From Scale-free to Erdos-Rényi Networks

Jesús Gómez-Gardeñes\textsuperscript{1,2} and Yamir Moreno\textsuperscript{1,3}

\textsuperscript{1}Instituto de Biocomputacion y Física de los Sistemas Complejos (BIFI), Universidad de Zaragoza, Zaragoza E-50009, Spain.
\textsuperscript{2}Departamento de Física de la Materia Condensada and Instituto de Ciencia de Materiales de Aragón (ICMA), Universidad de Zaragoza - CSIC, Zaragoza E-50009, Spain.
\textsuperscript{3}Departamento de Física Teórica, Universidad de Zaragoza, Zaragoza E-50009, Spain.

We analyze a model that interpolates between scale-free and Erdos-Rényi networks. The model introduced generates a one-parameter family of networks and allows to analyze the role of structural heterogeneity. Analytical calculations are compared with extensive numerical simulations in order to describe the transition between these two important classes of networks. Finally, an application of the proposed model to the study of the percolation transition is presented.

PACS numbers: 89.75.Fb, 05.70.Jk

I. INTRODUCTION

In the last several years, graph theory has experienced a burst of activity as many real systems can be represented and modeled as networks. A network is made up of vertices representing the interacting elements of the system and of edges that stand for the interactions among them. Network modeling comprises the analysis and characterization of the structure of networks as well as their modeling in terms of generic models aimed at reproducing the features found in real systems. The second important branch has to do with dynamics on networks. This has lately attracted the attention of many scientists as it is ultimately related with the functioning of the system that is being modeled.

In this paper, we deal with the first of these areas of research. The seminal paper by Barabási and Albert showed that many real world networks can not be described by graphs where the connectivity distribution (i.e., the probability that a given node has a given number of links) follows a Poisson-like distribution. Indeed, Barabási and Albert showed that most real networks are heterogeneous in the sense that the probability that a node is connected to another nodes follows a power-law distribution $P(k) \sim k^{-\gamma}$, where $\gamma$ usually lies between 2 and 3. These networks were termed scale-free networks.

Soon afterwards, many studies have dealt with the analysis and characterization of models that generate scale-free networks, along with other global and local topological properties found in real networked systems. In particular, models based on the mechanism of preferential attachment (PA), no matter if the network is growing or not, have been extensively studied in the last years. There are some models in which the PA rule is limited to a neighborhood due to geographic constraints or lack of global knowledge, or where its linear character is investigated. While today we have recognized that preferential attachment is not a necessary condition for the formation of scale-free networks, it seems to be clear that it is an important mechanism. Indeed, most of the existing models intrinsically incorporate a preferential attachment like rule. On the other hand, uniform random linking of nodes on growing networks gives rise to networks where the degree distribution decays exponentially fast with the degree $k$, thus producing homogeneous networks with a well defined (and meaningful) average value for $k$.

The combination of the two rules, i.e., uniform and preferential linking, have been also analyzed in several models for interpolating between scale-free and exponential networks. For instance, Liu et al. have studied a model in which the probability of establishing new links goes as a linear combination of both in such a way that a new link is established between a node $i$ and a new one proportionally to $(1-p)k_i + p$, where $p$ weights the contribution of the two mechanisms. However, in previous models of this sort, there is an assumption that does not apply always. It has to do with the fact that the network always grows around a single component of connected nodes and uniform or preferential links from the emerging nodes are always made with elements belonging to this unique cluster. This single component grows linearly in time until it reaches the size of the network. Since there are no clusters of nodes other than the giant component, the models can not account for phenomena such as the coalescence of small networks into a larger one, nor for situations in which more than one node is added to a preexisting structure at each time step, features that may be relevant in social, economic and other networked systems.

