

On the continuous-time limit of the Barabási–Albert random graph

Angelica Pachon¹, Federico Polito² & Laura Sacerdote²

¹Faculty of Computing, Engineering and Science, University of South Wales, UK

²Mathematics Department “G. Peano”, University of Torino, Italy

February 25, 2020

Abstract

We prove that, via an appropriate scaling, the degree of a fixed vertex in the Barabási–Albert model appeared at a large enough time converges in distribution to a Yule process. Using this relation we explain why the limit degree distribution of a vertex chosen uniformly at random (as the number of vertices goes to infinity), coincides with the limit distribution of the number of species in a genus selected uniformly at random in a Yule model (as time goes to infinity). To prove this result we do not assume that the number of vertices increases exponentially over time (linear rates). On the contrary, we retain their natural growth with a constant rate superimposing to the overall graph structure a suitable set of processes that we call the *planted model* and introducing an ad-hoc sampling procedure.

Keywords: Barabási–Albert model; Preferential attachment random graphs; Planted model; Discrete- and continuous-time models; Yule model.

MSC2010: 05C80, 90B15, 60J80.

1 Introduction

One of the most popular models for network growth is the preferential attachment model proposed by Barabási and Albert [2] to describe the web graph growth. In this model a newly created vertex is connected to one of those already present in the graph with a probability proportional to their degrees. An important characteristic related to a non-degenerate preferential attachment growth mechanism is the presence of a power-law distribution for the asymptotic degree of a vertex selected uniformly at random. This property is actually observed on World Wide Web data [2, 14, 21]. Furthermore, power-law distributions also occur frequently in other real-world phenomena and many of them are strictly related to the preferential attachment paradigm [7, 9, 10, 20, 27]. This fact determines an increasing interest on the Barabási–Albert model (BA model in the following) and for random graphs growing with preferential attachment rules in general. Indeed, there is an already extensive literature analyzing this class of random graphs. See for recent references Chapter 8 in [32] and the papers cited therein; we also recall [12, 22, 24]. The typical techniques considered, also implemented in the latter papers, are mainly of combinatorial type and based on the analysis of the expectation of specific functions of the degree or in-degree, together with concentration inequalities [5, 8, 11]. Other methods involve continuum and discrete approaches to study large but finite growing graphs [23] and an embedding of the random graph processes into a continuous setting involving a sequence of pure birth continuous Markov chains (see [1, 30, 3]). The technique of embedding a discrete sequence of random variables in continuous time processes is known for almost fifty years. When it is used on random graph processes, asymptotic results about properties of the vertices are obtained through an efficient use of branching process methods (see e.g. [30, 3]). Despite its generality, the application of this technique is not straightforward when the considered graph corresponds to the Barabási–Albert model whose growth allows the simultaneous birth of $m \geq 1$ new links. To deal with this problem it would be necessary either to develop suitable “ad-hoc” coupling

techniques or merge vertices. For this last procedure see [3] for preferential attachment networks. In continuous time there are other well-known probabilistic models which are clearly related to preferential attachment random graphs. Among them the Yule model for macroevolution [33]. The relationship linking the Yule model with some discrete-time preferential attachment models is not straightforward [28] but can be exploited to effectively study discrete-time preferential attachment random graphs. In [28] we showed how a specific discrete-time model with preferential attachment, the Simon model [31], is related (in a sense of weak convergence) to a set of Yule models.

The aim of this paper is to analyze specific aspects of the limiting behaviour of the BA model. To pursue this goal we follow and improve the methodology of [28] and relate the BA and the Yule models. Specifically, we couple the degree growth process of a vertex to a set of Markov processes, and we introduce the planted tree, an auxiliary branching structure superimposed to the random graph. Then, by means of this, we establish that the degree of a vertex chosen uniformly at random converges in distribution to the size of a uniformly chosen genus in a m -Yule model (a Yule model characterized by an infinite sequence of independent Yule processes, each starting with m individuals). We underline that we do not describe the dynamics of the degree of fixed vertices and the dynamics of the growth of the number of vertices with a given degree at the same time. Instead, we create a separate mechanism to describe the dynamics of the growth of the number of vertices with a given degree that does not need the Markov property of the degree processes.

We underline how this approach may prove to be useful in other cases as well, for example when the embedding method does not apply. Embedding techniques are problematic for more general preferential attachment models which are non-Markov, i.e., in which the emergence of future connections to existing vertices does not depend solely on the present state of their degree (for instance the connections could be affected by time delays of the random intervals at which the degree of a vertex changes, see [4] and the references therein). Other cases in which the approach of [28] and this paper might be applied are models with more general preferential attachment functions, such as models involving individual fitness [6] and/or aging [16], but also models in which hybrid rules are considered such as the uniform/preferential attachment [29].

The paper is organized as follows. The BA, the Yule and the m -Yule models are described in Section 2. In Section 3, the main results are presented (the related proofs are contained Section 5) together with a heuristic motivation of their validity. Summarizing the main results briefly, Theorem 3.1 shows that when infinitely many vertices have already appeared in the BA model, the degree distribution of a vertex appearing subsequently coincides with the distribution of the number of individuals in a Yule process starting with m initial individuals. Note that this result is consistent with the known related result in the case of preferential attachment trees, that is when $m = 1$ (see for example [30]). Theorem 3.2 proves the convergence to the same limit distribution of the degree of a vertex chosen uniformly at random in the BA model (when the number of vertices diverges) and of the size of a genus chosen uniformly at random in an m -Yule model when time goes to infinity. We also prove that in the BA model the proportion of vertices with a given degree k converges in probability (as the number of vertices diverges) to the probability that the degree of a vertex chosen uniformly at random is equal to k . The exact form of the limit distribution of Theorem 3.2 is then given in Proposition 3.1. Furthermore, the above results can be extended to preferential attachment random graphs for which Lemma 5.1 holds. This is mentioned in Remarks 3.4 and 3.5. In Section 4, a method for the sampling procedure of a random vertex in a general random graph model is proposed together with the notion of planted model. This method is a key tool to prove our main results. More specifically, this procedure is used to prove the relation between a randomly selected vertex in the BA model and a genus chosen uniformly at random from one of the m -Yule models which in turn is chosen uniformly at random from the set of all m -Yule models present in the planted model. Finally, as recalled above, Section 5 contains the proofs of the above-mentioned results and the necessary auxiliary lemmas.

2 Preliminaries

2.1 The Barabási-Albert model

In [2], the preferential attachment paradigm was proposed for the first time to model the growth of the World Wide Web. To do so the authors introduced a random graph model in which the vertices were added to the graph one at a time and joined to a fixed number of existing vertices, selected with probability proportional to their degree. In such a model the vertices represented the web pages and the edges their links. In [2] the model is described as follows:

Starting with a small number (m_0) of vertices, at every time step add a new vertex with m ($\leq m_0$) edges that link the new vertex to m different vertices already present in the system. To incorporate preferential attachment, assume that the probability that a new vertex will be connected to a vertex i depends on the connectivity k_i of that vertex, so it would be equal to $k_i/\sum_j k_j$. Thus, after t steps the model leads to a random network with $t + m_0$ vertices and mt edges.

The model was then defined in rigorous mathematical terms by Bollobás et al. [5]. However, in this paper we follow a large part of the literature in referring to the above model as the Barabási–Albert model even though it should be more correctly named after the authors of [5]. Here we recall their definition for the growth of the random graph process $(G_m^t)_{t \geq 1}$.

Definition 2.1. *For each $m \geq 1$ and for every $n \in \mathbb{N}$, the process $(G_m^t)_{t \geq 1}$ is such that,*

1. *at time $t = n(m + 1) + 1$ a new vertex v_{n+1} is added;*
2. *for $i = 2, \dots, m + 1$, at each time $t = n(m + 1) + i$ an edge from v_{n+1} to v is added with v chosen with the following probabilities:*

$$\mathbb{P}(v_{n+1} \rightarrow v) = \begin{cases} \frac{d(v, t-1)}{2(mn + i - 1) - 1}, & v \neq v_{n+1}, \\ \frac{d(v, t-1) + 1}{2(mn + i - 1) - 1}, & v = v_{n+1}. \end{cases} \quad (2.1)$$

In (2.1) $d(v, t)$ denotes the degree of the vertex v in G_m^t .

We explicitly underline that $(G_m^t)_{t \geq 1}$ starts at time $t = 1$ with a single vertex, v_1 , without loops. However, since at time $t = 2$ the only existing vertex is v_1 , then a loop is produced.

2.2 The Yule model

In this section we recall a classical continuous-time stochastic process which will be proven in the following to be strictly related to the BA random graph described in the preceding section. To avoid misunderstandings, we denote here by $T \in \mathbb{R}^+$ the continuous-time variable, while $t \in \mathbb{N}^* = \{1, 2, \dots\}$ indicates the discrete time.

The model we are concerned with was introduced in 1925 by Yule [33] to describe the macroevolution of a population characterized by the presence of different genera and species belonging to them. In order to describe it we first recall the well-known definition of a Yule process, i.e. a linear pure birth process in continuous time. A Yule model will then be defined in terms of a collection of independent Yule processes of possibly different birth intensities.

Definition 2.2. *A Yule process $\{N(T)\}_{T \geq 0}$ is a counting process in continuous time with state space \mathbb{N}^* , having initial condition $N(0) = g$, $g \geq 1$, almost surely and infinitesimal transition probabilities*

$$\mathbb{P}(N(T+h) = k + \ell \mid N(T) = k) = \begin{cases} k\lambda h + o(h), & \ell = 1, \\ o(h), & \ell > 1, \\ 1 - k\lambda h + o(h), & \ell = 0, \end{cases} \quad (2.2)$$

where $\lambda > 0$ is the birth intensity and $h > 0$.

This process describes the growth of the size of a population in which, during any short time interval of length h each member has probability $\lambda h + o(h)$, independently one another, to create a new individual. Note that the probability of simultaneous births is $o(h)$.

Yule [33] proposes to use independent copies of this process to model the growth of the number of species belonging to each separate genus. In turn, the evolution of the appearing genera is modelled by a further Yule process characterized by a possibly different birth intensity, say β , and independent of the former. The stochastic process determined by the combination of these two types of Yule processes is now known as a Yule model:

Definition 2.3. *A Yule model describes the growth of the number of genera and species according to the following rules:*

1. *genera (each comprising a single species) appear as a Yule process $\{N_\beta(T)\}_{T \geq 0}$ of parameter β with one genus at time $T = 0$ almost surely;*
2. *each time a new genus appears, a copy of a Yule process of parameter λ with a single initial progenitor starts. Those copies are independent one another and of the process of appearance of genera. Each copy models the evolution of species belonging to the same genus.*

In this paper we also consider an m -Yule model (denoted by $\{Y_{\lambda,\beta}^m(T)\}_{T \geq 0}$), that is a process similar to a classical Yule model but in which the birth processes describing the evolution of the species belonging to each genus start from $m \in \mathbb{N}^*$ initial species almost surely. To underline the initial condition we will add a superscript m to the Yule process counting the number of species for each genus: $\{N_\lambda^m(T)\}_{T \geq 0}$. We explicitly remark that the letter m , used for the initial value of the m -Yule model was already used to indicate the number of edges from a vertex in the BA model. This choice is not a coincidence, in the next sections we will in fact show that, as soon as we create a correspondance between the two models, the initial value $N_\lambda^m(0)$, is determined by the parameter m of the BA model. Finally, we would like to point out that the 1-Yule model coincides with the original Yule model of [33].

3 Main results

In order to better introduce our main results, let us first describe a heuristic approach explaining the relation between the discrete time process for the degree growth of a fixed vertex and a Yule process. In the BA model, m directed edges sequentially connect each new vertex to the others with probabilities proportional to the degrees of the existing vertices. Thus, at the time at which there are n vertices, that is at time $t = n(m + 1)$, we have mn directed edges, and by the preferential attachment rule we have approximately

$$\mathbb{P}[d(v, (n + 1)(m + 1)) = k + 1 \mid d(v, n(m + 1)) = k] \approx \frac{km}{2mn} = \frac{k}{2n}, \quad (3.1)$$

where $d(v, t)$ is the degree of v in the BA model. The approximation done in (3.1) consists in connecting all the m edges simultaneously instead of sequentially, that is, we consider m chances of increasing the degree of v from k to $k + 1$. Furthermore, we neglect the increase of the number of vertices during the random time interval between the instants at which the degree of v changes from k to $k + 1$. By formula (3.1) the distribution of this random time interval is geometric with parameter $k/(2n)$. In this approximations, when $n \rightarrow \infty$ we obtain a convergence to an exponential random variable of parameter $k\lambda$, with $\lambda = 1/2$. Moreover, neglecting also the possibility of loops, the initial degree of v_i , $i \geq n$, turns out to be equal to m . These two observations suggest for large values of n to approximate the distribution of the degree of a vertex in the BA model by the distribution of the number of individuals in a Yule process with parameter $\lambda = 1/2$ and initial condition $N_\lambda(0) = m$. In Theorem 3.1, below, we make rigorous the above heuristics by proving that the process describing the degree of a fixed vertex in the BA model converges in distribution to the number of individuals in a Yule process with initial size m . Further, much interest is towards the study of the asymptotic degree of a vertex chosen uniformly at random. In Theorem 3.2 we show that the BA model is related to a sequence of suitably scaled m -Yule models. Exploiting this

relation we prove that the asymptotic degree distribution of a vertex chosen uniformly at random in the BA model coincides with the asymptotic distribution of the size of a genus chosen uniformly at random in the m -Yule model.

