Abstract. Correlations may affect propagation processes on complex networks. To analyze their effect, it is useful to build ensembles of networks constrained to have a given value of a structural measure, such as the degree–degree correlation $r$, but that are random in other respects and preserve the degree sequence. This can be done through Monte Carlo optimization procedures. Meanwhile, when tuning $r$, other network properties may concomitantly change. Thus, in this work we analyze, for the $r$-ensembles, the impact of $r$ on properties such as transitivity, branching and characteristic path lengths, that are relevant when investigating spreading phenomena on these networks. The present analysis is performed for networks with degree distributions of two main types: either localized around a typical degree (with exponentially bounded asymptotic decay) or broadly distributed (with power-law decay). Size effects are also investigated.

Keywords: random graphs, networks
1. Introduction

Complex networks are realistic substrates for simulating many social and natural phenomena. To address the influence of network topology, primarily, different classes of degree distributions \( P(k) \) can be considered. Meanwhile, for a given distribution of degrees, correlations may give rise to important network structure effects on the process studied \([1]–[6]\). These structural effects may have important consequences; for instance, correlations may shift an epidemic threshold \([6]\). Although correlation effects may be absent in some cases \([7]\), in other ones, they cannot be neglected.

Despite there being efficient algorithms for generating networks with fixed degree–degree correlations \([8]\), real joint probabilities of two or more degrees measured in networks of moderate size may be noisy and hard to model. Then, operationally, average nearest neighbor degree distributions \([9]\) or single quantity measures are used. Although other variants have been defined in the literature \([10, 11]\), as the quantifier of the tendency of adjacent vertices to have similar or dissimilar degrees, we will consider the standard measure of (linear) degree–degree correlations, namely, the assortativity (Pearson) coefficient \([12]\)

\[
r = \frac{\langle kk' \rangle_e - \langle k \rangle_e^2}{\langle k^2 \rangle_e - \langle k \rangle_e^2},
\]

(1)

where \( \langle \cdots \rangle_e \) denotes average over edges and \( k \) and \( k' \) are the degrees of vertices at each ends of an edge. Despite this coefficient being known to present some drawbacks \([13]\), it is a standard and commonly used quantity, and hence worth analyzing. Moreover, it has the advantage of being a single-valued measure, and so easier to control than other multi-valued quantities.

To analyze the influence of correlations, and of any other structural feature, it is useful to build ensembles of networks having that property, while keeping the sequence of degrees fixed. As will be described in section 2, such ensembles can be achieved by means of a suitable rewiring, performed through a standard simulated annealing Monte Carlo (MC) procedure to minimize a given energy-like quantity (the maximum entropy ensemble approach), a function of the graph property to be controlled (\( r \) in our case) \([14]–[16]\). Once we have tuned \( r \), it is important to characterize how other network properties are altered.
as by-products. Some interdependences among certain network properties have already been numerically shown in the literature, for real as well as for artificial graphs [15]. Analytical relations have also been derived [17]–[19]. In view of its crucial role in spreading phenomena [20], we will focus here on the effect of $r$ over typical distance measures as well as on the branching and transitivity of links.

We consider, as a measure of the average separation between nodes, the average path length [21]. In the subsequent calculations we use the expression

$$L = \frac{\sum_{i=1}^{n} \langle L_i \rangle N_i (N_i - 1)}{\sum_{i=1}^{n} N_i (N_i - 1)},$$

where $n$ is the number of (disconnected) clusters and $N_i$ is the number of nodes in cluster $i$. Moreover, with $d_{kj}$ the distance (number of edges along the shortest path) between nodes $k$ and $j$ (taking $d_{kj} = 0$ if the nodes do not belong to the same cluster), then

$$\langle L_i \rangle = \frac{\sum_{i,j,k=1}^{N_i} d_{kj}}{N_i (N_i - 1)}.$$

Alternatively, in order to avoid the issue of the divergence of the distance between disconnected nodes, we consider the inverse, $1/E$, of the so-called efficiency [22]

$$E = \frac{1}{N(N-1)} \sum_{1 \leq i,j \leq N} \frac{1}{d_{ij}},$$

where $N$ is the number of nodes. It represents a harmonic mean instead of the arithmetic one. We also compute the diameter $D = \max \{d_{ij}\}$.

The transitivity of links can be measured via the clustering coefficient [23, 24]

$$C = \frac{6n_\Delta}{\sum_{i=1}^{N} k_i (k_i - 1)},$$

where $n_\Delta$ is the number of triangles and $k_i$ is the degree of node $i$. We also considered the mean value, $\bar{C}$, of the local clustering coefficient $C_i$, defined as $C_i = 2e_i/(k_i(k_i - 1))$, where $e_i$ is the number of connections between the neighbors of vertex $i$ [21]. We took $C_i = 0$ when $k_i = 0$ or 1. Other measures that arise in the decomposition of $r$ [17] will also be considered.

