

# DIRECTED PREFERENTIAL ATTACHMENT MODELS: LIMITING DEGREE DISTRIBUTIONS AND THEIR TAILS

TOM BRITTON,\* *Stockholm University*

## Abstract

The directed preferential attachment model is revisited. A new exact characterization of the limiting in- and out-degree distribution is given by two *independent* pure birth processes that are observed at a common exponentially distributed time  $T$  (thus creating dependence between in- and out-degree). The characterization gives an explicit form for the joint degree distribution, and this confirms previously derived tail probabilities for the two marginal degree distributions. The new characterization is also used to obtain an explicit expression for tail probabilities in which both degrees are large. A new generalized directed preferential attachment model is then defined and analyzed using similar methods. The two extensions, motivated by empirical evidence, are to allow double-directed (i.e. undirected) edges in the network, and to allow the probability of connecting an ingoing (outgoing) edge to a specified node to also depend on the out-degree (in-degree) of that node.

*Keywords:* Preferential attachment; directed network; birth processes; tail distribution.

2010 Mathematics Subject Classification: Primary 60K40  
Secondary 05C80

## 1. Introduction and models

The (undirected) preferential attachment model (PA) is a random network model defined by Barabási and Albert [1] and later, more rigorously, by Bollobás *et al.* [3]. To start off, the network consists of one single node without any edge. At each time step  $k = 1, 2, \dots$ , a new node with  $m$  (a fixed integer parameter in the model) new edges connected to it is added. Each of the new edges of the node is connected, independently, to existing nodes, and the probability of connecting to a specific node with current degree  $i$  is proportional to  $i$ . Two novel features, as compared to most other network models at the time, were that it was defined sequentially, thus with nodes having different ages, and that the degree distribution of nodes in a large network were shown to have power-law tails rather than exponentially decaying tail probabilities.

In 2003, Bollobás *et al.* [2] defined a related model, but now for a network in which edges are directed rather than undirected. As in the undirected model, edges/nodes are entered at each discrete time step. However, now one of three different possibilities may happen: (1) either a new node with an *outgoing* edge is added (with probability  $\alpha$ ), or (2) a new directed *edge* but no new node is added (with probability  $\beta$ ), or (3) a node with an *ingoing* edge is added (with probability  $\gamma = 1 - \alpha - \beta$ ). In the first event, the edge connects to (i.e. points at) a node  $u$  having current in- and out-degree  $i$  and  $j$  with probability proportional to  $i + \delta_1$ . In the second

---

Received 5 October 2018; revision received 2 October 2019.

\* Postal address: Department of Mathematics, Stockholm University, SE-106 91 Stockholm, Sweden. Email address: [tom.britton@math.su.se](mailto:tom.britton@math.su.se)

event the edge starts from  $u$  with probability proportional to  $j + \delta_O$  and, independently, ends at  $u$  with probability proportional to  $i + \delta_I$ . In the third and final event the edge starts from  $u$  with probability proportional to  $j + \delta_O$ . The procedure of adding nodes/edges is initiated by having one single node without edges (however, as shown in [2], the exact starting configuration is not important for the limiting degree distribution as long as it is finite).

The model has four parameters:  $\alpha, \beta, \delta_I, \delta_O$  (recall that  $\gamma = 1 - \alpha - \beta$ ). It is important that  $\delta_I > 0$  and  $\delta_O > 0$ , otherwise nodes born with in-degree 0 will never get positive in-degree, and similarly for the remaining nodes born with out-degree 0. In order to avoid some less interesting special cases of the model (which would require special attention in the analyses) we will also assume that  $\alpha > 0$  and  $\gamma > 0$ . We term this model the directed preferential attachment (DPA) model.

In [2] it was shown that  $E(X_i(n)) = \varphi_i n + o(n)$ ,  $E(Y_j(n)) = \phi_j n + o(n)$ , and  $E(M_{ij}(n)) = f_{ij} n + o(n)$ , where  $X_i(n)$  denotes the number of nodes with in-degree  $i$ ,  $Y_j(n)$  denotes the number of nodes with out-degree  $j$ , and  $M_{ij}(n)$  denotes the number of nodes having in-degree  $i$  and out-degree  $j$ , at time step  $n$ . Bollobás *et al.* [2] did not obtain explicit expressions for  $\varphi_i, \phi_j$ , or  $f_{ij}$ , but instead derived limiting properties for large  $i$  and  $j$  (sending either, but not both, off to infinity for  $f_{ij}$ ). More specifically, they showed that  $\varphi_i \sim i^{-(1+1/c_1)}$  and  $\phi_j \sim j^{-(1+1/c_0)}$ , where

$$c_1 = \frac{\alpha + \beta}{1 + \delta_I(\alpha + \gamma)}, \tag{1}$$

$$c_0 = \frac{\gamma + \beta}{1 + \delta_O(\alpha + \gamma)} \tag{2}$$

( $a_i \sim b_i$  meaning  $a_i/b_i \rightarrow c$  for some  $0 < c < \infty$  as  $i \rightarrow \infty$ ). As for the bivariate distribution  $f_{ij}$  it was shown that, having one of  $i$  and  $j$  fixed and letting the other tend to infinity,  $f_{ij} \sim i^{-x_1}$  (as  $i \rightarrow \infty$  and  $j$  fixed) and  $f_{ij} \sim j^{-x_0}$  (as  $j \rightarrow \infty$  and  $i$  fixed), where  $x_1 = 1 + 1/c_1 + c_0(\delta_O + I_{(\gamma\delta_O=0)})/c_1$ , and  $x_0 = 1 + 1/c_0 + c_1(\delta_I + I_{(\alpha\delta_I=0)})/c_0$  (note that there is a small misprint the first time the expressions appear in [2]). The method used when proving the results is by analyzing the evolution of the vector-valued processes  $(X_0(n), X_1(n), \dots)$ ,  $(Y_0(n), Y_1(n), \dots)$ , and  $(M_{10}(n), M_{01}(n), \dots)$ , and by analyzing the limiting partial differential equations. Samorodnitsky *et al.* [8] took this one step further and derived an exact integral characterization of the joint generating function  $\varphi(x, y)$  of  $\{f_{ij}\}$ . Using this characterization they were also able to prove that the joint distribution  $\{f_{ij}\}$  has jointly regularly varying tails with a specified tail measure.

In the current paper we analyze the same DPA model but using a different method. Instead we analyze the evolution of the in- and out-degree of *one* randomly selected node (born before  $n$ ) up until time step  $n$ . If we let  $p_{ij}^{(n)}$  denote the probability that the degree of this node equals  $(i, j)$  at time step  $n$ , it follows that  $\lim_n p_{ij}^{(n)} = cf_{ij}$  (the constant  $c$  is only there because the  $f_{ij}$  of [2] will not sum to unity). Using this alternative method we show, by means of weak convergence and a certain time transformation, that the evolution of degrees of a randomly selected node, in the limit, converges to the evolution of two independently evolving Markov birth processes, which are both stopped at a common exponentially distributed time  $T$ . We now define this limiting process.

