

# A PREFERENTIAL ATTACHMENT PROCESS APPROACHING THE RADO GRAPH

RICHARD ELWES

*School of Mathematics, University of Leeds, UK (r.h.elwes@leeds.ac.uk)*

(Received 7 May 2019; first published online 27 February 2020)

*Abstract* We consider a simple preferential attachment graph process, which begins with a finite graph and in which a new  $(t + 1)$ st vertex is added at each subsequent time step  $t$  that is connected to each previous vertex  $u \leq t$  with probability  $d_u(t)/t$ , where  $d_u(t)$  is the degree of  $u$  at time  $t$ . We analyse the graph obtained as the infinite limit of this process, and we show that, as long as the initial finite graph is neither edgeless nor complete, with probability 1 the outcome will be a copy of the Rado graph augmented with a finite number of either isolated or universal vertices.

*Keywords:* preferential attachment; Rado graph; scale-free network; Erdős–Rényi graph

2010 *Mathematics subject classification:* Primary 05C80  
Secondary 05C63

## 1. Introduction

Since its introduction by Barabási and Albert in [1], the mechanism of preferential attachment (PA) has been highly influential among scientists seeking to model real-world networks. In PA processes, a new vertex is introduced at each time step, which is then connected to each pre-existing vertex with a probability depending on the current degree of that vertex. The study of PA and related processes thus presents a new challenge in the mathematics of random graphs, differing from the classical approach of the Erdős–Rényi school who principally study structures arising from the following simpler process: at each time step, introduce a new vertex and connect it to each previous vertex with some fixed probability  $p$ .

Thus, an important question is whether the mathematics of PA processes can be developed to the same advanced level as the Erdős–Rényi theory. Recall that this theory has two distinct facets. First, researchers have analysed in great detail the finite graphs which emerge. Here, questions of interest include the emergence of a giant component and the degree distribution of the vertices, and analyses are typically highly sensitive to the value of  $p$ . See [2] for a comprehensive discussion of such matters.

The second angle of approach is to understand the infinite limit of the process. In this case, a remarkable theorem of Erdős and Rényi guarantees that, irrespective of the value of  $p \in (0, 1)$ , the resulting graph will, with probability 1, be isomorphic to the following important structure.

**Definition 1.1.** *The Rado graph* is a graph on a countably infinite set of vertices satisfying the following: given any disjoint finite sets of vertices  $U$  and  $V$ , there exists a vertex  $v$  connected to each vertex in  $V$  and none in  $U$ .

This graph exhibits many interesting properties. To start with (and justifying our use of the definite article), it is *countably categorical*, meaning that any two graphs obeying the definition are automatically isomorphic. Beyond this, all finite and countably infinite graphs isomorphically embed in it. For these and other reasons, this structure is an object of central importance in several branches of mathematics. See [5] for a recent survey.

This division between the finite and infinite is equally applicable to the study of PA processes. On the finite side, a good amount of progress has been made. Notably, in [7] and [6], Dereich and Mörters analyse the following family of PA processes: at stage  $t$ , we have a directed graph  $DG(t)$  into which a new vertex  $t + 1$  is introduced. For each previous vertex  $u$ , an edge from  $t + 1$  to  $u$  is introduced independently with probability  $f(I_u(t))/t$ , where  $f$  is a fixed sublinear function and  $I_u(t)$  is the indegree of  $u$  in  $DG(t)$ . The model considered in this paper (Definition 2.1 below) is essentially a linear version of theirs, although we consider our structures as *undirected* graphs. Dereich and Mörters successfully extract a great deal of valuable information about this process, including the distributions of in- and outdegrees.

This follows earlier work from Bollobás *et al.* [3] and Bollobás and Riordan [4] on analysing networks arising from a PA process in which a fixed number  $m$  of edges are added at each stage and are connected to previous vertices with probability directly proportional to their degrees.

The infinite limits of such processes have received less analysis, and this is our focus here. Our entry point is the paper of Kleinberg and Kleinberg [10], in which a process is studied whereby a single vertex and a constant number  $C$  of edges are added at each time step, with each new edge starting at the new vertex and with endpoint independently chosen among the pre-existing vertices with probability proportional to their degree. Thus, these structures are analysed as *directed multigraphs* in that each edge has a direction, and there may exist two or more edges sharing the same start and end-points. (Loops from a vertex to itself are not permitted, however.)

Kleinberg and Kleinberg show that, in each of the cases  $C = 1$  and  $C = 2$ , there is, up to isomorphism, a unique infinite limiting structure, which the process approaches with probability 1. (They also show that the analogous result fails for  $C \geq 3$ .)

