Communities, clustering phase transitions, and hysteresis: Pitfalls in constructing network ensembles

David Foster,* Jacob Foster, and Maya Paczuski
Complexity Science Group, University of Calgary, Calgary, Canada T2N 1N4

Peter Grassberger
NIC, Forschungszentrum Jülich, D-52425 Jülich, Germany
and Complexity Science Group, University of Calgary, Calgary, Canada T2N 1N4

Ensembles of networks are used as null models in many applications. However, simple null models often show much less clustering than their real-world counterparts. In this paper, we study a "biased rewiring model" where clustering is enhanced by means of a fugacity as in the Strauss (or "triangle") model, but where the number of links attached to each node is strictly preserved. Similar models have been proposed previously in Milo *et al.* [Science **298**, 824 (2002)]. Our model exhibits phase transitions as the fugacity is changed. For regular graphs (identical degrees for all nodes) with degree $k > 2$ we find a single first order transition. For all nonregular networks that we studied (including Erdös-Rényi, scale-free, and several real-world networks) multiple jumps resembling first order transitions appear. The jumps coincide with the sudden emergence of "cluster cores:" groups of highly interconnected nodes with higher than average degrees, where each edge participates in many triangles. Hence, clustering is not smoothly distributed throughout the network. Once formed, the cluster cores are difficult to remove, leading to strong hysteresis. To study the cluster cores visually, we introduce *q -clique adjacency plots*. Cluster cores constitute robust communities that emerge spontaneously from the triangle generating process, rather than being put explicitly into the definition of the model. All the quantities we measured including the modularity, assortativity, clustering and number of four and five-cliques exhibit simultaneous jumps and are equivalent order parameters. Finally, we point out that cluster cores produce pitfalls when using the present (and similar) models as null models for strongly clustered networks, due to strong hysteresis which leads to broken ergodicity on realistic sampling time scales.

I. INTRODUCTION

Networks find wide use as models of complex systems. They have been applied fruitfully to social [1–3], technological [4], physical [5–7], and biological [8] phenomena. The nodes in a network represent components of the system and links connecting nodes represent interactions. Various measures are thought to capture significant overall structural features of networks. In this paper, we examine five properties of particular interest: the degree sequence [9], which counts the number of nodes in the network with k links; the clustering coefficient [10], which measures the tendency of connected triples of nodes to form triangles; the number of q cliques, which counts complete subgraphs with q nodes [11]; the assortativity [12], which measures the tendency of nodes to connect to other nodes of similar degree; and the modularity [13], which measures the tendency of nodes in networks to form tightly connected communities. Their formal definitions are recalled in Sec. II.

Models of network *ensembles* formalize and guide expectations about real-world networks [14]. The most famous are the Erdös-Rényi (ER) model of random networks [15], and the scale-free Barabási-Albert model [16] of networks growing by preferential attachment. Comparison with an *a priori* realistic "null" model can also indicate which features of a

real network are expected based on the null model, and which are surprising and thus of interest, as in motif search [17]. In the latter context, the most popular ensemble is the *configuration model* [18] and related variants [19], in which all networks with a given number of nodes and a given degree sequence are sampled with equal probability. One problem with the configuration model is that it shows far too little clustering—limiting its usefulness. This problem is especially acute when determining statistical significance of motifs and other local structures associated with clustering, for instance, in protein interaction networks [20].

To remedy this situation, a number of other models have been introduced that lead to both broad degree distributions and clustering. Some combine a preferential attachment mechanism with an explicit transitive rule for triangle formation [3,21–23], while others use a predefined degree distribution (as in the configuration model) together with tunable clustering [24,25].

A more general approach is to study network ensembles based on a partition function governed by a Hamiltonian [26–28]. A model where clustering is enhanced by means of a fugacity term in a Hamiltonian was introduced by Strauss and studied in detail in [29,30]. In the Strauss model, both the density of links and of triangles are controlled by fugacities. The model is a generalization of the Erdös-Rényi model with fixed edge *probability*, not with fixed edge number. A strong first-order phase transition [30] appears in the Strauss model separating a phase with weak clustering from a phase

*ventres@gmail.com

where nearly all edges condense into a single densely connected cluster consisting of high-degree nodes. Below the phase transition the number of triangles increases linearly with the triangle fugacity for large networks [29,30]. The phase transition is often seen as a flaw, as it does not admit moderate clustering observed in most real networks [31].

Here, we introduce and analyze the *biased rewiring model* (BRM). As in the configuration model, we fix the exact degree sequence—accounting for quenched heterogeneity of individual nodes. But as in the Strauss model, we control the average number of closed triangles by a Hamiltonian [28] containing a conjugate fugacity, or triangle bias β. Fixing the degree sequence prevents the extreme condensation of edges seen in the Strauss model, so we might *a priori* hope to achieve smooth control of clustering. Indeed, a similar model with a different Hamiltonian was proposed in [17].

