

Approaching the thermodynamic limit in equilibrated scale-free networks

B. Waclaw,¹ L. Bogacz,² and W. Janke^{1,3}

¹*Institut für Theoretische Physik, Universität Leipzig, Postfach 100920, 04009 Leipzig, Germany*

²*Department of Information Technologies, Faculty of Physics, Astronomy and Applied Computer Science, Jagellonian University, Reymonta 4, 30-059 Kraków, Poland*

³*Centre for Theoretical Sciences (NTZ), Universität Leipzig, Emil-Fuchs-Straße 1, 04105 Leipzig, Germany*

(Received 28 July 2008; published 23 December 2008)

We discuss how various models of scale-free complex networks approach their limiting properties when the size N of the network grows. We focus mainly on equilibrated networks and their finite-size degree distributions. Our results show that the position of the cutoff in the degree distribution, k_{cutoff} , scales with N in a different way than predicted for $N \rightarrow \infty$; that is, subleading corrections to the scaling $k_{\text{cutoff}} \sim N^\alpha$ are strong even for networks of order $N \sim 10^9$ nodes. We observe also a logarithmic correction to the scaling for degenerated graphs with the degree distribution $\pi(k) \sim k^{-3}$. On the other hand, the distribution of the maximal degree k_{max} may have a different scaling than the cutoff and, moreover, it approaches the thermodynamic limit much faster. We argue that $k_{\text{max}} \sim N^{\alpha'}$ with an exponent $\alpha' = \min[\alpha, 1/(\gamma-1)]$, where γ is the exponent in the power law $\pi(k) \sim k^{-\gamma}$. We also present some results on the cutoff function and the distribution of the maximal degree in equilibrated networks.

DOI: 10.1103/PhysRevE.78.061125

PACS number(s): 05.50.+q, 05.70.-a, 89.75.-k

I. INTRODUCTION

Recent progress in understanding the structure and function of complex networks [1] has been largely influenced by the application of statistical methods of modern physics. The statistical mechanics of networks [2–8], if restricted to structural properties, deals with two classes of problems. In the first one, one considers networks being in a sort of equilibrium [2,4–6] where the concept of statistical ensembles applies in a natural way. Oppositely, growing networks are usually treated by rate-equation formalism [7,8]. Both approaches, however, share one common thing: the majority of considered models is analytically solvable only in the thermodynamic limit. This causes several difficulties. First of all, a question arises about the statistical equivalence of different ensembles of networks such as canonical or grand-canonical ensemble [2]. Second, considering dynamical processes on networks, one can ask how they are influenced by approaching the thermodynamic limit in different ways [9].

The third question, which we want to address in this paper, is how fast is the convergence towards the limiting values of some network features, for different models? This question is much more important than the analogous question in more traditional physics dealing with atoms or molecules, where the number of degrees of freedom is typically of order 10^{23} . Here, on the contrary, the largest observed system is the World Wide Web (WWW) with $N \sim 10^9$ – 10^{10} nodes. Many real networks are much smaller than this, typically having only 10^3 – 10^4 nodes. Therefore, finite-size corrections to solutions obtained in the thermodynamic limit are much stronger and in many situations cannot be neglected.

As a most prominent example we will consider the degree distribution $\pi(k)$, which is of most importance for characterizing the network as well as many dynamical processes taking place on it. Finite-size effects are especially strong for scale-free networks, for which $\pi(k) \sim k^{-\gamma}$ exhibits a power law. Examples are the Internet, the WWW and many social

and biological networks. Since the power law cannot extend to infinity for finite size N , the degree distribution $\pi(k)$ must decrease rapidly above some characteristic value $k = k_{\text{cutoff}}$. The cutoff scales typically as $k_{\text{cutoff}} \sim N^\alpha$, where the exponent α depends on the network type. For $2 < \gamma \leq 3$ when the second moment is divergent in the limit of large N , the cutoff appears either directly, or via moments of $\pi(k)$ in many applications, for instance, in percolation or infection spreading [10], transport models [11], synchronization processes, and others (see Ref. [12]). Because qualitative differences may show up for various values of α , it is important to know the proper scaling of k_{cutoff} with N .

For a growing network such as the Barabási-Albert (BA) model [13] or its generalization, the growing network with redirection (GNR) model [8], the scaling is well established [14,15]: for any $\gamma > 2$, the exponent $\alpha = 1/(\gamma-1)$. Moreover, it can be shown that for N large enough $\pi(k)$ can be approximated as $\pi_\infty(k)w(k/N^\alpha)$, where $\pi_\infty(k)$ is the degree distribution for $N \rightarrow \infty$ and $w(x)$ is some function independent of N . This approximation holds even for quite small networks of order 10^3 nodes because the convergence towards the limiting distribution is fast.

The situation is not so clear for equilibrated networks; that is, networks in which evolution is governed by rewiring of existing connections rather than by adding new nodes, and thus can be regarded as being in a sort of equilibrium. Although the scaling exponent α has been estimated for some models [16–19], the cutoff function $w(x)$ has not been determined so rigorously as for growing networks.

In this paper we will show that the predicted scaling is far from being true even for quite large networks. The paper is organized as follows. In Sec. II we discuss three models of equilibrated networks whose large- N behavior we want to study. Section III starts with basic concepts of how to calculate finite-size degree distributions and how to extract the behavior of the cutoff function. Then we consider two models, for which precise, semianalytical results are available for

sizes up to $N \sim 10^9$. In Sec. IV we discuss a more complicated model and show how to simulate it on a computer. Section V is devoted to a relation between the cutoff and the maximal degree. The paper is closed with a short summary in Sec. VI.

II. MODELS AND THEIR PROPERTIES IN THE THERMODYNAMIC LIMIT

We shall start from defining three different models of equilibrated networks, whose finite-size properties will be further examined. The word “equilibrated” means that networks (graphs in mathematical language) are maximally random under given constraints. These constraints are what defines the statistical ensemble of networks. The statistical ensemble consists of a set of states—graphs $\{g_i\}$, and a set of their statistical weights $\{W(g_i)\}$. This means that every graph g has a probability of occurrence proportional to $W(g)$. Every physical quantity X is then defined to be the average over the set of graphs: $\langle X \rangle = \sum_g W(g)X(g) / \sum_g W(g)$. Changing the set of graphs and/or the weights, one can obtain different models of random graphs. Three such ensembles will be considered in this paper and are defined below. In all cases, we assume that the graphs are undirected, have labeled nodes, fixed number of nodes N and links L , and that the weight $W(g)$ of every labeled graph has a product form

$$W(g) = \prod_{i=1}^N p(k_i), \quad (1)$$

where k_i is the degree of the i th node of graph g and $p(k)$ is some arbitrary node weight function. The product measure (1) turns out to be very convenient to obtain the desired degree distributions in the thermodynamic limit by tuning the function $p(k)$. In what follows, we will assume that all models have the same distribution $\pi(k)$ in the thermodynamic limit. Different models will be specified by restricting the set of possible graph shapes and/or $p(k)$.

