

December 16, 2011

Entropy rate of nonequilibrium growing networks

Kun Zhao

Northeastern University

Arda Halu

Northeastern University

Simone Severini

University College London

Ginestra Bianconi

Northeastern University

Recommended Citation

Zhao, Kun; Halu, Arda; Severini, Simone; and Bianconi, Ginestra, "Entropy rate of nonequilibrium growing networks" (2011). *Physics Faculty Publications*. Paper 524. <http://hdl.handle.net/2047/d20004301>

This work is available open access, hosted by Northeastern University.

Entropy rate of nonequilibrium growing networks

Kun Zhao,¹ Arda Halu,¹ Simone Severini,² and Ginestra Bianconi¹

¹*Department of Physics, Northeastern University, Boston, 02115 Massachusetts, USA*

²*Department of Computer Science, and Department of Physics & Astronomy, University College London, WC1E 6BT London, United Kingdom*

(Received 5 September 2011; published 16 December 2011)

New entropy measures have been recently introduced for the quantification of the complexity of networks. Most of these entropy measures apply to static networks or to dynamical processes defined on static complex networks. In this paper we define the entropy rate of growing network models. This entropy rate quantifies how many labeled networks are typically generated by the growing network models. We analytically evaluate the difference between the entropy rate of growing tree network models and the entropy of tree networks that have the same asymptotic degree distribution. We find that the growing networks with linear preferential attachment generated by dynamical models are exponentially less than the static networks with the same degree distribution for a large variety of relevant growing network models. We study the entropy rate for growing network models showing structural phase transitions including models with nonlinear preferential attachment. Finally, we bring numerical evidence that the entropy rate above and below the structural phase transitions follows a different scaling with the network size.

DOI: [10.1103/PhysRevE.84.066113](https://doi.org/10.1103/PhysRevE.84.066113)

PACS number(s): 89.75.Hc, 89.75.Da

I. INTRODUCTION

Complex networks describe the structure of many biological, social, and technological systems [1–5]. Recently new entropy measures have been introduced for the quantification of complexity of networks [6–24]. Methods for quantifying complexity are not only valuable from the theoretical point of view, but may also lead to important operational interpretations. This new framework has the potential to resolve one of the outstanding problems in statistical mechanics of complex systems. In fact, it opens the way for a new information theory of complex network topologies that will provide an evaluation of the information encoded in complex networks.

The entropy of network ensembles quantifies the number of graphs with given structural features such as degree distribution, degree correlations, community structure, or spatial embedding [7–14]. This quantity has been shown to be very useful for inference problems defined on networks, and it has been successfully applied to the problem of assessing the significance of features for network structure [15]. Other entropy measures of quantum mechanical nature have been derived by mapping the network either to a density matrix or to a quantum state [17–21]. These entropies, defined on single networks, set a path for the application of tools of quantum information theory to describe the complexity of single networks and to introduce new kind of network parameters (for example, by considering the notion of correlations and subsystems). Finally, the entropy rate of random walks [22–24] is extensively studied on networks, and it predicts how evenly the random walk spreads in the network. For many applications, indeed, one would wish to bias the random walk to construct maximally entropic random walks.

All these definitions of entropy of networks consider static networks, eventually explored by a dynamical process. In this paper we define and evaluate the entropy rate of growing network models. The literature in the field of growing network models generating scale-free networks is very large [1–5]. By studying the entropy rate of these models we aim at quantifying the number of typical networks that are generated

by these models, and we compare this number with the number of networks that is possible to construct with the same degree distribution. In particular we focus on trees to allow for an analytic treatment of the problem. Trees are networks in which no cycle is allowed. The maximal number of possible tree networks generated by a growing network model scales like $N!$ where N is the number of nodes (and links) in the network. The minimal number of tree networks generated by a growing network model is one, corresponding to the formation of a star or of a linear chain. The entropy rate of growing scale-free networks lies between these two limiting values. Understanding the value of the entropy of graphs is informative because it describes the complexity of the growing network models. In fact, the value of the entropy will quantify with a unique number the size of the space of typical networks generated by the growing network model. The smaller is the entropy rate of the networks the more complex the network structural properties implied by the growing model. In particular, it is essential to determine the scaling with N of the entropy rate, and in the case in which the entropy rate is not constant but depends on N it is important to evaluate the subleading terms that encode for the topology of the networks also for other entropy measures [8,9,20].

The main model of growing scale-free networks is the Barabási-Albert (BA) model [25], which generates scale-free networks with power-law exponent $\gamma = 3$. The BA networks are known to have weak degree correlations due to their causal structure; the growing network model with initial attractiveness [26] of the nodes and the fitness model [27] have more significant correlations. To quantify these correlations different measures have been introduced, such as the average degree of the neighbor of the nodes or the degree correlation matrix. Still we lack a way to quantify how much information is encoded in growing network models with respect to the networks constructed by the configuration model with the same degree distribution.