Besides detecting interdependences among structural properties, it is also important to know how these properties depend on the system size $N$. We will analyze these issues for two main classes of degree distributions (Poisson and power-law tailed). We will also investigate real network degree sequences.

2. Networks and ensembles

For each class of networks, we will consider different values of the size, $N$, and the mean degree, $\langle k \rangle$, within realistic ranges.

We consider, as a paradigm of the class of networks with a peaked distribution of degrees, with all its moments finite, the random network of Erdős and Rényi [25]. Within this model, a network with $N$ nodes is assembled by selecting $M$ different pairs of nodes at
Random degree–degree correlated networks

random and linking each pair. The resulting distribution of links is the Poisson distribution

\[ P(k) = e^{-\langle k \rangle} (\langle k \rangle)^k / k! \]

where the mean degree is \( \langle k \rangle = 2M/N \).

We also analyze networks of the power-law type, i.e., with \( P(k) \sim k^{-\gamma} \), \( \gamma > 2 \), corresponding to a wide distribution of degrees, with power-law tails. Then, moments of order \( n \geq \gamma - 1 \) are divergent. We built power-law networks by means of the configuration model [26]. Following this procedure, one starts by choosing \( N \) random numbers \( k \), drawn from the degree distribution \( P(k) \). They represent the number of edges coming out from each node, where these edges have one end attached to the node and another still open. Second, two open ends are randomly chosen and connected such that, although multiple connections are allowed, self-connections are not. This second stage is repeated until each node attains the connectivity attributed in the first step. If eventually an edge has an open end, then it is discarded. However, for large networks, the fraction of discarded edges is negligible. To draw the set of numbers \( k \) with probability \( P(k) = \mathcal{N} k^{-\gamma} \), with \( k_{\text{min}} \leq k \leq k_{\text{max}} \) (the normalization factor hence being \( \mathcal{N} = 1 / \sum_{k_{\text{min}}} k_{\text{max}}^{-\gamma} \)), we used the inverse transform algorithm [27]. Notice that \( k_{\text{max}} \leq N - 1 \) and \( k_{\text{max}} \gg k_{\text{min}} \); thus we determined \( k_{\text{min}} \) to fit the selected value of \( \langle k \rangle \) (within a tolerance of at most 1%), such that

\[
\langle k \rangle = \frac{\sum_{k_{\text{min}}}^{k_{\text{max}}} k^{-\gamma+1}}{\sum_{k_{\text{min}}}^{k_{\text{max}}} k^{-\gamma}} \approx \frac{\gamma - 1}{\gamma - 2} \frac{k_{\text{max}}^{2-\gamma} - k_{\text{min}}^{2-\gamma}}{k_{\text{max}}^{1-\gamma} - k_{\text{min}}^{1-\gamma}} \approx \frac{\gamma - 1}{\gamma - 2} k_{\text{min}}.
\]

It is worth mentioning that the value \( k = N - 1 \) is not usually achieved, the natural cutoff being \( k_c \sim N^{1/(\gamma - 1)} \) [28].

In order to attain a desired value of \( r \), we follow a standard rewiring approach. We want to build an ensemble of networks \( \{G\} \) with a given value of \( r \) (\( r \)-ensemble) but that are maximally random in other aspects, i.e., making as few assumptions as possible about the distribution \( P(G) \). Then, we use an exponential random graph model such that the set of networks \( \{G\} \) has distribution \( P(G) \propto e^{-H(G)} \), where \( H(G) \) is a Hamiltonian or energy-like quantity [14]. In order to get an \( r \)-ensemble with \( r = r_* \), we consider [15]

\[
H(G) = \beta (|r - r_*|),
\]

where \( \beta \) is a real parameter. The ensemble can be simulated by means of an MC procedure: at each step, a rewiring attempt is accepted with probability \( \min\{1, e^{-[H(G') - H(G)]}\} \).

