Article

Analysis of Basic Features in Dynamic Network Models

Pedro J. Zufiria [1,2,*] and Iker Barriales-Valbuena [1]

[1] Depto. Matemática Aplicada a las TIC, ETSI Telecomunicación, Universidad Politécnica de Madrid, Avda. Complutense 30, E-28040 Madrid, Spain; iker.barriales.valbuena@gmail.com

[2] Information Processing and Telecommunications Center (IPTC), Universidad Politécnica de Madrid, E-28040 Madrid, Spain

* Correspondence: pedro.zufiria@upm.es

Abstract: Time evolving Random Network Models are presented as a mathematical framework for modelling and analyzing the evolution of complex networks. This framework allows the analysis over time of several network characterizing features such as link density, clustering coefficient, degree distribution, as well as entropy-based complexity measures, providing new insight on the evolution of random networks. First, some simple dynamic network models, based only on edge density, are analyzed to serve as a baseline reference for assessing more complex models. Then, a model that depends on network structure with the aim of reflecting some characteristics of real networks is also analyzed. Such model shows a more sophisticated behavior with two different regimes, one of them leading to the generation of high clustering coefficient/link density ratio values when compared with the baseline values, as it happens in many real networks. Simulation examples are discussed to illustrate the behavior of the proposed models.

Keywords: complex networks; stochastic modelling; entropy; estimation

1. Introduction

A large variety of complex systems can be analyzed by constructing a model that relies on some network structure [1–4]. The model may be dynamical, meaning that the values of some (state) variables do change with time and, depending on the nature of such variables, we can have different types of network models. The first type corresponds to dynamic graphs that follow evolution laws defined explicitly on the network [5–8]; the second type gathers dynamical systems where the state variables are defined on a network [9,10]; finally, the third type refers to co-evolution models that combine evolving networks and dynamical systems. In the first and third type, the underlying network structure changes with time, defining a time-varying or evolving network [11,12]. In the present work, we first characterize the basic features of some simple models of evolving networks whose evolution does not depend on network structure; the time evolution of these features serves as a reference baseline signature of the behavior of simple models. Then, a model that makes use of network structure is proposed to reflect some real network characteristics. The analysis of this model shows several regimes that indicate a sophisticated behavior; for some regime, the network reaches a high clustering coefficient/link density ratio [13] (when compared to the ratio values of baseline signatures), a common feature in many real networks.

The paper is organized as follows: Section 2 presents the general framework for Dynamic Network Models and their characterization via some basic features, whereas entropy measures are shown in Section 3. Section 4 analyzes some simple evolution models whose basic features' time evolution serves as a behavior reference baseline. More elaborated evolution models that depend on network

structure are studied in Section 5. Simulations of Section 6 comparatively illustrate the time evolution of the different features for the proposed models. Finally, concluding remarks are summarized in Section 7.

2. Characterization of Network Sequences via Basic Features

Following [12], discrete-time network evolution over time can be generally defined by a random sequence or trajectory $\{G_t\}_{t=0,1,...}$, where each G_t can take values g from $\mathcal{G}$, being $\mathcal{G}$ the set of all possible networks. The analysis of $\{G_t\}_{t=0,1,...}$ can be framed by considering it as a stochastic process, whose full characterization may be very complex. In the following, we present some basic features that help for a partial characterization of such stochastic process.

Time Evolution of Network Features

In some cases, we may be interested in the evolution of some quantifiable properties or features, f, of the network, defined as follows (see [14] for details):

$$f : \mathcal{G} \longrightarrow \mathbb{R}^l, \tag{1}$$
$$g \longrightarrow f = f(g), \tag{2}$$

where $f(g)$ is the function that computes such quantifiable property (number of links, number of triangles, connectivity, degree of nodes, entropy of degree distribution, etc.) in graph g.

Note that, when $\mathcal{G}$ is endowed with a probability space, then, under some regularity assumptions on f, this function defines a random vector. Therefore, the sequence $f(G_t) \in \mathbb{R}^l$ defines a vector stochastic process that can be analyzed using standard stochastic process techniques. In the following analysis, we will focus on several of these properties such as the number of links, number of triangles, the connectivity and the degree distribution entropy (a scalar summary of the distribution vector). Since for these cases $l = 1$, the study will boil down to the analysis of scalar stochastic processes. A basic analysis would estimate, for instance, the deterministic sequence of expected values $E[f(G_t)]$.

In the following section, we focus on different entropy measures that can also be employed for characterizing the stochastic process $\{G_t\}_{t=0,1,...}$.

3. Entropy Measures for Stochastic Processes

The stochastic process $\{G_t\}_{t=0,1,...}$ is an indexed sequence of random variables, which can be completely characterized until time instant $t = T$ by its joint probability distribution

$$P(G_0, G_1, ..., G_T). \tag{3}$$

This joint distribution may be quite complex to study and, therefore, we may acquiesce in characterizing part of it. For instance, if we consider G_i for a fixed time $t = i$, this snapshot of the process, also called a cross sectional variable, can be represented by a "static" model such as the ones studied in [14], fully characterized by the marginal distribution of g_i. Accordingly, when considering entropy measures for characterizing a stochastic process, different distributions associated with such process can be considered, as developed below.

3.1. Snapshot Entropy and Entropy of Network Features

The simplest approach focuses on the entropy analysis of cross sectional variables G_i. Hence, one can define the snapshot entropy of index i, $H(G_i)$, of a stochastic process as the entropy of the i-th variable G_i of the process

$$H(G_i) = - \sum_{g \in \mathcal{G}} p(G_i = g) \log p(G_i = g). \tag{4}$$

When considering a network feature f, the entropy of the associated random variable $F_i = f(G_i)$ satisfies the condition

$$H(G_i) = H(G_i \mid F_i) + H(F_i), \qquad (5)$$

and therefore

$$H(F_i) = H(f(G_i)) \leq H(G_i), \qquad (6)$$

where the equality holds only if f is an injection.

Note that $H(G_i)$ in (4) is not to be confused with the feature mentioned in Section 2 called *degree distribution entropy*, associated with a concrete sample of G_i. For a more detailed explanation of degree distributions in static models, see [14].

The computation of $H(G_i)$, when performed for every $i \in \{0, 1, \ldots, \}$, would lead to a deterministic time series $\{H_t\}_{t=0,1,\ldots}$ as an alternative partial characterization of the stochastic process $\{G_t\}_{t=0,1,\ldots}$.