Here we propose to quantify the number of typical tree graphs generated by the nonequilibrium growing network

models [25–31] as a proxy of their complexity. This quantity can be used to measure the fraction of networks of given degree sequence that is generated by growing network models and to quantify in this way the complexity of growing network models. Moreover growing network models such as the Bianconi and Barabási fitness model [27,30] and the nonlinear preferential attachment model of Krapivsky and Redner [28,29] or the growing network model with aging of the nodes [31] are known to undergo structural phase transitions as a function of their parameters. Here we interestingly see that these phase transitions are characterized by a sharp drop of the entropy rate and strong finite size effects indicating that the network is reduced to a more ordered state.

The paper is structured as follows. In Sec. II we define the Gibbs entropy of networks with a given degree distribution. In Sec. III we introduce the necessary material for studying the entropy rate of growing trees. First, we recall the main growing network models. Then we obtain minimum and maximum bounds to the entropy. In Sec. IV we study growing trees with stationary degree distribution. In particular, we consider the BA model, initial attractiveness, Bianconi-Barabási fitness model, and networks with structural phase transitions. We draw brief conclusions in Sec. V.

II. GIBBS ENTROPY OF NETWORKS WITH A GIVEN DEGREE DISTRIBUTION

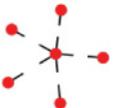
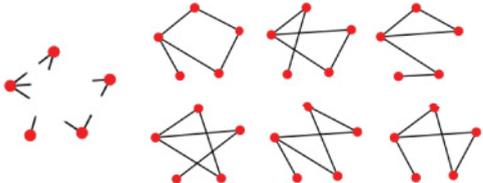
The Gibbs entropy $\Sigma[\{k_i\}]$ [7–11] of a network ensemble with given (graphical) degree sequence $\{k_i\}$ [32,33] is

$$\Sigma[\{k_i\}] = \frac{1}{N} \log \mathcal{N}[\{k_i\}], \quad (1)$$

where $\mathcal{N}[\{k_i\}]$ is the number of networks with the specified degree sequence and N is the number of labeled nodes $i = 1, 2, \dots, N$. The Gibbs entropy depends on the number of links but also on the specific details of the degree sequence. In Table I we give two illustrative examples for two degree sequences compatible with five links but defining ensembles of networks with different entropy.

It turns out that the ensemble of networks having a given degree distribution is a type of *microcanonical network*

TABLE I. The configuration of networks with degree sequence $\{1,1,1,1,5\}$ (top, $\mathcal{N}[\{k_i\}] = 1$) and $\{1,2,2,2,3\}$ (bottom, $\mathcal{N}[\{k_i\}] = 6$).

Degree Sequence	Networks	Entropy
		$\Sigma[\{k_i\}] = 0$
		$\Sigma[\{k_i\}] \neq 0$

ensemble satisfying a large number of hard constraints (the degree of each node is fixed). It is also possible to construct *canonical network ensembles* similar to what happens in classical statistical mechanics when one distinguishes the microcanonical and canonical ensembles according to the fact that the energy is perfectly conserved or conserved in average. A canonical network ensemble with given expected degree sequence is an ensemble of graphs in which the degree of each node is distributed as a Poisson variable with given expected degree $\{\bar{k}_i\}$. The entropy of the canonical network ensemble is the logarithm of the typical number of networks in the ensemble. This entropy $S[\{\bar{k}_i\}]$ is given by

$$S[\{\bar{k}_i\}] = -\frac{1}{N} \left[\sum_{ij} p_{ij} \log p_{ij} - \sum_{ij} (1 - p_{ij}) \log(1 - p_{ij}) \right], \quad (2)$$

where p_{ij} indicates the probability that a node i is linked to a node j . We can evaluate the entropy of a maximally random network ensemble with given expected degree distribution $\{\bar{k}_i\}$ by maximizing the entropy $S[\{k_i\}]$ with respect to p_{ij} under the conditions

$$\bar{k}_i = \sum_j p_{ij}. \quad (3)$$

In this way we get for the marginal probabilities p_{ij} [10]

$$p_{ij} = \frac{\theta_i \theta_j}{1 + \theta_i \theta_j}, \quad (4)$$

where θ_i are related to the lagrangian multipliers, or “hidden variables,” fixed by the constraints given by Eq. (3). In particular, for the uncorrelated network model in which $\bar{k}_i < \sqrt{\langle k \rangle N}$ and $p_{ij} = \frac{\bar{k}_i \bar{k}_j}{\langle k \rangle N}$, the Shannon entropy network ensemble takes a direct form [9]

$$S[\{\bar{k}_i\}] = \frac{1}{2} \langle k \rangle [\log(\langle k \rangle N) - 1] - \frac{1}{N} \sum_i (\ln \bar{k}_i - 1) \bar{k}_i. \quad (5)$$