Equilibrated simple graphs. We consider graphs without self- and multiple connections. Denoting by $\pi_\infty(k)$ the degree distribution for $N \rightarrow \infty$, we have

$$\pi_\infty(k) = \frac{p(k)}{k!} e^{A+Bk}, \quad (2)$$

as follows from Refs. [2,20]. The constants A, B are chosen to have $\sum_k \pi_\infty(k) = 1$ and $\sum_k k \pi_\infty(k) = 2L/N = \bar{k}$. Equation (2) can be used in two different ways. First, it tells us what $\pi_\infty(k)$ would be for given $p(k)$ and the average degree \bar{k} . Second, it allows us to calculate the weight $p(k)$, which will produce the desired degree distribution $\pi_{\text{des}}(k)$ in the thermodynamic limit. In the latter case, if N, L are chosen so that the average degree \bar{k} is equal to $\langle k \rangle = \sum_k k \pi_{\text{des}}(k)$, then $B=0$ and the choice $p(k) = \pi_{\text{des}}(k)k!$ implies that $\pi_\infty(k) = \pi_{\text{des}}(k)$.

Equilibrated multigraphs. We again assume the product weight (1), but now we also accept degenerated graphs; that is, graphs with multiple and self-connections. It can be shown [2,17,21] that the partition function of the system, being the sum over all configurations, assumes the form

$$Z(N, L) = \sum_{k_1=0}^{\infty} \dots \sum_{k_N=0}^{\infty} \frac{p(k_1)}{k_1!} \dots \frac{p(k_N)}{k_N!} \delta_{2L, k_1 + \dots + k_N}, \quad (3)$$

so it is equivalent to that of the balls-in-boxes model [22] or the zero-range-process model [23] with weights $p(k)/k!$. The same formula (2) as for simple graphs holds for the degree distribution in the limit $N \rightarrow \infty$.

Equilibrated trees. This ensemble consists of all labeled, connected tree graphs with N nodes. In addition, we assume that one node is distinguished by a “stem” attached to it, which is convenient from a mathematical point of view, but it does not change the large- N behavior. Such “planted” trees can be treated in a special way and a number of quantities can be calculated analytically [19,24]. For instance, in the thermodynamic limit the degree distribution is given by

$$\pi_\infty(k) = \frac{p(k)}{(k-1)!} e^{A+Bk}. \quad (4)$$

Note that we have $(k-1)!$ in the denominator, in contrast to Eq. (2). Now, the desired degree distribution $\pi_{\text{des}}(k)$ can be obtained only if $\sum_k k \pi_{\text{des}}(k) = 2$, because the average degree for trees approaches 2 for $N \rightarrow \infty$. When this criterion is fulfilled, then assuming $p(k) = \pi_{\text{des}}(k)(k-1)!$, one obtains $\pi_\infty(k) = \pi_{\text{des}}(k)$.

All these models share one common property: the degree distribution $\pi_\infty(k)$ in the thermodynamic limit is proportional to $p(k)$. Therefore, by choosing $p(k) \sim k!k^{-\gamma}$ (for graphs) or $\sim (k-1)!k^{-\gamma}$ (for trees), one can make these networks scale-free. In this paper we shall stick to the following choice for the degree distribution in all models

$$\pi_\infty(k) = \frac{(\gamma-1)\Gamma(2\gamma-3)}{\Gamma(\gamma-2)} \frac{\Gamma(k+\gamma-3)}{\Gamma(k+2\gamma-3)} \sim k^{-\gamma}, \quad (5)$$

for $k > 0$, and $\pi_\infty(0) = 0$. This is precisely the degree distribution $\pi_\infty(k)$ in the GNR model of a growing tree mentioned above. The average degree for this distribution is $\langle k \rangle = 2$, therefore the equilibrated trees will approach $\pi_\infty(k)$ in a natural way if $p(k) = \pi_\infty(k)(k-1)!$. As to equilibrated simple graphs and multigraphs, one then has to ensure that $p(k) = \pi_\infty(k)k!$ and $\bar{k} = 2L/N \rightarrow 2$ for $N \rightarrow \infty$, which can be simply done by assuming $L=N$. The purpose of choosing this particular distribution is that since finite-size effects in the GNR model are known [15], we can compare what happens if networks are equilibrated but have the same $\pi_\infty(k)$ as growing ones.

III. DEGREE DISTRIBUTION FOR A FINITE NETWORK

In the previous section we discussed the behavior of the three models in the thermodynamic limit. Now we shall ask what the degree distribution looks like for $N < \infty$. Assume that $\pi_N(k)$ and $\pi_\infty(k)$ are degree distributions for finite N and $N \rightarrow \infty$, respectively. It is convenient to write the finite-size distribution $\pi_N(k)$ as a product of $\pi_\infty(k)$ being N independent, and some cutoff function $w(N, k)$ depending explicitly on the size N ,

$$\pi_N(k) = \pi_\infty(k)w(N,k). \quad (6)$$

The function $w(N,k)$ is model dependent. In Ref. [15] it has been found that for the GNR model, moments of $w(N,k)$ scale as follows:

$$\mu_m \equiv \sum_k w(N,k)k^m \propto N^{\alpha(m+1)}[1 + O(N^{-\alpha})], \quad (7)$$

with the cutoff exponent α defined in the Introduction and equal to $\alpha = 1/(\gamma - 1)$. This means that for sufficiently large N the cutoff function depends effectively on a single rescaled variable $x = k/N^\alpha$,

$$w(N,k) \equiv w(k/N^\alpha). \quad (8)$$

For the GNR model, it is possible to find an explicit form of $w(x)$ for some values of γ . It is usually complicated, but it always has the large- x behavior

$$\ln w(x) \sim -x^{1/(1-\alpha)} = -x^\eta, \quad (9)$$

where we defined the exponent $\eta = 1/(1-\alpha)$. Formulas (7)–(9) were verified numerically [21] for N of order $10^3 - 10^4$, and an excellent agreement has been found. In this paper we ask: to what extent these relations are valid for equilibrated networks? In particular, if the scaling (8) still holds, the moments μ_m will behave as $\mu_m \sim N^{\alpha(m+1)}$ for very large networks. The convergence towards this asymptotic behavior can be, however, different from that of growing networks. Therefore, we can relax the constraint that the sub-leading term decays with the power α and generally expect the following large- N behavior:

$$\mu_m(N) \equiv aN^{\alpha(m+1)}(1 + bN^{-\beta}), \quad (10)$$

with some constants $a > 0$, $b \neq 0$, and $\beta > 0$, so that the sub-leading correction decays like $N^{-\beta}$ with β not necessarily equal to α , as it was in Eq. (7).