Rewiring steps are performed by randomly selecting two edges that connect the vertices \( a \), \( b \) and \( c \), \( d \), respectively, and substituting those two links by new ones connecting \( a \) and \( b \), \( c \) and \( d \) [29]. Movements yielding double links are forbidden. Notice that this process preserves the connectivity of each node. We start the simulation by taking \( \beta = 0 \) (during at most 100 MC steps (MCS), where each MCS corresponds to \( N \) attempts). The effect of this stage is basically to destroy multiple edges. We did not notice any clear hysteresis effect like those observed when controlling, instead, the number of triangles with a different Hamiltonian [30]. Subsequently, \( \beta \) is increased (in increments of \( \Delta \beta = 1000 \)), at each 50 MCS, until \( r \) stabilizes, typically attaining the prescribed value \( r_* \). Then, the quantities of interest are calculated and the whole process is repeated, starting with a new degree sequence. For power-law degree distributions, we observed that the process is non-ergodic; hence we computed the sample mean and standard deviation over 100 realizations of the described protocol. We checked that the choice of other expressions for \( H(G) \), vanishing at \( r_* \), did not significantly affect the results and just affected the convergence time.
3. Results

Let us start by reporting the effects of \( r \) on the clustering coefficient \( C \).

For the Poisson case, we depict in figure 1(a) the behavior of \( C \) as a function of \( r \) for a fixed number of nodes \((N = 8000)\) and different values of the mean connectivity \( \langle k \rangle \).

Very small values of \( C \) emerge. The transitivity \( C \) increases monotonically with \( r \). This is consistent with the results of [15] (restricted to \( r \geq 0 \)) for such networks. We observe two regimes with a crossover at \( r \approx 0.5 \): a very slight increase with \( r \) below the crossover and a more pronounced one in the region above it. The existence of two regimes could be related to the asymmetric character of \( r \), which does not measure assortativity and disassortativity on the same grounds. Below the crossover, \( C \) linearly increases with \( \langle k \rangle \) by about one order of magnitude within the analysis range. Meanwhile, above the crossover, \( C \) remains of the same order when the average connectivity increases, even for small \( \langle k \rangle \) (also see the inset of figure 1(a) where \( C \) is plotted versus \( \langle k \rangle \) for selected values of \( r \)).

In section 4, we will discuss these issues in more detail. For the mean local clustering coefficient \( \bar{C} \), we obtained a dependence on \( r \) qualitatively similar to that observed for the clustering coefficient \( C \). However, the increase of \( \bar{C} \) with \( \langle k \rangle \) is linear for any fixed \( r \). For \( r = 0 \), \( C = \bar{C} = \langle k \rangle / N \), as expected.

In figure 1(b), size effects are exhibited for \( \langle k \rangle = 4 \), representative of the other values considered. As the number of nodes increases, \( C \) decays as \( C \sim 1/N \) for all \( r \) (as depicted in the inset). Therefore, in the \( r \)-ensemble of Poisson networks, transitivity is only a finite size effect and vanishes in the infinite network (thermodynamic) limit with the same asymptotic law \( C \sim 1/N \) as for an uncorrelated random graph [31].

Figure 1. Clustering coefficient \( C \) as a function of \( r \) for Poisson networks. (a) \( N = 8000 \) and different values of \( \langle k \rangle \) indicated in the figure. The graph shows a monotonic increase of \( C \) with \( r \). There are two regimes: a very slight increase of \( C \) with \( r \) below \( r \approx 0.5 \) and a more pronounced one above it. (b) \( \langle k \rangle = 4 \) and different numbers of nodes \( N \), also indicated in the figure. As the number of nodes increases, \( C \) decays with the asymptotic law \( C \sim 1/N \), characteristic of uncorrelated random graphs. Standard errors are about 10%. Dotted lines are a guide to the eyes. The insets show \( C \) versus \( \langle k \rangle \) (a) and \( N \) (b) for selected values of \( r \) (\( -0.6, 0.0, 0.6 \) and 0.8).
For the power-law class, the range of allowed values of $r$ is restricted. That is, values of $r$ arbitrarily different from zero cannot be attained in typical realizations of the MC protocol described in section 2. In order to determine the typical maximal (minimal) values, $r_{\text{max}}$ ($r_{\text{min}}$), we imposed $r_s = 1$ ($-1$) and detected the stationary values of $r$. The time evolution of $r$ for $r_s = 1$ ($-1$) is illustrated in figure 2(a) for $\gamma = 3.5$, $N = 5000$ and $\langle k \rangle \approx 4.00 \pm 0.04$. Notice the large deviations amongst the steady values of different realizations, mainly for the upper bound. We verified that this picture does not change by implementing other definitions of $H(G)$ in equation (7), e.g., $\beta |r - r_s|^\alpha$, with $\alpha \neq 1$. Average extreme values (over 100 samples, after $3 \times 10^4$ MCS) are displayed in figure 2(b), as a function of the system size, for different values of $\gamma$. For fixed size, the lower $\gamma$, the narrower the allowed interval of $r$. In fact, in networks constrained to a given degree sequence, structural limitations (or correlations) arise: either multiple connections or disassortative two-vertex correlations [32]. For instance, the exclusion of multiple connections hampers the natural tendency that hubs connect among them, hence diminishing the assortativity. This effect is more pronounced for smaller $\gamma$. For fixed $\gamma$, the interval shrinks with system size, for $\gamma < 3$, due to the divergence of fluctuations in the large $N$ limit [13]. In [9], a similar restriction was also observed for $2 < \gamma < 3$, although instead of the average connectivity, $k_{\text{min}}$ was kept constant ($k_{\text{min}} = 6$). In that case, it was reported that the upper and lower bounds both tend to zero, and hence $r \to 0$ in the infinite network limit. In fact, we observe that in that interval of $\gamma$ (e.g., $\gamma = 2.5$), both bounds are negative and as $N$ increases the allowed interval collapses to a negative value that tends to zero. We noticed a restriction in the correlation bounds for $\gamma > 3$ too. As $N$ grows, the lower bound also increases towards zero or at least to a small finite value. Simultaneously, the upper bound seems more stable;
Figure 3. Clustering coefficient $C$ as a function of $r$, as in figure 1 but for power-law networks with $\gamma = 4.0$. (a) $N = 8000$ and different values of $\langle k \rangle$. In the assortative region, $C$ reaches larger values than in the Poisson class. (b) $\langle k \rangle = 4$ and different numbers of nodes $N$. Note that for assortative networks a finite degree of clustering seems to persist for large networks. Standard errors reach 50% for the lowest values of $C$. In the inset, missing values are due to the limitation in attaining the prescribed values of $r$.