The Gibbs entropy Σ of a microcanonical ensemble of networks with degree sequence $\{k_i\}$ with $k_i = \bar{k}_i$ is given by [11]

$$\Sigma[\{k_i\}] = S[\{k_i\}] - \Omega[\{k_i\}], \quad (6)$$

where $\Omega[\{k_i\}]$ is the entropy of large deviations of the canonical ensemble

$$\Omega[\{k_i\}] = -\frac{1}{N} \log \left[\sum_{a_{ij}} p_{ij}^{a_{ij}} (1 - p_{ij})^{1-a_{ij}} \prod_i \delta(\sum_j a_{ij}, k_i) \right], \quad (7)$$

where $\{a_{ij}\}$ is the adjacency matrix of the network. In particular the matrix element a_{ij} of the adjacency matrix is given by $a_{ij} = 1$ if a link is present between node i and node j , while $a_{ij} = 0$ otherwise. By replica methods and the cavity method [8,11] it is possible to derive the given expression for $\Omega[\{k_i\}]$,

$$\Omega[\{k_i\}] = -\frac{1}{N} \sum_i \log \pi_{k_i}(k_i), \quad (8)$$

where $\pi_r(n)$ is the Poisson distribution with $\langle n \rangle = r$. In particular, for the uncorrelated network model in which $k_i < \sqrt{\langle k \rangle N}$ and $p_{ij} = \frac{k_i k_j}{\langle k \rangle N}$, the Gibbs entropy network ensemble takes a direct form [9]

$$\Sigma[\{k_i\}] = \frac{1}{2} \langle k \rangle [\log(\langle k \rangle N) - 1] - \frac{1}{N} \sum_i (\ln k_i - 1) k_i - \frac{1}{2N} \sum_i \log(2\pi k_i) - \frac{1}{4N} \left(\frac{\langle k^2 \rangle}{\langle k \rangle} \right)^2. \quad (9)$$

We might as well define the Gibbs entropy $\Sigma[\{N_k\}]$ of networks with given degree distribution $\{N_k\}$. Since the number of graphs with given degree distribution $\mathcal{N}[\{N_k\}]$ is just given by

$$\mathcal{N}[\{N_k\}] = \mathcal{N}[\{k_i\}] \frac{N!}{\prod_k N_k!}, \quad (10)$$

it follows that

$$\Sigma[\{N_k\}] = \Sigma[\{k_i\}] - \sum_k \frac{N_k}{N} \log \left(\frac{N_k}{N} \right). \quad (11)$$

III. ENTROPY RATE OF GROWING TREES

Many networks are nonstatic, but they are growing by the addition of new nodes and links. A major class of growing networks are growing trees in which at each time a new node and a new link is added to the network. In the last 10 years, many growing network models have been proposed. Special attention has been addressed to growing network models generating scale-free networks. In fact, these stylized models explain the basic mechanism according to which many growing natural networks develop the universally found scale-free degree distribution. The fundamental model for scale-free growing network is the BA model [25], which generates networks with degree distribution $P(k) \sim k^{-\gamma}$ and $\gamma = 3$. This model is based on two ingredients: growth of the network and preferential attachment, meaning that nodes with large degree are more likely to acquire new links. In this paper we consider this model and other different significant variations to this model, including different additional mechanisms as initial attractiveness of the nodes [26], fitness of the nodes [27,30], nonlinear preferential attachment [28,29], and aging of the nodes [31]. Some of these models as explained below undergo structural phase transitions to be studied by statistical mechanics methods.

A. Growing network models

In the growing scale-free network model we start from two nodes linked together, at each time $t = 1, 2, \dots$:

- (1) We add a new node $i = t + 2$.
- (2) We link the new node to a node i_t of the network chosen with probability

$$\Pi(i_t) = \frac{A_{i_t}}{\mathcal{N}}, \quad (12)$$

where $\mathcal{N} = \sum_{i=1}^{t+1} A_i$

- (3) The number of nodes in the network is $N = t + 2$.

As a function of the choice of A_j different networks model are defined. In particular, we consider the following growing network models:

