Random Graphs Associated to some Discrete and Continuous Time Preferential Attachment Models

Angelica Pachon, Federico Polito & Laura Sacerdote
Mathematics Department “G. Peano”, University of Torino, Italy
March 23, 2015

Abstract

We give a common description of Simon, Barabási–Albert, and II-PA growth models, by introducing suitable random graph processes with preferential attachment mechanisms. Through the II-PA model, we prove the conditions for which the asymptotic degree distribution of the Barabási–Albert model coincides with the asymptotic in-degree distribution of the Simon model. Furthermore, we show that when the number of vertices in the Simon model (with parameter \( \alpha \)) goes to infinity, it behaves as a Yule model with parameters \((\lambda, \beta) = (1 - \alpha, 1)\). As a byproduct of our analysis, we prove the explicit expression of the in-degree distribution for the II-PA model, given without proof in [12].

Keywords: Preferential attachment; Random graph growth; Discrete and continuous time models; Stochastic processes.

MSC2010: 05C80, 90B15.

1 Introduction

A large group of networks growth models can be classified as preferential attachment models. In the simplest preferential attachment mechanism an edge connects a newly created node to one of those already present in the network with a probability proportional to the number of their edges.

Typically what is analyzed for these models are properties related both to the growth of the number of edges for each node and to the growth of the number of nodes.

After the seminal paper by Barabási and Albert [11], models admitting a preferential attachment mechanism have been successfully applied to the growth of different real world networks, such as, amongst others, physical, biological or social networks. The typical feature revealing a preferential attachment growth mechanism is the presence of power-law distributions, e.g., for the degree (or in-degree) of a node selected uniformly at random.

Despite its present success, the preferential attachment paradigm is not new. In fact it dates back to a paper by Udny Yule [13], published in 1925 and regarding the development of a theory of macroevolution. Specifically the study concerned the time-continuous process of creation of genera and the evolution of species belonging to them. Yule proved that when time goes to infinity, the limit distribution of the number of species in a genus selected uniformly at random has a specific form and exhibits a power-law behavior in its tail. Thirty years later, the Nobel laureate Herbert A. Simon proposed a time-discrete preferential attachment model to describe the appearance of new words in a large piece of a text. Interestingly enough, the limit distribution of the number of occurrences of each word, when the number of words diverges, coincides with that of the number of species belonging to the randomly chosen genus in the Yule model, for a specific choice of the parameters. This fact explains the designation Yule–Simon distribution that is commonly assigned to that limit distribution.

Furthermore, it should be noticed that Barabási–Albert model exhibits an asymptotic degree distribution that equals the Yule-Simon distribution in correspondence of a specific choice of the parameters and still presents power-law characteristics for more general choices of the parameters. The same happens also for other preferential attachment models.
Yule, Simon and Barabási–Albert models share the preferential attachment paradigm that seems to play an important role in the explanation of the scale-freeness of real networks. However, the mathematical tools classically used in their analysis are different. This makes difficult to understand in which sense models producing very similar asymptotic distributions are actually related one another. Although often remarked and heuristically justified, no rigorous proofs exist clarifying conditions for such result. Different researchers from different disciplines, for example theoretical physicists and economists asked themselves about the relations between Simon, Barabási–Albert, Yule and also some other models closely related to these first three (sometimes confused in the literature under one of the previous names). Partial studies in this direction exist but there is still a lack of clarifying rigorous results that would avoid errors and would facilitate the extension of the models.

The existing results refer to specific models and conditions but there is not a unitary approach to the problem. For instance, in [4], the authors compared the distribution of the number of occurrences of a different word in Simon model, when time goes to infinity, with the degree distribution in the Barabási–Albert model, when the number of vertices goes to infinity. In [16], an explanation relating the asymptotic distribution of the number of species in a random genus in Yule model and that of the number of different words in Simon model appears. More recently, following a heuristic argument, Simkin and Roychowdhury [15] gave a justification of the relation between Yule and Simon models.

The aim of this paper is to study rigorously the relations between these three models. A fourth model, here named II-PA model (second preferential attachment model), will be discussed in order to better highlight the connection between Simon and Barabási–Albert models. The basic idea at the basis of our study is to make use of random graph processes theory to deal with all the considered discrete-time models and to include in this analysis also the continuous-time Yule model through the introduction of a suitable discrete-time process converging to it.

The random graph process approach was used by Barabási and Albert to define their preferential attachment model of World Wide Web [1]. At each discrete-time step a new vertex is added together with $m$ edges originating from it. The end points of these edges are selected with probability proportional to the current degree of the vertices in the network. Simulations from this model show that the proportion of vertices with degree $k$ is $c_m k^{−\gamma}$, with $\gamma$ close to 3 and $c_m > 0$ independent of $k$. A mathematically rigorous study of this model was then performed by Bollobás, Riordan, Spencer and Tusnady [3] making use of random graph theory. The rigorous presentation of the model allowed the authors to prove that the proportion of vertices with degree $k$ converges in probability to $m(m + 1)B(k, 3)$ as the number of vertices diverges, where $B(x, y)$ is the Beta function.

Here we reconsider all the models of interest in a random graph process framework. In Section 2 we introduce the necessary notations and basic definitions. Then, in Section 3 we present the four preferential attachment models of interest, i.e. Simon, II-PA, Barabási–Albert and Yule models, through a mathematical description that makes use of the random graphs approach. Such a description allows us to highlight an aspect not always well underlined: the asymptotic distributions that in some cases coincide do not always refer to the same quantity. For instance, the Barabási–Albert model describes the degree of the vertices while II-PA considers the in-degree. In Section 4 we also discuss the historical context and the list of available mathematical results for each model. The proposed point of view by means of random graphs processes then permits us to prove the novel results presented in Section 3. The theorems described and proved there clarify the relations between the considered asymptotic distributions of the different models, specifying for which choice of the parameters these distributions coincide and when they are not related.

In the concluding Section 5 we summarize the proved results and we illustrate with a diagram the cases in which the considered models are actually related.

2 Definitions and mathematical background

In this section we introduce some classical definitions, theorems and mathematical tools we will use in the rest of the paper.

Let define a graph $G = (V, E)$ as an ordered pair comprising a set of vertices $V$ with a set of edges or lines $E$ which are 2-elements subsets of $V$, so $E \subseteq V \times V$. A graph $G$ is directed
if its edges are directed, i.e., if for every edge \((i,j) \in E, (i,j) \neq (j,i)\), otherwise \(G\) is called an undirected graph.

We say that \(G\) is a random graph, if it is a graph selected according with a probability distribution over a set of graphs, or it is determined by a stochastic process that describes the random evolution of the graph in time. A stochastic process generating a random graph is called a random graph process. In other words, a random graph process is a family \((G_t)_{t \in T}\) of random graphs (defined on a common probability space) where \(t\) is interpreted as time and \(T\) can be either countable or uncountable.

A loop is an edge that connects a vertex to itself. The in-degree of a vertex \(v\) at time \(t\), denoted by \(d(v,t)\), is the number of incoming edges (incoming connections). Similarly, the degree of a vertex \(v\) at time \(t\), denoted by \(d(v,t)\), is the total number of incoming and outgoing edges at time \(t\) (when an edge is a loop, it is counted twice). In this paper we also use the term directed loop to indicate a loop that counts one to the in-degree.

The random graphs studied in this paper are random graph processes starting at time \(t = 0\), without any edge neither vertex, growing monotonically by adding at each discrete step either a new vertex or some directed edges between the vertices already present, according to some law \(\mathbb{P}(v_i \to v_j) = \mathbb{P}((i, j) \in G_t)\).

We focus here on the analysis of the number of vertices with degree or in-degree \(k\) at time \(t\), which we denote by \(N_{k,t}\) and \(\bar{N}_{k,t}\), respectively. In particular we are interested in the asymptotic degree or in-degree distribution of a random vertex, i.e., in the proportion \(N_{k,t}/V_t\) or \(\bar{N}_{k,t}/V_t\), as \(t\) goes to infinity, where \(V_t\) denotes the total number of vertices at time \(t\). We will add an upper index to \(N_{k,t}\) or \(\bar{N}_{k,t}\), for instance \(N_{k,t}^{Simon}\), to indicate the process to which we refer, if necessary.

Furthermore, we will make use of the following standard notation: for (deterministic) functions \(f = f(t)\) and \(g = g(t)\), we write \(f = O(g)\) if \(\lim_{t \to \infty} f/g\) is bounded, \(f \sim g\) if \(\lim_{t \to \infty} f/g = 1\), and \(f = o(g)\) if \(\lim_{t \to \infty} f/g = 0\).

One of the methods used in the literature to study the asymptotic behavior of \(N_{k,t}/V_t\) or \(\bar{N}_{k,t}/V_t\) is to prove that these random processes concentrate around their expectations. In order to do this, the Azuma and Hoeffding inequality is applied, when possible (see also [2], page 93).

Lemma 2.1 (Azuma and Hoeffding inequality [9]). Let \((X_t)_{t=0}^{\infty}\) be a martingale with \(|X_s - X_{s-1}| \leq c\) for \(1 \leq s \leq t\) and \(c\) a positive constant. Then
\[
\mathbb{P}(\{|X_t - X_0| > x\}) \leq \exp(-x^2/2c^2t). \tag{2.1}
\]

One of the first to use this approach in preferential attachment random graphs studies were Bollobás, et. al in [3]. Here we apply this approach to study different random graph processes. In Section 3 we illustrate this technique by analyzing the Simon model, reporting the corresponding computations for the Barabási–Albert model.

### 3 Preliminaries: Preferential attachment models

As stressed in the introduction, a number of models that make use of “preferential attachment” mechanisms are present in the literature. Here we consider some of them rigorously introducing the corresponding random graph processes with the aim to allow a comparison of their features. To this aim, it helps to present the most known models using a common notation. We first discuss the case of discrete time preferential attachment models, specifically Simon and Barabási–Albert models, and also a model that we call here the II-PA model (second preferential attachment model) that will help us to understand the relation between Simon and Barabási–Albert models. Moreover, we also discuss a continuous time preferential attachment model, the Yule model, which is defined in terms of independent homogeneous linear birth processes. We rigorously prove that this model can be related with Simon, and hence Barabási–Albert models.

The Barabási–Albert model presented in [1] omits some necessary details to be formulated in terms of a random graph process. Here we follow its description detailed as in Bollobás et. al [3] where the rules for the growth of the random graph not mentioned in [1] are given. Furthermore, in order to make easier the understanding of each model, we follow the same
scheme for its presentation, eventually specifying the absence of some results when not yet available.

Our scheme considers:
1. The mathematical description of the associated graph structure and its growth law.
2. The historical context motivating the first proposal of the model and some successive applications.
3. Available results on the degree or in-degree distribution with particular reference to power law behavior. We collect both, theorems and simulation results.

3.1 Simon model

1. Mathematical description: The Simon model can be described as a random graph process in discrete time \((G^\alpha_t)_{t \geq 1}\), so that \(G^\alpha_t\) is a directed graph which starts at time \(t = 1\) with a single vertex \(v_1\) and a directed loop. Then, given \(G^\alpha_t\), form \(G^{\alpha+1}\) by either adding with probability \(\alpha\) a new vertex \(v_i\) with a directed loop, \(i \leq t + 1\), or adding with probability \((1-\alpha)\) a directed edge between the last added vertex \(v\) and \(v_j\), \(1 \leq j \leq t\), where the probability of \(v_j\) to be chosen is proportional to its in-degree, i.e.,

\[
P(v \rightarrow v_j) = (1 - \alpha)\frac{\vec{d}(v_j, t)}{t}, \quad 1 \leq j \leq t.
\]

In Figure 1 we illustrate the growth law of this graph.