In this paper, we analyze a model that interpolates between Erdos-Rényi (ER) and scale-free (SF) networks as far as the degree distribution is concerned through a tunable parameter. By construction, new links are not always established with nodes previously incorporated to the network. We explore analytically and numerically the time behavior of nodes attachment as well as of the
degree evolution. We find that, depending on the interplay between uniform and preferential linking, the transition from an ER like network to an SF one is smooth or more abrupt. We finally discuss other topological properties and perform a numerical percolation study that highlights the differences in their structure. The model presented here is useful as it provides a unique recipe to go progressively from homogeneous to heterogeneous topologies as well as for exploring the interplay between them.

II. THE MODEL

The model introduced in this work generates a one-parameter family of complex networks. This parameter, $\alpha \in [0,1]$, measures the degree of heterogeneity of the final networks. Let us assume the final size of the network to be $\Omega$. The network is generated in the following way:

(i) Start from a fully connected network of $m_0$ nodes and a set $\mathcal{U}(0)$ of $(\Omega - m_0)$ unconnected nodes.

(ii) At each time step choose a new node $j$ from $\mathcal{U}(0)$.

(iii) This node makes a link in two ways:

(a) With probability $\alpha$ it links to any other node $i$ of the whole set of $\Omega - 1$ nodes with uniform probability

\[ \Pi_i^{\text{uniform}} = (\Omega - 1)^{-1}. \]

(b) With probability $(1 - \alpha)$ establish a link following a preferential attachment strategy, that is, the probability for any other node $i$ to attach to node $j$ is a function of its connectivity as,

\[ \Pi_i^{PA} = F(k_i), \]

where different choices for the functional form of $F(x)$ are analyzed below.

(iv) Repeat $m$ times step (iii) for the same node $j$.

(v) Repeat $\mathcal{U}(0) = (\Omega - m_0)$ times steps (ii) to (iv).

A schematic plot of the linking procedure at step (iii) is shown in Fig. 1. The above rules allow for the coexistence of two classes of nodes. On one hand, there are nodes with at least one link. This set will be referred to henceforth as the connected set $N(t)$. On the other hand, there is another set $\mathcal{U}(t)$ of isolated nodes such that its size is $\Omega - N(t)$. At variance with other models in which there are only nodes with connectivity different from zero and thus the connected component grows linearly with time, the above rules allows the addition of more than one node to the set $N(t)$ as a result of random linking. Therefore, we expect the time dependency of $N(t)$ to be highly non trivial.

III. NETWORK GROWTH AND DEGREE EVOLUTION

In order to describe the evolution of the nodes degree, one has to consider the functional form of $F(x)$ for the preferential attachment probability. However, we can take into account some previous considerations that do not depend on the particular form of $F(x)$.

First of all, it is useful to consider two kind of links in order to analyze the model. Namely, the ones that arise from a uniform random choice, $k^u$, and the remaining, $k^{pa}$, corresponding to the implementation of the preferential attachment rule. The dynamics of $k^u$ is completely independent of the dynamics of the PA links, $k^{pa}$, but the opposite is not necessarily true. From this, it follows that the probability that one node has $k^u$ uniform links, $P^u(k^u)$, is a Poisson distribution with $\langle k^u \rangle = 2am$.

\[ P^u(k^u) = \frac{(2am)^{k^u} e^{-2am}}{k^u!} \]

As a consequence, we will concentrate on analyzing the growth dynamics of the PA links for the studied models.

It is particularly interesting to study at this point how uniform random linking affects the evolution of the connected set since this is completely independent on the specific PA rule considered. This feature represents one of the main differences between the studied model and other previous mechanisms used to generate growing networks. That is, in our model nodes are not incorporated to the connected set at a constant rate (like e.g.
in the standard Barabási-Albert model) due to the possibility of adding new nodes from $\mathcal{U}(t)$ by applying uniform linking at time $t$ and therefore the set $\mathcal{U}(t) \neq \mathcal{U}(0) - t$. We can easily derive the evolution of the connected set size, $N(t) = \Omega - \mathcal{U}(t)$, for any value of the parameter $\alpha$. For this, we consider the growth of the connected set at each time step, i.e., when a new node of $\mathcal{U}(0)$ throws its $m$ links

$$N(t+1) = N(t) + \frac{\Omega - N(t)}{\Omega - (t + m_0)} + \alpha m \left(1 - \frac{N(t)}{\Omega}\right). \quad (4)$$