- (1) If we take $A_i = \delta_{i,1}$, we get a star graph.
- (2) If we take $A_i = \delta_{i,t+1}$, we get the linear chain.
- (3) If we take $A_i = 1$, we get a maximally random connected and growing tree.
- (4) If we take $A_i = k_i$, where k_i is the degree of the node i , we get the BA model [25].
- (5) If we take $A_i = k_i - 1 + a$, with $a < 1$, we get a generalized BA model with initial attractiveness of the nodes [26].
- (6) If we assign to each node a fitness value η_i from a distribution $\rho(\eta) = 1$ and $\eta \in (0,1)$ and we take $A_i = \eta_i k_i$, we get the Bianconi-Barabási fitness model [27].
- (7) If we take $A_i = k_i^{\gamma'}$, we get the nonlinear preferential attachment model of Krapivsky-Redner [28,29]. This network model undergoes a gelation phenomenon for $\gamma' > 1$. Namely, there is an emergence of a single dominant node linked to almost every other node. For $\gamma' > 2$, there is a finite probability that the dominating node is the first node of the network.
- (8) If we assign to each node a fitness value $\eta_i = e^{-\beta \epsilon_i}$, with ϵ_i drawn from a distribution $g(\epsilon) \propto \epsilon^\kappa$ and $\kappa > 0$, and we take $A_i = \eta_i k_i$, we get as a function of β the so-called Bose-Einstein condensation in complex networks of Bianconi and Barabási [30]. When this happens, for $\beta > \beta_c$ one node with high fitness is connected to a finite fraction of other nodes in the network.
- (9) If we take $A_i = (t - t_i)^{-\alpha} k_i$ where t_i indicates the time at which the node i has joined the network, we get the preferential attachment model with aging of the sites of Dorogovtsev-Mendes [31]. In this growing network the power-law exponent γ of the degree distribution is diverging as $\gamma \simeq \frac{1}{c_1} \frac{1}{1-\alpha}$ when $\alpha \rightarrow 1^-$. For $\alpha > 1$ the network is exponential and becomes more and more similar to a linear chain.

B. Entropy rate

The growing connected trees are fully determined by the sequence of symbol $\mathcal{X} = (i_1, i_2, \dots, i_N)$ where i_t is the node linked at time t to the node $i = t + 2$. In order to evaluate the entropy rate of growing networks it is sufficient to determine the entropy rate of the sequence (i_1, i_2, \dots, i_t) :

$$h(t, \mathcal{X}) = - \sum_{i_t} P(i_t | i_1, i_2, \dots, i_{t-1}) \log P(i_t | i_1, i_2, \dots, i_{t-1}), \quad (13)$$

where $P(i_t | i_1, i_2, \dots, i_{t-1})$ is the conditional probability that the node i_t is chosen at time t given the history of the process. The entropy of the process evaluating how many networks are typically constructed by the growing network process is $S(\{\mathcal{X}\})$

$$S(\{\mathcal{X}\}) = - \sum_{\{i_1, i_2, \dots, i_t\}} P(i_1, i_2, \dots, i_t) \log P(i_1, i_2, \dots, i_t). \quad (14)$$

C. Maximal and minimal bound to the entropy rate of growing network trees

It is instructive to study the limits of the entropy rate of connected growing trees. The minimal entropy rate is given by the entropy rate of the star or of the linear chain. Indeed, by taking $A_i = \delta_{1,i}$ we have that the entropy rate is zero. The growing network model becomes deterministic, and it gives rise to a unique star network with the center on the node $i = 1$. The entropy rate of a linear chain $A_i = \delta_{i,t+1}$ is also zero by a similar argument, and the model generates a unique linear chain network structure. On the other hand, the maximal entropy rate is given by the maximally random growing connected trees that is generated by taking $A_i = 1$ and $\Pi_i = 1/(t + 1)$. For this process the entropy rate is given by

$$h(t, \mathcal{X}) = \log(t + 1). \quad (15)$$

Therefore this entropy rate increases logarithmically with time, and the probability of each tree with $N = t + 2$ nodes is given by

$$P(N) = \frac{1}{(N - 1)!}. \quad (16)$$

Therefore $S(\{\mathcal{X}\}) = \log[(N - 1)!]$. This is the maximal entropy of a growing connected tree.

IV. GROWING TREES WITH STATIONARY DEGREE DISTRIBUTION

For growing network models with stationary degree distributions there are simple relations between $h(t, \mathcal{X})$ and $S(\{\mathcal{X}\})$. Let us define the entropy rate

$$H(\mathcal{X}) = \lim_{N \rightarrow \infty} \frac{1}{N} [S(\{\mathcal{X}\}) - \log[(N - 1)!]]. \quad (17)$$

For a growing tree network with stationary degree distribution, by the recursive application of chain rule $P(i_1, i_2, \dots, i_t) = P(i_t | i_1, i_2, \dots, i_{t-1}) P(i_1, i_2, \dots, i_{t-1})$ we can easily get

$$H(\mathcal{X}) = \lim_{N \rightarrow \infty} \frac{1}{N} \left\{ \sum_{n=1}^{N-2} h(n, \mathcal{X}) - \log[(N - 1)!] \right\}. \quad (18)$$

If the entropy rate of growing networks H is a constant, it means that the number of graphs generated by the growing network model has a dominating term that goes like $N!$ and a subleading term that is exponential with the number of nodes N . On the other hand, if $H = -\infty$ it means that the number of networks generated by the growing network model increases with the number of nodes in the network N at most exponentially. Usually the typical number of labeled networks generated by growing network models with convergent degree distribution is less than the number of networks with the same degree distribution. In order to evaluate the ratio between these two cardinalities, we introduce here the difference Δ between the Gibbs entropy $\Sigma[\{N_k\}]$ of the network with the same degree distribution and the entropy of the networks generated by the growing network model. Therefore Δ is

$$\Delta = \lim_{N \rightarrow \infty} \left\{ \Sigma[\{N_k\}] - \frac{1}{N} \log[(N - 1)!] - H(\mathcal{X}) \right\}. \quad (19)$$

The larger the value of Δ is, the smaller the fraction of networks generated by the growing model compared with the networks generated by the configuration model. This implies that the larger is Δ the more complex the networks generated by the growing models are. In fact, these networks need the dynamics of the networks to implicitly force the networks to satisfy a more stringent set of structural conditions beyond the degree distribution.

