J. Appl. Probab. **59,** 299–324 (2022)
*doi:*10.1017/jpr.2021.59

GIANT DESCENDANT TREES, MATCHINGS, AND INDEPENDENT SETS IN AGE-BIASED ATTACHMENT GRAPHS

HÜSEYIN ACAN,* *Drexel University*
ALAN FRIEZE,** *Carnegie Mellon University*
BORIS PITTEL,*** *Ohio State University*

Abstract

We study two models of an age-biased graph process: the δ-version of the preferential attachment graph model (PAM) and the uniform attachment graph model (UAM), with m attachments for each of the incoming vertices. We show that almost surely the scaled size of a breadth-first (descendant) tree rooted at a fixed vertex converges, for $m = 1$, to a limit whose distribution is a mixture of two beta distributions and a single beta distribution respectively, and that for $m > 1$ the limit is 1. We also analyze the likely performance of two greedy (online) algorithms, for a large matching set and a large independent set, and determine – for each model and each greedy algorithm – both a limiting fraction of vertices involved and an almost sure convergence rate.

Keywords: Random; preferential attachment graph; asymptotics

2020 Mathematics Subject Classification: Primary 05C80
Secondary 05C05; 60C05; 05C30

1. Introduction

It is widely accepted that graphs/networks are an inherent feature of life today. The classical models $G_{n,m}$ and $G_{n,p}$ of Erdős and Rényi [17] and Gilbert [22], respectively, lacked some salient features of observed networks. In particular, they failed to have a degree distribution that decays polynomially. Barabási and Albert [3] suggested the preferential attachment model (PAM) as a more realistic model of a 'real-world' network. There was a certain lack of rigor in [3], and later Bollobás, Riordan, Spencer, and Tusnády [11] gave a rigorous definition.

Many properties of this model have been studied. Bollobás and Riordan [9] studied the diameter and proved that with high probability (w.h.p.) PAM with n vertices and $m > 1$ attachments for every incoming vertex has diameter $\approx \log n/ \log\log n$. An earlier result by Pittel [34] implied that for $m = 1$ w.h.p. the diameter of PAM is of exact order $\log n$. Bollobás and Riordan [7, 8] studied the effect on component size of deleting random edges from PAM and showed that it is quite robust w.h.p. The degree distribution was studied by Móri [30, 31], Flaxman, Frieze, and Fenner [18], and Berger, Borgs, Chayes, and Saberi [4]. Peköz, Röllin,

Received 20 November 2020; revision received 23 June 2021.
* Postal address: Department of Mathematics, Drexel University, Philadelphia, PA 19104, USA. Email: huseyin.acan@drexel.edu
** Postal address: Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA. Email: alan@random.math.cmu.edu
*** Postal address: Department of Mathematics, Ohio State University, Columbus, OH 43210, USA. Email: bgp@math.ohio-state.edu

and Ross [32] established convergence, with rate, of the joint distribution of the degrees of finitely many vertices. Acan and Hitczenko [1] found an alternative proof, without rate, via a memory game. Pittel [36] used the Bollobás–Riordan pairing model to approximate, with explicit error estimate, the degree sequence of the first $n^{m/(m+2)}$ vertices, $m \geq 1$, and proved that, for $m > 1$, PAM is connected with probability $\approx 1 - O((\log n)^{-(m-1)/3})$. Random walks on PAM have been considered in the work of Cooper and Frieze [14, 15]. In [14] there are results on the proportion of vertices seen by a random walk on an evolving PAM and [15] determines the asymptotic cover time of a fully evolved PAM. Frieze and Pegden [20] used random walk in a 'local algorithm' to find vertex 1, improving results of Borgs *et al.* [12]. The mixing time of such a walk was analyzed by Mihail, Papadimitriou, and Saberi [28], who showed rapid mixing. Interpolating between Erdős–Rényi and preferential attachment, Pittel [35] considered the birth of a giant component in a graph process G_M on a fixed vertex set, when G_{M+1} is obtained by inserting a new edge between vertices i and j with probability proportional to $[\deg(i) + \delta] \cdot [\deg(j) + \delta]$, with $\delta > 0$ being fixed. Confirming a conjecture of Pittel [35], Janson and Warnke [25] recently determined the asymptotic size of the giant component in the supercritical phase in this graph model.

The paragraph above gives a small sample of results on PAM that can be related to its role as a model of a real-world network. It is safe to say that PAM has now been accepted into the pantheon of random graph models that can be studied from a purely combinatorial perspective. For example, Cooper, Klasing, and Zito [16] studied the size of the smallest dominating set and Frieze, Pérez-Giménez, Prałat, and Reiniger [21] studied the existence of perfect matchings and Hamilton cycles.

One source of our inspiration was the work of Móri [30, 31]; see also Katona and Móri [27] and van der Hofstad [23]. They were able to construct a family of martingales in the form of factorial products, with arguments being the degrees of individual vertices. This allowed them to analyze the limiting behavior of vertex degrees in a δ-version of PAM. In this paper we construct a new factorial-type martingale with one argument being the total size of a 'descendant' subtree. This is a generalization of the martingale for $\delta = 0$, found by Pittel [36].

2. Our results

For each of the two models, PAM and UAM, we study the descendant tree of a given vertex v; it is a *maximal* subtree rooted at v and formed by increasing paths starting at v. The number of vertices in this subtree is a natural influence measure of vertex v. We also analyze the performance of two online greedy algorithms, for finding a large matching set and for a large independent set. Both algorithms are well known, and the classical random graphs are a gold mine for problems on the expected efficiency of these and similar algorithms, thanks to their homogeneity and independence. The PAM and UAM models are inherently more difficult, due to the rather limited scope of these properties. We carry out this analysis in the context of the PAM graph process described in [11] and its extension taken from [23, Chapter 8], and the UAM graph process; see [2] and [21].

The PAM graph process, δ-extension. Vertex 1 has m loops, so its degree is $2m$ initially. Recursively, vertex $t + 1$ has m edges, and it uses them one at a time either to connect to a vertex $x \in [t]$ or to loop back on itself.

To describe the transition probabilities, let $d_{t,i-1}(x)$ denote the degree of vertex x just before the ith edge of vertex $t + 1$ arrives. Let w denote the random receiving end of the ith edge

https://doi.org/10.1017/jpr.2021.59 Published online by Cambridge University Press

emanating from vertex $t+1$. *Conditioned on $G_{m,\delta}(t)$ and the previous $i-1$ edges emanating from vertex $t+1$*, we have

$$\mathbb{P}(w=x)=\begin{cases}\dfrac{d_{t,i-1}(x)+\delta}{(2m+\delta)t+2i-1+i\delta/m} & \text{if } x\in[t],\\[2ex] \dfrac{d_{t,i-1}(t+1)+1+i\delta/m}{(2m+\delta)t+2i-1+i\delta/m} & \text{if } x=t+1.\end{cases} \tag{2.1}$$

Thus one-step transition from $G_{m,\delta}(t)$ to $G_{m,\delta}(t+1)$ is the m-long sequence of choices made by the m edges emanating from vertex $t+1$, and given $G_{m,\delta}(t)$, these choices form a Markov sequence.

For $m=1$, writing $d_t(x)$ for $d_{t,0}(x)$, the probabilities above can be written a little more simply in the form

$$\mathbb{P}(w=x)=\begin{cases}\dfrac{d_t(x)+\delta}{(2+\delta)t+(1+\delta)} & \text{if } x\in[t],\\[2ex] \dfrac{1+\delta}{(2+\delta)t+(1+\delta)} & \text{if } x=t+1.\end{cases} \tag{2.2}$$

Bollobás and Riordan [9] discovered the following coupling between $\{G_{m,0}(t)\}_t$ for $m>1$ and $\{G_{1,0}(mt)\}_t$. Start with the $\{G_{1,0}(t)\}$ random process and let the vertices be $v_1, v_2, \ldots$. To obtain $\{G_{m,0}(t)\}$ from $\{G_{1,0}(mt)\}$:

(1) collapse the first m vertices $v_1, \ldots, v_m$ into the first vertex w_1 of $G_{m,0}(t)$, the next m vertices $v_{m+1}, \ldots, v_{2m}$ into the second vertex w_2 of $G_{m,0}(t)$, and so on;

(2) keep the full record of the multiple edges and loops formed by collapsing the blocks $\{v_{(i-1)m+1}, \ldots, v_{im}\}$ for each i.

By doing this collapsing indefinitely we get the jointly defined Bollobás–Riordan graph processes $\{G_{m,0}(t)\}$ and $\{G_{1,0}(mt)\}$. The beauty of the δ-extended Bollobás–Riordan model is that, similarly, this collapsing operation applied to the process $\{G_{1,\delta/m}(mt)\}$ delivers the process $\{G_{m,\delta}(t)\}$ [23].

(For the reader's convenience we present the explanation in Appendix A.)

Remark 2.1. Note that the process is well-defined for $\delta \geq -m$, since for such δ all the probabilities defined in (2.1) are non-negative and add up to 1. For $m=1$ and $\delta=-1$, it is easy to see from (2.2) that there is no loop in the graph except the loop on the first vertex. Hence a vertex $u>1$ starts with degree 1 and then its degree does not change, since as long as $d_t(u)+\delta=0$ the vertex cannot attract any neighbors (again from (2.2)). As a result, in this case the graph is a star centered at vertex 1. It follows from the above coupling that $G_{m,-m}(t)$ is also a star centered at vertex 1, and the key problems we want to solve have trivial solutions in this extreme case.

The UAM graph process. Conceptually close to the preferential attachment model is the uniform attachment model (UAM). In this model, vertex $t+1$ selects uniformly at random (repetitions allowed) m vertices from the set $[t]$ and attaches itself to these vertices. (See [2] for connectivity and bootstrap percolation results.) This model can be thought of the limit of the PAM model as $\delta \to \infty$ except that loops are not allowed in this case.

2.1. Number of descendants

Fix a positive integer r and let $X(t)$ denote the number of descendants of r at time t. Here r is a descendant of r, and x is a descendant of r if and only if x chooses to attach itself to at least one descendant of r in step x. In other words, if we think of the graph as a directed graph with edges oriented towards the smaller vertices, vertex x is a descendant of r if and only if there is a directed *decreasing* path from x to r. In [36] $X(t)$ was proposed as an influence measure of vertex r at time t.

We prove two theorems.

Theorem 2.1. *Suppose that r is fixed, $m = 1$ and $\delta > -1$, and set $p_X(t) := X(t)/t$. Then almost surely (i.e. with probability 1) $\lim p_X(t)$ exists, and its distribution is the mixture of two beta distributions, with parameters*

$$a = 1,\ b = r - \frac{1}{2+\delta} \quad \textit{and} \quad a = \frac{1+\delta}{2+\delta},\ b = r,$$

weighted by

$$\frac{1+\delta}{(2+\delta)r-1} \quad \textit{and} \quad \frac{(2+\delta)(r-1)}{(2+\delta)r-1}$$

respectively. Consequently almost surely $\lim_{t\to\infty} p(t) > 0$.

Remark 2.2. The two beta distributions in the theorem result from two possible scenarios for vertex r; if there is a loop on r we get the first beta distribution, otherwise we get the second one.

Note.

(i) The proof is based on a new family of martingales

$$M_\ell(t) := \frac{(X(t) + \gamma/(2+\delta))^{(\ell)}}{(t+\beta)^{(\ell)}},$$

where $(z)^{(\ell)}$ stands for the rising factorial. This family definitely resembles the martingales used in [30, 31] for the individual vertices' degrees. For instance, if $D_j(t)$ is the degree of vertex j at time $t \geq j$, then for some deterministic $\gamma_k(t)$, $Z_{j,k}(t) := \gamma_k(t)(D_j(t)+\delta)^{(k)}$ is a martingale [23]. However, $M_\ell(t)$ depends on $X(t)$, a *global* parameter of the PAM graph. Our initial proof was quite technical. We owe a debt of gratitude to an anonymous referee who was able to condense our argument to a few lines.

