Article

Network Rewiring in the *r*-*K* Plane

Maria Letizia Bertotti and Giovanni Modanese *

Faculty of Science and Technology, Free University of Bozen-Bolzano, 39100 Bolzano, Italy;
marialetizia.bertotti@unibz.it
* Correspondence: giovanni.modanese@unibz.it

check for
updates

Abstract: We generate correlated scale-free networks in the configuration model through a new rewiring algorithm that allows one to tune the Newman assortativity coefficient r and the average degree of the nearest neighbors K (in the range $-1 \leq r \leq 1$, $K \geq \langle k \rangle$). At each attempted rewiring step, local variations Δr and ΔK are computed and then the step is accepted according to a standard Metropolis probability $\exp(\pm \Delta r / T)$, where T is a variable temperature. We prove a general relation between Δr and ΔK, thus finding a connection between two variables that have very different definitions and topological meaning. We describe rewiring trajectories in the r-K plane and explore the limits of maximally assortative and disassortative networks, including the case of small minimum degree ($k_{min} \geq 1$), which has previously not been considered. The size of the giant component and the entropy of the network are monitored in the rewiring. The average number of second neighbors in the branching approximation $\bar{z}_{2,B}$ is proven to be constant in the rewiring, and independent from the correlations for Markovian networks. As a function of the degree, however, the number of second neighbors gives useful information on the network connectivity and is also monitored.

Keywords: scale-free networks; network rewiring; network assortativity; network entropy

1. Introduction

Rewiring algorithms are often employed in network science to build "synthetic networks" for the mathematical modeling of dynamics or diffusion processes [1–5]. Usually, rewiring algorithms preserve the degree distribution of the network while changing the degree correlations and other topological features.

The "configuration model" [6] is a well established generalization of the random networks of Erdös-Renyi which yields uncorrelated networks having a pre-assigned (typically scale-free) degree distribution. It is known, however, that assortative and disassortative correlations play an important role in dynamics and diffusion on networks [7–11]. For this reason some algorithms have been devised, which are able to perform a degree-conserving rewiring while modifying the pair correlations in the direction of increasing assortativity or disassortativity.

It is also possible to rewire the network in order to change its clustering coefficient (see [12] and references) or other metrics, but in this work the focus is on assortativity and disassortativity as measured by the Newman coefficient r, and on the average nearest neighbors degree function $k_{nn}(k)$, or better on its network average $K \equiv \langle k_{nn}(k) \rangle = \sum_k P(k) k_{nn}(k)$. Here, $P(k)$ is the degree distribution, $P(h|k)$ denotes the conditional probability for a node of degree k to be connected to a node of degree h, and $k_{nn}(k) = \sum_k h P(h|k)$.

The algorithm by Xulvi-Brunet and Sokolov [13] is quite efficient for generating networks that are maximally assortative or maximally disassortative, or even have an intermediate r coefficient, if a tunable return probability is inserted in the rewiring criterion. It does not allow, however, any direct control of the degree correlations e_{jk} or $P(h|k)$, the r coefficient or the k_{nn} function. (We recall that e_{jk} is

the probability for a node with excess degree j, i.e., degree $(j + 1)$, to be connected with a node with excess degree k.).

The rewiring method proposed by Newman [1] allows in principle to generate ensembles of networks displaying, on average, any "target" two-point correlations assigned through an e_{jk} matrix compatible with the given degree distribution. There exist several recipes for the construction of such matrices in the case of scale-free networks [1,14,15].

We have recently proposed a new algorithm [16] that is equivalent to the algorithm of [13] when applied to maximally assortative or disassortative networks, but allows at each step the control of the variation Δr in the Newman coefficient, and therefore permits the introduction of a rewiring "temperature" T in order to tune the return probability via a standard Metropolis update.

It is known that in many classes of networks, no simple method for the creation of random networks of the class exists [3]. In these cases a Monte Carlo method is usually invoked. The desired network class is determined by an energy function, which is defined on a network configuration. The lower the function is, the more desirable are the properties of the network. For a good performance of the algorithm, the energy function should be calculable locally. That is, the change in energy between configurations should depend only on the changed links and possibly some near neighbors, so calculating the new energy should not require recalculating the energy for the entire network. Our proposed method improves previous methods since it satisfies these properties by using Δr as energy variation.

In order to increase the convergence probability of Metropolis algorithms it is sometimes attempted to start with a high temperature, making all moves probable and covering the entire configuration space, and gradually decrease the temperature in order to approach the real energy minimum (procedure of "simulated annealing"). This is also possible in principle with our method, but it should be stressed that in general rewiring algorithms do not sample uniformly through spaces of networks (see for instance [17–19] and references), and that the more features one wants to preserve in the rewiring, the less successful sampling is. Therefore our rewiring algorithm, like the other cited algorithms for changing network correlations, cannot be used as a randomization procedure. Another reason for this is that these algorithms may generate multiple copies of isomorphic random networks, and so cause over-counting in the sampling.

One of the aims of this work is to clarify the relations existing among these rewiring methods and the asymptotic constraints on maximally assortative and disassortative networks found by Menche et al. [20]. Using our "Δr-rewiring" we have also been able to see the effects of extreme assortativity and disassortativity in networks with many nodes of small degree. These were not considered by the authors of [20], who took as a minimum degree a typical value $k_0 = 6$ and therefore found in general highly connected networks with a very large giant component.

The study of complex networks is often motivated by the interest for the dynamics of some diffusion problem on top of them. Clearly, if a mean-field approximation of the dynamics or diffusion on the network is sufficient for one's purposes, then the corresponding equations can be written, analyzed, or solved in terms of the excess-degree correlations e_{jk} or in terms of the degree distribution $P(k)$ plus the conditional probabilities $P(h|k)$ (the so-called probabilistic or "Markovian" description of the network). If, on the other hand, a full realization of the network is needed (e.g., for simulations or stochastic modeling, or because one wants to take into account the effect of correlations beyond the second order), then several issues arise, for example:

- Is it possible to build any desired assortative or disassortative network, defined at the probabilistic level through a suitable "theoretical" e_{jk} matrix, using a Newman rewiring, at least at the ensemble level? What is in this respect the role of the asymptotic constraints on r? Can the asymptotic constraints tell us in advance that a certain theoretical e_{jk} is impossible to be implemented in a real network?

- Among the networks obtained through a Xulvi–Brunet–Sokolov rewiring or our Δr rewiring, will one find the desired assortative or disassortative network? If yes, with what accuracy is this possible, compared to the Newman rewiring?
- How do the results (and their level of fluctuations and uncertainty) change if we modify the degree distribution, and especially the probability of the nodes with the lowest degree? Will a giant component always be present? If the network is much fragmented, what are the consequences for diffusion processes?

