Features and heterogeneities in growing network models

Luca Ferretti,[1,*] Michele Cortelezzi,[2] Bin Yang,[1,3,4] Giacomo Marmorini,[5,6] and Ginestra Bianconi[7]

[1]*Centre de Recerca en AgriGenòmica, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain*
[2]*Dipartimento di Fisica, Università di Pisa, Largo Bruno Pontecorvo 3, 56127 Pisa, Italy*
[3]*Department of Food and Animal Science, Veterinary School, Universitat Autònoma de Barcelona, Bellaterra 08193, Spain*
[4]*Key Laboratory for Animal Biotechnology of Jiangxi Province and the Ministry of Agriculture of China,*
Jiangxi Agricultural University, Nanchang 330045, China
[5]*Department of Physics, Tokyo University of Science, Tokyo 162-8601, Japan*
[6]*Research and Education Center for Natural Sciences, Keio University, Kanagawa 223-8521, Japan*
[7]*Department of Physics, Northeastern University, Boston, Massachusetts 02115, USA*

Many complex networks from the World Wide Web to biological networks grow taking into account the heterogeneous features of the nodes. The feature of a node might be a discrete quantity such as a classification of a URL document such as personal page, thematic website, news, blog, search engine, social network, etc., or the classification of a gene in a functional module. Moreover the feature of a node can be a continuous variable such as the position of a node in the embedding space. In order to account for these properties, in this paper we provide a generalization of growing network models with preferential attachment that includes the effect of heterogeneous features of the nodes. The main effect of heterogeneity is the emergence of an "effective fitness" for each class of nodes, determining the rate at which nodes acquire new links. The degree distribution exhibits a multiscaling behavior analogous to the the fitness model. This property is robust with respect to variations in the model, as long as links are assigned through effective preferential attachment. Beyond the degree distribution, in this paper we give a full characterization of the other relevant properties of the model. We evaluate the clustering coefficient and show that it disappears for large network size, a property shared with the Barabási-Albert model. Negative degree correlations are also present in this class of models, along with nontrivial mixing patterns among features. We therefore conclude that both small clustering coefficients and disassortative mixing are outcomes of the preferential attachment mechanism in general growing networks.

I. INTRODUCTION

In the last 10 years statistical mechanics has made great advances [1–6] in the understanding of the dynamics and the characteristic structural properties of complex networks. These findings shed light on the universal organization principles beyond a large variety of biological, social, communication, and technological systems. Recently, much attention has been addressed to spatial networks [7] in which the links are determined by the "similarity" or proximity of the nodes in the physical or hidden space in which the networks are embedded. Spatial networks [7] are found in communication [8], transportation [9,10], and even social networks [11]. The role of space in complex networks significantly affects the dynamical properties of the graphs changing their navigability properties [8,12], the critical behavior of the Ising model [13], or the epidemic spreading [7].

However, the similarity between two nodes might also correspond to a modular organization of the network [14]. The general networks models that consider a modular structure are block models [15] and multifractal models [16], that are intrinsically static models of modular networks. Nevertheless several networks are simultaneously growing and developing a modular structure. For example, the World Wide Web contains pages falling into different classes with widely different features (personal pages, thematic websites, news, blogs,

search engines, or social networks), and links between any two pages are influenced by their qualities as well as by the specific classes to which they belong.

Moreover many molecular networks, for example, protein-protein interaction networks, transcription networks, and coexpression networks, are scale-free [17] and growing but have in addition a relevant modular structure. In coexpression networks there is a good correspondence between network modules and biological functions as characterized by gene ontology or pathway enrichment analyses [18]. Connectivity between and within modules depends on tissue and species [19] and should be taken into account in realistic models of coexpression networks.

In light of these results it is necessary to understand how similarity, spatial embedding, modular structure, or other features might change the nowadays classic description of growing networks following preferential attachment [20]. In this mechanism, new nodes are added at a constant rate and connected to existing nodes of the network. The probability $\Pi(i)$ that a new node connects to a node i is proportional to the degree k_i of the node i, i.e., $\Pi(i) \propto k_i$. The preferential attachment mechanism has been directly measured for many complex networks and remains a successful explanation for their scale-free degree distribution. Nevertheless, pure preferential attachment, as implemented in the Barabási-Albert (BA) model [20], has drawbacks: For example, it generates a clustering coefficient that is small compared with the ones observed in real networks and follows a "first-mover-advantage"

_____
*luca.ferretti@gmail.com

mechanisms to the extent that older nodes are systematically associated with larger degree. Several rules for network growth giving rise to an effective preferential attachment mechanism have been studied to overcome these limitations (see reviews in Refs. [1,21,22]).

Shortly after the seminal paper by Barabási and Albert [20], it was recognized that heterogeneity between nodes is an important ingredient for more realistic models, destroying the age-degree correlation present in the BA model. The first proposal in this direction was the addition of node quality or "fitness" to the BA model [23,24]. This model by Bianconi and Barabási paved the way for the study of a wider class of models with other node features such as position in space [25–29]. Some properties of preferential attachment networks on metric spaces, such as the degree and link length distributions, were derived analytically by two of the authors [30]. The result is that both fitness [23] and space [30] give rise to networks with multiscaling in the degree distribution (i.e., a sum of power laws).

In this paper we provide a general analysis of growing networks with both preferential attachment and features. The preferential attachment probability from a new node to node i is proportional to the degree k_i of the node multiplied by a generic positive function of the features h of the new node and h_i of the node i:

$$\Pi(i) \propto \sigma(h_i,h)k_i. \tag{1}$$

The features can have different interpretations: They can represent spatial embedding of the network or discrete features of the nodes defining a modular structure of the network, or they can represent fitness describing the higher ability of some nodes to acquire new links. The space of feature and the connection function $\sigma(h_i,h)$ are fixed and do not coevolve with the network.

We give a full account of the model determining the degree distribution, clustering, and assortativity. We derive the general expression for the asymptotic degree distribution in the rate equation approach. The form of this distribution is a convolution of power laws depending on an "effective fitness" of the node, which is determined by the similarity matrix through a self-consistent equation. We also discuss the conditions under which the rate equation approach breaks down and the fate of the network in these cases.

We derive the general expression for the asymptotic clustering coefficient, showing that clustering always decreases as an inverse power of the network size and therefore disappears in the thermodynamic limit. Assortativity in node degree and features are also studied in these networks, both numerically and analytically. Node degree correlations are negative as in the BA model: Disassortativity increases with the heterogeneity of the nodes and decreases slowly with the network size.

Finally we allow for several variations on the model, including addition and rewiring of links, and show that the results are robust with respect to these variations as long as the connections are assigned through preferential attachment.

II. DEGREE DISTRIBUTION OF GROWING NETWORKS WITH FEATURES

A. The model

We present a general class of models for networks with preferential attachment and features. In these models each node has a feature h randomly chosen from a set S with probability $p(h)$. We call h_i and k_i the feature and degree of the ith node. At each time step a node with m links is added to the network. These links are connected to existing nodes with probability

$$\Pi(i) = \frac{\sigma(h_i,h)k_i}{\sum_j \sigma(h_j,h)k_j}, \tag{2}$$

where h corresponds to the feature of the new node and $\sigma(h',h)$ is a (positive) connection function from h to h'. Such a model is therefore completely defined by the functions $p(h)$ and $\sigma(h',h)$.

In modular or community models, denoting by N_c the number of communities, h is an integer number in $1, \dots, N_c$ while $p(1), \dots, p(N_c)$ denote the relative sizes of the different communities. In the fitness model, h is the quality of the nodes, that is, a real positive number. In spatial models, h is the spatial position of the node; for example, in models on a plane, $h = (x,y)$ and $p(x,y)$ is the node density.

B. Degree distribution

We derive the degree distribution through the rate equation approach introduced by Krapivsky *et al.* [22], modified as in Ref. [30]. The equation for $N_k(h)$, which is the average number of nodes with feature h and degree k, is

$$N_k(h,t+1) = N_k(h,t) + \delta_{k,m}p(h) - m \cdot \sum_{l \in S}$$

$$\times \frac{\sigma(h,l)[kN_k(h,t) - (k-1)N_{k-1}(h,t)]}{\sum_{j \in S} \sigma(j,l) \sum_{k'=m}^{\infty} k'N_{k'}(j,t)} p(l), \tag{3}$$

where the $\delta_{k,m}$ term in the left-hand side accounts for the birth of new nodes. If the features are continuous, the number of nodes is substituted with the number density in the equation above, and the sums over the features with integrals. More generally, if S is not a finite set, $p(h)$ denotes a probability measure over S, and the sum $\sum_{h \in S} p(h)f(h)$ should be read as $\int_S p(h)f(h)$, while $N_k(h,t)$ for fixed t is a finite measure over S.

We assume a linear scaling with time for the quantity

$$\sum_{l \in S} \sigma(l,h) \sum_{k'=m}^{\infty} k'N_{k'}(l,t) = mC(h)t + o(t), \tag{4}$$

where we neglect finite-size corrections contained in the $o(t)$ term, which depend on the initial nodes of the network [31]. Then we can define $n_k(h,t) = N_k(h,t)/t$ and rewrite it as

$$n_k(h,t+1)\left(1 + \frac{1}{t}\right) = n_k(h,t) + \frac{\delta_{k,m}p(h)}{t} - \frac{q(h)}{t}[kn_k(h,t) - (k-1)n_{k-1}(h,t)], \tag{5}$$

where $q(h)$ plays the same role as the (average) fitness of the node [23,24] and is defined as

$$q(h) = \sum_{l \in S} \frac{\sigma(h,l)}{C(l)} p(l), \qquad (6)$$

where $C(h)$ is determined by solving asymptotically Eq. (5), obtaining

$$n_k(h) = \frac{p(h)}{q(h)} \frac{\Gamma[m + q(h)^{-1}]\Gamma(k)}{\Gamma[k + 1 + q(h)^{-1}]\Gamma(m)}$$

$$\simeq \frac{p(h)}{q(h)m} \left(\frac{k}{m}\right)^{-(1+q(h)^{-1})} \qquad (7)$$

and substituting in the definition of $C(h)$ to obtain

$$C(h) = \sum_{l \in S} \sigma(l,h) \frac{p(l)}{1 - q(l)}. \qquad (8)$$

We can join (6) and (8) in a single functional equation for $q(h)$:

$$q(h) = \sum_{l \in S} \frac{\sigma(h,l)p(l)}{\sum_{j \in S} \sigma(j,l)\frac{p(j)}{1-q(j)}}. \qquad (9)$$

From the point of view of the degree distribution, this class of models is equivalent to the fitness model of Bianconi and Barabási [23,24], but in this case the fitness distribution is determined by $p(h)$ and $\sigma(h,l)$ through Eq. (9). The resulting degree distribution is

$$n_k = \sum_{h \in S} n_k(h) \simeq \sum_{h \in S} \frac{p(h)}{q(h)m} \left(\frac{k}{m}\right)^{-(1+q(h)^{-1})} \qquad (10)$$

so the distribution is a sum of power laws similarly to the fitness model, and for a regular distribution of h and $q(h)$ it typically reduces to a power law with logarithmic corrections [23].