Figure 1: Construction of \((G^\alpha_t)_{t \geq 1}\). (a) Begin at time 1 with one single vertex and a directed loop. (b) Suppose some time has passed, in this case, the picture corresponds to a a realization of the process at time \(t = 4\). (c) Given \(G^\alpha_t\) form \(G^\alpha_{t+1}\) by either adding with probability \(\alpha\) a new vertex \(v_3\) with a directed loop, or adding a directed edge with probability given by (3.1).

2. Historical context: In [16], Simon considered a model to describe the growth of a text that is being written such that a word is added at each time \(t \geq 1\). Different words correspond to different vertices and repeated words to directed edges in the previous description. Simon introduced the two following conditions: For \(\alpha \in (0, 1)\),

(a) \(P\) \((t + 1)\) th word has not yet appeared at time \(t\) \(= \alpha\)

(b) \(P\) \((t + 1)\) th word has appeared \(k\) times at time \(t\) \(= (1 - \alpha)k\vec{N}_{k,t}/t\),

where \(\vec{N}_{k,t}\) is the number of different words that have appeared exactly \(k\) times at time \(t\), or the number of vertices that have exactly \(k\) incoming edges (i.e., in-degree \(k\)) at time \(t\) in \(G^\alpha_t\). Thus, at time \(t + 1\) either with probability \(\alpha\) a new word appears (i.e., a new vertex \(v_i\), \(i \leq t + 1\), with a directed loop appears), or with probability \((1 - \alpha)\) the word is not new, and if it has appeared \(k\) times at time \(t\), a directed edge is added. The starting point of this edge is the last vertex that has appeared in \(G^\alpha_t\), while its end point is selected with probability (3.1) that corresponds in this case to \(k/t\).

3. Available results: Simon was interested in getting results for the proportion of vertices that have exactly in-degree \(k\), with respect to the total number of vertices \(V_t\) at time \(t\). Thus, he proved asymptotic results for \(\mathbb{E}\vec{N}_{1,t}/\mathbb{E}V_t\) as \(t \rightarrow \infty\).

Next, we will give a brief synopsis of the computations made by Simon in [16]. The idea is to condition on what has happened until time \(t\) and compute the expected value at time \(t + 1\). For \(k = 1\) it holds

\[
\mathbb{E}\vec{N}_{1,t+1} = \alpha + \left(1 - \frac{(1 - \alpha)}{t}\right)\mathbb{E}\vec{N}_{1,t},
\]

(3.2)
and, for \( k > 1 \),
\[
\mathbb{E} \tilde{N}_{k,t+1} = \frac{1-\alpha}{t} \left[ (k-1) \mathbb{E} \tilde{N}_{k-1,t} - k \mathbb{E} \tilde{N}_{k,t} \right] + \mathbb{E} \tilde{N}_{k,t}.
\]
(3.3)

Simon solved (3.2) and (3.3) (see also [7], pages 98–99) to get, as \( t \to \infty \),
\[
\frac{\mathbb{E} \tilde{N}_{1,t}}{t} \to \frac{\alpha}{2-\alpha}.
\]
(3.4)

and for \( k > 1 \),
\[
\frac{\mathbb{E} \tilde{N}_{k,t}}{t} \to \frac{\alpha}{1-\alpha} \frac{\Gamma(k) \Gamma\left(1+\frac{1}{1-\alpha}\right)}{\Gamma\left(k+1+\frac{1}{1-\alpha}\right)}.
\]
(3.5)

where \( \Gamma \) is the gamma function.

Observe now that the number of vertices appeared until time \( t \), \( V_t \sim \text{Bin}(t, \alpha) \), so \( \mathbb{E} V_t = \alpha t \). Hence, using this and \( 3.4 \) and \( 3.5 \) for \( k = 1 \),
\[
\frac{\mathbb{E} \tilde{N}_{1,t}}{t} \to \frac{1}{2-\alpha}.
\]
(3.6)

and for \( k > 1 \),
\[
\frac{\mathbb{E} \tilde{N}_{k,t}}{\mathbb{E} V_t} \to \frac{1}{1-\alpha} \frac{\Gamma(k) \Gamma\left(1+\frac{1}{1-\alpha}\right)}{\Gamma\left(k+1+\frac{1}{1-\alpha}\right)} = \frac{1}{1-\alpha} B\left(k, 1 + \frac{1}{1-\alpha}\right).
\]
(3.7)

as \( t \to \infty \), where \( B(x, y) \) is the Beta function.

Now, let \( \mathcal{G}_r \) and \( \mathcal{G}_s \) denote the \( \sigma \)-fields generated by the appearance of directed edges up to time \( r \) and \( s \) respectively, \( r \leq s \leq t \). Since \( \mathbb{E} [\mathbb{E} (\tilde{N}_{k,t} | \mathcal{G}_s) \mid \mathcal{G}_r] = \mathbb{E} (\tilde{N}_{k,t} | \mathcal{G}_r) \), then, \( Z^{\text{Simon}}_t = \mathbb{E} (\tilde{N}_{k,t} | \mathcal{G}_r) \) is a martingale, such that \( Z^{\text{Simon}}_t = \tilde{N}_{k,t} \) and \( Z^{\text{Simon}}_t = \mathbb{E} \tilde{N}_{k,t} \). Furthermore, observe that at each unit of time, say \( s \), either a new vertex appears or the last one added, is attaching to another existing vertex \( v_j \), \( j \leq s \), but note this does not effect the in-degree of \( v \neq v_j \), or the probabilities these vertices will choose later, so it yields that \( |Z^{\text{Simon}}_t - Z^{\text{Simon}}_s| \leq 1 \). Then, it is possible to use Azuma–Hoeffding’s inequality (2.1), and obtain that for every \( \epsilon_t > t^{-1/2} \) (for example take \( \epsilon_t = \sqrt{\ln t/t} \),
\[
\mathbb{P} \left( \left| \frac{\tilde{N}_{k,t}}{t} - \frac{\mathbb{E} \tilde{N}_{k,t}}{t} \right| \geq \epsilon_t \right) \leq \exp \left( -\frac{(t\epsilon_t)^2}{2t} \right) \to 0.
\]
(3.8)

Now, using Chebyshev’s inequality, for every \( \epsilon_t > 0 \), such that \( t\epsilon_t^2 \to \infty \) as \( t \to \infty \),
\[
\mathbb{P} \left( \left| \frac{V_t}{t} - \frac{\mathbb{E} V_t}{t} \right| \geq \epsilon_t \right) \leq \frac{t\alpha(1-\alpha)}{t^2 \epsilon_t^2} \to 0.
\]
(3.9)

Hence, by (3.8) and (3.9), \( \tilde{N}_{k,t}/t \to \mathbb{E} \tilde{N}_{k,t}/t \) and \( V_t/t \to \mathbb{E} V_t/t \) in probability.

Finally, since \( \mathbb{E} \tilde{N}_{k,t}/t \) and \( \mathbb{E} V_t/t \) converge as \( t \) goes to infinity to the constant values \( 3.4 \) and \( 3.5 \) respectively, and because \( V_t \) is a random variable with binomial distribution, \( \text{Bin}(t, \alpha) \), then by properties of convergence in probability we obtain that
\[
\frac{\tilde{N}_{k,t}}{V_t} \to \frac{1}{1-\alpha} \frac{\Gamma(k) \Gamma\left(1+\frac{1}{1-\alpha}\right)}{\Gamma\left(k+1+\frac{1}{1-\alpha}\right)} = \frac{1}{1-\alpha} B\left(k, 1 + \frac{1}{1-\alpha}\right),
\]
(3.10)
in probability.
3.2 The II-PA model (second preferential attachment model)

1. Mathematical description: In [12] a different model is analyzed. In that paper it is called Yule model and described in discrete time. The model is defined also as a preferential attachment model but in this case at each time step a new vertex is added with exactly \( m + 1 \) directed edges, \( m \in \mathbb{Z}^+ \). These edges start from the new vertex and are directed towards any of the previously existing vertices according to a preferential attachment rule. To define formally a random graph process, we can think for a moment at an increasing time rescaled by \( 1/(m+1) \) so that at each unit of time \( n, m + 1 \) scaling time steps happen. Let \((\tilde{G}_t^m)_{t \geq 1}\) be a random graph process such that for all \( n \in \mathbb{Z}^+ \cup \{0\}\),

(a) at time \( t = n(m+1) + 1 \) add a new vertex \( v_{n+1} \) with a directed loop (it does count one for the in-degree), and

(b) for \( i = 2, \ldots, m+1 \) at each time \( t = n(m+1) + i \) add a directed edge from \( v_{n+1} \) to \( v_j \), \( 1 \leq j \leq n+1 \), with probability

\[
P(v_{n+1} \rightarrow v_j) = \frac{\tilde{d}(v_j, t-1)}{t-1}.
\]

(3.11)

Note then that \((\tilde{G}_t^m)_{t \geq 1}\) starts at time \( t = 1 \) with a single vertex and one directed loop.

In Figure 2 we illustrate the growth law of this graph.

![Figure 2: Construction of \((\tilde{G}_t^m)_{t \geq 1}\) for \( m = 3 \).](image)

(a) Begin at time 1 with one single vertex and a directed loop. (b) Suppose some time has passed, in this case, the picture corresponds to a realization of the process at time \( t = 3 \). Keep in mind that here \( m = 3 \) and therefore \( m - 1 = 2 \) directed edge are added to the graph by preferential attachment rule (but at this point the only possible choice is the vertex \( v_1 \)). (c) Here time is \( t = 5 \). A new vertex \( v_2 \) already appeared at time \( t = 4 \) together with a directed loop. At time 5 instead the first of the \( m - 1 \) edges that must be added to the graph is chosen (red dashed directed edges) by means of the preferential attachment probabilities (3.11).

2. Historical context: In [12], Newman describes this model in terms of genus and species as follows.

"Species are added to genera by “speciation”, the splitting of one species into two, [...]. If we assume that this happens at some stochastically constant rate, then it follows that a genus with \( k \) species in it will gain new species at a rate proportional to \( k \), since each of the \( k \) species has the same chance per unit time of dividing in two. Let us further suppose that occasionally, say once every \( m \) speciation events, the new species produced is, by chance, sufficiently different from the others in its genus as to be considered the founder member of an entire new genus. (To be clear, we define \( m \) such that \( m \) species are added to preexisting genera and then one species forms a new genus. So \( m+1 \) new species appear for each new genus and there are \( m+1 \) species per genus on average.)"

This description is linked to the model proposed by Simon; the difference is that the original Simon model does not fix \( m \) speciation events, instead it assumes that the number of speciation events is random and follows a probability distribution \( \text{Geo}(\alpha) \), with \( 0 < \alpha < 1 \).
3. **Available results:** Note that the number of vertices with in-degree equal to \( k \) is equivalent to the number of genera that have \( k \) species, thus, the number of vertices with in-degree equal to \( k \) at time \( t = n(m + 1) \), corresponds to the number of genera that have \( k \) species, when the number of genera is \( n \).