As this work was progressing, we gradually realized that such issues are hard to solve in general terms. We found partial answers, but we chose to leave most of them for a forthcoming publication. Still the efforts described in this work have already led to some useful spin-offs. In an attempt to obtain a better characterization of the rewiring process, we have represented the state of the network in an *r-K* plane trying to use *K* as a "coordinate" independent from *r*. This brought us first to a revision of the meaning and of the properties of *K* (Sections 2, 2.1, and 2.2) and then to establish a new general relation between the variations of *r* and *K* in a rewiring step (Sections 2.3 and 2.4). This relation has been proven theoretically and verified numerically through the rewiring code. The code has also allowed to guess some related properties of the quantity $\bar{z}_{2,B}$ (average number of second neighbors in the branching approximation), which have been proven theoretically in Section 3, although they are not strictly necessary for the rest of the present work.

In Section 4, we describe the main features of the rewiring code and the procedure for the calculation of the entropy of the generated networks. In Section 5.1 we describe the "rewiring trajectories" in the *r-K* plane obtained for some different values of the scale-free exponent γ ($\gamma = 2.25$, 2.5, 2.75, 3) starting from uncorrelated networks and performing an assortative or disassortative rewiring at a low temperature, i.e., with small return probability. For some of the cases a qualitative description is given of the "super-assortative" and "super-disassortative" asymptotic networks generated. Section 5.2 describes the preliminary results of the assortative rewiring in equilibrium at variable temperature *T*, with a plot showing the case $\gamma = 2.5$, as a function of *T*, the values of the entropy *S*, of $\langle r \rangle$, $\langle K \rangle$, and the size of the giant component. Section 6 contains our conclusions.

2. The Function "Average Degree of the First Neighbors" K_N

The quantity $\langle k_{nn}(k) \rangle_N$, introduced by Boguñá et al. [21], is defined as:

$$\langle k_{nn}(k) \rangle_N = K_N = \sum_{k=1}^{n} P(k) k_{nn}(k) \,. \tag{1}$$

We shall denote it for simplicity *K*, or K_N, since for a given type of network it depends only on its size, namely on the number *N* of nodes (related in turn to the maximum degree *n*, for scale-free networks, through the Dorogovtsev–Mendes criterion as detailed in Equation (3)). Since $k_{nn}(k)$ amounts to the average degree of the first neighbors of a node of degree *k* and $P(k)$ is the probability that such a node is present, K_N is the average degree of the first neighbors taken over the entire network, or better the average degree of the first neighbors of a randomly chosen node. Generally speaking, K_N is strongly related to the diffusion properties of the network.

The definition above is probabilistic, and used for Markovian networks and for applications to mean-field equations on these networks. If we have complete knowledge of the network, we can compute *K* exactly by just looking at the first neighbors of each node and computing a total average of their degrees (see Section 4).

The authors of [21] prove that as a function of *N*, K_N is diverging when $N \to \infty$ in a scale-free network with exponent $2 < \gamma \le 3$, for any kind of correlations (at least when the function $k_{nn}(k)$ has a certain form). This property is employed to conclude that in the "thermodinamic" limit of large *N*, phenomena of epidemic diffusion always propagate to the entire network, no matter how small the contagion probability is ("absence of epidemic threshold"). The intuitive reason is that although the

average number of neighbors $\langle k \rangle$ tends to a constant for large N, the average degree of these neighbors tends to infinity. This means that each node is very close to a hub from which the epidemics can easily spread.

In the following two sub-sections we give some examples of computation of the function K_N in Markovian networks, as an introduction to the results obtained through the rewiring of real networks.

2.1. Uncorrelated Networks

For an uncorrelated network we obtain for K_N a simple expression. We have in this case:

$$P(h|k) = \frac{hP(h)}{\langle k \rangle} \, .$$

Therefore $k_{nn}(k)$ does not depend on k:

$$k_{nn}(k) = \frac{1}{\langle k \rangle} \sum_{h=1}^{n} h^2 P(h) = \frac{\langle k^2 \rangle}{\langle k \rangle} \, ,$$

and

$$K_N = \sum_{k=1}^{n} k_{nn}(k) P(k) = \frac{\langle k^2 \rangle}{\langle k \rangle} \geq \langle k \rangle \, , \tag{2}$$

where we have used the normalization condition $\sum_{k=1}^{n} P(k) = 1$ and the last inequality is due to the fact that in general $\langle k^2 \rangle \geq \langle k \rangle^2$.

The inequality $K_N \geq \langle k \rangle$ expresses the well-known property that in an uncorrelated network, from the point of view of one node looking at its first neighbors, on the average "my friends have more friends than me" (because their average degree is K_N and my degree is k). As we shall see, this property is also numerically confirmed for correlated networks, at least in the scale-free case.

The dependence from N in the expression of Equation (2) for K_N arises as follows. First note that when we consider a finite network with maximal degree n, the normalization condition of $P(k)$ is $\sum_{k=1}^{n} P(k) = 1$. Therefore the properly normalized degree distribution $P(k)$ for a scale-free network is:

$$P(k) = c_{n,\gamma} k^{-\gamma}, \quad c_{n,\gamma}^{-1} = \sum_{k=1}^{n} k^{-\gamma} \, .$$

The quantities $\langle k \rangle$ and $\langle k^2 \rangle$, respectively equal to $\sum_{k=1}^{n} kP(k)$ and $\sum_{k=1}^{n} k^2 P(k)$, depend on n through the factor $c_{n,\gamma}$ and the upper limit of the sum. However, in the limit of large n the factor $c_{n,\gamma}$ tends to a constant and $\sum_{k=1}^{n} kP(k)$ is convergent for $2 < \gamma \leq 3$; the dependence of K_N on n comes from the divergent series $\sum_{k=1}^{n} k^2 P(k)$.

Approximating with an integral the dependence of the series on its upper limit, we obtain for large n:

$$\langle k^2 \rangle \sim \int_1^n k^{2-\gamma} dk \sim n^{3-\gamma} \quad \text{for } 2 < \gamma < 3 \, ,$$

and

$$\langle k^2 \rangle \sim \ln(n) \quad \text{for } \gamma = 3 \, .$$

Of course, for k close to 1 the integral is not a good approximation of the series; furthermore, if the sum starts from a value $k_{min} > 1$, the factor $c_{n,\gamma}$ has a substantial dependence on k_{min} (see Table 1), as we shall later see in some examples. Here, however, we are interested into the divergent dependence of K on n.

In order to relate the maximum degree n to the number of nodes N we make recourse to the integral criterium of Dorogovtsev–Mendes [22], which states that the probability to have in the network a node with degree in the range $(n, +\infty)$ must be equal to 1, implying:

$$\int_n^\infty c_{n,\gamma}k^{-\gamma}dk = \frac{1}{N}.$$

From this we obtain the known relation:

$$\frac{\gamma - 1}{c_{n,\gamma}}n^{\gamma-1} = N. \tag{3}$$

An example of exact values of the various quantities involved is given in Table 1.

Table 1. An example of values of K for Markovian networks with 1000 nodes in dependence on the scale-free exponent γ and the minimum degree k_{min}, in the uncorrelated case, assortative case (Vazquez–Weigt recipe, Equation (4), with $r = 0.5$) and disassortive case (Porto–Weber recipe as employed in [23]).