3.2. Trajectory Entropy

Furthermore, one can study the entropy of a whole time period evolution of the process, seen as a sequence of $T + 1$ variables. We define the trajectory entropy (H_0^T) of a $T + 1$-length time period of a stochastic process, as the entropy of the joint probability $P(G_0, G_1, ..., G_T)$:

$$H_0^T = H(G_0, \ldots, G_T) = - \sum_{\mathcal{G}^{T+1}} p(g_0, g_1, \ldots g_T) \log p(g_0, g_1, \ldots g_T). \qquad (7)$$

If all G_i are independent variables, then:

$$H_0^T = \sum_{i=0}^{T} H(G_i). \qquad (8)$$

Note that, in general, as T increases, H_0^T may increase unbounded.

3.3. Normalized Asymptotic Entropy

Finally, one may want to characterize the entropy rate as a normalized entropy measure independent of T, which globally characterizes the asymptotic behavior of the stochastic process. This entropy rate is defined as

$$H^R = \lim_{T \to \infty} \frac{1}{T+1} H_0^T, \qquad (9)$$

whenever such limit exists. Alternatively, we can also compute

$$H'^R = \lim_{T \to \infty} H(G_T \mid G_{T-1}, G_{T-2}, ..., G_1, G_0), \qquad (10)$$

again when this limit does exist. For strongly stationary processes, both measures (9) and (10) do exist and they are equal.

After presenting these measures, some basic evolution models are illustrated in the next section.

4. Basic Evolution Models with a Fixed Number of Nodes: Evolution of Number of Links

Let us consider G_V the set of all networks (or graphs) $g_i = (V, E_i)$ having a fixed set of nodes $V = \{v_1, \ldots, v_N\}$, with $|V| = N$; each $g_i \in G$ is then characterized by its corresponding set of links $E_i \subset E$ with E being determined by V as the set of all pairs of nodes ($|E| = \binom{N}{2} = M$).

In this framework, any evolution process $\{G_t\}_{t=0,1,\ldots}$ is characterized by the sequence of the corresponding $\{E_t\}_{t=0,1,\ldots}$. In addition, since $g_i \in G_V$ can be represented via its corresponding binary adjacency matrix $A(g_i) \in \mathbb{R}^n \times \mathbb{R}^n$, the evolution process can also be characterized as a sequence of adjacency matrices $\{A(g_t)\}_{t=0,1,\ldots} = \{A_t\}_{t=0,1,\ldots}$.

4.1. Evolution of the Number of Links

In general, a complete characterization of $\{G_t\}_{t=0,1,...}$ will be very cumbersome. Alternatively, we can partially characterize such process by considering

$$f : \mathcal{G} \longrightarrow \{0, 1, \ldots M\}, \tag{11}$$

$$g_i \longrightarrow f(g_i) = |E_i| = \|A_i\|_1 = m_i, \tag{12}$$

where f is the function that computes the number of links in the network. We can partition the set $\mathcal{G}_V$ into equivalence classes C_k, $k = 0, \ldots M$ so that each class C_k gathers all graphs containing k links: $C_k = \{g_i \in G : f(g_i) = k\}$. Then, we can define a stochastic process $\{M_t\}_{t=0,1...}$ with each $M_t \in \{0, 1, \ldots, M\}$ which characterizes the transition between classes, and whose state space represents such equivalence classes (hence, we identify C_k with state k).

In general, for a given instant of time i, based on (5), we have that the snapshot entropy of G_i and the entropy of M_i will satisfy

$$H(G_i) = H(G_i \mid M_i) + H(M_i) \tag{13}$$

and this relationship will help to characterize G_i via the analysis of M_i. Therefore, the following proposed models will be partially characterized by analyzing the associated stochastic process, $M_t \in \{0, 1, \ldots, M\}$, for the evolution of the number of links.

4.2. A Simple Structure Independent Evolution Model

We define a simple network evolution process that may serve as a reference baseline for comparison purposes. Given g_t (equivalently, E_t or A_t), the next time step network g_{t+1} is generated by randomly selecting a pair of nodes $(v_i, v_j) \in E$ so that, if there exists a link between them (i.e., $(v_i, v_j) \in E_t$), such link is removed ($E_{t+1} = E_t \setminus \{(v_i, v_j)\}$) and, if there is no link between the nodes (i.e., $(v_i, v_j) \notin E_t$), then it is created ($E_{t+1} = E_t \cup \{(v_i, v_j)\}$). Note that if we consider the adjacency matrix representation A_t, at each stage of time, an element of the matrix A_t is randomly chosen so that its value is changed (from 0 to 1 or vice versa) to derive A_{t+1}.

Note that the evolution law is determined by the number of links of g_t. Therefore, as mentioned above, we will start the analysis of this evolution model by characterizing the time evolution of the number of links. The corresponding $M_t \in \{0, 1, \ldots, M\}$ satisfies:

$$P(M_{t+1} = 1 \mid M_t = 0) = 1, \tag{14}$$

$$P(M_{t+1} = M - 1 \mid M_t = M) = 1, \tag{15}$$

and for $i \in \{1, \ldots, M-1\}$:

$$P(M_{t+1} = j \mid M_t = i) = \begin{cases} 0, & \text{if } j = i \text{ or } |j - i| > 1, \\ \frac{i}{M}, & \text{if } j = i - 1, \\ \frac{M-i}{M}, & \text{if } j = i + 1. \end{cases} \tag{16}$$

This process is a Markov chain with transition probability matrix

$$
P = \begin{bmatrix}
0 & \frac{1}{M} & 0 & \cdots & 0 & 0 & 0 \\
1 & 0 & \frac{2}{M} & \cdots & 0 & 0 & 0 \\
0 & \frac{M-1}{M} & 0 & \cdots & 0 & 0 & 0 \\
0 & 0 & \frac{M-2}{M} & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & \frac{M-2}{M} & 0 & 0 \\
0 & 0 & 0 & \cdots & 0 & \frac{M-1}{M} & 0 \\
0 & 0 & 0 & \cdots & \frac{2}{M} & 0 & 1 \\
0 & 0 & 0 & \cdots & 0 & \frac{1}{M} & 0
\end{bmatrix}, \tag{17}
$$

which is known as the Ehrenfest model [15], and which can be similarly interpreted as representing an urn with white and black balls, where we randomly select a ball and change it by another ball with different color, hence representing a sort of discrete-time birth–death Markov process [16] but with finite number of states (two boundary conditions). Many discrete distributions have been obtained by studying urn models and Markov processes [17–19]. Note that these models can be seen as a reference baseline since they do not exploit the network structure properties (i.e., the relative location of white balls and black balls).