Let \( N_{k,t} \) be the number of vertices with in-degree equal to \( k \) in \( (G_t^m)_{t \geq 1} \). In [12] an heuristic analysis of the II-PA model shows that the proportion of vertices that have exactly in-degree \( k \), with respect to the total number of vertices at time \( t = n(m + 1) \), is in the limit

$$
\lim_{n \to \infty} \frac{N_{k,t}}{n} = \frac{(1 + 1/m)\Gamma(k)\Gamma(2 + 1/m)}{\Gamma(k + 2 + 1/m)} = (1 + 1/m)B(k, 2 + 1/m), \quad (3.12)
$$

We prove this result in Theorem 4.2.

### 3.3 Barabási–Albert model

1. **Mathematical description:** In [3], Bollobás, Riordan, Spencer, and Tusnady make the Barabási and Albert model precise in terms of a random graph process. We follow their description in this paragraph. Add at each time step a new vertex with the Barabási and Albert model precise in terms of a random graph process. We follow

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-

   In [3], Bollobás, Riordan, Spencer, and Tusnady make

   A vailable results:

   We can also define this model in a similar manner as we did for the II-PA model. Think-
To write a mathematical description of the process given above it is necessary to clarify some details. First, since the model starts with \( m_0 \) vertices and none edges, then the vertices degree are initially zero, so the probability that the new vertex is connected to a vertex \( i \), \( 1 \leq i \leq m_0 \), is not well defined. Second, to link the new vertex to \( m \) different vertices already present, it should be necessary to repeat \( m \) times the experiment of choosing an old vertex, but the model does not say anything on changes of attachment probabilities at each time, i.e. it is not explained if the \( m \) old vertices are simultaneously or sequentially chosen. These observations were made by Bollobás, Riordan, Spencer, and Tusnády in [4], where after noted the problems in the Barabási–Albert model, they give an exact definition of a random graph process that fits to that description.

3. Available results: In [1], Barabási and Albert obtain through simulation that after many time steps the proportion of vertices with degree \( k \) obeys a power law \( Ck^{-\gamma} \), where \( C \) is a constant and \( \gamma = 2.9 \pm 0.1 \), and by a heuristic argument they suggest that \( \gamma = 3 \). Let \( N_{k,t} \) be the number of vertices with degree equal to \( k \) in \( (G_n^m)_{t \geq 1} \). In [3] Bollobás, Riordan, Spencer, and Tusnády analyse mathematically this model. Their first result is that, for \( t = n(m+1) \), i.e., when the total number of vertices is \( n \), and \( m \leq k \leq m+n^{1/15} \) (the bound \( k \leq m+n^{1/15} \) is chosen to make the proof as easy as possible),

\[
\frac{\mathbb{E}N_{k,t}}{n} \sim \frac{2m(m+1)}{k(k+1)(k+2)} = \alpha_k,
\]

uniformly in \( k \).

The authors consider \( F_s \), the \( \sigma \)-field generated by the appearance of directed edges up to time \( s \), \( s \leq t \), and define \( Z_s = \mathbb{E}(N_{k,t} | F_s) \) and see it is a martingale satisfying \( |Z_s - Z_{s-1}| \leq 2 \), \( Z_t = N_{k,t} \) and \( Z_0 = \mathbb{E}N_{k,t} \), \( k = 1, 2, \ldots \) (at time \( t = 0 \) the random graph is the empty graph). Using Azuma–Hoeffding inequality (2.1) they obtain that

\[
\mathbb{P}\left( \left| \frac{N_{k,t}}{n} - \frac{\mathbb{E}N_{k,t}}{n} \right| \geq \sqrt{\ln t/t} \right) \leq \exp \left( - \frac{\ln t}{8} \right) \rightarrow 0,
\]

as \( t \) goes to infinity. Hence, it follows that, for every \( k \) in the range \( m \leq k \leq m+n^{1/15} \),

\[
\frac{N_{k,t}}{n} \rightarrow \alpha_k,
\]

in probability. Thus, the proportion of vertices with degree \( k \),

\[
\frac{N_{k,t}}{n} \rightarrow m(m+1)B(k,3)
\]

(3.15)

in probability as \( t \rightarrow \infty \). Note that

\[
\frac{2m(m+1)}{k(k+1)(k+2)} = m(m+1)B(k,3).
\]

Furthermore, since the Beta function satisfies the asymptotics \( B(x,y) \rightarrow x^{-y} \) for \( x \) large enough, then \( N_{k,t}/n \sim m(m+1)k^{-3} \) as \( k \rightarrow \infty \) and obeys a power law for large values of \( k \), with \( \gamma = 3 \) as Barabási and Albert suggested. Hence, it is proved mathematically that when vertices are added to the graph one at a time and joined to a fixed number of existing vertices selected with probability proportional to their degree, the degree distribution follows a power law behavior only in the tail (for \( k \) big enough), with an exponent \( \gamma = 3 \).

3.4 Yule model

Differently from the previous models, this model evolves in continuous time. We do not describe this model in terms of random graph processes, however in subsection 4.2 we discuss its relation with Simon model and conclude that the Yule model can be interpreted as a continuous time limit of Simon model, a model with a random graph interpretation.
1. **Mathematical description:** In the description of the Yule model we use $T$ to denote continuous time, instead of $t$ that denotes discrete time, i.e. $T \in \mathbb{R}^+ \cup \{0\}$ and $t \in \mathbb{Z}^+$. Consider a population starting at time $T = 0$ with one individual. As time increases, individuals may give birth to new individuals independent of each other at a constant rate $\lambda > 0$, i.e., during any short time interval of length $h$ each member has probability $\lambda h + o(h)$ to create an offspring. Since there is no interaction among the individuals, then if at epoch $T$ the population size is $k$, the probability that an increase takes place at some time between $T$ and $T + h$ equals $k \lambda h + o(h)$. Formally, let $N(T)$ be the number of individuals at time $T$ with $N(0) = 1$, then if $N(T) = k$, $k \geq 1$, the probability of a new birth in $(T, T + h)$ is $k \lambda h + o(h)$, and the probability of more than one birth is $o(h)$, i.e.,

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

Thus, $\{N(T)\}_{T \geq 0}$ is a pure birth process and with the initial condition $\Pr(N(0) = k) = \delta_{k,1}$, this linear birth process is called the Yule process.

Consider now two independent Yule processes, $\{N_\beta(T)\}_{T \geq 0}$ and $\{N_\lambda(T)\}_{T \geq 0}$, with parameters $\beta > 0$ and $\lambda > 0$ respectively, such that when a new individual appears in the process with parameter $\beta$, a new Yule process with parameter $\lambda$ starts. In a random graph context, a Yule model can be characterized through Yule processes of different parameters as described in the following. The first Yule process denoted by $\{N_\beta(T)\}_{T \geq 0}$, $\beta > 0$, accounts for the growth of the number of vertices. As soon as the first vertex is created, a second Yule process, $\{N_\lambda(T)\}_{T \geq 0}$, $\lambda > 0$, starts describing the creation of in-links to the vertex. The evolution of the number of in-links for the successively created vertices, proceeds similarly. Specifically, for each of the subsequent created vertices, an independent copy of $\{N_\lambda(T)\}_{T \geq 0}$, modeling the appearance of the in-links is initiated.

Let us define $Y_0 = 0$ and for $k \geq 1$,

$$
Y_k = \inf\{T : N_\lambda(T) = k + 1\},
$$

so that $Y_k$ is the time of the $k$th birth, and $W_k^* = Y_k - Y_{k-1}$ is the waiting time between the $(k-1)$th and the $k$th birth. In a Yule process it is well-known that the waiting times $W_k^*$, $k \geq 1$, are independent, each exponentially distributed with parameter $\lambda k$. Conversely, it is possible to reconstruct $\{N_\lambda(T)\}_{T \geq 0}$ from the knowledge of the $W_k^*$, $k \geq 1$, by defining

$$
Y_k = \sum_{j=1}^{k} W_j^*, \quad N_\lambda(T) = \min\{k : Y_k > T\}. \tag{3.16}
$$

Thus if the $W_k^*$ are independently distributed exponential random variables, of parameter $\lambda j$, then $\{N_\lambda(T)\}_{T \geq 0}$ is a Yule process of parameter $\lambda$.

2. **Historical context:** Yule in [13] observed that the distribution of species per genus in the evolution of a biological population typically presents a power law behavior, thus, he proposed a stochastic model to fit these data. In the original paper [13] the process is described as follows:

*Let the chance of a species “throwing” a specific mutation, i.e., a new species of the same genus, in some small assigned interval of time be $p$, and suppose the interval so small that $p^2$ may be ignored compared with $p$. Then, putting aside generic mutations altogether for the present, if we start with $N$ prime species of different genera, at the end of the interval we will have $N(1 - p)$ which remain monotype and $Np$ genera of two species. The new species as well as the old can now throw specific mutation.*

Yule proceeded to the limit, taking the time interval $\Delta T$ as indefinitely but the number of such intervals $n$ as large, so that $n\Delta T = T$ is finite, and he wrote $p = \lambda \Delta T = \lambda T/n$. Yule
not only studied this process. In [18], he furthermore constructed a model of evolution by considering two independent Yule processes, one for species with a constant rate $\lambda > 0$ and the other for new genera (each of them composed by a single species) created at a constant rate $\beta > 0$. In other words, at time $T = 0$ the process starts with a single genus composed by a single species. As time goes on, new genera (each composed by a single species) develop as a Yule process of parameter $\beta$, and simultaneously and independently new species evolve as a Yule process with rate $\lambda$. Furthermore, since a new genus appears with a single species, then each time a genus births, a Yule process with rate $\lambda$ starts.

3. Available results: Let $N_g(T)$ and $N_s(T)$, $T \geq 0$, be the counting processes measuring the number of genera and species created until time $T$, respectively. It is well-known that the probability distribution of the number of individual in a Yule process with parameter $\lambda$ is geometric, $\text{Geo}(e^{-\lambda T})$. Thus, the distribution of the number of species $N_s(T)$ in a genus during the interval of time $[0, T]$ is

$$P(N_s(T) = k) = e^{-\lambda T} (1 - e^{-\lambda T})^{k-1}, \quad k \geq 1, \quad T \geq 0. \tag{3.17}$$

On the other hand, it is also known that by conditioning on the number of genera present at time $T$, the random instants at which creation of novel genera occurs are distributed as the order statistics of iid random variables with distribution function $P(T \leq \tau) = e^{\beta \tau} - 1$, $e^{-\beta \tau} - 1$, $0 \leq \tau \leq t$. \tag{3.18}

(see [10] and the references therein). The authors in [10] take account that the homogeneous linear pure birth process lies in the class of the so-called processes with the order statistic property, see [11], and use [5, 8, 13] and [17] to get (3.18).

Thus, let $N_T$ be the size of a genus chosen uniformly at random at time $T$. Then,

$$P(N_T = k) = \int_0^T P(N_s(T) = k | N_s(\tau) = 1)P(T \in d\tau)$$

$$= \int_0^T e^{-\lambda (T - \tau)}(1 - e^{-\lambda (T - \tau)})^{k-1} \frac{e^{\beta \tau}}{e^{\beta T} - 1} d\tau$$

$$= \frac{\beta}{1 - e^{-\beta T}} \int_0^T e^{-\beta \tau} e^{-\lambda \tau} (1 - e^{-\lambda \tau})^{k-1} dy. \tag{3.19}$$

The interest now is in the limit behavior when $T \to \infty$:

$$\lim_{T \to \infty} P(N_T = k) = \beta \int_0^\infty e^{-\beta \tau} e^{-\lambda \tau} (1 - e^{-\lambda \tau})^{k-1} dy. \tag{3.20}$$

Letting $\rho = \beta/\lambda$ it is possible to recognize the integral as a beta integral to obtain (see [18], page 39)

$$\lim_{T \to \infty} P(N_T = k) = \rho \frac{\Gamma(k+1)}{\Gamma(1+\rho)} = \rho B(k, 1+\rho), \quad k \geq 1. \tag{3.21}$$

4 Main Results

4.1 Relations between Simon, II-PA and Barabási-Albert models

In [4], Bornholdt and Ebel pointed out that the asymptotic power law of the Barabási–Albert model with $m = 1$ coincides with that of the Simon model characterized by $\alpha = 1/2$ (see (3.10) and (3.15)). From this observation they suggested that for $m = 1$ Barabási–Albert model could be mapped to the subclass of Simon models with $\alpha = 1/2$.