We proceed in a similar spirit, pinning down the limiting structure up to isomorphism. However, in our model, the number of edges added at each stage is not prescribed, but is itself a random variable. Specifically, our model connects the new vertex  $t + 1$  to each previous vertex  $u \leq t$  with probability  $d_u(t)/t$ , where  $d_u(t)$  is the degree of  $u$  at time  $t$ . (As mentioned above, in the current model, edges are directionless, and parallel edges are not permitted.)

Thus, our graphs are denser than those of Kleinberg and Kleinberg: where the expected number of edges in their graphs is linear in the number of vertices, in our case it is quadratic (this is made explicit in Lemma 4.1 below).

In Theorem 2.3 below, we establish that, as long as the initial graph is neither complete nor edgeless, our process will with probability 1 approach the Rado graph, or a modification of it in which a finite number of universal or isolated vertices are incorporated.

In §5, we outline an adaptation of the foregoing machinery to a variation of the model with probability of attachment given by  $\lambda \cdot d_u(t)/t$  for some constant  $\lambda \in (0, 1]$ . We make the case (omitting the delicate conditional probability considerations) that if  $\lambda < 1$ , with probability 1 the limiting structure will be non-isomorphic to an augmented Rado graph.

In [8] (an as yet unpublished work developed simultaneously with the current paper), the author analyses a family of PA processes in which parallel edges are permitted and in which the number of edges added at each stage is prescribed. Specifically, it is shown that if  $f(t)$  edges are added at time  $t$ , where  $f$  is asymptotically bounded above and below by linear functions in  $t$ , then with probability 1 the process will approach the natural multigraph analogue of the Rado graph. (We refer the reader to [8, Definition 2.1] for the formal definition.)

## 2. The process

**Definition 2.1.** Let  $G'$  be any finite graph containing at least two vertices. We take its vertex set to be  $\{0, 1, \dots, v'\}$ , and we let  $G(t) := G'$  for all  $t \leq v'$ . For  $t \geq v'$ , we create a new graph  $G(t+1)$  by introducing a new vertex  $t+1$  which is connected to each previous vertex  $u \leq t$  with probability  $p_u(t+1) = d_u(t)/t$ , where  $d_u(t)$  is the degree of vertex  $u$  in  $G(t)$ .

Notice that, in the current model, as in [7], but in contrast to [8] and [11], the number of edges added at each stage is itself a random variable.

**Remark 2.2.**  $G(t)$  contains  $t+1$  vertices, and thus  $d_u(t)$  may take any value between 0 and  $t$ . Now if  $d_u(t') = 0$  for any  $t'$ , then automatically  $d_u(t) = 0$  for all  $t \geq u$ . Likewise, if  $d_u(t') = t'$ , then  $d_u(t) = t$  for all  $t \geq u$ . Thus isolated vertices remain isolated, and universal vertices (i.e., vertices connected to every other) remain universal. Of course, no graph can contain both.

We shall call isolated and universal vertices **non-standard** and all others **standard**.

Our interest is the infinite limit of this process  $G(\infty)$ , and we shall prove the following theorem.

### Theorem 2.3.

1. If  $G'$  is complete, then  $G(\infty)$  is complete.
2. If  $G'$  is edgeless, then so too is  $G(\infty)$ .
3. For any other  $G'$ , with probability 1, the infinite limit  $G(\infty)$  is isomorphic to one of the following:
  - the Rado graph, augmented with a finite number of isolated vertices; or
  - the Rado graph, augmented with a finite number of universal vertices.

The first two clauses of Theorem 2.3 are immediate from Remark 2.2 above, so we concentrate on the third, and we make the standing assumption that  $G'$  is neither complete nor edgeless (from which it immediately follows that no stage is complete or edgeless).

Since, by Remark 2.2, non-standard vertices can be recognized as soon as they appear and are of little interest, we might amend the process by immediately discarding each non-standard vertex when it appears. Theorem 2.3 then guarantees that we will obtain the Rado graph as the infinite limit of this modified process.

It will be convenient to amend the model as follows: we colour each edge as described in Definition 2.1 black, and we introduce the new rule that every pair of distinct vertices in  $G(t)$  is connected by a white edge if and only if it is not connected by a black edge. (Thus the white graph is the complement of the black graph. We imagine that, on white paper, these edges will become invisible.) Now let  $d_u^w(t)$  be the white degree of the vertex  $u$  at time  $t$ , and let  $d_u^b(t)$  be its black degree. This introduces a useful symmetry to the process.

**Remark 2.4.** At any stage  $(t)$ , for any vertex  $u \leq t$ , it holds that  $d_u^w(t) + d_u^b(t) = t$ . Thus, in  $G(t + 1)$ , the probability that vertex  $u$  is connected to vertex  $t + 1$  with a white edge is precisely  $1 - (d_u^b(t)/t) = d_u^w(t)/t$ .