however its asymptotic behavior is not neat yet, even having considered up to $N > 10^5$. Moreover, as $N$ increases, it takes longer to attain steady states.

The allowed interval of $r$ is quite restricted for scale-free networks. However, we still analyzed systematically cases with $\gamma > 3$ ($\gamma = 3.5$, 4.0 and 4.5), yielding finite second moment. Even in this range, the accessible interval of $r$ is limited; thus, we proceeded as follows. If the desired $r_*$ is not attained, within a tolerance of $10^{-3}$, in $2 \times 10^4$ MCS, that instance is discarded and we carry out a new trial. If we do not attain 100 successes in 200 trials, the procedure is interrupted. Alternatively to using the configuration model, we also started from networks generated by preferential attachment [33]; similar results were obtained.

The outcomes for the power-law class with $\gamma = 4.0$ are displayed in figure 3. Standard errors are larger in the power-law case, likely due to the variability in the tails of the distribution of links from sample to sample. Outcomes for the other two values of $\gamma$ studied (3.5 and 4.5) display features similar to those of the case $\gamma = 4.0$ used as an illustrative example, despite the third moment becoming divergent at $\gamma = 4.0$. Two regimes are also observed, with the crossover now closer to $r = 0$, but some qualitative differences appear in comparison to the Poisson case. $C$ rapidly increases with $r$, attaining, for assortative networks, larger values than in the Poisson class. These large values are due to the inclusion of highly connected nodes, absent in the Poisson networks, that for large $r_*$ tend to gain links among themselves, contributing strongly to $r$ and also to $C$.

With respect to finite size effects, below the crossover the small non-null $C$ is again due only to the finite size of the network. However, in the assortative region (above the crossover), it seems that a finite degree of clustering persists for large networks (see the inset of figure 3(b)), in contrast to the Poisson case and to the disassortative region. In fact, notice that, when $N$ increases by one order of magnitude, $C$ decreases also by one
Figure 4. Distance measures $1/E$, $L$, and diameter $D$, as a function of $r$, for Poisson (open symbols) and $\gamma = 4.0$ power-law (filled symbols) networks. In all cases $N = 8000$ and $\langle k \rangle = 4.00 \pm 0.04$. The inset is a zoom of the main plot. At first sight, the two types of network appear to display similar values of $1/E$ and $L$, for a given $N$ and $\langle k \rangle$. The diameter is more sensitive to the type of network and is more affected by $r$ in the Poisson case.

order of magnitude in the disassortative region, while $C$ remains of the same order in the assortative interval. Even if $C$ vanishes in the infinite size limit, since the decay is very slow, then an effective clustering would remain in moderate, or even large, size networks. We will discuss the interplay between $C$ and $r$ further in section 4. For the mean local clustering coefficient $\bar{C}$, a qualitatively similar dependence on $r$ is observed, but with smaller values. Moreover, $\bar{C}$ increases linearly with $\langle k \rangle$, in the analysis range, for any fixed $r$, not only for disassortative networks, and $\bar{C}$ decays with $N$ for any $r$. For $r = 0$, $C = \bar{C} = (\langle k^2 \rangle - \langle k \rangle)^2/(N \langle k \rangle^3)$ [19], as expected.