(ii) With high probability $G_{1,\delta}$ is a forest of $\Theta(\log t)$ trees rooted at vertices with loops. For the preferential attachment tree (no loops), Janson [24] recently proved that the scaled sizes of the *principal* subtrees, those rooted at the root's children and ordered chronologically, converge almost surely (a.s.) to the GEM-distributed random variables. His techniques differ significantly from ours.

For $m > 1$ we use Theorem 2.1 to prove a somewhat surprising result that, for fixed r, almost surely all but a vanishingly small fraction of vertices are descendants of the vertex r (see [24]).

Theorem 2.2. *Let r be fixed, $m > 1$ and $\delta > -m$, and let $p_X(t) = X(t)/t$, $p_Y(t) = Y(t)/(2mt)$, where $Y(t)$ is the total degree of the descendants of r at time t. Then almost surely* $\lim_{t\to\infty} p_X(t) = \lim_{t\to\infty} p_Y(t) = 1$.

https://doi.org/10.1017/jpr.2021.59 Published online by Cambridge University Press

For the case of UAM we have the following result.

Theorem 2.3. *Consider the UAM graph process $G_m(t)$. Given a fixed $r > 1$, let $X(t)$ be the cardinality of the descendant tree rooted at vertex r, and let $p_X(t) := X(t)/t$.*

(i) *For $m = 1$, almost surely* $\lim p_X(t)$ *exists, and it has the beta distribution with parameters 1 and r, which is the distribution of the minimum of r independent* $[0, 1]$*-uniforms. Consequently a.s.* $\liminf_{t\to\infty} p_X(t) > 0$.

(ii) *For $m > 1$, almost surely* $\lim_{t\to\infty} p_X(t) = 1$.

2.2. Greedy matching algorithm

We analyze a greedy matching algorithm; a.s. it delivers a surprisingly large matching set even for relatively small m. This algorithm generates the increasing sequence $\{M(t)\}$ of partial matchings on the sets $[t]$, with $M(1) = \emptyset$. Suppose that $X(t)$ is the set of unmatched vertices in $[t]$ at time t. If $t + 1$ attaches itself to a vertex $u \in X(t)$, then $M(t + 1) = M(t) \cup \{\{u, t + 1\}\}$, otherwise $M(t + 1) = M(t)$. (If $t + 1$ chooses multiple vertices from $X(t)$, then we pick one of those as u arbitrarily.)

First consider the PAM graph. Let

$$h(z) = h_{m,\delta}(z) := 2\left[1 - \left(\frac{m+\delta}{2m+\delta}\right)z\right]^m - z - 1,$$

and let $\rho = \rho_{m,\delta}$ be the unique solution of $h(z) = 0$ in the interval $[0, 1]$. (We have $\rho_{m,\delta} \in (0, 1)$ if $\delta > -m$.)

Theorem 2.4. *Let $M(t)$ and $X(t)$, respectively, be the set of greedy matchings and the set of uncovered vertices at time t, and let $x(t) = X(t)/t$. For any $\delta > -m$ and $\alpha < 1/3$, almost surely*

$$\lim_{t\to\infty} t^\alpha \max\{0, x(t) - \rho_{m,\delta}\} = 0.$$

In consequence, the greedy matching algorithm a.s. finds a sequence of nested matchings $\{M(t)\}$, where the number of vertices in $M(t)$ is asymptotically at least $(1 - \rho_{m,\delta})t$.

Remark 2.3. Observe that $\rho_{m,-m} = 1$, which makes it plausible that the maximum matching size is minuscule compared to t. In fact, by Remark 2.1, $G_{m,-m}(t)$ is the star centered at vertex 1 and hence the maximum matching size is 1.

Remark 2.4. Consider the case $\delta = 0$. Let $r_m := 1 - \rho_{m,0}$; some values of r_m are

$$\begin{aligned} r_1 &= 0.5000, & r_2 &= 0.6458, & r_5 &= 0.8044,\\ r_{10} &= 0.8863, & r_{20} &= 0.9377, & r_{70} &= 0.9803. \end{aligned} \tag{2.3}$$

With a bit of calculus, we obtain $r_m = 1 - 2m^{-1} \log 2 + O(m^{-2})$.

Theorem 2.5. *Let $M(t)$ denote the greedy matching set after t steps of the UAM process. Let r_m denote the unique positive root of $2(1 - z^m) - z = 0$, i.e. $r_m = 1 - m^{-1} \log 2 + O(m^{-2})$. Then, for any $\alpha < 1/3$, almost surely*

$$\lim_{t\to\infty} t^\alpha \left|\frac{2|M(t)|}{t} - r_m\right| = 0.$$

https://doi.org/10.1017/jpr.2021.59 Published online by Cambridge University Press

Some values of r_m in this case are

$$r_1 = 0.6667, \quad r_2 = 0.7808, \quad r_5 = 0.8891,$$
$$r_{10} = 0.9386, \quad r_{20} = 0.9674, \quad r_{35} = 0.9809.$$

2.3. Greedy independent set algorithm

The algorithm generates an increasing sequence of independent sets $\{I(t)\}$ on vertex sets $[t]$. Namely, $I(1) = \{1\}$, and $I(t+1) = I(t) \cup \{t+1\}$ if $t+1$ does not select any of the vertices in $I(t)$; if it does, then $I(t+1) = I(t)$. $I(t)$ is also a dominating set for the PAM/UAM graph with vertex set $[t]$; indeed, if a vertex $\tau \in [t] \setminus I(t)$ did not have any neighbor in $I(t)$, then vertex τ would have been added to $I(\tau - 1)$ at step τ. (Pittel [33] analyzed the performance of this algorithm applied to the Erdős–Rényi random graph with a large but fixed vertex set.)

For the PAM case we prove the following.

Theorem 2.6. *Let w_m denote the unique root of $(1-w)^m - w$ in $(0, 1)$. For any*

$$\chi \in \left(0, \min\left\{\frac{1}{3}, \frac{2m+2\delta}{3(2m+\delta)}\right\}\right),$$

almost surely

$$\lim_{t\to\infty} t^{\chi}\left|\frac{|I(t)|}{t} - w_m\right| = 0.$$

Remark 2.5. Thus the limiting scaled size of the greedy independent set does not depend on δ, but the convergence rate does.

For the UAM case we prove an almost identical result.

Theorem 2.7. *Let w_m be the unique positive root of $-w + (1-w)^m$ in $(0, 1)$. Then, for any $\alpha < 1/3$, almost surely*

$$\lim_{t\to\infty} t^{\alpha}\left|\frac{|I(t)|}{t} - w_m\right| = 0.$$

Remark 2.6. Let w_m be the unique positive root of $(1-w)^m - w$ in $(0, 1)$. As $m \to \infty$,

$$w_m = \frac{\log m}{m} + \mathrm{O}\left(m^{-1}\log\log m\right).$$

So, for m large and both models, a.s. for all large enough t the greedy algorithm delivers an independent set containing a fraction $\sim \log m/m$ of all vertices in $[t]$. It was proved in [21] that for each large t with probability $1 - \mathrm{o}(1)$, the fraction of vertices in the largest independent set in the PAM graph process and in the UAM graph process is at most $(4 + \mathrm{o}(1))\log m/m$ and $(2 + \mathrm{o}(1))\log m/m$, respectively. Since $I(t)$ is dominating, our results prove a.s. existence for all large t of relatively small dominating sets, of cardinality $\le (1+\varepsilon_m)t\log m/m$, $(\varepsilon_m \to 0)$.

Remark 2.7. Let $\mathcal{I}(t)$ denote the cardinality of the largest independent set. We conjecture that for each of the processes there exists a corresponding constant $c(m)$ such that a.s. $\lim_{t\to\infty} t^{-1}\mathcal{I}(t) = c(m)$; of course, $c(m) \ge w_m$.

Notice that, for the UAM process, the existence of a *deterministic* function $c(m,t)$, bounded away from 0 and 1 as $t \to \infty$, such that the distribution of $\mathcal{I}(t)$ is sharply concentrated around $tc(m,t)$, is not difficult to prove. Indeed, $\mathcal{I}(t)$ is determined by $t-1$ *independent* selections of

https://doi.org/10.1017/jpr.2021.59 Published online by Cambridge University Press

m neighbors among the older vertices made by vertices $2, \ldots, t$. Obviously, $\mathcal{I}(t)$ meets the following two conditions: (1) changing the outcome of any such selection can affect $\mathcal{I}(t)$ by at most 1; (2) if $\mathcal{I}(t) \geq s$, then there are s vertices that form an independent set, which depends entirely on selections made by these vertices. Using Talagrand's general inequality [37] (see also [29]), we have, for $z \leq \mathbb{E}[\mathcal{I}(t)]$,

$$\mathbb{P}\Big(|\mathcal{I}(t) - \mathbb{E}[\mathcal{I}(t)]| > z + 3\sqrt{\mathbb{E}[\mathcal{I}(t)]}\Big) \leq 2\exp\left(-\frac{z^2}{16\mathbb{E}[\mathcal{I}(t)]}\right).$$

That does it, since by Theorem 2.7 $\mathbb{E}[\mathcal{I}(t)] \geq \mathbb{E}[I(t)] \geq 0.9 w_m t$ when t is large enough and, with exponentially high probability, we have $\mathcal{I}(t) \leq (\sigma + o(1))t$, where $\sigma = \sigma(m) \in (0, 1)$ is the root of $\sigma \log \sigma + (1-\sigma)\log(1-\sigma) + \sigma^2 m/2$; see Lemma 9 of [21].

3. Descendant trees

Instead of referring the reader back to Section 2, we start this, and other proof sections, by formulating each claim in full.

3.1. Proof of Theorem 2.1

In this subsection we prove Theorem 2.1 restated below.

Theorem 2.1. *Suppose that $m = 1$ and $\delta > -1$, and set $p_X(t) := X(t)/t$. Then almost surely (i.e. with probability 1) $\lim p_X(t)$ exists, and its distribution is the mixture of two beta distributions, with parameters*

$$a = 1,\ b = r - \frac{1}{2+\delta} \quad \textit{and} \quad a = \frac{1+\delta}{2+\delta},\ b = r,$$

weighted by

$$\frac{1+\delta}{(2+\delta)r - 1} \quad \textit{and} \quad \frac{(2+\delta)(r-1)}{(2+\delta)r-1}$$

respectively. Consequently a.s. $\lim_{t\to\infty} p_X(t) > 0$.

Proof. For $t \geq r > 1$, let $X(t) = X_{m,\delta}(t) = X_{m,\delta}(t, r)$ and $Y(t) = Y_{m,\delta}(t) = Y_{m,\delta}(t, r)$ denote the size and total degree of the vertices in the descendant set rooted at r; so $X(r) = 1$ and $Y(r) \in [m, 2m]$, where m (resp. $2m$) is attained when vertex r forms no loops (resp. forms m loops) at itself. Introduce $p_Y(t) = Y(t)/(2mt)$. This notation will be used in the proof of Theorem 2.2 as well, but of course $m = 1$ in Theorem 2.1.