The left stochastic, tri-diagonal, irreducible matrix P of Equation (17) has period 2, but it has a unique eigenvector associated with eigenvalue $\lambda = 1$. This eigenvector defines the stationary distribution of the process, denoted by M_s, and it can be easily proved that such distribution is binomial:

$$
P(g_s \in C_k) = P(M_s = k) = \binom{M}{k} \left(\frac{1}{2}\right)^M, \tag{18}
$$

so that taking a snapshot of the process for large t is equivalent to generating a sample from the Gilbert model [20] with $p = \frac{1}{2}$ or, equivalently, the uniform model with maximum entropy (see [14] for details). Note that, given a number of links $M_s = k$, the distribution of $G_i \mid (M_s = k)$ is uniform, each link having probability $\frac{1}{|C_k|} = \frac{1}{\binom{M}{k}}$. Hence, considering (18), the entropy expression provided in (13) becomes

$$
H(G_i) = H(M_s) + H(G_i \mid M_s) = -\sum_k p(k) \log p(k) - \sum_k p(k) \log \frac{1}{\binom{M}{k}} \tag{19}
$$

$$
= -\sum_k p(k) \log \frac{p(k)}{\binom{M}{k}} = M \cdot \log 2 = M = \binom{N}{2}, \tag{20}
$$

measuring the entropy in bits.

Concerning the entropy of M_t, it is known that Ehrenfest model snapshot (relative) entropy at time t, defined in terms of the Kullback–Leibler divergence between the distribution and the steady state equilibrium distribution

$$
H_{\text{rel}}(t) = -D_{KL}(P(t)||P_s) = -\sum_{k=0}^{M} P(M_t = k), \log \frac{P(M_t = k)}{P(M_s = k)} \tag{21}
$$

is non-decreasing in time as approaches the maximum value zero, upon the so called *H-Theorem* [21].

4.3. Extensions of the Model for Asymmetric Evolution

One can extend the symmetric model provided in (17) with the aim of considering cases in which the network may have an uneven tendency to increase or decrease the number of edges.

Let us consider the following transition behavior from g_t to g_{t+1}: we start selecting a pair of nodes in network g_t; if the selected pair already has an associated link, such link is removed with probability p_r, whereas, if such pair does not have an associated link, a link is added between such pair of nodes

with probability p_a. If no change (removal or addition) happens, the process is repeated until the network undergoes some modification, which is registered in g_{t+1}.

Again, if we focus the analysis on the time evolution of the number of links, M_t, the corresponding transition matrix becomes:

$$
P(p_r, p_a) = \begin{bmatrix}
0 & \frac{p_r}{p_r+(M-1)p_a} & 0 & \cdots & 0 & 0 & 0 \\
1 & 0 & \frac{2p_r}{2p_r+(M-2)p_a} & \cdots & 0 & 0 & 0 \\
0 & \frac{(M-1)p_a}{p_r+(M-1)p_a} & 0 & \cdots & 0 & 0 & 0 \\
0 & 0 & \frac{(M-2)p_a}{2p_r+(M-2)p_a} & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & \frac{(M-2)p_r}{(M-2)p_r+2p_a} & 0 & 0 \\
0 & 0 & 0 & \cdots & 0 & \frac{(M-1)p_r}{(M-1)p_r+p_a} & 0 \\
0 & 0 & 0 & \cdots & \frac{2p_a}{(M-2)p_r+2p_a} & 0 & 1 \\
0 & 0 & 0 & \cdots & 0 & \frac{p_a}{(M-1)p_r+p_a} & 0
\end{bmatrix}. \tag{22}
$$

The analysis of this system can be simplified if we denote $\frac{p_r}{p_a} = u$ the *unbalance* coefficient, since the matrix can be reformulated as

$$
P(u) = \begin{bmatrix}
0 & \frac{u}{u+M-1} & 0 & \cdots & 0 & 0 & 0 \\
1 & 0 & \frac{2u}{2u+M-2} & \cdots & 0 & 0 & 0 \\
0 & \frac{M-1}{u+M-1} & 0 & \cdots & 0 & 0 & 0 \\
0 & 0 & \frac{M-2}{2u+M-2} & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & \frac{(M-2)u}{(M-2)u+2} & 0 & 0 \\
0 & 0 & 0 & \cdots & 0 & \frac{(M-1)u}{(M-1)u+1} & 0 \\
0 & 0 & 0 & \cdots & \frac{2}{(M-2)u+2} & 0 & 1 \\
0 & 0 & 0 & \cdots & 0 & \frac{1}{(M-1)u+1} & 0
\end{bmatrix}. \tag{23}
$$

If $u < 1$, the model has more tendency to add links than to remove them, and vice versa for $u > 1$. The analysis and interpretation of the network behavior can be performed either way due to such symmetry. For instance, if $u < 1$, the model can be interpreted as characterizing the following behavior: if the selected pair in g_t has an associated link, this link is removed with probability u; if the pair does not have an associated link, then a link is added. Again, the selection procedure is repeated until a link is either removed or added, defining g_{t+1}.

It can be proved that the resulting stationary distribution has the form:

$$
P_u(M_s = k) = \frac{\binom{M}{k}\dfrac{k \cdot u + M - k}{M \cdot u^k}}{\displaystyle\sum_{i=0}^{M} \binom{M}{i}\dfrac{i \cdot u + M - i}{M \cdot u^i}}, \quad u \in \mathbb{R}^+, \tag{24}
$$

which can be seen as a generalization of the binomial distribution $\mathrm{Bin}(\frac{1}{2}, M)$ via the new parameter u.

Repeating a similar procedure to (19) and (20), the corresponding G_i entropy can be computed as

$$
H_u(G_i) = H_u(M_s) + H(G_i \mid M_s) = -\sum_k p_u(k) \log p_u(k) - \sum_k p_u(k) \log \frac{1}{\binom{M}{k}}
$$

$$
= -\sum_k p_u(k) \log \frac{p_u(k)}{\binom{M}{k}},
$$

which for $u = 1$ becomes $H_{u=1}(G_i) = \binom{N}{2} = M$.