Consider a bivariate birth process  $X(t) = (X_I(t), X_O(t))$  with time-homogeneous birth rates

$$P(X(t+h) = (i+1, j) \mid X(t) = (i, j)) = \lambda_{ij}^I h + o(h),$$

$$P(X(t+h) = (i, j+1) \mid X(t) = (i, j)) = \lambda_{ij}^O h + o(h),$$

$$P(X(t+h) = (i, j) \mid X(t) = (i, j)) = 1 - \lambda_{ij}^I h - \lambda_{ij}^O h + o(h),$$

where the jump intensities are given by

$$\lambda_{ij}^I = \lambda_i^I = (i + \delta_I)c_I, \tag{3}$$

$$\lambda_{ij}^O = \lambda_j^O = (j + \delta_O)c_O, \tag{4}$$

with  $c_I$  and  $c_O$  defined in (1) and (2).

Because the two jump rates depend only on the first and second coordinate respectively, the two coordinate processes, reflecting in- and out-degree, evolve *independently*. Assume that the process is started either in state (0, 1) or (1, 0):

$$P(X(0) = (0, 1)) = \frac{\alpha}{\alpha + \gamma}, \text{ and}$$

$$P(X(0) = (1, 0)) = \frac{\gamma}{\alpha + \gamma}.$$

In Section 2 we make use of this process and show that the limiting-degree distribution of the DPA model is identical to that of this process observed after an exponentially distributed time, which enables the computation of its bivariate tail distribution. Another advantage with this new characterization is that it easily extends to related, but more realistic, preferential attachment models, including the one defined below.

Recently, and independently, Wang and Resnick [10] studied the special case of a DPA model where  $\beta = 0$  so that a node is always added. The focus of their paper is slightly different from that of the current paper. Wang and Resnick studied convergence properties of empirical tail degrees, estimates of marginal tail power laws using the Hill estimator, and the degree growth rate of a specific node as the network increases. Here we focus on deriving explicit expressions for the limiting bivariate degree distribution and its tail, and extending the model (see below).

The DPA model of [2] has two main features which may be criticized from the point of realism. The first is that the two different degrees of a node evolve independently in the sense that the rate/probability of acquiring an additional ingoing edge (thus increasing the in-degree by 1) depends on the current in-degree of the node, but not on its out-degree, and vice versa. A more realistic model would in many situations be to let the probability that an added directed edge points to a node having current in- and out-degree  $(i, j)$  be proportional to  $i + cj + \delta_I$ , and similarly that the probability that a new edge points out from this node to be proportional to  $di + j + \delta_O$  (where  $c$  and  $d$  are non-negative model parameters; in the original DPA model  $c = d = 0$ ). This will allow for a stronger dependence between in- and out-degree.

A second, perhaps more important, feature that the original DPA model can be criticized for is that the fraction of edges that are ‘double directed’ (or equivalently undirected), i.e. pairs of nodes for which there exist directed edges going both ways between them, will be negligible. In many empirical networks having directed edges, the fraction  $\rho$  of directed edges for which the reciprocal edge is also present is far away from 0, typically in the interval (0.2, 0.8) (cf. [7], and Spricer and Britton [9] who considered another model for a partially directed random network). This can be achieved if we modify the DPA model by simply stating that, at each time step (when a directed edge is added), the corresponding reciprocal edge is added with probability  $\rho$ .

From the reasoning above we now define what we call the generalized directed preferential attachment (GDPA) model.

**Definition 1.** (*The generalized directed preferential attachment (GDPA) model.*) The process is started at  $k = 0$  with a single node without any edges. At each discrete time step  $k = 1, 2, \dots$ ,

one of three different events can happen: (1) with probability  $\alpha$  a new node with an edge pointing out from this node is added, (2) with probability  $\beta$  a new directed edge without nodes is added, and (3) with the remaining probability  $\gamma = 1 - \alpha - \beta$  a new node with a directed edge pointing at the new node is added. In the first and second cases, the probability that the new edge points to a specific existing node with in- and out-degree  $(i, j)$  is proportional to  $i + cj + \delta_I$ . In the second and third cases, the probability that the new edge points out from a specific node with in- and out-degree  $(i, j)$  is proportional to  $di + j + \delta_O$ . Finally, the added directed edge is made reciprocal (i.e. double-directed) with probability  $\rho$ , independently in each time step.

The GDPA model has seven parameters:  $\alpha, \beta, \delta_I, \delta_O, c, d$ , and  $\rho$  (recall that  $\gamma = 1 - \alpha - \beta$  and hence is not a free parameter). Note that in the DPA model it was crucial that  $\delta_I > 0$  and  $\delta_O > 0$ , since otherwise a node would only have positive in- or out-degree. For the GDPA model this is no longer necessary when  $c > 0$  and  $d > 0$ . For this reason it is possible to reduce the number of parameters to five by assuming  $\delta_I = \delta_O = 0$ . Further, the GDPA model is identical to the DPA model when instead  $c = d = \rho = 0$ .

## 2. Main results

We now state our main result characterizing the limiting-degree distribution of the directed preferential attachment model in terms of our simple two-dimensional birth process  $X(\cdot)$  (defined in the previous section) evaluated after an exponentially distributed time. Recall that we assume  $\alpha, \gamma, \delta_I$ , and  $\delta_O$  to all be strictly positive (the interesting case) to avoid special cases.

**Theorem 2.1.** *Let  $p_{ij}^{(n)}$  denote the probability that a randomly selected node in the DPA model after  $n$  steps has in- and out-degree  $i$  and  $j$  respectively. Let  $X(t)$  denote the bivariate birth process defined above, and  $T$  an independent  $\text{Exp}(1)$  random variable. Then*

$$p_{ij}^{(n)} \rightarrow p_{ij} := \mathbb{P}(X(T) = (i, j)).$$

**Remark.** Let  $N(n)$  denote the number of nodes after  $n$  steps in the DPA model (which follows a  $1 + \text{Bin}(n, \alpha + \gamma)$  distribution). It then directly follows that  $M_{ij}(n)/N(n)$ , the fraction of nodes having in- and out-degree  $(i, j)$ , converges in probability to  $p_{ij}$ .

The explicit distribution of  $X(T)$  can be derived in two ways. The first is by studying the embedded discrete-time random walk which, if currently in state  $(k, \ell)$ , goes to state  $k + 1, \ell$  with probability  $\lambda_k^I / (\lambda_k^I + \lambda_\ell^O + 1)$ , to state  $k, \ell + 1$  with probability  $\lambda_\ell^O / (\lambda_k^I + \lambda_\ell^O + 1)$ , or gets stuck for ever in state  $(k, \ell)$  with the remaining probability  $1 / (\lambda_k^I + \lambda_\ell^O + 1)$ . These jump probabilities follow directly from the Markov property: the jump probabilities equal the rate of jumping to the considered state divided by the overall jump rate. The probability  $\mathbb{P}(X(T) = (i, j))$  is then the probability that this random walk gets stuck in state  $(i, j)$ . Note that there are many different paths, in general having different probabilities, ending in state  $(i, j)$ . This derivation technique will be applied when analyzing the GDPA model.