Here

$$Y(t) = \begin{cases} 2X(t) & \text{if } r \text{ looped on itself,} \\ 2X(t) - 1 & \text{if } r \text{ selected a vertex in } [r-1]. \end{cases}$$

$\big($In particular, $p_Y(t) = p_X(t) + O\big(t^{-1}\big)$.$\big)$ So, by (2.2),

$$\begin{aligned} \mathbb{P}(X(t+1) = X(t) + 1 \mid \circ) &= \frac{Y(t) + \delta X(t)}{(2+\delta)t + (1+\delta)} \\ &= \begin{cases} \dfrac{(2+\delta)X(t)}{(2+\delta)t + (1+\delta)} & \text{if } r \text{ looped on itself,} \\ \dfrac{(2+\delta)X(t) - 1}{(2+\delta)t + (1+\delta)} & \text{if } r \text{ selected a vertex in } \ [r-1], \end{cases} \end{aligned}$$

https://doi.org/10.1017/jpr.2021.59 Published online by Cambridge University Press

where '∘' indicates conditioning on prehistory (particularly $X(t)$ and $Y(t)$). Thus we are led to consider the process $X(t)$ such that

$$\mathbb{P}(X(t+1)=X(t)+1 \mid \circ)=\frac{(2+\delta)X(t)+\gamma}{(2+\delta)t+(1+\delta)},$$
$$\mathbb{P}(X(t+1)=X(t) \mid \circ)=1-\mathbb{P}(X(t+1)=X(t)+1 \mid \circ),$$

$\gamma=0$ if r looped on itself, $\gamma=-1$ if r selected a vertex in $[r-1]$. *So* γ *is determined by* $G_{1,\delta}(t)$. Letting $\beta=(1+\delta)/(2+\delta)$, the above equation can be written as

$$\begin{aligned}\mathbb{P}(X(t+1)=X(t)+1 \mid \circ)&=\frac{X(t)+\gamma/(2+\delta)}{t+\beta},\\ \mathbb{P}(X(t+1)=X(t) \mid \circ)&=1-\frac{X(t)+\gamma/(2+\delta)}{t+\beta}.\end{aligned} \tag{3.1}$$

For $\delta=0$ the following claim was proved in [36]. We let $z^{(\ell)}$ denote the rising factorial $\prod_{j=0}^{\ell-1}(z+j)$.

Lemma 3.1. *Let* $\beta=(1+\delta)/(2+\delta)$ *and* $Z(t)=X(t)+\gamma/(2+\delta)$. *Then, conditioned on the attachment record during the time interval* $[r, t]$, *i.e. starting with the attachment decision by vertex r, we have*

$$\mathbb{E}[Z(t+1)^{(\ell)} \mid \circ]=\left(\frac{t+\beta+\ell}{t+\beta}\right)Z(t)^{(\ell)}.$$

Consequently

$$M_\ell(t):=\frac{Z(t)^{(\ell)}}{(t+\beta)^{(\ell)}}$$

is a martingale.

We thank an anonymous referee for the following proof, which is much shorter than our original proof.

Proof. We have

$$\begin{aligned}\mathbb{E}[(Z(t+1)^{(\ell)} \mid \circ]&=(Z(t)+1)^{(\ell)}\frac{Z(t)}{t+\beta}+Z(t)^{(\ell)}\left(1-\frac{Z(t)}{t+\beta}\right)\\ &=(Z(t)+\ell)\,Z(t)^{(\ell)}\frac{1}{t+\beta}+Z(t)^{(\ell)}-Z(t)^{(\ell)}\frac{Z(t)}{t+\beta}\\ &=Z(t)^{(\ell)}\left(\frac{\ell+t+\beta}{t+\beta}\right).\end{aligned}$$

The second part follows from this immediately. □

To identify the $\lim_{t\to\infty}p(t)$, recall that the classical beta probability distribution has density

$$f(x;a, b)=\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}(1-x)^{b-1}, \quad x\in(0, 1),$$

parametrized by two parameters $a>0$, $b>0$, and moments

$$\int_0^1 x^\ell f(x;a, b)\,\mathrm{d}x=\prod_{j=0}^{\ell-1}\frac{a+j}{a+b+j}.$$

https://doi.org/10.1017/jpr.2021.59 Published online by Cambridge University Press

We can now complete the proof of Theorem 2.1. By Lemma 3.1, we have, for all $t \geq r$, $\mathbb{E}[M_\ell(t) \mid \gamma] = M_\ell(r)$, or explicitly

$$\mathbb{E}\left[\frac{(X(t)+\gamma/(2+\delta))^{(\ell)}}{(t+\beta)^{(\ell)}} \,\Big|\, \gamma\right] = \frac{(1+\gamma/(2+\delta))^{(\ell)}}{(r+\beta)^{(\ell)}}.$$

Since $|\gamma/(2+\delta)| \leq 1$ and $X(t) \leq t$, we obtain

$$|M_\ell(t)| \leq \left[\max\left(\frac{t+1}{t+\beta},\ \frac{t+\ell}{t+\beta+\ell-1}\right)\right]^\ell \Longrightarrow \sup_{t\geq 0} |M_\ell(t)| \leq \max(1,\ \beta^{-\ell}).$$

Therefore, for every $\ell \geq 1$, by the martingale convergence theorem, conditioned on γ, a.s. there exists

$$\lim_{t\to\infty} \frac{(X(t)+\gamma/(2+\delta))^{(\ell)}}{(t+\beta)^{(\ell)}} =: \mathcal{M}_{\gamma,\ell} \leq \max(1,\ \beta^{-\ell}), \quad \mathbb{E}[\mathcal{M}_{\gamma,\ell}] = \frac{(1+\gamma/(2+\delta))^{(\ell)}}{(r+\beta)^{(\ell)}}. \tag{3.2}$$

Since $X(t) \leq t$, it follows from (3.2) that a.s. there exists $\lim_{t\to\infty} X(t)/t = \lim_{t\to\infty} p_X(t)$, and

$$\lim_{t\to\infty} \mathbb{E}[p_X(t)^\ell] = \mathbb{E}[\mathcal{M}_{\gamma,\ell}] = \frac{(1+\gamma/(2+\delta))^{(\ell)}}{(r+\beta)^{(\ell)}} = \prod_{j=0}^{\ell-1} \frac{1+\gamma/(2+\delta)+j}{r+\beta+j}.$$

The sequence of the right-hand side products is the sequence of moments of a unique distribution, which is the beta distribution with parameters $1+\gamma/(2+\delta)$ and $r+\beta-1-\gamma/(2+\delta)$. By the definition of γ and (2.2), we have

$$\mathbb{P}(\gamma = 0) = \frac{1+\delta}{(2+\delta)(r-1)+(1+\delta)} = \frac{1+\delta}{2r-1+\delta r}.$$

We conclude that $\lim_{t\to\infty} p_X(t)$ has the distribution which is the mixture of the two beta distributions, with parameters

$$a = 1,\ b = r - \frac{1}{2+\delta} \quad \text{and} \quad a = \frac{1+\delta}{2+\delta},\ b = r,$$

weighted by

$$\frac{1+\delta}{(2+\delta)r-1} \quad \text{and} \quad \frac{(2+\delta)(r-1)}{(2+\delta)r-1}$$

respectively. This completes the proof of Theorem 2.1. □

3.2. Proof of Theorem 2.2

In this subsection we prove Theorem 2.2 restated below.

Theorem 2.2. *Let $m > 1$ and $\delta > -m$, and let $p_X(t) = X(t)/t$, $p_Y(t) = Y(t)/(2mt)$, where $Y(t)$ is the total degree of the descendants of r at time t. Then almost surely $\lim_{t\to\infty} p_X(t) = \lim_{t\to\infty} p_Y(t) = 1$.*

For the proof of this theorem, we need tractable formulas/bounds for the conditional distribution of $Y(t+1) - Y(t)$. The existence of loops in a preferential attachment graph is a minor

https://doi.org/10.1017/jpr.2021.59 Published online by Cambridge University Press

nuisance for stating exact conditional probabilities. Hence we will start with the scenario that we do not get a loop on vertex $t+1$, in which case we can write the exact equations easily and we will then argue that the effect of loops is negligible. Let $\mathcal{E}_m(t)$ be the event that vertex $t+1$ has no loop. Conditioned on $G_{m,\delta}(t)$ and $\mathcal{E}_m(t)$, by the transition probabilities given in (2.1), vertex $t+1$ chooses a vertex $x \in [t]$ with probability

$$\frac{d_{t,i-1}(x)+\delta}{(2m+\delta)t+i-1}.$$

In the following lemma we use '$\mid \circ$)' to denote conditioning on $G_{m,\delta}(t)$ and $\mathcal{E}_m(t)$.

Lemma 3.2. *For* $a \in [m]$,

$$\mathbb{P}(Y(t+1)=Y(t)+m+a \mid \circ)=\binom{m}{a}\frac{(Y(t)+\delta X(t))^{(a)}\cdot(2mt-Y(t)+\delta(t-X(t))^{(m-a)}}{((2m+\delta)t)^{(m)}}$$

and

$$\mathbb{P}(Y(t+1)=Y(t)\mid \circ)=\frac{(2mt-Y(t)+\delta(t-X(t))^{(m)}}{((2m+\delta)t)^{(m)}}.$$

Proof. Let $V(t)$ denote the descendants of r at time t. Vertex $t+1$ selects, in m steps, a sequence $\{v_1, \ldots, v_m\}$ of m vertices from $[t]$, with t choices for every selection. Introduce $\mathbf{I}=(\mathbb{I}_1, \ldots, \mathbb{I}_m)$, where $\mathbb{I}_i$ is the indicator of the event $\{v_i \in V(t)\}$. The total vertex degree of $[t]$ (resp. $V(t)$) right before step i is $2mt+i-1$ (resp. $Y(t)+\mu_i$, $\mu_i := |\{j<i\colon \mathbb{I}_j=1\}|$). Conditioned on $G_{m,\delta}(t)$, $\mathcal{E}_m(t)$, and the outcomes of the previous $i-1$ steps, we have

$$\mathbb{P}(\mathbb{I}_i=1)=\frac{Y(t)+\delta X(t)+\mu_i}{2mt+\delta t+i-1},$$

$$\mathbb{P}(\mathbb{I}_i=0)=\frac{2mt-Y(t)+\delta(t-X(t))+i-1-\mu_i}{2mt+\delta t+i-1}.$$

Therefore a sequence $(j_1, \ldots, j_m)$ will be the outcome of the loopless m-step selection with probability

$$\begin{aligned}\mathbb{P}(\mathbf{I}=(j_1, \ldots, j_m))\\ &=\prod_{i\in[m]}((2m+\delta)t+i-1)^{-1}\\ &\quad\times\prod_{i\colon j_i=1}(Y(t)+\delta X(t)+\mu_i)\times\prod_{i\colon j_i=0}(2mt-Y(t)+\delta(t-X(t))+i-1-\mu_i).\end{aligned}$$

For $j_1+\cdots+j_m=a\in[m]$, the second product above equals $(Y(t)+\delta X(t))^{(a)}$ and the third product equals $(2mt-Y(t)+\delta(t-X(t))^{(m-a)}$ so that

$$\mathbb{P}(\mathbf{I}=(j_1, \ldots, j_m))=\frac{(Y(t)+\delta X(t))^{(a)}(2mt-Y(t)+\delta(t-X(t))^{(m-a)}}{((2m+\delta)t)^{(m)}}.$$

Since the total number of admissible sequences $(j_1, \ldots, j_m)$ with $j_1+\cdots+j_m=a$ is $\binom{m}{a}$, we obtain the first formula in Lemma 3.2. The second formula is the case of $\mathbb{P}(\mathbf{I}=(0, \ldots, 0))$. □

Proof of Theorem 2.2. Lemma 3.2 gives the conditional probabilities for $Y(t+1)-Y(t)$ in the event that there is no loop on vertex $t+1$. Now let us evaluate the probability that there is

https://doi.org/10.1017/jpr.2021.59 Published online by Cambridge University Press

no loop on vertex $t+1$. For $0 \le i \le m$, let $\mathcal{E}_i(t)$ denote the event that no loop has been formed on vertex $t+1$ in the first i steps when vertex $t+1$ is choosing its neighbors. On the event $\mathcal{E}_{i-1}(t)$, as the ith edge incident to $t+1$ is about to attach its second end to a vertex in $[t] \cup \{t+1\}$, the total degree of vertices in $[t]$ is $2mt+i-1$ $(1 \le i \le m)$. So, by the transition probabilities given in (2.1),

$$\mathbb{P}(\mathcal{E}_i(t) \mid \mathcal{E}_{i-1}(t)) = \frac{2mt+i-1+t\delta}{2mt+2(i-1)+t\delta+1+i\delta/m}.$$

