Phi_t^{g^\ell}(Z)=\sum_{s=\tau_0}^{t-1}\sum_{k=2}^{\ell_0}P_s(U_s,U_s+e_k)C_t(\ell,k)^2.$$

We set, for each $1\le\ell\le\ell_0$,

$$S^\ell=225\frac{\tau-1}{\ell^2}+10^{18}\psi^2\log^{13}(\psi\tau),$$

and

$$T_S^\ell = \inf\{t \geq \tau_0 : \Phi_t^\ell(Z) > S^\ell\}.$$

An application of Theorem 2.5 gives

$$\mathbb{P}\left(\left(\sup_{\tau_0 \leq t \leq \tau} |\widetilde{M}_t^\tau(\ell)| > 3\sqrt{\log(\psi\tau)S^\ell}\right) \wedge (T_S^\ell \geq \tau)\right) \leq \frac{2}{\psi^2\tau^2}$$

for each $\tau \geq \tau_0$ and $2 \leq \ell \leq \ell_0$.

Now let $\widehat{T}$ be the infimum of the times $s \geq \tau_0$ such that either there is a vertex of degree at least $\psi\sqrt{s-1}$, or, for some $k \leq \ell_0$,

$$\left|U_s(k) - \frac{2s}{k(k+1)}\right| > 45\frac{\sqrt{s\log(\psi s)}}{k} + 4 \times 10^9 \psi \log^7(\psi s).$$

The next step is to prove the following result, analagous to Lemma 4.2.

Lemma 7.1. *For all $t \leq \widehat{T}$, and all $\ell = 2, \ldots, \ell_0$, $\Phi_{t-1}^\ell(Z) \leq S^\ell$.*

The result implies that $T_S^\ell \geq \widehat{T}$ for all $\ell = 2, \ldots, \ell_0$.

Now, for each $\tau \geq \tau_0$ and each $\ell \geq 2$, set

$$\delta_\tau(\ell) = 45\frac{\sqrt{\tau\log(\psi\tau)}}{\ell} + 3 \times 10^9 \psi \log^7(\psi\tau).$$

We have

$$\mathbb{P}\left(\left(\sup_{\tau_0 \leq t \leq \tau} |\widetilde{M}_t^\tau(\ell)| > \delta_\tau(\ell)\right) \wedge (T_S^\ell \geq \tau)\right) \leq \frac{2}{\psi^2\tau^2}.$$

In particular, together with Lemma 7.1 and the fact that $\widetilde{M}_\tau^\tau = F_\tau$ for each $\tau \geq \tau_0$, this implies that

$$\mathbb{P}\left((|F_\tau(\ell)| > \delta_\tau(\ell)) \wedge (\widehat{T} \geq \tau)\right) \leq \frac{2}{\psi^2\tau^2}.$$

We next use this inequality to show that, with probability at least $3/\psi$, for all $\tau \geq \tau_0$ and $\ell = 2, \ldots, \ell_0$, either $|F_\tau(\ell)| \leq \delta_\tau(\ell)$ or $\tau \geq \widehat{T}$. Similarly to the proof of Theorem 1.1, this leads to the conclusion that $\mathbb{P}(\widehat{T} < \infty) \leq 4/\psi$, which is the desired result.

8 More complex preferential attachment models

In this section, we discuss some of the issues we confront when extending this proof to other models of preferential attachment.

A first extension would cover the model which again generates a random tree, where now an arriving vertex chooses an existing vertex v as a neighbour with probability proportional to $X(v) + \beta$, where $X(v)$ is the degree of vertex v, and β is a fixed constant. For such a model, the expected degree of a vertex at time t grows as $Ct^{1/(2+\beta)}$, and the expected number of vertices of degree ℓ at time t behaves as $Ct/\ell^{3+\beta}$. When attempting to follow the proof in this paper to establish concentration results, the main difficulty is in finding a suitable analogue of Lemma 5.1, giving bounds on the error function playing the role of $e_j(s)$.

Another well-studied variant is to have each arriving vertex select some fixed number m of neighbours (with replacement), instead of just one. The main difficulty introduced by this variation is that we have to account for the possibility that some existing vertex has its degree increased by more than one at each step, and that the recurrence relations do not have such clean forms.

In the full Cooper-Frieze model (see [6], [5]), the number of new edges added at each step is a random variable. Indeed, with some probability, no new vertex is added,

and some edges are added between existing vertices, chosen either uniformly or via preferential attachment. This means that the numbers of vertices and edges present at time t are no longer determined, causing further complications in the application of our method.

We do believe that all of these problems can be overcome, and that our method can be used to analyse general Cooper-Frieze models. We also hope that the method will find further applications in the analysis of other random processes.