In the above equation the second term on the right accounts for the probability that the new node (which is throwing its $m$ links) of $\mathcal{U}(0)$ does not belong already to the connected set at time $t$ (due to the possible uniform links obtained from previous nodes of $\mathcal{U}(0)$ already connected to the connected set $N(t)$). Besides, the third term on the right describes the probability that any uniform link thrown by the node is directed to a node belonging to $\mathcal{U}(t)$. These two terms account for the growth rate of the connected set. We can consider that both time and $N(t)$ are continuous variables and make the time step small enough in order to obtain the corresponding ODE associated to eq. (4), whose solution is given by

$$N(t) = \Omega + (t + m_0 - \Omega)e^{-\alpha mt/\Omega}. \quad (5)$$

The agreement between this calculation and Monte Carlo simulations is shown in Fig. 2 for different values of $\alpha$ and a preferential attachment as described in what follows (model A). It is worth noting the highly nonlinear behavior of $N(t)$, at variance with models in which its size changes at a constant rate.

We formulate below two different ways to implement the preferential attachment rule, which give rise to different behaviors. In both models we will consider that the PA probability of a node $j$ depends only on the PA links of the node, $k^\text{PA}_j$. This new separation between PA links and uniform ones introduces a higher differentiation between the two simultaneous kinds of link dynamics implemented here allowing us to manipulate (as shown below) the degree of correlation between them. The two models interpolate between scale-free and Erdos-Rényi topologies but the structural transition is quite different (as we will show in section V).

A. MODEL A

In this first model we shall study a preferential attachment rule strongly correlated with the simultaneous uniform random linking. First, we consider that the PA probability of a node $i$ is linear with the incoming PA degree of the node, $k^\text{PA}_i$, that is, those links received by $i$ when other node launches (in average) $(1 - \alpha)m$ links following the PA rule. This particularity of the PA rule was already considered by Dorogovtsev et al. [20]. Besides, we consider that when a node is introduced in the connected component (because either it is chosen at random by any node or it is launching its $m$ outgoing links over the rest of nodes) it has an initial attractiveness (or fitness) $A$. In other words, each node has an associated parameter $A_i$ that is zero if the node $i$ is not in the connected set and is $A_i = A$ if $i$ is linked to other nodes (i.e., it belongs...
to \( N(t) \). We further consider that the attractiveness \( A_i \) enters linearly in the preferential linking probability of node \( i \). With these two ingredients, the expression for \( \Pi^{PA}_j \) is given by

\[
\Pi^{PA}_j = \frac{\hat{k}^{pa}_j + A_j}{\sum_{j \in \Omega} (\hat{k}^{pa}_j + A_j)},
\]

(6)

The introduction of the fitness \( A \) correlates the PA rule with the uniform linking in the sense that the more links are established uniformly (the higher \( A \)), the more new nodes with \( \hat{k}^{pa}_i = 0 \) are incorporated to the connected set from \( U(t) \) and hence (by the presence of \( A \) in the PA probability) the more candidates to obtain PA links are available. This can be observed from the evolution of the connected set \( N(t) \), when \( \alpha \) is high there are a lot of nodes added into \( N(t) \) at the early stage of the network construction so that the potential growth of the PA degree of the former members of the connected set is strongly weakened. In order to confirm these heuristic considerations we derive the mean field evolution for the incoming PA degree of a node \( i \), \( \hat{k}^{pa}_i \)

\[
\frac{d\hat{k}^{pa}_i}{dt} = (1 - \alpha) m \frac{\hat{k}^{pa}_i + A}{(1 - \alpha) m t + A N(t)},
\]

(7)