The other way to derive the distribution is by integrating over possible values of  $T$ , and using that the two components of the continuous-time bivariate processes evolve independently. Let  $X^{(0,1)}(t) = (X_1^0(t), X_0^1(t))$  and  $X^{(1,0)}(t) = (X_1^1(t), X_0^0(t))$  denote the bivariate birth process described above, but where we condition on the starting state being  $(0, 1)$  and  $(1, 0)$

respectively (remember that the probabilities for these two starting points are  $\alpha/(\alpha + \gamma)$  and  $\gamma/(\alpha + \gamma)$  respectively). We then have

$$\begin{aligned}
 p_{ij} = & \frac{\alpha}{\alpha + \gamma} \int_0^\infty P(X_1^0(t) = i)P(X_0^1(t) = j)e^{-t} dt \\
 & + \frac{\gamma}{\alpha + \gamma} \int_0^\infty P(X_1^1(t) = i)P(X_0^0(t) = j)e^{-t} dt, \tag{5}
 \end{aligned}$$

and the following result gives explicit expressions for the bivariate degree distribution (using the notation  $a_k = a(a - 1) \cdots (a - k + 1)$ ).

**Theorem 2.2.** *The marginal distributions of the two birth processes conditioned on their starting values  $r$  are given by*

$$\begin{aligned}
 P(X_1^r(t) = i) &= \frac{(\delta_I + i - 1)_{i-r}}{(i - r)!} e^{-c_1(\delta_I+r)t} (1 - e^{-c_1t})^{i-r}, \quad i = r, r + 1, \dots, \\
 P(X_0^r(t) = j) &= \frac{(\delta_O + j - 1)_{j-r}}{(j - r)!} e^{-c_0(\delta_O+r)t} (1 - e^{-c_0t})^{j-r}, \quad j = r, r + 1, \dots
 \end{aligned}$$

The marginal distributions of the stopped birth processes conditioned on their starting value  $r$  are given by

$$\begin{aligned}
 P(X_1^r(T) = i) &= \frac{(\delta_I + i - 1)_{i-r}}{(\delta_I + \frac{1}{c_1} + i)_{i-r+1}}, \\
 P(X_0^r(T) = j) &= \frac{(\delta_O + j - 1)_{j-r}}{(\delta_O + \frac{1}{c_0} + j)_{j-r+1}}.
 \end{aligned}$$

The joint degree distribution  $p_{ij} = P(X(T) = (i, j))$  is given by

$$\begin{aligned}
 p_{ij} = & \frac{\alpha}{\alpha + \gamma} \frac{(\delta_I + i - 1)_i(\delta_O + j - 1)_{j-1}}{i!(j - 1)!} \sum_{k=0}^i \sum_{\ell=0}^{j-1} (-1)^{k+\ell} \frac{\binom{i}{k} \binom{j-1}{\ell}}{c_1(\delta_I + k) + c_0(\delta_O + \ell) + 1} \\
 & + \frac{\gamma}{\alpha + \gamma} \frac{(\delta_I + i - 1)_{i-1}(\delta_O + j - 1)_j}{(i - 1)!j!} \sum_{k=0}^{i-1} \sum_{\ell=0}^j (-1)^{k+\ell} \frac{\binom{i-1}{k} \binom{j}{\ell}}{c_1(\delta_I + k) + c_0(\delta_O + \ell) + 1}.
 \end{aligned}$$

Besides having an explicit (albeit long) form for the limiting-degree distribution  $\{p_{ij}\}$ , its characterization as a stopped bivariate birth process making independent jumps component-wise gives hope for a richer analysis of the tail behavior of the two degrees. For example, it is not hard to confirm the earlier stated results of [2].

**Corollary 2.1.** (Bollobás et al. [2]) *For large in-degrees,  $p_i := \sum_j p_{ij} \sim i^{-(1+1/c_1)}$ ; for large out-degrees,  $q_j := \sum_i p_{ij} \sim j^{-(1+1/c_0)}$ . For fixed  $j$  and large  $i$ ,  $p_{ij} \sim i^{-x_1}$ ; for fixed  $i$  and large  $j$ ,  $p_{ij} \sim j^{-x_0}$ ; where  $x_1 = 1 + 1/c_1 + c_0\delta_O/c_1$  and  $x_0 = 1 + 1/c_0 + c_1\delta_I/c_0$ .*

**Remark.** The original theorem of Bollobás *et al.* contains an additional term for  $x_1$  and  $x_0$ , but these vanish when restricting the parameters  $\delta_I$ ,  $\delta_O$ ,  $\alpha$ , and  $\gamma$  to all being strictly positive, as we have done (the interesting case).

Our new result concerns the tail probability  $p_{ij}$  when both  $i \rightarrow \infty$  and  $j \rightarrow \infty$ . More specifically, let  $i = n$  and  $j = \lfloor sn^r \rfloor$  for some  $0 < s, r < \infty$ , where  $\lfloor x \rfloor$  is the integer part of  $x$ . We have the following theorem.

**Theorem 2.3.** *The tail probabilities  $p_{ij}$ , for  $i = n$  and  $j = \lfloor sn^r \rfloor$ , satisfy*

$$p_{n, \lfloor sn^r \rfloor} \sim L_n n^{\delta_I - 1 + r(\delta_O - 1) - (c_I \delta_I + \delta_O c_O + 1) \max(1/c_I, r/c_O)}$$

$$= \begin{cases} L_n n^{-(1 + \delta_O(c_O/c_I - r) + r + 1/c_I)} & \text{if } r \leq c_O/c_I, \\ L_n n^{-(1 + \delta_I(r c_I/c_O - 1) + r + r/c_O)} & \text{if } r \geq c_O/c_I, \end{cases}$$

where  $L_n \sim n^{\varepsilon_n}$  ( $\varepsilon_n \rightarrow 0$ ) is a slowly varying function. The factor  $s$  hence has no effect on the tail exponents. The choice of  $r$  which maximizes the tail probability is given by

$$\arg \max_{\{r \geq 0\}} p_{n, \lfloor sn^r \rfloor} \rightarrow \begin{cases} 0 & \text{if } \delta_O < 1, \\ c_O/c_I & \text{if } \delta_O > 1. \end{cases}$$

If  $\delta_O = 1$  any value of  $r$  between 0 and  $c_O/c_I$  gives the same asymptotic tail.

The limiting-degree distribution of the original (undirected) PA model of [1] can also be derived using a similar methodology. Here, the limiting process is even simpler: a one-dimensional pure birth process  $Z(t)$  starting at  $Z(0) = m$  and having birth rate  $\lambda_i = i/2$  that is stopped at  $T \sim \text{Exp}(1)$ . The limiting distribution has previously been derived using other methods.