Figure 1 represents smoothed probability mass functions for the baseline, theoretical given by (24) and empirical (based in simulations) with $p_a = 0.3$ and $p_r = 1$. Note that asymmetry of the u value generates a probability function with less entropy than the corresponding to the baseline mass function.
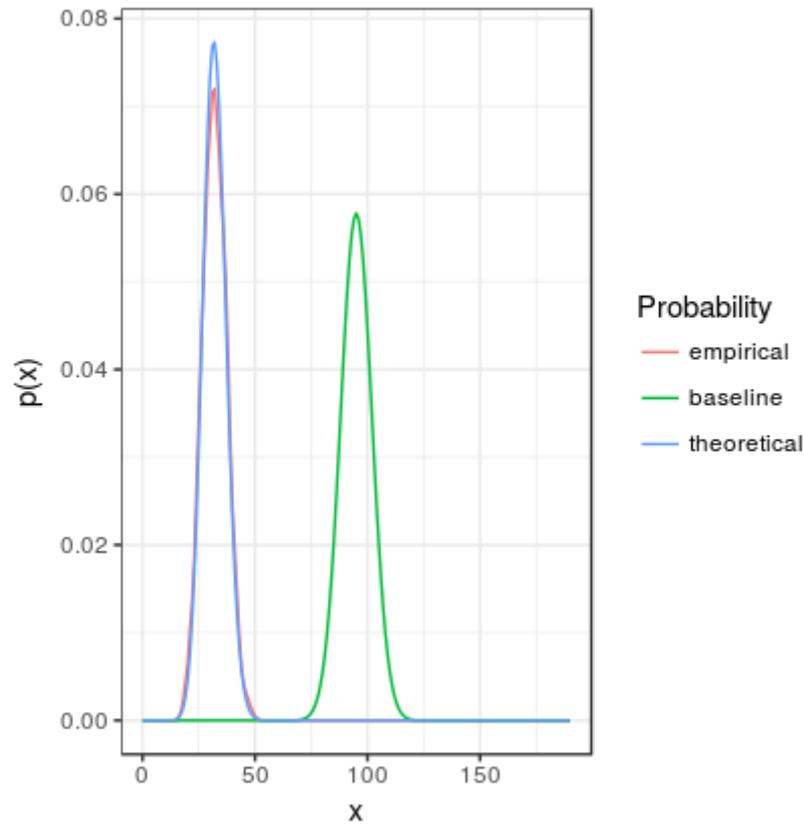

Figure 1. Simple and extended asymmetric models. Comparison among smoothed probability mass functions: baseline, theoretical and empirical for $p_a = 0.3$ and $p_r = 1$.

Alternative Simple Model

Another simple model could assume that, whenever an existing edge is selected to be removed, it is removed with probability $p_r \in [0,1]$, whereas, alternatively, a new edge is randomly added. The transition matrix of the corresponding $M_t \in \{0, 1, \ldots, M\}$ for the number of links would be

$$
P = \begin{bmatrix}
0 & \frac{p_r}{M} & 0 & \cdots & 0 & 0 & 0 \\
1 & 0 & \frac{2p_r}{M} & \cdots & 0 & 0 & 0 \\
0 & 1 - \frac{p_r}{M} & 0 & \cdots & 0 & 0 & 0 \\
0 & 0 & 1 - \frac{2p_r}{M} & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & \frac{(M-2)p_r}{M} & 0 & 0 \\
0 & 0 & 0 & \cdots & 0 & \frac{(M-1)p_r}{M} & 0 \\
0 & 0 & 0 & \cdots & 1 - \frac{(M-2)p_r}{M} & 0 & 1 \\
0 & 0 & 0 & \cdots & 0 & 1 - \frac{(M-1)p_r}{M} & 0
\end{bmatrix}. \tag{25}
$$

Note that an equivalent symmetric model can be defined as follows. If the selected pair of nodes does not have an associated link, we add such a link with probability p_a; otherwise, an existing link is removed.

It can be proved that the resulting stationary distribution has the form:

$$
P_{p_r}(M_s = k) = \begin{cases} \dfrac{1}{1 + \displaystyle\sum_{i=1}^{M} \dfrac{M \cdot (M - p_r) \cdots (M - (i-1)p_r)}{i!\, p_r^i}} & \text{if } k = 0, \\[2em] \dfrac{\dfrac{M \cdot (M - p_r) \cdots (M - (k-1)p_r)}{k!\, p_r^k}}{1 + \displaystyle\sum_{i=1}^{M} \dfrac{M \cdot (M - p_r) \cdots (M - (i-1)p_r)}{i!\, p_r^i}} & \text{if } k \in \{1, \ldots, M\}, \end{cases}
\tag{26}
$$

which can be seen as another generalization of the binomial distribution $\text{Bin}(\frac{1}{2}, M)$ via the new parameter $p_r \in [0, 1]$. Again, the network snapshot entropy can be computed as

$$
H_{p_r}(G_i) = -\sum_k p_{p_r}(k) \log \frac{p_{p_r}(k)}{\binom{M}{k}}.
\tag{27}
$$

Both models (23) and (25) provide respectively stationary distributions (24) and (26), which, in general, are not binomial. Therefore, if we take a snapshot of these stationary distributions, the resulting network will follow a new static model, different from the standard known reference models for static networks.

Note that again these models can be interpreted as urn-derived finite state discrete-time birth-death models, in the sense that they do not incorporate network structural information, but only the total number of links. In other words, these models do not differentiate among networks that belong to the same equivalence class C_k, i.e., they are structure independent.

The time evolution of the expected value for the number of links, the clustering coefficient, the connectivity and the sample degree distribution entropy define a vector time series that can be employed as a signature that characterizes the evolution models. The signature of the above considered structure independent models can be employed as a reference baseline to assess more complex behaviors.

In Section 6, these signature quantities are estimated via simulation procedures.

5. Evolution Models Depending on Network Structure: Evolution of Clustering Coefficient

Usually, the evolution of networks depends not only on the number of links but also on the network structure. To illustrate this idea, we will analyze the behavior of models whose dynamics depend on the fact that triangles are going to be created or deleted in the network; then, the evolution of the clustering coefficient will be an essential feature to be considered.

Let us consider, for instance, an extension of the asymmetric model of Section 4.3 where the probability of a given link to be added (or removed) may depend on the fact that a triangle will or will not be generated (or eliminated) when adding (or removing) such link. Precisely, the transition behavior from g_t to g_{t+1} is defined as follows:

1. a pair of nodes in network g_t is uniformly selected.
2. (a) If the selected pair already has an associated link, such link is removed

   i. with probability $p_{r,n}$ when the selected nodes have at least one common neighbor (hence, at least one triangle will be deleted), or
   ii. with probability $p_{r,nn}$ when the selected nodes do not have a common neighbor (no triangle will be deleted);

   (b) if the selected pair does not have an associated link, a link is added between these nodes

   i. with probability $p_{a,n}$ when the selected nodes have at least one common neighbor (hence, some new triangles will be generated), or
   ii. with probability $p_{a,nn}$ when the selected nodes do not have common neighbor (no triangle will be generated).

3. If no change (removal or addition) happens, the process is repeated until the network undergoes some modification, which is registered in g_{t+1}.

The discrete process g_t provided by this model remains invariant to a common scaling of all probability values, provided the proportion among them is preserved. Hence, such dynamical model behavior can be reformulated as a function of, for instance, the following three parameters, $\alpha = \frac{p_{a,n}}{p_{a,nn}}$, $\beta = \frac{p_{r,nn}}{p_{r,n}}$ and $u_1 = \frac{p_{r,n}}{p_{a,n}}$; note that α and β measure the strength for favoring the creation and preservation of triangles, respectively, and the meaning of u_1 will become clear below. Note that this model is general enough to represent dynamic networks having a tendency to either create (and preserve) or eliminate triangles. In the specific case when α and β are larger than 1, triangle creation (and preservation) are promoted.