A. Multigraphs

We shall start from multigraphs. As we said, we assume that the average degree \bar{k} is chosen to ensure $\sum_k k \pi_\infty(k) = \bar{k}$. This means that for a given number of nodes N , the number of links $L = L(N)$ is fixed. In the case of the distribution (5), $L = N$. The partition function (3) becomes a function of N only and can be rewritten as

$$Z(N) = \oint \frac{dz}{2\pi i} z^{-1-N\bar{k}} F^N(z), \quad (11)$$

where the contour of integration encircles zero and $F(z)$ denotes the generating function for $\pi_\infty(k)$,

$$F(z) = \sum_{k=0}^{\infty} \pi_\infty(k) z^k. \quad (12)$$

The degree distribution for finite N can be calculated as follows [24]:

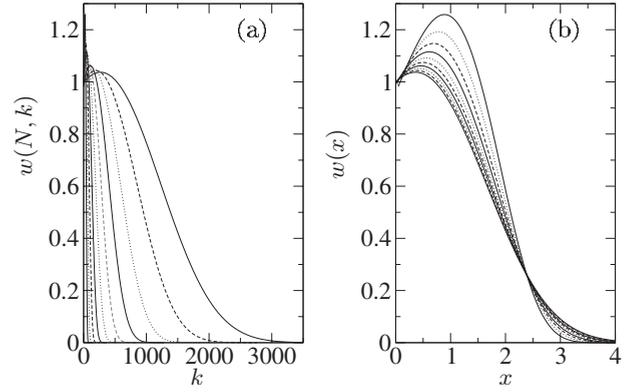


FIG. 1. (a) Plots of $w(N,k)$ for multigraphs of sizes between 100 (the narrowest peak) and 51 200 (the widest peak). (b) The same data but plotted against the rescaled variable $x = k/(N \ln N)^{1/2}$. The highest peak shows data for $N = 100$.

$$\begin{aligned} \pi_N(k) &= \frac{p(k)}{NZ(N)} \frac{\partial Z(N)}{\partial p(k)} = \pi_\infty(k) \frac{\oint \frac{dz}{2\pi i} z^{k-1-N\bar{k}} F^{N-1}(z)}{\oint \frac{dz}{2\pi i} z^{-1-N\bar{k}} F^N(z)} \\ &\equiv \pi_\infty(k)w(N,k), \end{aligned} \quad (13)$$

where the cutoff function $w(N,k)$ is given by the above ratio of contour integrals. The integrals cannot be, in general, performed analytically for arbitrary N . However, the integration can be easily done numerically for different distributions $\pi_\infty(k)$ and sizes N as follows. We choose the contour of integration to be $z = re^{i\phi}$, with $\phi \in (-\pi, \pi)$ and r smaller than the radius of convergence of $F(z)$. Because the integrated function becomes concentrated around zero for large N , one does not need to integrate over the whole range of ϕ . We have written a procedure in MATHEMATICA, which finds the value of r , for which the integrated function has the broadest maximum at $\phi = 0$. This speeds up the convergence of numerical integration. Then, the procedure searches for ϕ_{\max} for which the function falls to a sufficiently small part (typically 10^{-10}) of its maximal value. Then the function is integrated by means of the adaptive method over the range $(-\phi_{\max}, \phi_{\max})$. Let us consider first the distribution (5) for $\gamma = 3$, when it reduces to

$$\pi_\infty(k) = \frac{4}{k(k+1)(k+2)}, \quad (14)$$

for which

$$F(z) = \frac{z(3z-2) - 2(z-1)^2 \ln(1-z)}{z^2}, \quad (15)$$

so that the radius of convergence is one. Using the above numerical procedure we have calculated $w(N,k)$ for $N = 100, 200, \dots, 51\,200$, shown in Fig. 1(a). The theoretical value of α , predicted for this model, should be $1/2$ [17]; the same value comes from the correspondence to the zero-range process [25]. Therefore, for large N , plots of $w(N,k)$ for different sizes should collapse into a single curve in the rescaled variable $x = k/N^\alpha = k/N^{1/2}$. To check how fast the scal-

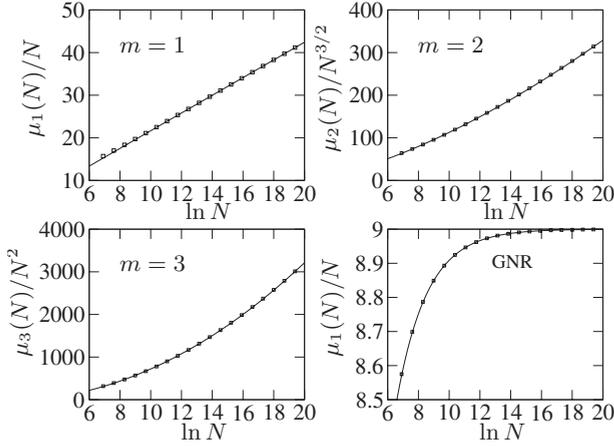


FIG. 2. Plots of $\mu_m(N)/N^{(m+1)/2}$ calculated from Eq. (16) (squares), for degenerated graphs with $\gamma=3$ and for the GNR tree network with the same $\pi_\infty(k)$ for comparison. Solid lines display solid functions: Eq. (10) for lower-right panel, and Eq. (18) for other panels, with parameters fitted to data points.