**Theorem 2.4.** (Dorogovtsev *et al.* [4].) *Let  $p_i^{(n)}$  denote the probability that the degree of a randomly selected individual after  $n$  steps has degree  $i$  in the (undirected) PA model, and let  $Z(t)$  and  $T$  be defined as above. Then*

$$p_i^{(n)} \rightarrow p_i := P(Z(T) = i) = \frac{2m(m + 1)}{i(i + 1)(i + 2)}, \quad i = m, m + 1, \dots$$

We now consider the generalized directed preferential attachment (GDPA) model. Using very similar methods as when proving Theorem 2.1, it can be shown that for the GDPA model the limiting degree distribution can also be characterized by a continuous-time bivariate ‘birth process’ which is stopped and observed after an exponentially distributed random time  $T$ . We now describe the limiting process to which the degree distribution of the GDPA model converges. Let  $Y(t) = (Y_1(t), Y_0(t))$  be a time-homogeneous bivariate Markov birth process, but where now simultaneous births of the two components are also possible. The three different birth rates  $\beta_{ij}^I$ ,  $\beta_{ij}^O$ , and  $\beta_{ij}^{I+O}$  are defined by

$$\begin{aligned} \beta_{ij}^I &= (i + cj + \delta_I)(1 - \rho) \frac{\alpha + \beta}{(1 + \rho)(1 + c) + \delta_I(\alpha + \gamma)} \\ &=: (i + cj + \delta_I)(1 - \rho)g_I, \\ \beta_{ij}^O &= (di + j + \delta_O)(1 - \rho) \frac{\gamma + \beta}{(1 + \rho)(1 + d) + \delta_O(\alpha + \gamma)} \\ &=: (di + j + \delta_O)(1 - \rho)g_O, \\ \beta_{ij}^{I+O} &= (i + cj + \delta_I)\rho g_I + (di + j + \delta_O)\rho g_O, \end{aligned} \tag{6}$$

$g_I$  and  $g_O$  being the respective ratio expressions. The third rate means that  $P(Y(t + h) = (i + 1, j + 1) \mid Y(t) = (i, j)) = \beta_{ij}^{I+O} h + o(h)$ . The process is started in one of the states  $(0, 1)$ ,  $(1, 0)$ , or  $(1, 1)$  with respective probability  $(1 - \rho)\alpha/(\alpha + \gamma)$ ,  $(1 - \rho)\gamma/(\alpha + \gamma)$ ,  $\rho$ . As before,

let  $T \sim \text{Exp}(1)$  be an independent exponentially distributed time. We then have the following theorem.

**Theorem 2.5.** *Let  $\pi_{ij}^{(n)}$  denote the probability that a randomly selected node in the GDPA model after  $n$  steps has in- and out-degree  $i$  and  $j$  respectively. Let  $Y(t)$  denote the bivariate birth process defined above, and  $T$  an independent  $\text{Exp}(1)$  random variable. Then*

$$\pi_{ij}^{(n)} \rightarrow \pi_{ij} := \text{P}(Y(T) = (i, j)).$$

**Remark.** Just as in Theorem 2.1, it follows that the fraction of nodes having in- and out-degree  $(i, j)$  converges in probability to  $\pi_{ij}$ .

The limiting probabilities  $\{\pi_{ij}\}$  can be computed by summing over all paths, starting from one of  $(0, 1)$ ,  $(1, 0)$ , or  $(1, 1)$ , and ending in  $(i, j)$ , where each jump either increases the in-degree by 1, the out-degree by 1, both degrees by 1, or makes a complete stop. If currently in state  $(k, \ell)$ , the process jumps to  $(k + 1, \ell)$  with probability  $\beta_{k,\ell}^I/\sigma_{k,\ell}$ , to state  $(k, \ell + 1)$  with probability  $\beta_{k,\ell}^O/\sigma_{k,\ell}$ , to state  $(k + 1, \ell + 1)$  with probability  $\beta_{k,\ell}^{I+O}/\sigma_{k,\ell}$ , or makes a complete final stop with probability  $1/\sigma_{k,\ell}$ , where  $\sigma_{k,\ell} = \beta_{k,\ell}^I + \beta_{k,\ell}^O + \beta_{k,\ell}^{I+O} + 1$  (the different  $\beta$ s were defined in (6)).

As an illustration,

$$\begin{aligned} \pi_{11} = & \frac{(1 - \rho)\alpha}{\alpha + \gamma} \times \frac{\beta_{0,1}^I}{\beta_{0,1}^I + \beta_{0,1}^O + \beta_{0,1}^{I+O} + 1} \times \frac{1}{\beta_{1,1}^I + \beta_{1,1}^O + \beta_{1,1}^{I+O} + 1} \\ & + \frac{(1 - \rho)\gamma}{\alpha + \gamma} \times \frac{\beta_{1,0}^O}{\beta_{1,0}^I + \beta_{1,0}^O + \beta_{1,0}^{I+O} + 1} \times \frac{1}{\beta_{1,1}^I + \beta_{1,1}^O + \beta_{1,1}^{I+O} + 1} \\ & + \rho \times \frac{1}{\beta_{1,1}^I + \beta_{1,1}^O + \beta_{1,1}^{I+O} + 1}. \end{aligned}$$

The first factor in each row is the probability of starting in  $(0, 1)$ ,  $(1, 0)$ , and  $(1, 1)$  respectively. For higher degrees there will be many more paths to sum over. Starting at  $(0, 1)$  and ending in  $(1, 2)$  can, for example, happen in three different ways, either first jumping to  $(1, 1)$  followed by a jump to  $(1, 2)$ , or first jumping to  $(0, 2)$  and then to  $(1, 2)$ , or jumping directly to  $(1, 2)$ .

As for the tail probabilities of the GDPA model  $\{\pi_{ij}\}$ , it should be possible to derive them using a similar analysis as for the tails of the DPA model (Corollary 2.1 and Theorem 2.3). However, the fact that the two components no longer evolve independently makes the analysis more involved, and its tail behavior remains an open problem.

### 3. Proofs

#### 3.1. Proof of Theorem 2.1

The proof consists of two parts. First we show that the evolution of the degrees of a randomly selected node up to step  $n$  converges to the evolution of the degrees of a continuous-time bivariate birth process  $D = (D_I, D_O)$ , born at  $U \sim U[0, 1]$ , having *time-inhomogeneous* birth rates  $\lambda_{ij}^I/t$  and  $\lambda_{ij}^O/t$  respectively, and observed at time  $t = 1$ . Then, using a time transformation, we show that this limiting distribution has the same distribution as the *time-homogeneous* bivariate process  $X = (X_I, X_O)$  of the theorem, having birth rates  $\lambda_{ij}^I$  and  $\lambda_{ij}^O$  respectively, born at time 0 and observed at time  $T \sim \text{Exp}(1)$ .

Recall that  $p_{ij}^{(n)}$  is the probability that a randomly chosen node after  $n$  time steps has in- and out-degree  $(i, j)$ . At the start there are no edges, and at each time step one edge is added, so there are  $n$  edges at this time. Further, at the start there is one node, and at each time step a new node is added with probability  $\alpha + \gamma$ , so the number of nodes at time step  $n$ , denoted  $N(n)$ , is  $1 + \text{Bin}(n, \alpha + \gamma)$ . For large  $n$  this is well approximated by  $(\alpha + \gamma)n$ , which is done below.