(with the initial condition \( \hat{k}^{pa}_i(t_0) = 0 \)). Obviously, in the limit \( \alpha = 0 \) we recover the mean field equation for the Generalized Dorogovtsev model [20] (which, when \( A = m \), describes the Barabási-Albert model). For \( \alpha \neq 0 \) the influence of the uniform random linking is evident from the presence of \( N(t) \). The number of nodes that start to have the above dynamics at some time \( t_0 \) is \( dN(t)/dt \) evaluated at time \( t = t_0 \) which for \( \alpha \neq 0 \) is not constant as we have seen in the previous calculation of \( N(t) \). The solution of (7) is then given by

\[
\frac{\hat{k}^{pa}_i(t = \Omega)}{A} = -1 + \exp \left[ (1 - \alpha) m \int_{t_0}^{\Omega} \frac{dt}{(1 - \alpha) m t + A N(t)} \right].
\]

(8)

We have solved numerically eq. (8) in order to obtain \( \hat{k}^{pa}_i(t = \Omega) \) (or \( k^{pa}_i(t = \Omega) = \hat{k}^{pa}_i(t = \Omega) + \alpha m \)) as a function of \( t_0 \). This function, along with the number of nodes which are incorporated to the connected set at time \( t_0 \), gives the degree distribution of the PA links. We have compared the results given by eq. (8) for different values of \( \alpha \) with the corresponding ones obtained by performing Monte Carlo simulations of the model (averaging over \( 10^5 \) networks for each value of \( \alpha \)). The results, plotted in Fig. 9 show a very good agreement for the mean field model and the numerical network construction. As expected, the sooner a node is incorporated to the connected set the higher its final PA degree. However, as discussed above, one can observe that this gain of the oldest nodes becomes less important when the value of \( \alpha \) grows due to the combination of two effects: (i) the application of the PA rule becomes less frequent and (ii) the fast growth of the connected set tends to make more homogeneous the PA probability of the nodes.

B. MODEL B

In the second proposal the two different linking processes are completely independent. For this, we consider that \( \Pi^{PA}_j \) is a linear function of the (incoming and outgoing) links that appear as a product of the application of the PA rule. Then, \( k^{pa}_i \) will be zero until it launches its \( \alpha m \) PA links over the rest of the nodes, i.e. regardless of \( k^{pa}_i \). Then, the mean field equation for the evolution of \( k^{pa}_i \) is given by

\[
\frac{dk^{pa}_i}{dt} = (1 - \alpha) m \frac{k^{pa}_i}{2(1 - \alpha) m t + m_0},
\]

(9)

with the initial condition \( k^{pa}_i(t_0) = (1 - \alpha) m \) and \( t_0 \) being the time when node \( i \) launches its \( m \) links. Solving the above equation yields

\[
k^{pa}_i(t) = (1 - \alpha)m \left( \frac{t}{t_0} \right)^{1/2}.
\]

(10)

Because the nodes launch their links at a constant rate (one node per time step), it is easy to obtain the degree distribution \( P(k^{pa}) \)

\[
P(k^{pa}) = 2(1 - \alpha)^2 m^2 (k^{pa})^{-3},
\]

(11)

which is simply a power law distribution with a Barabási-Albert exponent regardless of the value of \( \alpha \). On the other hand, the relative weight of the power law with respect to the Poisson distribution in the total degree distribution \( P(k) \) will be obviously affected by \( \alpha \) (as the prefactor in the above equation suggests).

IV. NETWORK PROPERTIES

In this section we discuss the transition from SF to ER networks in terms of the global topological features of the networks. We have performed Monte Carlo simulations of the two models and compared how the relevant topological measures evolve as a function of \( \alpha \). We are interested in obtaining how the different correlations between the uniform and PA linking rules affect several structural measures. To do this, we have studied the behavior of three magnitudes that behave very different in the two known limiting cases (SF and ER networks), namely: the degree distribution \( P(k) \), the average shortest path length \( \langle L \rangle \) and the second moment of the degree distribution \( \langle k^2 \rangle \).