ing is approached when N grows, instead of using the cutoff function directly, it is better to consider its moments μ_m , which can be calculated as follows:

$$\mu_m(N) = \frac{\oint \frac{dz}{2\pi i} z^{-1-N\bar{k}} F^{N-1}(z) f_m(z)}{\oint \frac{dz}{2\pi i} z^{-1-N\bar{k}} F^N(z)}, \quad (16)$$

where the auxiliary function $f_m(z)$ is defined as

$$f_m(z) = \left(z \frac{d}{dz} \right)^m \frac{1}{1-z}. \quad (17)$$

We evaluated numerically the integrals in Eq. (16) using the same procedure as for the cutoff function, for $m=1, 2, 3$ and sizes $N=(1, 2, \dots, 2^{19}) \times 1000$. The maximal value of N is $\approx 5 \times 10^8$. In Fig. 2 we show plots of $\mu_m(N)/N^{(m+1)/2}$, where the subleading behavior has been exposed by dividing the data by the asymptotic form of $\mu_m \sim N^{\alpha(m+1)}$. Since N spans several orders of magnitude, the horizontal axis has a logarithmic scale in order to make all points visible. The first moment $\mu_1(N)/N$ for the GNR model of the growing tree network, calculated as in Ref. [15], is also plotted for comparison. It approaches its maximal value as $\sim N^{-0.5}$, which agrees with Eq. (7). But the moments for the degenerated graph show quite distinct behavior, namely, they grow with $\ln N$ linearly or faster. This means that the subleading, powerlike term $N^{-\beta}$ from Eq. (10) might turn into a leading behavior, which takes place for $\beta \rightarrow 0$ and leads to a multiplicative, logarithmic correction. To check this, we fitted the formula

$$\mu_m(N)/N^{(m+1)/2} = A_m (\ln N)^{C_m} + B_m \quad (18)$$

to data points in Fig. 2. The fits are represented there as solid lines. The ratio of obtained values $C_1:C_2:C_3 = 1.07:1.54:2.08 \approx 1:1.5:2$ suggests that the moments behave as

$$\mu_m(N) \sim (N \ln N)^{(m+1)/2}. \quad (19)$$

In other words, the cutoff scales as $\sim (N \ln N)^{1/2}$. This is also confirmed in Fig. 1(b) where the cutoff function is plotted in the rescaled variable $x=k/(N \ln N)^{1/2}$. According to our knowledge, this logarithmic correction has not been observed before. Its origin cannot lie in the distribution $\pi_\infty(k)$ only, because the growing networks with the same $\pi_\infty(k)$ do not have it, but it is the property of equilibrated graphs. Indeed, one can predict this scaling analytically, studying the cutoff function

$$w(N, k) \equiv W(N, k)/W(N, 0), \quad (20)$$

where

$$W(N, k) = \oint \frac{dz}{2\pi i} z^{k-1-N\bar{k}} F^N(z), \quad (21)$$

with $\bar{k}=2$ and $F(z)$ given by Eq. (15). Following the lines of Sec. VI from Ref. [25] one can argue that the function under the integral is localized around $z=1$ and can be expanded at this point. Choosing the contour of integration $z=e^{i\phi}$ we obtain

$$W(N, k) \equiv \int_{-\epsilon}^{\epsilon} \frac{d\phi}{2\pi} \exp \left\{ N \left[\frac{ik\phi}{N} + \phi^2(3 - i\pi + \ln \phi^2) \right] \right\}, \quad (22)$$

with ϵ small enough. Inserting now $k=x(N \ln N)^{1/2}$ and $\phi = t(N \ln N)^{-1/2}$ we have

$$W(N, x(N \ln N)^{1/2}) \equiv (N \ln N)^{-1/2} \int \frac{dt}{2\pi} e^{itx - t^2 + \dots}, \quad (23)$$

where higher terms are of order $\ln \ln N / \ln N$ and can be neglected for large N . Then the above formula becomes a Gaussian integral and hence the cutoff function reads

$$w(N, k) \approx \exp \left[-\frac{1}{4} \left(\frac{k}{(N \ln N)^{1/2}} \right)^2 \right], \quad (24)$$

with the predicted scaling.

Let us now go to $\gamma > 3$. In this region, the exponent α should again be $1/2$ [25]. We calculated the first three moments $\mu_m(N)$ for $\gamma=3.5$ from Eq. (16) for the same sizes N as previously. Next, we fitted formula (10) to the data points. The best fit gives $\beta \approx 0.24$ for all three moments. This value is significantly greater than zero, thus the correction is now clearly of the powerlike type. But the exponent $\beta \approx 1/4$ is small enough to produce large corrections even for moderate-size networks. For example, if $N=10^4$, the correction is of order $0.1b$, with b being usually much larger than 1. But β grows with γ , thus the convergence to the thermodynamic limit becomes faster, e.g., for $\gamma=4$ we have estimated $\beta > 0.4$.

We also investigated the range $2 < \gamma < 3$. Again, we calculated moments for different values of N , for $\gamma=2.5$ and fitted the formula (10). The theory [25] predicts $\alpha=1/(\gamma-1)=2/3$ for this case. The agreement between the data and Eq. (10) is very good for $\beta \approx 0.33$.

At last, in order to check whether the formula (9) for the large- x behavior of the cutoff function holds for multigraphs, we fitted a function $A + \ln(1+Bx) - (x/C)^D$ to $\ln w(x)$ obtained from numerical integration for $N=100, 200, 51\ 200$, for $\gamma=2.5, 3, 3.5$. The assumed form of $w(x)$ approximates the measured cutoff functions very well. The D 's obtained for different sizes N tend to some limiting values, which we found to be $D_{N \rightarrow \infty} = 2.0$ for $\gamma=2.5$, 1.9 for $\gamma=3$, and 3.1 for $\gamma=3.5$, with uncertainties of order 0.1. These values are in good agreement with the exponent $\eta = 1/(1-\alpha)$ from Eq. (9), which gives 2, 2, and 3, respectively, and with theoretical results for the equivalent model in Ref. [25].

To summarize our findings for the degenerated graphs: For $\gamma=3$ we have observed logarithmic corrections to the scaling $k_{\text{cutoff}} \sim N^\alpha$. The point $\gamma=3$ is the critical one; for $\gamma \neq 3$, corrections are powerlike, with the exponent approaching zero for $\gamma \rightarrow 3$. The corrections are thus very strong for $\gamma \approx 3$, even for very large networks. We have also examined the large- x behavior of the cutoff function $w(x)$ and showed that it agrees with Eq. (9) derived for growing networks. This suggests some universality, but we will see later that it does not hold for simple graphs.