We think that (3.10) and (3.15) cannot be compared since they give the asymptotic of different random variables. Indeed the first refers to the in-degree and the second to the degree distribution. However we do believe that may exist a relation between these two models, which needs to be explained.
In this section we discuss rigorous arguments that allow us to clarify the relation between Simon and Barabási–Albert models. To this aim we make use also of the II-PA model. We first relate Barabási–Albert and II-PA models and then II-PA and Simon models. This double step is made necessary by the different quantities described by these models.

Now, we are ready to formulate our results.

**Theorem 4.1.** Let \( m = 1 \). Then, the in-degree distribution of the II-PA model and the degree distribution of the Barabási–Albert model at time \( t \), \( t \geq 2 \), are the same, i.e., if at time \( t \) there are \( n \) vertices in the processes, then for any \( k \in \mathbb{Z}^+ \),

\[
\frac{\bar{N}_{k,t}}{n} = \frac{N_{k,t}}{n},
\]

where \( \bar{N}_{k,t} \) and \( N_{k,t} \) denote the number of vertices with in-degree and degree equal to \( k \) in \((G^t_1)\) and \((G^1_t)\) at time \( t \), respectively.

**Proof.** We follow the mathematical description of the II-PA and Barabási–Albert models in terms of the random graph processes \((G^t_m)_{n \geq 1}\) and \((G^1_m)_{n \geq 1}\), presented in Sections 3.2 and 3.3 respectively. Let us divide each unit of time in two sub-units. At each instant of time \( t = 2n + 1 \) a new vertex \( v_{n+1} \) is created in both models; in the II-PA model this vertex is created together with a directed loop. Furthermore, at each time \( t = 2n + 2 = 2(n+1) \) an edge (a directed edge in the II-PA model) is added from \( v_{n+1} \) to \( v_j \), \( j \leq n + 1 \), with probabilities given by (3.11) and (3.14) for the II-PA and Barabási–Albert models, respectively. Hence our thesis corresponds to show that (3.11) and (3.14) coincide under our hypotheses.

We see that the denominator for both probabilities (3.11) and (3.14) is \( 2n+1 \), and although the two numerators count different quantities, the in-degree for the II-PA and the degree for Barabási–Albert models, their values also coincide. This is easy to check when \( v_j = v_{n+1} \) and the directed edge created at time \( t = 2n + 1 \) is to \( v_{n+1} \). In fact the numerators of (3.11) and (3.14) become both one. Let us now show that the two numerators coincide also when \( v_j \neq v_{n+1} \).

Let us suppose \( v_j \neq v_{n+1} \), and let \( t = 2(n+1) + 1 = 2n+1 \). Observe that in the Barabási–Albert model the degree of \( v_j \) at time \( t = 2n+1 \), \( d(v_j, 2n+1) \), \( j \in \{1, 2, \ldots, n\} \), is the sum of the number of incoming edges from time \( t = 2j + 1 \) (when \( v_{j+1} \) is added) to time \( 2n + 1 \), plus the degree corresponding to the edge added at time \( t = 2j \) from \( v_j \), that is two if the edge was a loop and one otherwise. On the other hand, in the II-PA model the in-degree of \( v_j \) at time \( t = 2n+1 \), \( \bar{d}(v_j, 2n+1) \), \( j \in \{1, 2, \ldots, n\} \) is the sum of the number of incoming edges added in the interval of time \( t \in [2j + 1, 2n+1] \) (so this part coincide with Barabási–Albert model), plus the in-degree corresponding to the directed edge added at time \( t = 2j \) from \( v_j \). Thus, if it is a directed loop to \( v_j \), the in-degree of \( v_j \) at time \( t = 2j \) is two (since when \( v_j \) appeared, it did together with a directed loop), otherwise the in-degree is one. This conclude the proof and (3.11) and (3.14) coincide.

**Remark 4.1.** The proof of Theorem 4.1 enlightens the advantage given by the re-definition of existing models in terms of random graph processes. In particular this reading shows immediately that the two models can be related only when \( m = 1 \).

**Remark 4.2.** The in-degree distribution of the II-PA model and the degree distribution of the Barabási–Albert model are different when \( m > 1 \). In fact take for example \( m = 2 \) and suppose the first directed edge from \( v_{n+1} \) is not a loop, i.e., a vertex \( v_j \), \( 1 \leq j \leq n \) is chosen. Then, at time \( t = 3n + 1 \), \( d(v_{n+1}, 3n+1) = 1 \) in the II-PA model, while \( d(v_{n+1}, 3n+1) = 2 \) in the Barabási–Albert model. Thus at time \( t = 3n + 2 \), (3.11) and (3.14) are different because the corresponding numerators differ.

Next we discuss the relation between Simon and the II-PA models, which allows us to relate Barabási–Albert and Simon models. To obtain this result we prove a theorem establishing that the asymptotic in-degree distribution of the II-PA and Simon models coincide when \( m = 1 \). First we introduce the following definition.

**Definition 4.1.** We say that a vertex \( v_i \) appears “complete” when it has appeared in the process together with all the directed edges originated from it. Thus, at time \( t = (m+1) \), the II-PA model has exactly \( n \) “complete” vertices.

Now we are ready to enunciate the theorem.
Theorem 4.2. Let \( m \in \mathbb{Z}^+ \) be fixed and let \( \alpha = 1/(m+1) \). If \( (\tilde{G}_{m}^t)_{t \geq 1} \) is the random graph process defining the II-PA model, \( N_{II-PA}^{m,t} \) the number of vertices with in-degree equal \( k \), \( k \geq 1 \), and \( k = o(n/m) \), at time \( t \), and \( \bar{N}(k, n) \) the number of vertices with in-degree \( k \) when there are exactly \( n \) complete vertices. Then, at time \( t = o(n+1) \)

\[
\frac{N_{II-PA}^{m,t}}{n} = \frac{\bar{N}(k, n)}{n} \rightarrow \frac{(1 + 1/m)\Gamma(k)(2 + 1/m)}{\Gamma(k + 2 + 1/m)} \quad (4.1)
\]

almost surely as \( n \rightarrow \infty \).

Remark 4.3. From Theorem 4.1, 4.2, and 4.3, it follows that the asymptotic degree distribution of Barabási–Albert model coincides with the asymptotic in-degree distribution of Simon model only when \( m = 1 \) and \( \alpha = 1/(m+1) \), so that \( \alpha = 1/2 \). We conjecture that some other properties of Barabási–Albert model when \( m = 1 \), for example the diameter, should be also related with the analogous features of Simon model when \( \alpha = 1/2 \).

Remark 4.4. Observe that (7,7) coincides with (5,13). Thus, the previous theorem gives a rigorous formalization to the heuristic result in [12].

Before proving Theorem 4.2 we need to prove the following lemmas.

Lemma 4.1. Let \( r, s, t \in \mathbb{Z}^+ \) and \( b \in \mathbb{R} \) such that \( |b/r| < 1 \), then

\[
\prod_{r=s+1}^{t} \left( 1 - \frac{b}{r} \right) = \left( \frac{s}{t} \right)^b \left( 1 + O\left( \frac{t-s}{st} \right) \right) .
\]

Proof. Since \( |b/r| < 1 \), then using Taylor expansion for \( \ln(1 - b/r) \) we get

\[
\prod_{r=s+1}^{t} \left( 1 - \frac{b}{r} \right) = \exp \left[ \sum_{r=s+1}^{t} \left( -\frac{b}{r} + O\left( \frac{b^2}{r^2} \right) \right) \right] .
\]

Now, by Euler–Maclaurin it is possible to obtain that (see [14])

1. \( \sum_{r=1}^{t} \frac{1}{r} = \ln t + 1 - \int_{1}^{t} \frac{y - \lfloor y \rfloor}{y^2} dy \),
2. \( \sum_{r=1}^{t} \frac{1}{r^2} = \frac{1}{t} - 2 \int_{1}^{t} \frac{1 - \lfloor y \rfloor}{y^2} dy \).

Using these expressions and the fact that \( y - 1 \leq \lfloor y \rfloor \leq y \), where \( \lfloor y \rfloor \) indicates the integer part of \( y \), we obtain

\[
\ln t - \ln s - \frac{t-s}{st} < \sum_{r=1}^{t} \frac{1}{r} < \ln t - \ln s, \quad (4.2)
\]

\[
\frac{t-s}{st} - \frac{2(t-s)^2}{(st)^2} < \sum_{r=1}^{t} \frac{1}{r^2} < \frac{t-s}{st}, \quad (4.3)
\]

or, \( \sum_{r=1}^{t} 1/r = \ln t - \ln s - \lfloor \delta_1 \rfloor \), where \( \lfloor \delta_1 \rfloor < (t-s)/(st) \), and \( \sum_{r=1}^{t} 1/r^2 = (t-s)/(st) - \lfloor \delta_2 \rfloor \), where \( \lfloor \delta_2 \rfloor < (t^2 - s^2)/(st^2) \). Thus,

\[
\prod_{r=s+1}^{t} \left( 1 - \frac{b}{r} \right) = \exp \left[ b \ln \left( \frac{s}{t} \right) + O\left( \frac{t-s}{st} \right) \right] = \left( \frac{s}{t} \right)^b \left( 1 + O\left( \frac{t-s}{st} \right) \right) .
\]

Lemma 4.2. Let \( N_{II-PA}^{m,t} \) and \( \bar{N}(k, n) \) be as in Theorem 4.2. Then,

- for \( m = 1 \) and \( k = 1 \),

\[
\mathbb{E}\bar{N}(1, n+1) = \left( 1 - \frac{1}{(m+1)(n+1) - 1} \right) + \left( 1 - \frac{1}{(m+1)(n+1) - 1} \right) \mathbb{E}\bar{N}(1, n); \quad (4.4)
\]
• for \( m > 1 \) and \( k = 1 \),
\[
\mathbb{E} \tilde{N}(1, n + 1) = 1 + \left( 1 - \frac{m}{(n+1)(n+1) - 1} \right) \mathbb{E} \tilde{N}(1, n) + O\left( \frac{1}{n} \right);
\]  
(4.5)

• for \( m = 1 \) and \( k \geq 2 \),
\[
\mathbb{E} \tilde{N}(k, n + 1) = \frac{(k-1)\mathbb{E} \tilde{N}(k-1, n)}{(n+1)(n+1) - 1} + \left( 1 - \frac{k}{(n+1)(n+1) - 1} \right) \mathbb{E} \tilde{N}(k, n)
\]
\[
= \frac{(k-1)\mathbb{E} \tilde{N}(k-1, n)}{2(n+1) - 1} + \left( 1 - \frac{k}{2(n+1) - 1} \right) \mathbb{E} \tilde{N}(k, n);
\]  
(4.6)

• for \( m > 1 \) and \( k \geq 2 \),
\[
\mathbb{E} \tilde{N}(k, n + 1) = \frac{(k-1)m\mathbb{E} \tilde{N}(k-1, n)}{(n+1)(n+1) - 1} + \left( 1 - \frac{km}{(n+1)(n+1) - 1} \right) \mathbb{E} \tilde{N}(k, n) + O\left( \frac{k}{n} \right).
\]  
(4.7)

Proof. Let \( m = 1 \) and \( k = 1 \) we start at time \( t = (m+1)n = 2n \), i.e., when there are exactly \( n \) complete vertices. To see what happens when exactly \((n+1)\) complete vertices appear, we need to check what happens in two steps of the process, at time \( 2n+1 \), when deterministically appears a new vertex with a directed loop, and at time \( 2n+2 = 2(n+1) \), when a new directed edge is added by preferential attachment, and the last vertex added becomes complete. Thus, conditioning on what happens until time \( t+1 \), we have

\[
\mathbb{E}(N_{1,t+1}^{PA}) = \mathbb{E} \left[ \left( \frac{N_{1,t}^{PA}}{t+1} + 1 \right) + N_{1,t}^{PA} \left( \frac{N_{1,t}^{PA}}{t+1} + 1 \right) \right] = \left( 1 - \frac{1}{t+1} \right) + \left( 1 - \frac{1}{t+1} \right) \mathbb{E}(N_{1,t}^{PA}).
\]  
(4.8)

Thus, if \( \tilde{N}(k, n) \) denotes the number of vertices with in-degree \( k \) when there are exactly \( n \) complete vertices in the process, then we can write the previous equation as \( (4.8) \).