γ	k_{min}	N	n	$c_{n,\gamma}$	$\langle k \rangle$	K (unc.)	K (ass.)	K (dis.)	r_{dis}
2.5	1	1000	93	1.34	1.79	7.43	4.61	8.52	−0.088
2.5	4	1000	15	0.0896	6.23	7.39	6.81	20.1	−0.140
2.75	1	1000	43	1.26	1.50	3.63	2.56	5.15	−0.062
2.75	4	1000	6	0.0413	4.64	4.77	4.70	16.0	−0.120

2.2. Correlated Networks

In order to obtain K_N for a correlated network we must compute numerically the sum over k starting from an explicit expression for $k_{nn}(k)$, if known, or else expressing also $k_{nn}(k)$ as a sum, according to its definition $k_{nn}(k) = \sum_{h=1}^n hP(h|k)$.

A simple formula which defines assortative correlations has been proposed by Vazquez and Weigt [14] and has been employed in [24] for diffusion studies. It is a linear combination of an uncorrelated term and a totally assortative term proportional to δ_{hk}, namely:

$$P(h|k) = (1-r)\frac{hP(h)}{\langle k \rangle} + r\delta_{hk}, \tag{4}$$

where r ranges from 0 to 1 and coincides with the Newman assortativity coefficient.

From this matrix one obtains:

$$k_{nn}(k) = (1-r)\frac{\langle k^2 \rangle}{\langle k \rangle} + rk,$$

whence

$$K_N = (1-r)\frac{\langle k^2 \rangle}{\langle k \rangle} + r\langle k \rangle.$$

It follows that for fixed n (which also fixes $\langle k \rangle$ and $\langle k^2 \rangle$), when $r \to 1$ one has $K_N \to \langle k \rangle$, corresponding to the fact that in the case of extreme assortativity each node is only connected with other nodes having the same degree. We shall show in forthcoming work, however, that for a real scale-free network this limit is purely hypothetical, because the function $k_{nn}(k)$ cannot increase linearly for large k, but eventually must decrease.

Further evaluations of K_N as a function of N for assortative networks, built using a different set of $P(h|k)$ matrices [15,25], and for disassortative networks will be given elsewhere.

In any case, for fixed N the value of K is quite useful in characterizing the network and depends strongly on the type of correlations, on the scale-free exponent and on the minimum degree. A first example is given in Table 1 for Markovian networks. Then in Sections 5.1 and 5.2 we investigate the behavior of K for real rewired networks. An exact direct calculation of K from the list of links of a real

network can be efficiently implemented and compared with the average value of K obtained from the k_{nn} function.

2.3. Local Variation of K

The variation of K in a rewiring step is obtained using its definition as the average degree of the first neighbors of each node, averaged over the whole network. Let A, B, C, D denote respectively the degrees of the nodes a, b, c, d involved in the rewiring (see Figure 1). Let σ_a denote the sum of the degrees of the first neighbors of the node a, which are not involved in the rewiring, and similarly define $\sigma_b, \sigma_c, \sigma_d$. Before the rewiring the averages s_a, s_b, s_c, s_d of the degrees of the first neighbors of a, b, c, d are:

$$s_a = \frac{\sigma_a + B}{A}; \quad s_b = \frac{\sigma_b + A}{B}; \quad s_c = \frac{\sigma_c + D}{C}; \quad s_d = \frac{\sigma_d + C}{D}.$$

After the rewiring, these quantities become:

$$s_a = \frac{\sigma_a + C}{A}; \quad s_b = \frac{\sigma_b + D}{B}; \quad s_c = \frac{\sigma_c + A}{C}; \quad s_d = \frac{\sigma_d + B}{D}.$$

Therefore the change in the total average K is:

$$\Delta K = \frac{1}{N}\left[\left(\frac{C}{A} + \frac{D}{B} + \frac{A}{C} + \frac{B}{D}\right) - \left(\frac{B}{A} + \frac{A}{B} + \frac{D}{C} + \frac{C}{D}\right)\right]. \tag{5}$$
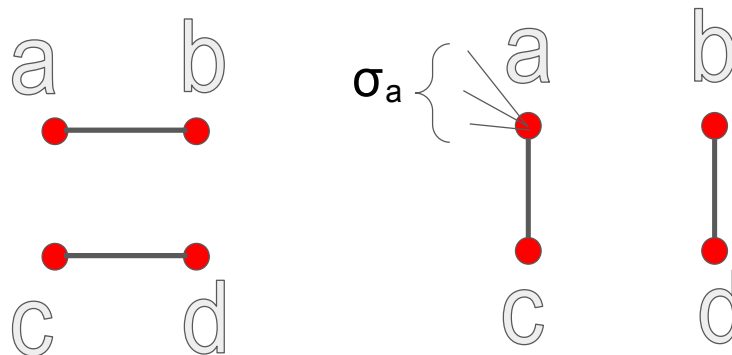

Figure 1. Scheme of a rewiring which involves the nodes a, b, c, d and leaves their degrees A, B, C, D unchanged. Before the rewiring the links are $(a, b), (c, d)$. After the rewiring the links are $(a, c), (b, d)$. Each node has in general other neighbors, not depicted; the sum of the degrees of the neighbors of a not involved in the rewiring is denoted in the text as σ_a, and similarly for b, c, d.

2.4. Relation between the Local Variations of K and r

In previous work [16] we found an expression for the local variation of the Newman coefficient in a rewiring of the same kind as in Figure 1. The variation is given by:

$$\Delta r = \frac{2(-AB - CD + AC + BD)}{L\sigma_q^2}, \tag{6}$$

where L is the number of links in the network and σ_q^2 is the denominator of the fraction which defines r, namely:

$$r = \frac{1}{\sigma_q^2} \sum_{j,k=0}^{n-1} jk(e_{jk} - q_j q_k), \tag{7}$$

$$\sigma_q^2 = \sum_{k=0}^{n-1} k^2 q_k - \left(\sum_{k=0}^{n-1} k q_k \right)^2 . \tag{8}$$

Here e_{jk} is the probability of finding in the network a link between nodes with excess-degrees j and k and $q_k = \sum_{j=0}^{n-1} e_{jk}$ is the excess-degree distribution. Note that q_k and σ_q^2 depend on the degree distribution but not on the correlations. Also note that in [16] a slightly different notation is used, in which A, B, C, D denote directly the excess-degrees. However, since:

$$-(A-1)(B-1) - (C-1)(D-1) + (A-1)(C-1) + (B-1)(D-1) = -AB - CD + AC + BD,$$

we can safely use Equation (6) according to the conventions of this paper, where A, B, C, D are the degrees and not the excess-degrees.