Actually the fitness model itself is a particular example of such a model with a feature $h = \eta \in [0,1]$ distributed as $\rho(\eta)$ and a connection probability $\sigma(h,l) = h = \eta$. Then Eq. (9) reduces to

$$q(\eta) = \int_0^1 dy \, \frac{\eta}{\int_0^1 dz \, \frac{\rho(z)}{1-q(z)}z} \rho(y) = \frac{\eta}{\int_0^1 dz \, \frac{\rho(z)}{1-q(z)}z} \equiv \frac{\eta}{C}, \qquad (11)$$

where C depends on the whole distribution of η and is determined by the usual consistency equation

$$1 = \int_0^1 dz \, \frac{\rho(z)}{C/z - 1}; \qquad (12)$$

therefore the model reduces to the fitness model with $q(\eta) = \eta/C$. Also the spatial network models discussed in Refs. [25–28,30,32] are special cases of the above model, with the position playing the role of feature.

Note that since the sum of all node degrees $\sum_i k_i(t)$ should be equal to twice the total number of links mt, and since its mean is given by $\langle \sum_i k_i(t) \rangle = \sum_h mp(h)t/(1 - q(h))$, $q(h)$ should satisfy an additional identity

$$\sum_{h \in S} \frac{p(h)}{1 - q(h)} = 2. \qquad (13)$$

However, this identity is not new but can be derived from (6) and (8) by substituting the definition of $C(h)$ in the numerator of the identity $1 = \sum_h p(h)C(h)/C(h)$ and rearranging.

Interestingly, many simple models are based on "symmetric" features; that is, all the values of the features are equivalent. In other words, there is a group of bijective transformations T_α from S to itself such that the distribution and connection function are invariant {that is, $p(T_\alpha(h)) = p(h)$ and $\sigma(T_\alpha(h),T_\alpha(h')) = \sigma(h,h')$}, and moreover the action of the group is transitive [that is, for every pair h,h' there is a transformation T_β mapping h in h', $T_\beta(h) = h'$]. We denote these models as homogeneous models. In this case, if all the sums in the above equations are convergent, the symmetry implies $q(T(h)) = q(h)$, and transitivity implies that all q are the same, then from Eq. (13) we obtain immediately $q = 1/2$. This means that all homogeneous models have the same degree distribution $n_k \sim k^{-3}$ of the Barabási-Albert model. We will see some examples in Sec. II C. We can actually extend the argument to a slightly more general condition, following Ref. [33]: If the quantities $\sum_{l \in S} p(l)\sigma(l,h)$ and $\sum_{l \in S} p(l)\sigma(h,l)$ are equal and independent of h, then the degree distribution is the same as the BA model.

For nonhomogeneous models, Eq. (9) often need to be solved numerically. For discrete features, a solution can be obtained by root-finding methods. For continuous features, an effective way to solve this kind of equations was presented in Ref. [30].

C. Examples

1. Bipartite networks

A bipartite network is usually composed by two classes of nodes ($h = 1,2$) with links connecting only nodes of different classes. To build a scale-free bipartite network, we choose $\sigma_{(1,1)} = \sigma_{(2,2)} = 0$. We assume that a node can belong to class 1 or 2 with probabilities p_1, p_2 (with $p_1 + p_2 = 1$). Note that in this model the nonzero terms of the connection function can be redefined as $\sigma_{(1,2)} = \sigma_{(2,1)} = 1$, so the function is actually symmetric.

The qualities of the two classes from Eq. (9) are then

$$q_1 = p_2, \quad q_2 = p_1, \qquad (14)$$

and the degree distribution is given by $n_k = \frac{p_1}{m(1-p_1)} \left(\frac{k}{m}\right)^{-(2-p_1)/(1-p_1)} + \frac{(1-p_1)}{mp_1} \left(\frac{k}{m}\right)^{-(1+p_1)/p_1}$. Note that if $p_1 = p_2 = 1/2$, the model is actually homogeneous and $n_k \sim k^{-3}$ as expected from our general arguments.

Similar models have been applied to the human sexual networks, which have a bipartite heterosexual component [34].

2. A network with asymmetric connection function

Another simple but interesting example can be obtained by assuming two kinds of nodes, "central" and "periferic" (with probabilities p_C and $p_P = 1 - p_C$) such that two periferic nodes are never connected (i.e., the connection function satisfies $\sigma_{(P,P)} = 0$). The other terms of the connection function can be always redefined as $\sigma_{(C,C)} = \sigma_{(C,P)} = 1$, so its only parameter is $\bar{\sigma} \equiv \sigma_{(P,C)} \in [0,+\infty]$. This parameter controls the asymmetry of the connection function.

The relevant solution of Eq. (9) with this connection function is

$$q_C = 1 - \frac{2(1 - \bar{\sigma})p_C}{1 + p_C - 3\bar{\sigma} + \sqrt{(1 + p_C - 3\bar{\sigma})^2 + 8\bar{\sigma}(1 - \bar{\sigma})}}, \tag{15}$$

$$q_P = \bar{\sigma} \cdot \frac{p_C - 1 - \bar{\sigma} + \sqrt{(1 + p_C - 3\bar{\sigma})^2 + 8\bar{\sigma}(1 - \bar{\sigma})}}{1 + p_C - 3\bar{\sigma} + \sqrt{(1 + p_C - 3\bar{\sigma})^2 + 8\bar{\sigma}(1 - \bar{\sigma})}}. \tag{16}$$

In the limit $\bar{\sigma} \to \infty$, the model reduces to a bipartite network and the solution to $q_C = 1 - p_C$, $q_P = p_C$ as expected.

3. Community structure

To model scale-free networks with community structure [35], a simple possibility is to label each community by a feature h_c and choose a connection function with $\sigma(h_c, h_c) > \sigma(h_c, h_{c'})$. In the simplest model, all communities have the same size and connect randomly to the other communities, with some preference for self-connections. [The corresponding connection function is $\sigma(h_c, h_c) = 1$ for all h_c, while $\sigma(h_c, h_{c'}) = \bar{\sigma}$ for all pairs $h_c \neq h_{c'}$.] In this case the model is actually homogeneous and has the same degree distribution as the BA model, independently of the number of communities and the value of $\bar{\sigma}$.

4. Hierarchical structure and navigable networks

Navigable networks are often based on a hierarchical structure [36], with a connection probability that depends on the distance on a tree representing the hierarchical levels and the nodes. Scale-free navigable networks can be easily built by choosing a symmetric connection function depending only on the distances on the tree. (Note that these models are actually spatial models, since a tree is an ultrametric space.)

Even if there is a lot of interesting structure in these networks, the degree distribution follows the simple multiscaling behavior in Eq. (10). In particular, the simplest cases of binary or n-ary trees (or more general trees where the length and the number of branches splitting from a single branch depend only on the level), with nodes located at the top of the terminal branches, have the usual degree distribution $n_k \sim k^{-3}$, since these trees are homogeneous spaces.

5. Modular structure with fitness

As a final example, we discuss a model with both modular structure and fitness. This model can be considered a simplified model of the WWW. We assume that each page is assigned to some category n according to type, content, and functionality. Each category could have a different relative size π_n and a distribution of page fitness $\rho_n(\eta)$. Moreover, the relative importance of different categories for a page of category q is given by the weights $w_{n,q}$, with $\sum_q w_{n,q} = 1$. (For a page of category n, the weights $w_{n,q}$ affect its probability of being linked by pages from other categories.) The network evolves as follows: At each time, a new node is added to the network and assigned to the nth category with probability π_n, then its fitness is randomly extracted from $\rho_n(\eta)$. The node is connected to the existing nodes according to a probability proportional to
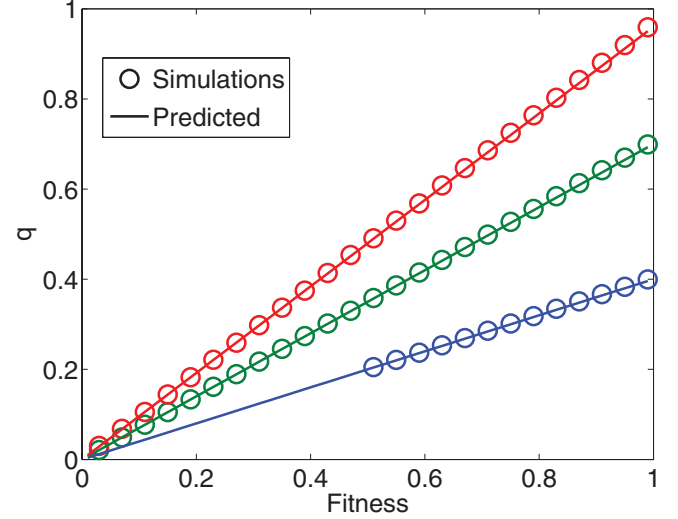

FIG. 1. (Color online) Node qualities as a function of fitness in a model with hierarchical modules A, B, and C and fitness. Continuous lines represent the theoretical predictions. The hierarchy is A < B < C, and nodes in each module can connect only to nodes in the same or higher modules with equal probabilities. The fitness distributions are $\rho_A(\eta) = U_{(1/2,1)}(\eta)$, $\rho_B(\eta) = U_{(0,1)}(\eta)$ and $\rho_C(\eta) = 2\eta$. Simulated networks have size $N = 10^7$ and initial node degree $m = 5$. Qualities are measured as $\hat{q} = \log(\langle k(t) \rangle / \langle k(t/2) \rangle)/\log 2$ as detailed in Ref. [30].

the fitness η, the weight w, and the degree k of the nodes:

$$\Pi(i) = \frac{\eta_i w_{n_i,n} k_i}{\sum_j \eta_j w_{n_j,n} k_j}, \tag{17}$$

so the model dynamics follows Eq. (2) with $h = (n, \eta)$, $\sigma(h_i, h) = \eta_i w_{n_i,n}$, and $p(h) = \pi_n \rho_n(\eta)$. Similar models appear in a natural way in the study of many systems. In this model, beyond the fitness, the dynamics of a node is influenced by the relative size of its category π_n as well as by the weights $w_{n,q}$ and the sizes of other categories π_q. In the WWW example, there are millions of blogs but only a few search engines, and there is a good probability that a blog links a search page; this explains the different connectivity and degree distribution of these categories.

The node qualities for this model from Eq. (9) are equivalent to a modified fitness model $q(\eta, n) = \eta / \gamma_n$, where the coefficients γ_n solve the nonlinear equations

$$\frac{1}{\gamma_n} = \sum_m w_{n,m} \pi_m \left[\sum_l w_{l,m} \pi_l \int d\eta \frac{\rho_l(\eta)}{\eta^{-1} - \gamma_l^{-1}} \right]^{-1}. \tag{18}$$

Simulation results are in very good agreement with the numerical solution of these equations, as shown in Fig. 1.