Let us analyze now the influence of $r$ on network characteristic lengths. The dependence of the measures $1/E$, $L$, and $D$ on $r$ is depicted in figure 4, for Poisson and power-law distributed networks, with $N = 8000$ and $\langle k \rangle \simeq 4$. $1/E$ and $L$ have close values, systematically shifted. In a first approximation the two types of network yield similar values of $1/E$ (and hence also $L$), given $N$ and $\langle k \rangle$. However, the diameter $D$ is more dependent on the type of network. It is larger and is more strongly affected by $r$ in the homogeneous Poisson case.

Plots of $1/E$ versus $r$ for different values of $N$ and $\langle k \rangle$ are shown in figures 5 and 6 for Poisson and power-law networks respectively. In both cases, the networks display the small-world property [21] (even smaller in the power-law case) with a slow (logarithmic) increase with $N$ and a smooth decrease with $\langle k \rangle$ (see the insets of figures 5 and 6).

In the Poisson case, the mean path does not depend on $r$ significantly when $\langle k \rangle$ is not too small ($\geq 6$), as indicated by the relatively flat plots in figure 5(a). Only for small $\langle k \rangle$ are there important effects for assortative correlations. For instance, for $\langle k \rangle \simeq 4$ (figure 5(b)), $1/E$ increases by about two units from $r \simeq 0$ to $r \simeq 1$, for all the analysis sizes. This effect is still larger for $\langle k \rangle = 2$, where $L$ increases by a factor of about 2 in the same interval of $r$, as shown in figure 5(a) for $N = 8000$. 

doi:10.1088/1742-5468/2013/02/P02024
Figure 5. Mean distance $1/E$ as a function of $r$ for Poisson networks. (a) $N = 8000$ and different values of $\langle k \rangle$ indicated in the figure. The networks exhibit the small-world property. (b) $\langle k \rangle = 4.00 \pm 0.04$ and different numbers of nodes $N$ indicated in the figure. The effects of $r$ on the mean path are significant only for small $\langle k \rangle$ because of the fragmentation of the network.

Figure 6. Mean distance $1/E$ as a function of $r$ as in figure 5 but for a power-law network with $\gamma = 4.0$. In this case, paths are shorter than in the Poisson nets.

In order to further interpret these results, we investigated the cluster structure of the resulting rewired networks. We measured the size of the largest cluster (let us call it $N_1$), the number $n$ of clusters, and the average size $S$ of clusters different from the largest one. The plots are presented in figure 7 for $\langle k \rangle = 2$ and 4. For $\langle k \rangle = 4$, the relative size of the largest cluster (circles) is about 0.98 for most of the range of $r$; notice however that it slightly decays towards $|r| = 1$ (which is more evident for $\langle k \rangle = 2$). As $\langle k \rangle$ increases, the number of fragments rapidly decays and the average size $S$ (triangles) tends to unity, meaning that only single nodes are disconnected (recall also that $P(0) = e^{-\langle k \rangle}$). Therefore the increase of the mean distance towards $|r| = 1$, observed for low $\langle k \rangle$ in figure 5, may simply reflect the fragmentation of the network. Clearly, for high values of the assortativity, the network tends to fragment into groups of vertices that have the same degree. Moreover,
Figure 7. Cluster analysis. Plots of $N_1/N$, where $N_1$ is the size of the largest cluster (circles), average size of finite size clusters, $S$ (squares), and number of clusters, $n$ (triangles), as a function of $r$ for Poisson networks with $\langle k \rangle = 2$ (open symbols) and 4 (filled symbols). The fragmentation of the network observed for high values of the assortativity reflects the grouping of vertices with the same degree. The figure shows outcomes for $N = 8000$. Outcomes for sizes $N = 4000$, $8000$ and $16\,000$ all collapse onto single curves (not shown).

For values of $\langle k \rangle$ approximately larger than 1, the percolation analysis performed by Noh on Poisson networks [16] shows that the size of the largest cluster is smaller for assortative networks than for disassortative and neutral ones. Meanwhile, as $\langle k \rangle$ increases, the fraction of vertices that do not belong to the largest cluster becomes negligible, although more slowly for a more assortative network. Therefore, in such a large $\langle k \rangle$ limit, the mean distance remains invariant under changes of $r$. Hence, transport processes modeled in these networks may suffer an important impact when $r$ is large and $\langle k \rangle$ small. The longer the typical separation between nodes, the slower the propagation.