A. The entropy rate of the BA model

We consider the BA model, we take $A_i = k_i$ and $\Pi_i = \frac{k_i}{2(N-1)}$, and therefore

$$P(i_t | i_1, i_2, \dots, i_{t-1}) = \frac{k_{i_t}}{2(N-1)}. \quad (20)$$

The BA model, asymptotically in time has a degree distribution that converges to the value N_k given by

$$N_k = \frac{4N}{k(k+1)(k+2)}. \quad (21)$$

Therefore, asymptotically in time the entropy rate of the BA model is

$$h(t = N - 2, \mathcal{X}) \rightarrow - \sum_{k=1}^{\infty} N_k \frac{k}{2(N-1)} \log \left[\frac{k}{2(N-1)} \right]. \quad (22)$$

Hence, the entropy rate $h(\mathcal{X})$ increases in time as the logarithm of the number of nodes in the network, but it has a subleading term that is constant in time and depends on the degree sequence, i.e.,

$$\begin{aligned} h(t = N - 2, \mathcal{X}) &\rightarrow \log(N - 1) + \log(2) \\ &\quad - \sum_{k=1}^{N-1} \log(k) \frac{2}{(k+1)(k+2)} \\ &\rightarrow \log(N - 1) - 0.51(0). \end{aligned} \quad (23)$$

The entropy rate $H(\mathcal{X})$ in the limit $N \rightarrow \infty$ is therefore given by

$$H(\mathcal{X}) \simeq -0.51 \dots \quad (24)$$

We note here that the degree distribution N_k is known to have interesting finite size effects [2,34], in addition to the asymptotic scaling Eq. (21). Here we checked that the value of the entropy rate is not modified by these corrections up to the significant digit we have considered. Finally, in order to compare the number of networks generated by the BA model with the network that we can construct with the same degree sequence, we evaluate the value of Δ in the thermodynamic limit. This can be written as

$$\begin{aligned} \Delta &= \lim_{N \rightarrow \infty} \left\{ \frac{1}{N} \sum_{t=1}^{N-1} \log(N/t) \right. \\ &\quad - \frac{1}{2N} \sum_k N_k [k \log(k) + \log(2\pi k)] \\ &\quad \left. - \sum_k \frac{N_k}{N} \log \left(\frac{N_k}{N} \right) + 1 \right\} \simeq 0.9(1). \end{aligned} \quad (25)$$

B. The entropy rate of the growing network model with initial attractiveness

If we take $A_i = k_i - 1 + a$, the network generated is scale free with power-law exponent $\gamma = 2 + a$ [26]. The probability to choose the node i_t given the history of the process is therefore given by

$$P(i_t | i_1, i_2, \dots, i_{t-1}) = \frac{A_i}{\mathcal{N}} = \frac{k_i - 1 + a}{(a + 1)(N - 1)}. \quad (26)$$

Asymptotically in time the degree distribution for trees converges to the value [2,26]

$$N_k = N(1 + a) \frac{\Gamma(1 + 2a)\Gamma(k + a - 1)}{\Gamma(a)\Gamma(k + 1 + 2a)}. \quad (27)$$

From this, the entropy rate $h(t, \mathcal{X})$ is asymptotically

$$h(t = N - 2, \mathcal{X}) \rightarrow - \sum_{k=1}^{\infty} N_k \frac{k - 1 + a}{(a + 1)(N - 1)} \times \log \left[\frac{k - 1 + a}{(a + 1)(N - 1)} \right], \quad (28)$$

which can be simplified as

$$h(t = N - 2, \mathcal{X}) \rightarrow \log(N - 1) + \log(a + 1) - \sum_{k=1}^{N-1} \log \left[(k - 1 + a) \frac{\Gamma(1 + 2a)\Gamma(k + a)}{\Gamma(a)\Gamma(k + 1 + 2a)} \right]. \quad (29)$$

In the limit $N \rightarrow \infty$, the entropy rate $H(\mathcal{X})$ is

$$H(\mathcal{X}) = \log(a + 1) - \sum_{k=1}^{N-1} \log \left[(k - 1 + a) \frac{\Gamma(1 + 2a)\Gamma(k + a)}{\Gamma(a)\Gamma(k + 1 + 2a)} \right]. \quad (30)$$