D. Breakdown of the rate equation approach

If the space S of features has a finite number of elements, the approach presented here gives rise to a finite number of nonlinear Eqs. (6) and (8) in the variables $q(h), C(h)$. If these equations admit a single solution, as we expect, then the degree distribution follows Eq. (10).

If the number of features is infinite, Eqs. (6) and (8) could involve divergent sums (or integrals). In particular, there are

two situations where the rate equations break down: (1) the sums of the connection function are divergent, that is, the connection function is not measurable, or (2) all sums converge but there is no solution to the self-consistency equations. We discuss these two scenarios in the next sections.

1. Heterogeneity-driven attachment

If the sums of the connection function are divergent, links attach preferentially to some nodes not chosen on the basis of preferential attachment but belonging to the sets of features with divergent sums of the connection function. This can give rise to exponential tails or condensation or other behavior, depending on the form of the connection function. As an example, spatial models with a divergent connection function near $d = 0$ show a behavior similar to nearest-neighbor attachment and consequently an exponential tail [25,30]. Another example is given by networks on a flat space of dimension D with uniform node density and a connection function $\sigma(x,x')$ that is an inverse power law in the distance between the node and a point $\hat{x}$, for example, $\sigma(x,x') = d(x,\hat{x})^{-k}$ with $k > D$. In this case the divergence is localized around $\hat{x}$, prompting condensation on the nodes closer to $\hat{x}$ since these nodes get most of the new connections.

2. Bose-Einstein condensation

Even if all the sums of the connection function are convergent, it is still possible to find cases where the equations for $q(h)$ do not admit solutions. In fitness models, the lack of solution of the self-consistent equations is a signal of Bose-Einstein condensation of links [24] on the nodes of highest fitness.

In more general models, condensation occurs on nodes close to a feature h_c determined as follows. Each $h \in S$ corresponds to an element of the set M of measures $\mu_h(h') = \sigma(h,h')p(h')$. Conversely, given a Radon metric on M, each point in its closure $\bar{M}$ can be mapped to a feature belonging to an "extended" space $\bar{S}$. The Radon metric on $\bar{M}$ induces a metric on $\bar{S}$, which can therefore be thought of as the closure of S: The elements of $\bar{S}$ are "limit points" of S and are the candidates for condensation. In particular, the location h_c of the condensate can be found by considering the addition of a single node with variable $\bar{h} \in \bar{S}$ to the network and maximizing its asymptotic link share $n_{\bar{h}}$:

$$h_c = \arg\max_{\bar{h} \in \bar{S}} n_{\bar{h}}. \qquad (19)$$

As a simple example, in the fitness model with $S = [0,\eta_{\max})$, the set M contains the measures $\mu_\eta(\eta') = \eta p(\eta')$, and its closure corresponds to $\bar{S} = [0,\eta_{\max}]$, which is the closure of S. In this case $n_{\bar{\eta}}$ is zero unless $\bar{\eta} = \eta_{\max}$, because the consistency equation with an additional node has no finite solution for other values of $\bar{\eta}$, and therefore condensation occurs near $\eta_{\max}$ as expected.

III. CLUSTERING AND ASSORTATIVITY

A. Clustering

We define the average clustering coefficient C_{clust} of a network as $C_{\text{clust}} = 3 \cdot n_{\text{triangles}}/n_{\text{triples}}$. It is well known that the clustering coefficient of the Barabási-Albert model decreases as the network size t increases, converging to zero in the thermodynamic limit [37]. As we show in the next sections, this property is quite general, being shared by all heterogenous models under some conditions on the convergence of sums of $\sigma(h,h')$. This mean that heterogeneity or features cannot account for nonvanishing clustering coefficients observed in most real networks.

1. Clustering in the Barabási-Albert and in the homogeneous model

Both the average number of triangles and the average number of triples can be easily computed from the preferential attachment rule if we assume that node degrees follow the continuum equation for the mean degree $k_i(t) = m(t/t_i)^{q(h_i)}$ [38]. This approach has been applied to the BA model in Ref. [39]. Here we generalize it to models with features. First, we summarize the computation for the BA case. The asymptotic number of triangles is given by

$$n_{\text{triangles}} \sim \frac{m^2(m-1)\ln^3 t}{48}. \qquad (20)$$

Denoting by $t_A < t_B < t_C$ the birth times of a triplet of nodes, BA networks contain three kind of triples: $A \leftarrow B \leftarrow C, A \leftarrow C \rightarrow B$, and $B \rightarrow A \leftarrow C$. Since older nodes have the highest degrees and are therefore the most attractive under preferential attachment, for large t almost all triples are of the last kind, and their number is given by

$$n_{\text{triples}} \sim \frac{m^2 t \ln t}{2}, \qquad (21)$$

therefore obtaining the known result for the asymptotic clustering coefficient

$$C_{\text{clust}}^{\text{BA}} \sim \frac{(m-1)}{8}\frac{\ln^2 t}{t}. \qquad (22)$$

Details of the calculations can be found in Appendix A.

For homogeneous models ($q(h) = 1/2$) the same calculation is valid for the number of triples, while the number of triangles and therefore the clustering coefficient are multiplied by a factor dependent on the connection function:

$$\frac{C_{\text{clust}}}{C_{\text{clust}}^{\text{BA}}} = \left\langle \frac{\sigma(h_A,h_B)\sigma(h_A,h_C)\sigma(h_B,h_C)}{C(h_B)C(h_C)^2/8} \right\rangle_{A,B,C}. \qquad (23)$$

Features appear only in this factor. For example, bipartite networks have no triangles and the above factor is zero, giving $C_{\text{clust}} = 0$. Note that the above factor could be divergent: In this case our approach breaks down, and the asymptotic behavior of $C_{\text{clust}}(t)$ could change. However, we are not aware of any model of this kind with nonzero clustering in the limit $t \rightarrow \infty$ and a power-law tail in the degree distribution.

For a symmetric model with community structure, the clustering coefficient is

$$\frac{C_{\text{clust}}^{\text{com}}}{C_{\text{clust}}^{\text{BA}}} = \frac{N_c + 3N_c(N_c-1)\bar{\sigma}^2 + N_c(N_c-1)(N_c-2)\bar{\sigma}^3}{(1 + \bar{\sigma}(N_c-1))^3}. \qquad (24)$$

An additional remark applies to spatial networks with short-range interactions and networks with community structure.
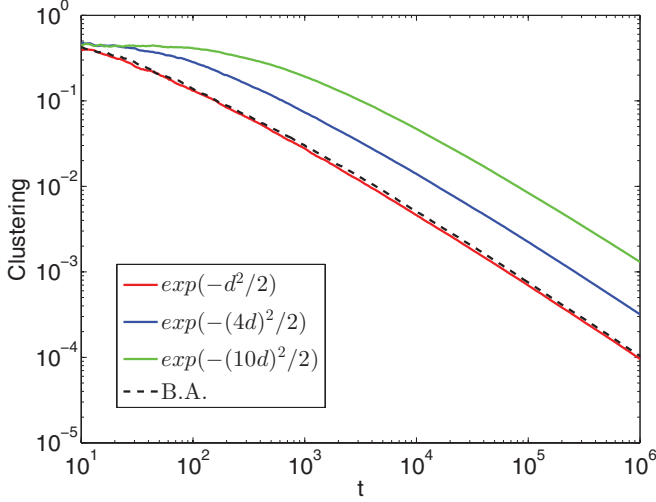
FIG. 2. (Color online) Plot of the clustering coefficient as a function of the size of the network, for networks on the two-dimensional flat disk of radius $r = 1$ with Gaussian connection functions.

At short times, these models resemble geometric random graph models, which exhibit strong clustering [27]. However, when the average number of nodes in the interaction volume increases, the clustering coefficient decreases as a consequence of the preferential attachment dynamics. At longer times, clustering falls according to Eqs. (22) and (23) but remains higher for spatial networks than for the Barabási-Albert model, as is shown in Fig. 2.

2. Clustering in general models

For general models with some $q(h) \neq 1/2$, the calculation is similar to the homogeneous case. The dominant contribution comes from triples and triangles with a vertex of maximum fitness $q_M = q(h_M) > 1/2$ with $h_M = \arg\max_{h \in S} q(h)$. The leading contribution to the average number of triangles is

$$n_{\text{triangles}} \sim \frac{m^2(m-1)p(h_M)}{(2q_M - 1)^2} t^{2q_M - 1} \ln t$$
$$\cdot \sum_{h_B, h_C} \frac{\sigma(h_M, h_B)\sigma(h_M, h_C)\sigma(h_B, h_C)p(h_B)p(h_C)}{C(h_B)C(h_C)^2},$$
(25)

while the number of triples is

$$n_{\text{triples}} \sim \frac{m^2 p(h_M)}{2(2q_M - 1)} t^{2q_M},$$
(26)

and then the clustering coefficient is

$$C_{\text{clust}} \sim \frac{6(m-1)}{2q_M - 1} \frac{\ln t}{t}$$
$$\cdot \sum_{h_B, h_C} \frac{\sigma(h_M, h_B)\sigma(h_M, h_C)\sigma(h_B, h_C)p(h_B)p(h_C)}{C(h_B)C(h_C)^2}.$$
(27)

This expression is valid as long as the sum inside it is finite. So the asymptotic clustering is slightly smaller than in the BA model.

From this general result we can extract the clustering for the fitness model of Bianconi and Barabási [23] in the fit-get-rich phase:

$$C_{\text{clust}}^{\text{BB}} \sim \frac{6(m-1)\langle \eta \rangle}{C^2(2-C)} \frac{\ln t}{t},$$
(28)

assuming $\eta_{\max} = 1$ and $\rho(\eta_{\max}) > 0$.

B. Assortativity: Features

Nodes with different h are not randomly connected: in- and outgoing links connect preferentially nodes with some features. These preferences are embedded in the in- and out-distributions $f^{\text{IN}}(h_i, h')$ and $f^{\text{OUT}}(h_i, h')$, which are asymptotically defined by the equations

$$k_{i(h')}^{\text{IN}} = f^{\text{IN}}(h_i, h')k_i^{\text{IN}},$$
(29)

$$k_{i(h')}^{\text{OUT}} = f^{\text{OUT}}(h_i, h')k_i^{\text{OUT}},$$
(30)

where k_i^{IN} and k_i^{OUT} are the number of in- and outgoing links for the ith node (note that in these models $k_i^{\text{OUT}} = m$), while $k_{i(h')}^{\text{IN}}$ and $k_{i(h')}^{\text{OUT}}$ are the numbers of in- and outgoing links between the ith node and nodes with variable h'. From the definitions above, we have $\sum_{h'} f^{\text{IN}}(h_i, h') = \sum_{h'} f^{\text{OUT}}(h_i, h') = 1$.