Hence

$$\mathbb{P}(\mathcal{E}_m(t)) = \prod_{i=1}^{m} \frac{2mt+i-1+t\delta}{2mt+2(i-1)+t\delta+1+i\delta/m} = \prod_{i=1}^{m} (1-\mathrm{O}(1/t)) = 1-\mathrm{O}\big(t^{-1}\big). \tag{3.3}$$

It is clear from the proof of Lemma 3.2 that, given $X(t)$ and $Y(t)$,

$$\mathbb{P}_m(a) := \binom{m}{a} \frac{(Y(t)+\delta X(t))^{(a)}(2mt-Y(t)+\delta(t-X(t))^{(m-a)}}{((2m+\delta)t)^{(m)}}$$

is a probability distribution of a random variable D, a 'rising-factorial' counterpart of the binomial distribution $\mathcal{D} = \mathrm{Bin}(m, p = Y(t)/2mt)$. (We have not seen this distribution in the literature.) Define the falling factorial $(x)_\ell = x(x-1)\cdots(x-\ell+1)$. It is well known that $\mathbb{E}[(\mathcal{D})_\mu] = (m)_\mu p^\mu$, $(\mu \le m)$. For D we have

$$\begin{aligned}
\mathbb{E}[(D)_\mu] &= \sum_a (a)_\mu \mathbb{P}_m(a) \\
&= \frac{(m)_\mu\, (Y(t)+\delta X(t))^{(\mu)}}{((2m+\delta)t)^{(\mu)}} \cdot \sum_{a \ge \mu} \binom{m-\mu}{a-\mu} \\
&\quad \times \frac{(Y(t)+\delta X(t)+\mu)^{(a-\mu)}((2m+\delta)t+\mu-(Y(t)+\delta X(t)+\mu))^{((m-\mu)-(a-\mu))}}{(2mt+\mu)^{(m-\mu)}} \\
&= \frac{(m)_\mu\, (Y(t)+\delta X(t))^{(\mu)}}{((2m+\delta)t)^{(\mu)}},
\end{aligned} \tag{3.4}$$

since the sum over $a \ge \mu$ is $\sum_{\nu \ge 0} \mathbb{P}_{m-\mu}(\nu) = 1$.

By Lemma 3.2, (3.3), and (3.4),

$$\begin{aligned}
\mathbb{E}[Y(t+1)-Y(t) \mid \circ] &= \sum_{a=1}^{m} (a+m)\mathbb{P}_m(a) + \mathrm{O}\big(t^{-1}\big) \\
&= \left(\sum_{a=1}^{m} a\mathbb{P}_m(a) + \sum_{a=1}^{m} m\mathbb{P}_m(a)\right) + \mathrm{O}\big(t^{-1}\big) \\
&= \frac{m(Y(t)+\delta X(t))}{(2m+\delta)t} + m\left(1 - \frac{((2m+\delta)t-Y(t)-\delta X(t))^{(m)}}{((2m+\delta)t)^{(m)}}\right) + \mathrm{O}\big(t^{-1}\big)
\end{aligned} \tag{3.5}$$

https://doi.org/10.1017/jpr.2021.59 Published online by Cambridge University Press

and

$$\mathbb{E}[X(t+1) - X(t) \mid \circ] = 1 - \frac{((2m+\delta)t - Y(t) - \delta X(t))^{(m)}}{((2m+\delta)t)^{(m)}} + \mathrm{O}\big(t^{-1}\big),$$

where '$\circ$' indicates conditioning on $G_{m,\delta}(t)$. (Note that $X(t+1) = X(t)$ if and only if $Y(t+1) = Y(t)$.)

To continue the proof of Theorem 2.2, we note first that $mX(t) \le Y(t) \le 2mX(t)$. These inequalities follow from the fact that every time a vertex joins the descendant set, the total degree of the descendant set increases by an amount $m + a$ for some $a \in [m]$. Let

$$p(t) := \frac{2m}{2m+\delta}\, p_Y(t) + \frac{\delta}{2m+\delta}\, p_X(t),$$

where $p_Y(t) = Y(t)/(2mt)$ and $p_X(t) = X(t)/t$ as defined before. By the definition of $p(t)$ and the above inequalities, we have

$$\frac{p_X(t)}{2} \le p_Y(t) \le p_X(t) \Longrightarrow \frac{m+\delta}{2m+\delta}\, p_X(t) \le p(t) \le p_X(t);$$

in particular, $p(t) \in [0, 1]$ since $\delta \ge -m$. We will also need

$$\frac{((2m+\delta)t - Y(t) - \delta X(t))^{(m)}}{((2m+\delta)t)^{(m)}} = (1 - p(t))^m + \mathrm{O}\big(t^{-1}\big).$$

So, using (3.5), we compute

$$\begin{aligned} \mathbb{E}[p_Y(t+1) \mid \circ] &= \mathbb{E}\left[\frac{Y(t+1)}{2mt} \cdot \frac{t}{t+1} \,\Big|\, \circ\right] \\ &= \frac{t}{t+1}\left(p_Y(t) + \frac{1}{2t}\big[1 + p(t) - (1-p(t))^m\big] + \mathrm{O}\big(t^{-2}\big)\right) \\ &= p_Y(t) + q_Y(t), \\ q_Y(t) &:= \frac{1}{2(t+1)}\big[1 + p(t) - 2p_Y(t) - (1-p(t))^m\big] + \mathrm{O}\big(t^{-2}\big). \end{aligned} \tag{3.6}$$

Likewise

$$\begin{aligned} \mathbb{E}[p_X(t+1) \mid \circ] &= p_X(t) + q_X(t), \\ q_X(t) &= \frac{1}{t+1}\big[1 - p_X(t) - (1-p(t))^m\big] + \mathrm{O}\big(t^{-2}\big). \end{aligned} \tag{3.7}$$

Multiplying (3.6) by $2m/(2m+\delta)$ and (3.7) by $\delta/(2m+\delta)$, and adding them, we obtain

$$\begin{aligned} \mathbb{E}[p(t+1) \mid \circ] &= p(t) + q(t), \\ q(t) &:= \frac{m+\delta}{(2m+\delta)(t+1)}\big[1 - p(t) - (1-p(t))^m\big] + \mathrm{O}\big(t^{-2}\big). \end{aligned} \tag{3.8}$$

https://doi.org/10.1017/jpr.2021.59 Published online by Cambridge University Press

From the first line in (3.8) it follows that

$$\sum_{t=1}^{\tau} \mathbb{E}[p(t+1)] = \sum_{t=1}^{\tau} \mathbb{E}[p(t)] + \sum_{t=1}^{\tau} \mathbb{E}[q(t)],$$

implying that

$$\limsup_{\tau\to\infty} \sum_{t\le\tau} \mathbb{E}[q(t)] \le \limsup_{\tau\to\infty} \mathbb{E}[p(\tau+1)] \le 1.$$

Since $1 - z - (1-z)^m \ge 0$ on $[0, 1]$, the second line in (3.8) implies that $|q(t)| \le q(t) + \mathrm{O}\big(t^{-2}\big)$. Since $\sum_t t^{-2} < \infty$, we see that $\sum_t \mathbb{E}[|q(t)|] < \infty$.

So a.s. there exists $Q := \lim_{\tau\to\infty} \sum_{1\le t\le\tau} q(t)$, with $\mathbb{E}[|Q|] \le \sum_t \mathbb{E}[|q(t)|] < \infty$, that is, a.s. $|Q| < \infty$. Introducing $Q(t+1) = \sum_{\tau\le t} q(\tau)$, we see from (3.8) that $\{p(t+1) - Q(t+1)\}_{t\ge1}$ is a martingale with $\sup_t |p(t+1) - Q(t+1)| \le 1 + \sum_{\tau\ge1} |q(\tau)|$. By the martingale convergence theorem we obtain that there exists an integrable $\lim_{t\to\infty} (p(t) - Q(t))$, implying that a.s. there exists a random $p(\infty) = \lim_{t\to\infty} p(t)$. Equation (3.8) also implies that

$$1 \ge \mathbb{E}[p(\infty)] = \frac{m+\delta}{2m+\delta} \sum_{t\ge1} \frac{1}{t+1} \mathbb{E}\big[1 - p(t) - (1-p(t))^m\big] + \mathrm{O}(1).$$

Since $m + \delta > 0$ and

$$\lim_{t\to\infty} \mathbb{E}\big[1 - p(t) - (1-p(t))^m\big] = \mathbb{E}\big[1 - p(\infty) - (1-p(\infty))^m\big],$$

and the series $\sum_{t\ge1} t^{-1}$ diverges, we obtain that $\mathbb{P}(p(\infty) \in \{0, 1\}) = 1$.

Recall that

$$p(t) \ge \frac{m+\delta}{2m+\delta}\, p_X(t).$$

If we show that a.s. $\liminf_{t\to\infty} p_X(t) > 0$, it will follow that a.s. $p(\infty) > 0$, whence a.s. $p(\infty) = 1$, implying (by $p(t) \le p_X(t)$) that a.s. $p_X(\infty)$ exists, and is 1, and consequently (by the formula for $p(t)$) a.s. $p_Y(\infty)$ exists, and is 1.

So let us prove that a.s. $\liminf_{t\to\infty} p_X(t) > 0$. Recall that we did prove the latter for $m = 1$. To transfer this earlier result to $m > 1$ we need to establish some kind of monotonicity with respect to m. The coupling described in Section 2 comes to the rescue!

Lemma 3.3. *For the coupled processes* $\{G_{m,\delta}(t)\}$ *and* $\{G_{1,\delta/m}(mt)\}$, *we have* $X_{m,\delta}(t, r) \ge m^{-1} X_{1,\delta/m}(mt, mr)$.

Proof. Let us simply write G_1 and G_m for the two graphs $G_{1,\delta/m}(mt)$ and $G_{m,\delta}(t)$, respectively. Similarly, write T_1 and T_m, respectively, for the descendant tree in $G_{1,\delta/m}(mt)$ rooted at mr and the descendant tree in $G_{m,\delta}(t)$ rooted at r. If $v_a \in T_1$, i.e. v_a is a descendant of mr, then for $b = \lceil a/m \rceil$ we have $w_b = \{v_{m(b-1)+i}\}_{i\in[m]} \ni v_a$, implying that w_b is a descendant of r in G_m, i.e. $w_b \in T_m$. (The converse is generally false: if w_b is a descendant of r, it does not mean that every $v_{m(b-1)+i}$, $(i \in [m])$, is a descendant of mr.) Therefore

$$X_{m,\delta}(t, r) = |V(T_m)| \ge m^{-1} |V(T_1)| = m^{-1} X_{1,\delta/m}(mt, mr). \qquad \square$$

Thus, to complete the proof of the theorem, i.e. for $\delta > -m$, we (a) use Theorem 2.1, to assert that for the process $\{G_{1,\delta/m}(t)\}$, a.s. $\lim_{t\to\infty} p_X(t) > 0$, (b) use Lemma 3.3, to assert that a.s. $\liminf_{t\to\infty} p_X(t) > 0$ for $\{G_{m,\delta}(t)\}$ as well. The proof of Theorem 2.2 is complete. $\square$

https://doi.org/10.1017/jpr.2021.59 Published online by Cambridge University Press

3.3. Proof of Theorem 2.3

Theorem 2.3. *Consider the UAM graph process $G_m(t)$. Given $r > 1$, let $X(t)$ be the cardinality of the descendant tree rooted at vertex r, and let $p_X(t) := X(t)/t$.*

(i) *For $m = 1$, almost surely* $\lim p_X(t)$ *exists and it has the same distribution as the minimum of* $(r-1)$ *independent* $[0, 1]$*-uniforms. Consequently a.s.* $\liminf_{t\to\infty} p_X(t) > 0$.