B. Equilibrated trees

In Ref. [24], the partition function $Z(N)$ for equilibrated trees is found to be

$$Z(N) = \oint \frac{dz}{2\pi i} z^{-N-1} Z_{\text{GC}}(z). \quad (25)$$

Here $Z_{\text{GC}}(z)$ is a grand-canonical partition function obeying the equation

$$Z_{\text{GC}}(z) = z\tilde{F}(Z_{\text{GC}}(z)), \quad (26)$$

with

$$\tilde{F}(z) = \sum_{k=0}^{\infty} \pi_{\infty}(k+1)z^k, \quad (27)$$

assuming that in the thermodynamic limit we want to obtain the degree distribution $\pi_{\infty}(k)$. With help of Eq. (26), the partition function (25) can be rewritten as

$$Z(N) = \oint \frac{dz}{2\pi i} z^{-N} \tilde{F}^{N-1}(z) [\tilde{F}(z) - z\tilde{F}'(z)]. \quad (28)$$

Because we assume $\pi_{\infty}(0)=0$, the function $\tilde{F}(z)$ equals $\frac{1}{z}F(z)$, with $F(z)$ as in Eq. (12) for multigraphs. This in turn leads to the following formula for $Z(N)$:

$$Z(N) = \frac{1}{N} \oint \frac{dz}{2\pi i} z^{-2N} F^N(z), \quad (29)$$

which is, up to a factor $1/N$, equivalent to the partition function (11) for multigraphs with $\bar{k}=2-1/N$. This proves the equivalence between equilibrated multigraphs and trees, provided that the average degree is properly tuned, which has been already reported [22]. In our case, when $L=N$ for multigraphs and therefore $\bar{k}=2$, a small difference in the values

of moments μ_m arise. This does not, however, change the fact that in the thermodynamic limit both models become fully equivalent, and that for any finite N the difference can be neglected.

This means that the degree distribution for equilibrated trees behaves exactly as for multigraphs; that is, we again have logarithmic corrections to the moments $\mu_m(N)$ for $\gamma=3$, and power-law corrections for $\gamma \neq 3$. This indicates also that the cutoff exponent $\alpha=1/2$ for $\gamma \geq 3$ and $\alpha=1/(\gamma-1)$ for $\gamma < 3$.

IV. MULTICANONICAL SIMULATIONS OF SIMPLE GRAPHS

So far we have considered equilibrated multi- and tree graphs. But real-world networks are usually simple graphs; that is, they do not have multiple and self-connections, and have loops. Unfortunately, for these reasons simple graphs are not accessible with the technique used before, because one does not know how to write the partition function as a single contour integral. Therefore, to obtain $\pi_N(k)$ one needs to turn to Monte Carlo (MC) techniques. Let us shortly describe here the general method, which serves for this purpose. Details can be found elsewhere [4,26]. We simulate graphs with a fixed number of nodes N and links L . Each new graph g_{t+1} is generated from the previous one g_t by rewiring a single link. The move is accepted with the Metropolis probability

$$P(g_t \rightarrow g_{t+1}) = \min \left\{ 1, \frac{W(g_{t+1})}{W(g_t)} \right\}, \quad (30)$$

which ensures that graphs are generated with correct weights $W(g)$ from Eq. (1). In order to restrict to simple graphs we reject moves introducing self- or multiple connections. This method allows for estimating $\pi(k)$ with good accuracy for N of order thousands, which is, however, too small for our purpose. Larger networks are not accessible in a reasonable computer time. To show this, assume that $\pi(k) \sim k^{-\gamma}$ and that we start from a network with all k 's much smaller than $k_{\text{cutoff}} \sim N^\alpha$. We assume that to obtain the experimental distribution $\pi_{\text{exp}}(k)$ with an acceptable level of noise, in the course of simulation each node has to change its degree between 1 and k_{cutoff} many times. Let us denote by T the first-passage time from $k=1$ to $k=k_{\text{cutoff}}$. Our algorithm changes k by ± 1 every step, so the process can be treated as a biased random walk in a potential $V(k) = -\ln \pi(k)$. Then T can be estimated as $T \sim N^{\alpha(\gamma+1)}$ for $\gamma > 2$, which is nothing more than the Arrhenius law $T \sim \exp[V(k_{\text{cutoff}})]$ with an additional correction factor.

If, however, one could "flatten" $\pi(k)$ by appropriate reweighting of graphs, then instead of a biased random walk one would deal with a simple random walk where $T \sim N^{2\alpha}$. This is always better than the previous estimate for the interesting range of γ . Moreover, by increasing the probability of graphs with $k > k_{\text{cutoff}}$, one could measure $\pi_{\text{exp}}(k)$ far above the cutoff. Here comes the idea of multicanonical simulations (MUCA) [27]; a similar idea has already been applied to graphs [28]. To apply MUCA in our case, we modify the weight function from Eq. (1) to

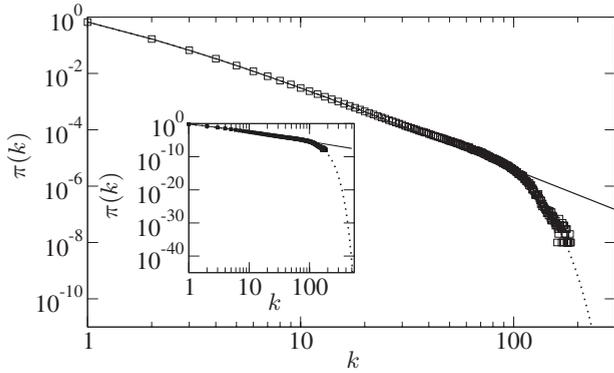


FIG. 3. Example of $\pi(k)$ for simple MC (squares) and MUCA (dotted line) for simple graphs with $N=10^3$ and $\gamma=3$. The solid line shows $\pi_\infty(k)$. Inset: The same for a broader range of k . MUCA simulation allows one to estimate the tail of $\pi_N(k)$ much better than simple MC.

$$W(g) = \frac{r(k_1)}{\pi_\infty(k_1)} \prod_{i=1}^N p(k_i), \quad (31)$$

where $r(k_1)$ is some function depending on the degree of one node, say the first one, and is chosen so that the distribution $\pi_{\text{flat}}(k)$, measured now only for the first node, is flat. The factor $1/\pi_\infty(k_1)$ gives an additional advantage of flattening the distribution at no cost for $k \ll k_{\text{cutoff}}$. To obtain the true distribution $\pi_{\text{exp}}(k)$ one multiplies $\pi_{\text{flat}}(k)$ by $\pi_\infty(k)/r(k)$. Since all nodes are statistically equivalent in equilibrated graphs, this procedure must give the same result as standard MC simulations. The only question is to find the optimal $r(k)$. In our simulations, we applied the iterative method from Ref. [29], which calculates $r(k)$ “on the fly” from sampled data. Because changing $r(k)$ during the simulation violates detailed balance and thus can change statistical weights, after obtaining a sufficiently flat histogram of

$\pi_{\text{flat}}(k)$ we fix $r(k)$ and perform a simple MC run with $W(g)$ given by Eq. (31). To speed up the simulation we divide the whole range of k , for which we want to determine $\pi(k)$, into overlapping subranges of size ≈ 50 and then glue results. We validated the method for graphs by comparing it with direct MC simulations for $N \leq 1000$ (see Fig. 3) and for multigraphs by comparing with the exact results from Sec. III A.