Theorem 3.1. *Let $z(i, w) : \mathbb{N}^* \times \mathbb{R}^+ \rightarrow \mathbb{N}$, $w \in \mathbb{R}^+$ be a function such that $c(w) := \lim_{i \rightarrow \infty} z(i, w)/i$ exists finite, with $c(w) : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ increasing in w . Let $b \geq 1$ and $w_1 < w_2 < \dots < w_b$ be positive real numbers. We have that*

$$\begin{aligned} & \lim_{i \rightarrow \infty} \mathbb{P} [d(v_i, (i + z(i, w_1))(m + 1)) = k_1, \dots, d(v_i, (i + z(i, w_b))(m + 1)) = k_b] \quad (3.2) \\ &= \mathbb{P} [N_{1/2}^m(\log(1 + c(w_1))) = k_1, \dots, N_{1/2}^m(\log(1 + c(w_b))) = k_b] \\ &= \prod_{\ell=1}^b \binom{k_\ell - 1}{k_\ell - k_{\ell-1}} e^{-\frac{k_\ell - 1}{2} \log\left(\frac{1+c(w_\ell)}{1+c(w_{\ell-1})}\right)} \left(1 - e^{-\frac{1}{2} \log\left(\frac{1+c(w_\ell)}{1+c(w_{\ell-1})}\right)}\right)^{k_\ell - k_{\ell-1}}. \end{aligned}$$

Here $w_0 = 0$, $k_0 = m$, and $m \leq k_1 \leq \dots \leq k_b \in \mathbb{N}^*$.

Remark 3.1. *Notice that, for $\ell = 1, \dots, b$, the required time change $i \mapsto i + z(i, w_\ell)$ behaves asymptotically as the linear function $i \mapsto i + ic(w_\ell)$. Moreover, the logarithm of its slope, $1 + c(w_\ell)$, is the time at which the Yule process is evaluated. Regarding the existence of the function $z(i, w_\ell)$, possible choices can be $z(i, w_\ell) = \lfloor iw_\ell \rfloor$ or $z(i, w_\ell) = \lfloor (i - 1)w_\ell \rfloor$.*

Remark 3.2. *Theorem 3.1 states that the joint distribution of the degrees of v_i at times $z(i, w_1)(m + 1), \dots, z(i, w_b)(m + 1)$ after its first appearance in $(G_m^t)_{t \geq 1}$, converges, as $i \rightarrow \infty$, to the joint distribution of the number of individuals of a Yule process with m initial individuals and parameter $\lambda = 1/2$, evaluated at the times $\log(1 + c(w_1)), \dots, \log(1 + c(w_b))$.*

Theorem 3.2. *Consider an m -Yule model $\{Y_{1/2,1}^m(T)\}_{T \geq 0}$, and let \mathcal{N}_T^m be the size of a genus chosen uniformly at random at time T in $\{Y_{1/2,1}^m(T)\}_{T \geq 0}$. Consider the random graph process $(G_m^t)_{t \geq 1}$ defining the BA model with $N_{k,t}$ vertices with degree k . Let $d(V_t)$ be the degree of a vertex chosen uniformly at random at time t in (G_m^t) . Then, for $t = n(m + 1)$ we have*

$$p_k := \lim_{n \rightarrow \infty} \mathbb{P}(d(V_t) = k) = \lim_{T \rightarrow \infty} \mathbb{P}(\mathcal{N}_T^m = k), \quad k \geq m, \quad (3.3)$$

and for $C > m\sqrt{8}$,

$$\mathbb{P}\left(\max_k \left| \frac{N_{k,t}}{n} - \mathbb{P}(d(V_t) = k) \right| \geq C \sqrt{\frac{(m+1) \log(n(m+1))}{n}}\right) = o(1). \quad (3.4)$$

Furthermore, as $n \rightarrow \infty$, $N_{k,t}/n \rightarrow p_k$ in probability.

Using the previous theorem and directly exploiting the properties of the m -Yule model we are able to recover the well-known result for the asymptotic degree distribution of the BA random graph.

Proposition 3.1. *Consider an m -Yule model $\{Y_{1/2,1}^m(T)\}_{T \geq 0}$ and the size \mathcal{N}_T^m of a genus chosen uniformly at random at time T from it as in Theorem 3.2. Then,*

$$p_k = m(m + 1)B(k, 3), \quad k \geq m, \quad (3.5)$$

where $B(a, b)$ is the Beta function.

Remark 3.3. *Notice that the distribution (3.5) coincides with the degree distribution of the BA model [5].*

Remark 3.4. *In Section 5 we prove the technical Lemma 5.1 on the behaviour of the degree process for a fixed vertex. Theorems 3.1, 3.2 and Proposition 3.1 can also be proved for any random graph process for which such Lemma 5.1 holds. Notice that if necessary, Lemma 5.1 can be extended to the case in which the constant b_2 can be taken equal to zero.*

Remark 3.5. An alternative example in which Lemma 5.1 still holds is the “independent” model: for each newly added vertex its m edges are connected to old vertices independently one another. Formally, in Definition 2.1 replace (2.1) by

$$\mathbb{P}(v_{n+1} \rightarrow v) = \begin{cases} \frac{d(v, n(m+1))}{2(mn+1)-1}, & v \neq v_{n+1}, \\ \frac{d(v, n(m+1))+1}{2(mn+1)-1}, & v = v_{n+1}. \end{cases} \quad (3.6)$$

In Section 5, which is devoted to the proofs, Remark 5.1 explains why Lemma 5.1 holds for the independent model.

4 Sampling a random vertex

Before going through the proofs of Theorem 3.1 and Theorem 3.2, we introduce here the general notion of *planted model* and a fundamental procedure we will make use in the next section to prove the relationship between two random quantities in the BA model and in the m -Yule model. In particular, we will put in relation the degree of a vertex chosen uniformly at random in the BA model and the number of species of a genus chosen uniformly at random from one of the m -Yule models, also chosen uniformly at random from the set of all m -Yule models in the planted model.

For a greater generality we consider the case in which the number of edges added each time a vertex appears, form a sequence $\{M_j\}_{j \geq 1}$ of random variables taking values in \mathbb{N}^* almost surely. This result can be easily specialized to the case of the BA model, that is $M_j = m$ a.s. for every j . An example is the random graph related to Simon model (see [28]): it can be related to a Yule model where the above random variables are independent and geometrically distributed.

In order to describe the sampling procedure, we introduce first a model that we call the *planted model* for the random graph $(\mathcal{G}_t)_{t \geq 1}$. The idea underlying the planted model is to superimpose a tree structure on the graph which is independent of the degree processes.

We start by noting that, at each time of the form $\mathfrak{T}_i = \sum_{r=1}^i (M_r + 1)$, the graph $\mathcal{G}_{\mathfrak{T}_i}$ has exactly i vertices, $i \in \mathbb{N}^*$. We refer to them as the *planted vertices*. Let us now consider the value i to be fixed; to obtain the tree structure at the following times \mathfrak{T}_{n+1} , $n \geq i$, we attribute to v_{n+1} the role of child of a vertex chosen uniformly at random from the set of the existing vertices $\{v_1, v_2, \dots, v_n\}$. Iterating this procedure we obtain chains of successive offsprings of each of the planted vertices $\{v_1, v_2, \dots, v_i\}$. Further, we call a vertex v that appeared after v_j , $j = 1, \dots, i$, a descendant of v_j if both v and v_j belong to the same ancestral line. We order the descendant of v_j by renaming v as $v_{j,\ell}$, if v is the ℓ -th descendant of v_j and denote v_j as $v_{j,1}$, that is, v_j is in turn, its first descendant. In this way we construct i birth processes in discrete time, $\{b(v_j, \mathfrak{T}_n)\}_{n \geq i}$, $j = 1, \dots, i$. Here $b(v_j, \mathfrak{T}_n)$, $j = 1, \dots, i$, $n \geq i$, is the total number of descendants of v_j at time \mathfrak{T}_n . Table 1 shows an example of the construction of the planted model. Note that we have:

- $b(v_j, \mathfrak{T}_i) = 1$, $j = 1, \dots, i$;
- $\mathbb{P}[b(v_j, \mathfrak{T}_{n+1}) = k + 1 \mid b(v_j, \mathfrak{T}_n) = k] = k/n$, $k \geq 1$, $n \geq i$, $j = 1, \dots, i$.

The second equality holds because at time \mathfrak{T}_{n+1} , $n \geq i$, each existing vertex in the set $\{v_1, v_2, \dots, v_n\}$ may give birth to a new one with probability $1/n$.

Note that, for a fixed $i \geq 1$, the planted model is defined for $n \geq i$. Thus for example, given a value of i there is no process $\{b(v_j, \mathfrak{T}_n)\}_{n \geq i}$ with $j > i$, because j has to be an element of $\{1, \dots, i\}$. The dynamic of the planted model then proceeds for $n \geq i$. Finally, note that the i discrete-time birth processes are exchangeable.

4.0.1 Sampling from the planted model

Consider the following procedure. Given a realization of $\{b(v_j, \mathfrak{T}_n)\}_{n \geq i}$, $j = 1, \dots, i$,

1. choose one of the i discrete-time birth processes with probability proportional to the number of its vertices;

n		$b(v_1, \mathfrak{T}_n)$		$b(v_2, \mathfrak{T}_n)$	\dots		$b(v_i, \mathfrak{T}_n)$
i	$v_1 = v_{1,1}$	1	$v_2 = v_{2,1}$	1		$v_i = v_{i,1}$	1
$i+1$		1	$v_{i+1} = v_{2,2}$	2			1
$i+2$		1		2	\dots	$v_{i+2} = v_{i,2}$	2
$i+3$	$v_{i+3} = v_{1,2}$	2		2			2
$i+4$		2	$v_{i+4} = v_{2,3}$	3			2
\dots	\dots	\dots	\dots	\dots	\dots	\dots	\dots

Table 1: (First line): The construction starts with i discrete-time birth processes at time \mathfrak{T}_i , each one with one individual. (Second line): At time \mathfrak{T}_{i+1} a new vertex v_{i+1} appears. The vertex v_{i+1} is assigned as a child to one of the existing vertices $\{v_1, v_2, \dots, v_i\}$ with probability $1/i$. In this table the appearing vertex v_{i+1} becomes a child of v_2 and consequently it is renamed as $v_{2,2}$, that is the second individual in the birth process relative to v_2 . (Next lines): At times \mathfrak{T}_{n+1} , $n \geq i$, the vertex v_{n+1} appears and it is assigned to one of the existing vertices with probability $1/n$. Observe that in this example $b(v_2, \mathfrak{T}_{i+4}) = 3$. Given this information, $\mathbb{P}[b(v_2, \mathfrak{T}_{i+5}) = 4] = 3/(i+4)$.

2. choose a vertex uniformly at random among those belonging to the realization of the selected birth process.

Our focus is on the selected vertex $v_{j,\ell}$, $j = 1, \dots, i$, $\ell = 1, \dots, b(v_j, \mathfrak{T}_n)$ and on the selected birth process. Let W be the index of the birth process chosen. Plainly, W takes values in $\{1, \dots, i\}$ almost surely.

Proposition 4.1. *It holds,*

1. $\mathbb{P}(W = j) = 1/i$, $j \in \{1, \dots, i\}$,
2. $\mathbb{P}(\{v_{j,\ell} \text{ is selected}\}) = 1/n$.