Let \( m > 1 \) and \( k = 1 \). Now we need a bit more attention, since we have to consider two different situations, when \( t \) is multiple of \((m+1)\) and when \( t \) is not. In the first situation \( t \) has the form \( t = n(m+1) \), so we are in the instant of time when there are exactly \( n \) complete vertices, and as we did above, to see what happens later we check what happens in the two subsequent steps of the process, at time \( n(m+1) + 1 \) when a deterministic event happens, a new vertex with a directed loop appears, and at time \( n(m+1) + 2 \) when something probabilistic happens, a new directed edge is added by preferential attachment. In the first case equation \( (4.9) \) still holds. In the second situation observe that if \( m > 1 \), in order to see complete the vertex added at time \( t = n(m+1) + 1 \), we have to check what happens from \( n(m+1) + 1 \) until \( n(m+1) + (m+1) = (n+1)(m+1) \), when this vertex becomes complete. Thus, when \( t \) is not multiple of \((m+1)\) we have the following equation.

\[
\mathbb{E}(N_{1,t}^{PA}) = \mathbb{E} \left[ \left( \frac{N_{1,t}^{PA}}{t} + 1 \right) + \left( \frac{N_{1,t}^{PA}}{t} - 1 \right) \right] = \left( 1 - \frac{1}{t} \right) \mathbb{E}(N_{1,t}^{PA}).
\]  
(4.9)

Now, we may use simultaneously \( (4.8) \) and \( (4.9) \) to get the corresponding equation of what happens in \((m+1)\) steps of the process. We start at time \( t = (n+1)(m+1) - 1 \), so at time \( t+1 \) the process will have exactly \((n+1)\) complete vertices, and since \( t \) is not multiple of \((m+1)\), we need to begin using \( (4.5) \) \((m-1)\) times, and then use \( (4.6) \). Iterating \( m \) times, we obtain

\[
\mathbb{E}(N_{1,t+1}^{PA}) = \left[ 1 + \mathbb{E}(N_{1,t}^{PA}) \right] \prod_{j=0}^{m-1} \left( 1 - \frac{1}{t-j} \right)
\]
\[
= \left[ 1 + \mathbb{E}(N_{1,t-m}^{PA}) \right] \prod_{r=1}^{l} \left( 1 - \frac{1}{r} \right)
\]
\[
= \left( 1 - \frac{m}{t} + O\left( \frac{1}{t^2} \right) \right) \mathbb{E}(N_{1,t}^{PA}),
\]  
(4.10)
where we have used in the the last two steps that \( r = t - j \) and Lemma 4.1. Finally, using the notation \( \tilde{N}(1,n) \), and since \( \tilde{N}(1,n)/n \leq 1 \), we get (4.15).

The cases \( k = 2 \) and \( k > 2 \) need to be first considered separately, and in each of these we need to analyze when \( m = 1 \) and when \( m > 1 \). Then we will show that the equations for \( k = 2 \) and \( k > 2 \) admit a general form, that include the cases \( k \geq 2 \).

Let \( m = 1 \). Analogously as we did when \( m = 1 \) and \( k = 1 \), consider the time \( t = n(m + 1) \), i.e., when there are exactly \( n \) complete vertices. To account for what happens until when \( (n + 1) \) complete vertices appear, we need to recognize two steps of the process. Indeed,

\[
E(\tilde{N}_{k,t+2}^{\text{II-P A}}) = E\left[ (\tilde{N}_{k,t}^{\text{II-P A}} + 1)\left( \frac{\tilde{N}_{k-1,t}^{\text{II-P A}}}{t+1} + (\tilde{N}_{k,t}^{\text{II-P A}} - 1)\frac{2\tilde{N}_{k,t}^{\text{II-P A}}}{t+1} \right) \right]
\]

and for \( k > 2 \),

\[
E(\tilde{N}_{k,t+2}^{\text{II-P A}}) = E\left[ (\tilde{N}_{k,t}^{\text{II-P A}} + 1)\left( \frac{(k-1)\tilde{N}_{k-1,t}^{\text{II-P A}}}{t+1} + (\tilde{N}_{k,t}^{\text{II-P A}} - 1)\frac{k\tilde{N}_{k,t}^{\text{II-P A}}}{t+1} \right) \right]
\]

Note that in the last line of (4.11) and (4.12), we have replaced \( \tilde{N}_{k,t+1}^{\text{II-P A}} \) with \( \tilde{N}_{k,t}^{\text{II-P A}} \). In fact if \( t = n(m + 1) \), then at time \( t + 1 \) the process just adds deterministically a new vertex with in-degree one, so when \( k \geq 2 \), \( \tilde{N}_{k,t+1}^{\text{II-P A}} = \tilde{N}_{k,t}^{\text{II-P A}} \), as well as \( \tilde{N}_{k+1,t+1}^{\text{II-P A}} = \tilde{N}_{k,t}^{\text{II-P A}} + 1 \). Using this observation we can express (4.11) and (4.12) as a single equation holding for \( k \geq 2 \) and \( m = 1 \). Using the notation \( \tilde{N}(1,n) \) it can be written as (4.13). Let \( m > 1 \). Once more we need to consider when \( t \) is multiple of \( (m + 1) \), and when it is not. When \( t = n(m + 1) \) we obtain again (4.11) and (4.12) for \( k = 2 \) and \( k > 2 \), respectively, while if \( t \) is not multiple of \( (m + 1) \) and \( k \geq 2 \) it holds

\[
E(\tilde{N}_{k,t+1}^{\text{II-P A}}) = E\left[ (\tilde{N}_{k,t}^{\text{II-P A}} + 1)\left( \frac{(k-1)\tilde{N}_{k-1,t}^{\text{II-P A}}}{t+1} + (\tilde{N}_{k,t}^{\text{II-P A}} - 1)\frac{k\tilde{N}_{k,t}^{\text{II-P A}}}{t+1} \right) \right]
\]

Now, in order to get the corresponding equation for what happens in \( (m+1) \) steps, i.e., during the time interval from when there are \( n \) vertices until when there are \( (n + 1) \) vertices, it is necessary to use (4.12) and (4.13) simultaneously. In the same manner as we did for \( k = 1 \), we take \( t = (n+1)(m+1) - 1 \), so that at time \( t + 1 \) the process will have exactly \( (n + 1) \) complete vertices. Since \( t \) is not multiple of \( (m + 1) \), we need to begin using (4.13) \( (m-1) \) times, and then use (4.12). Thus, iterating \( m \) times, after some algebra we obtain that for any \( k \geq 2 \),

\[
E(\tilde{N}_{k,t+1}^{\text{II-P A}}) = \sum_{i=0}^{m-1} \frac{(k-1)E(\tilde{N}_{k-1,t+1-i}^{\text{II-P A}})}{t-i} \prod_{j=0}^{i-1} \left( 1 - \frac{k}{t-j} \right)
\]
where the empty product (i.e., when \( i = 0 \)) is equal to unity. Let now \( r = t - j \), and since \( i \leq m \) and \( m \) is fixed, then by Lemma 4.1 we have

\[
\prod_{j=0}^{i-1} \left( 1 - \frac{k}{t-j} \right) = \prod_{r=t-(i-1)}^{t} \left( 1 - \frac{ki}{t} \right) = 1 - \frac{ki}{t} + O\left( \frac{k^2}{t^2} \right). \tag{4.15}
\]

Moreover, observe that \( \mathbb{E}(\bar{N}_{1,1}^{(n)}) - \mathbb{E}(\bar{N}_{1,1}^{(m-1)}) \) is fixed, then by Lemma 4.1 we have

\[
\prod_{i=0}^{m-1} \mathbb{E}(\bar{N}_{1,1}^{(n)}) = \sum_{i=0}^{m-1} \mathbb{E}(\bar{N}_{1,1}^{(n)}) (1 + \frac{i}{t}) + O\left( \frac{1}{t} \right). \tag{4.16}
\]

Then using (4.15) and (4.16) and noting that \( \frac{(k-1)i}{t-i} \leq 1 \), we can write (4.13) as

\[
\mathbb{E}(\bar{N}_{1,1}^{(n)}) = \frac{(k-1)m\mathbb{E}(\bar{N}_{1,1}^{(n)})}{t} + \left( 1 - \frac{km}{t} \right) \mathbb{E}(\bar{N}_{1,1}^{(m-1)}) + O\left( \frac{k}{t} \right), \tag{4.17}
\]

and using the notation \( \bar{N}(k,n) \) we obtain (4.7).

Theorem 4.2 gives the limiting value to which \( \bar{N}(k,n)/n \) converges when \( n \) goes to infinity. However, before proving the limit, we need to argue that such limit exists.