The distributions f^{IN} and f^{OUT} are positive. f^{IN} can be obtained from the continuum equations for $k_{i(h')}^{\text{IN}}$,

$$\frac{dk_{i(h')}^{\text{IN}}}{dt} = \frac{\sigma(h_i, h')p(h')(k_i^{\text{IN}} + m)}{C(h')t},$$
(31)

by plugging in Eq. (29) and comparing with the continuum equation for k_i^{IN},

$$\frac{dk_i^{\text{IN}}}{dt} = q(h_i) \frac{k_i^{\text{IN}} + m}{t},$$
(32)

giving as a result

$$f^{\text{IN}}(h_i, h') = \frac{\sigma(h_i, h')p(h')}{C(h')q(h_i)},$$
(33)

while f^{OUT} can be obtained from Eq. (2) by substituting k_i with the mean of the total degree of nodes of feature h':

$$f^{\text{OUT}}(h_i, h') = \frac{\sigma(h', h_i)p(h')}{C(h_i)[1 - q(h')]}.$$
(34)

From these distributions it is easy to find the fraction of links between nodes with variables h', h'':

$$\varphi(h', h'') = \frac{p(h')p(h'')}{1 + \delta_{h',h''}}$$
$$\cdot \left\{ \frac{\sigma(h', h'')}{C(h'')[1 - q(h')]} + \frac{\sigma(h'', h')}{C(h')[1 - q(h'')]} \right\}$$
(35)

that should be compared with the null value of the same quantity, obtained by a random rearrangement of links that preserves degree (i.e., unassortative connections):

$$\varphi_0(h', h'') = \frac{p(h')p(h'')}{1 + \delta_{h',h''}} \frac{1}{2[1 - q(h')][1 - q(h'')]}.$$
(36)

As an example, consider the simplest model with community structure in Sec. II C3. Denote the number of communities by N_c and the only parameter of the connection function by

$\bar{\sigma} = \sigma(h,h')/\sigma(h,h)$ for $h' \neq h$. The null distribution of links is given by $\varphi_0(h',h') = 1/N_c^2$, $\varphi_0(h',h'') = 2/N_c^2$ for $h'' \neq h'$, while the actual assortativity depends on the parameter $\bar{\sigma}$:

$$\varphi(h',h'') = \begin{cases} \frac{1}{N_c^2} \cdot \left[1 + (\bar{\sigma} - 1)\frac{N_c-1}{N_c} \right]^{-1} & \text{if} \quad h' = h'' \\ \frac{2}{N_c^2} \cdot \left[1 + \left(\frac{1}{\bar{\sigma}} - 1\right)\frac{1}{N_c} \right]^{-1} & \text{if} \quad h' \neq h'' \end{cases} \tag{37}$$

so links are randomly distributed between features if $\bar{\sigma} = 1$, while the mixing is assortative [that is, $\varphi(h,h)/\varphi_0(h,h) > \varphi(h',h'')/\varphi_0(h',h'')$] for $\bar{\sigma} < 1$ and disassortative for $\bar{\sigma} > 1$.

For the Bianconi-Barabási fitness model, the distributions are

$$\varphi_0(\eta,\eta') = \frac{\rho(\eta)\rho(\eta')}{2\left(1 - \frac{\eta}{C}\right)\left(1 - \frac{\eta'}{C}\right)}, \tag{38}$$

$$\frac{\varphi(\eta,\eta')}{\varphi_0(\eta,\eta')} = 2\left[\frac{\eta}{C}\left(1 - \frac{\eta'}{C}\right) + \frac{\eta'}{C}\left(1 - \frac{\eta}{C}\right) \right], \tag{39}$$

and therefore the system shows disassortative mixing, since the ratio $\varphi(\eta,\eta')/\varphi_0(\eta,\eta')$ is higher between high and low fitness than between similar fitnesses.

C. Assortativity: Degree correlations

Nontrivial connection functions $\sigma(h,h')$ and node distributions $p(h)$ do not only induce assortative mixing between nodes with different variables, but affect also the degree-degree correlations between neighbors, as shown in Ref. [40] for the BA model. The BA model is disassortative [41]; however, in models with features, the pattern of (disassortative) mixing between nodes with different degree is influenced by the hidden space and the connection function.

1. Analytical results: Continuum approximation

A first understanding of degree correlations can be obtained by assuming deterministic evolution of node degree $k = m(t/t_0)^q$. In this approximation, average nearest-neighbor degree is given by

$$\langle k_{\text{NN}}(k)\rangle = \sum_h \frac{n_k(h)}{\sum_j n_k(j)} \sum_i \int_m^\infty dk_{\text{NN}} \, k_{\text{NN}} P(k_{\text{NN}},i|k,h), \tag{40}$$

where $P(k_{\text{NN}},i|k,h)$ is the probability that if we choose a random neighbor of a random node with features h and degree k, the node chosen has degree k_{NN} and features i. If we denote the birth times of these nodes by t_0 and $t_{0\text{NN}}$, respectively, then if $t_{0\text{NN}} < t_0$ [i.e., $(k_{\text{NN}}/m)^{1/q(i)} > (k/m)^{1/q(h)}$] this probability can be obtained from the connection probability $m\frac{\sigma(i,h)k_{\text{NN}}(t_0)}{mC(h)t_0}$ multiplied by $1/k$ (which accounts for the random neighbor chosen), $p(i)$ (the probability that a node of feature i is born at time $t_{0\text{NN}}$) and $dt_{0\text{NN}}/dk_{\text{NN}}$ (the Jacobian of the mapping from k_{NN} to $t_{0\text{NN}}$). The final result is given by $\frac{dt_{0\text{NN}}}{dk_{\text{NN}}} \frac{p(i)\sigma(i,h)k_{\text{NN}}(t_0)}{kC(h)t_0}$. In the other case, if $t_{0NN} > t_0$, the relevant connection probability is $m\frac{\sigma(h,i)k(t_{0\text{NN}})}{mC(i)t_{0\text{NN}}}$. Substituting the present values of k and k_{NN},
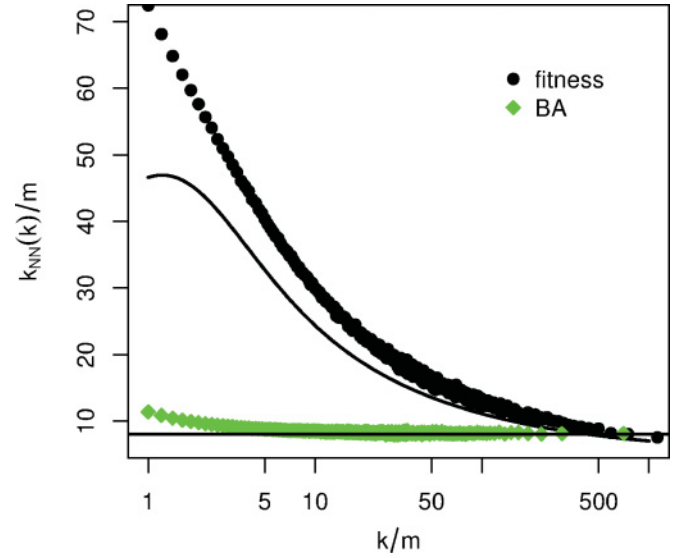

FIG. 3. (Color online) Plot of $\langle k_{\text{NN}}(k)\rangle$ for the BA model and for a fitness model with $\rho(\eta) = \frac{1}{2}\delta(\eta - 1/2) + \frac{1}{2}\delta(\eta - 1)$. Results are averaged from 20 simulations of size $N = 10^7$ and initial degree $m = 5$. Continuous lines correspond to the predictions from continuum equation (41).

we obtain

$$P(k_{\text{NN}},i|k,h) = \begin{cases} \frac{\sigma(i,h)p(i)}{mq(i)C(h)}(k/m)^{\frac{1-q(i)-q(h)}{q(h)}}(k_{\text{NN}}/m)^{-1/q(i)} \\ \qquad \text{for} \quad k_{\text{NN}} > m\left(\frac{k}{m}\right)^{q(i)/q(h)}; \\ \frac{\sigma(h,i)p(i)}{mq(i)C(i)}(k_{\text{NN}}/m)^{-1-q(h)/q(i)} \\ \qquad \text{for} \quad k_{\text{NN}} < m\left(\frac{k}{m}\right)^{q(i)/q(h)}. \end{cases} \tag{41}$$

For the BA model, we have $P(k_{\text{NN}}|k) = mk_{NN}^{-2}$, and the average nearest-neighbor degree $\langle k_{\text{NN}}(k)\rangle = m\log(t)/2$ is independent of k, so the model shows no assortativity at all in this approximation. On the other hand, models with features tend to be disassortative, i.e., $\langle k_{\text{NN}}(k)\rangle$ decreases with k. However, this is not true for all classes of nodes: for example, low quality nodes tend to be assortative; i.e., they attach more often to low-quality nodes with similar degree.

In Fig. 3 we compare the above results (41) with the actual values of $\langle k_{\text{NN}}(k)\rangle$ in the BA model and in a simple fitness model. The continuum approach describes correctly the qualitative pattern of degree correlations in these models, even if it does not fully account for their disassortativity.

It is also possible to estimate the asymptotic behavior of the assortativity coefficient in the same approximation. We denote the maximum quality by $q_M = \max_h q(h)$. As in the BA model, the coefficient C_{ass} tends to zero in the thermodynamic limit $t \to \infty$. However, its asymptotic behavior is $C_{\text{ass}} \sim t^{q_M-1}$, and since $q_M \simeq 1$ in many models with features, its decrease is typically very slow compared to the BA model ($C_{\text{ass}} \sim t^{-1/2}$).

2. Analytical results: Rate equation approach

Assortativity in node degree can be easily estimated from the matrix $n_{k,l}$, defined as the fraction of links in the network joining a node of degree k with a younger node of degree l. In models with features, we define $n_{k,l}^{(h,i)}$ as the fraction

of links joining a node of degree k and variable h with a younger node of degree l and variable i. The matrix $n_{k,l}^{(h,i)}$ takes into account correlations between both degrees and features. Asymptotically, this matrix can be obtained as in Ref. [40], by solving the difference equation

$$
\begin{aligned}
n_{k,l}^{(h,i)}&[1 + q(h)k + q(i)l] \\
&= q(h)(k-1)n_{k-1,l}^{(h,i)} + q(i)(l-1)n_{k,l-1}^{(h,i)} \\
&\quad + \delta_{l,m}\frac{\sigma(h,i)p(i)}{C(i)}(k-1)n_{k-1}(h).
\end{aligned}
\tag{42}
$$

This equation implies that $n_{k,l}^{(h,i)} \propto F[q(h),q(i)] \cdot \sigma(h,i)p(i)p(h)/C(i)$ where $F[q(h),q(i)]$ is a function of the qualities, in agreement with the form (35) for the assortative mixing between features. The matrix $n_{k,l}$ can then be obtained as $n_{k,l} = \sum_{h,i} n_{k,l}^{(h,i)}$.