Proof. It immediately follows from the exchangeability of the i discrete-time birth processes. \square

Remark 4.1. *The suggested algorithm is a way to select a vertex uniformly at random from $\mathcal{G}_{\mathfrak{T}_n}$, $n \geq i$, and refers to a given realization of the i birth processes $\{b(v_j, \mathfrak{T}_n)\}_{n \geq i}$, $j = 1, \dots, i$. Averaging on all possible realizations of the i birth processes we actually select a vertex uniformly at random: we first choose one of the i birth processes belonging to the planted model with uniform probability, then we select a vertex among those belonging to the chosen birth process again with uniform probability.*

5 Proofs

We first give a brief outline of the proofs of the main results described in Section 3. Regarding Theorem 3.1, we start by showing that the transition probabilities of the degree process of a fixed vertex with sufficiently large index in the BA model are bounded above and below (Lemma 5.1). With these bounds we construct two Markov processes coupled with the original degree process (Lemma 5.2 and Corollary 5.1). In Lemma 5.3, we exploit this coupling to conclude that the finite-dimensional distribution of the degree of a vertex in the BA model converges to the finite-dimensional distribution of the number of individuals in a Yule process with initial population size equal to m . Notice that this result is consistent with the analysis of preferential attachment trees performed through continuous-time branching processes (see e.g. [30, 3]).

To prove Theorem 3.2 we proceed according to the following steps. First, by making use of the planted model and the sample procedure from the planted model described in Section 4, we make explicit the relationship between the deterministic appearance of new vertices in the BA model and the random appearance of new births in a continuous-time Yule process. The key point is that, by Theorem 4.1, the choice of a vertex with uniform distribution in the BA model is equivalent to choosing a birth process from the planted model with uniform distribution and then choosing a vertex among those belonging to the selected birth process, again with uniform distribution. Then,

in Lemma 5.4 we prove that the number of individuals in each birth process of the planted model, say $\{b_i^j\}_{i \geq 1}$, where $b_i^j = \{b(v_j, n(m+1))\}_{n \geq i}$, $1 \leq j \leq i$, converges in distribution as $i \rightarrow \infty$, to the size of a Yule process with parameter $\beta = 1$ and with one initial progenitor.

5.1 Auxiliary lemmas and the proof of Theorem 3.1

The proof of Theorem 3.1 can be summarized in two main steps. Within the structure of the BA model we first identify two different counting processes in discrete time, one for the appearing of in-links of each specific vertex and the other related to the creation of new vertices. Then, we prove that these two processes converge to the two birth processes which are at the basis of the definition of an m -Yule model.

Before starting the construction of the process for the appearance of in-links of a fixed vertex we introduce the following definition.

Definition 5.1. *We say that a vertex v_i appears “complete” when it has appeared in the BA random graph process together with all the directed edges originated from it.*

Note that the degree of a complete vertex is at least m , and at time $t = n(m+1)$, the BA model has for the first time exactly n complete vertices.

Next we determine how the degree of a fixed vertex v_i , for a sufficiently large i , changes during the time until a new complete vertex appears.

Lemma 5.1. *Let $(G_m^t)_{t \geq 1}$ be the random graph process defining the BA model and let $d(v_i, t)$ denote the degree of an existing vertex v_i at time t , $i \leq n$. Given that $d(v_i, n(m+1)) = k$, $n > k \geq m$, for sufficiently large i there exist constants $b_1 > b_2 > 0$ and $c_1, c_2 > 0$ such that*

$$\begin{aligned} \frac{k}{2(n+1)} + c_2 \left(\frac{k}{n}\right)^2 &< \mathbb{P}[d(v_i, (n+1)(m+1)) = k+1 | d(v_i, n(m+1)) = k] \\ &< \frac{k}{2n} + c_1 \left(\frac{k}{n}\right)^2, \end{aligned} \quad (5.1)$$

and, for $m > 1$,

$$b_2 \left(\frac{k}{n}\right)^2 \leq \mathbb{P}[k+2 \leq d(v_i, (n+1)(m+1)) \leq k+m | d(v_i, n(m+1)) = k] \leq b_1 \left(\frac{k}{n}\right)^2. \quad (5.2)$$

Furthermore,

$$\mathbb{P}[d(v_{n+1}, (n+1)(m+1)) = m] = \prod_{\ell=2}^{m+1} \left(1 - \frac{1}{2(mn + \ell - 1) - 1}\right) = 1 - \Theta(1/n), \quad (5.3)$$

where we make use of the asymptotic Big Theta notation [17].

Proof. Our aim is to determine the change of degree of a fixed vertex during the time interval $(n(m+1), (n+1)(m+1)]$, i.e., during the time interval necessary to switch from n to $(n+1)$ complete vertices.

Let us fix $t = n(m+1)$ and follow the graph growth during the considered interval. At time $n(m+1) + 1$ a new vertex v_{n+1} (without edges) appears. Then from time $n(m+1) + 2$ to time $(n+1)(m+1)$ a directed edge from v_{n+1} to an existing vertex v_i , $i \leq n+1$, is added. The vertex v_i is chosen with probability given by (2.1). Let $Y_{v_i}^n$ be the total number of incoming edges to v_i , $i \leq n$, added to v_i during the time interval $(n(m+1), (n+1)(m+1)]$. Note that $\mathbb{P}[d(v_i, (n+1)(m+1)) = k + \ell | d(v_i, n(m+1)) = k] = \mathbb{P}[Y_{v_i}^n = \ell | d(v_i, n(m+1)) = k]$, $\ell = 0, \dots, m$. To estimate the latter conditional probabilities for a sufficiently large i we distinguish the cases $Y_{v_i}^n = 0$, $Y_{v_i}^n = 1$, and $Y_{v_i}^n \geq 2$.

In the first case, considering the probabilities (2.1) we have

$$\mathbb{P}[Y_{v_i}^n = 0 | d(v_i, n(m+1)) = k] = \prod_{\ell=2}^{m+1} \left(1 - \frac{k}{2(mn + \ell - 1) - 1}\right). \quad (5.4)$$

Since $\prod_{\ell=2}^{m+1} \left(1 - \frac{k}{2(mn+\ell-1)}\right) \leq \left(1 - \frac{k}{2(mn+m)-1}\right)^m$, we get the upper bound for (5.4),

$$\left(1 - \frac{k}{2(mn+m)-1}\right)^m = 1 - \frac{mk}{2m(n+1)-1} + O\left(\frac{k^2}{n^2}\right). \quad (5.5)$$

Furthermore, since $\prod_{\ell=2}^{m+1} \left(1 - \frac{k}{2(mn+\ell-1)}\right) \geq \left(1 - \frac{k}{2(mn+1)-1}\right)^m$, we get the lower bound

$$\left(1 - \frac{k}{2(mn+1)-1}\right)^m = 1 - \frac{mk}{2mn+1} + O\left(\frac{k^2}{n^2}\right). \quad (5.6)$$

Now we move first to the third case. We observe that if $m = 1$ then $\mathbb{P}(Y_{v_i}^n \geq 2 \mid d(v_i, n(m+1)) = k) = 0$. Thus, we calculate such probability for $m > 1$ only. Furthermore, since we do not need a closed form of $\mathbb{P}(Y_{v_i}^n \geq 2 \mid d(v_i, n(m+1)) = k)$, we limit ourselves to estimate its order of magnitude. For each $y = 2, \dots, m$, the event $\{Y_{v_i}^n = y\}$ means that v_i gets y new incoming edges joining v_i at the times $t = n(m+1) + \ell$, $\ell = 2, \dots, m+1$. Given the value of the degree of v_i at time $t-1$, considering (2.1), a new edge is attached to v_i at time $t = n(m+1) + \ell$ with probability

$$p_{v_i}^{n,\ell} := \frac{d(v_i, n(m+1) + \ell - 1)}{2(mn + \ell - 1) - 1}. \quad (5.7)$$

Let Ω be the space of all sequences of m dichotomous independent experiments, performed at times $t = n(m+1) + \ell$, $\ell = 2, \dots, m+1$, with exactly y successes. Assume that $p_{v_i}^{n,\ell}$, $\ell = 2, \dots, m+1$, are the probabilities of success. Note that the cardinality of Ω is equal to that of the set of all y -combinations from a given set of m distinct elements, i.e. $|\Omega| = \binom{m}{y}$. Take the set $\{2, \dots, m+1\}$ and consider its y -combinations, say $C_y = \{e_1, \dots, e_{\binom{m}{y}}\}$ (e.g. ordered by their smallest element). For each $e \in C_y$, let $e(j)$ denote the position of the j -th success in e , $j = 1, \dots, y$. We have,

$$\begin{aligned} \mathbb{P}(Y_{v_i}^n = y \mid d(v_i, n(m+1)) = k) &= \sum_{e \in C_y} \prod_{j=1}^y p_{v_i}^{n,e(j)} \prod_{\ell \in \{2, \dots, m+1\}, \ell \notin e(1), \dots, e(y)} (1 - p_{v_i}^{n,\ell}) \\ &= \binom{m}{y} \Theta\left(\frac{k^y}{n^y}\right) \left[1 - \Theta\left(\frac{k}{n}\right)\right]^{m-y} \\ &= \binom{m}{y} \Theta\left(\frac{k^y}{n^y}\right) \sum_{\ell=0}^{m-y} \binom{m-y}{\ell} (-1)^\ell \Theta\left(\frac{k^\ell}{n^\ell}\right) \\ &= \Theta\left(\frac{k^y}{n^y}\right), \quad 2 \leq y \leq m. \end{aligned} \quad (5.8)$$

Hence,

$$\mathbb{P}(Y_{v_i}^n \geq 2 \mid d(v_i, n(m+1)) = k) = \Theta\left(\frac{k^2}{n^2}\right). \quad (5.9)$$

Finally, by (5.6) and (5.9) we obtain that $\mathbb{P}(Y_{v_i}^n = 1 \mid d(v_i, n(m+1)) = k)$ is at most

$$1 - \left[1 - \frac{mk}{2mn+1} + O\left(\frac{k^2}{n^2}\right)\right] - \Theta\left(\frac{k^2}{n^2}\right) < \frac{k}{2n} + O\left(\frac{k^2}{n^2}\right), \quad (5.10)$$

and by (5.5) and (5.9), $\mathbb{P}(Y_{v_i}^n = 1 \mid d(v_i, n(m+1)) = k)$ is at least

$$1 - \left[1 - \frac{mk}{2m(n+1)-1} + O\left(\frac{k^2}{n^2}\right)\right] - \Theta\left(\frac{k^2}{n^2}\right) > \frac{k}{2(n+1)} + O\left(\frac{k^2}{n^2}\right). \quad (5.11)$$

Therefore, for sufficiently large i we have that, for each $k \geq m$, there exist $b_1 > b_2 > 0$ such that (5.9) gives (5.2), and $c_1, c_2 > 0$ such that (5.10) and (5.11) give (5.1).

In order to determine (5.3), let $X_{v_{n+1}}^n$ be the number of incoming edges from v_{n+1} to itself during the time interval $(n(m+1), (n+1)(m+1)]$, that is the number of loops. Note that during this period $X_{v_{n+1}}^n$ can be at most equal to m , since at time $n(m+1) + 1$ no edge is added. Thus, by (2.1), the probability of no loops for v_{n+1} during such time interval is given by

$$\mathbb{P}(X_{v_{n+1}}^n = 0) = \prod_{i=2}^{m+1} \left(1 - \frac{1}{2(mn+i-1)-1}\right) = [1 - \Theta(1/n)]^m = 1 - \Theta(1/n). \quad (5.12)$$

If the number of loops for v_{n+1} is zero, this is equivalent to say that when v_{n+1} appears complete, its degree is equal to m . Thus, by (5.12) we can write $\mathbb{P}[d(v_{n+1}, (n+1)(m+1)) = m] = 1 - \Theta(1/n)$, so that the proof is complete. \square

Remark 5.1. For the independent model described in Remark 3.5, note that the left-hand side of formula (5.4) is equal to (5.6), formula (5.7) is equal to $p_{v_i}^{n,2}$ for all $\ell = 2, \dots, m+1$, $\mathbb{P}(Y_{v_i}^n = 1 \mid d(v_i, n(m+1)) = k)$ is equal to (5.10) and

$$\mathbb{P}[d(v_{n+1}, (n+1)(m+1)) = m] = \left(1 - \frac{1}{2(mn+1)-1}\right)^m = 1 - \Theta(1/n). \quad (5.13)$$

Now we consider the degree process $\{d(v_i, n(m+1))\}_{n \geq i}$, indexed by n , where $d(v_i, n(m+1))$ satisfies (5.1), (5.2) and (5.3). Let $E = \{m, m+1, \dots\}$ be the state space of the process $\{d(v_i, n(m+1))\}_{n \geq i}$ and let $\mathcal{M}(E)$ be the class of probability measures on the space E endowed with the σ -algebra $\mathcal{F} = \mathcal{P}(E)$, the power set of E . The degree process $\{d(v_i, n(m+1))\}_{n \geq i}$ is defined on the product space $(E^\infty, \mathcal{F}^\infty) = (\times_{n=i}^\infty E, \otimes_{n=i}^\infty \mathcal{F})$. The process from time i to time $i+h$, $\{d(v_i, n(m+1))\}_{n=i}^{i+h}$ takes values in the product space $(E^h, \mathcal{F}^h) = (\times_{n=i}^{i+h} E, \otimes_{n=i}^{i+h} \mathcal{F})$. The elements of the spaces (E^h, \mathcal{F}^h) and $(E^\infty, \mathcal{F}^\infty)$ will be denoted by $x^{i+h} = (x_i, x_{i+1}, \dots, x_{i+h})$ and $x^\infty = (x_i, x_{i+1}, \dots)$, respectively. We say that $x^{i+h} \leq y^{i+h}$ if and only if $x_{i+j} \leq y_{i+j}$ for all $0 \leq j \leq h$.