After some algebraic manipulations it is possible to express the variation ΔK in Equation (5) in terms of Δr, thus establishing a relation between two quantities which have very different definitions and topological meaning. We find:

$$\Delta K = -\frac{L \sigma_q^2}{2N} \frac{(AD + BC)}{ABCD} \Delta r . \tag{9}$$

This holds for each rewiring. The factor $-L\sigma_q^2/(2N)$ is fixed for a given degree distribution, while the factor $(AD + BC)/(ABCD)$ clearly depends on the nodes involved; we only know a priori that it is always positive, and as a consequence ΔK is always opposite to Δr and $\Delta K = 0$ if and only if $\Delta r = 0$. For the rewiring trajectories described in Section 5.1 the ratio $\Delta K/\Delta r$, averaged over many rewirings, turns out to be approximately constant for a given degree distribution; see data in Table 2.

Table 2. An example of the numerical factors involved in the relation of Equation (9) between ΔK and Δr, for scale-free networks with a fixed number of nodes. These factors explain why the rewiring trajectories in the *r-K* plane have different slopes. (See Figures 2 and 3 and the description in Section 5.1.) N is the number of nodes, L the number of links. σ_q^2 is the variance of the excess-degree distribution (Equation (8)). $\langle (AD + BC)/(ABCD) \rangle_{rew}$ denotes the average along an assortative rewiring trajectory with low temperature $T = 10^{-6}$. The hubs of the networks are defined by the cumulative probability method.

γ	N	L	σ_q^2	$\frac{L\sigma}{2N}$	$\left\langle \frac{AD+BC}{ABCD} \right\rangle_{rew}$
2.25	715	798	193	108	0.24
2.5	715	630	81.6	36	0.31
2.75	715	536	30.2	11.3	0.35
3	715	481	12.4	4.2	0.43

3. The Average Number of Second Neighbors $\bar{z}_{2,B}$

The condition for the existence of a giant component in an uncorrelated network with arbitrary degree distribution has been first found by Molloy and Reed [26] and is expressed by the inequality:

$$\langle k^2 \rangle - 2\langle k \rangle > 0 . \tag{10}$$

Later the same condition has been proven by Newman, Strogatz, and Watts with the method of the generating functions, which also allows one to find the size of the giant component [27]. In [27] the inequality in Equation (10) is reformulated in an intuitive way by stating that the giant component exists when $\bar{z}_2 > \langle k \rangle$, where $\bar{z}_2$ is the average number of second neighbors and $\langle k \rangle$ (the average degree) can also be interpreted as the average number of first neighbors. A crucial underlying assumption is that the network is locally a branching structure; moreover, being the network uncorrelated, one

supposes that there are no preferences in linking behavior depending on the node degrees and that therefore it makes sense to consider total averages like $\bar{z}_2$ and $\langle k \rangle$.

We are now going to define a quantity that is closely related to $\bar{z}_2$ and will show that starting from the intuitive "percolation" condition $\bar{z}_2 > \langle k \rangle$, Equation (10) can be immediately obtained without using the generating functions.

We call this quantity "average number of second neighbors in the branching approximation" and denote it by $\bar{z}_{2,B}$. It is a network average like $\bar{z}_2$, but includes by definition multiple counting in the case of shared second neighbors. More precisely, if one node a_0 has a second neighbor b in common with other h nodes $a_1, \ldots, a_h$, then b is counted $h + 1$ times in the average $\bar{z}_{2,B}$. The two quantities $\bar{z}_2$ and $\bar{z}_{2,B}$ coincide if the network is a pure branching structure, without nodes that have second neighbors in common with other nodes.

According to this definition, $\bar{z}_{2,B}$ can be obtained as:

$$\bar{z}_{2,B} = \sum_k P(k)k[k_{nn}(k) - 1], \tag{11}$$

because the probability for a node to have degree k is $P(k)$, the node has k first neighbors and their average degree is $k_{nn}(k)$.

Let us compute $\bar{z}_{2,B}$ for an uncorrelated network:

$$\bar{z}_{2,B} = \sum_k P(k)k[k_{nn}(k) - 1] = \sum_k P(k)k\left[\frac{\langle k^2 \rangle}{\langle k \rangle} - 1\right] = \tag{12}$$

$$= \left[\frac{\langle k^2 \rangle}{\langle k \rangle} - 1\right]\langle k \rangle = \langle k^2 \rangle - \langle k \rangle \quad \text{(uncorrelated network)}. \tag{13}$$

This expression gives the Molloy–Reed condition if we require $\bar{z}_2 > \langle k \rangle$ and admit that the network is a locally branching structure such that $\bar{z}_2 \simeq \bar{z}_{2,B}$.

Equation (11) for $\bar{z}_{2,B}$ is interesting in itself and we have used our code for the configuration model with rewiring in order to test it. The code generates a list, called the "Friends" list, of the first neighbors of each node. The list is updated and used in many parts of the program, for instance after the first wiring of the stubs, in order to check that the nodes' degrees match the prescribed degree distribution. It is also used at the end of the rewiring cycles, in order to find the giant component of the final network, and possibly for the numerical solution of diffusion equations in first or second closure approximation. It is straightforward to use the Friends list to also obtain the number of second neighbors of each node, because the degrees of the nodes do not change in the rewiring and are stored in a vector "Degrees[i]", with $i = 1, \ldots, N$, fixed from the degree distribution before the wiring. The contribution to $\bar{z}_{2,B}$ from each node is obtained as the sum of (*degree of each friend $-$ 1*). The total network average $\bar{z}_{2,B}$ is the sum of the contributions of all nodes, divided by N. One can check that the exact value obtained in this way is well approximated by the probabilistic value in Equation (11).

Somewhat unexpectedly, the exact obtained value of $\bar{z}_{2,B}$ is accurately reproduced in each simulation with the same degree distribution, signaling that it is not affected by the rewiring. In fact, the following two properties hold, which are not difficult to prove but cannot be found in the literature, to the best of our knowledge.

Property 1 (Property of $\bar{z}_{2,B}$). *For Markovian networks which satisfy the Network Closure Condition $hP(k|h)P(h) = kP(h|k)P(k)$, $\bar{z}_{2,B}$ does not depend on the correlations but only on the degree distribution, and it is equal to the value $\langle k^2 \rangle - \langle k \rangle$ obtained for an uncorrelated network having that degree distribution (Equation (13)).*

In fact the first term on the r.h.s. in Equation (11), namely $\sum_k P(k)kk_{nn}(k)$, is equal to $\langle k^2 \rangle$, as already noted in [21]:

$$\sum_k P(k)kk_{nn}(k) = \sum_k P(k)k \sum_h hP(h|k) =$$

$$= \sum_k \sum_h P(k)khP(h|k) = \sum_k \sum_h hhP(k|h)P(h) =$$

$$= \sum_h h^2 P(h) \sum_k P(k|h) = \langle k^2 \rangle$$

(because $\sum_k P(k|h) = 1$). The second term is equal to $\langle k \rangle$, which is fixed if the degree distribution is fixed.