(ii) *For $m > 1$, almost surely* $\lim_{t\to\infty} p_X(t) = 1$.

Proof. By the definition of the UAM process, we have

$$\mathbb{P}(X(t+1) = X(t) + 1 \mid \circ) = 1 - (1 - p_X(t))^m. \tag{3.9}$$

Case (i). Consider $m = 1$. For $r = 1$, we have $p(t) \equiv 1$. Consider $r \geq 2$. Equation (3.9) is the case $\delta = -1$, $\gamma = 0$ of (3.1). By Lemma 3.1, we claim that

$$M(t) := \frac{(X(t))^{(\ell)}}{t^{(\ell)}}$$

is a martingale. So arguing as in the proof of Theorem 2.1, we obtain that almost surely (a.s.) $\lim p_X(t) = p_X(\infty)$ exists, and the limiting distribution of $p_X(\infty)$ is a beta distribution with parameters $a = 1$ and $b = r$. That is, the limiting density is $r(1-x)^{r-1}$, $x \in [0, 1]$. Therefore a.s. $\lim p_X(t) > 0$.

Case (ii). Consider $m > 1$. Clearly $G_1(t) \subset G_m(t)$. Therefore a.s. $\liminf p_X(t) > 0$ as well. Furthermore, it follows from (3.9) that

$$\mathbb{E}[p_X(t+1) \mid \circ] = p_X(t) + \frac{1}{t+1}\left[1 - p_X(t) - (1 - p_X(t))^m\right]$$

(as before, '$\mid \circ$)' means conditioning on $G_m(t)$), which is a special case of (3.8), with $O\big(t^{-2}\big)$ dropped. So we obtain that $\mathbb{P}(p(\infty) \in \{0, 1\}) = 1$, which in combination with $\mathbb{P}(p_X(\infty) > 0) = 1$ implies that $\mathbb{P}(p_X(\infty) = 1) = 1$. □

4. A technical lemma

In this section we will prove Lemma 4.1. We need the following Chernoff bound for its proof (see e.g. [26, Theorem 2.8]).

Theorem 4.1. *If $X_1, \ldots, X_n$ are independent Bernoulli random variables, $X = \sum_{i=1}^n X_i$, and $\lambda = \mathbb{E}[X]$, then*

$$\mathbb{P}(|X - \lambda| > \varepsilon\lambda) < 2\exp\left(-\varepsilon^2\lambda/3\right) \quad \textit{for all } \varepsilon \in (0, 3/2).$$

Lemma 4.1. *Let $\{X(t)\}_{t\geq 0}$ be a sequence of random variables such that $X(0) = 0$ and $X(t+1) - X(t) \in \{0, 1\}$. Let $x(t) = X(t)/t$ and, using '$\mid \circ$)' to denote conditioning on $\{x(s) \colon s \leq t\}$, let us assume*

$$\mathbb{E}[x(t+1) - x(t) \mid \circ] \leq \frac{h(x(t))}{t} + O\big(t^{-2}\big), \tag{4.1}$$

where h is a continuous, strictly decreasing function with $h(0) > 0$ and $h(1) < 0$, so that $h(x)$ has a unique root $\rho \in (0, 1)$. Assume also that $h'(x) < -1$ in $(0, 1)$. Then, for any $\gamma < 1/3$, almost surely

$$\lim_{t\to\infty} t^\gamma \max\{0, x(t) - \rho\} = 0.$$

https://doi.org/10.1017/jpr.2021.59 Published online by Cambridge University Press

Lemma 4.2. (Extensions of Lemma 4.1.) *Lemma 4.1 can be extended in a couple of ways as follows.*

(a) *If the hypothesis $X(t+1) - X(t) \in \{0, 1\}$ in Lemma 4.1 is replaced with $X(t+1) - X(t) \in \{-1, 1\}$, then the conclusion of Lemma 4.1 still holds. This follows from minor modifications in the proof of Lemma 4.1.*

(b) *If the inequality sign in (4.1) is replaced with an equality sign, i.e. under the condition*

$$\mathbb{E}[x(t+1) - x(t) \mid \circ] = \frac{h(x(t))}{t} + \mathrm{O}\big(t^{-2}\big),$$

we have the following conclusion: for any $\gamma < 1/3$, almost surely

$$\lim_{t\to\infty} t^{\gamma}(x(t) - \rho) = 0.$$

Proof of Lemma 4.2(b). First of all, by Lemma 4.1, we have $\lim_{t\to\infty} t^{\gamma} \max\{0, x(t) - \rho\} = 0$ almost surely. Second, let $g(z) = -h(1-z)$, so that $g(0) > 0$ and $g(1) < 0$, and in $(0, 1)$, we have $g'(z) = h'(1-z) < -1$. Letting $y(t) = 1 - x(t)$,

$$\begin{aligned}\mathbb{E}[y(t+1) - y(t) \mid \circ] &= \mathbb{E}[(1 - x(t+1)) - (1 - x(t)) \mid \circ] \\ &= -\frac{h(x(t))}{t} + \mathrm{O}\big(t^{-2}\big) \\ &= \frac{g(y(t))}{t} + \mathrm{O}\big(t^{-2}\big).\end{aligned}$$

Applying Lemma 4.1 with $X_1(t) = t - X(t)$, and then switching back to $X(t)$, we see that $\lim_{t\to\infty} t^{\gamma} \max\{0, \rho - x(t)\} = 0$ almost surely, as well. □

Proof of Lemma 4.1. Let $\varepsilon = \varepsilon_t := t^{-1/3} \log t$. We will show

$$\mathbb{P}(x(t) > \rho + \varepsilon) \le \exp\left(-\Theta\left(\log^3 t\right)\right). \tag{4.2}$$

Once we show (4.2), the Borel–Cantelli lemma gives

$$\mathbb{P}(x(t) - \rho > t^{-1/3} \log t \text{ infinitely often}) = 0,$$

which proves what we want. Let us prove (4.2).

For $T \in [0, t)$, let $\mathcal{E}_T$ be the event that $\{x(t) > \rho + \varepsilon$ and T is the last time such that $X(\tau) \le (\rho + \varepsilon/2)\tau\}$, that is,

$$X(T) \le (\rho + \varepsilon/2)T, \quad x(\tau) > \rho + \varepsilon/2\tau \text{ for all } \tau \in (T, t), \quad x(t) > \rho + \varepsilon.$$

Since $X(t+1) - X(t) \in \{0, 1\}$, we have

$$X(T) + t - T \ge X(t) > t(\rho + \varepsilon).$$

Using $X(t) = tx(t)$ above, we get

$$T(\rho + \varepsilon/2) + t - T > t(\rho + \varepsilon),$$

implying

$$t - T > \frac{t\varepsilon}{2(1-\rho)}. \tag{4.3}$$

https://doi.org/10.1017/jpr.2021.59 Published online by Cambridge University Press

We conclude that

$$\{x(t) > \rho + \varepsilon\} \subseteq \bigcup_{T=1}^{s} \mathcal{E}_T, \quad s = s(t) := t - \left\lceil \frac{t\varepsilon}{2(1-\rho)} \right\rceil.$$

Now let us fix a $T \in [0, s]$ and bound $\mathbb{P}(\mathcal{E}_T)$. The main idea of the proof is that, as long as $x(\tau) > \rho$, by (4.1), the process $\{x(\tau)\}$ has a negative drift.

Let ξ_τ denote the indicator of the event $\{x(\tau-1) > \rho + \varepsilon/2$ and $X(\tau) = X(\tau-1)+1\}$ and let $\mathcal{Z}_T := \xi_{T+2} + \cdots + \xi_t$. On the event $\mathcal{E}_T$, the sum $\mathcal{Z}_T$ counts the total number of upward unit jumps ($X(\tau) - X(\tau-1) = 1$, $\tau \in [T+2, t]$) and therefore

$$X(T+1) + \mathcal{Z}_T = X(t) \geq t(\rho + \varepsilon).$$

Since $X(T+1) \leq X(T) + 1 \leq T(\rho + \varepsilon/2) + 1$, we must have

$$Z_T > (\rho + \varepsilon)(t - T), \quad Z_T := 1 + \mathcal{Z}_T.$$

Writing $p(x(\tau)) := \mathbb{P}(X(\tau+1) = X(\tau) + 1 \mid \circ)$,

$$\begin{aligned}
\mathbb{E}[x(\tau+1) \mid \circ] &= p(x(\tau))\frac{X(\tau)+1}{\tau+1} + (1 - p(x(\tau)))\frac{X(\tau)}{\tau+1} \\
&= \frac{p(x(\tau))}{\tau+1} + \frac{\tau x(\tau)}{\tau+1} \\
&= x(\tau) + \frac{p(x(\tau)) - x(\tau)}{\tau+1},
\end{aligned}$$

so that $p(x(\tau)) = h(x(\tau)) + x(\tau) + \mathrm{O}\left(\tau^{-1}\right)$.

Recall that for $\tau \geq T+1$ we have $x(\tau) > \rho + \varepsilon/2$. Since $h'(x) < -1$ in (0, 1), the sum $h(x) + x$ is decreasing in (0, 1). Hence, conditioning on the full record (up to and including time τ),

$$\begin{aligned}
\mathbb{P}(\xi_{\tau+1} = 1 \mid \circ) &= \mathbb{P}(X(\tau+1) = X(\tau) + 1 \mid \circ) \\
&= h(x(\tau)) + x(\tau) + \mathrm{O}\left(\tau^{-1}\right) \\
&< h(\rho + \varepsilon/2) + \rho + \varepsilon/2 + \mathrm{O}\left(\tau^{-1}\right) \\
&= h(\rho) + (\varepsilon/2) \cdot h'(y) + \rho + \varepsilon/2 + \mathrm{O}\left(\tau^{-1}\right) \quad \text{for some } y \in (\rho, \rho + \varepsilon/2) \\
&< \rho + \mathrm{O}\left(\tau^{-1}\right).
\end{aligned}$$

Hence the sequence $\{\xi_\tau\}$ is stochastically dominated by the sequence of *independent* Bernoulli random variables B_τ with parameters $\min\left(\rho + \mathrm{O}\left(\tau^{-1}\right), 1\right)$. Consequently Z_T is stochastically dominated by $1 + \sum_{j=T+2}^{t} B_j$, and

$$\lambda := \sum_{j=T+2}^{t} \mathbb{E}[B_j] = \rho(t - T) + \mathrm{O}(\log t).$$

For the choice of ε we have, (4.3) gives

$$(\rho + \varepsilon)(t - T) \geq (1 + \varepsilon/2)\lambda.$$

https://doi.org/10.1017/jpr.2021.59 Published online by Cambridge University Press

Thus, by the Chernoff bound in Theorem 4.1 and using (4.3),

$$\begin{aligned}\mathbb{P}(\mathcal{E}_T) &\le \mathbb{P}(Z_T > (t-T)(\rho+\varepsilon))\\ &\le \mathbb{P}(1+B_{T+2}+\cdots+B_t > (t-T)(\rho+\varepsilon))\\ &\le \mathbb{P}(1+B_{T+2}+\cdots+B_t > (1+\varepsilon/2)\lambda)\\ &\le \exp\left(-\Theta(\varepsilon^2(t-T))\right) \le \mathrm{e}^{-\Theta\left(\log^3 t\right)}.\end{aligned}$$

Using the union bound on T, we complete the proof of (4.2) and of the lemma. □

Note. In the next sections we turn to the two greedy algorithms for the PAM and UAM Markov graph processes. We will continue using the symbol '$\mid \circ$)' to denote conditioning on a current graph $\mathrm{G}(t)$.