To prove Theorem 3.1, we proceed through two steps:

1. we define two Markov processes, on the same probability space as $\{d(v_i, n(m+1))\}_{n \geq i}$, determined by suitable Markov kernels and we show that those two processes bound from above and below the degree process of the BA model (Lemma 5.2 and Corollary 5.1);
2. we prove that those two processes, each of them evaluated at a convenient time, converge in distribution as $i \rightarrow \infty$ (and therefore as $n \rightarrow \infty$) to a unique process evaluated at a unique time (Lemma 5.3).

With respect to the first step, for any $m > 1$, we define p_{n+1} to be a positive function on $E^n \times E$, measurable with respect to $\mathcal{F}^n \otimes \mathcal{F}$, and given by

$$p_{n+1}(x^n, x) = p_{n+1}(x^n, x_n + \ell) = \begin{cases} \frac{x_n}{2(n+1)} + c_2 \left(\frac{x_n}{n}\right)^2, & \ell = 1, \\ b_2 \left(\frac{x_n}{n}\right)^2, & \ell = 2, \\ 1 - \frac{x_n}{2(n+1)} - (b_2 + c_2) \left(\frac{x_n}{n}\right)^2, & \ell = 0, \\ 0, & \text{otherwise.} \end{cases} \quad (5.14)$$

Note that this function depends only on x_n , the last element of x^n , and n . Then we define the following Markov transition kernel K_{n+1}^p from $E^n \times \mathcal{F}$ into $[0, 1]$:

$$K_{n+1}^p(x^n, B) = \sum_{x \in B} p_{n+1}(x^n, x), \quad x^n \in E^n, B \in \mathcal{F}. \quad (5.15)$$

The mapping $B \rightarrow K_{n+1}^p(x^n, B)$ is a measure $P_{n+1} \in \mathcal{M}(E)$ for every $x^n \in E^n$. Similarly we define the function

$$r_{n+1}(z^n, z) = r_{n+1}(z^n, z_n + \ell) = \begin{cases} \frac{z_n}{2n} + c_1 \left(\frac{z_n}{n}\right)^2, & \ell = 1, \\ b_1 \left(\frac{z_n}{n}\right)^2, & \ell = m, \\ 1 - \frac{z_n}{2n} - (b_1 + c_1) \left(\frac{z_n}{n}\right)^2, & \ell = 0, \\ 0, & \text{otherwise,} \end{cases} \quad (5.16)$$

that we associate to the Markov transition kernel K_{n+1}^r , where $B \rightarrow K_{n+1}^r(z^n, B)$, is a measure $R_{n+1} \in \mathcal{M}(E)$ for every $z^n \in E^n$. Note that if $x_i \leq z_i$, $b_2 < b_1$ and letting $c_1 > c_2$, then by (5.14) and (5.16), $x_n \leq z_n$, $n > i$. Take $y_n \in E$ such that, $x_n \leq y_n \leq z_n$, $n \geq i$. Then, from (5.14) and (5.16), there exists a function $q_{n+1}(y^n, y) = q_{n+1}(y^n, y_n + \ell)$, such that:

$$\begin{aligned} p_{n+1}(x^n, x_n + 1) &< q_{n+1}(y^n, y_n + 1) < r_{n+1}(z^n, z_n + 1), \\ p_{n+1}(x^n, x_n + 2) &< \sum_{\ell=2}^m q_{n+1}(y^n, y_n + \ell) < r_{n+1}(z^n, z_n + m), \\ q_{n+1}(y^n, y_n) &= 1 - \sum_{\ell=1}^m q_{n+1}(y^n, y_n + \ell), \end{aligned} \tag{5.17}$$

whenever $x_n \leq y_n \leq z_n$, $n \geq i$. We associate this function to a further Markov transition kernel K_{n+1}^q in the same way as K_{n+1}^p and K_{n+1}^r , where $B \rightarrow K_{n+1}^q(y^n, B)$ is a measure $Q_{n+1} \in \mathcal{M}(E)$ for every $y^n \in E^n$.

In order to prove that there exist two processes bounding respectively from above and below the degree process of the BA model we first need the following result:

Lemma 5.2. *Let X_i , Y_i , and Z_i , be random variables on E with distributions P_i , Q_i and R_i , respectively, and satisfying $\mathbb{P}(X_i = Y_i = Z_i) = 1$. Then there exist random variables X_{n+1} , Y_{n+1} , and Z_{n+1} , $n \geq i$, taking values in E , such that the conditional distributions of X_{n+1} given $X_n = x_n$, Y_{n+1} given $Y_n = y_n$, and Z_{n+1} given $Z_n = z_n$, are exactly $p_{n+1}(x^n, \cdot)$, $q_{n+1}(y^n, \cdot)$, and $r_{n+1}(z^n, \cdot)$, respectively. Moreover,*

$$\mathbb{P}(X_n \leq Y_n \leq Z_n, n = i, i + 1, \dots) = 1. \tag{5.18}$$

Proof. We seek to prove a stochastic ordering for $K_{n+1}^p(x^n, \cdot)$, $K_{n+1}^q(y^n, \cdot)$ and $K_{n+1}^r(z^n, \cdot)$. For this aim, take the set $B \in \mathcal{F}$ such that $B := \{b, b + 1, \dots\}$, where b is any integer $b \geq m$. Then,

$$K_{n+1}^p(x^n, B) = \sum_{j \geq b} p_{n+1}(x^n, j) = \begin{cases} 1, & b \leq x_n, \\ \frac{x_n}{2(n+1)} + (c_2 + b_2) \left(\frac{x_n}{n}\right)^2, & b = x_n + 1, \\ b_2 \left(\frac{x_n}{n}\right)^2, & b = x_n + 2, \\ 0, & b \geq x_n + 3, \end{cases} \tag{5.19}$$

$$K_{n+1}^q(y^n, B) = \sum_{j \geq b} q_{n+1}(y^n, j) = \begin{cases} 1, & b \leq y_n, \\ \sum_{i=\ell}^m q_{n+1}(y^n, y_n + \ell), & b = y_n + \ell, \\ & \ell = 1, \dots, m - 1, \\ q_{n+1}(y^n, y_n + m), & b = y_n + m, \\ 0, & b \geq y_n + m + 1, \end{cases} \tag{5.20}$$

and

$$K_{n+1}^r(z^n, B) = \sum_{j \geq b} r_{n+1}(z^n, j) = \begin{cases} 1, & b \leq z_n, \\ \frac{z_n}{2n} + (c_1 + b_1) \left(\frac{z_n}{n}\right)^2, & b = z_n + 1, \\ b_1 \left(\frac{z_n}{n}\right)^2, & b = z_n + 2, \\ 0, & b \geq z_n + 3. \end{cases} \tag{5.21}$$

Since $X_i = Y_i = Z_i$ a.s., $b_2 < b_1$, and $c_2 < c_1$, then by (5.14), (5.16) and (5.17), we obtain that $x_n \leq y_n \leq z_n$, $n \geq i$. Thus comparing the three kernels (5.19), (5.20) and (5.21), we have

$$K_{n+1}^p(x^n, B) \leq K_{n+1}^q(y^n, B) \leq K_{n+1}^r(z^n, B).$$

Equivalently, $K_{n+1}^p(x^n, \cdot)$ is stochastically smaller than $K_{n+1}^q(y^n, \cdot)$, and the latter is in turn stochastically smaller than $K_{n+1}^r(z^n, \cdot)$, whenever $x^n \leq y^n \leq z^n$. To show that (5.18) holds we finally apply Theorem 2 in [19]. \square

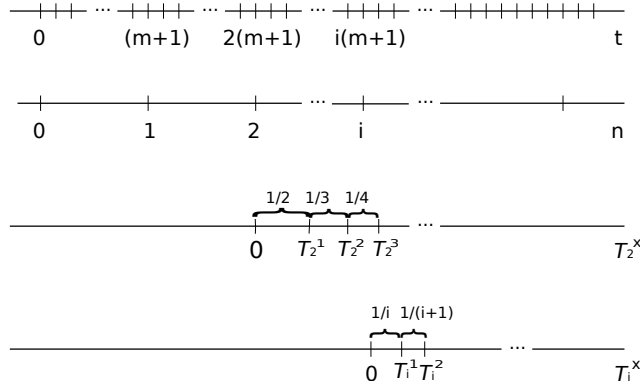


Figure 1: The first line represents the time axis of the BA model. The second line shows the number of complete vertices in the BA model. The third and fourth lines correspond to the partitions of $(0, T_2^x]$ and $(0, T_i^x]$, respectively.

Let us consider the process $\{d(v_i, n(m+1))\}_{n \geq i}$ and its probability space $(\Omega, \mathcal{A}, \mathbb{P})$. On the same probability space let us define two Markov processes $\{d^1(v_i, n(m+1))\}_{n \geq i}$ and $\{d^2(v_i, n(m+1))\}_{n \geq i}$, with their initial states such that $\mathbb{P}(d^1(v_i, i(m+1)) = d(v_i, i(m+1)) = d^2(v_i, i(m+1))) = 1$, and transition probabilities given by (5.16) and (5.14), respectively.

Corollary 5.1. *For i sufficiently large, there exist versions $\{\tilde{d}^1(v_i, n(m+1))\}_{n \geq i}$, $\{\tilde{d}^2(v_i, n(m+1))\}_{n \geq i}$, $\{\tilde{d}(v_i, n(m+1))\}_{n \geq i}$ of the processes $\{d^1(v_i, n(m+1))\}_{n \geq i}$, $\{d^2(v_i, n(m+1))\}_{n \geq i}$ and $\{d(v_i, n(m+1))\}_{n \geq i}$, respectively, such that*

$$\mathbb{P}[\tilde{d}^2(v_i, n(m+1)) \leq \tilde{d}(v_i, n(m+1)) \leq \tilde{d}^1(v_i, n(m+1)), n = i, i+1, \dots] = 1. \quad (5.22)$$

Proof. It immediately follows by applying Lemma 5.2 to $\{d^1(v_i, n(m+1))\}_{n \geq i}$, $\{d^2(v_i, n(m+1))\}_{n \geq i}$, and $\{d(v_i, n(m+1))\}_{n \geq i}$. \square

Lemma 5.3. *Let $\{\tilde{d}(v_i, n(m+1))\}_{n \geq i}$, be the process of Corollary 5.1, $w \in \mathbb{R}^+$ and let $z(i, w) : \mathbb{N}^* \times \mathbb{R}^+ \rightarrow \mathbb{N}^*$ be a function such that $c(w) := \lim_{i \rightarrow \infty} z(i, w)/i$ exists finite, where $c(w) : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ is an increasing function in w . Let $b \geq 1$ and $w_1 < w_2 < \dots < w_b$ be positive real numbers. Then, the random vector*

$$(\tilde{d}(v_i, (i + z(i, w_1))(m+1)), \dots, \tilde{d}(v_i, (i + z(i, w_b))(m+1)))$$

converges in distribution to $(N_{1/2}^m(\log(1+c(w_1))), \dots, N_{1/2}^m(\log(1+c(w_b))))$ as $i \rightarrow \infty$. Here $N_{1/2}^m(T)$, $T \geq 0$, is the number of individuals at time T in a Yule process with parameter $1/2$ and m initial individuals.

Proof. In order to prove the convergence, we make use of the processes $\{\tilde{d}^1(v_i, n(m+1))\}_{n \geq i}$ and $\{\tilde{d}^2(v_i, n(m+1))\}_{n \geq i}$ and of their behaviour as i goes to infinity. We focus now only on the process $\{\tilde{d}^1(v_i, n(m+1))\}_{n \geq i}$ as the case of $\{\tilde{d}^2(v_i, n(m+1))\}_{n \geq i}$ can be treated analogously.

Let i be fixed and let $T_i^0 = 0$. For every $x \geq 1$ we introduce the times $T_i^x = \sum_{n=i}^{i+x-1} 1/n$. In this way we obtain a partition of $(0, T_i^x]$,

$$(0, T_i^x] = (0, T_i^1] \cup (T_i^1, T_i^2] \cup \dots \cup (T_i^{x-1}, T_i^x]. \quad (5.23)$$

The intervals of this partition have lengths $h_n = 1/n$, $n = i, i+1, \dots, i+x-1$ (see Figure 1).