Property 2 (Property of $\bar{z}_{2,B}$). $\bar{z}_{2,B}$ *does not change in a binary rewiring which preserves the degree distribution.*

This is a direct consequence of the definition of the rewiring (see Figure 1). Let σ_a denote, as before, the sum of the degrees of the first neighbors of the node a which are not involved in the rewiring, and similarly define σ_b, σ_c, σ_d. Let A, B, C, D denote the degrees of the nodes. The total number of second neighbors of the four nodes involved in the rewiring is equal, before the rewiring, to:

$$z_{2,B}^{before}(a,b,c,d) = (\sigma_a - A + B - 1) + (\sigma_b - B + A - 1) + (\sigma_C - C + D - 1) + (\sigma_D - D + C - 1).$$

After the rewiring we have:

$$z_{2,B}^{after}(a,b,c,d) = (\sigma_a - A + C - 1) + (\sigma_b - B + D - 1) + (\sigma_C - C + A - 1) + (\sigma_D - D + B - 1),$$

and the two quantities are equal.

These properties offer a strong indication for the absence of epidemic threshold in correlated large scale-free networks with $2 < \gamma \le 3$. In fact, in this range of γ, $\bar{z}_{2,B}$ diverges when $N \to \infty$. Denoting by λ the contagion probability for one single contact, consider an infected node randomly chosen, thus with the average degree $\langle k \rangle$. The probability that the node infects one of its neighbors is $\langle k \rangle \lambda$, which tends to zero if λ is very small. However, the probability that the node infects one of its second neighbors is (for a locally branching structure) equal to $\bar{z}_{2,B}\lambda^2$, and this quantity can stay finite even as $\lambda \to 0$, because of the divergence of $\bar{z}_{2,B}$, independently from the degree correlations.

4. The Rewiring Algorithm

In the first part of the algorithm, the "wiring" part, we generate a series of $2L$ "stubs" with given degree distribution, like in any implementation of the configuration model. The exact procedure for assigning the degrees to the hubs has two possible alternatives and has been described in [16]. In the first alternative ("cumulative hubs"), the probability of hubs for which $P(k) \cdot N < 1$ is cumulated with increasing k until it exceeds 1, at which point the hub is created and the accumulation starts again. In the second alternative ("random hubs"), hubs of degree k are created entirely at random with probability $P(k) \cdot N$.

For a description of the linking of the stubs the reader is referred to [16]. Details of this procedure are little relevant because the extensive rewiring that follows cancels any memory of the initial wiring scheme.

When we perform a rewiring step, we choose at random two links in the current list of links describing the network, say (a, b) and (c, d) (nodes are identified by a sequential number in the range $1, \ldots, N$, so a, b, c, d denote this number). With a probability of 50% we exchange a and b, to avoid any asymmetry, and then we build the new links (a, c), (b, d). In order to avoid the formation of loops, the rewiring is not performed if $a = c$ or $b = d$. The formation of multi-links (more than one link between the same two nodes) is avoided through a check of the adjacency matrix $[A_{ab}]$, which is

computed from the list of links before the rewiring cycle and updated after each rewiring according to formulas:

$$A_{ab} \rightarrow A_{ab} - 1; \quad A_{cd} \rightarrow A_{cd} - 1; \quad A_{ac} \rightarrow A_{ac} + 1; \quad A_{bd} \rightarrow A_{bd} + 1$$

(plus the symmetrical variations for A_{ba} etc.).

The knowledge of the adjacency matrix also allows one to compute the Shannon entropy of the network (see [28] and references) by ensemble-averaging A over sub-cycles. For instance, consider a typical rewiring cycle with 100 sub-cycles of 10^4 steps each, for a network with $N = 10^3$. The values of A_{ab}, with $a, b = 1, \ldots, N$ at the end of each sub-cycle are averaged to compute S according to the formula:

$$S = - \sum_{a,b=1}^{N} \bar{A}_{ab} \ln \bar{A}_{ab} + (1 - \bar{A}_{ab}) \ln(1 - \bar{A}_{ab}); \quad \bar{A}_{ab} = \langle A_{ab} \rangle_{sub-cycles}. \tag{14}$$

The number of sub-cycles is increased until the result stabilizes. The averages of r and K are also computed in the same way. A long rewiring process of this kind is used to compute S and other quantities as functions of the temperature; see results in Section 5.2.

When the rewiring temperature is very low one can actually observe that the value of r changes by 10^{-5} or less, and the entropy becomes very small. This is because when the network is very close to its maximum possible assortativity or disassortativity, almost all rewiring steps are rejected, changes in r are very small and the adjacency matrix remains practically constant (thus $\bar{A}_{ab}$ is either very close to 0 or 1, with small contribution to S).

In its present form the algorithm has been applied to relatively small networks ($N < 10^4$ nodes), because the double sum involved in the calculation of the entropy has N^2 terms and the average adjacency matrix $\bar{A}$ requires much memory. For such values of N the part of the code which performs the rewiring is quite fast, taking at most a few seconds to complete. If one wants to apply the rewiring to larger networks without computing the entropy, a check of the adjacency matrix A will still be needed (especially in the assortative case) in order to avoid the formation of multi-link. However, in that case A should be handled as a sparse integer matrix, thus lowering the memory requirements.

5. Results

5.1. Trajectories of Assortative and Disassortative Rewiring in the r-K Plane at Low T

Let us first describe the "trajectories" (see examples in Figures 2 and 3), which are generated in the *r-K* plane when we make a rewiring of the assortative kind, i.e., with the rewiring step always accepted when $\Delta r > 0$ and accepted with probability $\exp(-|\Delta r|/T)$ when $\Delta r < 0$. The rewiring cycles described in this section are relatively short, in comparison to those of Section 5.2, in which we need to reach equilibrium, and T can be large.

The starting point of each trajectory is a network generated through a completely random rewiring. Note that especially for the lower values of γ this "configuration model" network is not uncorrelated, but displays some structural disassortativity.

We observe at the beginning a rapid increase of r along the trajectory. For instance, for a network with 1500 nodes, in Figure 2 the data points represent the values of r and K during 80 sub-cycles of 100 rewirings each. The temperature T is chosen to be quite low, compared to the magnitude order of the variation Δr. The latter is of the order of 10^{-3} as deduced from Equation (6) and Table 2. With these values and units, a temperature of 10^{-6} is sufficiently low to give a very small return probability in the Metropolis algorithm and to make the networks evolve quickly and almost without fluctuations towards their maximum possible assortativity.

The points on the trajectory are seen to converge to a final spot, and typically if we perform 50 sub-cycles of 10,000 steps (not shown in the figure), which take about one second to be completed, the values of r and K become constant up to the sixth decimal digit. This value, however, depends on

the degrees of the hubs effectively present in the network and is reproducible in different runs only if the hubs are generated with the method of cumulative probability (see Section 4).