Fix  $s, 0 \leq s \leq 1$ , and consider a node that entered the network at time step  $\lfloor sn \rfloor$  (the integer part of  $sn$ ). For  $s \leq t \leq 1$  we define the bivariate process  $(D_I^{(n)}(t), D_O^{(n)}(t))$  as the in- and out-degree of that node at time step  $\lfloor tn \rfloor$ . The starting value of our node may either be  $(0, 1)$  or  $(1, 0)$  depending on whether it entered through the event 1 or event 3 (if event 2 happens, no new node is added). The probabilities are hence given by

$$P(D_I^{(n)}(s) = 0, D_O^{(n)}(s) = 1) = \frac{\alpha}{\alpha + \gamma} = 1 - P(D_I^{(n)}(s) = 1, D_O^{(n)}(s) = 0).$$

Assume that at time  $t$  our process equals  $(D_I^{(n)}(t), D_O^{(n)}(t)) = (i, j)$ , and let  $\Delta t = 1/n$ . We first compute the probability/intensity that our process will increase its in-degree by 1. This happens if either event 1 happens and our node is selected as the node to point at, or that event 2 happens and our node is selected as the node to point at. The probability for this is hence

$$\begin{aligned} P((D_I^{(n)}(t + \Delta t), D_O^{(n)}(t + \Delta t)) = (i + 1, j) \mid (D_I^{(n)}(t), D_O^{(n)}(t)) = (i, j)) \\ = (\alpha + \beta) \frac{i + \delta_I}{\lfloor tn \rfloor + \delta_I N(\lfloor tn \rfloor)} + o_p(1/n) \\ = (\alpha + \beta) \frac{i + \delta_I}{t(1 + \delta_I(\alpha + \gamma))} \Delta t + o_p(\Delta t), \end{aligned}$$

where the terms  $o_p(\cdot)$  tend to 0 in probability as the argument tends to 0. For the out-degree we obtain the similar result

$$\begin{aligned} P((D_I^{(n)}(t + \Delta t), D_O^{(n)}(t + \Delta t)) = (i, j + 1) \mid (D_I^{(n)}(t), D_O^{(n)}(t)) = (i, j)) \\ = (\beta + \gamma) \frac{j + \delta_O}{\lfloor tn \rfloor + \delta_O N(\lfloor tn \rfloor)} + o_p(1/n) \\ = (\beta + \gamma) \frac{j + \delta_O}{t(1 + \delta_O(\alpha + \gamma))} \Delta t + o_p(\Delta t). \end{aligned}$$

It can in fact also happen that our node increases both its in- and out-degree. This happens in the case that event 2 happens and our node is chosen both for the start and the end of the edge (thus creating a loop). This event should in fact also be excluded from the above probabilities. However, this happens with a probability proportional to  $1/n^2 = (\Delta t)^2$ , so these events will not occur in the limit.

We have thus shown that the jump rates of  $(D_I^{(n)}, D_O^{(n)})$  converge to those of  $(D_I, D_O)$ . For a more general Markov jump process (e.g. taking any values in  $\mathbb{R}^2$ ) this corresponds to the generator converging to the generator of the limiting process. From this result we can use theory for density-dependent jump Markov processes to conclude that  $(D_I^{(n)}, D_O^{(n)})$  converges in the limit as  $n \rightarrow \infty$  to the continuous-time stochastic process  $(D_I, D_O)$  on  $[s, 1]$  (see [5, Theorem 4.2.5, p. 167], and also Exercise 4.8 on p. 262 therein, which considers the current situation except for being one-dimensional). The limiting process only makes increases by one unit, and never in both coordinates at the same time: it is a bivariate time-inhomogeneous Markov birth process. An important observation is that the birth rate for the in-degree only



depends on the current value of the in-degree  $i$  and not on the current out-degree  $j$ , and similarly for the out-degree, which means that the two degrees evolve independently. We can hence drop this dependence in our notation and write  $\lambda_i^I$  and  $\lambda_j^O$  as defined in (3) and (4). It hence follows that our process  $(D_I^{(n)}(t), D_O^{(n)}(t))$  converges to a continuous-time, time-inhomogeneous bivariate Markov birth process, with birth intensities given by  $\lambda_i^I/t$  and  $\lambda_j^O/t$  respectively. The process starts at time  $s$  with degrees  $(0, 1)$  or  $(1, 0)$ , with respective probabilities  $\alpha/(\alpha + \gamma)$  and  $\gamma/(\alpha + \gamma)$ .

Finally, our randomly chosen node has the same probability  $\alpha + \gamma$  of being born at any time point between 1 and  $n$ , implying that the birth time  $s$  of the limiting process is  $U[0, 1]$ . We have thus shown that the distribution  $\{p_{ij}^{(n)}\}$  of the in- and out-degree of a randomly selected node after step  $n$  in the GPA model converges (as  $n \rightarrow \infty$ ) to the distribution of  $(D_I(1), D_O(1))$ , where this process is started at a uniformly distributed time  $s$  with starting configuration as stated above, and where the two degrees evolve according to the rates given above until time  $t = 1$ .

The second part of the proof consists of showing that this distribution equals the one stated in the theorem. We do this by looking at the accumulated intensities. Since the jump rates of  $(D_I, D_O)$  are of the form  $\lambda_{ij} \times 1/t$ , and the jump rates of  $(X_I, X_O)$  equal  $\lambda_{ij}$  for the same  $\lambda_{ij}$ , it suffices to show that the accumulated rates agree for a fixed parameter  $\lambda$ , say.

The accumulated rate for the  $D$  process, starting at time  $s \sim U[0, 1]$  and evolving until time 1, equals

$$\int_0^1 \left( \int_s^1 \lambda \frac{1}{t} dt \right) ds = \lambda.$$

The accumulated rate for the  $X$  process starting at time 0 and evolving up until time  $T \sim \text{Exp}(1)$  equals

$$\int_0^\infty \left( \int_0^t \lambda ds \right) e^{-t} dt = \lambda.$$

The two processes hence have the same accumulated jump rates and start from the same initial condition, which implies that they have the same distribution at the observation points. Another way to obtain this result is to make a time transformation of  $D$ : first we reverse time and let it evolve from 0 to  $s \sim U[0, 1]$  with rate  $\lambda/(1 - s)$ , and then transform time:  $s' = \log(1 - s)$ . These two changes combined lead to the rates and limits of the  $X$  process. This completes the proof.

### 3.2. Proof of Theorem 2.2

For the first part of the theorem we show the result for  $X_I^r$ , the proof for  $X_O^r$  being identical. We only need the result for  $r = 0$  and  $r = 1$ , but the result holds for any  $r$ . Fix  $r$ . We use induction. The jump rate of  $X_I$  if currently in state  $i$  equals  $\lambda_i^I = (i + \delta_I)c_I$ . We start with  $i = r$ , meaning that there must be no jump between 0 and  $t$ :

$$P(X_I^r(t) = r) = e^{-\lambda_i^I t} = e^{-(r + \delta_I)c_I t},$$

which agrees with the theorem. We now assume the expression for  $P(X_I^r(t) = k)$  is as stated in the theorem for  $k = 1 \dots, i - 1$  for all  $t$ . Then

$$\begin{aligned} P(X_I^r(t) = i) &= \int_0^t P(X_I^r(s) = i - 1) \lambda_{i-1}^I e^{-\lambda_i^I(t-s)} ds \\ &= \frac{(\delta_I + i - 1)_{i-r}}{(i - 1 - r)!} c_I e^{-(\delta_I + i)c_I t} \int_0^t e^{(i-r)c_I s} (1 - e^{-c_I s})^{i-1-r} ds \end{aligned}$$

$$\begin{aligned}
 &= \frac{(\delta_I + i - 1)_{i-r}}{(i - 1 - r)!} c_I e^{-(\delta_I + i)c_I t} \sum_{j=0}^{i-1-r} \binom{i-1-r}{j} \frac{e^{(i-r-j)c_I t} - 1}{c_I(i-r-j)} (-1)^j \\
 &= \frac{(\delta_I + i - 1)_{i-r}}{(i - 1 - r)!} \frac{e^{-(\delta_I + i)c_I t}}{i - r} \left( \sum_{j=0}^{i-1-r} \binom{i-1-r}{j} e^{(i-r-j)c_I t} (-1)^j + (-1)^{i-r} \right) \\
 &= \frac{(\delta_I + i - 1)_{i-r}}{(i - 1 - r)!} \frac{e^{-(\delta_I + i)c_I t}}{i - r} (e^{c_I t} - 1)^{i-r} \\
 &= \frac{(\delta_I + i - 1)_{i-r}}{(i - r)!} e^{-(\delta_I + r)c_I t} (1 - e^{-c_I t})^{i-r},
 \end{aligned}$$

which proves the first part of the theorem. The third equality is obtained by expanding  $(1 - e^{-c_I t})^{i-1-r}$  in its binomial terms and integrating each term, and the fourth equality comes from summing the ‘-1’ terms.