We introduce the point process $\{\mathfrak{N}^{1,i}(T)\}_{T \geq 0}$, jumping at times T_i^x , $x \geq 1$, and determined by the following rules:

1. At time $T = 0$ the process starts with an initial random number of individuals supported on $\{m, m+1, \dots, 2m\}$ and with distribution only asymptotically degenerate on m , i.e.

$$\mathbb{P}(\mathfrak{N}^{1,i}(0) \neq m) = 1 - \prod_{\ell=2}^{m+1} \left(1 - \frac{1}{2(mi + \ell - 1)}\right) = O(1/i). \quad (5.24)$$

2. The transition probabilities of this point process coincide with (5.16) when $n = i + x - 1$ and $z_n = k$, for every fixed $k \in \mathbb{N}^*$, $k \geq m$, and $x \geq 1$. We write these probabilities here using asymptotic notation. For each $x \geq 1$, let $h_{i,x} = T_i^x - T_i^{x-1} = 1/(i + x - 1)$, then we can write

$$\mathbb{P}[\mathfrak{N}^{1,i}(T_i^x) = k + \ell \mid \mathfrak{N}^{1,i}(T_i^{x-1}) = k] = \begin{cases} \frac{k}{2}h_{i,x} + o(h_{i,x}), & \text{if } \ell = 1, \\ o(h_{i,x}), & \text{if } \ell = m, \\ 1 - \frac{k}{2}h_{i,x} + o(h_{i,x}), & \text{if } \ell = 0, \\ 0, & \text{otherwise.} \end{cases} \quad (5.25)$$

Observe that the sample paths of $\{\mathfrak{N}^{1,i}(T)\}_{T \geq 0}$ and those of $\{\tilde{d}^1(v_i, n(m+1))\}_{n \geq i}$ are non-decreasing right-continuous and integer-valued step functions. However, the lengths of the steps in $\{\tilde{d}^1(v_i, n(m+1))\}_{n \geq i}$ always equal unity, while those of $\{\mathfrak{N}^{1,i}(T)\}_{T \geq 0}$ admit the rational values $h_{i,x}$.

Using the well-known relation $\sum_{n=1}^M 1/n = \log(M) + \gamma + O(1/M)$, where γ is the Euler–Mascheroni constant, we have that

$$T_i^x = \log\left(1 + \frac{x}{i-1}\right) + O(1/i),$$

so, if $z(i, w_\ell) \geq 1$ for $\ell = 1, \dots, b$,

$$T_i^{z(i, w_\ell)} = \log\left(1 + \frac{z(i, w_\ell)}{i-1}\right) + O(1/i) \rightarrow \log(1 + c(w_\ell)), \quad (5.26)$$

as $i \rightarrow \infty$.

Analogously, we introduce the times $\mathcal{T}_i^y = \sum_{n=i}^{i+y-1} 1/(n+1)$, $y \geq 1$, and $\mathcal{T}_i^0 = 0$. We divide $(0, \mathcal{T}_i^y]$ into y disjoint subintervals of lengths $h_{i,y}^* = 1/(i+y)$,

$$(0, \mathcal{T}_i^y] = (0, \mathcal{T}_i^1] \cup (\mathcal{T}_i^1, \mathcal{T}_i^2] \cup \dots \cup (\mathcal{T}_i^{y-1}, \mathcal{T}_i^y].$$

We introduce the point process $\{\mathfrak{N}^{2,i}(T)\}_{T \geq 0}$, jumping at times \mathcal{T}_i^y , $y \geq 1$, and determined by the following properties:

1. This process starts with an initial random number of individuals supported on $\{m, m+1, \dots, 2m\}$ and such that

$$\mathbb{P}(\mathfrak{N}^{2,i}(0) \neq m) = 1 - \prod_{\ell=2}^{m+1} \left(1 - \frac{1}{2(mi + \ell - 1)}\right) = O(1/i). \quad (5.27)$$

2. Its transition probabilities coincide with (5.14) when $n = i + y - 1$ and $x_n = k$, for every fixed $k \in \mathbb{N}^*$, $k \geq m$, and $y \geq 1$. Hence, for each $y \geq 1$, $h_{i,y}^* = \mathcal{T}_i^y - \mathcal{T}_i^{y-1} = 1/(i+y)$, we write

$$\mathbb{P}[\mathfrak{N}^{2,i}(\mathcal{T}_i^y) = k + \ell \mid \mathfrak{N}^{2,i}(\mathcal{T}_i^{y-1}) = k] = \begin{cases} \frac{k}{2}h_{i,y}^* + o(h_{i,y}^*), & \ell = 1, \\ o(h_{i,y}^*), & \ell = 2, \\ 1 - \frac{k}{2}h_{i,y}^* + o(h_{i,y}^*), & \ell = 0, \\ 0, & \text{otherwise.} \end{cases} \quad (5.28)$$

Then we get that $\mathcal{T}_i^y = \log(1 + y/i) + O(1/i)$. Therefore, if $z(i, w_\ell) \geq 1$ for $\ell = 1, \dots, b$,

$$\mathcal{T}_i^{z(i, w_\ell)} = \log\left(1 + \frac{z(i, w_\ell)}{i}\right) + O(1/i) \rightarrow \log(1 + c(w_\ell)), \quad (5.29)$$

as $i \rightarrow \infty$.

Note that $\mathfrak{N}^{1,i}(T_i^x)$ and $\mathfrak{N}^{2,i}(T_i^y)$ have the same law and initial condition as $\tilde{d}^1(v_i, (i+x)(m+1))$ and $\tilde{d}^2(v_i, (i+y)(m+1))$, $x, y \geq 0$, respectively. In addition, by (5.24) and (5.27), these processes start with m initial individuals, as $i \rightarrow \infty$.

We emphasize that by (5.26) and (5.29) both $T_i^{z(i, w_\ell)}$ and $\mathcal{T}_i^{z(i, w_\ell)}$ converge to the same time $T_\ell = \ln(1 + c(w_\ell))$, $\ell = 1, \dots, b$. Moreover, as i increases, (5.25) and (5.28) are closer and closer to the infinitesimal transition probabilities of a Yule process (see (2.2)) with intensity $1/2$. Since the transition probabilities and the initial condition determine uniquely the finite-dimensional distributions of a Markov process, then as $i \rightarrow \infty$, the finite-dimensional distribution of $\mathfrak{N}^{1,i}(T_i^x)$ and $\mathfrak{N}^{2,i}(T_i^y)$, converge to the finite-dimensional distribution of a Yule process with intensity $1/2$. In other words,

$$\left(\mathfrak{N}^{1,i}(T_i^{z(i, w_1)}), \dots, \mathfrak{N}^{1,i}(T_i^{z(i, w_b)}) \right) \rightarrow \left(N_{1/2}^m(T_1), \dots, N_{1/2}^m(T_b) \right), \quad (5.30)$$

in distribution, as $i \rightarrow \infty$, where $N_{1/2}^m(T)$ is the number of individuals of a Yule process at time T with intensity $1/2$ and initial population size equal to m . Analogously,

$$\left(\mathfrak{N}^{2,i}(T_i^{z(i, w_1)}), \dots, \mathfrak{N}^{2,i}(T_i^{z(i, w_b)}) \right) \rightarrow \left(N_{1/2}^m(T_1), \dots, N_{1/2}^m(T_b) \right), \quad (5.31)$$

in distribution. To rigorously prove (5.30) and (5.31) it is enough to focus on (5.30) only, as (5.31) follows in a similar way. Recall that $\{\mathfrak{N}^{1,i}(T)\}_{T \geq 0}$ starts with an initial random integer number of individuals $R \in \{m, m+1, \dots, 2m\}$. Let $\{\mathfrak{N}_m^{1,i}(T)\}_{T \geq 0}$ denote the process $\{\mathfrak{N}^{1,i}(T)\}_{T \geq 0}$ conditioned to $\mathfrak{N}^{1,i}(0) = m$. Set $L_{i,m} = 0$, while for $k > m$ define

$$L_{i,k} = \min\{T \geq L_{i,k-1} : \mathfrak{N}_m^{1,i}(T) > k-1\}. \quad (5.32)$$

Since the jumps of the process $\{\mathfrak{N}_m^{1,i}(T)\}$ are of size either 1 or m , then observe the following: if at time $L_{i,k}$ the jump is of size 1, i.e., if $\mathfrak{N}_m^{1,i}(L_{i,k}) - \mathfrak{N}_m^{1,i}(L_{i,k-1}) = 1$, then $\mathfrak{N}_m^{1,i}(L_{i,k}) = k$. On the other hand, if $\mathfrak{N}_m^{1,i}(L_{i,k}) - \mathfrak{N}_m^{1,i}(L_{i,k-1}) = m$, then it follows that $\mathfrak{N}_m^{1,i}(L_{i,k}) = k + m - 1$, and we have

$$L_{i,k} = L_{i,k+1} = \dots = L_{i,k+m-1}. \quad (5.33)$$

Let now $U_{i,j} = L_{i,j+1} - L_{i,j}$, $j \geq m$. If (5.33) holds,

$$U_{i,k} = U_{i,k+1} = \dots = U_{i,k+m-1} = 0. \quad (5.34)$$

Note that we can write

$$L_{i,k} = \sum_{j=m}^{k-1} U_{i,j}, \quad k > m. \quad (5.35)$$

In addition if we consider only the times $L_{i,k}$, such that $L_{i,k} < L_{i,k+1}$, $k \geq m$, then we can reconstruct $\mathfrak{N}_m^{1,i}(T)$ for every $T \geq 0$:

$$\mathfrak{N}_m^{1,i}(T) = k, \quad \text{if } L_{i,k} \leq T < L_{i,k+1}. \quad (5.36)$$

Now we are going to write the finite-dimensional distributions of the original process $\{\mathfrak{N}^{1,i}(T)\}_{T \geq 0}$. Consider the times $0 = T_0 < T_1 < T_2 < \dots < T_b$, where $T_\ell = \ln(1 + c(w_\ell))$, $\ell = 1, \dots, b$, and let $c \in \mathbb{R}^b$. Taking into account the initial position of the process, we see that the random vector $(\mathfrak{N}^{1,i}(T_1), \dots, \mathfrak{N}^{1,i}(T_b))$ has, over \mathbb{R}^b , the joint distribution

$$\mathbb{P}[(\mathfrak{N}^{1,i}(T_1), \dots, \mathfrak{N}^{1,i}(T_b)) = c] = \mathbb{P}[(\mathfrak{N}_m^{1,i}(T_1), \dots, \mathfrak{N}_m^{1,i}(T_b)) = c] + \varepsilon_1, \quad (5.37)$$

where $\varepsilon_1 \leq O(1/i)$ by (5.24). Let $m = k_0 \leq k_1 \leq \dots \leq k_b \in \mathbb{N}^*$. By the Markov property we obtain

$$\mathbb{P}[\mathfrak{N}_m^{1,i}(T_\ell) = k_\ell, \ell = 1, \dots, b] = \prod_{\ell=1}^b \mathbb{P}[\mathfrak{N}_m^{1,i}(T_\ell) = k_\ell \mid \mathfrak{N}_m^{1,i}(T_{\ell-1}) = k_{\ell-1}], \quad (5.38)$$

where $k_0 = m$. Observe that by (5.35) and (5.36), we can write the conditional probabilities $\mathbb{P}[\mathfrak{N}_m^{1,i}(T_\ell) = k_\ell \mid \mathfrak{N}_m^{1,i}(T_{\ell-1}) = k_{\ell-1}]$ as follows. For $\ell = 1$ we have

$$\begin{aligned} \mathbb{P}[\mathfrak{N}_m^{1,i}(T_1) = k_1 \mid \mathfrak{N}_m^{1,i}(0) = m] &= \mathbb{P}[L_{i,k_1} \leq T_1] - \mathbb{P}[L_{i,k_1+1} \leq T_1] \\ &= \mathbb{P}\left[\sum_{j=m}^{k_1-1} U_{i,j} \leq T_1\right] - \mathbb{P}\left[\sum_{j=m}^{k_1} U_{i,j} \leq T_1\right], \end{aligned} \quad (5.39)$$

while for $\ell = 2, \dots, b$,

$$\begin{aligned} \mathbb{P}[\mathfrak{N}_m^{1,i}(T_\ell) = k_\ell \mid \mathfrak{N}_m^{1,i}(T_{\ell-1}) = k_{\ell-1}] & \\ &= \mathbb{P}[L_{i,k_\ell} \leq T_\ell < L_{i,k_\ell+1} \mid L_{i,k_{\ell-1}} \leq T_{\ell-1} < L_{i,k_{\ell-1}+1}] \\ &= \mathbb{P}[L_{i,k_\ell} - L_{i,k_{\ell-1}+1} \leq T_\ell - T_{\ell-1} < L_{i,k_\ell+1} - L_{i,k_{\ell-1}} \mid \mathfrak{N}_m^{1,i}(T_{\ell-1}) = k_{\ell-1}] \\ &= \mathbb{P}[L_{i,k_\ell} - L_{i,k_{\ell-1}+1} \leq T_\ell - T_{\ell-1} \mid \mathfrak{N}_m^{1,i}(T_{\ell-1}) = k_{\ell-1}] \\ &\quad - \mathbb{P}[L_{i,k_\ell+1} - L_{i,k_{\ell-1}} \leq T_\ell - T_{\ell-1} \mid \mathfrak{N}_m^{1,i}(T_{\ell-1}) = k_{\ell-1}] \\ &= \mathbb{P}\left[\sum_{j=k_{\ell-1}+1}^{k_\ell-1} U_{i,j} \leq T_\ell - T_{\ell-1} \mid \mathfrak{N}_m^{1,i}(T_{\ell-1}) = k_{\ell-1}\right] \\ &\quad - \mathbb{P}\left[\sum_{j=k_{\ell-1}}^{k_\ell} U_{i,j} \leq T_\ell - T_{\ell-1} \mid \mathfrak{N}_m^{1,i}(T_{\ell-1}) = k_{\ell-1}\right]. \end{aligned} \quad (5.40)$$