**Lemma 4.3.** Let \( \bar{N}(k,n) \) be as in Theorem 4.2. Then, there exist values \( N_1(k) > 0 \) and \( N_2(k) > 0 \) such that

\[
\frac{\bar{N}(k,n)}{n} \rightarrow N_1(k) \quad \text{a.s.,}
\]

and

\[
\frac{n\bar{N}(k,n)}{(m+1)n} \rightarrow N_2(k) \quad \text{a.s.}
\]

**Proof.** We make use of supermartingale’s convergence theorem (see [2], Theorem 35.5) and equations (4.1), (4.3), (4.5) and (4.7). Consider first (4.1) and (4.3) and observe that since \( \bar{N}(1,n)/(n+1)(m+1) - 1 \) \leq 1,

\[
\mathbb{E}\bar{N}(1,n) + 1 \leq \mathbb{E}\bar{N}(1,n) + 1 + O\left( \frac{k}{n} \right), \tag{4.18}
\]

while for (4.6) and (4.7),

\[
\mathbb{E}\bar{N}(k,n) + 1 \leq \mathbb{E}\bar{N}(k,n) + 1 + O\left( \frac{k}{n} \right). \tag{4.19}
\]

Let \( H_n \) be the filtration generated by the process \( \{\bar{N}(k,n), \bar{N}(k-1,n)\}_{n} \) until time \( n \), i.e., \( H_n := \sigma(\bar{N}(k,n), \bar{N}(k-1,n); 0 \leq j \leq n) \). If \( k = 1 \), let \( Z(1,n) = (\bar{N}(1,n) - n)/n \), then by (4.18),

\[
\mathbb{E}\bar{N}(1,n) + 1 \leq \mathbb{E}\bar{N}(1,n) + 1 + O\left( \frac{k}{n} \right). \tag{4.20}
\]

as \( \bar{N}(1,n) \) is \( H_n \)-measurable. Hence, \( \{Z(1,n)\}_n \) is a supermartingale and in order to apply supermartingale convergence theorem to \( \{Z(1,n)\}_n \), it remains to prove that

\[
\sup_n \mathbb{E}(\mathbb{E}(Z(1,n))) < \infty.
\]

This is true as

\[
\mathbb{E}(\mathbb{E}(Z(1,n))) = \mathbb{E}(\mathbb{E}(Z(1,n)|H_{n-1})) \leq \frac{1}{n}(\mathbb{E}\bar{N}(1,n-1) + 1 + n) < \infty, \tag{4.21}
\]

\[15\]
having used that \( \bar{N}(k,n)/n \leq 1 \), for any \( n \geq 1 \).

When \( k \geq 2 \), i.e., for (4.10) and (4.17), note first that if \( f(n) = O(k/n) \), then there exists \( M > 0 \) such that \( |f(n)| = M k/n + |\delta| \), where \( |\delta| < k/n \). Since \( k/n \leq 1 \), then there exists \( M \) such that \( |f(n)| \leq M + 1 \). Thus, take \( Z(k,n) = \bar{N}(k,n) - n(c + 1)/n \), with \( c = M + 1 \), then by (4.18) we also get that \( \{Z(k,n)\}_n \) is a supermartingale, and similarly as we did above we also show that \( \sup_n \mathbb{E}(|Z(k,n)|) < \infty \). In this manner we have proved that \( Z(k,n) \) converges almost surely, thus \( \bar{N}(k,n)/n \) converges almost surely. In perfect analogy we can prove that \( m k Z(k,n)/(m + 1) \) converges almost surely, and thus obtain that \( m k \bar{N}(k,n)/(m + 1) \) also converges almost surely.

In order to determine such a limit, we still need to prove the following lemma.

**Lemma 4.4.** Let \( p_k := \lim_{n \to \infty} \mathbb{E}\bar{N}(k,n)/n \). Then,

\[
p_k = \frac{(1 + 1/m)(k)(2 + 1/m)}{(k + 2 + 1/m)}, \quad k \geq 1.
\]  

**Proof.** Observe that for a function \( f(k) \),

\[
\frac{m f(k)}{(n + 1)(m + 1) - 1} = \frac{m f(k)}{n(m + 1)} \left(1 - \frac{m}{(n + 1)(m + 1) - 1} \right) = \frac{m f(k)}{n(m + 1)} + O\left(\frac{f(k)}{n^2}\right).
\]  

By using (4.24) we can write the equations (4.24), (4.25), (4.26) and (4.27) as follows. For \( m = 1, k = 1 \),

\[
\mathbb{E}\bar{N}(1,n + 1) = 1 + \left(1 - \frac{1}{n(m + 1)}\right)\mathbb{E}\bar{N}(1,n) + O\left(\frac{1}{n}\right),
\]  

for \( m > 1, k = 1 \),

\[
\mathbb{E}\bar{N}(1,n + 1) = 1 + \left(1 - \frac{m}{n(m + 1)}\right)\mathbb{E}\bar{N}(1,n) + O\left(\frac{1}{n}\right),
\]  

for \( m = 1, k \geq 2 \),

\[
\mathbb{E}\bar{N}(k,n + 1) = \frac{(k - 1)\mathbb{E}\bar{N}(k - 1,n)}{n(m + 1)} + \left(1 - \frac{k}{n(m + 1)}\right)\mathbb{E}\bar{N}(k,n) + O\left(\frac{1}{n}\right),
\]  

and for \( m > 1, k \geq 2 \),

\[
\mathbb{E}\bar{N}(k,n + 1) = \frac{m(k - 1)\mathbb{E}\bar{N}(k - 1,n)}{n(m + 1)} + \left(1 - \frac{mk}{n(m + 1)}\right)\mathbb{E}\bar{N}(k,n) + O\left(\frac{k}{n}\right).
\]

Looking at (4.24), (4.25), (4.26) and (4.27), we remark that they can be unified as

\[
\mathbb{E}\bar{N}(k,n + 1) = g(k - 1,n) + \left(1 - \frac{b}{n}\right)\mathbb{E}\bar{N}(k,n) + \epsilon_n,
\]

where \( b = km/(m + 1) \), \( g(0,n) = 1, g(k - 1,n) = (mk/n(m + 1))\bar{N}(k,n) \) for \( k \geq 2 \), and \( \epsilon_n = O(1/n) \) if \( m = 1 \) and of order \( O(k/n) \) if \( m > 1 \). We underline that \( k \) is function of \( n \) and hence in general \( O(k/n) \) can be different of \( O(1/n) \).

Note now that when the first complete vertex appears, it has in-degree equal to \( (m + 1) \), so \( \bar{N}(k,1) = 0 \) for any \( k \neq (m + 1) \), and \( \bar{N}(m + 1,1) = 1 \). Iterating (4.28) we have, if \( k \neq (m + 1) \),

\[
\mathbb{E}\bar{N}(k,n + 1) = \sum_{i=0}^{n-1} g(k - 1,n - i) \prod_{j=0}^{i-1} \left(1 - \frac{b}{n - j}\right) + \sum_{i=0}^{n-1} \epsilon_{n-i},
\]

while, if \( k = m + 1 \),

\[
\mathbb{E}\bar{N}(k,n + 1) = \sum_{i=0}^{n-1} g(k - 1,n - i) \prod_{j=0}^{i-1} \left(1 - \frac{b}{n - j}\right) + \sum_{i=0}^{n-1} \epsilon_{n-i}.
\]
To solve (4.30), let $s = n - i$ and $r = n - j$ so that

$$
\prod_{j=0}^{i-1} \left(1 - \frac{b}{n-j}\right) = \prod_{r=s+1}^{n} \left(1 - \frac{b}{r}\right),
$$

then observe that if $s < \lfloor b \rfloor$,

$$
\prod_{r=s+1}^{n} \left(1 - \frac{b}{r}\right) = \prod_{r=\lfloor b \rfloor + 1}^{n} \left(1 - \frac{b}{r}\right),
$$

which is equal to 0 if $b = \lfloor b \rfloor$ or to

$$
(-1)^{\lfloor b \rfloor} \prod_{i=1}^{\lfloor b \rfloor - i + 1} \prod_{r=\lfloor b \rfloor + 1}^{n} \left(1 - \frac{b}{r}\right),
$$

if $b \neq \lfloor b \rfloor$. Applying Lemma 4.3 (note that to apply this lemma is necessary to have $b/r < 1$, i.e., $r \geq \lfloor b \rfloor + 1$) we have

$$
\prod_{r=s+1}^{n} \left(1 - \frac{b}{r}\right) = \begin{cases} 0, & s < \lfloor b \rfloor, b = \lfloor b \rfloor, \\ O\left(\frac{\lfloor b \rfloor}{n}\right), & s < \lfloor b \rfloor, b \neq \lfloor b \rfloor, \\ 1 + O\left(\frac{\lfloor b \rfloor}{n}\right), & s \geq \lfloor b \rfloor. \end{cases}
$$

Using this and (4.29), formula (4.30) can be written as

$$
\mathbb{E}\tilde{N}(k, n + 1) = \sum_{s=\lfloor b \rfloor}^{n} g(k-1, s) \left(\frac{s}{n}\right)^b \left(1 + O\left(\frac{n-s}{sn}\right)\right) + O\left(\frac{\lfloor b \rfloor}{n}\right) + \varepsilon,
$$

(4.31)

where the error term $\varepsilon$ is of order $O(\ln n)$ if $m = 1$ and of order $O(k \ln n)$ if $m > 1$. It is not difficult to see that, following a similar procedure, we can get the same solution for (4.29).

Now, by Lemma 4.3 we know that there exist some $N_1(k) > 0$ and some $N_2(k) > 0$, such that $\tilde{N}(k, n)/n \to N_1(k)$ and $mk\tilde{N}(k, n)/[n(m + 1)] \to N_2(k)$ almost surely. Thus, by the dominated convergence theorem $p_k := \lim_{n \to \infty} \mathbb{E}\tilde{N}(k, n)/n$, and for $k \geq 2, g(k-1) := \lim_{n \to \infty} g(k-1, n), n$, exist.

Note that $g(k-1) = m(k-1)p_{k-1}/(m+1)$, and let us write $g(k-1, n) = g(k-1) + O(\varepsilon_n)$, where $\varepsilon_n \to 0$ as $n \to \infty$. Hence,

$$
\sum_{s=\lfloor b \rfloor}^{n} g(k-1, s) \left(\frac{s}{n}\right)^b = \frac{g(k-1)}{n^b} \sum_{s=\lfloor b \rfloor}^{n} s^b + \frac{1}{n^b} \sum_{s=\lfloor b \rfloor}^{n} O(\varepsilon_n) s^b,
$$

(4.32)

and using that $\sum_{s=\lfloor b \rfloor}^{n} s^b = n^{b+1}/(b+1) + o(n^{b+1})$ (see 3.II of [14]), we obtain

$$
\frac{\mathbb{E}\tilde{N}(k, n + 1)}{n + 1} = \frac{\frac{g(k-1)}{n^b}}{\frac{g(k-1)}{n^b}} + O\left(\frac{\ln n}{n}\right), \quad m = 1,
$$

$$
\frac{\mathbb{E}\tilde{N}(k, n + 1)}{n + 1} = \frac{\frac{g(k-1)}{n^b}}{\frac{g(k-1)}{n^b}} + O\left(\frac{\ln n}{n}\right), \quad m > 1.
$$

(4.33)

When $m > 1$ we need more restrictions on $k$ in order to determine the limit of $\mathbb{E}\tilde{N}(k, n + 1)/(n + 1)$. Indeed it should satisfy that $k \ln n/n \to 0$ as $n \to \infty$. Thus, by (4.33),

$$
p_k = \lim_{n \to \infty} \frac{\mathbb{E}\tilde{N}(k, n + 1)}{n + 1} = \begin{cases} \frac{m+1}{m(k+1)}, & k = 1, \\ \frac{m(k-1)}{m(k+1)}, & k > 1. \end{cases}
$$

(4.34)

Solving (4.34) recursively we get (4.22).
Proof of Theorem 4.2. We follow the approach of Dorogovtsev, Mendes, and Samukhin [6], that uses master equations for the expected value of the number of vertices with in-degree $k$. To obtain the exact equations we need to consider two stages. For the first one we consider what happens in one step of the process, during which the number of vertices of in-degree $k$ can be increased by counting also some vertices coming from those having previously in-degree $(k-1)$ or in-degree $(k+1)$, and then we consider what happens in $(m+1)$ steps, thus obtaining the change of the vertices in-degree in an interval of time starting when the process has $n$ vertices, until it has $(n+1)$ vertices. This part corresponds to finding the equations (4.3), (4.4), and (4.7) given in Lemma 4.2. For the second stage, we iterate the previous equations with respect to $n$ and obtain the limit of $\mathbb{E}N(n)/n$ as $n \to \infty$. This part was proved in Lemma 4.3 (determining (4.22)).