If at this stage we compute the Shannon entropy S according to Equation (14), we find $S = 0$ exactly. This means that at this temperature the assortative rewiring causes the network in the long run to evolve towards a unique "asymptotic" state.
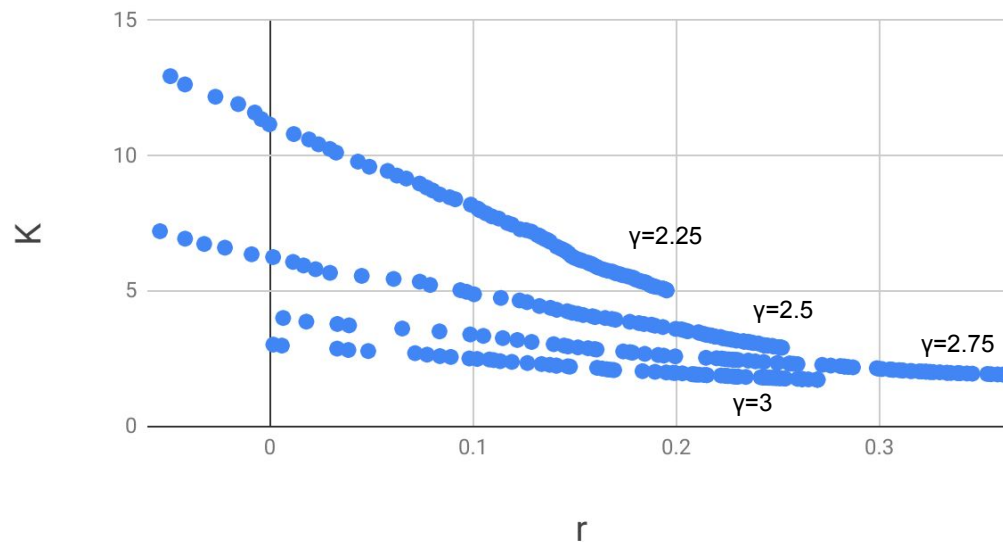

Figure 2. Trajectories of assortative rewiring in the *r-K* plane. Each trajectory begins from the left (initial network obtained by uncorrelated rewiring) and converges to a maximally assortative network on the right. Between two dots on the same trajectory there are 100 rewiring steps with return probability $\exp(-\Delta r/T)$, $T = 10^{-6}$. The number of nodes is $N = 1500$.
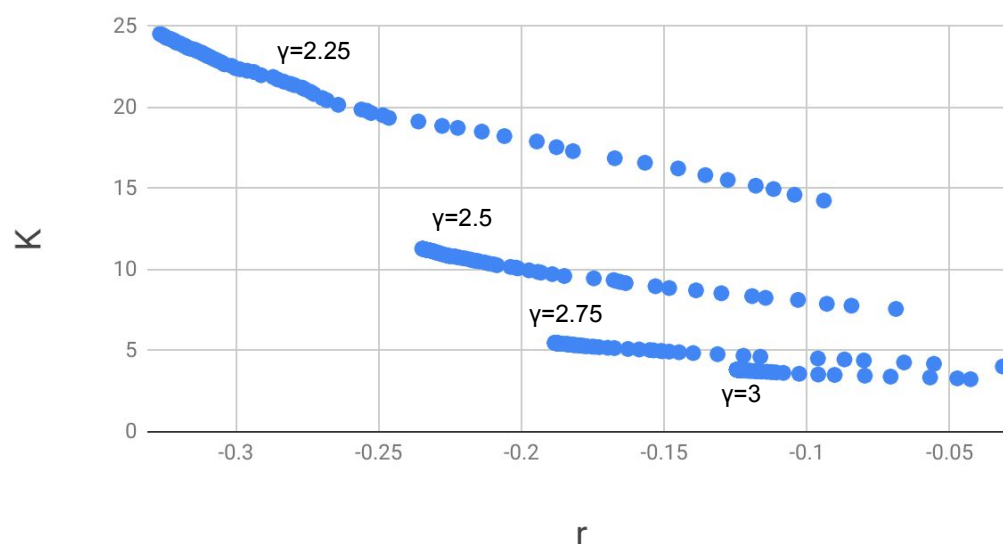
Figure 3. Trajectories of disassortative rewiring in the *r-K* plane. Each trajectory begins from the right (initial network obtained by uncorrelated rewiring) and converges to a maximally disassortative network on the left. Between two dots on the same trajectory there are 100 rewiring steps with return probability $\exp(\Delta r/T)$, $T = 10^{-6}$. The number of nodes is $N = 1500$.

The maximum and minimum values (asymptotic values) that can be obtained for the r coefficient of scale-free networks have been extensively studied in [20]. Those results, however, are referred to networks whose minimum degree is relatively large, typically $k_{min} = 6$; such networks are usually completely connected (giant component equal to 100%). In this work we take instead $k_{min} = 1$, because we are also interested in the network "fragmentation" caused by the strong assortativity and disassortativity. The giant component we find can be very small, down to 10% or less. Most of the asymptotic networks consist, in the assortative case, of isolated couples made of nodes of degree $k = 1$, and then of long chains with variable length, open or closed, made of nodes of degree $k = 2$. Such structures can be easily identified in a complete graphical representation of the network or more "economically" in the table of the values of the number z_2 of second neighbors of each node. For instance, for each node of degree 1 belonging to a couple we obviously have $z_2 = 0$, and for each node of degree 2 belonging to a chain we have $z_2 = 2$ (except at the ends of the chain). Another less obvious occurrence observed in strongly disassortative networks is shown in Figure 4 and involves nodes of degree 3.
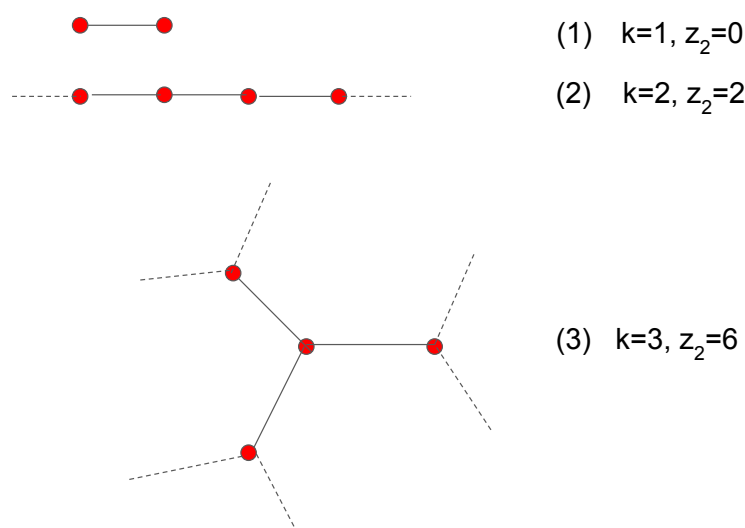

(1)   k=1, z_2=0

(2)   k=2, z_2=2

(3)   k=3, z_2=6

Figure 4. Basic assortative building blocks classified through k (node degree) and z_2 (number of second neighbors). Looking at the table which gives for each node of a strongly assortative network the number z_2 of its second neighbors, we find a large number of nodes with $k = 1$, $z_2 = 0$ (corresponding to Graph (1), isolated couples), of nodes with $k = 2$, $z_2 = 2$ (corresponding to Graph (2), chains), and of nodes with $k = 3$, $z_2 = 6$ (corresponding to Graph (3), linked 3-stars).