When $a \rightarrow 1$, the solution reduces to the solution of the BA model. In Fig. 1 we plot the value of $H = H(\mathcal{X})$ versus a calculated by Eq. (30) using an upper cutoff for the degree $k_i < K \forall i = 1, \dots, N$. As the parameter $a \rightarrow 0$, the entropy rate decreases, indicating that the network model generates an exponentially smaller number of networks. Also the Gibbs entropy of scale-free networks decreases as long as the power-law exponent converges toward 2, i.e., in the limit $\gamma \rightarrow 2$. In order to evaluate the change in the ratio of networks generated by the growing network model to the number of possible networks with the same degree distribution, in Fig. 2 we plot Δ as a function of a . As $a \rightarrow 0$ and $\gamma \rightarrow 2$ the number of networks generated by the growing network models is a smaller function of the total number of networks that is possible to build with the same degree distribution. This is an indication and quantification of the importance of correlations generated by the growing network model with given initial attractiveness a .

C. The entropy rate of the Bianconi-Barabási fitness model

If the kernel A_i is heterogeneous and specifically given by $A_i = \eta_i k_i$, the model is called the Bianconi-Barabási fitness

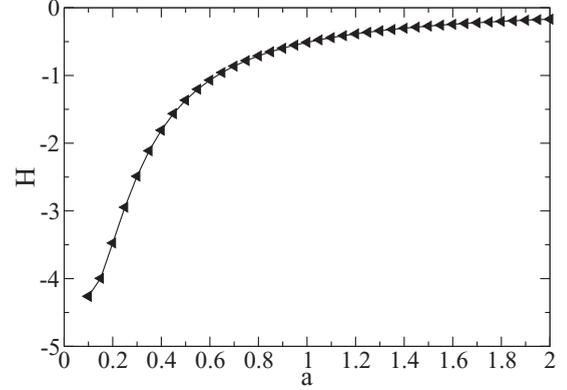


FIG. 1. The entropy rate H calculated for the growing network model with initial attractiveness [26] as a function of a and evaluated by Eq. (30) using a maximal degree equal to $K = 10^7$.

model [27]. The probability to choose the node i_t given the history of the process is therefore given by

$$P(i_t | i_1, i_2, \dots, i_{t-1}) = \frac{A_i}{\mathcal{N}} = \frac{\eta_i k_i}{\mu(N - 1)}, \quad (31)$$

where $\mu(N - 1) = \sum_{i=1}^{N-1} \eta_i k_i$. The degree distribution $N_k(\eta)$ for nodes of fitness η , asymptotically in time, converges to [5]

$$N_k(\eta) = \frac{N\mu\rho(\eta)}{\eta} \frac{\Gamma(k)\Gamma(1 + \mu/\eta)}{\Gamma(k + 1 + \mu/\eta)}, \quad (32)$$

where $\rho(\eta)$ is the distribution of η . Given the analytic solution of the model [5,27], μ is determined by the self-consistent relation

$$\int_0^{\eta_0} \rho(\eta)(\mu/\eta - 1)^{-1} d\eta = 1. \quad (33)$$

We consider here the case of uniform distribution of the fitness, i.e., $\rho(\eta) = 1$ with $\eta \in (0, 1)$. Therefore the entropy rate is

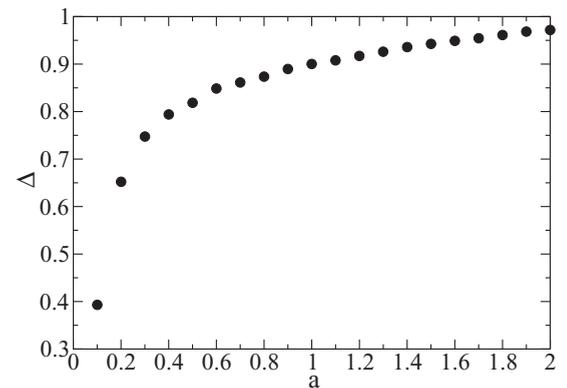


FIG. 2. The value of Δ calculated for the growing network model with initial attractiveness [26] as a function of a evaluated for networks of $N = 50000$ nodes and more than 20 realizations of the process.

given by

$$h(t = N - 2, \mathcal{X}) \rightarrow - \sum_{k=1}^{N-1} \int_0^1 N_k(\eta) \frac{\eta k}{\mu(N-1)} \log \left[\frac{\eta k}{\mu(N-1)} \right], \quad (34)$$

which gives

$$H(\mathcal{X}) = -1.59 \dots \quad (35)$$

D. Entropy rate for growing network models with structural phase transitions

We have measured the entropy rate for three growing network models showing a phase transition:

(1) The Krapivsky-Redner model [28,29] with

$$h(t, \mathcal{X}) = - \sum_{i=1}^t \frac{k_i^{y'}}{\mathcal{N}} \log \left(\frac{k_i^{y'}}{\mathcal{N}} \right) \quad (36)$$

and $\mathcal{N} = \sum_{i=1}^t k_i^{y'}$.