References

[1] Athreya, K. B., Ghosh, A. P. and Sethuraman, S.: Growth of preferential attachment random graphs via continuous-time branching processes. *Proc. Indian Acad. Sci. (Math. Sci.)* **118**, (2008), 473–494. MR-2450248

[2] Barabási, A.-L. and Albert, R.: Emergence of scaling in random networks. *Science*, (1999), 509–512. MR-2091634

[3] Bollobás, B. and Riordan, O.: Mathematical results on scale-free graphs. *Handbook of Graphs and Networks* (S. Bornholdt and H. Schuster, eds.), Wiley-VCH, Berlin, (2002).

[4] Bollobás, B., Riordan, O., Spencer, J. and Tusnády, G.: The degree sequence of a scale-free random graph process. *Rand. Struct. Alg.* **18**, (2001), 279–290. MR-1824277

[5] Cooper, C.: Distribution of vertex degree in web-graphs. *Combin. Probab. Comput.* **15**, (2006), 637–661. MR-2248318

[6] Cooper, C. and Frieze, A.: A general model of web graphs. *Rand. Struct. Alg.* **22**, (2003), 311–335. MR-1966545

[7] Darling, R. W. R. and Norris, J. R.: Differential equation approximations for Markov chains. *Probability Surveys* **5**, (2008), 37–79. MR-2395153

[8] Deijfen, M., van den Esker, H., van der Hofstad, R. and Hooghiemstra, G.: A preferential attachment model with random initial degrees. *Arkiv för Matematik* **47**, (2007), 41–72. MR-2480915

[9] Dereich, S. and Mörters, P.: Random networks with sublinear preferential attachment: Degree evolutions. *Elec. J. Probab.* **14**, (2009), 1222–1267. MR-2511283

[10] Ethier, S. N. and Kurtz, T.: *Markov processes: characterisation and convergence.* Wiley, (1986). MR-0838085

[11] Janson, S.: Asymptotic degree distribution in random recursive trees. *Rand. Struct. Alg.* **26**, (2005), 69–83. MR-2116576

[12] Lu, J. and Feng, Q.: Strong consistency of the number of vertices of given degrees in nonuniform random recursive trees. *Yokohama Math. J.* **45**, (1998), 61–69. MR-1617833

[13] Luczak, M. J.: A quantitative law of large numbers via exponential martingales. In *Stochastic Inequalities and Their Applications.* Progress in Probability **56**, (2003), 93–111, Birkhäuser. MR-2073429

[14] Luczak, M.J. and McDiarmid, C.: On the maximum queue length in the supermarket model. *Ann. Probab.* **34**, (2006), 493–527. MR-2223949

[15] Luczak, M. J. and Norris, J. R.: Strong approximation for the supermarket model. *Ann. Appl. Probab.* **15**, (2005), 2038–2061. MR-2152252

[16] Mahmood, H. M. and Smythe, R. T.: A survey of recursive trees. (Ukrainian); translated from *Teor. Imovir. Mat. Stat.* **51** (1994), 1–29. *Theory Probab. Math. Statist.* **51** (1995), 1–27. MR-1445048

[17] Mahmood, H. M., Smythe, R. T. and Szymański, J.: On the structure of random plane-oriented recursive trees and their branches, *Rand. Struct. Alg.* **4**, (1993), 151–176. MR-1206674

[18] McDiarmid, C.: Concentration. In *Probabilistic Methods for Algorithmic Discrete Mathematics.* (M. Habib, C. McDiarmid, J. Ramirez and B. Reed, eds.) 195–248. Springer – Verlag, Berlin, (1998). MR-1678578

[19] Móri, T. F.: The maximum degree of the Barabási-Albert random tree. *Combin. Probab. Comput* **14**, (2005), 339–348. MR-2138118

[20] Norris, J. R. *Markov chains.* Reprint of 1997 original. Cambridge Series in Statistical and Probabilistic Mathematics, 2. Cambridge University Press, Cambridge, 1998. xvi+237 pp. ISBN: 0-521-48181-3. MR-1600720

[21] Rudas, A., Toth, B. and Valko, B.: Random trees and general branching processes. *Rand. Struct. Alg.* **31**, (2007), 186–202. MR-2343718

[22] Szymański, J.: On a nonuniform random recursive tree. Random graphs '85 (Poznan, 1985), 297–306, North-Holland Math. Stud., 144, North-Holland, Amsterdam. (1987) MR-0930497

[23] Szymański, J.: Concentration of vertex degrees in a scale-free random graph process. *Rand. Struct. Alg.* **26**, (2005), 224–236. MR-2116583

[24] Talagrand, M.: Concentration of measure and isoperimetric inequalities in product spaces. *Publ. Math. Institut des Hautes Études Scientifiques* **81**, (1995), 73–205. MR-1361756

[25] Yule, G.: A mathematical theory of evolution based on the conclusions of Dr. J. C. Willis *Phil. Trans. Roy. Soc. London* **B213**, (1925), 21–87.