In power-law networks, for fixed $N$ and $\langle k \rangle > 2$, there is an interval of $r$ where paths are shorter than in the Poisson nets (figure 4), and still shorter as $\gamma$ decreases (not shown). Moreover $1/E$ becomes more sensitive to the coefficient $r$ (smile shape), in the region where plots are flat for Poisson networks. Notice also that minimal mean paths occur for slightly assortative correlations ($r \gtrsim 0$), slowly increasing with $N$ (figure 6(b)). The analysis of clusters for $\gamma = 4.0$ reveals that for $\langle k \rangle \geq 4$ there is a single cluster of size $N$, for all $r$. Only for $\langle k \rangle = 2.0$ did we observe fragmentation with $N_1/N \simeq 0.7 - 0.8$, $n/N \simeq 0.06$, $S \simeq 4$ for all $N > 2000$ (plots not shown).

For the mean path $L$, we observed qualitatively similar outcomes, although shifted to slightly higher values, as illustrated in figure 4.

We also applied the rewiring procedure described in section 2 to real degree sequences. Networks were symmetrized and edge weights were ignored. In figure 8 we depict the behavior of $C$ and $1/E$ versus $r$, for the PGP (encrypted communication network using the Pretty Good Privacy encryption algorithm) largest component [34], P2P (Gnutella peer-to-peer network) [35], PGR (electrical power grid of the western United States), an example of a small-world network [36] and APC (astrophysics collaboration network) [31].

doi:10.1088/1742-5468/2013/02/P02024
Figure 8. $C$ and $1/E$ versus $r$, for real networks. Original values, before rewiring, are also indicated (filled symbols). PGP (Pretty Good Privacy encrypted communication network) [34]: $N = 10,680$, $\langle k \rangle \approx 4.55$; P2P (Gnutella peer-to-peer network) [35]: $N = 10,876$, $\langle k \rangle \approx 7.35$; PGR (power grid) [36], $N = 4,941$, $\langle k \rangle \approx 2.67$; APC (astrophysics collaboration) [31]: $N = 16,706$, $\langle k \rangle \approx 14.5$. Besides the different details of real degree sequences, we can interpret the main features of these nets in terms of those observed for the Poisson and power-law classes.

First notice that in all cases the clustering $C$ is much larger in real networks than in the randomized ($r$-ensemble) ones, as already observed for other examples in [15]. The mean distance is also typically larger in the real networks. An exception is presented by the P2P network, characterized by a value of $1/E$ typical of the $r$-ensemble. Rewired real networks display some of the typical behaviors observed for the artificial cases. Let us make some remarks arising from comparisons. (i) PGR (power grid) displays plots of $C$ versus $r$ and $1/E$ versus $r$ that are in good accord with those observed for similar parameters $\langle k \rangle$ and $N$ of the Poisson case. In fact its degree distribution decays exponentially. (ii) P2P presents power-law decay of the degree distributions, for $k > 10$, with exponent close to $\gamma = 4$. Both plots of P2P are in agreement with those obtained for the $\gamma = 4$ class with similar values of $N$ and $\langle k \rangle$, despite the distributions only having in common the power-law tail. (iii) PGP (of size similar to P2P) has a degree distribution that decays with exponent $\gamma < 3$ for $k < 50$ and $\gamma \approx 4$ beyond [34]. The plot for $1/E$ versus $r$ presents larger values of $1/E$ than P2P, consistent with its $\langle k \rangle$. However, the plot for $C$ versus $r$ of PGP deviates from the standard behavior, presenting larger values of $C$ that increase with $r$ in a single regime. The interval of allowed values of $r$ is sensitive also to features other than the tails. These deviations can be attributed to different initial power-law regimes. (iv) Finally, APC has a power-law decay with exponent $\gamma \approx 1$ and exponential cutoff for $k > 50$ [31]. The low and constant plot of $1/E$ versus $r$ is expected for a network with large $\langle k \rangle$, almost independently of the class of degree distribution. The large values of $C$ are also consistent with heterogeneous distributions with large $\langle k \rangle$.

Then, despite the different details of real degree distributions, the main observed features can be interpreted in terms of the analysis classes with corresponding values of parameters $\langle k \rangle$ and $N$.

doi:10.1088/1742-5468/2013/02/P02024
4. Discussion and final remarks

For all classes of networks considered, $C$ increases with $r$ in the whole allowed range of $r$. This behavior has already been observed in [15], where only positive values of $r$ were analyzed and also in [18], although different definitions of clustering and correlation were used. However, we observed that, in the $r$-ensemble, a non-vanishing clustering coefficient $C$ is typically due to finite size effects, such that, in the large size limit, network transitivity vanishes as $1/N$. In contrast, for power-law networks characterized by $r$ above a threshold, a significantly non-null transitivity arises, apparently persisting for large $N$.