Now to the second part of the theorem. As before, we only show it for  $X_I^r(T)$  (the proof for  $X_O^r(T)$  being identical). If currently in state  $k$ , the rate at which  $X_I$  gives birth equals  $\lambda_k^I = (\delta_I + k)c_I$  and the rate at which the process stops equals 1 ( $T \sim \text{Exp}(1)$ ). The probability for a birth before a stop is hence  $\lambda_k^I / (\lambda_k^I + 1)$ . Since the process is Markovian we hence have

$$P(X_I^r(T) = i) = \left( \prod_{k=r}^{i-1} \frac{\lambda_k^I}{\lambda_k^I + 1} \right) \frac{1}{\lambda_i^I + 1},$$

where the last factor comes from requiring that the process stops when in state  $k$ . Simple manipulation of this expression, using that  $\lambda_k^I = (\delta_I + k)c_I$ , gives the desired result.

Now to the final part of the theorem. We know that the starting configuration of  $X$  is either  $(0, 1)$  or  $(1, 0)$  with probabilities  $\alpha / (\alpha + \gamma)$  and  $\gamma / (\alpha + \gamma)$ , so by conditioning on the starting configuration and the stopping time  $T \sim \text{Exp}(1)$ , and using that  $X_I(t)$  and  $X_O(t)$  are independent given the starting configuration, we get

$$\begin{aligned}
 p_{ij} = P(X(T) = (i, j)) &= \frac{\alpha}{\alpha + \gamma} \int_0^\infty P(X_I^0(t) = i) P(X_O^1(t) = j) e^{-t} dt \\
 &\quad + \frac{\gamma}{\alpha + \gamma} \int_0^\infty P(X_I^1(t) = i) P(X_O^0(t) = j) e^{-t} dt.
 \end{aligned}$$

From the first part of the corollary it follows that the first integral above equals

$$\frac{(\delta_I + i - 1)_i (\delta_O + j - 1)_{j-1}}{i!(j - 1)!} \int_0^\infty e^{-\delta_I c_I t} (1 - e^{-c_I t})^i e^{-(\delta_O + 1)c_O t} (1 - e^{-c_O t})^{j-1} dt.$$

By expanding both  $(1 - e^{-c_I t})^i$  and  $(1 - e^{-c_O t})^{j-1}$ , and integrating term by term, the first sum of the theorem is obtained. The second term is treated identically.

### 3.3. Proof of Corollary 2.1

We have that  $p_i = P(X_I(T) = i) = \alpha / (\alpha + \gamma) P(X_I^0(T) = i) + \gamma / (\alpha + \gamma) P(X_I^1(T) = i)$ , and these probabilities are given in Theorem 2.2:

$$P(X_I^r(T) = i) = \frac{1}{(\delta_I + \frac{1}{c_I} + i)} \frac{(\delta_I + i - 1)_{i-r}}{(\delta_I + \frac{1}{c_I} + i - 1)_{i-r}}.$$

It is easy to show that, as  $i \rightarrow \infty$ ,  $(a + i)_{i-r}/(b + i)_{i-r} \sim i^{a-b}$ . This implies that

$$P(X_1^r(T) = i) \sim i^{-(1+1/c_1)},$$

and since the same expression holds for  $r = 0$  and  $r = 1$  it also holds for  $p_i$ . The proof for  $q_j$  is obtained similarly. We now fix  $j$  and study  $p_{ij} = P(X(T) = (i, j))$  for large  $i$ . We start by considering  $j = 0$ , and hence look at  $p_{i0}$  for large  $i$ . For  $j = 0$ , the process must start in state  $(1, 0)$  and the out-degree birth process must never have a birth, so

$$\begin{aligned} p_{i0} &= \frac{\gamma}{\alpha + \gamma} \left( \prod_{k=1}^{i-1} \frac{\lambda_k^I}{\lambda_k^I + \lambda_0^O + 1} \right) \frac{1}{\lambda_i^I + \lambda_0^O + 1} \\ &= \frac{(\delta_I + i - 1)_{i-1}}{(\delta_I + \frac{1}{c_1} + \frac{c_O}{c_1} \delta_O + i - 1)_{i-1}} \sim i^{-(1+\frac{1}{c_1} + \frac{c_O}{c_1} \delta_O)}, \end{aligned}$$

where the second equality is obtained by plugging in the expressions for  $\lambda_k^I$  and  $\lambda_0^O$ , and the asymptotic size follows from the above stated property of  $(a + i)_{i-r}/(b + i)_{i-r}$ .

For another fixed  $j \geq 1$  we now show that it is of the same order. In order to reach  $(i, j)$  where  $i$  is large and  $j$  is fixed (and hence small in comparison with  $i$ ) the  $j$  births of the out-degree process can occur for different values of the in-degree process. However, since the birth rate for the in-degree process increases with  $i$ , the probability that any out-degree birth takes place when the in-degree is greater than  $n$  is  $o(1/n)$ . So, the more likely paths ending in  $(i, j)$  (where  $j$  is fixed and small and  $i$  is large) are where the out-degree births take place when the in-degree is small. We now compute the probability for one such path, and since there are only a fixed number of such paths, and all these paths have probability of the same order, we conclude that the overall probability  $p_{ij}$  is of the same order as the probability of one such (likely) path.

We assume that the initial configuration is  $(0, 1)$  (the other case is done equivalently). We now compute the probability of the path, namely first having all the  $j$  out-degree births, followed by all the in-degree births:

$$\begin{aligned} P((0, 1) \rightarrow (0, 2) \cdots (0, j) \rightarrow (1, j) \rightarrow (2, j) \rightarrow \cdots (i, j)) \\ = \prod_{k=1}^{j-1} \frac{\lambda_k^O}{\lambda_0^I + \lambda_k^O + 1} \prod_{k=0}^{i-1} \frac{\lambda_k^I}{\lambda_k^I + \lambda_j^O + 1} \frac{1}{\lambda_i^I + \lambda_j^O + 1}. \end{aligned}$$

For fixed  $j$ , the first product is a constant, and the second product and the last factor are of order  $i^{-(1+\frac{1}{c_1} + \frac{c_O}{c_1} \delta_O)}$ , which is shown in the same way as was done for  $p_{i0}$ . So, since  $j$  is fixed, together with the fact that  $j$  births of the out-degree process happen when the in-degree is small with large probability, proves the last statement of the theorem.