### 3. The standard vertices

**Remark 3.1.** For a fixed vertex  $u$ , we can consider the edges of each colour emanating from  $u$  as black or white balls within a Pólya urn. Our process then is identical to selecting, at each stage, a ball from the urn uniformly at random and then returning it along with a second ball of the same colour. This observation would allow us to appeal to known facts about the limiting proportions of black balls in the urn as given by a beta distribution  $\text{Beta}(d_u^b(t_0), d_u^w(t_0))$ . See, for instance, [9]. From this, we could easily derive the main result of this section, Proposition 3.9.

Nevertheless, we opt to provide a self-contained elementary argument, since we shall not need the details of the beta distribution, and since also the main technical ingredient of our argument, Proposition 3.8, will, in any case, be required for separate purposes in §4.

**Lemma 3.2.** *For any standard vertex  $u$  in any stage  $G(t_0)$ , the probability that  $u$  never receives another black edge is 0. The same goes for white edges.*

**Proof.** Suppose that, at time  $t_0$ ,  $u$  has black degree  $d(t_0) = D > 0$ . The probability that  $u$  never receives another black edge is therefore

$$\prod_{t=t_0}^{\infty} \left(1 - \frac{D}{t}\right).$$

We shall show that this is 0. Taking logarithms, it is therefore enough to show that

$$\sum_{t=t_0}^{\infty} \ln \left(1 + \frac{D}{t - D}\right)$$

diverges to  $\infty$ . This follows from the divergence of the harmonic series since, for all small enough  $x$ , we know that  $\ln(1 + x) > \frac{1}{2}x$ .

An identical argument applies after replacing ‘black’ with ‘white’. □

**Corollary 3.3.** *Given any state of the graph  $G(t_0)$  containing any standard vertex  $u$ , with probability 1 it will be true that  $d_u^b(t) \rightarrow \infty$  and  $d_u^w(t) \rightarrow \infty$  as  $t \rightarrow \infty$ .*

**Proof.** This follows automatically from Lemma 3.2 by the countable additivity of the probability measure. □

**Definition 3.4.** Given a vertex  $u$  and a colour  $c \in \{b, w\}$ , we define

$$U_u^c(t + 1) := d_u^c(t + 1) - d_u^c(t),$$

$$X_u^c(t) := \frac{d_u^c(t)}{t}.$$

We shall suppress the subscript  $u$  and superscript  $c$  in the above when they are clear from context.

**Lemma 3.5.** *For any vertex  $u$  and colour  $c \in \{b, w\}$ , we have the following.*

- (i) *For all  $t$ ,  $X_u^c(t) \in [0, 1]$ . If  $u$  is standard, then  $X_u^c(t) \in (0, 1)$ .*
- (ii)  *$(X_u^c(t))_{t \geq u}$  is a martingale.*
- (iii) *With probability 1, there exists  $x_u^c \in [0, 1]$  such that  $d_u^c(t)/t \rightarrow x_u^c$ .*
- (iv) *Furthermore,  $x_u^b + x_u^w = 1$ .*

**Proof.** The first part is immediate from the definitions. Writing  $X$  for  $X_u^c$ , we show that  $X(t)$  is a martingale: i.e.,

$$\begin{aligned} \mathbf{E}\left(X(t + 1) \mid X(t) = \frac{d}{t}\right) &= \frac{1}{t + 1} \mathbf{E}(d(t + 1) \mid d(t) = d) \\ &= \frac{1}{t + 1} \left(d + \frac{d}{t}\right) = \frac{d}{t} = X(t). \end{aligned}$$

Part (iii) follows from parts (i) and (ii) by Doob’s convergence theorem. Part (iv) follows too, since  $X^b(t) + X^w(t) = 1$  by Remark 2.4. □

**Remark 3.6.** By Lemma 3.5(i), the martingale  $(X_u^c(t))_{t \geq u}$  is bounded in  $\mathcal{L}_2$ .

Our next goal is to show (in Proposition 3.9 below) that with probability 1 we have  $0 < x_u^c < 1$ . Towards this, we recall some martingale machinery from [10], which we express more generally for subsequent reuse. First, recall the Kolmogorov–Doob inequality.

**Theorem 3.7 (Kolmogorov–Doob inequality).** *Suppose that  $(\tilde{Z}(n))_{n \in \mathbf{N}}$  is a non-negative submartingale and  $\alpha > 0$ . Then, for any  $N \in \mathbf{N}$ ,*

$$\mathbf{P}\left(\max_{n \leq N} \tilde{Z}(n) \geq \alpha\right) \leq \frac{1}{\alpha} \mathbf{E}(\tilde{Z}(N)).$$

**Proposition 3.8.** *Suppose that  $(Z(t))_{t \in \mathbf{N}}$  is a martingale with limit  $z$  such that, with probability 1, there exist  $\alpha, A > 0$  and  $t_1 \geq 3$  (each of whose values may be given by random variables) so that:*

- (i)  $Z(t) \in (0, 1)$  for all  $t \geq t_1$ ;
- (ii)  $t \cdot Z(t) > \alpha$  for all  $t \geq t_1$ ;
- (iii) for all  $t > m \geq t_1$ ,

$$\mathbf{E}(Z(t + 1)^2 - Z(t)^2 | Z(m)) < \frac{A}{t^2} \cdot Z(m);$$

- (iv)  $\beta := 8A/\alpha < 1$ .