Observe also that if $\mathfrak{N}_m^{1,i}(T_{\ell-1}) = k_{\ell-1}$, $\ell \geq 1$, then the random variable $U_{i,k_{\ell-1}}$ is strictly positive, while for $j > k_{\ell-1}$, $U_{i,j} \geq 0$. We focus on the limit distribution only of $L_{i,k_1} = \sum_{j=m}^{k_1-1} U_{i,j}$ as the others follow similarly. Recall that the process $\{\mathfrak{N}^{1,i}(T)\}_{T \geq 0}$ jumps at times of the form $T_i^x = \sum_{n=1}^{i+x-1} 1/n$, $x \geq 1$, and let $z(i, w_1)$ be such that $\lim_{i \rightarrow \infty} T_i^{z(i, w_1)} = T_1 = \ln(1 + c(w_1))$, and hence $|T_1 - T_i^{z(i, w_1)}| < 1/(i + z(i, w_1))$, for i large enough. Consider the interval $(0, T_1] \times \dots \times (0, T_1]$ and the partition $(0, T_i^{z(i, w_1)}] \times \dots \times (0, T_i^{z(i, w_1)}]$ in \mathbb{R}^{k_1-m} . Then

$$\begin{aligned} \mathbb{P}\left[\sum_{j=m}^{k_1-1} U_{i,j} \leq T_1\right] &\sim \mathbb{P}[(U_{i,m}, \dots, U_{i,k_1-1}) \in B_{T_i^{z(i, w_1)}}] \\ &= \sum_{B_{T_i^{z(i, w_1)}}} p(u_m, \dots, u_{k_1-1}), \end{aligned} \quad (5.41)$$

where $B_{T_i^{z(i, w_1)}} = \{(u_m, \dots, u_{k_1-1}) : u_m + \dots + u_{k_1-1} \leq T_i^{z(i, w_1)}\}$ and $p(u_m, \dots, u_{k_1-1})$ is the joint probability function of $(U_m, \dots, U_{i,k_1-1})$. By conditioning, the right side of (5.41) can be written as

$$\begin{aligned} \sum_{u_m=1}^{z(i, w_1)} \sum_{u_{m+1}=u_m}^{z(i, w_1)} \dots \sum_{u_{k_1-2}=u_{k_1-3}}^{z(i, w_1)} &\mathbb{P}[U_{i,m} = T_i^{u_m}] \mathbb{P}[U_{i,m+1} = T_i^{u_{m+1}} - T_i^{u_m} \mid U_{i,m} = T_i^{u_m}] \\ &\times \dots \times \mathbb{P}\left[U_{i,k_1-2} = T_i^{u_{k_1-2}} - T_i^{u_{k_1-3}} \mid \sum_{j=m}^{k_1-3} U_{i,j} = T_i^{k_1-3}\right] \\ &\times \mathbb{P}\left[U_{i,k_1-1} \leq T_i^{z(i, w_1)} - T_i^{u_{k_1-2}} \mid \sum_{j=m}^{k_1-2} U_{i,j} = T_i^{k_1-2}\right]. \end{aligned} \quad (5.42)$$

By (5.25) and since the event $\{U_{i,m} = T_i^{u_m}\}$ means that at times $T_i^1, \dots, T_i^{u_m-1}$, the process $\mathfrak{N}_m^{1,i}(T)$

did not jump, but at time $T_i^{u_m}$ it did,

$$\begin{aligned}\mathbb{P}[U_{i,m} = T_i^{u_m}] &= \left[\prod_{\ell=1}^{u_m-1} \left(1 - \frac{m}{2} h_{i,\ell} + o(h_{i,\ell}) \right) \right] \left(\frac{m}{2} h_{i,u_m} + o(h_{i,u_m}) \right) \\ &= \frac{m}{2} h_{i,u_m} \exp \left(\sum_{\ell=1}^{u_m-1} \ln \left(1 - \frac{m}{2} h_{i,\ell} + o(h_{i,\ell}) \right) \right) + o(h_{i,u_m}).\end{aligned}\quad (5.43)$$

Using Taylor expansion for $\ln(1-y)$ and e^y , and rearranging the terms we obtain that

$$\mathbb{P}[U_{i,m} = T_i^{u_m}] = \frac{m}{2} h_{i,u_m} \exp \left(-\frac{m}{2} T_i^{u_m} \right) + \text{Err}_1, \quad (5.44)$$

where $\text{Err}_1 \leq O(h_{i,u_m}^2)$. Next, we calculate $\mathbb{P}[U_{i,j} = T_i^{u_j} - T_i^{u_{j-1}} \mid L_{i,j} = T_i^{u_{j-1}}]$, $j = m+1, \dots, k_1-1$. Let $J_{i,j}$ be the size of the jump at time $L_{i,j}$. By (5.33) and (5.34), if $J_{i,j} > 1$, then $U_{i,j} = 0$, otherwise, if $J_{i,j} = 1$ then $U_{i,j} > 0$. In addition, by (5.25) and (5.32) we have that for every $x \geq 1$,

$$\mathbb{P}[J_{i,j} > 1 \mid L_{i,j} = T_i^x] = \mathbb{P}[\mathfrak{N}_m^{1,i}(T_i^x) = j + m - 1 \mid \mathfrak{N}_m^{1,i}(T_i^{x-1}) = j - 1] = o(h_{i,x}). \quad (5.45)$$

Thus, conditioning on $J_{i,j}$ and using (5.45) we get

$$\begin{aligned}\mathbb{P}[U_{i,j} = T_i^{u_j} - T_i^{u_{j-1}} \mid L_{i,j} = T_i^{u_{j-1}}] \\ = \mathbb{P}[U_{i,j} = T_i^{u_j} - T_i^{u_{j-1}} \mid L_{i,j} = T_i^{u_{j-1}}, J_{i,j} = 1] + \text{Err}_2,\end{aligned}\quad (5.46)$$

where $\text{Err}_2 \leq o(h_{i,u_j})$. Observe that the conditional event $\{U_{i,j} = T_i^{u_j} - T_i^{u_{j-1}} \mid L_{i,j} = T_i^{u_{j-1}}, J_{i,j} = 1\}$ means that at times $T_i^{u_{j-1}+1}, T_i^{u_{j-1}+2}, \dots, T_i^{u_j-1}$, the process $\mathfrak{N}_m^{1,i}(T)$ did not jump, but at time $T_i^{u_j}$ it did. Therefore, by (5.25) and applying similar arguments to those used to obtain (5.43) and (5.44) we have

$$\begin{aligned}\mathbb{P}[U_{i,j} = T_i^{u_j} - T_i^{u_{j-1}} \mid L_{i,j} = T_i^{u_{j-1}}, J_{i,j} = 1] \\ = \left[\prod_{\ell=u_{j-1}+1}^{u_j-1} \left(1 - \frac{j}{2} h_{i,\ell} + o(h_{i,\ell}) \right) \right] \left(\frac{j}{2} h_{i,u_j} + o(h_{i,u_j}) \right) \\ = \frac{j}{2} h_{i,u_j} \exp \left[-\frac{j}{2} (T_i^{u_j} - T_i^{u_{j-1}}) \right] + \text{Err}_3,\end{aligned}\quad (5.47)$$

where $\text{Err}_3 \leq O(h_{i,u_j}^2)$. Thus, by (5.46) and (5.47)

$$\mathbb{P}[U_{i,j} = T_i^{u_j} - T_i^{u_{j-1}} \mid L_{i,j} = T_i^{u_{j-1}}] = \frac{j}{2} h_{i,u_j} \exp \left[-\frac{j}{2} (T_i^{u_j} - T_i^{u_{j-1}}) \right] + \text{Err}_4, \quad (5.48)$$

where $\text{Err}_4 \leq O(h_{i,u_j}^2)$. Following these same steps we can also find that

$$\mathbb{P}[U_{i,j} > T_i^{u_j} - T_i^{u_{j-1}} \mid L_{i,j} = T_i^{u_{j-1}}] = \exp \left[-\frac{j}{2} (T_i^{u_j} - T_i^{u_{j-1}}) \right] + \text{Err}_5, \quad (5.49)$$

where $\text{Err}_5 \leq O(h_{i,u_j}^2)$. Now, by substituting (5.44), (5.48) and (5.49) in (5.42) we arrive at

$$\begin{aligned}\mathbb{P} \left[\sum_{j=m}^{k_1-1} U_{i,j} \leq T_1 \right] &\sim \sum_{u_m=1}^{z(i,w_1)} \sum_{u_{m+1}=1}^{z(i,w_1)} \dots \sum_{u_{k_1-2}=1}^{z(i,w_1)} h_{i,u_m} h_{i,u_{m+1}} \frac{m}{2} \exp \left[-\frac{m}{2} T_i^{u_m} \right] \\ &\times \frac{(m+1)}{2} \exp \left[\frac{-(m+1)}{2} (T_i^{u_{m+1}} - T_i^{u_m}) \right] \mathbb{1}_{\{T_i^{u_{m+1}} \geq T_i^{u_m}\}} \\ &\times \dots \times h_{i,u_{k_1-2}} \frac{(k_1-2)}{2} \exp \left[\frac{-(k_1-2)}{2} (T_i^{u_{k_1-2}} - T_i^{u_{k_1-3}}) \right] \\ &\times \left[1 - \exp \left(\frac{-(k_1-1)}{2} (T_i^{z(i,w_1)} - T_i^{u_{k_1-2}}) \right) \right] \mathbb{1}_{\{T_i^{z(i,w_1)} \geq T_i^{u_{k_1-2}}\}} + \text{Err}_6,\end{aligned}\quad (5.50)$$

where $\text{Err}_6 \leq O\left(h_{i,z(i,w_1)} \prod_{\ell=m}^{k_1-2} h_{i,\ell}\right)$. Taking the limit as $i \rightarrow \infty$, the right-hand side of (5.50) becomes

$$\begin{aligned} & \int_0^{T_1} \int_{y_{m+1}}^{T_1} \cdots \int_{y_{k_1-2}}^{T_1} \frac{m(m+1)}{2} \frac{\exp\left(-\frac{m}{2}y_m\right) \exp\left(-\frac{m+1}{2}(y_{m+1}-y_m)\right)}{2} \\ & \quad \times \cdots \times \frac{(k_1-2)}{2} \exp\left(-\frac{k_1-2}{2}(y_{k_1-2}-y_{k_1-3})\right) \\ & \quad \times \left[1 - \exp\left(-\frac{k_1-1}{2}(T_1-y_{k_1-2})\right)\right] dy_{k_1-2} \cdots dy_m, \end{aligned} \quad (5.51)$$

as the right side of (5.50) is the Riemann sum for the above definite integral. We observe that this integral also corresponds to $\mathbb{P}[Y_m + \cdots + Y_{k_1-1} \leq T_1]$, where the random variables $Y_j, j = 1, \dots, k_1-1$, are independent and exponentially distributed with parameter $j/2$, respectively. Therefore,

$$\lim_{i \rightarrow \infty} \mathbb{P}\left[\sum_{j=m}^{k_1-1} U_{i,j} \leq T_1\right] = \mathbb{P}[Y_m + \cdots + Y_{k_1-1} \leq T_1], \quad (5.52)$$

and in an analogous way we also obtain

$$\lim_{i \rightarrow \infty} \mathbb{P}\left[\sum_{j=m}^{k_1} U_{i,j} \leq T_1\right] = \mathbb{P}[Y_m + \cdots + Y_{k_1} \leq T_1]. \quad (5.53)$$