5. Greedy matching algorithm

Recall that the greedy matching algorithm (for either of two graph models) generates the increasing sequence $\{M(t)\}$ of partial matchings on the sets $[t]$, with $M(1)=\emptyset$. Given $M(t)$, let

$$\begin{aligned}X(t) &:= \text{ number of unmatched vertices at time } t,\\ Y(t) &:= \text{ total degree of unmatched vertices at time } t,\\ U(t) &:= \text{ number of unmatched vertices selected by } t+1 \text{ from } [t]\setminus M(t),\\ x(t) &:= X(t)/t,\\ y(t) &:= Y(t)/(2mt).\end{aligned}$$

5.1. The PAM case

Theorem 2.4. *Let $X(t)$ be the number of unmatched vertices at time t in the greedy matching algorithm. For $\delta > -m$, let $\rho_{m,\delta}$ be the unique root in $(0, 1)$ of*

$$h(z) = h_{m,\delta}(z) := 2\left[1 - \left(\frac{m+\delta}{2m+\delta}\right)z\right]^m - z - 1. \tag{5.1}$$

Then, for any $\alpha < 1/3$, almost surely

$$\lim_{t\to\infty} t^{\alpha} \max\{0, x(t) - \rho_{m,\delta}\} = 0.$$

In consequence, the greedy matching algorithm a.s. finds a sequence of nested matchings $\{M(t)\}$, where the number of vertices in $M(t)$ is asymptotically at least $(1-\rho_{m,\delta})t$.

Proof. Notice first that for $\delta > -m$ the function $h(z)$ is decreasing on $(0, 1)$ and $h(z)=0$ does have a unique solution in the same interval.

We will prove our claim first for a slightly different model that does not allow any loops other than at the first vertex. In this model, vertex 1 has m loops, and the ith edge of vertex $t+1$ attaches to $u\in[t]$ with probability

$$\frac{d_{t,i-1}(u)+\delta}{2mt+2(i-1)+t\delta}.$$

https://doi.org/10.1017/jpr.2021.59 Published online by Cambridge University Press

Loops not allowed except at vertex 1. In this case, since each degree is at least m, we have $Y(t) \geq mX(t)$ and hence $y(t) \geq x(t)/2$. Also, since

$$X(t+1) = \begin{cases} X(t)+1 & \text{if } U(t)=0, \\ X(t)-1 & \text{if } U(t)>0, \end{cases}$$

we have

$$\mathbb{E}[X(t+1) \mid \circ] = X(t) + \mathbb{P}(U(t)=0 \mid \circ) - \mathbb{P}(U(t)>0 \mid \circ). \tag{5.2}$$

Since $\mathbb{P}(\text{vertex } t+1 \text{ has some loop}) = \mathrm{O}\big(t^{-1}\big)$, using $Y(t) \geq mX(t)$ in the last step below, by Lemma 3.2 we get

$$\begin{aligned}
\mathbb{P}(U(t)=0 \mid \circ) &= \mathbb{P}(U(t)=0 \text{ and vertex } t+1 \text{ has no loop} \mid \circ) + \mathrm{O}\big(t^{-1}\big) \\
&= \big(1-\mathrm{O}\big(t^{-1}\big)\big)\frac{(2mt - Y(t) + \delta t - \delta X(t))^{(m)}}{(2mt+\delta t)^{(m)}} + \mathrm{O}\big(t^{-1}\big) \\
&= \frac{(2mt + \delta t - Y(t) - \delta X(t))^{m}}{(2mt+\delta t)^{m}} + \mathrm{O}\big(t^{-1}\big) \\
&= \left(1 - \frac{2m}{2m+\delta}\, y(t) - \frac{\delta}{2m+\delta}\, x(t)\right)^m + \mathrm{O}\big(t^{-1}\big) \\
&\leq \left(1 - \frac{m+\delta}{2m+\delta}\, x(t)\right)^m + \mathrm{O}\big(t^{-1}\big).
\end{aligned} \tag{5.3}$$

Using (5.2) and (5.3) gives

$$\begin{aligned}
\mathbb{E}[x(t+1) \mid \circ] &\leq x(t) + \frac{1}{t}\left[2\left(1 - \frac{m+\delta}{2m+\delta}\, x(t)\right)^m - x(t) - 1\right] + \mathrm{O}\big(t^{-2}\big) \\
&= x(t) + \frac{1}{t} h(x(t)) + \mathrm{O}\big(t^{-2}\big),
\end{aligned} \tag{5.4}$$

where $h(z)$ is as defined in (5.1). Note that $X(0)=0$ and $h(z)$ satisfies the conditions given in Lemmas 4.1 and 4.2, namely $h(0)>0$, $h(1)<0$, and $h'(z) < -1$ for $z \in (0, 1)$. The conclusion of the theorem follows from the first part of Lemma 4.2 in this case.

Loops allowed everywhere. The above analysis is carried over to this more complicated case via an argument similar to that for the descendant trees in Section 2.1. Here is a proof sketch. First, the counterpart of (5.3) is

$$\begin{aligned}
&\mathbb{P}(\{U(t)=0\} \cap \{\text{no loops at } t+1\} \mid \circ) \\
&\quad = \Pi_m(t) \prod_{j=0}^{m-1} \left(\frac{2mt - Y(t) + \delta t - \delta X(t) + j}{2mt + 2j + 1 + \delta t + (j+1)\,\delta/m}\right) \\
&\quad \leq \Pi_m(t)\left[\left(1 - \frac{m+\delta}{2m+\delta}\, x(t)\right)^m + \mathrm{O}\big(t^{-1}\big)\right] \\
&\quad = \big(1-\mathrm{O}\big(t^{-1}\big)\big)\left[\left(1 - \frac{m+\delta}{2m+\delta}\, x(t)\right)^m + \mathrm{O}\big(t^{-1}\big)\right] \\
&\quad = \left(1 - \frac{m+\delta}{2m+\delta}\, x(t)\right)^m + \mathrm{O}\big(t^{-1}\big);
\end{aligned}$$

see (3.3) for $\Pi_m(t)$. Therefore we again obtain (5.4). The rest of the proof remains the same. □

https://doi.org/10.1017/jpr.2021.59 Published online by Cambridge University Press

Remark 5.1. Let $r = r_{m,\delta} := 1 - \rho_{m,\delta}$, where $\rho_{m,\delta}$ is the unique root in (0, 1) of

$$h(z) = h_{m,\delta}(z) := 2\left[1 - \left(\frac{m+\delta}{2m+\delta}\right)z\right]^m - z - 1.$$

Then r is the unique root in (0, 1) of

$$f(z) = f_{m,\delta}(z) := 2 - z - 2\left(\frac{m}{2m+\delta} + \frac{m+\delta}{2m+\delta}\,z\right)^m.$$

Thus, by Theorem 2.4, we have

$$\liminf\,(1 - x(t)) \geq r$$

almost surely, where $1 - x(t)$ is the fraction of the vertices in $M(t)$. See (2.3) for various r values when $\delta = 0$.

Remark 5.2. When $\delta \to \infty$, the function $f_{m,\delta}(z)$ as defined above converges to $2 - z - 2z^m$ in (0, 1). So it is plausible that for the case of uniform attachment model, the number of vertices in $M(t)$ is asymptotically rt, where r is the unique root of $2 - z - 2z^m$. This is in fact the case, as shown in the next theorem.

5.2. The UAM case

Theorem 2.5. *Let $M(t)$ denote the greedy matching set after t steps of the UAM process. Let r_m denote the unique positive root of $2(1 - z^m) - z = 0$, i.e. $r_m = 1 - m^{-1}\log 2 + \mathrm{O}\big(m^{-2}\big)$. Then, for any $\alpha < 1/3$, almost surely*

$$\lim_{t\to\infty} t^\alpha \left|\frac{2|M(t)|}{t} - r_m\right| = 0.$$

Proof. Let $X(t) = t - 2|M(t)|$ as before. In particular, we have $X(0) = 0$ and $X(1) = 1$. At each step $t \geq 2$ we check the edges incident to vertex t. If some of the edges end at vertices that do not belong to $M(t)$, then we choose the largest (youngest) of those vertices, say w, and set $M(t) := M(t-1) \cup \{(t, w)\}$ and $X(t) = X(t-1) - 1$. Otherwise $M(t) = M(t-1)$ and $X(t) = X(t-1) + 1$. Let $x(t) = X(t)/t$ be the fraction of unmatched vertices after step t. Then $X(t)$ is a Markov chain with

$$\mathbb{P}(X(t+1) - X(t) = 1 \mid X(t)) = (1 - x(t))^m,$$

since for $X(t+1) - X(t) = 1$ to happen, each of the m choices made by vertex $t+1$ must lie outside of $M(t)$, the probability of each such choice is $1 - x(t)$, and the choices are independent of each other. With the remaining probability vertex $t+1$ chooses at least one of m vertices from $X(t)$, in which case X(t) decreases by 1. Consequently

$$\mathbb{E}[X(t+1) \mid \circ] = X(t) + (1 - x(t))^m - \big(1 - (1 - x(t))^m\big) = X(t) + 2(1 - x(t))^m - 1.$$

Dividing both sides by $t + 1$, we obtain

$$\mathbb{E}[x(t+1) \mid \circ] = x(t) + \frac{h(x(t))}{t+1} - \mathrm{O}\Big(1/t^2\Big),$$

where $h(z) = 2(1 - z)^m - z - 1$. The function h meets the conditions of the second part of Lemma 4.2. Hence $\lim_{t\to\infty} t^\gamma\,|x(t) - \rho_m| = 0$ a.s. On the other hand, if ρ_m is the unique root of

https://doi.org/10.1017/jpr.2021.59 Published online by Cambridge University Press

h in (0, 1), then $r_m := 1 - \rho_m$ is the unique root of $2(1 - z^m) - z$ in (0, 1). Since $x(t) - \rho_m = (1 - 2|M(t)|/t) - (1 - \rho_m) = \rho_m - 2|M(t)|/t$, we also have a.s.

$$\lim_{t\to\infty} t^{\gamma} \left|2|M(t)|/t - r_m\right| = 0.$$

This completes the proof. □

6. Analysis of greedy independent set algorithm

The algorithm, for both the PAM and UAM cases, generates the increasing sequence of independent sets $\{I(t)\}$ on the sets $[t]$, with $I(1) := \{1\}$. If vertex $t + 1$ does not select a single vertex from t, we set $I(t + 1) = I(t) \cup \{t + 1\}$; otherwise $I(t + 1) := I(t)$. Given $I(t)$, let

$$\begin{aligned} X(t) &:= \text{number of vertices} \le t \text{ outside of the current independent set } I(t),\\ Y(t) &:= \text{total degree of these outsiders},\\ Z(t) &:= \text{number of insiders selected by outsiders by time } t,\\ U(t) &:= \text{number of insiders selected by vertex } t+1,\\ x(t) &:= \frac{X(t)}{t}, \quad y(t) := \frac{Y(t)}{2mt}, \quad z(t) = \frac{Z(t)}{mt}, \quad i(t) = \frac{|I(t)|}{t}. \end{aligned}$$

Since each insider selects only among outsiders, the total degree of insiders is $m|I(t)| + Z(t)$ and

$$Y(t) = 2mt - m|I(t)| - Z(t) \Longrightarrow y(t) = 1 - (i(t) + z(t))/2.$$

6.1. The PAM case

Theorem 2.6. *Let w_m denote the unique root of $-w + (1 - w)^m$ in* (0, 1). *For any*

$$\chi \in \left(0, \min\left\{\frac{1}{3}, \frac{2m + 2\delta}{3(2m + \delta)}\right\}\right),$$

almost surely

$$\lim_{t\to\infty} t^{\chi} \left|\frac{|I(t)|}{t} - w_m\right| = 0.$$

Proof. By the definition of the algorithm, we have

$$|I(t + 1)| = \begin{cases} |I(t)| + 1 & \text{with probability } \mathbb{P}(U(t) = 0 \mid \circ),\\ |I(t)| & \text{with probability } \mathbb{P}(U(t) > 0 \mid \circ). \end{cases}$$