The analysis of this model can be complex since the existence and characterization of a stationary behavior may depend on the network size (number of nodes), the selected parameter values and the initial state g_0. Note that M_t cannot be directly defined anymore via a transition matrix of the type of (22), (23) or (25), since the future evolution of such M_t depends not only on its actual value but also on some structural properties (i.e., the existence and location of triangles) of g_t. In addition, the existence and form of a limiting stationary distribution for M_t may be a complicated issue to deal with.

Two Regimes of Behavior

If we assume that g_t follows approximately a Gilbert model, the probability $p_{ij,t}$ of any pair of nodes (v_i, v_j) to have a common neighbor (i.e., they may take part of a triangle) depends on the ratio between the number of links m_t and the total number of node pairs $M = \binom{N}{2}$ in g_t:

$$p_{ij,t} = 1 - \left(1 - \left(\frac{m_t}{M}\right)^2\right)^{N-2}. \tag{28}$$

The value of $p_{ij,t}$ is very sensitive to the link density $d = \frac{m_t}{M}$. For large values of N, we have that, if $m_t \sim o(N^{\frac{3}{2}})$, then $p_{ij,t}$ remains small, approaching value 1 otherwise. Therefore, the behavior of the model presented in Section 5 may be approximated considering two possible regimes.

1. Regime 1. For large enough $m_t \sim \Omega(N^{\frac{3}{2}})$, the probability of creating or deleting triangles is not negligible and the dynamics of the system are approximately governed by a model following (23) with $u_1 = \frac{p_{r,n}}{p_{a,n}}$.

2. Regime 2. For small $m_t \sim o(N^{\frac{3}{2}})$, the probability of creating or deleting triangles is small and the dynamics of the system are approximately governed by a model following (23) with $u_2 = \frac{p_{r,nn}}{p_{a,nn}}$.

Hence, within each regime, the dynamics can be approximated via the baseline model (23).

We now focus on the analysis for the cases where triangle creation and preservation is favored, meaning that both α and β would be larger than 1. Then, $u_2 = \alpha \cdot \beta \cdot u_1 > u_1$ and several behaviors can be found depending on the concrete selected values for u_1, u_2 and α (or β). Precisely, some regime may be transitory, or both may coexist as stationary behaviors depending on the selected initial conditions. In general, regime 1 will be more common since it fits with a wide range of possible values for u_1; only if u_1 is very large (note that it would imply a huge u_2) may we start with a g_0 satisfying regime 1 condition, but the expected value of stationary distribution M_s for such u_1 may correspond to regime 2, so that the system may end up in such second regime. On the other hand, since regime 2 corresponds to a narrow range of small values of M_t, a very large value of u_2 will be required for such regime to show up as stationary; if u_2 is not large enough, even if we start with a g_0 satisfying regime 2 condition, the expected value of stationary distribution M_s for such u_2 may lie in the range of values corresponding to regime 1, so that the system may end up in such first regime. Finally, both regimes may coexist with u_2 large enough and u_1 small enough so that the respective expected values of stationary distributions M_s correspond to each one of the regimes. Note that, if α and β are large enough to favor triangle creation

and preservation, u_1 and u_2 may differ in some orders of magnitude allowing a natural coexistence of both regimes.

In the next section, different simulations are performed to characterize the time evolution of some basic features (e.g., the expected value for the number of links, the clustering coefficient, the connectivity and the sample degree distribution) for the dynamics models presented above.

6. Simulations for the Time Evolution of Features

Numerical simulations have been performed to characterize the time evolution of the number of links, the clustering coefficient and the entropy of the sample degree distribution for the extended model defined by (23) and the structure dependent model presented in Section 5.

6.1. Extended Asymmetric Model

We begin by characterizing the extended model defined by (23). Figure 2 shows the evolution (starting from the empty graph) of the relative number of edges (i.e., edge density $d = \frac{m_t}{M}$), the clustering coefficient and the samples degree distribution entropy of a graph that evolves following the extended model defined by (23) with $p_a = 0.3$ and $p_r = 1$. The estimations of relative number of edges and clustering coefficient converge to the same stationary value as the iteration number increases; hence, their ratio converges to one, this value being a reference baseline signature of structure independent models. Note that the variance of the clustering coefficient is significantly larger than the variance corresponding the relative number of edges. The estimated degree distribution presents also a significant variance.
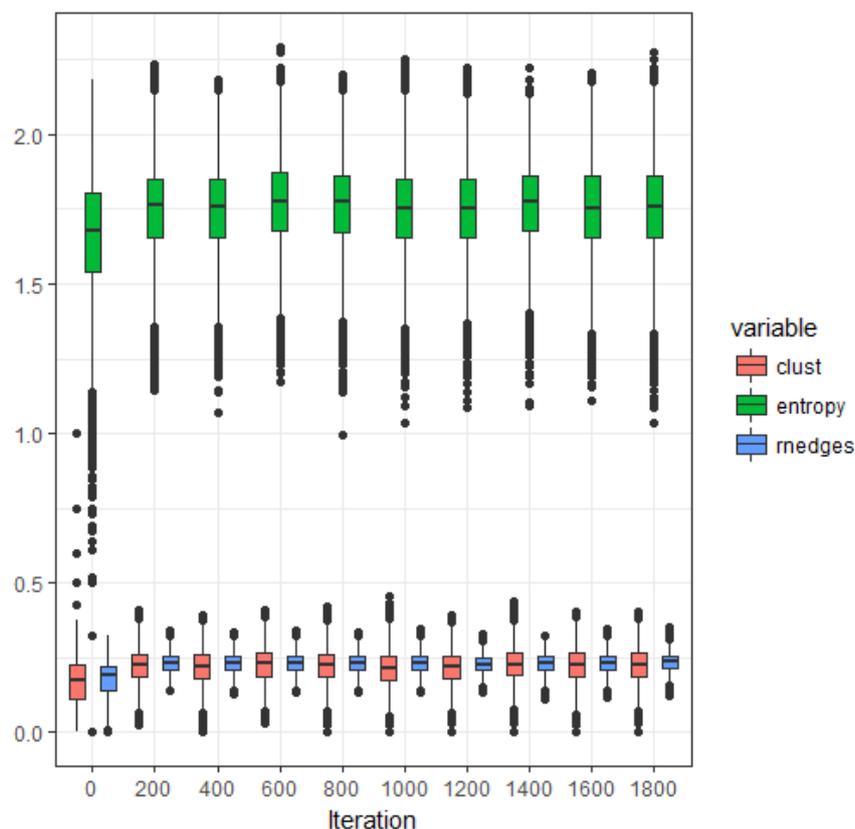

Figure 2. Extended asymmetric model. Estimated expected values of relative number of edges (rnedges), clustering coefficient (clust) and sample degree distribution entropy (entropy), as a function of the iteration number ($p_a = 0.3$ and $p_r = 1$).