Degree distribution - The degree distribution evolution is clearly different for the two models. In Fig. 10 we have plotted the degree distribution and the rank-degree relation for both models. The rank-degree relation provides a useful tool for analysing the degree heterogeneity of the networks [3] and thus it is helpful when looking at the transition between ER and SF networks. As can be observed from Figs. 10(a) and 10(c) the correlated model A shows a smooth transition from the power law (\( \alpha = 0 \))
FIG. 4: (Color online). Monte Carlo results for the degree distribution $P(k)$ and rank-degree relation for several values of $\alpha$. (a) and (c) show the results for model A revealing a progressive increase of the tails decaying rate when $\alpha \to 1$. The results for model B ((b) and (d)) show how the decaying rate is not affected by $\alpha$. The networks were generated with the following parameters $\Omega = 10^5$ and $m = m_0 = 3$ ($A = 3$ for model A).

to the Poisson distribution ($\alpha = 1$). The decay of the tails ($k >> 1$) of the degree distribution and the rank-degree relation becomes progressively faster as $\alpha$ grows revealing the decrease of the exponent of $P_{pa}(k_{pa})$ as expected from the results obtained by the analytical insights developed for model A. For model B the transition is completely different as it is shown in Figs. 4(b) and 4(d). In both representations the decaying rate of the tails is independent of $\alpha$ and the transition to the Poisson distribution is much more apparent for low values of $k$. In this sense one can conclude that highly connected nodes persist along the transition of model B while for model A the heterogeneity is progressively lost.

Average shortest path length - The different evolution of the degree distributions observed above suggests to look at how the average shortest path length behaves along the two paths of interpolation. It is well known that the existence of high degree nodes makes the network more compact due to the possibility of finding shortcuts between nodes going through the hubs. Hence, the persistence of highly connected nodes determines the small diameter of the scale-free network. The results obtained are shown in Fig. 5(a). As expected, the average shortest path length as a function of $\alpha$ grows slower for model B because the probability of finding hubs is higher than for the networks generated using model A for the same value of $\alpha$.

Second moment of $P(k)$ - In order to obtain a quantitative measure of the evolution of the degree heterogeneity for the two models it is convenient to measure the second moment of the degree distribution, $\langle k^2 \rangle$. This magnitude diverges (in the thermodynamic limit $\Omega \to \infty$) for scale-free networks with exponents between 2 and 3. So, we expect a decrease of the heterogeneity on the path to ER graphs. As can be observed from Fig. 5(b), model A shows a faster decrease of $\langle k^2 \rangle$ as expected from the study of the degree distribution while for model B the transition is much smoother revealing again the persistence of highly connected nodes along the path to the ER limit.

As for other properties like the clustering coefficient and degree-degree correlations we have checked that they remain unchanged irrespective of the value $\alpha$ and whether model A or B is implemented.
The results clearly manifest the two different transitions of the models regarding the heterogeneity evolution along the interpolating path. The averaged networks had the ER limit. This kind of behavior is precisely what makes different behaviors of dynamical processes that take place on top of them [21, 22, 23, 24, 25, 26]. For instance, epidemic spreading processes show a natural threshold below which the epidemic cannot spread for ER graphs, while this threshold is absent in the thermodynamic limit in SF networks with a diverging second moment [24, 25, 26]. This kind of behavior is precisely what makes scale-free networks so special.

We have implemented a percolation process on top of the networks generated by models A and B. It is aimed at simulating the random failures of a fraction \( f \) of nodes [21, 22]. By computing the size of the giant connected component, one can characterize the percolation transition. Here, however, we are not interested in the transition per se, but on the influence of the topological features unraveled in the previous section on the size of the giant component of the network. As expected, the behavior of the two models is different when the limit of ER graphs is approached for the same values of \( \alpha \).