Then  $\mathbf{P}(z > 0) = 1$ .

**Proof.** The argument is essentially contained in [10], but we include it here for completeness. Given any  $n > m \geq t_1$ , define

$$\tilde{Z}_m(n) := (Z(n) - Z(m))^2.$$

Thus, by our third hypothesis, for all  $n > m$ ,

$$\begin{aligned} \mathbf{E}(\tilde{Z}_m(n) | Z(m)) &= \sum_{t=m}^{n-1} \mathbf{E}(Z(t + 1)^2 | Z(m)) - \mathbf{E}(Z(t)^2 | Z(m)) \\ &< A \cdot Z(m) \cdot \sum_{t=m}^{\infty} \frac{1}{t^2} < \frac{2A}{m} \cdot Z(m) \end{aligned}$$

since  $m \geq 3$ .

Notice too that, for fixed  $m$ , the sequence  $(\tilde{Z}_m(n))_n$  forms a submartingale.

Now, beginning at any time  $t_0$ , we may define a sequence of times:  $n_0 = t_0$ . Let  $n_{i+1}$  be the least  $n \geq n_i$  (if any exists) such that  $Z(n) < \frac{1}{2}Z(n_i)$ . Otherwise,  $n_{i+1} = \infty$ .

We next apply the Kolmogorov–Doob inequality (Theorem 3.7) to  $\tilde{Z}_{n_i}(n)$ .

$$\begin{aligned} \mathbf{P}(n_{i+1} < \infty | n_i < \infty) &= \mathbf{P}\left(\min_{n \geq n_i} Z(n) < \frac{1}{2}Z(n_i) | Z(n_i)\right) \\ &\leq \mathbf{P}\left(\max_{n \geq n_i} \tilde{Z}_{n_i}(n) > \frac{1}{4}Z(n_i)^2 | Z(n_i)\right) \\ &= \lim_{N \rightarrow \infty} \mathbf{P}\left(\max_{n: N \geq n \geq n_i} \tilde{Z}_{n_i}(n) > \frac{1}{4}Z(n_i)^2 | Z(n_i)\right) \\ &\leq \frac{4}{Z(n_i)^2} \cdot \lim_{N \rightarrow \infty} \mathbf{E}(\tilde{Z}_{n_i}(N) | Z(n_i)). \end{aligned}$$

The above holds if we start at any time  $t_0$ . However, with probability 1, the time  $t_1$  exists as described. In the case  $t_0 \geq t_1$ , we find that

$$\mathbf{P}(n_{i+1} < \infty | n_i < \infty) < \frac{8A}{Z(n_i) \cdot n_i} < \beta.$$

Since  $\beta < 1$ , it follows that the event that every  $n_i$  is finite has probability zero, giving the result. □

**Proposition 3.9.** *Given any standard vertex  $u$ , we have  $\mathbf{P}(x_u^c > 0) = \mathbf{P}(x_u^c < 1) = 1$ .*

**Proof.** As usual, we write  $X$  for  $X_u^c$  and  $x$  for  $x_u^c$ .

The result will follow by Proposition 3.8 applied to  $X(t)$ , once we have verified its hypotheses. The first follows by the assumption that  $u$  is standard. For the second, note that, by Corollary 3.3 with probability 1, we will see  $d_u^c(t_0) \geq 16$  (say) for some  $t_0$ . Thus we may take  $\alpha = 16$ .

For the third, recall that  $U(t + 1) := d(t + 1) - d(t)$ . So  $U(t + 1)$  is a Bernoulli variable with

$$\mathbf{E}(U(t + 1) | d(t) = d) = \mathbf{E}(U(t + 1)^2 | d(t) = d) = \frac{d}{t}.$$

Since  $d(t + 1)^2 = d(t)^2 + 2d(t)U(t + 1) + U(t + 1)^2$ ,

$$\begin{aligned} \mathbf{E}(d(t + 1)^2 | d(t) = d) &= d^2 + 2d \cdot \frac{d}{t} + \frac{d}{t} \\ &< \frac{d}{t} + d^2 \left(1 + \frac{1}{t}\right)^2 \\ &= X + X^2(t + 1)^2. \end{aligned}$$

Now

$$\begin{aligned} \mathbf{E}(X(t + 1)^2 | d(t) = d) &= \frac{1}{(t + 1)^2} \cdot \mathbf{E}(d(t + 1)^2 | d(t) = d) \\ &< \frac{1}{(t + 1)^2} X(t) + X(t)^2. \end{aligned}$$

Since  $\mathbf{E}(X(t) | X(m)) = X(m)$ , by the law of total expectation,

$$\mathbf{E}(X(t + 1)^2 - X(t)^2 | X(m)) < \frac{1}{t^2} X(m). \tag{3.1}$$

Taking  $A = 1$  gives the result. □

We may now prove the first part of Theorem 2.3.