Figure 3 represents the estimated expected value of the number of edges as a function of iteration number (starting from the empty graph) and parameter u. Due to the uniform nature of $P(G_i \mid M_i)$, the behavior of the clustering coefficient follows a similar behavior; again, this clustering coefficient/link density ratio value close to one is a reference baseline signature for these types of structure independent models.
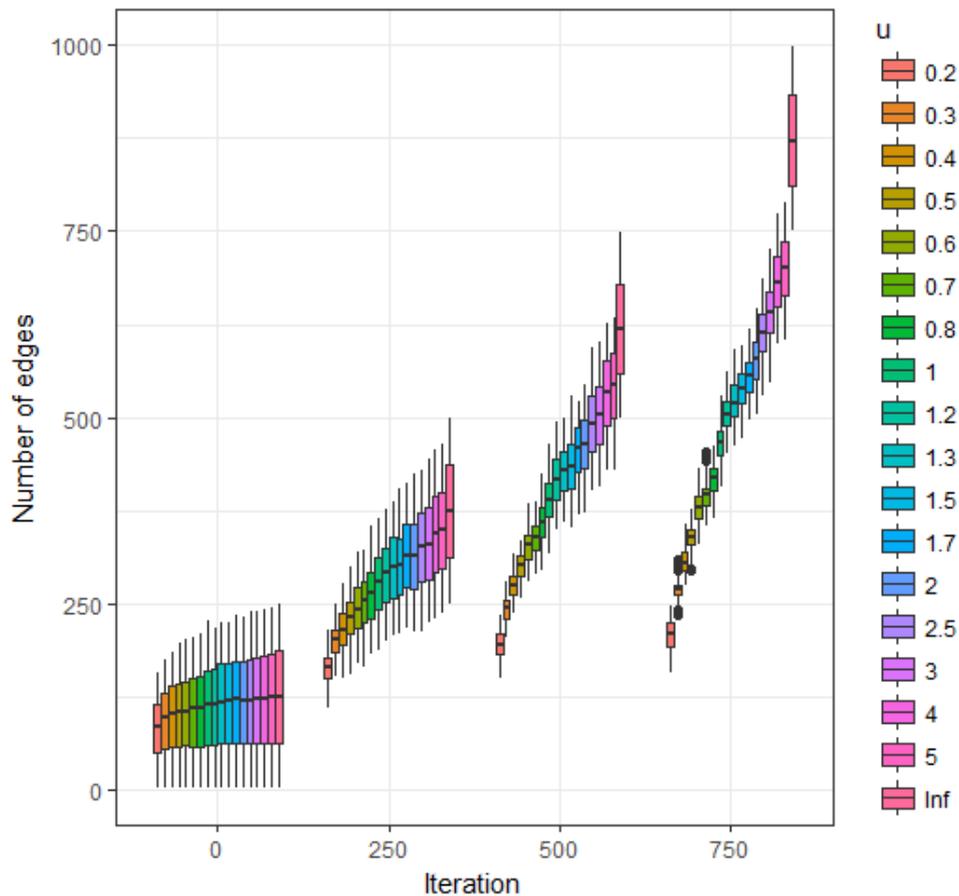

Figure 3. Extended asymmetric model. Estimated expected value of number of edges as a function of u at iterations 250, 500, 750 and 1000.

Figure 4 represents the estimated expected value of the sample degree distribution entropy as a function of iteration number (starting from the empty graph) and parameter u. Larger values are obtained for $u = 1$ as also illustrated in Figure 1.
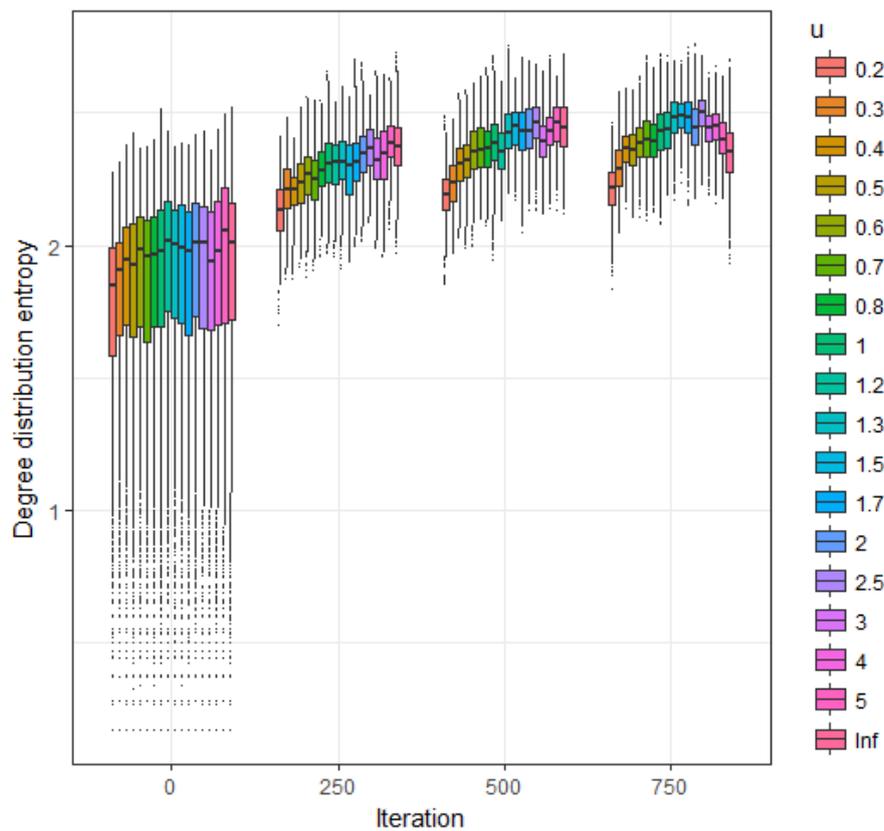
Figure 4. Extended asymmetric model. Estimated expected value of the sample degree distribution entropy as a function of u at iterations 250, 500, 750 and 1000.