### 3.4. Proof of Theorem 2.3

The exact expression for  $p_{ij}$  is given in (5). It is easy to show that for large  $i = n$ ,  $P(X_1^1(t) = n) \sim P(X_1^0(t) = n)$ , and for  $j = \lfloor sn^r \rfloor$ ,  $P(X_0^1(t) = \lfloor sn^r \rfloor) \sim P(X_0^0(t) = \lfloor sn^r \rfloor)$ , which implies that starting in either state  $(0, 1)$  or  $(1, 0)$  has no effect on the tail behavior. Further, it is known and easily proven that  $(a + n)_n/n! \sim n^a$ . From these two observations, together with the exact expression (5), we have

$$p_{n, \lfloor sn^r \rfloor} \sim n^{\delta_I - 1} (sn^r)^{\delta_O - 1} \int_0^\infty e^{-(c_1 \delta_I + c_O \delta_O + 1)t} (1 - e^{-c_1 t})^n (1 - e^{-c_O t})^{sn^r} dt. \tag{7}$$

We now analyze the integral in (7), writing  $b = c_1\delta_1 + c_0\delta_0 + 1$  and using Laplace’s method. The integral in (7) can be written as

$$\int_0^\infty e^{-nf_n(t)} dt, \quad \text{where } f_n(t) = (b/n)t - \log(1 - e^{-c_1t}) - sn^{r-1} \log(1 - e^{-c_0t}).$$

Clearly,  $f_n(t) > 0$  for  $t > 0$ , and as  $t \rightarrow 0$  or  $t \rightarrow \infty$ ,  $f_n(t) \rightarrow +\infty$ . From this it follows that the main contribution to the integral comes for  $t$  values close to  $t_n^{(\min)}$  for which  $f_n(t)$  is minimized. We hence Taylor expand  $f_n(t)$  around  $t_n^{(\min)}$ :

$$\begin{aligned} f_n(t) &= f_n(t_n^{(\min)}) + (t - t_n^{(\min)})f'_n(t_n^{(\min)}) + \frac{(t - t_n^{(\min)})^2}{2}f''_n(t_n^{(\min)}) + o((t - t_n^{(\min)})^2) \\ &= f_n(t_n^{(\min)}) + 0 + \frac{(t - t_n^{(\min)})^2}{2}f''_n(t_n^{(\min)}) + o((t - t_n^{(\min)})^2). \end{aligned}$$

By differentiating  $f_n(t)$ , it follows that, for large  $n$ ,  $t_n^{(\min)} = \max(1/c_1, r/c_0) \log n + o(\log n)$  and  $f'_n(t_n^{(\min)}) := f'_n(t_n^{(\min)}) = \frac{\log n}{n} b \max(1/c_1, r/c_0) + o(\log(n)/n)$ . Further,  $f''_n(t) = c_1^2 e^{-c_1t} + n^{r-1} c_0^2 e^{-c_0t}$ , and  $f''_n(t_n^{(\min)}) \sim c_1^2 n^{-c_1 \max} + c_0^2 n^{r-1-c_0 \max} \sim n^{-1}$ , where  $\max := \max(1/c_1, r/c_0)$ . Going back to the integral we hence have

$$\int_0^\infty e^{-nf_n(t)} dt \sim e^{-nf_n(t_n^{(\min)})} \int_0^\infty e^{-n(t-t_n^{(\min)})^2 f''_n(t_n^{(\min)})/2} dt \sim e^{-nf_n(t_n^{(\min)})};$$

the last step is by identifying the normal density. We have the following approximation for (7):

$$p_{n, \lfloor sn^r \rfloor} \sim L_n n^{\delta_1-1} (sn^r)^{\delta_0-1} e^{-nf_n(t_n^{(\min)})} \sim L_n n^{\delta_1-1} n^{r(\delta_0-1)} n^{-(c_1\delta_1+c_0\delta_0+1) \max(1/c_1, r/c_0)}.$$

In the last asymptotic equivalence we have dropped the constant term  $s^{\delta_0-1}$ , and replaced  $n f_n(t_n^{(\min)})$  by its expansion derived above. The remainder term becomes the factor  $L_n \sim n^{o(\log n)} = n^{\varepsilon_n}$ , with  $\varepsilon_n \rightarrow 0$  as  $n \rightarrow \infty$ , which is a slowly varying function and thus does not affect the exponents. This proves the first statement of the theorem.

As for the second statement, we have that  $p_{n, \lfloor n^r \rfloor} \sim L_n n^{\delta_1-1+r(\delta_0-1)-b \max(1/c_1, r/c_0)}$ . The  $r$  value which maximizes this is hence the same  $r$  which maximizes  $g(r) := r(\delta_0 - 1) - b \max(1/c_1, r/c_0)$ . For  $r < c_0/c_1$  we have  $g'(r) = \delta_0 - 1$ , and for  $r > c_0/c_1$ ,  $g'(r) = -1 - c_1\delta_1/c_0 - 1/c_0 < 0$ , using that  $b = c_1\delta_1 + c_0\delta_0 + 1$ . So, if  $\delta_0 < 1$  then the maximum is obtained at  $r = 0$ , and if  $\delta_0 > 1$  then  $g(r)$  is maximized for  $r = c_0/c_1$ . Finally, if  $\delta_0 = 1$ , then  $g(r)$  is constant up to  $r = c_0/c_1$ , after which it has a negative derivative, so the maximum is obtained for any  $r \in [0, c_0/c_1]$ , which completes the proof.

### 3.5. Proof of Theorem 2.4

We start by specifying the original PA model for an undirected network once again. Suppose that at time  $k = 0$  the network consists of one single node without any edge. For each  $k = 1, 2, \dots, n$  one node with  $m$  edges is added, and each of the edges is attached, independently, to a node  $u$  with degree  $i$  with probability proportional to  $i$ . We now prove the result previously derived using other methods by Dorogovtsev *et al.* [4]. After step  $n$ , let  $p_i^{(n)} = P(\text{random node has degree } i)$ . We want to prove that

$$p_i^{(n)} \rightarrow p_i := \frac{2m(m+1)}{i(i+1)(i+2)}, \quad i = m, m+1, \dots$$

As  $n$  tends to infinity we approximate the PA model by a continuous-time process by speeding up time, similarly to what was done with the DPA model. For each  $n$  we speed up time by letting new nodes enter after a time step  $1/n$  rather than after one unit of time. So, for each  $n$  we define the whole preferential attachment process up to step  $n$ :  $X(1), \dots, X(n)$  by  $X^{(n)}(t)$ ,  $0 \leq t \leq 1$ , where  $X^{(n)}(t) := X(\lfloor tn \rfloor)$  (as before,  $\lfloor tn \rfloor$  denotes the integer part of  $tn$ ).

More specifically, let's consider a node selected randomly among the  $n + 1$  nodes at time 1 (after step  $n$  in the original time scale), and analyze the distribution of its degree at time 1. This degree will depend on when it was born (= entered the network), and since one node enters at each time unit in the original time scale, the birth time is  $U[0, 1]$  on the new time scale.