In any case, since rewiring in the $r$-ensemble turns $C$ typically small, transitivity does not seem to contribute for attaining the prescribed value of $r$. To identify the factors that contribute to $r$, it is useful to rewrite equation (1). Recalling that $\langle k^n \rangle_e = \langle k^{n+1} \rangle / \langle k \rangle$ [37], then equation (8) becomes

$$r = \frac{\langle k \rangle^2 \langle kk' \rangle_e - \langle k^2 \rangle^2}{\langle k \rangle \langle k^3 \rangle - \langle k^2 \rangle^2}. \quad (8)$$

Expression (9) is determined by the three first raw moments of $P(k)$, and also by $B$ and $C$ which are the quantities embodying the information on the linear degree–degree correlations.

For the Poisson distributed networks, by taking into account the analytical expressions for the moments of $P(k)$, it is straightforward to see that

$$r = B - \langle k \rangle + C. \quad (10)$$

Clearly, disassortative correlations are favored by vanishing $C$. Only for positive $r$ is the growth of $C$ convenient, but $B - \langle k \rangle$ can vary in a wider interval than $C$ (twice the size in this case). The existence of two regimes in the increase of $C(r)$, observed in figure 1, is consistent with this picture. In other words, equation (10) indicates that, in the process of rewiring of a Poisson network to attain $r_\star$, as soon as $P(k)$ is conserved and $C$ remains very small, then $r$ is ruled predominantly by $B$.

For the power-law distributed networks, some qualitatively similar effects occur, as long as the relation of $r$ with $B$ and $C$ is always linear and $C$ is constrained to a narrower interval than $B$. The formation of triangles also in this case contributes only for assortative values of $r$ (above the crossover), with values of $C$ larger than in the Poisson case but still small. The existence of two regimes in the increase of the branching must drive rewiring. In contrast to the Poisson case, the other terms in equation (9), related to the moments of
Random degree–degree correlated networks

$P(k)$, might have a crucial influence on $r$ because of the divergence, in the infinite network limit, of the $n$th moments for $\gamma \geq n + 1$.

Let us analyze the large $k_{\text{max}}$ (and hence $N$) limit, setting aside the marginal (logarithmic) cases. Here, we use the result $k_{\text{max}} \sim N$. However, if we used instead $k_c \sim N^{1/(\gamma - 1)}$, the conclusions would remain the same. Considering the expressions for the moments (e.g., equation (6)), one has, for $3 < \gamma < 4$: $r \sim [B - \mathcal{O}(1)]/\mathcal{O}(N^{4-\gamma})$, while for $2 < \gamma < 3$: $r \sim [B - \mathcal{O}(N^{3-\gamma})]/\mathcal{O}(N)$, meaning that $\mathcal{O}(x^a) \sim ax^a$, with $a > 0$. To approach the lower limit $r = -1$, one must have minimal $B$. If it is of order greater than the other terms in the numerator, then one cannot have negative $r$, because $B$ is non-negative and it will dominate the numerator. Thus, negative $r$ can arise only if $B$ is of the same or lower order. But in that case $r \to 0$ in the large $N$ limit. This explains why the lower bound $r_{\text{min}}$ tends to 0 when $\gamma \leq 4$ (see figure 2(b)). Along these lines, however, $r_{\text{min}}$ is not expected to vanish when $\gamma > 4$, but to tend to a small finite value. Similarly, to attain a non-null upper bound of $r$, $B$ needs to grow like the denominator; otherwise, the upper bound will be negative and also vanish when $N \to \infty$, leading to the collapse of the upper bound too. However, this does not necessarily happen if $B$ is driven to grow enough during rewiring, which is what seems to happen according to figure 2(b).

The connection between $r$ and distance measures is not so direct analytically. Numerical results showed that, for networks with localized distribution of links, changing $r$ modifies the mean path length significantly only when correlations are assortative ($r > 0.5$) and $\langle k \rangle$ small. These changes could be related to the induced fragmentation, that diminishes with increasing $\langle k \rangle$. Then, the impact of $r$ becomes less important as $\langle k \rangle$ increases. Meanwhile, the influence on the diameter is more dramatic. In power-law networks, the modification of the mean path length by $r$ is a bit more marked even if fragmentation is absent for $\langle k \rangle \geq 4$, while the diameter is not greatly affected. In both cases, the modification of characteristic lengths that occurs when varying $r$ may affect transport processes and should be taken into account when interpreting or designing numerical experiments on these networks.