To check how the cutoff scales with N , we simulated graphs with $N=625, 1250, 2500, \dots, 40\,000$ nodes, and $\gamma=2.5, 3, 3.5$, for the same $\pi_\infty(k)$ as before. In each case we performed 10^{10} to 10^{11} MC steps (in units of single-link rewirings) for every subrange of k , after fixing $r(k)$. Each simulation was repeated three to four times in order to estimate statistical errors. We then calculated the moments μ_m of the cutoff function and checked if one can fit a simple power law $\mu_m \cong aN^{\alpha(m+1)}$ to data points. The results are shown in Table I, left, together with reduced χ^2 values and significance levels (“ Q ” values) for the fits. We see that none of these cases can be accepted with the confidence level 0.05, which is commonly used in hypotheses testing. On the other hand, if we assume a more sophisticated form of $\mu_m(N)$ with the subleading correction from Eq. (10) and, according to Ref. [16], with $\alpha(\gamma \geq 3) = 1/2, \alpha(\gamma = 2.5) = 1/(5-\gamma) = 0.4$, the agreement is very good [see Table I, right]. We thus conclude that the subleading term plays an important role for moderate sizes. In Fig. 4 we show plots of $\mu_1(N)$ and both fitted functions. The subleading correction coefficient b is very large. For instance, for $\gamma=3.5$ the correction is of order 0.1 for $N \sim 10^6$, much larger than for multigraphs, perhaps due to much stronger structural constraints. Therefore, simple graphs approach the thermodynamic limit even slower than equilibrated trees or degenerated graphs.

Finally, we examined the large- x behavior of $w(x)$ using the same method as for multigraphs. We found that the exponent η in Eq. (9) is not $1/(1-\alpha)$ as for multigraphs but approximately 2.0 for all values of γ . This agrees with the results of Ref. [16] that the cutoff is always Gaussian in

TABLE I. Results of fitting of the two different power laws (without and with a correction) to the measured moments μ_m for simple graphs. For each γ , three consecutive rows present results for $m=1, 2, 3$. “Parameters” indicate the set of free parameters during the fitting procedure; Q is the significance level of the fit. If α is marked as “fixed,” its value is 0.4 for $\gamma=2.5$ and 0.5 for the two other cases.

γ	$\mu_m = aN^{\alpha(m+1)}$ parameters a, α			$\mu_m = aN^{\alpha(m+1)}(1 + bN^{-\beta})$ parameters a, b, β , fixed α			
	α	Reduced χ^2	Q	β	b	Reduced χ^2	Q
2.5	0.366	4.5	<0.001	0.39	5.4	2.40	0.05
	0.373	3.5	0.004	0.42	8	1.99	0.09
	0.376	2.9	0.013	0.44	11	1.72	0.14
3	0.392	2.64	0.021	0.34	30	0.94	0.44
	0.400	2.41	0.034	0.40	80	0.87	0.48
	0.405	2.34	0.039	0.47	200	0.86	0.49
3.5	0.409	6.1	<0.001	0.38	22	1.09	0.36
	0.420	6.8	<0.001	0.48	56	0.73	0.57
	0.428	9.2	<0.001	0.56	125	0.51	0.72

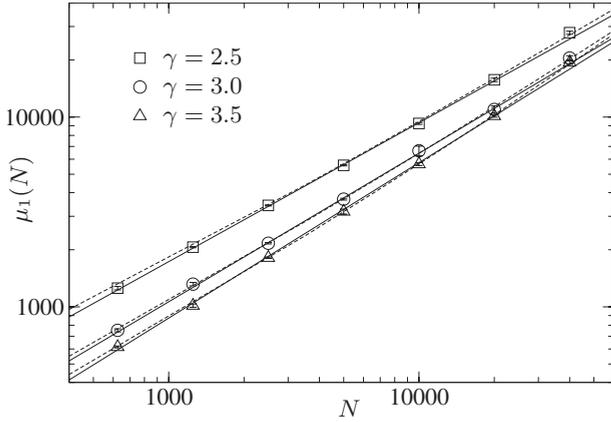


FIG. 4. Plots of $\mu_1(N)$ for simple graphs with $\gamma=2.5, 3, 3.5$. Solid and dotted lines are fits from Table I, left and right, respectively.

simple graphs, but it means also that η is not universal.

V. CUTOFFS AND THE DISTRIBUTION OF MAXIMAL DEGREE

So far we have studied the cutoffs in the degree distribution averaged over the ensemble of random graphs of a certain kind. The value $k_{\text{cutoff}} \sim N^\alpha$ tells where the power law ends and $\pi_N(k)$ starts to fall off rapidly. One can also pose the question: how does the maximal degree, k_{max} , scale with N ? Let α and α' be exponents in the power laws $k_{\text{cutoff}} \sim N^\alpha$ and $k_{\text{max}} \sim N^{\alpha'}$, respectively. The exponents α, α' may, but do not have to be equal, which has not been emphasized strongly enough in the literature on complex networks. To show this, let us consider the hypothetical network where all degrees k_1, \dots, k_N are independent random variables taken from the distribution $p(k) \sim k^{-\gamma}$. The degree distribution $\pi(k)$ is then equal to $p(k)$. In a real, finite network, degrees are (at least slightly) correlated due to structural constraints and $\pi(k)$ has a cutoff. In our simplified model, we can mimic these constraints by introducing the cutoff explicitly: $p(k) \sim k^{-\gamma} w(\frac{k}{N^\alpha})$, with some exponent α . Let us focus on a particular form of the cutoff,

$$w(x) = \exp(-ax^\eta), \quad (32)$$

with parameters a, η . This form of the cutoff function seems to be very general: In the case of growing networks it is an asymptotic behavior for large k , we also checked numerically that it holds for equilibrated graphs. For large N , the distribution $P(k_{\text{max}})$ of the maximal degree,

$$P(k_{\text{max}}) \approx Np(k_{\text{max}}) \left(\int_1^{k_{\text{max}}} p(k) dk \right)^{N-1}, \quad (33)$$

takes its maximal value at k_{max}^* being a solution to

$$p'(k_{\text{max}}^*) \int_1^\infty p(k) dk \equiv -Np^2(k_{\text{max}}^*), \quad (34)$$

which for the assumed cutoff gives

TABLE II. Values of exponents α, η , and scaling of k_{max}^* for degenerated and simple graphs, for different power laws $\pi(k) \sim k^{-\gamma}$. Assumed scalings are $k_{\text{cutoff}} \sim N^\alpha$, $\ln w(x) \sim -x^\eta$. The exponents α, η have been confirmed numerically in the previous section. The formulas for k_{max}^* are obtained from Eq. (35).