**Proposition 3.10.** *The subgraph of  $G(\infty)$  comprising the standard vertices is isomorphic to the Rado graph.*

**Proof.** Let  $U = \{u_i : i \leq n\}$  and  $V = \{v_j : j \leq m\}$  be disjoint finite sets of standard vertices, and let  $t_0$  be any stage of the process. Our initial goal is to establish the existence of a *witness* for  $(U, V)$ , i.e., a vertex adjacent to each  $u_i$  and no  $v_j$ , which emerges at some time  $t_1 > t_0$ .

For each  $u_i$  (respectively,  $v_j$ ), by Lemma 3.5 and Proposition 3.9 there exists  $x_i$  (respectively,  $y_j$ ) in  $(0, 1)$  representing the limiting proportion of neighbours of  $u_i$  ( $v_j$ ) among all vertices. Since a vertex  $w > \max\{u_i, v_j : i \leq n, j \leq m\}$  is connected to each  $u_i$  or  $v_j$  independently of whether it is connected to the others, it follows with probability 1 that the limiting proportion of witnesses for  $(U, V)$  is given by  $\prod_{i=1}^n x_i \cdot \prod_{j=1}^m (1 - y_j) > 0$ .

Thus with probability 1, there are infinitely many such witnesses, and thus at least one arises after stage  $t_0$ .

To complete the proof, we enumerate all possible pairs of disjoint finite sets  $(U_i, V_i)_{i \in \mathbf{N}}$ . Then let  $t_0$  be the first stage at which  $(U_0, V_0)$  receives a witness and let  $t_{i+1}$  be the first stage strictly after  $t_i$  at which  $(U_{i+1}, V_{i+1})$  receives a witness.

Then, by the argument above and the countable additivity of the probability measure, with probability 1 every  $t_i$  is finite, as required. □

#### 4. The non-standard vertices

Recall our standing assumption that  $G'$  (and thus every stage of the process) is neither complete nor edgeless. Thus we remain under the hypotheses of part 3 of Theorem 2.3. Our goal in this section is Proposition 4.6, where we show that the non-standard vertices in  $G(\infty)$  remain finite with probability 1. As in the previous section, we employ the machinery of martingales.

**Lemma 4.1.** *Write  $\mathcal{E}^c(t)$  for the number of edges of colour  $c$  in  $G(t)$ . Define  $Y^c(t) := \mathcal{E}^c(t)/(t(t + 1))$ . Then:*

- (i)  $Y^c(t) \in (0, 1)$ ;
- (ii)  $(Y^c(t))_t$  is a Martingale; and
- (iii)  $Y^c(t) \rightarrow y^c$  for some  $y^c \in [0, 1]$ .

**Proof.** Omitting the superscript  $c$ , since  $G(t)$  is not edgeless (in either colour) we know that  $Y(t) > 0$ . Also,  $\mathcal{E}(t) < \binom{t}{2}$ , which implies that  $Y(t) < 1$ . Now,

$$\mathbf{E}(d_{t+1}(t + 1) | G(t)) = \sum_{u=0}^t \frac{d_u(t)}{t} = \frac{2\mathcal{E}(t)}{t}.$$

Also  $\mathcal{E}(t + 1) = \mathcal{E}(t) + d_{t+1}(t + 1)$ . It follows that

$$\mathbf{E}(\mathcal{E}(t + 1) | G(t)) = \left( \frac{t + 2}{t} \right) \cdot \mathcal{E}(t).$$

Therefore  $\mathbf{E}(Y(t + 1) | Y(t)) = Y(t)$  and thus  $Y$  is a martingale. Part (iii) follows by Doob’s convergence theorem. □

**Remark 4.2.** The Martingale  $Y(t)$  converges in  $\mathcal{L}_2$ .

Our next aim is to show that with probability 1, in fact,  $y > 0$ . This will again follow from Proposition 3.8, once we have shown that its hypotheses hold. The first is part (i) of Lemma 4.1. We turn our attention to the second hypothesis.



**Proposition 4.3.** *Given any state of the graph  $G(t_0)$  and  $c \in \{b, w\}$ , with probability 1 there exist  $\alpha > 0$  and  $t_1 \geq t_0$  such that, for all  $t \geq t_1$ ,*

$$t \cdot Y^c(t) > \alpha.$$

**Proof.** We suppress the superscript  $c$ , as usual. First, notice that, by definition,  $t \cdot Y(t) = 1/(t + 1) \cdot \mathcal{E}(t)$ .