In principle, Eq. (42) can be solved by generating function methods discussed in Appendix B. The general solution for $m = 1$ is

$$
\begin{aligned}
n_{k,l}^{(h,i)} &= \frac{\sigma(h,i)p(i)}{C(i)q(h)}\sum_{c=2}^{\infty}(c-1)n_{c-1}(h) \\
&\quad \cdot \sum_{j=0}^{l-1}(-1)^j\frac{1}{\alpha_j}\binom{l-1}{j}\sum_{n=0}^{k-c}(-1)^n \\
&\quad \cdot \frac{\Gamma(\alpha_j+1+k-c-n)}{(\alpha_j+k-n)\Gamma(\alpha_j-n)\Gamma(n+1)\Gamma(k-c-n+1)},
\end{aligned}
\tag{43}
$$

where $\alpha_j = \frac{1+(j+1)q(i)}{q(h)}$. However, extracting information from this solution is difficult.

The existence of degree correlations can be also shown simply by looking at the scaling of the quantity $n_{k,l}^{(h,i)}$ for $k \gg l$ and $k \ll l$. In the BA model with $m = 1$, the scaling is $n_{k,l} \sim k^{-2}l^{-2}$ and kl^{-5}, respectively [40], which is different from the naive $k^{-2}l^{-2}$ expected in the absence of degree correlations.

The scaling $k^{-2}l^{-2}$ for $k \gg l$ can be understood from the following simple arguments. The degree of young nodes is dominated by outgoing connections, which select random nodes with probability proportional to $kn_k \sim k^{-2}$. On the other hand, the attractiveness of old nodes decays with time as $t^{-1/2}$, and therefore the distribution of the linked nodes [assuming a deterministic evolution $\bar{k} = l(t) \sim t^{-1/2}$ as a function of the birth time t] is $dt(l)/\sqrt{t(l)} = dl/l^2$.

To obtain the scaling for $n_{k,l}^{(h,i)}$ in these models with $m = 1$, we approximate the difference equation with the corresponding differential equation $n_{k,l}^{(h,i)} = q(h)\partial(kn_{k,l}^{(h,i)})/\partial k + q(i)\partial(ln_{k,l}^{(h,i)})/\partial l$ and match the solution with the exact boundary conditions $n_{k,1}^{(h,i)}$ (for $k \gg l$) and $n_{2,l}^{(h,i)}$ (for $l \gg k$). The computation is outlined in Appendix B. The result is

$$
n_{k,l}^{(h,i)} \sim \begin{cases} k^{-1/q(h)}l^{-1-q(h)/q(i)} & \text{for}\quad k \gg l \\ kl^{-1-(1+2q(h))/q(i)} & \text{for}\quad k \ll l \end{cases},
\tag{44}
$$

which is quite different from the null scaling $k^{-1/q(h)}l^{-1/q(i)}$. This shows that degree correlations exist also in heterogeneous model. Note that the scaling for $k \gg l$ is consistent with the continuum approximation (41). The same scaling for $k \gg l$

and l fixed can also be found directly from the generating function (B3) using Tauberian theorems. For $m > 1$, the scaling for $k \gg l$ is the same as for $m = 1$, while the scaling for $l \gg k$ changes to $n_{k,l}^{(h,i)} \sim k^m l^{-1-(1+(m+1)q(h))/q(i)}$.

The BA model is slightly disassortative in degree, and the addition of features does not change this property. This is already apparent from the numerical results in the previous sections but can be also understood from the above scaling properties. In fact, in the BA model the asymptotic ratio between $n_{k,l}$ and its null value scales between 1 for $k \gg l$ and $(k/l)^3 \sim 0$ for $k \ll l$, therefore decreasing while k and l get closer. In homogeneous models ($q = 1/2$), like the symmetric communities model, the scaling is the same as in the BA model. On the other side, in models with different qualities (like the Bianconi-Barabási fitness model) the pattern of degree correlations is nontrivial, as already observed in the previous section: For example, old and well-connected nodes with features h are preferentially linked to younger nodes of features i and similar degree if qualities are low [$q(i) + q(h) < 1$], but they link instead to younger nodes of low degree if qualities are high [$q(i) + q(h) > 1$].

3. Numerical results: Spatial networks

As an example, we consider spatial networks with preferential attachment [30]. In these networks S corresponds to a metric space, and the connection function depends only on the distance $\sigma(x,y) = \tilde{\sigma}[d(x,y)]$. We simulated network growth on disks in two-dimensional spaces with constant curvature (sphere, flat, and hyperbolic space) and calculated both the assortativity coefficient C_{ass} [41] and the average nearest-neighbor degree $\langle k_{\text{NN}}(k)\rangle$. We show the results in Table I and Figs. 4, 5, and 6.

It is apparent that the short-range connections and the spatial structure make the network slightly more disassortative, because hubs tend to be sparse (close hubs compete between them and reduce their degree). Hyperbolic spaces at strong
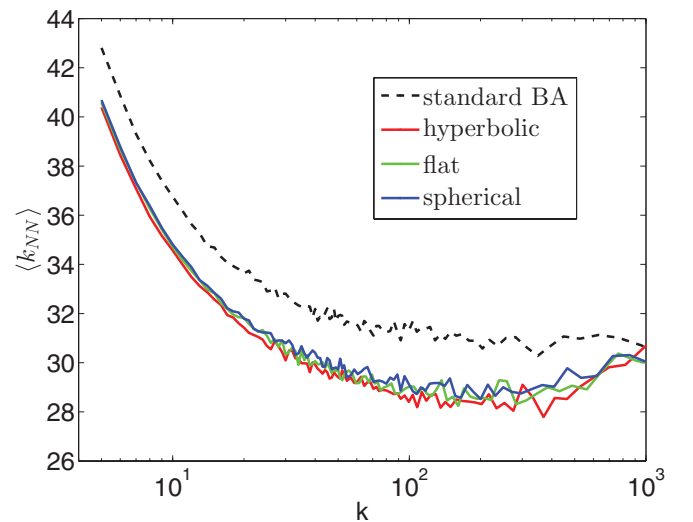
FIG. 4. (Color online) Plot of average neighbor degree $\langle k_{NN}(k)\rangle$ (averaged over 50 simulations) for networks of size 2×10^5 on disks of radius $r = 1.5$ in two-dimensional spaces of curvature $+1$, 0, -1 with connection function $\tilde{\sigma}(d) = e^{-16d/r}$.

TABLE I. Values of the assortativity coefficient C_{ass} (averaged over 50 simulations) for spatial networks of size 2×10^5 on disks of radius $r = 1$ in different two-dimensional spaces and for different connection functions $\tilde{\sigma}(d)$. (Note that the only parameter of the space is actually the product of the curvature and the squared radius.) The corresponding assortativity coefficient for the Barabási-Albert model is $C_{ass}^{BA} = -0.0112$.

Space Curvature	Spherical		Flat	Hyperbolic		
	2.25	1	0	−1	−2.25	−16
$\tilde{\sigma} = e^{-16d/r}$	−0.0139	−0.0142	−0.0145	−0.0147	−0.0151	−0.0187
$\tilde{\sigma} = e^{-4d/r}$	−0.0144	−0.0151	−0.0155	−0.0162	−0.0169	−0.0136
$\tilde{\sigma} = \theta\left(1 - \frac{d}{r}\right)$	−0.0160	−0.0158	−0.0152	−0.0149	−0.0149	−0.0909

curvature show even higher disassortativity because of their almost starlike connectivity.

Also, the generally low values of the assortativity coefficients suggest that they decrease with time as happens in the BA model.

IV. GENERALIZATIONS

The Barabási-Albert model is based purely on addition of nodes and preferential attachment. Realistic models can include many other ingredients: addition of extra links, rewiring and removal of links, directed links, variable initial degree, attachment functions that are only asymptotically linear in k, etc. Some generalizations of the BA model are reviewed in Refs. [1,21].

In this section we present several variations on the heterogeneous models with preferential attachment presented in Sec. II A. For most of these generalizations, the degree distribution is a sum of power laws, showing that scale-free or multiscaling behavior is a robust feature of preferential attachment models. The results are generally similar when different generalizations are combined.

The consistency equations for $q(h)$ in these models can be obtained through a rate equation approach or, more easily, by using the continuum approach, i.e., the deterministic

evolution of node degree $dk_i/dt = q(h_i)k_i/t$ or equivalently $k_i(t) = k_i(t_0)(t/t_0)^{q(h_i)}$ [38]. The corresponding equations for $q(h)$ are exact, as explained in Appendix C.

A. Heterogeneity in initial degree

We consider a model of growing networks with the usual preferential attachment rule (2). However, new nodes have a initial degree $m(h)$ that depends on their feature h. The continuum equation for the node degree is

$$\frac{dk_i}{dt} = \sum_h p(h)m(h) \frac{\sigma(h_i,h)k_i}{\sum_l \sigma(h_l,h)k_l}. \quad (45)$$

The degree distribution is given by equation

$$n_k = \sum_{h \in S} \frac{p(h)}{q(h)m(h)} \left[\frac{k}{m(h)}\right]^{-[1+q(h)^{-1}]} \theta[k - m(h)] \quad (46)$$

with $q(h)$ satisfying the consistency equations

$$q(h) = \sum_{l \in S} \frac{\sigma(h,l)p(l)m(l)}{C(l)}, \quad C(h) = \sum_{l \in S} \frac{\sigma(l,h)p(l)m(l)}{1 - q(l)}. \quad (47)$$

Interestingly, if $\sigma(h,h') = 1$, the model is equivalent to a variation on the BA model with a random initial degree for
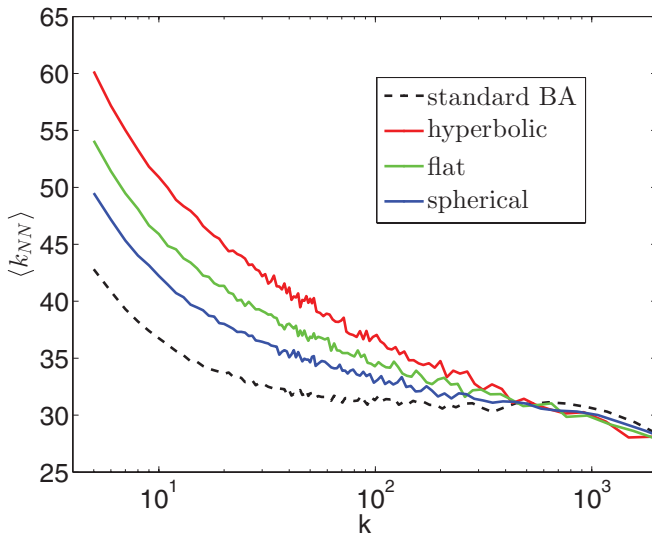