(2) The Bianconi-Barabási model showing a Bose-Einstein condensation in complex networks [30] with

$$h(t, \mathcal{X}) = - \sum_{i_i} \frac{e^{-\beta \epsilon_i} k_i}{\mathcal{N}} \log \left(\frac{e^{-\beta \epsilon_i} k_i}{\mathcal{N}} \right) \quad (37)$$

and $\mathcal{N} = \sum_{i=1}^t e^{-\beta \epsilon_i} k_i$.

(3) The Dorogovtsev-Mendes model with aging of the nodes [31] with

$$h(t, \mathcal{X}) = - \sum_i \frac{\tau_i^{-\alpha} k_i}{\mathcal{N}} \log \left(\frac{\tau_i^{-\alpha} k_i}{\mathcal{N}} \right), \quad (38)$$

where $\tau_i = t - t_i$ is the age of node i and $\mathcal{N} = \sum_{i=1}^t \tau_i^{-\alpha} k_i$.

In Fig. 3 the entropy rate $H(\mathcal{X})$ is calculated by numerical simulations using

$$H(\mathcal{X}) = \frac{1}{N} \left\{ \sum_{n=1}^{N-2} h(n, \mathcal{X}) - \log[(N-1)!] \right\} \quad (39)$$

for a network of sufficiently large size N for the three models as a function of the parameters γ , β , and α , respectively. We show that at the transition point the scaling of H evaluated for a network of size N changes from constant to N -dependent behavior. In particular we checked that in the three cases $H \propto \log(N)$ indicating that as the network grows the typical number of networks that are generated scales only exponentially with N (and not like $N!$).

This behavior signifies a disordered-ordered phase transition in the topology of the network. In the Bose-Einstein condensation network model and in the Krapivsky-Redner model, below the phase transition, the network is dominated by a hub node that grabs a finite fraction of the nodes. In the aging model, below the phase transition, the network develops a structure more similar to a linear chain.

V. CONCLUSION

In conclusion, we have studied growing network models and their entropy rate. We have seen that the entropy rate of growing simple trees have maximal and minimal bound, and we have studied the entropy rate of scale-free tree networks. This entropy rate allows us to calculate the number of typical graphs generated by growing scale-free network models and to quantify their complexity by comparing this number to the total number of graphs with the same degree distribution. Although we have focused on trees, the definition of entropy rate can be easily extended to growing network models with cycles. However, the probabilities of adding two or more links at a given time should explicitly account for the fact that the new links must be distinct, a fact that induces a small correction to the simple preferential attachment. We have analyzed a variety of growing network models, and we have studied nonequilibrium growing network models showing structural phase transitions. By numerical investigations, we have shown that when a growing network model has a phase transition, the entropy rate changes its scaling with the system size indicating the disorder-to-order transition. In the future we believe that an integrated view of information theory of complex networks will provide a framework to extend quantitative measures of complexity to a large variety of network structures, models, and dynamics. The present work is a step in this direction.

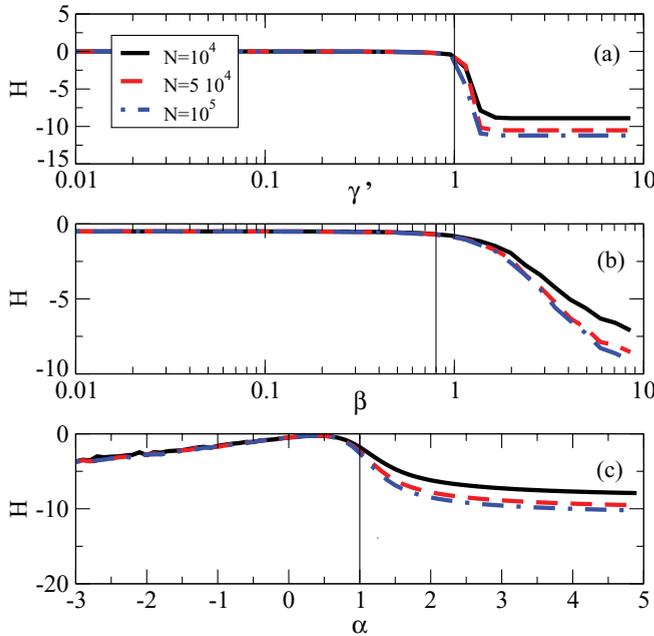


FIG. 3. (Color online) The entropy rate H is evaluated for the Krapivsky-Redner model [28,29] (A), for the “Bose-Einstein condensation in complex networks” of Bianconi-Barabási with $g(\epsilon) = 2\epsilon$, and $\epsilon \in (0,1)$, ($\kappa = 1$) [30] (B), and for the aging model [31] of Dorogovtsev-Mendes (C). The data are averaged over N_{run} different realizations of the network. We took $N_{\text{run}} = 100$ for simulations with $N = 10^4$ and $N_{\text{run}} = 30$ otherwise. Above the structural phase transition indicated with the solid line, the entropy rate H strongly depends on N .