From this and (5.39) we get

$$\begin{aligned} \lim_{i \rightarrow \infty} \mathbb{P}[\mathfrak{N}_{1,i}(T_1) = k_1 \mid \mathfrak{N}_{1,i}(0) = m] &= \lim_{i \rightarrow \infty} \mathbb{P}[\mathfrak{N}_m^{1,i}(T_1) = k_1] \\ &= \mathbb{P}\left[\sum_{j=m}^{k_1-1} Y_j \leq T_1\right] - \mathbb{P}\left[\sum_{j=m}^{k_1} Y_j \leq T_1\right] \\ &= \mathbb{P}\left[\sum_{j=m}^{k_1-1} Y_j \leq T_1 < \sum_{j=m}^{k_1} Y_j\right]. \end{aligned} \quad (5.54)$$

In a Yule process with parameter λ and starting with m initial individuals, the interarrival or sojourn times are independent random variables exponentially distributed with parameter $\lambda j, j \geq m$, respectively. Thus, (5.54) corresponds to

$$\mathbb{P}[N_{1/2}^m(T_1) = k_1 \mid N_{1/2}^m(0) = m], \quad (5.55)$$

where $N_{1/2}^m(T)$ is the number of individuals in a Yule process with parameter $1/2$ and starting with m initial individuals. Following analogous steps from (5.39) to (5.55) we also find that for $\ell = 2, \dots, b$,

$$\lim_{i \rightarrow \infty} \mathbb{P}[\mathfrak{N}_{1,i}(T_\ell) = k_\ell \mid \mathfrak{N}_{1,i}(T_{\ell-1}) = k_{\ell-1}] = \mathbb{P}[N_{1/2}^m(T_\ell) = k_\ell \mid N_{1/2}^m(T_{\ell-1}) = k_{\ell-1}]. \quad (5.56)$$

Consequently, from (5.37), (5.38), (5.54), (5.55), (5.56), and since the Yule process is Markov, we conclude that

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P}[\mathfrak{N}_m^{1,i}(T_\ell) = k_\ell, \ell = 1, \dots, b] &= \prod_{\ell=1}^b \mathbb{P}[N_{1/2}^m(T_\ell) = k_\ell \mid N_{1/2}^m(T_{\ell-1}) = k_{\ell-1}] \\ &= \mathbb{P}[N_{1/2}^m(T_\ell) = k_\ell, \ell = 1, \dots, b]. \end{aligned} \quad (5.57)$$

Now, once proven the convergence to the Yule process of intensity $1/2$, we immediately have that

$$\left(\bar{d}^1(v_i, (i + z(i, w_\ell))(m+1)), \ell = 1, \dots, b\right) \rightarrow \left(N_{1/2}^m(T_\ell), \ell = 1, \dots, b\right), \quad (5.58)$$

and

$$(\tilde{d}^2(v_i, (i + z(i, w_\ell))(m + 1)), \ell = 1, \dots, b) \rightarrow (N_{1/2}^m(T_\ell), \ell = 1, \dots, b), \quad (5.59)$$

in distribution, as $i \rightarrow \infty$.

Observe that at time $n(m + 1)$, $n \geq i$, by (5.22) the random variables $\tilde{d}^1(v_i, n(m + 1))$, $\tilde{d}(v_i, n(m + 1))$ and $\tilde{d}^2(v_i, n(m + 1))$ are almost surely ordered, that is

$$\begin{aligned} \mathbb{P}[\tilde{d}^2(v_i, (i + z(i, w_\ell))(m + 1)) \leq \tilde{d}(v_i, (i + z(i, w_\ell))(m + 1)) \\ \leq \tilde{d}^1(v_i, (i + z(i, w_\ell))(m + 1)), \ell = 1, \dots, b] = 1. \end{aligned} \quad (5.60)$$

This implies that for $k \geq m$,

$$\begin{aligned} \mathbb{P}(\tilde{d}^1(v_i, (i + z(i, w_\ell))(m + 1)) \leq k_\ell, \ell = 1, \dots, b) \\ \leq \mathbb{P}(\tilde{d}(v_i, (i + z(i, w_\ell))(m + 1)) \leq k_\ell, \ell = 1, \dots, b) \\ \leq \mathbb{P}(\tilde{d}^2(v_i, (i + z(i, w_\ell))(m + 1)) \leq k_\ell, \ell = 1, \dots, b). \end{aligned} \quad (5.61)$$

Thus, from (5.58), (5.59) and (5.61) we obtain the convergence in distribution of the random vector

$$(\tilde{d}(v_i, (i + z(i, w_\ell))(m + 1)), \ell = 1, \dots, b) \rightarrow (N_{1/2}^m(T_\ell), \ell = 1, \dots, b), \quad (5.62)$$

as $i \rightarrow \infty$. \square

Proof of Theorem 3.1. Using Lemma 5.1, Corollary 5.1 and Lemma 5.3 we obtain the convergence to the b -finite-dimensional distributions of a Yule process, for all $b \geq 1$. To obtain the exact formula we make use of the independence of the increments and of the distribution of the number of individuals in a Yule process with k_ℓ initial progenitors, $\ell = 0, \dots, b$. Thus,

$$\begin{aligned} \mathbb{P}[N_{1/2}^m(\log(1 + c(w_1))) = k_1, \dots, N_{1/2}^m(\log(1 + c(w_b))) = k_b] \\ = \prod_{\ell=1}^b \mathbb{P}\left(N_{1/2}^{k_{\ell-1}}\left(\log\left(\frac{1 + c(w_\ell)}{1 + c(w_{\ell-1})}\right)\right) = k_\ell\right). \end{aligned}$$

Finally, we use equation (3.5) in [15], Section XVII.3. \square

5.2 A lemma and the proof of Theorem 3.2

In this section we make use of the planted model described in Section 4. Notice that the BA random graph model corresponds to the case in which all the random variables M_j , $j \geq 1$, are concentrated on m , so that $\mathfrak{T}_n = n(m + 1)$ almost surely.

Formally, let $(G_m^t)_{t \geq 1}$ be the random graph process defining the BA model as in subsection 2.1. For each $1 \leq j \leq i$ consider the birth processes in discrete time $\{b(v_j, n(m + 1))\}_{n \geq i}$, with state space given by \mathbb{N}^* and determined by the transition probabilities

$$\mathbb{P}[b(v_j, (n + 1)(m + 1)) = k + \ell \mid b(v_j, n(m + 1)) = k] = \begin{cases} \frac{k}{n}, & \ell = 1, \\ 0, & \text{otherwise,} \end{cases} \quad (5.63)$$

and initial condition $b(v_j, i(m + 1)) = 1$ almost surely. Recall that i is taken large so that Lemma 5.1 holds.

Lemma 5.4. *Let $z(i, w) : \mathbb{N}^* \times \mathbb{R}^+ \rightarrow \mathbb{N}$ be a function such that $c(w) := \lim_{i \rightarrow \infty} z(i, w)/i$ exists finite, where $c(w) : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ is an increasing function in w , and let $w_1 < \dots < w_b$, $b \in \mathbb{N}^*$, be positive real numbers. For every $1 \leq j \leq i$ we have*

$$(b(v_j, (i + z(i, w_\ell))(m + 1)), \ell = 1, \dots, b) \rightarrow (N_1^1(\log(1 + c(w_\ell))), \ell = 1, \dots, b) \quad (5.64)$$

in distribution as $i \rightarrow \infty$, where $N_1^1(T)$ is the number of individuals of a Yule process at time T , with one initial individual and parameter 1.

Proof. For each $1 \leq j \leq i$, we prove convergence in distribution in the same way as we did in the proof of Lemma 5.3 for $\mathfrak{N}^{1,i}(T_i^x)$, but now with $\mathfrak{N}^1(0) = 1$, i.e. the process starts with only one individual, and transition probabilities given by (5.63). Therefore, as $i \rightarrow \infty$, the probabilities (5.63) become the infinitesimal transition probabilities of a Yule process with intensity 1, starting with one individual. Since the process is Markov, the transition probabilities and the initial condition determine uniquely the finite-dimensional distributions. \square

Remark 5.2. *To prove the first part of Theorem 3.2 we will make use of the result of Theorem 3.1. The idea is to take $n := n(i, w)$, a function of i and a positive real number w , such that, $i/n(i, w) \rightarrow 1/(1 + c(w))$ as $i \rightarrow \infty$, with $c(w)$ as in Theorem 3.1. Thus, $\lim_{w \rightarrow \infty} \lim_{i \rightarrow \infty} i/n(i, w) = 0$.*

Proof of Theorem 3.2. We start proving (3.3). Consider the BA model at time $t = n(m + 1)$, $n \geq i$, and the planted model of Section 4. Recall that in the planted model we have i discrete-time birth processes $\{b(v_j, n(m + 1))\}_{n \geq i}$, $j = 1, \dots, i$, which are exchangeable. By Theorem 4.1, the event of choosing a vertex uniformly at random in the BA model is equivalent to that of selecting first uniformly at random one of the i processes $\{b(v_j, n(m + 1))\}_{n \geq i}$, $j = 1, \dots, i$, and then choosing uniformly at random a vertex belonging to it. Therefore, the degree of V_t can be studied through the analysis of the degree of a random vertex chosen with uniform probability between the vertices in any of the i processes $\{b(v_j, n(m + 1))\}_{n \geq i}$, $j = 1, \dots, i$. Let V_t^j be a vertex chosen uniformly at random from the vertices in the j -th process $\{b(v_j, n(m + 1))\}_{n \geq i}$, and let $\epsilon(i, n)$ be a function we will use to measure the error. Using the notation of Section 4.0.1, where W denotes the index of the birth process chosen and Y_j is a random variable taking values in $\{1, 2, \dots, n - i + 1\}$ denoting the number of vertices in $b(v_j, n(m + 1))$, we have

$$\begin{aligned} \mathbb{P}[d(V_t) = k] &= \sum_{j=1}^i \mathbb{P}[d(V_t^j) = k, V_t^j \neq v_j, W = j] \\ &\quad + \sum_{j=1}^i \mathbb{P}[d(V_t^j) = k, V_t^j = v_j, W = j] \\ &= \mathbb{P}[d(V_t^1) = k \mid V_t^1 \neq v_1, W = 1] \sum_{j=1}^i \mathbb{P}[V_t^j \neq v_j, W = j] \\ &\quad + \sum_{j=1}^i \mathbb{P}[d(V_t^j) = k \mid V_t^j = v_j, W = j] \mathbb{P}[V_t^j = v_j, W = j] \\ &= \mathbb{P}[d(V_t^1) = k \mid V_t^1 \neq v_1, W = 1] + \epsilon(i, n). \end{aligned} \tag{5.65}$$

The last two equalities are obtained by considering the following two observations. First, permuting the labels of the i birth processes $\{b(v_j, n(m + 1))\}_{n \geq i}$, $j = 1, \dots, i$, will not change the distribution of the process of the new vertices and their degrees, thus for $j = 1, \dots, i$, we can write $\mathbb{P}[d(V_t^j) = k \mid V_t^j \neq v_j, W = j] = \mathbb{P}[d(V_t^1) = k \mid V_t^1 \neq v_1, W = 1]$. Second,

$$\begin{aligned} \sum_{j=1}^i \mathbb{P}[d(V_t^j) = k \mid V_t^j = v_j, W = j] \mathbb{P}[V_t^j = v_j, W = j] &\leq \sum_{j=1}^i \mathbb{P}[V_t^j = v_j, W = j] \\ &= \sum_{j=1}^i \sum_{\ell=1}^{n-i+1} \frac{1}{\ell} \frac{\ell}{n} \mathbb{P}(Y_j = \ell) = \frac{i}{n}, \end{aligned} \tag{5.66}$$

that is, $\epsilon(i, n) = O(i/n)$.

Note that the degree of the planted vertices behaves differently as they have appeared in the very early history of the graph evolution. Also, in the limit, the number of planted vertices becomes negligible compared to the total size of the graph.

Now take $n(i, w) = i + z(i, w)$, where $z(i, w)$ is defined as in Lemma 5.3, Lemma 5.4 and Theorem 3.1. As $i \rightarrow \infty$,

- by Lemma 5.4 we have that $b(v_1, n(m+1))$, converges in distribution to the size of a Yule process evaluated at time $T = \log(1 + c(w))$, with intensity 1 and starting with one initial individual;
- by Lemma 5.3, the degree of each vertex belonging to $\{b(v_1, n(m+1))\}$, given that it is different to v_1 , converges in distribution to the size of a Yule process with intensity $1/2$ and m initial individuals.