Acknowledgments

We acknowledge partial financial support from Brazilian Agency CNPq. The authors are grateful to Professor Thadeu Penna for having provided the computational resources of the Group of Complex Systems of the Universidade Federal Fluminense, Brazil, where some of the simulations were performed.

References

[1] Vazquez A, 2006 Phys. Rev. E 74 056101
[2] Small M, Xu X, Zhou J, Zhang J, Sun J and Lu J, 2008 Phys. Rev. E 77 066112
[3] Cohen R and Havlin S, 2003 Phys. Rev. Lett. 90 058701
[4] Xue Y, Wang J, Li L, He D and Hu B, 2010 Phys. Rev. E 81 037101
[5] Moreno Y, Gómez J B and Pacheco A F, 2003 Phys. Rev. E 68 035103(R)
[6] Gallos L K, Song C and Makse H A, 2008 Phys. Rev. Lett. 100 248701
[7] Boguñá M, Pastor-Satorras R, 2002 Phys. Rev. E 66 047104
[8] Boguñá M, Pastor-Satorras R and Vespignani A, 2003 Phys. Rev. Lett. 90 028701
[9] Pusch A, Weber S and Porto M, 2008 Phys. Rev. E 77 017101
[10] Menche J, Valeriani A and Lipowsky R, 2010 Phys. Rev. E 81 046103
Random degree–degree correlated networks

[10] Newman M E J, 2003 Phys. Rev. E 67 026126
[11] Xulvi-Brunet R and Sokolov I M, 2004 Phys. Rev. E 70 066102
[12] Newman M E J, 2002 Phys. Rev. Lett. 89 208701
[13] Serrano M A, Boguñá M, Pastor-Satorras R and Vespignani A, Correlations in complex networks, 2007 Large Scale Structure and Dynamics of Complex Networks: From Information Technology to Finance and Natural Sciences ed G Caldarelli and A Vespignani (Singapore: World Scientific)
[14] Park J and Newman M E J, 2004 Phys. Rev. E 70 66117
[15] Foster D V, Foster J G, Grassberger P and Paczuski M, 2011 Phys. Rev. E 84 066117
[16] Noh J D, 2007 Phys. Rev. E 76 026116
[17] Estrada E, 2011 Phys. Rev. E 84 047101
[18] Serrano M and Boguñá M, 2006 Phys. Rev. E 74 056114
[19] Dorogovtsev S N, 2004 Phys. Rev. E 69 027104
[20] Newman M E J, 2001 Proc. Nat. Acad. Sci. 98 404
[21] Watts D J and Strogatz S H, 1998 Nature 393 440
[22] Boccaletti S, Latora V, Moreno Y, Chavez M and Hwang D-U, 2006 Phys. Rep. 424 175
Latora V and Marchiori M, 2001 Phys. Rev. Lett. 87 198701
[23] Barrat A and Weigt M, 2000 Eur. Phys. J. B 13 547
[24] Newman M E J, Strogatz S H and Watts D J, 2001 Phys. Rev. E 64 026118
[25] Erdős P and Rényi A, 1959 Publ. Math. 6 290
Albert R and Barabási A L, 2001 Rev. Mod. Phys. 74 47
[26] Molloy M and Reed B, 1995 Random Struct. Algorithms 6 161
Newman M E J, 2010 Networks: An Introduction (New York, NY: Oxford University Press)
[27] Ross S M, 2006 Simulation (New York: Academic)
Newman M E J and Barkema G T, 1999 Monte Carlo Methods in Statistical Physics (Oxford: Oxford University Press)
[28] Dorogovtsev S N and Mendes J F F, 2002 Adv. Phys. 51 1079
[29] Maslov S and Sneppen K, 2002 Science 296 910
[30] Foster D, Foster J and Paczuski M, 2010 Phys. Rev. E 81 046115
[31] Newman M E J, 2001 Proc. Nat. Acad. Sci. USA 98 404
Data at www-personal.umich.edu/mejn/netdata/
[32] Park J and Newman M E J, 2003 Phys. Rev. E 68 026112
[33] Dorogovtsev S N, Mendes J F F and Samukhin A N, 2000 Phys. Rev. Lett. 85 4633
[34] Boguñá M, Pastor-Satorras R, Díaz-Guilera A and Arenas A, 2004 Phys. Rev. E 70 056122
[35] Data can be found at http://snap.stanford.edu/data/p2p-Gnutella04.html
[36] Watts D J and Strogatz S H, 1998 Nature 393 440 data can be found at www-personal.umich.edu/mejn/netdata/
[37] Dorogovtsev S N, Ferreira A L, Goltsev A V and Mendes J F F, 2010 Phys. Rev. E 81 031135

Data at www-personal.umich.edu/mejn/netdata/