6.2. Structure Depending Model

We now illustrate the behavior of some models of the type presented in Section 5. A network with $n = 100$ has been considered, where different parameter values and initial conditions have been tested. In the following, we indicate some cases that illustrate the different behaviors:

1. If $p_{an} = 0.5$, $p_{ann} = 0.05$, $p_{rn} = 0.1$ and $p_{rnn} = 1$ (or, equivalently, $u_1 = 20$, $u_2 = 0.2$, $\alpha = \beta = 10$) the system always converges to a stationary behavior in regime 1, with link density $d = \frac{m_t}{M} \approx 0.8$, and clustering coefficient $cc \approx 0.8$ (hence the ratio $\frac{cc}{d}$ approaches 1 as it happens in Erdős–Renyi [22] or Gilbert models).
2. If $p_{an} = 0.01$, $p_{ann} = 0.005$, $p_{rn} = 0.5$ and $p_{rnn} = 1$ (or, equivalently, $u_1 = 200$, $u_2 = 50$, $\alpha = \beta = 2$), the system always converges to a stationary behavior in regime 2, with link density $d = \frac{m_t}{M} \approx 0.005$, and oscillating clustering coefficient between 0 and 0.2 (hence the ratio $\frac{cc}{d}$ presents also large oscillations).
3. If $p_{an} = 0.05$, $p_{ann} = 0.005$, $p_{rn} = 0.1$ and $p_{rnn} = 1$ (or, equivalently, $u_1 = 200$, $u_2 = 2$, $\alpha = \beta = 10$), the system presents the two above illustrated regimes; and, depending on the initial condition, there is more or less probability to evolve within each one of these regimes:

    (a)　For g_0 following a Gilbert model with $p \gtrsim 0.0375$, the system is more likely to remain in regime 1 with link density $d = \frac{m_t}{M} \approx 0.32$, and clustering coefficient $cc \approx 0.32$ (hence the ratio $\frac{cc}{d}$ approaches 1 as it happens in ER or Gilbert models).
    (b)　For g_0 following a Gilbert model with $p \lesssim 0.0375$, the system is more likely to steadily remain (at least for a long time beyond the number of performed iterations) in regime 2 with link density $d = \frac{m_t}{M} \approx 0.008$, and clustering coefficient $cc \approx 0.3$. Hence, the ratio $\frac{cc}{d}$ evolves around 40; this large clustering coefficient in proportion to link density illustrates a very common feature of many real networks.

The two regime case is illustrated in Figures 5 and 6, where the estimated expected value of the density of edges and the clustering coefficient as a function of the iteration number are presented for different initial condition graphs g_0. Figure 5 illustrates the behavior when g_0 is either the complete or the empty graph. All simulations starting from the complete graph led to regime 1, whereas all simulations starting from the empty graph led to regime 2.
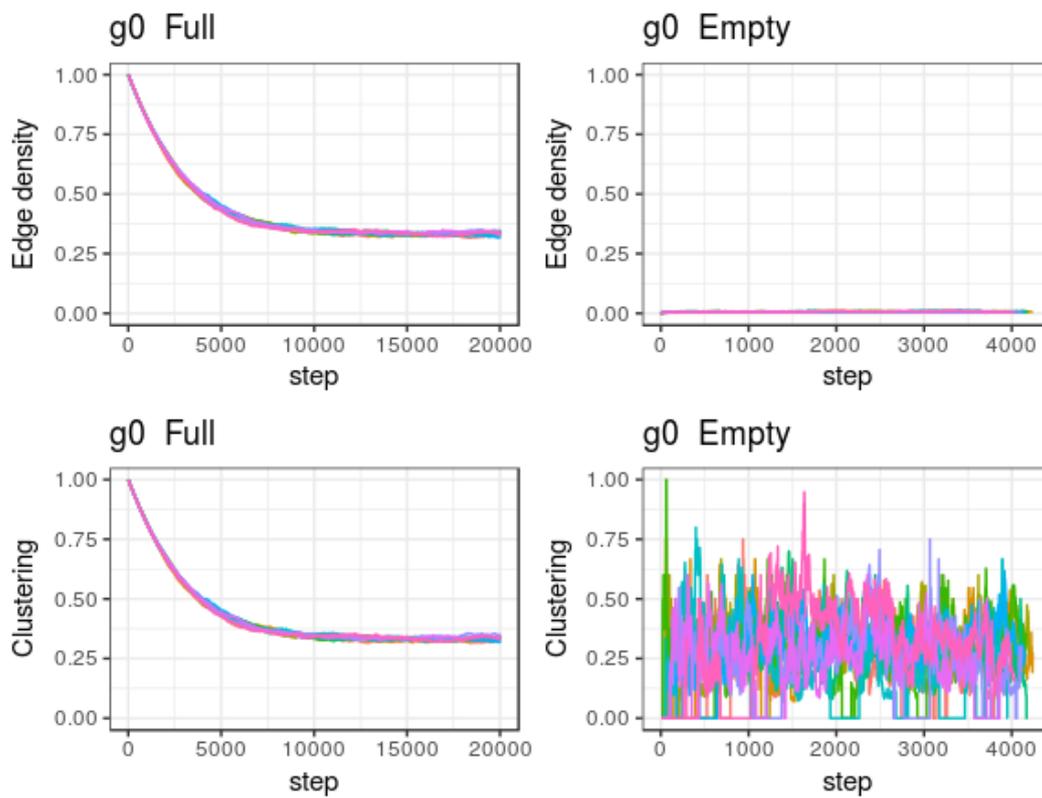

Figure 5. Structure depending model. Estimated expected value of number of edges and clustering coefficient as a function of number of steps, starting from complete and empty graphs, respectively. Different line colors correspond to different instantiations of the model.
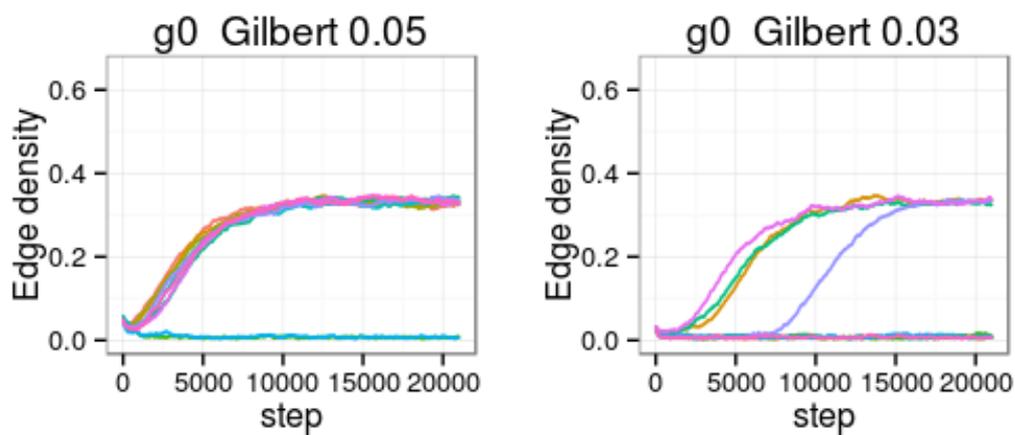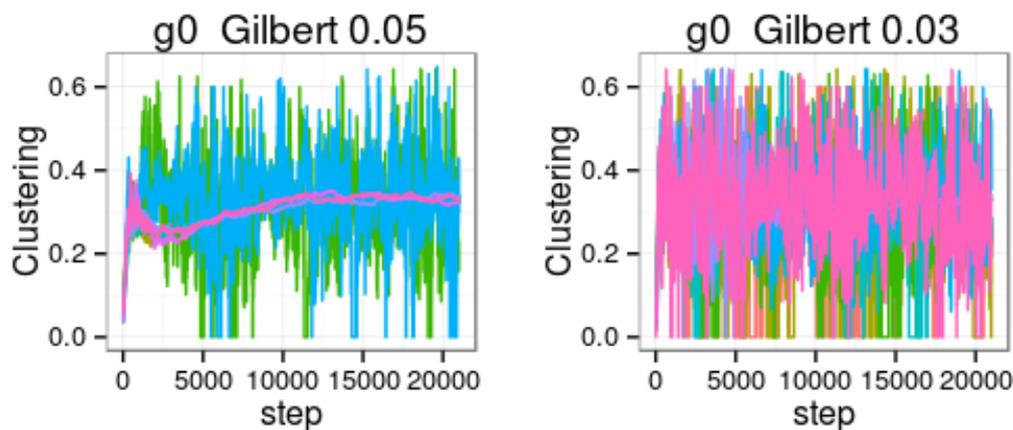