In the extreme assortative case the giant component contains all the major hubs, strongly connected among themselves. From the strongly connected core depart long chains (see an example in Figure 5) whose connection or disconnection to the core affects the size of giant component, but has very little influence on the value of r. In other words, a rewiring in which long chains are connected or disconnected at their ends can easily occur also at low temperature, with little effect on r and on the entropy.

In the extreme disassortative case, especially for γ close to 3, it may happen that the network is completely fragmented and there is no giant component. One example is shown in Figure 6. These kinds of networks have been studied theoretically in [29]. In addition, in this case the table of values of the number z_2 of second neighbors is useful in order to identify some patterns without the need to visualize the entire network. For instance, all the hubs of the network in Figure 6 have $z_2 = 0$ exactly.
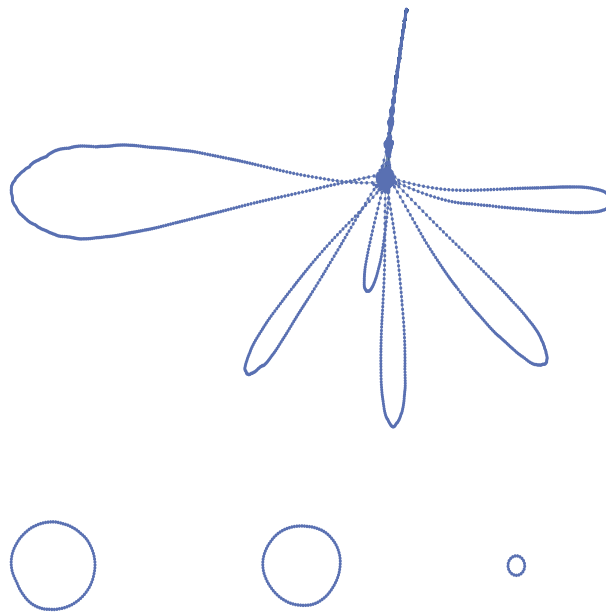
Figure 5. An example of a maximally assortative network with $\gamma = 3$, $k_{min} = 2$. The minimum degree has been taken greater than 1 in this example in order to avoid the formation of isolated couples, which make up the large majority of maximally assortative networks if $k_{min} = 1$. The length of the chains and their connection or disconnection to the the giant component have only a little effect on the total r coefficient, and can therefore vary in each realization, even at a very low temperature (here $T = 10^{-6}$).
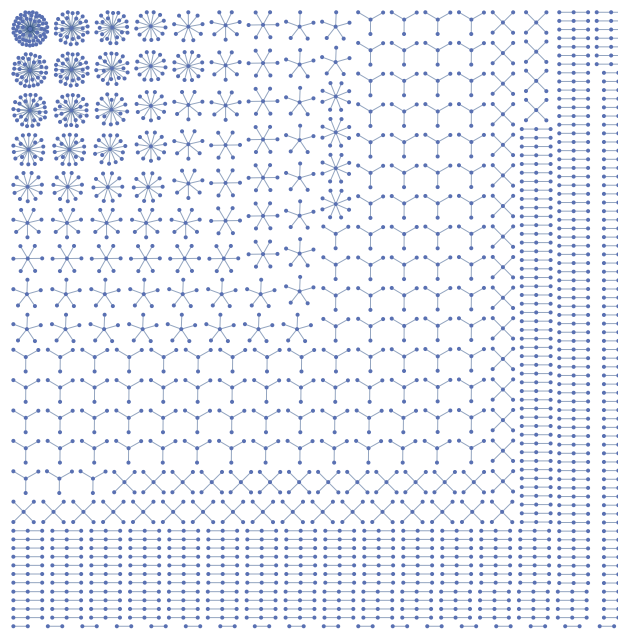
Figure 6. Maximally disassortative network with $\gamma = 2.75$, $k_{min} = 1$. The network is completely fragmented into "stars" (compare [29]). For all hubs the number of second neighbors is exactly zero.

5.2. Equilibrium Rewiring at Variable T

In this subsection we report some preliminary results from assortative rewiring cycles performed at variable temperature, in the range $10^{-5} \leq T \leq 10^{-2}$. These rewiring cycles consist of several long sub-cycles (typically 100 sub-cycles of 10^4 rewiring steps), such that the equilibrium is attained and entropy can be measured by averaging the adjacency matrix on each sub-cycle. We recall (see

Section 5.1) that the value of the temperature for the rewiring algorithm must refer to the magnitude order of the variations of *r*, and this depends in turn on the number of nodes.

The data are still quite noisy and more statistics need to be accumulated for different values of γ, followed by extension to the case of disassortative rewiring. One of the objectives of these measurements is to verify the conjecture of [28] that disassortative networks are entropically favored. Another interesting suggestion from the data in Figure 7 is that there is a correlation between *K* and the size of the giant component, as *T* varies. The exact anticorrelation between *r* and *K* according to Equation (9) is also evident.
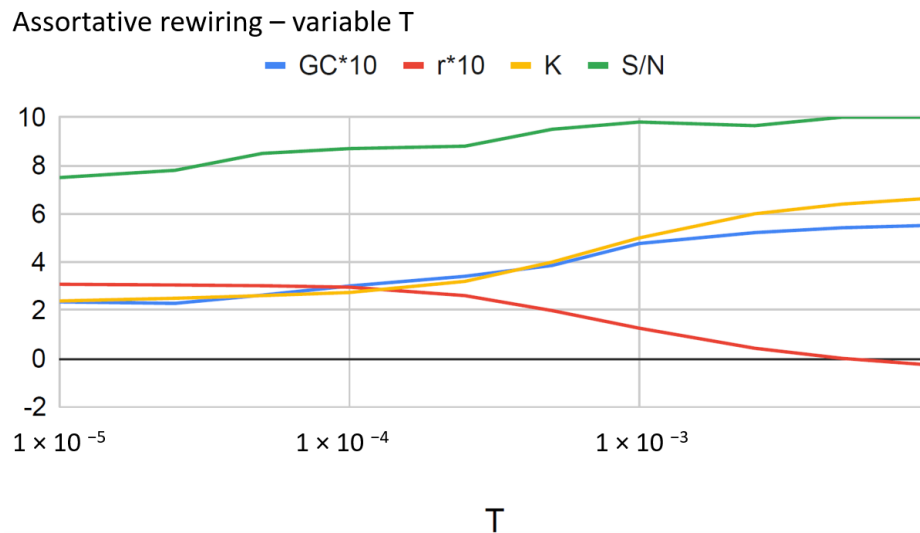

Figure 7. Assortative rewiring in equilibrium at different temperatures *T*. (Network of 1000 nodes, with $\gamma = 2.5$, $k_{min} = 1$.) The entropy per node S/N is shown, along with the average degree of the first neighbors *K*, the Newman assortativity coefficient *r* multiplied by 10 (note the slight structural disassortativity at high temperature), and the fractional size of the giant component multiplied by 10.