Let  $\varepsilon > 0$ . We show that suitable  $\alpha$  and  $t_1$  exist with probability  $> 1 - \varepsilon$ . Pick a standard vertex  $u$  in  $G(t_0)$  and pick  $i$  large enough that  $2^{-i} < \varepsilon$ . Then, by Proposition 3.9, with probability  $> 1 - \varepsilon$  we have  $x_u \geq 1/2^i \cdot 8/t_0$ . Thus, for some large enough  $t_1$  and all  $t \geq t_1$ , we have  $d_u(t) > (1/2^i)(4/t_0) \cdot t$  and thus  $\mathcal{E}(t) > (1/2^i)(4/t_0) \cdot t$  and  $(1/(t + 1))\mathcal{E}(t) > \alpha := (1/2^i)(2/t_0)$ .  $\square$

**Proposition 4.4.**

$$\mathbf{E}(Y(t + 1)^2|Y(t)) - Y(t)^2 < \frac{2}{t^3} \cdot Y(t).$$

**Proof.** Suppressing the superscript  $C$  and writing  $\mathcal{E}$  for  $\mathcal{E}(t)$  and  $d$  for  $d_{t+1}(t + 1)$ , we have

$$\begin{aligned} &\mathbf{E}(Y(t + 1)^2|Y(t)) - Y(t)^2 \\ &= \frac{1}{(t + 1)^2(t + 2)^2} \mathbf{E}((\mathcal{E} + d)^2|Y(t)) - \frac{1}{t^2(t + 1)^2} \mathcal{E}^2 \\ &= \frac{1}{t^2(t + 1)^2(t + 2)^2} (t^2(\mathcal{E}^2 + 2\mathcal{E}\mathbf{E}(d|Y(t))) + \mathbf{E}(d^2|Y(t)) - (t + 2)^2\mathcal{E}^2). \end{aligned}$$

Now, conditioning on  $\mathcal{E}(t)$ , the random variable  $d = d_{t+1}(t + 1)$  has a Poisson-binomial distribution with expectation  $2\mathcal{E}(t)/t$  and variance

$$\sum_{u=0}^t \left(1 - \frac{d_u(t)}{t}\right) \frac{d_u(t)}{t} < \sum_{u=0}^t \frac{d_u(t)}{t} = \frac{2\mathcal{E}(t)}{t},$$

meaning that

$$\begin{aligned} \mathbf{E}(d^2|Y(t)) &= \mathbf{E}(d|Y(t))^2 + \mathbf{Var}(d|Y(t)) \\ &< \frac{4\mathcal{E}^2}{t^2} + \frac{2\mathcal{E}}{t}. \end{aligned}$$

Thus

$$\begin{aligned} &\mathbf{E}(Y(t + 1)^2|Y(t)) - Y(t)^2 \\ &< \frac{1}{t^2(t + 1)^2(t + 2)^2} \left( t^2 \left( \mathcal{E}^2 + \frac{4\mathcal{E}^2}{t} + \frac{4\mathcal{E}^2}{t^2} + \frac{2\mathcal{E}}{t} \right) - (t + 2)^2 \mathcal{E}^2 \right) \end{aligned}$$

$$\begin{aligned}
 &= \frac{2t\mathcal{E}}{t^2(t+1)^2(t+2)^2} \\
 &= \frac{2}{(t+1)(t+2)^2} \cdot Y(t) \\
 &< \frac{2}{t^3} \cdot Y(t). \quad \square
 \end{aligned}$$

**Proposition 4.5.** *Given any state of the process  $G(t_0)$  and  $c \in \{b, w\}$ , with probability 1 we have  $y^c > 0$ .*

**Proof.** As discussed above, we proceed using Proposition 3.8. The first hypothesis comes from Lemma 4.1 and the second comes from Proposition 4.3. The third follows from Proposition 4.4 above: we may pick  $A$  as we like, say,  $A = \alpha/16$ . By increasing  $t_1$  by some predictable amount, we obtain that, for  $t \geq t_1$ ,

$$\mathbf{E}(Y(t+1)^2 - Y(t)^2 | Y(t)) < \frac{A}{t^2} \cdot Y(t). \tag{4.1}$$

The law of total expectation now gives the third hypothesis, and the fourth is immediate from our choice of  $A$ , giving the result.  $\square$

The following result completes the proof of Theorem 2.3.