FIG. 5. (Color online) Plot of average neighbor degree $\langle k_{NN}(k) \rangle$ (averaged over 50 simulations) for networks of size 2×10^5 on disks of radius $r = 1.5$ in two-dimensional spaces of curvature $+1$, 0, -1 with connection function $\tilde{\sigma}(d) = e^{-4d/r}$.
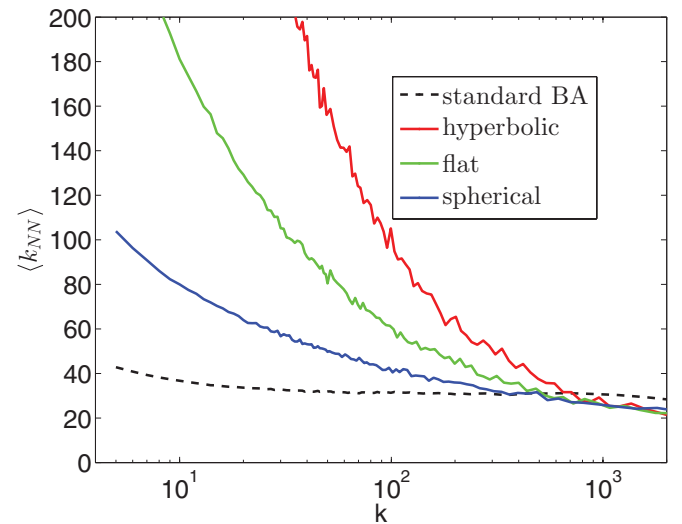

FIG. 6. (Color online) Plot of average neighbor degree $\langle k_{NN}(k) \rangle$ (averaged over 50 simulations) for networks of size 2×10^5 on disks of radius $r = 1.5$ in two-dimensional spaces of curvature $+1$, 0, -1 with connection function $\tilde{\sigma}(d) = \theta(1 - d/r)$.

the new nodes. In practice, the feature is the initial degree m of each node. Assume that the average initial connectivity $\bar{m}$ is finite. Then if the distribution of m decays faster than m^{-3}, the degree distribution of this model for $k \gg \bar{m}$ is the same as the BA model. Instead, if the distribution of m decays as a power law with exponent $-\alpha$ greater than -3, the sum in Eq. (46) gives an additional factor $k^{3-\alpha}$ and therefore $n_k \sim k^{-\alpha}$. More generally, we can consider a general variation on the BA model with degree distribution $p_0(k) \sim k^{-\gamma}$ for fixed m and modify this model to allow for a stochastic initial degree with distribution $p(m) \sim m^{-\alpha}$. In this case the degree distribution from Eq. (46) is $n_k \sim k^{-\min(\alpha,\gamma)}$, in agreement with the formal results in Ref. [42].

B. Heterogeneous links

In this model the attractiveness of a node does not depend only on the number of links, but also on the types of nodes to which they are attached. For examples, in copying vertex models or walking models [43], new nodes could preferentially explore existing nodes with a specific feature (e.g., search engines in the WWW example); therefore the effective preferential attachment dynamics would give an higher weight to the links that connect to this feature.

The preferential attachment probability is a positive linear combination of $k_{i(h)}$, which is the number of links between the ith node and nodes with variable h:

$$\Pi(i) = \frac{\sum_{h'} \sigma(h_i,h;h')k_{i(h')}}{\sum_l \sum_{h'} \sigma(h_l,h;h')k_{l(h')}}. \tag{48}$$

The distribution follows Eq. (10) with quality $q(h)$ defined by the set of consistency equations

$$q(h_i) = \sum_{h,h'} \frac{\sigma(h_i,h;h')f^{\mathrm{IN}}(h_i,h')p(h)}{C(h)}, \tag{49}$$

$$C(h) = \sum_{h',h''} \sigma(h'',h;h')p(h'')$$
$$\cdot \left[f^{\mathrm{OUT}}(h'',h') + f^{\mathrm{IN}}(h'',h')\frac{q(h'')}{1-q(h'')} \right]. \tag{50}$$

$$f^{\mathrm{OUT}}(h_i,h') = \sum_{h''} \frac{\sigma(h',h_i;h'')p(h')}{C(h_i)}$$
$$\cdot \left[f^{\mathrm{OUT}}(h',h'') + f^{\mathrm{IN}}(h',h'')\frac{q(h')}{1-q(h')} \right], \tag{51}$$

$$f^{\mathrm{IN}}(h_i,h') = \sum_{h''} \frac{p(h'')\sigma(h_i,h';h'')f^{\mathrm{IN}}(h_i,h'')}{q(h_i)C(h')}. \tag{52}$$

C. Shifted preferential attachment

The preferential attachment rule is modified by the addition of a positive term $a(h_i,h)$ independent of the degree but dependent on the features. Similar models without features were proposed in Refs. [40,44]:

$$\Pi(i) = \frac{\sigma(h_i,h)k_i + a(h_i,h)}{\sum_l \sigma(h_l,h)k_l + a(h_l,h)}. \tag{53}$$

The degree distribution for large k follows Eq. (10) with quality $q(h)$ defined by (6) and $C(h)$ defined by

$$C(h) = \sum_{h'} \left[\frac{\sigma(h',h)}{1-q(h')} + \frac{a(h',h)}{m} \right] p(h'). \tag{54}$$

D. Directed links

In this model the preferential attachment probability is proportional to the number of incoming links k^{IN}:

$$\Pi(i) = \frac{\sigma(h_i,h)\big[k_i^{\mathrm{IN}} + a(h_i)\big]}{\sum_l \sigma(h_l,h)\big[k_l^{\mathrm{IN}} + a(h_l)\big]}. \tag{55}$$

Note that the positive term $a(h)$ is needed to specify the initial attachment probability because initially $k^{\mathrm{IN}} = 0$. The results are similar to the shifted preferential attachment case, with $C(h)$ defined by

$$C(h) = \sum_{h'} \sigma(h',h) \left[\frac{q(h')}{1-q(h')} + \frac{a(h)}{m} \right] p(h'). \tag{56}$$

E. Addition of links

In this model, in addition to the usual growth rules, extra links are added at rate mr_+ and attached to nodes i,j according to the probability

$$\Pi_+(i,j) = \frac{\sigma_+(h_i,h_j)k_i k_j}{\sum_{r,s} \sigma_+(h_r,h_s)k_r k_s}. \tag{57}$$

This model can be solved similarly to the usual one under some extra assumptions, like the scaling

$$\sum_{r,s} \sigma_+(h_r,h_s)k_r k_s = m^2 C_+ t^2 + o(t^2). \tag{58}$$

The degree distribution follows Eq. (10) with quality $q(h)$ defined by

$$q(h_i) = \sum_h \left\{ \frac{\sigma(h_i,h)}{C(h)} + \frac{2r_+\sigma_+(h_i,h)}{C_+[1-q(h)]} \right\} p(h), \tag{59}$$

$$C(h) = \sum_{h'} \frac{\sigma(h',h)p(h')}{1-q(h')}, \tag{60}$$

$$C_+ = \sum_{h,h'} \frac{\sigma_+(h',h)p(h)p(h')}{[1-q(h)][1-q(h')]}. \tag{61}$$

F. Preferential rewiring in directed networks

Modifications to the usual preferential attachment growth of the node degree include also possible losses of links, because of either removal or rewiring [45]. A high rate of link removal or rewiring could result in a degree distribution with an exponential tail instead of the usual power-law tail. However, if the removal or rewiring process is not too fast, the resulting distribution is typically a power law with an exponent dependent on the rates of the different processes.

Several heterogeneous models with preferential rewiring and/or removal of links can be analyzed with the techniques of this paper. In these models the growth of node degrees can be characterized by an effective fitness $q(h)$, and the stochastic noise due to link addition and removal does not

change significantly the tail of the degree distribution, as explained in Appendix C. Here we present a simple example of such a model.

We consider a directed network growing under the same rules as model (2) and with rewiring taking place at rate mr. In each rewiring process, a random link is selected with probability proportional to $\sigma_-(h_{\rm in}, h_{\rm out})$, where $h_{\rm in}$ and $h_{\rm out}$ are the variables of the attached nodes. The ingoing end of the link is then detached and reattached to another node according to the probability

$$\Pi_+(i) = \frac{\sigma_+(h_i, h_{\rm out}) k_i^{\rm IN}}{\sum_j \sigma_+(h_j, h_{\rm out}) k_j^{\rm IN}}. \tag{62}$$

In this model, at least for large k, the degree distribution follows a multiscaling behavior similar to Eq. (10), given by

$$n_k \simeq \sum_{h \in S} \theta(q(h)) \frac{p(h)}{q(h)m} \left(\frac{k}{m}\right)^{-(1+q(h)^{-1})} \tag{63}$$

with quality $q(h)$ defined by the set of consistency equations

$$q(h_i) = \sum_h \frac{\sigma(h_i, h) p(h)}{C(h)} + \frac{r \Sigma_+(h_i)}{C_+}$$
$$- \sum_h \frac{r \sigma_-(h_i, h) f^{\rm IN}(h_i, h)}{C_-}, \tag{64}$$

$$C(h) = \sum_{h'} \frac{\sigma(h', h) p(h')}{1 - q(h')}, \tag{65}$$

$$C_+ = \sum_h \frac{\Sigma_+(h) q(h) p(h)}{1 - q(h)}, \tag{66}$$

$$C_- = \sum_{h, h'} \frac{\sigma_-(h', h) f^{\rm IN}(h', h) q(h') p(h')}{1 - q(h')}, \tag{67}$$

$$f^{\rm IN}(h_i, h) = \frac{\sigma(h_i, h) p(h)/C(h) + r \Sigma_+(h_i)/C_+}{q(h_i) + r \sigma_-(h_i, h)/C_-}, \tag{68}$$

$$\Sigma_+(h_i) = \sum_{h, h'} \frac{\sigma_+(h_i, h) \sigma_-(h', h) f^{\rm IN}(h', h) q(h') p(h')}{C_- [1 - q(h')]}. \tag{69}$$

In this model (and more generally in models including rewiring or removal of links) the quality $q(h)$ can also be negative, thus requiring the factor $\theta[q(h)]$ in Eq. (63).