- [1] R. Albert and A.-L. Barabási, *Rev. Mod. Phys.* **74**, 47 (2002).
- [2] S. N. Dorogovtsev and J. F. F. Mendes, *Evolution of Networks: From Biological Nets to the Internet and the WWW* (Oxford University Press, Oxford, 2003); M. E. J. Newman, *SIAM Rev.* **45**, 167 (2003).
- [3] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. U. Hwang, *Phys. Rep.* **424**, 175 (2006).
- [4] S. N. Dorogovtsev, *Lectures on Complex Networks* (Oxford University Press, Oxford, 2010).
- [5] M. E. J. Newman, *Networks: An Introduction* (Oxford University Press, 2010).
- [6] R. V. Solé and S. Valverde, in *Complex Networks*, edited by E. Ben-Naim, H. Frauenfelder, and Z. Toroczkai, Lecture Notes in Physics, Vol. 650 (Springer, Berlin, 2004), pp. 189–207.
- [7] G. Bianconi, *Europhys. Lett.* **81**, 28005 (2008).
- [8] G. Bianconi, A. C. C. Coolen, and C. J. Perez Vicente, *Phys. Rev. E* **78**, 016114 (2008).
- [9] G. Bianconi, *Phys. Rev. E* **79**, 036114 (2009).
- [10] K. Anand and G. Bianconi, *Phys. Rev. E* **80**, 045102 (2009).
- [11] K. Anand and G. Bianconi, *Phys. Rev. E* **82**, 011116 (2010).
- [12] A. Annibale, A. C. C. Coolen, L. P. Fernandes, F. Fraternali, and J. Kleinjung, *J. Phys. A* **42**, 485001 (2009).
- [13] E. S. Roberts, A. C. C. Coolen, and T. Schlitt, *J. Phys. A* **44**, 275002 (2011).
- [14] S. Johnson, J. J. Torres, J. Marro, and M. A. Muñoz, *Phys. Rev. Lett.* **104**, 108702 (2010).
- [15] G. Bianconi, P. Pin, and M. Marsili, *Proc. Natl. Acad. Sci. USA* **106**, 11433 (2009).
- [16] A. E. Teschendorff and S. Severini, *BMC Sys. Biol.* **4**, 104 (2010).
- [17] S. L. Braunstein, S. Ghosh, and S. Severini, *Ann. Comb.* **10**, 291 (2006).
- [18] S. L. Braunstein, S. Ghosh, T. Mansour, S. Severini, and R. C. Wilson, *Phys. Rev. A* **73**, 012320 (2006).
- [19] F. Passerini and S. Severini, in *Social Structures in Communication Networks*, edited by John Symons and Jorge Lou, IJATS, Vol. 1, Issue 4.
- [20] S. Garnerone, P. Giorda, and P. Zanardi, e-print [arXiv:1103.5989](https://arxiv.org/abs/1103.5989).
- [21] K. Anand, G. Bianconi, and S. Severini, *Phys. Rev. E* **83**, 036109 (2011).
- [22] J. Gómez-Gardeñes and V. Latora, *Phys. Rev. E* **78**, 065102(R) (2008).
- [23] R. Sinatra, J. Gómez-Gardeñes, R. Lambiotte, V. Nicosia, and V. Latora, *Phys. Rev. E* **83**, 030103(R) (2011).
- [24] Z. Burda, J. Duda, J. M. Luck, and B. Waclaw, *Acta Phys. Polon. B* **41**, 949 (2010).
- [25] A. L. Barabási and R. Albert, *Science* **286**, 509 (1999).
- [26] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, *Phys. Rev. Lett.* **85**, 4633 (2000).
- [27] G. Bianconi and A. L. Barabási, *Europhys. Lett.* **54**, 436 (2001).
- [28] P. L. Krapivsky, S. Redner, and F. Leyvraz, *Phys. Rev. Lett.* **85**, 4629 (2000).
- [29] P. L. Krapivsky and S. Redner, *Phys. Rev. E* **63**, 066123 (2001).
- [30] G. Bianconi and A. L. Barabási, *Phys. Rev. Lett.* **86**, 5632 (2001).
- [31] S. N. Dorogovtsev and J. F. F. Mendes, *Phys. Rev. E* **62**, 1842 (2000).
- [32] C. I. Del Genio, H. Kim, Z. Toroczkai, and K. E. Bassler, *PLoS ONE* **5**, e10012 (2010).
- [33] C. I. Del Genio, T. Gross, and K. E. Bassler, *Phys. Rev. Lett.* **107**, 178701 (2011).
- [34] C. Godreche, H. Grandclaude, and J. M. Luck, *J. Stat. Phys.* **137**, 1117 (2009).