First, we condition on the birth time  $s$ . Let  $D^{(n)}(t)$ ,  $s \leq t \leq 1$ , denote the degree of this node from birth until time 1 (when there are in total  $n + 1$  nodes). Since a node has degree  $m$  when it enters the network (except for the first node, which is selected with probability  $1/(n + 1)$  tending to 0) we have  $D^{(n)}(s) = m$ . We hence seek the limit  $\lim_n P(D^{(n)}(1) = i \mid D^{(n)}(s) = m) =: p_i(s)$ .

At any time step, the total number of nodes and the total number of edges is non-random. At time  $t$  ( $\lfloor tn \rfloor$  in the original time scale) there are  $\lfloor tn \rfloor + 1$  nodes and  $m \cdot \lfloor tn \rfloor$  edges. Since each edge is connected to two nodes, the sum of all degrees then equals  $2m \lfloor tn \rfloor$ .

The process  $D^{(n)}(t)$ ,  $s \leq t \leq 1$ , is a pure birth chain with possible jumps at the increments  $1/n$ , so in the limit it converges to a continuous-time time-inhomogeneous Markov birth process. Since  $m$  edges are added when a new node is added, the degree could of course increase by more than 1 at a given time instant, but the probability of increasing by more than one is  $O(1/n^2)$  and hence neglected. However, the probability of increasing by 1 is  $m$  times the probability that a specific edge connects to our node, plus terms of smaller order. We now compute the jump probability/intensity, which follows straightforwardly from the model definition. As before, let  $\Delta t = 1/n$ . We have

$$\lambda_i^{(n)}(t) = P(D^{(n)}(t + \Delta t) = i + 1 \mid D^{(n)}(t) = i) / \Delta t \approx m \frac{i}{2m \lfloor tn \rfloor} \times \frac{1}{1/n} \approx \frac{i/2}{t} =: \lambda_i/t,$$

where  $\lambda_i = i/2$  (and  $a_n \approx b_n$  means that  $a_n/b_n \rightarrow 1$  as  $n \rightarrow \infty$ ). Our limiting birth process is hence born at a uniform time  $s$ , and then evolves with birth rate  $\lambda_i/t$  up until  $t = 1$ . Similarly to the DPA model, this distribution is equal to the distribution of a time-homogeneous linear birth process  $Z(t)$  (a Yule process), born at time 0 in state  $m$ , having birth rate  $\lambda_i$ , and being stopped after an  $\text{Exp}(1)$  time  $T$ . A Yule process with birth rate  $\lambda_i = i/2$ , starting in state  $m$  and observed at time  $t$ , has distribution

$$P(Z(t) = k \mid Z(0) = m) = \binom{m-1}{k-1} e^{-mt/2} (1 - e^{-t/2})^{k-1}$$

(see, e.g., [6]).

Furthermore, if the birth process has intensities  $\lambda_k = k/2$  and  $T \sim \text{Exp}(1)$ , then

$$\begin{aligned} P(Z(T) = i \mid Z(0) = m) &= \frac{\lambda_m}{\lambda_m + 1} \cdots \frac{\lambda_{i-1}}{\lambda_{i-1} + 1} \frac{1}{\lambda_i + 1} \\ &= \frac{m/2}{m/2 + 1} \cdots \frac{(i-1)/2}{(i-1)/2 + 1} \frac{1}{i/2 + 1} = \frac{2m(m+1)}{i(i+1)(i+2)}, \end{aligned}$$

where the last equality is shown by induction.

### 3.6. Proof of Theorem 2.5

The proof of Theorem 2.5 is more or less identical to the proof of Theorem 2.1. Now the limiting process can have births of each type, but also a simultaneous birth of both types (corresponding to the event that a new node attaches a directed edge, in either direction, to an existing node and the edge is reciprocated, which happens with probability  $\rho$ ). This makes the limiting distribution  $P(Y(T) = (i, j))$  harder to derive in that there are more paths to reach the state  $(i, j)$ , but the proof goes through in the same way.

## 4. Conclusions and discussion

We have analyzed the directed preferential attachment model using a new approach and showed that the limiting-degree distribution can be characterized by two *independent* (except when starting either in state  $(0, 1)$  or  $(1, 0)$ ) birth processes that are observed at a common  $\text{Exp}(1)$  random time point, thus creating dependence. Beside shedding more light on the structure of the limiting-degree distribution, this method also allows analyses of the bivariate tail probabilities (where both in- and out-degree were assumed large), cf. Theorem 2.3.

We also extended the DPA model to a more general model, the GDPA model, where new directed edges select nodes to attach to, in a way that may depend on both types of degrees, and, perhaps even more important, that the network may have a substantial fraction of edges being bidirected (or, equivalently, undirected). The limiting-degree distribution of this process was derived using similar methods. Also, here the distribution may be described by a bivariate birth process, but now the birth processes no longer evolve independently, and both components may have births at the same point in time. As before, the limiting-degree distribution is the state of this process observed at an  $\text{Exp}(1)$  random time point. The limiting tail probabilities for the GDPA model may perhaps also be derived using similar methods to the DPA model, but this remains an open problem.

Beside deriving the tail distribution for the GDPA model it would be interesting to study other generalizations of the DPA network model. For instance, many empirical networks exhibit clustering, meaning that triangles are more common than, for instance, in the DPA model. It would be of interest to study related models that allow for such clustering to be present in the network, and to see if a similar alternative construction of the limiting tail behavior may be obtained.

## Acknowledgements

I thank Sid Resnick and Svante Janson for valuable feedback during the early stage of this work, and Anders Martin-Löf for help with Laplace's method. I am grateful to the Swedish Research Council (grant 2015-05015) for financial support.

## References

- [1] BARABÁSI, A.-L. AND ALBERT, R. (1999). Emergence of scaling in random networks. *Science* **286**, 509–512.
- [2] BOLLOBÁS, B., BORGES, C., CHAYES, J. AND RIORDAN, O. (2003). Directed scale-free graphs. In *Proc. 14th Ann. ACM-SIAM Symp. on Discrete Algorithms (SODA '03)*, 132–139.
- [3] 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.
- [4] DOROGOVTSSEV, S. N., MENDES, J. F. F. AND SAMUKHIN, A. N. (2000). Structure of growing networks with preferential linking. *Phys. Rev. Lett.* **85**, 4633.
- [5] ETHIER, S. N. AND KURTZ, T. G. (2005). *Markov Processes: Characterization and Convergence*. John Wiley, Hoboken, NJ.
- [6] DE LA FORTELLE, A. (2006). Yule process sample path asymptotics. *Electron. Commun. Probab.* **11**, 193–199.

- [7] LESKOVEC, J. AND KREVL, A. (2014) SNAP datasets, *Stanford Large Network Dataset Collection*. Available at <http://snap.stanford.edu/data>.
- [8] SAMORODNITSKY, G., RESNICK, S. I., TOWSLEY, D., DAVIS, R., WILLIS, A. AND WAN, P. (2016). Nonstandard regular variation of in-degree and out-degree in the preferential attachment model. *J. Appl. Prob.* **53**, 146–161.
- [9] SPRICER, K. AND BRITTON, T. (2015). The configuration model for partially directed graphs *J. Statist. Phys.* **161**, 965–985.
- [10] WANG, T. AND RESNICK, S. I. (2019). Degree growth rates and index estimation in a directed preferential attachment model. To appear in *Stoch. Process. Appl.* Available at <https://doi.org/10.1016/j.spa.2019.03.021>.