G. Fixing the connection probability between features

The last variation is a more radical departure from the heterogeneous models presented in this paper, since it is a modification of the attachment probability (2). In this model, a new node with feature h attaches to nodes with different variables h' according to a probability $\pi(h'|h)$ independent of the degrees. Then, once a feature h' is chosen at random according to $\pi(h'|h)$, a specific node i with $h_i = h'$ is chosen according to the usual preferential attachment rule. [If no such node exists, another variable h'' is chosen according to $\pi(h''|h)$.] Asymptotically, the overall probability is then

$$\Pi(i) = \pi(h_i|h) \frac{k_i}{\sum_k k N_k(h_i, t)}. \tag{70}$$

It is possible to define a quality $q(h)$ also for these models. Assuming the scaling

$$\sum_{\{j|h_j = h\}} k_j = m p(h) C(h) t + o(t), \tag{71}$$

we obtain $C(h) = 1/[1 - q(h)]$, and the quality can be obtained explicitly:

$$q(h_i) = \frac{\sum_h \pi(h_i|h) p(h)}{p(h_i) + \sum_h \pi(h_i|h) p(h)}. \tag{72}$$

The degree distribution follows the usual Eq. (10). This model works if the feature space S is discrete and finite, but can be generalized to continuous spaces through discretization, and Eq. (72) is valid with $p(h)$ and $\pi(h'|h)$ interpreted as probability densities.

V. CONCLUSIONS

The addition of node features to growing network models with preferential attachment is an important step toward realistic network modeling and results in a wide class of models, for which this paper provides several analytical results. In particular, this work shows that the power-law scaling of the degree distribution generated by preferential attachment is quite robust with respect to the heterogeneity between nodes. The main effect of heterogeneity is the emergence of an "effective fitness" $q(h)$ for each class of nodes, and therefore their degree distribution resembles the fitness model of Bianconi and Barabási [23,24].

Beyond the degree distribution, other network properties were studied. The clustering coefficient of these networks disappears for large network size, a property shared with the BA model. Negative degree correlations are also present in these models, along with nontrivial mixing patterns among features. Both small clustering coefficients and disassortative mixing are therefore outcomes of the preferential attachment mechanism in general growing networks.

The effect of the features h associated to each node has been presented as nonrandom, but the formalism applies to any kind of heterogeneity. In particular it is easy to include random variables or features with random effects as well, as long as their values do not change with time. In fact, any random effect can be parametrized by some extra random variables χ with a distribution $p_R(\chi)$. Then it is possible to redefine a nonrandom variable $\tilde{h} = (h, \chi)$ with frequency $\tilde{p}(\tilde{h}) = p(h) p_R(\chi)$. The connection function $\tilde{\sigma}(\tilde{h}, \tilde{h}')$ now takes into account both random and nonrandom components. So the formalism captures stochastic as well as deterministic node features.

Moreover, the growth of many scale-free networks is based on some local dynamics such as the vertex copying or duplication rules for growth of molecular networks. However, from the point of view of the link distribution, the local dynamics often results in an effective preferential attachment mechanism. Our methods and results on degree distribution and assortativity apply to these models as well, if we take the connection probability (2) as an effective dynamics for the growth of the node connectivities. Therefore the main results of this paper, i.e., power-law multiscaling of the degree distribution and disassortative mixing in degree, are generally

valid for models with effective preferential attachment. On the other hand, the clustering coefficient depends on the specific model and not only on the effective form (2) for the attachment probability, and therefore our proof that the clustering vanishes in the thermodynamic limit is valid only for models with pure preferential attachment.

In future work it would be very interesting to map the metric space implied by the class of growing network models discussed in this paper with the hidden metrics recently introduced to model complex networks in hyperbolic spaces [46]. Finally, an interesting extension of the model would be to include features that fluctuate in time, in order to determine how time-dependent heterogeneities affect the power-law behavior of the degree distribution and the other properties discussed here, and features that coevolve with the network, for example, spaces that expand with time while the node density remains constant.

ACKNOWLEDGMENTS

APPENDIX A: CLUSTERING IN THE BARABÁSI-ALBERT MODEL AND IN HETEROGENOUS MODELS

In a general heterogenous model, we consider a triplet of nodes born at times $t_A < t_B < t_C$ with qualities q_A, q_B, q_C. The average number of triangles can be found by integrating the probability that all three nodes are connected on the birth times t_A, t_B, t_C:

$$n_{\text{triangles}} = \int_1^t dt_C \int_1^{t_C} dt_A \int_{t_A}^{t_C} dt_B \frac{\sigma(h_A, h_C) m(t_C/t_A)^{q_A}}{mC(h_C)t_C}$$
$$\cdot \frac{\sigma(h_B, h_C)(m-1)(t_C/t_B)^{q_B}}{mC(h_C)t_C} \frac{\sigma(h_A, h_B) m(t_B/t_A)^{q_A}}{mC(h_B)t_B}. \quad (A1)$$

The average number of triples can be found in a similar way. These networks can contain three kind of triples: $A \leftarrow B \leftarrow C$, $A \leftarrow C \rightarrow B$, and $B \rightarrow A \leftarrow C$. Since each new node has m outgoing links, it increases the number of triples $A \leftarrow C \rightarrow B$ by a factor $m(m-1)/2$ and the number of triples $A \leftarrow B \leftarrow C$ by a factor m^2, independently of the model, so their total number is $m(3m-1)t/2$. The number of triples $B \rightarrow A \leftarrow C$ is given by the integral

$$n_{\text{triples } B \rightarrow A \leftarrow C}$$
$$= \int_1^t dt_C \int_1^{t_C} dt_A \int_{t_A}^{t_C} dt_B \frac{\sigma(h_A, h_C) m(t_C/t_A)^{q_A}}{mC(h_C)t_C}$$
$$\times \frac{\sigma(h_A, h_B) m(t_B/t_A)^{q_A}}{mC(h_B)t_B}. \quad (A2)$$

We give the full result for BA networks, since the other cases are straightforward (but cumbersome) generalizations. In the BA case we have $q_A = q_B = q_C = 1/2$, and the clustering

coefficient is

$$C_{\text{clust}}^{\text{BA}} = \frac{m(m-1)\ln^3 t}{8\left(mt\ln t - t + mt^{1/2}\right)}, \quad (A3)$$

including some finite corrections to the leading behavior $C(t) \propto \ln^2 t/t$.

APPENDIX B: ASSORTATIVITY IN HETEROGENEOUS MODELS

Following Ref. [40], the rate equation for $N_{k,l}^{(h,i)} = n_{k,l}^{(h,i)} \cdot mt$ is

$$\frac{dN_{k,l}^{(h,i)}}{dt} = \frac{q(h)}{t}\left[(k-1)N_{k-1,l}^{(h,i)} - kN_{k,l}^{(h,i)}\right]$$
$$+ \frac{q(i)}{t}\left[(l-1)N_{k,l-1}^{(h,i)} - lN_{k,l}^{(h,i)}\right]$$
$$+ \delta_{l,m}\frac{\sigma(h,i)p(i)}{C(i)t}(k-1)N_{k-1}(h,t), \quad (B1)$$

which is equivalent to Eq. (42) after substituting $n_{k,l}^{(h,i)}$ and rearranging. The solution can be found in terms of the generating function

$$g(x,y) = \sum_{k=m+1}^{\infty}\sum_{l=m}^{\infty} x^k y^l n_{k,l}^{(h,i)}, \quad (B2)$$

and the result is

$$g(x,y) = \frac{\sigma(h,i)p(i)}{C(i)q(h)}\frac{(1-x)^{[1+mq(i)]/q(h)}y^m}{x^{1/q(h)}}$$
$$\cdot \int_0^x \frac{dr\ f(r)r^{\frac{1+mq(i)-q(h)}{q(h)}}(1-r)^{-1-\frac{1}{q(h)}}}{\left[(1-x)^{\frac{q(i)}{q(h)}}yr^{\frac{q(i)}{q(h)}} + x^{\frac{q(i)}{q(h)}}(1-y)(1-r)^{\frac{q(i)}{q(h)}}\right]^m}, \quad (B3)$$

where $f(r)$ is the generating function of $(k-1)n_{k-1}(h)$:

$$f(r) = \frac{p(h)\Gamma[m + q(h)^{-1}]}{q(h)\Gamma(m)}\left(\frac{1-x}{x}\right)^{q(h)^{-1}-1}$$
$$\cdot \int_0^x dz\ z^{m-1}\left(\frac{z}{1-z}\right)^{q(h)^{-1}}. \quad (B4)$$

For $m = 1$ the generating function (B3) can be expanded into the solution (43).

For $m = 1$ it is easy to expand $g(x,y)$ in power of y and then apply Tauberian theorems to the functions $l!(\partial/\partial y)^l g(x,y)|_{y=0}$, obtaining the scaling $n_{k,l}^{(h,i)} \sim k^{-1/q(h)}l^{-1-q(h)/q(i)}$ for fixed l and large k.

Since it is not easy to further understand the behavior of $n_{k,l}^{(h,i)}$ from the above solution, we approximate the differences in square brackets in (B1) as derivatives $-\partial(kN_{k,l}^{(h,i)})/\partial k$ and $-\partial(lN_{k,l}^{(h,i)})/\partial l$, rearrange, and obtain the differential equation $n_{k,l}^{(h,i)} = q(h)\partial(kn_{k,l}^{(h,i)})/\partial k + q(i)\partial(ln_{k,l}^{(h,i)})/\partial l$. Its solutions have the form

$$n_{k,l}^{(h,i)} \sim f(k^{1/q(h)}l^{-1/q(i)})k^{-1-1/2q(h)}l^{-1-1/2q(i)}, \quad (B5)$$

where f is an arbitrary differentiable function.

Now we obtain the scaling of the exact boundary conditions $n_{k,m}^{(h,i)}$ and $n_{m+1,l}^{(h,i)}$. The first satisfies the equation

$$n_{k,m}^{(h,i)}[1 + q(h)k + q(i)m]$$
$$= q(h)(k-1)n_{k-1,m}^{(h,i)} + \frac{\sigma(h,i)p(i)}{C(i)}(k-1)n_{k-1}(h). \quad \text{(B6)}$$

Its solution is

$$n_{k,m}^{(h,i)} = \frac{\sigma(h,i)p(i)p(h)}{C(i)q(h)^2} \frac{\Gamma[m+q(h)^{-1}]\Gamma(k)}{\Gamma(m)\Gamma\{k+1+[1+mq(i)]/q(h)\}}$$
$$\cdot \sum_{j=m+1}^{k} \frac{\Gamma\{j+[1+mq(i)]/q(h)\}}{\Gamma[j+q(h)^{-1}]} \sim k^{-1/q(h)}. \quad \text{(B7)}$$

Matching the scaling in k with Eq. (B5) gives $f(k^{1/q(h)})k^{-1-1/2q(h)} \sim k^{-1/q(h)}$, so $f(x) \sim x^{q(h)-1/2}$ and therefore $n_{k,l}^{(h,i)} \sim k^{-1/q(h)}l^{1-q(h)/q(i)}$ for $l \ll k$, consistent with exact results.