6. Conclusions

The method of assortative and disassortative rewiring at variable *T* that we presented in this work appeared to be quite effective for the generation of correlated scale-free networks. At each step of the rewiring process, our algorithm permitted a control of the variations of the assortativity coefficient *r* and of the average degree of the nearest neighbors *K*. The two variations were actually connected through a general relation that we have proven in Section 2.4.

We also proved that the average number $\bar{z}_{2,B}$ of second neighbors in the branching approximation was constant in the rewiring. This property provided further evidence for the absence of epidemic threshold in scale-free networks with exponent γ in the range $2 < \gamma \leq 3$.

If we represented an assortative or disassortative rewiring process at low temperature (i.e., with low return probability) in a *r-K* plane, we obtained an almost linear trajectory converging towards a point, which represents the maximally assortative or disassortative network having the given degree distribution. The position of the trajectory in the plane and its (negative) slope depended on the exponent γ. In general, the value of *K* was smaller for assortative networks, compared to uncorrelated or disassortative networks having the same degree distribution. For a fixed value of *r*, *K* was larger when γ was smaller, therefore the trajectories with small γ lie in the upper part of the plane.

The features of super-assortative and super-disassortative networks were found to depend quite strongly, for a given γ, on the minimum and maximum degree present in the network.

Preliminary evaluations of the network entropy, the size GC of the giant component, *K* and *r* as functions of the rewiring temperature confirmed the exact anti-correlation between *r* and *K* and indicated a positive correlation between *K* and GC.

References

1. Newman, M. Mixing patterns in networks. *Phys. Rev. E* **2003**, *67*, 026126. [CrossRef] [PubMed]
2. Boccaletti, S.; Latora, V.; Moreno, Y.; Chavez, M.; Hwang, D.U. Complex networks: Structure and dynamics. *Phys. Rep.* **2006**, *424*, 175–308. [CrossRef]
3. Cohen, R.; Havlin, S. *Complex Networks: Structure, Robustness and Function*; Cambridge University Press: Cambridge, UK, 2010.
4. Pastor-Satorras, R.; Castellano, C.; Van Mieghem, P.; Vespignani, A. Epidemic processes in complex networks. *Rev. Mod. Phys.* **2015**, *87*, 925. [CrossRef]
5. Barabási, A.L. *Network Science*; Cambridge University Press: Cambridge, UK, 2016.
6. Newman, M. *Networks: An Introduction*; Oxford University Press: Oxford, UK, 2010.
7. Newman, M. Assortative mixing in networks. *Phys. Rev. Lett.* **2002**, *89*, 208701. [CrossRef] [PubMed]
8. Van Mieghem, P.; Wang, H.; Ge, X.; Tang, S.; Kuipers, F. Influence of assortativity and degree-preserving rewiring on the spectra of networks. *Eur. Phys. J. B* **2010**, *76*, 643–652. [CrossRef]
9. D'Agostino, G.; Scala, A.; Zlatić, V.; Caldarelli, G. Robustness and assortativity for diffusion-like processes in scale-free networks. *Europhys. Lett.* **2012**, *97*, 68006. [CrossRef]
10. Noldus, R.; Van Mieghem, P. Assortativity in complex networks. *J. Complex Netw.* **2015**, *3*, 507–542. [CrossRef]
11. Arcagni, A.; Grassi, R.; Stefani, S.; Torriero, A. Higher order assortativity in complex networks. *Eur. J. Oper. Res.* **2017**, *262*, 708–719. [CrossRef]
12. Alstott, J.; Klymko, C.; Pyzza, P.; Radcliffe, M. Local rewiring algorithms to increase clustering and grow a small world. *J. Complex Netw.* **2019**, *7*, 564–584. [CrossRef]
13. Xulvi-Brunet, R.; Sokolov, I. Reshuffling scale-free networks: From random to assortative. *Phys. Rev. E* **2004**, *70*, 066102. [CrossRef]
14. Vázquez, A.; Weigt, M. Computational complexity arising from degree correlations in networks. *Phys. Rev. E* **2003**, *67*, 027101. [CrossRef] [PubMed]
15. Bertotti, M.; Brunner, J.; Modanese, G. The Bass diffusion model on networks with correlations and inhomogeneous advertising. *Chaos Soliton Fract.* **2016**, *90*, 55–63. [CrossRef]
16. Bertotti, M.; Modanese, G. The configuration model for Barabasi-Albert networks. *Appl. Netw. Sci.* **2019**, *4*, 32. [CrossRef]
17. Ray, J.; Pinar, A.; Seshadhri, C. A stopping criterion for Markov chains when generating independent random graphs. *J. Complex Netw.* **2015**, *3*, 204–220. [CrossRef]
18. Gray, C.; Mitchell, L.; Roughan, M. Generating connected random graphs. *J. Complex Netw.* **2019**, *7*, 896–912. [CrossRef]
19. Cimini, G.; Squartini, T.; Saracco, F.; Garlaschelli, D.; Gabrielli, A.; Caldarelli, G. The statistical physics of real-world networks. *Nat. Rev. Phys.* **2019**, *1*, 58–71. [CrossRef]
20. Menche, J.; Valleriani, A.; Lipowsky, R. Asymptotic properties of degree-correlated scale-free networks. *Phys. Rev. E* **2010**, *81*, 046103. [CrossRef]
21. Boguñá, M.; Pastor-Satorras, R.; Vespignani, A. Epidemic spreading in complex networks with degree correlations. In *Statistical Mechanics of Complex Networks*; Springer: Berlin/Heidelberg, Germany, 2003; pp. 127–147.
22. Dorogovtsev, S.; Mendes, J. Evolution of networks. *Adv. Phys.* **2002**, *51*, 1079–1187. [CrossRef]
23. Silva, D.; Ferreira.; S.C.; Cota, W.; Pastor-Satorras, R.; Castellano, C. Spectral properties and the accuracy of mean-field approaches for epidemics on correlated power-law networks. *Phys. Rev. Res.* **2019**, *1*, 033024. [CrossRef]
24. Nekovee, M.; Moreno, Y.; Bianconi, G.; Marsili, M. Theory of rumour spreading in complex social networks. *Physica A* **2007**, *374*, 457–470. [CrossRef]

25. Bertotti, M.; Modanese, G. On the evaluation of the takeoff time and of the peak time for innovation diffusion on assortative networks. *Math. Comp. Model. Dyn.* **2019**, *25*, 482–498. [CrossRef]

26. Molloy, M.; Reed, B. A critical point for random graphs with a given degree sequence. *Random Struct. Algor.* **1995**, *6*, 161–180. [CrossRef]

27. Newman, M.; Strogatz, S.; Watts, D. Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E* **2001**, *64*, 026118. [CrossRef]

28. Johnson, S.; Torres, J.; Marro, J.; Muñoz, M. Entropic origin of disassortativity in complex networks. *Phys. Rev. Lett.* **2010**, *104*, 108702. [CrossRef] [PubMed]

29. Moreno, Y.; Vazquez, A. Disease spreading in structured scale-free networks. *Eur. Phys. J. B* **2003**, *31*, 265–271. [CrossRef]