Finally, we use Azuma–Hoeffding inequality (2.21) to obtain (4.4). Let $F_t$ be the natural filtration generated by the process $\{\mathcal{N}_{k,t}\}$ up to time $t$. Then, in the same way as it was explained for to Simon model, Section 3.1, it is easy to show that for $s \leq t$, $Z_{s,PA} = \mathbb{E}(\mathcal{N}_{k,t}\mid F_s)$ is a martingale such that, $|Z_{s,PA} - Z_{s-1,PA}| \leq 1$, $Z_{t,PA} = \mathbb{E}\mathcal{N}_{k,t}$ and $Z_{0,PA} = \mathbb{E}\mathcal{N}_{k,PA}$ (at time $t = 0$ the random graph is the empty graph). Thus by (2.21) we get that for every $\epsilon_n \gg 1/\sqrt{n}$, e.g. take $\epsilon_n = \sqrt{\ln n/n}$,

$$
P\left(\frac{\mathcal{N}_{k,t}}{n} - \mathbb{E}\mathcal{N}_{k,t} \geq \epsilon_n\right) \leq \exp\left(-\frac{(n\epsilon_n)^2}{2}\right) \to 0,$$

as $n$ goes to infinity. Here $t = n(m+1) + i$, for $i = 0, 1, \ldots, m$. Thus we obtain that for $t = n(m+1)$,

$$
\frac{\mathcal{N}_{k,t}}{n} \to \frac{(1+1/m)\Gamma(k)\Gamma(2+1/m)}{\Gamma(k+2+1/m)},
$$
in probability as $n \to \infty$. However, by Lemma 4.3 we actually have an almost sure convergence.

4.2 Relation between Simon and Yule models

Bearing in mind the construction of Yule model as explained in (3.16), we underline that the inter-event times of in-links appearance and those related to creation of new vertices, are exponentially distributed. In order to relate Yule and Simon models we investigate here the inter-events times characterizing the Simon model, showing that a suitable rescaling in the limit leads to exponentially random variables. The idea is to identify the two processes which conditionally describe the Simon model, such that these converge to two independent Yule processes, which define a Yule model.

The next theorem together with Remarks 4.5 and 4.7 allows us to recognize the first process inside a Simon model behaving asymptotically as a Yule process with parameter $(1-\alpha)$, while Theorem 4.4 and Remark 4.7 determine the second process which behaves asymptotically as a Yule process with parameter equal to one. The first process models how the vertices get new in-links, thus at each moment a new vertex appears, a process starts. On the other hand, the second process corresponds to how the vertices appear.

Let $(G_n)_{n \geq 1}$ be the random graph process associated to Simon model of parameter $\alpha$, as described in Subsection 3.1, and let $\{d(v_i,t)\}_{i \geq 1}$ be the in-degree process associated to the vertex $v_i$, which appears at time $t_i$, i.e., $t_i = \min\{t : d(v_i,t) = 1\}$ (note that $d(v_i, t_i) = 1$ as the vertex appears together with a directed loop in the model).

Our first focus is on the study of the distributions of the waiting times between the instant in which each vertex has in-degree $k$, till that in which it has in-degree $k + 1$. Formally, we study the distribution of the random variables $W_k = t_{k+1} - t_k - 1$, $k \geq 1$, where $t_j = \min\{t : d(v_i, t) = j + 1\}$ for $j = 0, 1, 2, \ldots$

**Theorem 4.3.** Let $z = \ln \left(1 + \frac{1}{k-1}\right)$, $k \geq 1$, $x > 0$. It holds

$$
|\mathbb{P}(W_k \leq x) - \mathbb{P}(Z_k \leq z)| < O\left(\frac{1}{t_{k-1}}\right),
$$

where $Z_k$ is an exponential random variable of parameter $(1-\alpha)k$. 

18
Remark 4.5. Theorem 4.3 states that for any $t^*$ large enough but fixed, 
\[ \Pr(W^1_k \leq x) - \Pr(Z^1_k \leq x) < O\left(\frac{1}{t^*}\right). \] (4.36)

$\forall j \geq \min\{i : t^*_i \geq t^*\}$, and for $k \geq 1$. This means that from a fixed but large time $t^*$, all the waiting times $W^1_k$ are approximately exponential random variables, with an error term smaller than $O(1/t^*)$.

Proof of Theorem 4.3 By the preferential attachment probabilities of $(G^*_t)_{t \geq 1}$, for $x \geq 1$, we have
\[
\Pr[W^1_k = x] = \frac{\tilde{d}(v_i, t^*_k - 1) (1 - \alpha)}{t^*_k + x - 1} \left(1 - \frac{\tilde{d}(v_i, t^*_k - 1) (1 - \alpha)}{t^*_k + x - 1}\right) \ldots \left(1 - \frac{\tilde{d}(v_i, t^*_k - 1) (1 - \alpha)}{t^*_k + x - 1}\right)
= \left(\frac{k(1 - \alpha)}{t^*_k + x - 1}\right) \left(1 - \frac{k(1 - \alpha)}{t^*_k + x - 1}\right) \ldots \left(1 - \frac{k(1 - \alpha)}{t^*_k + x - 1}\right)
= \left(\frac{k(1 - \alpha)}{t^*_k + x - 1}\right) \prod_{r = t^*_k + 1}^{t^*_k + x - 2} \left(1 - \frac{k(1 - \alpha)}{r}\right),
\] (4.37)
as $t^*_k + 1 > k$. Then $k(1 - \alpha)/r < 1$, so we can apply Lemma 4.1 to the product to obtain
\[
\prod_{r = t^*_k + 1}^{t^*_k + x - 2} \left(1 - \frac{k(1 - \alpha)}{r}\right)
= \left(\frac{t^*_k + 1 - 1}{t^*_k + x - 1}\right)^{k(1 - \alpha)} \left(1 + O\left(\frac{x}{(t^*_k + x - 1)^2 + wt^*_k}ight)\right).
\] (4.38)

Thus, using (4.37), (4.38), the Euler–Maclaurin formula,
\[
\sum_{j=1}^{n} j^{-s} = \frac{n^{s-1}}{s-1} - \frac{1}{s} \int_{1}^{n} \frac{|y|}{y^{s+1}} dy,
\] with $s \in \mathbb{R} \setminus \{1\}$ (see [11]) and the fact that $|y| \leq y$, we arrive at
\[
\Pr[W^1_k \leq x] = \sum_{w=1}^{x} \left(\frac{k(1 - \alpha)}{t^*_k + w - 1}\right) \left(\frac{t^*_k + 1 - 1}{t^*_k + x - 1}\right)^{k(1 - \alpha)} \left(1 + O\left(\frac{w}{(t^*_k + x - 1)^2 + wt^*_k}\right)\right)
= k(1 - \alpha)(t^*_k + 1 - 1)^{k(1 - \alpha)} \sum_{w=1}^{x} \left(\frac{1}{t^*_k + w - 1}\right)^{k(1 - \alpha) + 1} \left(1 + O\left(\frac{w}{(t^*_k + x - 1)^2 + wt^*_k}\right)\right)
< \left(1 + O\left(\frac{1}{t^*_k}\right)\right)k(1 - \alpha)(t^*_k + 1 - 1)^{k(1 - \alpha)} \sum_{j=t^*_k + 1}^{t^*_k + x - 1} \left(\frac{1}{j}\right)^{k(1 - \alpha) + 1}
= \left(1 + O\left(\frac{1}{t^*_k}\right)\right)k(1 - \alpha)(t^*_k + 1 - 1)^{k(1 - \alpha)}
\times \left(\frac{1}{(t^*_k + x - 1)^{k(1 - \alpha)}} - \frac{1}{(t^*_k + 1)^{k(1 - \alpha)}} + (k(1 - \alpha) + 1) \int_{t^*_k + 1}^{t^*_k + x - 1} \frac{|y|}{y^{k(1 - \alpha) + 2}} dy\right)
< \left(1 + O\left(\frac{1}{t^*_k}\right)\right)k(1 - \alpha)(t^*_k + 1 - 1)^{k(1 - \alpha)} \left[\left(\frac{1}{(t^*_k + 1)^{k(1 - \alpha)}} - \frac{1}{(t^*_k + x - 1)^{k(1 - \alpha)}}\right)\right]
= \left(1 + O\left(\frac{1}{t^*_k}\right)\right)\left[1 - \exp \left(- k(1 - \alpha) \ln \left(1 + \frac{x}{(t^*_k + 1 - 1)}\right)\right)\right] \left(1 + O\left(\frac{1}{t^*_k}\right)\right).
\] (4.39)

Thus, we get
\[
\Pr(W^1_k \leq x) - \left[1 - \exp \left(- k(1 - \alpha) \ln \left(1 + x/(t^*_k + 1 - 1)\right)\right)\right] < O\left(\frac{1}{t^*_k}\right).
\]
Then, by taking \( z = \ln(1 + x/(t'_{k-1} - 1)) \), it holds

\[
\left| \mathbb{P}(W_k^t \leq x) - \mathbb{P}(Z_k^t \leq z) \right| < O\left( \frac{1}{t'_{k-1}} \right),
\]

where \( Z_k^t \) is a random variable exponentially distributed with parameter \((1 - \alpha)k\). \( \square \)

**Remark 4.6.** Note that the knowledge of \( \{W_k^t\}, k \geq 1 \) is equivalent to the knowledge of \( \bar{d}(v_i, t) \), as \( \bar{d}(v_i, t) := \min\{k: \sum_{j=1}^{k} W_k^j > (t - t_0^i)\} \). Thus, due to Theorem 4.3, the process 
\( \{\bar{d}(v_j, t)\}_{t \geq t_0^j}, \forall j \geq \min\{i: t_0^i \geq t_0\} \), behaves asymptotically as a Yule process with parameter \((1 - \alpha)\).

Let us now consider the growth of the vertices in a Simon model. At each instant of time \( t \), a new vertex is created with a fixed probability \( \alpha \). This fact can be re-thought from a different perspective. Remember that in Simon model the number of vertices at time \( t \) is a random variable \( V(t) \), distributed Binomially, Bin\((t, \alpha)\), and that at each instant of time, one and only one vertex can appear. Then, conditionally on \( V(t) \), suppose that with probability proportional to the number of vertices, i.e., with probability \( \alpha/V(t) \), the new vertex is originated by the \( i \)th existing vertex. Thus, since there are \( V(t) \) vertices, the probability of the birth of a new vertex is \( V(t)/(\alpha/V(t)) = \alpha \).

Let \((G_k^t)_{t \geq 1}\) be the random graph process corresponding to Simon model (described in Subsection 3.1). For each vertex in this process, say \( v_i, i = 1, \ldots, V(t) \), let \( Y_k^i \) be the random variables \( Y_k^i := \ell_k^i - \ell_k^{i-1} \), for \( k = 1, 2, \ldots \), where \( \ell_k^0 = \min\{t: \bar{d}(v_i, t) = 1\} \), and \( \ell_k^i \) is the instant at which the \( i \)th vertex originates its \( k \)th vertex, \( k \geq 1 \). Hence \( Y_k^i \) represents the waiting time between the appearance of the \((k - 1)\)th and the \( k \)th vertex originated by the \( i \)th vertex.