The second boundary $n_{m+1,l}^{(h,i)}$ satisfies the equation

$$n_{m+1,l}^{(h,i)}[1 + q(h)(m+1) + q(i)l]$$
$$= q(i)(l-1)n_{m+1,l-1}^{(h,i)} + \delta_{l,m}\frac{\sigma(h,i)p(i)}{C(i)}mn_m(h). \quad \text{(B8)}$$

Its solution is

$$n_{m+1,l}^{(h,i)} = \frac{\sigma(h,i)p(i)p(h)m}{C(i)[m+q(h)^{-1}]q(h)q(i)}$$
$$\cdot \frac{\Gamma(l)\Gamma\{m+[1+(m+1)q(h)]/q(i)\}}{\Gamma(m)\Gamma\{l+1+[1+(m+1)q(h)]/q(i)\}}$$
$$\sim l^{-1-[1+(m+1)q(h)]/q(i)}, \quad \text{(B9)}$$

so the matching $f(l^{-1/q(i)})l^{-1-1/2q(i)} \sim l^{-1-[1+(m+1)q(h)]/q(i)}$ gives $f(x) \sim x^{1/2+(m+1)q(h)}$, and therefore $n_{k,l}^{(h,i)} \sim k^m l^{-1-[1+(m+1)q(h)]/q(i)}$ for $k \ll l$.

For the BA model [i.e., $q(h) = q(i) = 1/2$ with $m = 1$, we obtain $n_{k,l}^{(h,i)} \sim k^{-2}l^{-2}$ for $k \gg l$ and kl^{-5} for $k \ll l$, in agreement with exact results [40].

APPENDIX C: STOCHASTIC EFFECTS AND REMOVAL OF LINKS

Some of the models presented in this paper involve with not only addition of new links, but also removal of existing links. For example, the rewiring process is equivalent to removing and then adding a link to the network.

In these models, the continuum equation for the degree of a node [38] is simply $dk/dt = (q_+ - q_-)k/t$, where q_+ and q_- represent asymptotic growth and reduction coefficients coming from addition and deletion of links, respectively. The birth rate of nodes is constant, and therefore the continuum approach predicts that the degree distribution is the usual power law $n_k \propto k^{-1-1/(q_+-q_-)}$ expected for growing networks with preferential attachment. However, the continuum equation could be wrong in predicting the degree distribution or the consistency equations in these models. In this appendix we examine these aspect in more detail.

First, the consistency equations have typically the form

$$f(h)t = \left\langle \sum_i g(h_i)k_i(t) \right\rangle$$
$$\Updownarrow$$
$$f(h) = \sum_h g(h)p(h) \cdot \frac{1}{p(h)t} \sum_{\{i|h_i=h\}} \langle k_i(t) \rangle, \quad \text{(C1)}$$

where $\langle \ldots \rangle$ represents an average over realizations of the process. Therefore these equations depend only on the average degree $\langle k \rangle$ of the nodes. But the asymptotic evolution of the average degree $\langle k(t) \rangle$ is described precisely by the continuum equation for $k(t)$, and therefore the consistency equations predicted by the continuum approach are exact. For example, Eq. (8) can be obtained equivalently from the rate equation approach or the continuum approach, i.e., from the first or the second line of Eq. (7).

Second, while q_+ and q_- are positive by definition, there is no reason for $q_+ - q_-$ to be positive definite. In fact, there are models where all nodes have negative global quality $q_+ - q_- < 0$. The degree distribution of these models deviates significantly from the scale-free behavior, since on average the connectivity of all nodes decreases with time and nodes of high degree can appear only due to the effect of fluctuations, and therefore the tail of the distribution falls exponentially. (In particular, applying the Langevin equation approximation presented below, it should fall faster than $e^{-\sqrt{k}}$.) The exact degree distribution for these models will be presented elsewhere.

However, in the present paper we focus on models where at least some classes of nodes have positive quality $q_+ - q_- > 0$. In these models there are two kinds of nodes: those with $q_+ > q_-$ (and therefore positive average growth rate), for which the continuum approach could work, and those with $q_+ < q_-$ (and negative average growth rate), for which the continuum approach fails as described above. However, the tail of the latter nodes falls rapidly and do not contribute to the degree distribution, so it is sufficient to calculate the degree distribution using the continuum approximation for the nodes with $q_+ - q_- > 0$ and discard the others.

Third, even for nodes with $q_+ - q_- > 0$, rewiring and removal enhance the stochasticity of the process, potentially affecting the distribution. To take this noise into account, we promote the above equation to a Langevin equation obtained by adding a stochastic term $+\sqrt{(q_+ + q_-)k/t} \cdot \eta(t)$ with $\eta(t)$ a white Gaussian noise with variance 1. Defining $q = q_+ - q_-$ and $Q = q_+ + q_-$, we finally obtain the integrated Fokker-Planck equation,

$$\frac{\partial n(k,t)}{\partial t} = -\frac{n(k,t)}{t} - \frac{q}{t}\frac{\partial[kn(k,t)]}{\partial k} + \frac{Q}{2t}\frac{\partial^2[kn(k,t)]}{\partial k^2}, \quad \text{(C2)}$$

for the average degree distribution $n(k,t) = \int_1^t dt_0 P(k,t|t_0)/t$, where $P(k,t|t_0)$ is the distribution for a node born at time t_0. The rate equation approach leads to the same equation. The (normalizable) stationary solution of this equation can be derived as a power series:

$$n_k \propto k^{-1-q^{-1}}\left[1 + \sum_{n=1}^{\infty}(1+nq)\frac{\Gamma(n+q^{-1})^2}{\Gamma(n+1)}\left(-\frac{Q}{q}\right)^n k^{-n}\right]. \quad \text{(C3)}$$

This is an asymptotic series that can be resummed by Borel summation in the variable $z = k^{-1}$. The leading contribution to the Borel transform of the sum in (C3) behaves as $\mathcal{B}(z) \sim Qz$ for $z \to 0$; therefore for large k the Borel sum is Qk^{-1}, and the resulting degree distribution is $n_k \sim k^{-1-q^{-1}}[1 + O(Qk^{-1})]$, thereby confirming the validity of the simple continuum approach (without the diffusion term) for $k \gg Q$.

[1] R. Albert and A. Barabási, Rev. Mod. Phys. **74**, 47 (2002).

[2] G. Caldarelli, *Scale-Free Networks* (Oxford University Press, Oxford, 2007).

[3] V. Latora, Y. Moreno, M. Chavez, M. D. Hwang, and S. Boccaletti, Phys. Rep. **424**, 175 (2006).

[4] M. B. A. Barrat and A. Vespignani, *Dynamical Processes on Complex Networks* (Cambridge University Press, Cambridge, 2008).

[5] M. E. J. Newman, *Networks: An Introduction* (Oxford University Press, Oxford, 2010).

[6] S. N. Dorogovtsev, *Lectures on Complex Networks* (Oxford University Press, Oxford, 2009).

[7] M. Barthélemy, Phys. Rep. **499**, 1 (2011).

[8] F. Papadopoulos, D. Krioukov, M. Bogua, and A. Vahdat, in *INFOCOM, 2010 Proceedings IEEE* (IEEE, San Diego, CA, USA, 2010), pp. 1–9.

[9] G. Bianconi, P. Pin, and M. Marsili, Proc. Natl. Acad. Sci. USA **106**, 11433 (2009).

[10] K. K. A. B. L. Daqing and S. Havlin, Nat. Phys. **7**, 481 (2010).

[11] M. Boguñá, R. Pastor-Satorras, A. Díaz-Guilera, and A. Arenas, Phys. Rev. E **70**, 056122 (2004).

[12] J. Kleinberg, Nature (London) **406**, 845 (2000).

[13] G. Bianconi, Phys. Lett. A **303**, 166 (2002).

[14] S. Fortunato and M. Barthélemy, Proc. Natl. Acad. Sci. USA **104**, 36 (2007).

[15] E. Airoldi, D. Blei, S. Fienberg, and E. Xing, J. Mach. Learn. Res. **9**, 1981 (2008).

[16] G. Palla, L. Lovász, and T. Vicsek, Proc. Natl. Acad. Sci. USA **107**, 7640 (2010).

[17] A. Barabási and Z. Oltvai, Nat. Rev. Genet. **5**, 101 (2004).

[18] J. Stuart, E. Segal, D. Koller, and S. Kim, Science **302**, 249 (2003).

[19] B. Yang, A. Bassols, Y. Saco, and M. Pérez-Enciso, Genet. Sel. Evol. **43**, 28 (2011).

[20] A. Barabasi and R. Albert, Science **286**, 509 (1999).

[21] M. Newman, SIAM Rev. **45**, 167 (2003).

[22] P. L. Krapivsky, S. Redner, and F. Leyvraz, Phys. Rev. Lett. **85**, 4629 (2000).

[23] G. Bianconi and A. Barabási, Europhys. Lett. **54**, 436 (2001).

[24] G. Bianconi and A. L. Barabási, Phys. Rev. Lett. **86**, 5632 (2001).

[25] S. S. Manna and P. Sen, Phys. Rev. E **66**, 066114 (2002).

[26] R. Xulvi-Brunet and I. M. Sokolov, Phys. Rev. E **66**, 026118 (2002).

[27] M. Barthélemy, Europhys. Lett. **63**, 915 (2003).

[28] A. Santiago and R. Benito, Int. J. Mod. Phys. C **18**, 1591 (2007).

[29] A. Santiago and R. Benito, Europhys. Lett. **82**, 58004 (2008).

[30] L. Ferretti and M. Cortelezzi, Phys. Rev. E **84**, 016103 (2011).

[31] S. Cuenda and J. Crespo, Europhys. Lett. **95**, 38002 (2011).

[32] S. Yook, H. Jeong, and A. Barabási, Proc. Natl. Acad. Sci. USA **99**, 13382 (2002).

[33] J. Jordan, Adv. Appl. Prob. **42**, 319 (2010).

[34] G. Ergün, Physica A **308**, 483 (2002).

[35] M. Girvan and M. Newman, Proc. Natl. Acad. Sci. USA **99**, 7821 (2002).

[36] D. Watts, P. Dodds, and M. Newman, Science **296**, 1302 (2002).

[37] B. Bollobás and O. M. Riordan, in *Handbook of Graphs and Networks: From the Genome to the Internet*, edited by S. Bornholdt and H. G. Schuster (Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, FRG, 2005).

[38] A. Barabási, R. Albert, and H. Jeong, Physica A **272**, 173 (1999).

[39] A. Fronczak, P. Fronczak, and J. A. Hołyst, Phys. Rev. E **68**, 046126 (2003).

[40] P. L. Krapivsky and S. Redner, Phys. Rev. E **63**, 066123 (2001).

[41] M. E. J. Newman, Phys. Rev. E **67**, 026126 (2003).

[42] M. Deijfen, H. van den Esker, R. Van Der Hofstad, and G. Hooghiemstra, Arkiv för matematik **47**, 41 (2009).

[43] A. Vazquez, Europhys. Lett. **54**, 430 (2001).

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

[45] R. Albert and A. Barabási, Phys. Rev. Lett. **85**, 5234 (2000).

[46] D. Krioukov, F. Papadopoulos, A. Vahdat, and M. Boguñá, Phys. Rev. E **80**, 035101 (2009).