**Proposition 4.6.** *Given any state  $G(t_0)$ , the number of non-standard vertices in  $G(\infty)$  will be finite with probability 1.*

**Proof.** Let  $\varepsilon > 0$  and let  $\Delta$  be the event  $|\{t : d_t(t) = 0\}| < \infty$ . Our goal is to show that  $\mathbf{P}(\Delta) \geq 1 - \varepsilon$ . (This and everything that follows is conditioned upon  $G(t_0)$ , which we suppress.)

First, by Proposition 4.5, there exists  $\gamma > 0$  such that

$$\mathbf{P}(y < 2\gamma) < \varepsilon. \tag{4.2}$$

(The 2 is included simply for our convenience.) Next, by Proposition 4.4 above and the law of total expectation, for any  $t \geq t_1 \geq t_0$ ,

$$\mathbf{E}(Y(t+1)^2 - Y(t)^2 | Y(t_1)) < \frac{2}{t^3}.$$

Summing successive expressions, by the linearity of expectation, there is some  $C > 0$  so that, for all  $T \geq t$ , we see that

$$\mathbf{E}(Y(T+1)^2 - Y(t)^2 | Y(t_1)) < \frac{C}{t^2}.$$

Thus, by linearity and Remark 4.2,

$$\mathbf{E}(y^2 - Y(t)^2 | Y(t_1)) < \frac{C}{t^2}. \tag{4.3}$$

Furthermore, by Remark 4.2 again,  $\mathbf{E}(y \mid Y(t)) = Y(t)$  and thus

$$\mathbf{E}((y - Y(t))^2 \mid Y(t)) = \mathbf{E}(y^2 - Y(t)^2 \mid Y(t)).$$

So, by the law of total expectation and bound (4.3),

$$\mathbf{E}((y - Y(t))^2) < \frac{C}{t^2}.$$

Thus by Markov's theorem, for any  $\delta > 0$ ,

$$\mathbf{P}(|y - Y(t)| > \delta) < \frac{C}{\delta^2 t^2}.$$

Since  $\sum_{t=t_0}^{\infty} C/\delta^2 t^2 < \infty$ , by the Borel–Cantelli lemma, the number of  $t$  for which it holds that  $|y - Y(t)| > \delta$  will be finite with probability 1.

Taking  $\delta = \gamma$  and noticing that  $\mathcal{E}(t) < \gamma t^2 \Rightarrow Y(t) < \gamma$ , we define  $\Gamma$  to be the event  $|\{t : \mathcal{E}(t) < \gamma t^2\}| < \infty$ . Then

$$\mathbf{P}(\Gamma \mid y \geq 2\gamma) = 1. \tag{4.4}$$

Also, for any  $\xi > 0$  (the case  $\xi = \gamma$  is of primary interest),

$$\begin{aligned} \mathbf{E}(d_{t+1}(t+1) \mid \mathcal{E}(t) \geq \xi t^2) &= \mathbf{E}\left(\sum_{u=0}^t U_u(t) \mid \mathcal{E}(t) \geq \xi t^2\right) \\ &= \mathbf{E}\left(\sum_{u=0}^t \frac{d_u(t)}{t} \mid \mathcal{E}(t) \geq \xi t^2\right) \\ &= \frac{2 \cdot \mathbf{E}(\mathcal{E}(t) \mid \mathcal{E}(t) \geq \xi t^2)}{t} \geq 2\xi t. \end{aligned}$$

Since each  $U_u(t) \in [0, 1]$ , applying Hoeffding's inequality gives

$$\begin{aligned} \mathbf{P}(d_{t+1}(t+1) = 0 \mid \mathcal{E}(t) \geq \xi t^2) \\ \leq \mathbf{P}(\mathbf{E}(d_{t+1}(t+1)) - d_{t+1}(t+1) > \xi t \mid \mathcal{E}(t) \geq \xi t^2) \\ \leq e^{-(2\xi^2 t^2 / (t+1))} < e^{-\xi^2 t}. \end{aligned}$$

Furthermore, clearly,  $\mathbf{E}(y \mid \mathcal{E}(t) \geq \gamma t^2) \geq \mathbf{E}(y)$ , and thus  $\mathbf{P}(y \geq 2\gamma \mid \mathcal{E}(t) \geq \gamma t^2) \geq \mathbf{P}(y \geq 2\gamma)$ . Hence by Bayes' theorem,

$$\begin{aligned} \mathbf{P}(\mathcal{E}(t) \geq \gamma t^2 \mid y \geq 2\gamma) &= \frac{\mathbf{P}(y \geq 2\gamma \mid \mathcal{E}(t) \geq \gamma t^2)}{\mathbf{P}(y \geq 2\gamma)} \cdot \mathbf{P}(\mathcal{E}(t) \geq \gamma t^2) \\ &\geq \mathbf{P}(\mathcal{E}(t) \geq \gamma t^2). \end{aligned}$$