The above Yule processes describe an m -Yule model $\{Y_{1/2,1}^m(T)\}_{T \geq 0}$ of parameters $\lambda = 1/2$ and $\beta = 1$. For $i \rightarrow \infty$, the degree of V_t^1 given that $V_t^1 \neq v_1$, converges in distribution to the size of a genus chosen uniformly at random in the m -Yule model at time $T = \log(1 + c(w))$, given in turn that such a random genus is different to the first genus appeared, g_1 . Thus, if \mathcal{N}_T^m denotes the size of a genus G_T chosen uniformly at random at time T in $\{Y_{1/2,1}^m(T)\}$,

$$\lim_{i \rightarrow \infty} \mathbb{P}(d(V_t^1) = k \mid V_t^1 \neq v_1, W = 1) = \mathbb{P}(\mathcal{N}_{\log(1+c(w))}^m = k \mid G_{\log(1+c(w))} \neq g_1). \quad (5.67)$$

By (5.65) and (5.67),

$$\lim_{i \rightarrow \infty} \mathbb{P}(d(V_i) = k) = \mathbb{P}(\mathcal{N}_{\log(1+c(w))}^m = k \mid G_{\log(1+c(w))} \neq g_1) + \varepsilon(w), \quad (5.68)$$

where $\varepsilon(w) = O(1/(1+c(w)))$. Since $c(w)$ is an increasing function and a Yule process is supercritical, then

$$\lim_{w \rightarrow \infty} \mathbb{P}(\mathcal{N}_{\log(1+c(w))}^m = k) = \lim_{w \rightarrow \infty} \mathbb{P}(\mathcal{N}_{\log(1+c(w))}^m = k \mid G_{\log(1+c(w))} \neq g_1). \quad (5.69)$$

Therefore, by (5.68) and (5.69),

$$\lim_{w \rightarrow \infty} \lim_{i \rightarrow \infty} \mathbb{P}(d(V_i) = k) = \lim_{w \rightarrow \infty} \mathbb{P}(\mathcal{N}_{\log(1+c(w))}^m = k). \quad (5.70)$$

To prove (3.4) note that

$$\mathbb{E}\left(\frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\{d(v_i, t) = k\}}\right) = \frac{\mathbb{E}N_{k,t}}{n} = \frac{1}{n} \sum_{i=1}^n \mathbb{P}(d(v_i, t) = k) = \mathbb{P}(d(V_t) = k).$$

Let \mathcal{F}_t be the natural filtration generated by the process $\{N_{k,t}\}_{t \geq 1}$ up to time t , and define $Z_s = \mathbb{E}(N_{k,t} \mid \mathcal{F}_s)$. Observe that Z_s is a martingale as $\mathbb{E}[\mathbb{E}(N_{k,t} \mid \mathcal{F}_s) \mid \mathcal{F}_r] = \mathbb{E}(N_{k,t} \mid \mathcal{F}_r)$, for $r \leq s \leq t$. Considering that at each time interval $(s-1, s]$ a new vertex v_s appears and m directed edges from it are attached to existing vertices, then v_s is attached to at most m different vertices, say v^1, \dots, v^m . This does not affect neither the degree of the other existing vertices $w \neq v^1, \dots, v^m$, nor the attachment probabilities related to them. Thus, it follows that $|Z_s - Z_{s-1}| \leq 2m$. Since $Z_t = N_{k,t}$ and $Z_0 = \mathbb{E}N_{k,t}$, then by taking $x = C\sqrt{t \log t}$, with $C > m\sqrt{8}$ and applying the Azuma–Hoeffding inequality (see Lemma 4.1.3 in [13]), we obtain

$$\mathbb{P}\left(\left|\frac{N_{k,t}}{n} - \frac{\mathbb{E}N_{k,t}}{n}\right| > C\sqrt{\frac{(m+1) \log(n(m+1))}{n}}\right) \leq o\left(\frac{1}{n}\right). \quad (5.71)$$

Now observe that $N_{k,t} = 0$ when $k \geq n(m+1)$, $n \geq 1$. Therefore,

$$\begin{aligned} & \mathbb{P}\left(\max_k \left|\frac{N_{k,t}}{n} - \frac{\mathbb{E}N_{k,t}}{n}\right| > C\sqrt{\frac{(m+1) \log(n(m+1))}{n}}\right) \\ &= \mathbb{P}\left(\max_{k < n(m+1)} \left|\frac{N_{k,t}}{n} - \frac{\mathbb{E}N_{k,t}}{n}\right| > C\sqrt{\frac{(m+1) \log(n(m+1))}{n}}\right) \\ &\leq \sum_{k=1}^{n(m+1)-1} \mathbb{P}\left(\max_{k < t} \left|\frac{N_{k,t}}{n} - \frac{\mathbb{E}N_{k,t}}{n}\right| > C\sqrt{\frac{(m+1) \log(n(m+1))}{n}}\right). \end{aligned} \quad (5.72)$$

Concluding, by (5.71) we get the desired result. \square

5.3 Proof of Proposition 3.1

Proof. Let us consider an m -Yule model $\{Y_{1/2,1}^m(T)\}_{T \geq 0}$. It is known that by conditioning on the number of genera present at time T , the random times at which novel genera appear are distributed as the order statistics of i.i.d. random variables distributed with distribution function given by (see e.g. [25] or [26] and the references therein)

$$\mathbb{P}(\mathcal{T} \leq \tau) = \frac{e^\tau - 1}{e^T - 1}, \quad 0 \leq \tau \leq T. \quad (5.73)$$

As above, let \mathcal{N}_T^m denote the size of a genus chosen uniformly at random at time T . Then, for every $k \geq m$ and recalling the distribution of a Yule process starting with m initial individuals,

$$\begin{aligned} \mathbb{P}(\mathcal{N}_T^m = k) &= \int_0^T \mathbb{P}(N_{1/2}^m(T) = k \mid N_{1/2}^m(\tau) = m) \mathbb{P}(\mathcal{T} \in d\tau) \\ &= \int_0^T \binom{k-1}{m-1} e^{-m \frac{T-\tau}{2}} (1 - e^{-\frac{T-\tau}{2}})^{k-m} \frac{e^\tau}{e^T - 1} d\tau \\ &= \frac{1}{1 - e^{-T}} \int_0^T \binom{k-1}{m-1} e^{-y} e^{-m \frac{y}{2}} (1 - e^{-\frac{y}{2}})^{k-m} dy. \end{aligned} \quad (5.74)$$

By letting $z = 1 - e^{-\frac{y}{2}}$, we can write (5.74) as

$$\mathbb{P}(\mathcal{N}_T^m = k) = \frac{2}{1 - e^{-T}} \int_0^{1 - e^{-\frac{T}{2}}} \binom{k-1}{m-1} z^{k-m} (1 - z)^{m+1} dz. \quad (5.75)$$

Our interest is in the asymptotic behaviour when $T \rightarrow \infty$. In this case (5.75) reduces to

$$\begin{aligned} \lim_{T \rightarrow \infty} \mathbb{P}(\mathcal{N}_T^m = k) &= 2 \binom{k-1}{m-1} B(k - m + 1, m + 2) \\ &= m(m + 1) B(k, 3), \end{aligned} \quad (5.76)$$

where $B(a, b)$ denotes the Beta function. □

Acknowledgments

The authors would like to thank the anonymous Referee for the valuable suggestions which improved both the content and presentation of the paper.

F. Polito and L. Sacerdote have been supported by the projects *Memory in Evolving Graphs* (Compagnia di San Paolo/Università di Torino) and by INDAM (GNAMPA/GNCS).

References

- [1] K. B. Athreya. Preferential attachment random graphs with general weight function. *Internet Math.*, 4(4):401–418, 2007.
- [2] A. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.
- [3] S. Bhamidi. *Universal techniques to analyze preferential attachment trees: Global and Local analysis*. Available at www.unc.edu/~bhamidi/preferent.pdf, 2007.
- [4] M. Boguña and A. Serrano. Simulating non-Markovian stochastic processes. *Phys. Rev. E*, 90(4):042108, 2014.
- [5] B. Bollobás, O. Riordan, J. Spencer, and G. Tusnády. The Degree Sequence of a Scale-free Random Graph Process. *Random Struct. Algorithms*, 18(3):279–290, 2001.
- [6] C. Borgs, J. Chayes, C. Daskalakis, and S. Roch. First to market is not everything: An analysis of preferential attachment with fitness. In *Proceedings of the Thirty-ninth Annual ACM Symposium on Theory of Computing*, pages 135–144, New York, NY, USA, 2007. ACM.

- [7] A. Clauset, C. R. Shalizi, and M. E. J. Newman. Power-Law Distributions in Empirical Data. *SIAM Review*, 51(4):661–703, 2009.
- [8] C. Cooper and A. Frieze. A General Model of Web Graphs. *Random Struct. Algorithms*, 22(3):311–335, 2003.
- [9] D. J. de Solla Price. Networks of scientific papers. *Science*, 149:510–515, 1965.
- [10] D. J. de Solla Price. A general theory of bibliometric and other cumulative advantage processes. *J. Amer. Soc. Inform. Sci.*, 27:292–306, 1976.
- [11] M. Deijfen, H. van den Esker, R. van der Hofstad, and G. Hooghiemst. A preferential attachment model with random initial degrees. *Ark. Mat*, 47(1):41–72, 2009.
- [12] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin. Structure of Growing Networks with Preferential Linking. *Phys. Rev. Lett.*, 85:4633–4636, 2000.
- [13] R. Durrett. *Random Graph Dynamics*. Cambridge University Press, New York, NY, USA, 2006.
- [14] M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the internet topology. *SIGCOMM Comput. Commun. Rev.*, 29(4):251–262, 1999.
- [15] W. Feller. *An Introduction to Probability Theory and its Applications*, volume 1. Wiley, third edition, 1968.
- [16] A. Garavaglia, R. van der Hofstad, and G. Woeginger. The dynamics of power laws: Fitness and aging in preferential attachment trees. *Journal of Statistical Physics*, 168(6):1137–1179, 2017.
- [17] R. L. Graham, D. E. Knuth, and O. Patashnik. *Concrete mathematics. A foundation for computer science*. Addison-Wesley Publishing Company, Reading, MA, 1994.
- [18] N. Johnson and S. Kotz. *Urn Models and Their Application: An Approach to Modern Discrete Probability Theory*. John Wiley & Sons Inc, New York, NY, USA, 1977.
- [19] T. Kamae, U. Krengel, and G. O’Brien. Stochastic inequalities on partially ordered spaces. *Ann. Probab.*, 5(6):899–912, 1977.
- [20] A. Klaus, S. Yu, and D. Plenz. Statistical Analyses Support Power Law Distributions Found in Neuronal Avalanches. *PLoS ONE*, 6:19779, May 2011.
- [21] J. Kleinberg, R. Kumar, P. Raghavan, S. Rajagopalan, and A. S. Tomkins. The web as a graph: Measurements, models, and methods. In *Proceedings of the 5th Annual International Conference on Computing and Combinatorics, COCOON’99*, pages 1–17, Berlin, Heidelberg, 1999. Springer-Verlag.
- [22] P. L. Krapivsky and S. Redner. Organization of growing random networks. *Phys. Rev. E*, 63:066123, May 2001.
- [23] P. L. Krapivsky and S. Redner. Finiteness and fluctuations in growing networks. *Journal of Physics A: Mathematical and General*, 35(45):9517, 2002.
- [24] P. L. Krapivsky, S. Redner, and F. Leyvraz. Connectivity of growing random networks. *Phys. Rev. Lett.*, 85:4629–4632, Nov 2000.
- [25] P. Lansky, F. Polito, and L. Sacerdote. The role of detachment of in-links in scale-free networks. *Journal of Physics A: Mathematical and Theoretical*, 47(34):345002, 2014.
- [26] P. Lansky, F. Polito, and L. Sacerdote. Generalized nonlinear Yule models. *J. Stat. Phys.*, 165(3):661–679, 2016.
- [27] M. E. J. Newman. The structure and function of complex networks. *SIAM REVIEW*, 45:167–256, 2003.
- [28] A. Pachon, F. Polito, and L. Sacerdote. Random graphs associated to some discrete and continuous time preferential attachment models. *Journal of Statistical Physics*, 162(6):1608–1638, 2016.

- [29] A. Pachon, L. Sacerdote, and S. Yang. Scale-free behavior of networks with the copresence of preferential and uniform attachment rules. *Phys. D*, 371:1–12, 2018.
- [30] A. Rudas, B. Tóth, and B. Valkó. Random trees and general branching processes. *Random Struct. Algorithms*, 31(2):186–202, 2007.
- [31] H. A. Simon. On a class of skew distribution functions. *Biometrika*, 42(3–4):425–440, 1955.
- [32] R. van der Hofstad. *Random Graphs and Complex Networks*., volume 1. Cambridge University Press, Cambridge, 11 2016.
- [33] G. U. Yule. A Mathematical Theory of Evolution, Based on the Conclusions of Dr. J. C. Willis, F.R.S. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 213(402-410):21–87, 1925.