γ	Multigraphs			Simple graphs		
	α	η	k_{max}^*	α	η	k_{max}^*
(2,3)	$\frac{1}{\gamma-1}$	$\frac{\gamma-1}{\gamma-2}$	$N^{1/(\gamma-1)}$	$\frac{1}{5-\gamma}$	2	$N^{1/(5-\gamma)}(\ln N)^{1/2}$
3	$1/2 + \ln$	2	$N^{1/2}$	$1/2$	2	$N^{1/2}$
(3, ∞)	$1/2$	2	$N^{1/(\gamma-1)}$	$1/2$	2	$N^{1/(\gamma-1)}$

$$k_{\text{max}}^* \sim \begin{cases} N^{1/(\gamma-1)}, & \alpha \geq \frac{1}{\gamma-1} \\ N^\alpha (\ln N)^{1/\eta}, & \alpha < \frac{1}{\gamma-1}. \end{cases} \quad (35)$$

This means that, modulo the logarithmic correction, $\alpha' = \min(\alpha, 1/(\gamma-1))$. One can also show that $P(k_{\text{max}})$ for $\alpha > 1/(\gamma-1)$ is given by the Fréchet extreme value distribution: $P(x) \propto x^{-\gamma} e^{-x^{-\gamma+1}}$, while it approaches the Gumbel distribution: $P(x) \propto e^{-x-e^{-x}}$ for $\alpha < 1/(\gamma-1)$, in the properly rescaled variable $x = A + Bk_{\text{max}}$. Because these two distributions are narrow, k_{max}^* approximates also the mean value $\langle k_{\text{max}} \rangle$. We can learn two things from Eq. (35). First, k_{max}^* scales as a pure power of N with the same exponent as for k_{cutoff} , $\alpha' = \alpha$, only if $\alpha = 1/(\gamma-1)$. The distribution of k_{max} is then neither Fréchet nor Gumbel, but has a more complicated form, as follows from Eq. (33),

$$P(x) \equiv \mathcal{N} x^{-\gamma} w(x) \exp \left[-\mathcal{N} \int_x^\infty y^{-\gamma} w(y) dy \right], \quad (36)$$

where \mathcal{N} is a normalization coefficient and $x = k_{\text{max}}/N^{1/(\gamma-1)}$, and depends on the exact form of $w(x)$. Second, α' may be smaller than α ; that is, the maximal degree grows slower than the cutoff. At first sight this may appear counterintuitive. One could have expected that $\pi(k)$ for large k is dominated by the maximal degree distribution and thus k_{max} must not grow slower than k_{cutoff} .

The formula for k_{max}^* presented above works surprisingly well for real graphs where the cutoff stems from correlations between nodes degrees. In order to check Eq. (35), we performed multicanonical simulations and obtained distributions $P(k_{\text{max}})$ for multigraphs and simple graphs with $\gamma=2.5, 3$, and 3.5 . In light of what has been said in Sec. III B, the scaling of k_{max} for trees has to be identical to that for multigraphs, so it is not necessary to consider trees separately. The procedure was similar to that described in the previous section, with the only difference that we flatten $P(k_{\text{max}})$ and not $\pi(k)$. In each simulation we performed around 10^8 Monte Carlo steps after fixing the weight $r(k_{\text{max}})$.

Let us discuss first multigraphs. In Table II, left, we

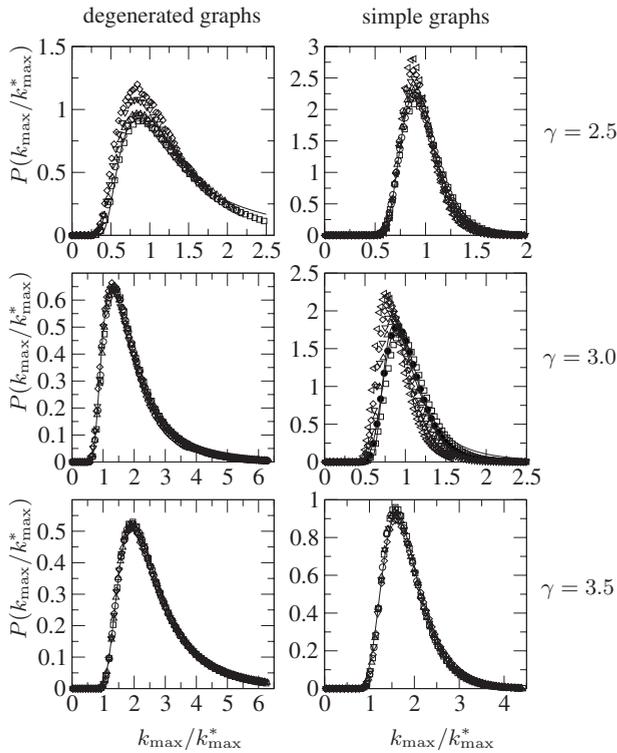


FIG. 5. Left: Rescaled plots of $P(k_{\max})$ for degenerated graphs, for $\gamma=2.5$ (top), 3 (middle), and 3.5 (bottom picture), and for different sizes ranging from $N=1000$ up to 32 000 (1000: square; 2000: circle; 4000: triangle up; 8000: triangle down; 16000: diamond; 32000: triangle left). Solid lines are Fréchet distributions $P(x) \propto x^{-\gamma} e^{-x^{-\gamma+1}}$ fitted to data points for one chosen size N (usually $N=4000$). Right: The same for simple graphs, lines are now: top and bottom—Gumbel; middle—Gumbel (thin line) and Fréchet (thick line), fitted to black circles ($N=2000$).

present values of α, η for various γ , which we confirmed numerically in previous sections. The fourth column shows scaling laws for k_{\max}^* predicted by means of Eq. (35). In Fig. 5 we plotted the experimental distributions $P(k_{\max})$, in the rescaled variable $x = k_{\max}/k_{\max}^*$. Plots for different N show good agreement of positions of the maximum. The distribution for $\gamma=3.5$ is also approximately Fréchet, which agrees with recent findings [30], the two for $\gamma=2.5, 3$ deviate slightly from Fréchet in the tail. The perfect scaling of k_{\max} means that the thermodynamic limit for the distribution of maximal degree is approached much faster than for $\pi(k)$, and

is reached already for $N \sim 10^4$ nodes; no subleading corrections are necessary.