By (4.4),

$$\mathbf{P}(\mathcal{E}(t) \geq \gamma t^2 \mid \Gamma \& y \geq 2\gamma) = \mathbf{P}(\mathcal{E}(t) \geq \gamma t^2 \mid y \geq 2\gamma),$$

so

$$\mathbf{P}(\mathcal{E}(t) \geq \gamma t^2 \mid \Gamma \& y \geq 2\gamma) \geq \mathbf{P}(\mathcal{E}(t) \geq \gamma t^2).$$

Likewise,

$$\begin{aligned} \mathbf{P}(d_{t+1}(t+1) = 0 \mid \mathcal{E}(t) \geq \gamma t^2 \& \Gamma \& y \geq 2\gamma) \\ \leq \mathbf{P}(d_{t+1}(t+1) = 0 \mid \mathcal{E}(t) \geq \gamma t^2) \\ < e^{-\gamma^2 t}. \end{aligned}$$

Now  $\sum_{t=t_0}^{\infty} e^{-\gamma^2 t}$  converges to a finite limit, and thus, by the Borel–Cantelli lemma again, we see that

$$\mathbf{P}(\Delta \mid \Gamma \& y \geq 2\gamma) = 1. \tag{4.5}$$

The result follows from (4.2), (4.4) and (4.5). □

### 5. Closing comments on generalizations

The current work considers only one model (modulo the initial graph), so it is natural to ask whether the result generalizes to related models, such as one with attachment probability  $\lambda \cdot \frac{d}{t}$  for some  $\lambda \in (0, 1]$ . We offer an informal argument that it is only in the case  $\lambda = 1$  that the infinite limit will be the Rado graph.

Much of the preceding theory goes through. In particular, one can define a martingale

$$X_u(t) := \frac{d(t)}{u \cdot \prod_{j=u}^{t-1} (1 + \frac{\lambda}{j})}.$$

Considering the limit of this martingale and the characterization of the  $\Gamma$ -function as  $\Gamma(\lambda) := \lim_{t \rightarrow \infty} (t^\lambda / \prod_{j=0}^t (1 + \lambda/j))$ , one may find  $C \geq 0$  so that, for all large enough  $t$ , we have  $d(t) \leq C \cdot t^\lambda$ .

Now consider (as per the proof of Proposition 3.10) a witness request  $(U, V)$ , where  $|U| = n$ . (For simplicity, we can take  $V = \emptyset$ .) For all large enough  $t$ , the probability of a suitable witness appearing at time  $t$  is bounded above by  $c \cdot t^{(\lambda-1)n}$  for some  $c > 0$ . Thus, omitting the intricacies of various conditional probability calculations, it is only in the case  $\lambda = 1$  that we expect to be able to guarantee the eventual appearance of such a witness. In particular, among the infinitely many witness requests  $(U, V)$ , where  $|U| > 1/(1 - \lambda)$ , with probability 1 at least one will fail to be satisfied, making the limit non-isomorphic to an augmented Rado graph.

### References

1. A.-L. BARABÁSI AND R. ALBERT, Emergence of scaling in random networks, *Science* **286**(5439) (1999), 509–512.
2. B. BOLLOBÁS, *Random graphs*, 2nd edn, Cambridge Studies in Advanced Mathematics (Cambridge University Press, 2001).
3. B. BOLLOBÁS, O. RIORDAN, J. SPENCER AND G. TUSNÁDY, The degree sequence of a scalefree random graph process, *Random Struct. Algorithms* **18**(3) (2001), 279–290.

4. B. BOLLOBÁS AND O. RIORDAN, The diameter of a scale-free random graph, *Combinatorica* **24**(1) (2004), 5–34.
5. P. CAMERON, The random graph, in *The mathematics of Paul Erdős II* (ed. R. L. Graham, J. Nešetřil and S. Butler), pp. 333–351 (Springer, Berlin–Heidelberg, 1997).
6. S. DEREICH AND P. MÖRTERS, Random networks with sublinear preferential attachment degree evolutions, *Electron. J. Probab.* **43**(14) (2009), 1222–1267.
7. S. DEREICH AND P. MÖRTERS, Random networks with sublinear preferential attachment: the giant component, *Ann. Probab.* **41**(1) (2013), 329–384.
8. R. ELWES, Preferential attachment processes approaching the rado multigraph, preprint (arXiv:1502.05618, 2015).
9. D. FREEDMAN, Bernard Friedman’s urn, *Ann. Math. Stat.* **36**(3) (1965), 956–970.
10. R. KLEINBERG AND J. KLEINBERG, Isomorphism and embedding problems for infinite limits of scale-free graphs. *Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms (SODA)*, pp. 277–286 (Society for Industrial and Applied Mathematics, 2005).
11. R. OLIVEIRA AND J. SPENCER, Connectivity transitions in networks with super-linear preferential attachment, *Internet Math.* **2**(2) (2005), 121–163.