As usual, we have analyzed the evolution of the size of the giant component of the network after a fraction \( f \) of the nodes (and hence their links) are removed from it. In this sense, it is relevant to study the relation \( G(f)/G_{ER}(f) \), where \( G(f) \) (\( G_{ER}(f) \)) is the size of the giant component of the network for \( \alpha \neq 1 \) (\( \alpha = 1 \)), since this will clearly unravel the two different approaches from the scale-free limit to the Erdös-Rényi network. As can be observed from Fig. 6 the differences between the SF and ER networks are relevant when \( f \) is high due to the different critical behaviors near the transition point [21, 22]. More interesting for our concerns is the evolution of the magnitude \( G(f)/G_{ER}(f) \) as a function of \( \alpha \). For high values of \( f \) we can clearly appreciate the fast approach from SF to ER supplied by model A (upper plot of Fig. 6). In fact, the size of the giant component is very similar to that corresponding to the ER graph for \( \alpha > 0.7 \). In contrast, the transition exhibited for the same values of \( f \) when model B is implemented reveals again a slower transition between SF and ER networks (Fig. 6 down). In particular one can observe that the...
values of $G(f)/G_{ER}(f)$ for $\alpha = 0.9$ are similar to those for $\alpha = 0.5$ in model A.

The two different behaviors observed can be explained in terms of the topological measures showed in the previous section, in particular with the behavior of the second moment of the degree distribution $\langle k^2 \rangle$ which is known to play a key role for understanding the different transitions in SF and ER networks. It is known that the critical value of $f$ relates to the moments of the distribution as $1 - f_c = 1/\kappa_0$, where $\kappa_0 = (\langle k^2 \rangle/\langle k \rangle) - 1$. From Fig. 5(b), one can see that for model A, the second moment of the degree distribution approaches the value obtained in ER limit for intermediate values of $\alpha$. This is no anymore the case for model B, in which the ratio $(\langle k^2 \rangle/\langle k^2 \rangle_{ER})$ approaches 1 only for values close to $\alpha = 1$. Given that the average connectivity $\langle k \rangle$ is the same for all values of $\alpha$ in both models, we should expect that the behavior of a percolation process near the critical point would be nearly the same to that of ER networks when $(\langle k^2 \rangle/\langle k^2 \rangle_{ER}) \rightarrow 1$. This is the situation for a broader range of $\alpha$ values in model A as suggested by Fig. 5(b). In other words, the different behaviors for the two networks families observed in Fig. 6 can be explained in terms of the speed at which $\langle k^2 \rangle$ convergences to its corresponding value in the ER limit.

The implemented percolation dynamics serves us to illustrate how the interpolating model presented in this work can be a useful tool for discovering what topological measures are fundamental to explain the different behaviors observed when more complex dynamics are studied. Moreover, real networks are not purely scale-free nor completely ER, so that their behavior in front of percolation like processes lies between these extreme cases as clearly shown in Fig. 6.

VI. CONCLUSIONS

In this paper, we have analyzed a model that interpolates between Erdos-Rényi and scale-free networks. The combination of uniform and preferential linking allows us to explore the whole path between the two limiting cases.

Acknowledgments

J. G.-G. acknowledge financial support of the MEC through a FPU grant. Y. M. is supported by MEC through the Ramón y Cajal Program. This work has been partially supported by the Spanish DGICYT Projects FIS2004-05073-C04-01 and FIS2005-00337.
We recall on the possibility that the connected set is temporarily composed by more than one connected component. This is the case for the initial stage of the network growth when $\alpha \to 1$. However, when $\Omega$ is high enough the final network is composed by a unique connected component.

[19] A. 303, 337 (2002).

[20] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, Phys. Rev. Lett. 85, 4633 (2000).

[21] D. S. Callaway, M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Phys. Rev. Lett. 85, 5468 (2000).

[22] R. Cohen, K. Erez, D. ben Avraham, and S. Havlin, Phys. Rev. Lett. 85, 4626 (2000)

[23] R. Cohen, K. Erez, D. ben Avraham, and S. Havlin, Phys. Rev. Lett. 86, 3682 (2001).

[24] R. Pastor-Satorras, and A. Vespignani. Phys. Rev. Lett. 86, 3200 (2001).

[25] R. Pastor-Satorras, and A. Vespignani. Phys. Rev. E 63, 066117 (2001).

[26] Y. Moreno, R. Pastor-Satorras, and A. Vespignani, Eur. Phys. J. B 26, 521 (2002).

[27] G. Grinstein, and R. Linsker, Proc. Nat. Acad. Sci. 102, 9948 (2005).