Now $U(t) = 0$ means that vertex $t + 1$ selects all m vertices from the outsiders set, so that

$$\begin{aligned} \mathbb{P}(U(t) = 0 \mid \circ) &= \frac{(Y(t) + \delta X(t))^{(m)}}{((2m + \delta)t)^{(m)}} + O\big(t^{-1}\big)\\ &= \left(\frac{2my(t) + \delta x(t)}{2m + \delta}\right)^m + O\big(t^{-1}\big)\\ &= \left(1 - \frac{m + \delta}{2m + \delta}\, i(t) - \frac{m}{2m + \delta}\, z(t)\right)^m + O\big(t^{-1}\big). \end{aligned}$$

The leading term in the first equality is the exact (conditional) probability of $\{U(t)=0\}$ when no loops at vertices other than the first vertex are allowed, and the extra $\mathrm{O}\big(t^{-1}\big)$ is for our more general case when the loops are admissible. So, using $i(t)=|I(t)|/t$, we have

$$\mathbb{E}[i(t+1)\mid\circ]=i(t)+\frac{1}{t+1}\left[-i(t)+\left(1-\frac{m+\delta}{2m+\delta}i(t)-\frac{m}{2m+\delta}z(t)\right)^{m}\right]+\mathrm{O}\big(t^{-2}\big). \quad (6.1)$$

Now $Z(t+1)=Z(t)+U(t)$, so that

$$\begin{aligned}\mathbb{E}[Z(t+1)\mid\circ]&=Z(t)+\mathbb{E}[U(t)\mid\circ]\\&=Z(t)+m\frac{I(t)(m+\delta)+Z(t)}{(2m+\delta)t}+\mathrm{O}\big(t^{-1}\big),\end{aligned}$$

and using $z(t)=Z(t)/mt$, we have

$$\mathbb{E}[z(t+1)\mid\circ]=z(t)+\frac{1}{t+1}\left[-z(t)+\left(i(t)\frac{m+\delta}{2m+\delta}+z(t)\frac{m}{2m+\delta}\right)\right]+\mathrm{O}\big(t^{-2}\big). \quad (6.2)$$

Introduce

$$w(t)=i(t)\frac{m+\delta}{2m+\delta}+z(t)\frac{m}{2m+\delta}.$$

Multiplying (6.1) and (6.2) by $(m+\delta)/(2m+\delta)$ and $m/(2m+\delta)$ and adding the products, we obtain

$$\begin{aligned}\mathbb{E}[w(t+1)\mid\circ]&=w(t)+\frac{f(w(t))}{t+1}+\mathrm{O}\big(t^{-2}\big),\\ f(w):=&\ \frac{m+\delta}{2m+\delta}\big[-w+(1-w)^{m}\big].\end{aligned} \quad (6.3)$$

The function f is qualitatively similar to the function h in the proof of Theorem 2.4. Indeed, $f(w)$ is strictly decreasing, with $f(0)=1$ and $f(1)=-1$. Therefore $f(w)$ has a unique root $w_m\in(0,\,1)$; it is not difficult to see that

$$w_m=\frac{\log m}{m}[1+\mathrm{O}(\log\log m/\log m)],\qquad m\to\infty.$$

Let us prove that for any

$$\alpha<c(m,\,\delta):=\min\left\{1,\,\frac{2m+2\delta}{2m+\delta}\right\}\quad(\subseteq(0,\,1])$$

and $A\geq A(\alpha)$ we have

$$\mathbb{E}\Big[(w(t)-w_m)^2\Big]\leq At^{-\alpha}, \quad (6.4)$$

First

$$\begin{aligned}|i(t+1)-i(t)|&=\left|\frac{|I(t+1)|}{t+1}-\frac{|I(t)|}{t}\right|\\&\leq\frac{1}{t+1}(|I(t+1)|-|I(t)|)+\frac{|I(t+1)|}{t+1}\\&\leq\frac{2}{t+1},\end{aligned} \quad (6.5)$$

https://doi.org/10.1017/jpr.2021.59 Published online by Cambridge University Press

and similarly

$$|z(t+1)-z(t)| \le \frac{2}{t+1}.$$

Therefore $|w(t+1)-w(t)| \le 2/(t+1)$, and consequently

$$(w(t+1)-w_m)^2 \le (w(t)-w_m)^2 + \frac{4}{t^2} + 2(w(t)-w_m)(w(t+1)-w(t)).$$

So, conditioning on prehistory, we have

$$\begin{aligned}\mathbb{E}\big[(w(t+1)-w_m)^2 \mid \circ\big] &\le (w(t)-w_m)^2 + 2(w(t)-w_m)\mathbb{E}[w(t+1)-w(t) \mid \circ] + \mathrm{O}\big(t^{-2}\big)\\ &= (w(t)-w_m)^2 + \frac{2(w(t)-w_m)}{t+1} f(w(t)) + \mathrm{O}\big(t^{-2}\big).\end{aligned}$$

By (6.3),

$$f'(w) \le -\frac{m+\delta}{2m+\delta}.$$

Therefore, with

$$c(m,\delta) = \frac{2m+2\delta}{2m+\delta},$$

the last inequality gives

$$\mathbb{E}\big[(w(t+1)-w_m)^2 \mid \circ\big] \le \bigg(1-\frac{c(m,\delta)}{t+1}\bigg)(w(t)-w_m)^2 + \mathrm{O}\big(t^{-2}\big),$$

leading to a recursive inequality

$$\mathbb{E}\big[(w(t+1)-w_m)^2\big] \le \bigg(1-\frac{c(m,\delta)}{t+1}\bigg)\mathbb{E}\big[(w(t)-w_m)^2\big] + \mathrm{O}\big(t^{-2}\big). \tag{6.6}$$

The bound (6.4) follows from (6.6) by a straightforward induction on t. Next we use (6.4) to prove that, for A_1 large enough,

$$\mathbb{E}\big[(i(t+1)-w_m)^2 \mid \circ\big] \le A_1 t^{-\alpha}. \tag{6.7}$$

Using (6.5), we have

$$(i(t+1)-w_m)^2 \le (i(t)-w_m)^2 + \frac{4}{t^2} + 2(i(t)-w_m)(i(t+1)-i(t)).$$

Consequently

$$\begin{aligned}\mathbb{E}\big[(i(t+1)-w_m)^2 \mid \circ\big] &\le (i(t)-w_m)^2 + 2(i(t)-w_m)\mathbb{E}[i(t+1)-i(t) \mid \circ] + \mathrm{O}\big(t^{-2}\big)\\ &= (i(t)-w_m)^2 + \frac{2(i(t)-w_m)}{t+1}\big[-i(t) + (1-w(t))^m\big] + \mathrm{O}\big(t^{-2}\big).\end{aligned}$$

Taking expectations of both sides, and using Cauchy's inequality and (6.4), we obtain

$$\begin{aligned}\mathbb{E}\big[(i(t+1)-w_m)^2\big] \le{}& \bigg(1-\frac{2}{t+1}\bigg)\mathbb{E}\big[(i(t)-w_m)^2\big]\\ &+ \frac{2}{t+1}\mathbb{E}^{1/2}\big[(i(t)-w_m)^2\big]\,\mathbb{E}^{1/2}\big[(w_m-(1-w(t))^m)^2\big] + \mathrm{O}\big(t^{-2}\big)\\ \le{}& \bigg(1-\frac{2}{t+1}\bigg)\mathbb{E}\big[(i(t)-w_m)^2\big] + \frac{2mA^{1/2}}{t^{1+\alpha/2}}\mathbb{E}^{1/2}\big[(i(t)-w_m)^2\big] + \mathrm{O}\big(t^{-2}\big),\end{aligned}$$

https://doi.org/10.1017/jpr.2021.59 Published online by Cambridge University Press

since $w_m = (1 - w_m)^m$, and

$$|(1 - w(t))^m - (1 - w_m)^m| \le m|w(t) - w_m|.$$

So it suffices to show the existence of A_1 such that

$$\left(1 - \frac{2}{t+1}\right)A_1 t^{-\alpha} + \frac{2mA^{1/2}A_1^{1/2}t^{-\alpha/2}}{t^{1+\alpha/2}} + \mathrm{O}\big(t^{-2}\big) \le A_1(t+1)^{-\alpha}$$

holds for $t > t_0$, where t_0 depends only on A and α. For large t, the above inequality becomes

$$2mA^{1/2}A_1^{1/2} + \mathrm{O}(t^{-1-\alpha}) \le A_1[(2-\alpha) + \mathrm{O}\big(t^{-1}\big)] + \mathrm{O}(t^{-1+\alpha}),$$

and

$$A_1 > A\left(\frac{2m}{2-\alpha}\right)^2$$

does the job. So (6.7) is proved. Therefore, by Markov's inequality,

$$\mathbb{P}(|i(t) - w_m| \ge t^{-\chi}) \le At^{-\alpha+2\chi} \to 0, \qquad \chi \in (0, \alpha/2). \tag{6.8}$$

This inequality already means that $i(t) \to w_m$ *in probability*. Let us show the considerably stronger statement that $i(t) \to w_m$ with probability 1, at least as fast as $t^{-\chi}$, for any given $\chi < 1/3$. Pick $\beta > 1$ and introduce a sequence $\{t_\nu\}$, $t_\nu = \lfloor \nu^\beta \rfloor$. By (6.8), we have

$$\sum_{\nu\ge1} \mathbb{P}(|i(t_\nu) - w_m| \ge t_\nu^{-\chi}) \le \sum_{\nu\ge1} A_1 t_\nu^{-\alpha+2\chi} = \mathrm{O}\left(\sum_{\nu\ge1} \nu^{-\beta(\alpha-2\chi)}\right) < \infty,$$

provided that $\beta > (\alpha - 2\chi)^{-1}$, which we assume from now. For such choice of β, by the Borel–Cantelli lemma with probability 1 for all but finitely many ν, we have $|i(t_\nu) - \rho| \le t_\nu^{-\chi}$. Let $t \in [t_\nu, t_{\nu+1}]$. By (6.5), we have

$$|i(t) - i(t_\nu))| = \mathrm{O}\left(\frac{t_{\nu+1} - t_\nu}{t_\nu}\right)$$

uniformly for all ν. So if $|i(t_\nu) - \rho| \le t_\nu^{-\chi}$, then (using $t_\nu = \Theta\big(\nu^\beta\big)$) we have, for $t \in [t_\nu, t_{\nu+1}]$,

$$\begin{aligned}
|i(t) - w_m| &\le t_\nu^{-\chi} + \mathrm{O}\left(\frac{t_{\nu+1} - t_\nu}{t_\nu}\right) \\
&= \mathrm{O}\left(\nu^{-\beta\chi} + \nu^{-1}\right) \\
&= \mathrm{O}\left(\nu^{-\min(\beta\chi, 1)}\right) \\
&= \mathrm{O}\left(t^{-\min(\chi, \beta^{-1})}\right).
\end{aligned}$$

Since $|i(t_\nu) - w_m| \le t_\nu^{-\chi}$ holds almost surely (a.s.) for all but finitely many ν, we see then that a.s. so does the bound $|i(t) - w_m| = \mathrm{O}(t^{-\min(\chi, \beta^{-1})})$ for all but finitely many t. Now, by taking β sufficiently close to $(\alpha - 2\chi)^{-1}$ from above, we can make $\min(\chi, \beta^{-1})$ arbitrarily close to $\min(\chi, \alpha - 2\chi)$ from below. It remains to notice that $\min(\chi, \alpha - 2\chi)$ attains its maximum $\alpha/3$ at $\chi = \alpha/3$. The proof of Theorem 2.6 is complete. □

https://doi.org/10.1017/jpr.2021.59 Published online by Cambridge University Press

6.2. The UAM case

Theorem 2.7. *Let w_m be the unique root of $-w + (1 - w)^m$ in (0, 1). Then, for any $\alpha < 1/3$, almost surely*

$$\lim_{t\to\infty} t^{\alpha}\left|\frac{|I(t)|}{t} - w_m\right| = 0.$$

The following remark is already stated as Remark 2.3.