**Theorem 4.4.** Let \( z = \ln(1+y/(\ell_{k-1}^i - 1)) \), \( k \geq 1 \), \( y > 0 \), and \( 0 < \varepsilon_i < 1 \) such that \( t\varepsilon_i^2 \to \infty \) as \( t \to \infty \). Then,

\[
\left| \mathbb{P}(Y_k^i \leq y) - \mathbb{P}(Z_k^i \leq z) \right| < O\left( \frac{1}{\ell_{k-1}^i t\varepsilon_i^{2}} \right), \tag{4.40}
\]

where \( Z_k^i \) is an exponentially distributed random variable of parameter one.

**Remark 4.7.** The previous theorem states that for any \( t\varepsilon_i^2 \) large enough but fixed,

\[
\left| \mathbb{P}(Y_k^i \leq y) - \mathbb{P}(Z_k^i \leq z) \right| < O\left( \frac{1}{t\varepsilon_i^2} \right), \tag{4.41}
\]

\( \forall j \geq \min\{i: \ell_k^j \geq \ell^* \} \), and for \( k \geq 1 \). In words it means that from a fixed but large time \( \ell^* \), all the waiting times \( Y_k^i \) are approximately exponential random variables, with an error term smaller than \( O\left( 1/((t\varepsilon_i^2)^2) \right) \). Thus, for \( \ell^* \) large enough we start to see a process of creation of vertices which is very close to a Yule process with parameter one.

**Proof of Theorem 4.4.** Note that if \( Y_k^i \) represents the waiting time between the \((k - 1)\)th and the \( k \)th vertex originated by \( v_i, i = 1, \ldots, V(t) \). Then the event \( \{Y_k^i = y\} \) is equivalent to the event \( \{X_{i,j}^k = 0, X_{i,j}^{k+1} = 0, \ldots, X_{i,y}^k = 1\} \), where \( X_{i,j}^k = 1 \), if \( v_i \) originates its \( k \)th vertex at time \( t'_{k-1} + j \), and \( X_{i,y}^k = 0 \), if not.

Now define the events \( E_i := \{t(\alpha - \varepsilon_i) \leq V(t) \leq t(\alpha + \varepsilon_i)\} \). By Chebyshev’s inequality we have \( \mathbb{P}(E_i) \leq \alpha(1 - \alpha)/\varepsilon_i^2 \), so \( \mathbb{P}(E_i) \to 1 \) if \( t\varepsilon_i^2 \to \infty \) as \( t \to \infty \). Then observe that

\[
\mathbb{P}(X_{i,j}^k = x) = \mathbb{P}(X_{i,j}^k = x \mid E_{i,k-1+j-1}^t) + O\left( \frac{1}{(t_{k-1}^i + j - 1)\varepsilon_i^2} \right),
\]

and

\[
\mathbb{P}(Y_k^i = y) = \left[ \mathbb{P}(X_{i,y}^k = 1 \mid E_{i,k-1+y-1}^t) \prod_{j=1}^{y-1} \mathbb{P}(X_{i,j}^k = 0 \mid E_{i,k-1+j-1}^t) \right] + O\left( \frac{1}{(t_{k-1}^i)\varepsilon_i^2} \right).
\]
Assuming that $\varepsilon_{i_{k-1}+x-1} > \varepsilon_{i_{k-1}+x-2} > \cdots > \varepsilon_{i_{k-1}}$, we get

\[
\mathbb{P}(Y_i^k = y) < \frac{\alpha}{(t_{k-1}^i + y - 1)(\alpha - \varepsilon_{i_{k-1}+y-1})} \left( 1 - \frac{\alpha}{(t_{k-1}^i + y - 2)(\alpha - \varepsilon_{i_{k-1}+y-2})} \right) \times \\
\cdots \times \left( 1 - \frac{\alpha}{t_{k-1}^i(\alpha - \varepsilon_{i_{k-1}})} \right) + O\left(\frac{1}{t_{k-1}^{i} \varepsilon_{i_{k-1}}^2}\right) \\
= \frac{\alpha}{(t_{k-1}^i + y - 1)(\alpha - \varepsilon_{i_{k-1}+y-1})} \prod_{r=i_{k-1}^i}^{i_{k-1}^i+y-2} \left( 1 - \frac{\alpha}{r(\alpha - \varepsilon_{r})} \right) + O\left(\frac{1}{t_{k-1}^{i} \varepsilon_{i_{k-1}}^2}\right),
\]

and

\[
\mathbb{P}(Y_i^k = y) > \frac{\alpha}{(t_{k-1}^i + y - 1)(\alpha - \varepsilon_{i_{k-1}+y-1})} \left( 1 - \frac{\alpha}{(t_{k-1}^i + y - 2)(\alpha - \varepsilon_{i_{k-1}+y-2})} \right) \times \\
\cdots \times \left( 1 - \frac{\alpha}{t_{k-1}^i(\alpha - \varepsilon_{i_{k-1}})} \right) + O\left(\frac{1}{t_{k-1}^{i} \varepsilon_{i_{k-1}}^2}\right) \\
= \frac{\alpha}{(t_{k-1}^i + y - 1)(\alpha - \varepsilon_{i_{k-1}+y-1})} \prod_{r=i_{k-1}^i}^{i_{k-1}^i+y-2} \left( 1 - \frac{\alpha}{r(\alpha - \varepsilon_{r})} \right) + O\left(\frac{1}{t_{k-1}^{i} \varepsilon_{i_{k-1}}^2}\right).
\]

Thus, in a similar manner as we did in the proof of Theorem 4.3 by using Lemma 4.1 and Euler–Maclaurin formula to (4.42) and (4.43), we find that

\[
\mathbb{P}(Y_i^k \leq y) - [1 - \exp(-\ln(1 + y/(t_{k-1}^i - 1)))] < O\left(\frac{1}{t_{k-1}^{i} \varepsilon_{i_{k-1}}^2}\right).
\]

Then, taking $z = \ln(1 + y/(t_{k-1}^i - 1))$, we obtain

\[
\left| \mathbb{P}(Y_i^k \leq y) - \mathbb{P}(Z_i^k \leq z) \right| < O\left(\frac{1}{t_{k-1}^{i} \varepsilon_{i_{k-1}}^2}\right),
\]

where $Z_i^k$ is an exponentially distributed random variable with parameter one. The last step consists of showing the convergence of the minimum of the variables $Y_i^k$ with respect to all the vertices $v_i$, $i = 1, \ldots, V(t)$, already present in the process. Therefore,

\[
\mathbb{P}\left( \min_i Y_i^k \leq y \right) = 1 - \left( 1 - \mathbb{P}(Y_i^k \leq y) \right)^{V(t)} \\
= 1 - \left( 1 - \mathbb{P}(Z_i^k \leq z) + \varepsilon \right)^{V(t)},
\]

where $\varepsilon < O\left(1/(t_{k-1}^i \varepsilon_{i_{k-1}}^2)\right)$. Then, we have

\[
\mathbb{P}\left( \min_i Y_i^k \leq y \right) = 1 - \left( 1 - \mathbb{P}(Z_i^k \leq z) \right)^{V(t)} + \varepsilon = \mathbb{P}\left( \min_i Z_i^k \leq z \right) + \varepsilon,
\]

which proves the thesis.

5 Discussion and conclusions

To compare the Barabási–Albert and Simon models, we considered a third model that we called here the II-PA model, first introduced in [12] with a different name. Then we gave a common description of the three models by introducing three different random graph processes related to them. This representation allowed us to clarify in which sense the three models...
can be related. For each fixed time, if $m = 1$, we proved that the Barabási–Albert and the II-PA models have exactly the same preferential attachment probabilities (Theorem 4.1). Furthermore, since in the first model the preferential attachment is meant with respect to the whole degree of each vertex while in the second case it is meant with respect only to the in-degree, the conclusion is that, for a uniformly selected random vertex, the degree distribution in the Barabási–Albert model equals the in-degree distribution in the Simon model. Note that $m = 1$ is the only case in which this is true.

Since the direct comparison between Barabási–Albert and Simon model is not possible we first compared II-PA model with Barabási–Albert model (Theorem 4.1), and then II-PA model with Simon model (Theorem 4.2). We underline that, even if the introduction of II-PA model was functional to the study of the connections between the Barabási–Albert and Simon models, this hybrid model is interesting in itself.

Regarding the connections between Simon and II-PA models, Theorem 4.2 shows that when time goes to infinity, the II-PA model has the same limiting in-degree distribution as that of the Simon model with parameter $\alpha = 1/(m + 1)$, for any $m \geq 1$. The proof uses the Azuma–Hoeffding concentration inequality and the supermartingale’s convergence theorem.

Combining these theorems we conclude that, in the limit, the Simon model has the same in-degree distribution as that of the Barabási–Albert model, for $\alpha = 1/2$ and $m = 1$. The existing relations between the three models are summarized in Figure 3.

On the other hand, Yule model is defined in continuous time. In Section 4.2 we gave a mathematical explanation of the reason why, when time goes to infinity the distribution of the size of a genus selected uniformly at random in the Yule model coincide with the in-degree distribution of Simon model. More precisely, Theorem 4.3 and Theorem 4.4 show that, as time flows, the Simon model approximates the behavior of a continuous time process that in fact corresponds to a Yule model with parameters $(1 - \alpha, 1)$. This result is obtained in probability.

Figure 3: The relations between Simon, II-PA, and Barabási–Albert (B–A in the picture) models. Note that II-PA and Barabási–Albert models can be put in relation for any time $t$ but just in the case $m = 1$. Instead, the connections between II-PA and Simon models and Simon and Barabási–Albert models, respectively, hold in the limit for $t$ going to infinity.

References

[1] A. Barabási and R. Albert. Emergence of Scaling in Random Networks. *Science*, 286(5439):509–512, 1999.

[2] P. Billingsley. *Probability and Measure*. Wiley, New York, NY, 3rd edition, 1995.

[3] 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.

[4] S. Bornholdt and H. Ebel. World Wide Web scaling exponent from Simon’s 1955 model. *Phys. Rev. E*, 64:035104, 2001.

[5] K.S. Crump. On Point Processes Having an Order Statistic Structure. *Sankhya, Series A*, 37(3):396–404, 1975.

[6] 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.
[7] R. Durrett. *Random Graph Dynamics*. Cambridge University Press, New York, NY, USA, 2006.

[8] P.D. Feigin. On the Characterization of Point Processes with the Order Statistic Property. *Journal of Applied Probability*, 16(2):297–304, 1979.

[9] S. Janson, T. Łuczak, and A. Ruciński. *Random Graphs*. Wiley, 2011.

[10] 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.

[11] M. Neuts and S.I. Resnick. On the times of birth in a linear birth process. *Journal of the Australian Mathematical Society*, 12:473–475, 1971.

[12] M.E.J. Newman. Power laws, Pareto distributions and Zipf’s law. *Contemporary Physics*, 46(5):323–351, 2005.

[13] P.S. Puri. On the Characterization of Point Processes with the Order Statistic Property without the Moment Condition. *Sankhya: Journal of Applied Probability*, 19(1):39–51, 1982.

[14] R. Sedgewick and P. Flajolet. *An Introduction to the Analysis of Algorithms*. Addison-Wesley, 2013.

[15] M.V. Simkin and V.P. Roychowdhury. Re-inventing Willis. *Physics Reports*, 502(1):1–35, 2011.

[16] H.A. Simon. On a class of skew distribution functions. *Biometrika*, 42(3–4):425–440, 1955.

[17] E.A. Thompson. *Human Evolutionary Trees*. Cambridge University Press, New York, NY, USA, 1975.

[18] 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.