We repeated the same procedure for graphs [see Table II, right, and Fig. 5]. Again, one sees very good agreement, except for $\gamma=3$, where one obtains better collapse for a slightly different value $k_{\max}^* \sim N^{0.45}$. The shape of $P(k_{\max})$ is better approximated by Gumbel than Fréchet distribution. Also for $\gamma=3.5$ and $N \leq 32\,000$, the distribution $P(k_{\max})$ deviates from Fréchet and is closer to Gumbel.

To conclude, we have shown that the distribution of the maximal degree exhibits the scaling $k_{\max} \sim N^{\alpha'}$ but, in general, with a different exponent than that in the cutoff. The large- N scaling is, however, approached much faster than in case of the degree distribution and subleading corrections can be safely neglected in most cases for networks of order 10^4 nodes. It is also worth mentioning that α' depends on both the exponents γ and α .

VI. SUMMARY

In this paper we considered finite-size effects in the degree distribution of equilibrated networks. We showed that the convergence towards the thermodynamic limit is very slow, thus in order to get reasonable estimation of the cutoff one has to consider subleading corrections to the scaling $k_{\text{cutoff}} \sim N^{\alpha}$. For multigraphs and $\gamma=3$, the correction turns into a leading behavior $\sim (N \ln N)^{1/2}$ and therefore does not vanish for $N \rightarrow \infty$. This has to be taken into account when comparing any numerical results for finite networks to that derived analytically for the leading behavior only. We checked also that the asymptotic behavior of the cutoff function is very simple: $\ln w(x) \sim -x^{\eta}$, but η is not universal among different classes of graphs. We argued that the maximal degree k_{\max} reaches the asymptotic power-law scaling faster than the cutoff but scales in many cases differently than k_{cutoff} .

The networks presented here do not have explicit degree-degree correlations. It would be interesting to check which of the results will survive the introduction of correlations, i.e., are universal to some extent.

ACKNOWLEDGMENTS

B.W. thanks A. Krzywicki and S. N. Majumdar for discussions. We are particularly indebted to Z. Burda for many useful comments. We also thank the EC-RTN Network “EN-RAGE” under Grant No. MRTN-CT-2004-005616 and the Alexander von Humboldt Foundation for support.

- [1] R. Albert and A.-L. Barabási, *Rev. Mod. Phys.* **74**, 47 (2002); S. N. Dorogovtsev and J. F. F. Mendes, *Adv. Phys.* **51**, 1079 (2002); M. E. J. Newman, *SIAM Rev.* **45**, 167 (2003); S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, *Phys. Rep.* **424**, 175 (2006).
 [2] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, *Nucl. Phys. B* **666**, 396 (2003).

- [3] J. Park and M. E. J. Newman, *Phys. Rev. E* **70**, 066117 (2004).
 [4] L. Bogacz, Z. Burda, and B. Waclaw, *Physica A* **366**, 587 (2006).
 [5] G. Bianconi, *Chaos* **17**, 026114 (2007).
 [6] M. Bauer and D. Bernard, e-print arXiv:cond-mat/0206150.
 [7] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, *Phys. Rev. Lett.* **85**, 4633 (2000).

- [8] P. L. Krapivsky and S. Redner, *Phys. Rev. E* **63**, 066123 (2001).
- [9] C. Castellano and R. Pastor-Satorras, *Phys. Rev. Lett.* **100**, 148701 (2008).
- [10] R. Pastor-Satorras and A. Vespignani, *Phys. Rev. Lett.* **86**, 3200 (2001); M. E. J. Newman, *Phys. Rev. E* **66**, 016128 (2002).
- [11] J. D. Noh, G. M. Shim, and H. Lee, *Phys. Rev. Lett.* **94**, 198701 (2005).
- [12] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, *Rev. Mod. Phys.* **80**, 1275 (2008).
- [13] R. Albert and A.-L. Barabási, *Science* **286**, 509 (1999).
- [14] P. L. Krapivsky and S. Redner, *J. Phys. A* **35**, 9517 (2002).
- [15] B. Waclaw and I. M. Sokolov, *Phys. Rev. E* **75**, 056114 (2007).
- [16] S. N. Dorogovtsev, J. F. F. Mendes, A. M. Povolotsky, and A. N. Samukhin, *Phys. Rev. Lett.* **95**, 195701 (2005).
- [17] Z. Burda and A. Krzywicki, *Phys. Rev. E* **67**, 046118 (2003).
- [18] M. Boguna, R. Pastor-Satorras, and A. Vespignani, *Eur. Phys. J. B* **38**, 205 (2004).
- [19] Z. Burda, J. D. Correia, and A. Krzywicki, *Phys. Rev. E* **64**, 046118 (2001).
- [20] Z. Burda, J. Jurkiewicz, and A. Krzywicki, *Physica A* **344**, 56 (2004).
- [21] B. Waclaw, e-print arXiv:0704.3702.
- [22] P. Bialas, Z. Burda, and D. Johnston, *Nucl. Phys. B* **493**, 505 (1997); **542**, 413 (1999).
- [23] M. R. Evans and T. Hanney, *J. Phys. A* **38**, R195 (2005).
- [24] P. Bialas, Z. Burda, J. Jurkiewicz, and A. Krzywicki, *Phys. Rev. E* **67**, 066106 (2003).
- [25] M. R. Evans, S. N. Majumdar, and R. K. P. Zia, *J. Stat. Phys.* **123**, 357 (2006).
- [26] L. Bogacz, Z. Burda, W. Janke, and B. Waclaw, *Comput. Phys. Commun.* **173**, 162 (2005).
- [27] W. Janke, *Physica A* **254**, 164 (1998).
- [28] A. Engel, R. Monasson, and A. K. Hartmann, *J. Stat. Phys.* **117**, 387 (2004).
- [29] W. Janke, in *Computer Simulations of Surfaces and Interfaces*, edited by B. Dünweg, D. P. Landau, and A. I. Milchev, NATO Science Series, II. Mathematics, Physics and Chemistry, Vol. 114 (Kluwer, Dordrecht, 2003), p. 137.
- [30] M. R. Evans and S. N. Majumdar, *J. Stat. Mech.: Theory Exp.* (2008) P05004.