Remark 6.1. Thus, the convergence rate aside, almost surely the greedy independent algorithm delivers a sequence of independent sets of asymptotically the same size as for the PAM case.

Proof. Let $i(t) = |I(t)|/t$. From the definition of the UAM process and the greedy independent set algorithm, we obtain

$$|I(t+1)| = \begin{cases} |I(t)| + 1 & \text{with conditional probability } (1 - i(t))^m, \\ |I(t)| & \text{with conditional probability } 1 - (1 - i(t))^m. \end{cases}$$

Therefore

$$\mathbb{E}[|I(t+1)| \mid \circ] = |I(t)| + (1 - i(t))^m,$$

or equivalently

$$\mathbb{E}[i(t+1) \mid \circ] = i(t) + \frac{1}{t+1}[-i(t) + (1 - i(t))^m].$$

The function $-x + (1 - x)^m$ differs by a constant positive factor from the function f in (6.3). The function f meets the conditions of Lemma 4.1, and r_m is a unique root of f. Therefore, for any $\alpha < 1/3$, a.s. $\lim_{t\to\infty} t^{\alpha}|i(t) - r_m| = 0$. □

Appendix A. Coupling $\{G_{m,0}(t)\}_t$ and $\{G_{1,0}(mt)\}_t$

In order to show that the coupling described in Section 2 really works, we can compute the probability that the ith edge of vertex w_{t+1} connects to vertex w_x in the coupling and compare it with the probability in (2.1). Let $\{G'_{m,\delta}(t)\}$ denote the process obtained by collapsing the vertices of $\{G_{1,\delta/m}(mt)\}$. Note that the $(mt + i)$th edge of the $\{G_{1,\delta/m}(mt)\}$ process becomes the ith edge of w_{t+1} after collapsing. Hence the ith edge of vertex w_{t+1} connects to w_x $(x \le t)$ if and only if the $(mt + i)$th edge of the $\{G_{1,\delta/m}(mt)\}$ process connects v_{mt+i} with one of the vertices $v_{m(x-1)+1}, \ldots, v_{mx}$. Let $d_{mt+i-1}(v_y)$ denote the degree of v_y $(y \le mt + i)$ just before the $(mt + i)$th edge of the $\{G_{1,\delta/m}\}$ process is drawn. Also, let $D_{t,i-1}(w_x)$ denote the degree of w_x at the exact same time. Hence, by definition,

$$D_{t,i-1}(w_x) = \begin{cases} \displaystyle\sum_{y=mx-m+1}^{mx} d_{mt+i-1}(v_y) & x \le t, \\ \displaystyle\sum_{y=mt+1}^{mt+i} d_{mt+i-1}(v_y) & x = t+1. \end{cases}$$

https://doi.org/10.1017/jpr.2021.59 Published online by Cambridge University Press

By (2.2), for $x \le t$, the probability that v_{mt+i} connects to one of the vertices $v_{m(x-1)+1}, \ldots, v_{mx}$ (equivalently, the probability that the ith edge of w_{t+1} connects to w_x) is

$$\frac{\sum_{y=mx-m+1}^{mx} (d_{mt+i-1}(v_y) + \delta/m)}{(2+\delta/m)(mt+i-1)+1+\delta/m} = \frac{\delta + \sum_{y=mx-m+1}^{mx} d_{mt+i-1}(v_y)}{\delta(t+i/m)+2mt+2i-1}$$
$$= \frac{\delta + D_{t,i-1}(w_x)}{\delta(t+i/m)+2mt+2i-1}.$$

Similarly, the probability that v_{mt+i} selects one of $v_{mt+1}, \ldots, v_{mt+i}$ (equivalently, the probability that the ith edge of w_{t+1} is a loop) is

$$\frac{1 + i\delta/m + \sum_{j=1}^{i-1} d_{mt+i-1}(v_{mt+j})}{\delta(t+i/m)+2mt+2i-1} = \frac{1+i\delta/m + D_{t,i-1}(w_{t+1})}{\delta(t+i/m)+2mt+2i-1}.$$

Note that the two probabilities above are the same as those in (2.1) if we replace $D_{t,i-1}(w_x)$ with $d_{t,i-1}(x)$. Moreover, the two processes, $\{G'_{m,\delta}(t)\}$ and $\{G_{m,\delta}(t)\}$ as defined by (2.1), both start with m loops on the first vertex, which implies $d_{1,0}(\cdot) = D_{1,0}(\cdot)$. This gives us that $\{G_{m,\delta}(t)\}$ and $\{G'_{m,\delta}(t)\}$ are equivalent processes, that is, at every stage they produce the same random graph.

Acknowledgements

We thank the referees for numerous critical comments that genuinely helped us to improve the presentation of our results. We are sincerely grateful to the Editors for an efficient refereeing/editorial process. We dedicate this work to the memory of Irina Pittel, the wife of the third coauthor.

Funding information

During the time of this research, the first coauthor was supported by NSF Fellowship (award no. 1502650). Research of the second author is supported in part by NSF Grant DMS 1661063.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] ACAN, H. AND HITCZENKO, P. (2016). On a memory game and preferential attachment graphs. *Adv. Appl. Prob.* **48**, 585–609.
[2] ACAN, H. AND PITTEL, B. (2019). On connectivity, conductance and bootstrap percolation for the uniformly random k-out, age-biased graph. *Random Structures Algorithms* **56**, 37–62.
[3] BARABÁSI, I. AND ALBERT, R. (1999). Emergence of scaling in random networks. *Science* **286**, 509–512.
[4] BERGER, N., BORGS, C., CHAYES, J. T. AND SABERI, A. (2014). Asymptotic behavior and distributional limits of preferential attachment graphs. *Ann. Prob.* **42**, 1–40.
[5] BOLLOBÁS, B. (2001). *Random Graphs*, 2nd edn. Cambridge University Press.
[6] BOLLOBÁS, B. AND RIORDAN, O. (2000). Linearized chord diagrams and an upper bound for the Vassiliev invariants. *J. Knot Theory Ramifications* **9**, 847–853.
[7] BOLLOBÁS, B. and RIORDAN, O. (2002). Mathematical results on scale-free random graphs. In *Handbook of Graphs and Networks: From the Genome to the Internet*, eds S. Bornholdt and H. G. Schuster, pp. 1–34. Wiley-VCH.

[8] BOLLOBÁS, B. AND RIORDAN, O. (2003). Robustness and vulnerability of scale-free random graphs. *Internet Math.* **1**, 1–35.
[9] BOLLOBÁS, B. AND RIORDAN, O. (2004). The diameter of a scale-free random graph. *Combinatorica* **4**, 5–34.
[10] BOLLOBÁS, B., BORGS, C., CHAYES, J. AND RIORDAN, O. (2003). Directed scale-free graphs. In *Proceedings of the 14th Annual ACM–SIAM SODA Conference*, pp. 132–139. ACM Press.
[11] BOLLOBÁS, B., RIORDAN, O., SPENCER, J. AND TUSNÁDY, G. (2001). The degree sequence of a scale-free random graph process. *Random Structures Algorithms* **18**, 279–290.
[12] BORGS, C., BRAUTBAR, M., CHAYES, J., KHANNA, S. AND LUCIER, B. (2012). The power of local information in social networks. In *International Workshop on Internet and Network Economics WINE 2012: Internet and Network Economics* (Lecture Notes Comput. Sci. **7695**), pp. 406–419. Springer, Berlin and Heidelberg.
[13] COMTET, L. (1974). *Advanced Combinatorics*. D. Reidel, Dordrecht.
[14] COOPER, C. AND FRIEZE, A. M. (2003). Crawling on web graphs. *Internet Math.* **1**, 57–90.
[15] COOPER, C. AND FRIEZE, A. M. (2007). The cover time of the preferential attachment graph. *J. Combinatorial Theory B* **97**, 269–290.
[16] COOPER, C., KLASING, R. AND ZITO M. (2005). Lower bounds and algorithms for dominating sets in web graphs. *Internet Math.* **2**, 275–300.
[17] ERDŐS, P. AND RÉNYI, A. (1959). On random graphs I. *Publ. Math. Debrecen* **6**, 290–297.
[18] FLAXMAN, A., FRIEZE, A. M. AND FENNER, T. (2005). High degree vertices and eigenvalues in the preferential attachment graph. *Internet Math.* **2**, 1–20.
[19] FRIEZE, A. M. AND KAROŃSKI, M. (2016). *Introduction to Random Graphs*. Cambridge University Press.
[20] FRIEZE, A. M. AND PEGDEN, W. (2017). Looking for vertex number one. *Ann. Appl. Prob.* **27**, 582–630.
[21] FRIEZE, A. M., PÉREZ-GIMÉNEZ, X., PRAŁAT, P. AND REINIGER, B. (2019). Perfect matchings and Hamiltonian cycles in the preferential attachment model. *Random Structures Algorithms* **54**, 258–288.
[22] GILBERT, E. N. (1959). Random graphs. *Ann. Math. Statist.* **30**, 1141–1144.
[23] VAN DER HOFSTAD, R. (2017). *Random Graphs and Complex Networks*, Vol. 1. Cambridge University Press.
[24] JANSON, S. (2019). Random recursive trees and preferential attachment trees are random split trees. *Combinatorics Prob. Comput.* **28**, 81–99.
[25] JANSON, S. AND WARNKE, L. (2021). Preferential attachment without vertex growth: emergence of the giant component. *Ann. Appl. Prob.* **31**, 1523–1547.
[26] JANSON, S., ŁUCZAK, T. AND RUCIŃSKI, A. (2000). *Random Graphs.* Wiley, New York.
[27] KATONA, Z. AND MÓRI, T. (2006). A new class of scale free random graphs. *Statist. Prob. Lett.* **76**, 1587–1593.
[28] MIHAIL, M., PAPADIMITRIOU, C. AND SABERI, A. (2006). On certain connectivity properties of the internet topology. *J. Comput. System Sci.* **72**, 239–251.
[29] MOLLOY, M. (1998). The probabilistic method. In *Probabilistic Methods for Algorithmic Discrete Mathematics* (Algorithms and Combinatorics **16**), pp. 1–35. Springer, Berlin and Heidelberg.
[30] MÓRI, T. (2003). On random trees. *Studia Sci. Math. Hungar.* **39**, 143–155.
[31] MÓRI, T. (2005). The maximum degree of the Barabási–Albert random tree. *Combinatorics Prob. Comput.* **14**, 339–348.
[32] PEKÖZ, E., RÖLLIN, A. AND ROSS, N. (2017). Joint degree distributions of preferential attachment random graphs. *Adv. Appl. Prob.* **49**, 368–387.
[33] PITTEL, B. (1982). On the probable behaviour of some algorithms for finding the stability number of a graph. *Math. Proc. Camb. Phil. Soc.* **92**, 511–526.
[34] PITTEL, B. (1994). Note on the heights of random recursive trees and random m-ary search trees. *Random Structures Algorithms* **5**, 337–347.
[35] PITTEL, B. (2010). On a random graph evolving by degrees. *Adv. Math.* **223**, 619–671.
[36] PITTEL, B. (2021) On Bollobás–Riordan random pairing model of preferential attachment graph. *Random Structures Algorithms* **58**, 691–725.
[37] TALAGRAND, M. (1995) Concentration of measure and isoperimetric inequalities in product spaces. *Institut des Hautes Études Scientifiques, Publications Mathématiques* **81**, 73–205.