Figure 6. *Cont.*

Figure 6. Structure depending model. Estimated expected value of number of edges and clustering coefficient as a function of the number of steps, starting from a sample of the Gilbert model with $p = 0.05$ and $p = 0.03$, respectively. Different line colors correspond to different instantiations of the model.

Figure 6 illustrates the behavior for two cases when g_0 is obtained as a sample of the Gilbert model with $p = 0.05$ and $p = 0.03$, respectively. A majority of the simulations starting from a Gilbert $p = 0.05$ graph led to regime 1, whereas a majority of the simulations starting from a Gilbert $p = 0.03$ graph led to regime 2. These simulations illustrate approximately the size of the stochastic domains corresponding to each regime.

7. Conclusions

Several basic models for dynamic networks, based only on edge density, have been initially proposed and analyzed in terms of the time evolution of the number of links, clustering coefficient, connectivity and entropy of the sample degree distribution; the evolution of these features helps to characterize the proposed models and provides a reference baseline signature to assess more complex behaviors. The proposed model involving network structure presents a more sophisticated behavior and, for some regime, it leads to the generation of a high clustering coefficient/link density ratio when compared with the reference baseline values. This result is promising for the design of network models with tunable clustering coefficient with the aim to replicate some real networks characteristics.

The proposed framework will serve to assess, in a systematic manner, the properties of existing models as well as future more complex models for time evolving networks.

Author Contributions: P.J.Z. developed the conceptualization and formal analysis, wrote the document and helped with the simulations. I.B.-V. developed the simulations and helped with the theoretical content and the writing of the paper.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Watts, D.J.; Strogatz, S.H. Collective dynamics of "small-world" networks. *Nature* **1998**, *393*, 440–442. [CrossRef] [PubMed]
2. Albert, R.; Barabási, A.L. Statistical mechanics of complex networks. *Rev. Mod. Phys.* **2002**, *74*, 47. [CrossRef]
3. Newman, M. *Networks: An Introduction*; Oxford University Press: Oxford, UK, 2010.

4. Newman, M.; Barabasi, A.L.; Watts, D.J. *The Structure and Dynamics of Networks*; Princeton University Press: Princeton, NJ, USA, 2011; Volume 19.

5. Dorogovtsev, S.N.; Mendes, J.F. *Evolution of Networks: From Biological Nets to the Internet and WWW*; Oxford University Press: Oxford, UK, 2013.

6. Siljak, D.D. Dynamic Graphs. *Nonlinear Anal. Hybrid Syst.* **2008**, *2*, 544–567. [CrossRef]

7. Kejžar, N.; Nikoloski, Z.; Batagelj, V. Probabilistic inductive classes of graphs. *J. Math. Sociol.* **2008**, *32*, 85–109. [CrossRef]

8. Holme, P.; Saramäki, J. Temporal networks. *Phys. Rep.* **2012**, *519*, 97–125. [CrossRef]

9. Pastor-Satorras, R.; Vespignani, A. Epidemic spreading in scale-free networks. *Phys. Rev. Lett.* **2001**, *86*, 3200. [CrossRef] [PubMed]

10. Rahmani, A.; Ji, M.; Mesbahi, M.; Egerstedt, M. Controllability of multi-agent systems from a graph-theoretic perspective. *SIAM J. Control Optim.* **2009**, *48*, 162–186. [CrossRef]

11. Dorogovtsev, S.N.; Mendes, J.F. Evolution of networks. *Adv. Phys.* **2002**, *51*, 1079–1187. [CrossRef]

12. Zufiria, P.J.; Barriales-Valbuena, I. Evolution models for dynamic networks. In Proceedings of the 2015 38th International Conference on Telecommunications and Signal Processing (TSP), Prague, Czech Republic, 9–11 July 2015; pp. 252–256.

13. Herrera, C.; Zufiria, P.J. Generating scale-free networks with adjustable clustering coefficient via random walks. In Proceedings of the 2011 IEEE Network Science Workshop (NSW), West Point, NY, USA, 22–24 June 2011; pp. 167–172.

14. Zufiria, P.J.; Barriales-Valbuena, I. Entropy Characterization of Random Network Models. *Entropy* **2017**, *19*, 321. [CrossRef]

15. Klein, M.J. Entropy and the Ehrenfest urn model. *Physica* **1956**, *22*, 569–575. [CrossRef]

16. Van Doorn, E.A.; Schrijner, P. Geometric ergodicity and quasi-stationarity in discrete-time birth-death processes. *ANZIAM J.* **1995**, *37*, 121–144.

17. Feller, W. *An Introduction to Probability Theory and Its Applications*; Wiley: New York, NY, USA, 1968; Volume 1.

18. Johnson, N.L.; Kemp, A.W.; Kotz, S. *Univariate Discrete Distributions*; John Wiley & Sons: New York, NY, USA, 2005; Volume 444.

19. Kemp, A.W. Steady-state Markov chain models for certain q-confluent hypergeometric distributions. *J. Stat. Plan. Inference* **2005**, *135*, 107–120. [CrossRef]

20. Gilbert, E.N. Random graphs. *Ann. Math. Stat.* **1959**, *30*, 1141–1144. [CrossRef]

21. Morimoto, T. Markov processes and the H-theorem. *J. Phys. Soc. Jpn.* **1963**, *18*, 328–331. [CrossRef]

22. Erdős, P.; Rényi, A. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.* **1960**, *5*, 17–60.