To our surprise we found that clustering cannot be smoothly controlled. To search for phase transitions, we plotted a number of network characteristics, such as the clustering coefficient, against β. In each case and for all nonregular graphs (i.e., graphs with a nontrivial degree distribution), we found a sequence of jumps that look like first order phase transitions (or large Barkhausen jumps in ferromagnets [32–34]). Associated with these jumps are significant hysteresis effects. Further, we found that high-degree nodes play a crucial role. Thus it is not surprising that a simpler scenario holds for regular graphs (same degree k for all nodes)— where only a single phase transition occurs for all $k>2$. For $k=2$ no transition occurs. Unfortunately it is only in this last case that we obtain exact analytic results. In all other cases our results are based on numerical simulations.

At each transition, all measured quantities including the clustering, number of four and five cliques, assortativity, and modularity exhibit pronounced, simultaneous jumps. Hence, any of these quantities can serve as an order parameter for these transitions. In order to visualize the structures responsible for these jumps we introduce q-clique adjacency plots, which measure the number of q-cliques each link participates in. Using the adjacency plots together with a community detection algorithm [35], we observe that for each jump a new strongly connected cluster core appears. The cores are made of higher than average degree nodes in which each link participates in many triangles. Indeed, clustering is not evenly distributed through the network but concentrated into cluster cores. Hence, these networks cannot be easily classified into strong or weak clustering overall [36]. While previous work has shown how assortativity leads to clustering [25], or communities lead to clustering [2], we find the converse—that clustering leads to communities as well as to assortativity. The cluster cores are tightly connected communities of high-degree nodes, thus contributing to both modularity and assortativity. Hence, a transitive relation might be sufficient to explain both community formation and assortativity in certain (e.g., social) networks.

Our BRM is similar to a null model introduced in [17], where the Hamiltonian was chosen to bias not toward a *larger* number of triangles, but toward a *specific* number. In order to achieve this reliably, one needs a fugacity which is larger than that in the BRM. In the limit of large fugacities this is similar to a model with a hard constraint. In general,

statistical models with hard constraints show slower relaxation and worse ergodic behavior than models with soft constraints [37]. Hence, we expect that hysteresis effects might be even more pronounced in the model of [17] than in our BRM and might render it less useful as a null model, even if the problem of phase transitions is hidden. For simplicity we shall in the following call the model of [17] "triangle conserving," although the name is not strictly correct. We find that for triangle conserving rewiring, important structures remain largely unchanged on extremely long time scales, requiring particular care when using the method. Indeed, this difficulty is also relevant to the conclusions in [38], where the number of closed triangles and open wedges with any specific triple of node degrees was fixed. In general, phase transitions, strong hysteresis, and persistent structures of highly connected nodes together present substantial pitfalls for null models of clustered networks. Sampling from the ensemble of networks with a desired clustering coefficient must be done carefully, with the knowledge that the details of the observed ensemble will depend on the precise sampling method used, and may not reflect much about the underlying ensemble.

In the next section we summarize background information, including the precise definitions of the models with unbiased rewiring and the Strauss model. The definition of the BRM and our numerical procedure is presented in Sec. II F. Our main results are found in Secs. III A–III C, while some results for the model of Milo *et al.* [17] are presented in Sec. III D Finally, Sec. IV contains our conclusions.

II. BACKGROUND

A. Degree sequences

The degree of a node is the number of links in which the node participates. The network's degree sequence $\{n_k | k=0,1\ldots k_{\max}\}$ counts the number of nodes in the network with degree k. In this paper we studied regular ($n_k = \delta_{k,k_0}$) and Erdös-Rényi (poissonian n_k) networks, and several real world networks with fat tails. Network properties often depend strongly on the degree sequence [9]. Thus real networks are often compared with null models, which preserve that property.

B. Clustering coefficient and q cliques

Three nodes are *connected* if at least two of the three possible links between them exist. If all three links exist, they form a *triangle*. The clustering coefficient [9,39,40] measures the "transitivity" of relationships in the network, i.e., the probability that three connected nodes are also a triangle. Denoting the number of triangles by n_Δ and the degree of node i by k_i, one has

$$C = \frac{3n_\Delta}{\frac{1}{2}\sum_{i=1}^{N}(k_i-1)k_i} = \frac{3n_\Delta}{N_3}, \qquad (1)$$

where N_3 is the number of connected triples in the network. If every relationship in the network is transitive, $C=1$; if no

relationships are transitive, $C=0$. Note that the denominator of Eq. (1) depends only on the degree sequence, and thus, $C \propto n_\Delta$ in any ensemble with fixed degrees. In addition to C, we can also define similar higher order clustering coefficients based on q cliques, i.e., on complete subgraphs with q nodes, as

$$C_q = \frac{q n_{q\text{-clique}}}{\sum\limits_{i=1}^{N} \binom{k_i}{q-1}}, \qquad (2)$$

where $n_{q\text{-clique}}$ is the number of q-cliques in the network. Notice that $C=C_3$. As we shall see, we can use any C_q as an order parameter for the phase transitions discussed in this work.

C. Assortativity

The assortativity r measures the tendency for nodes in the network to be linked to other nodes with similar degree. It is defined as the Pearson correlation coefficient between the degrees of nodes that are joined by a link [12].

$$r = \frac{L\sum\limits_{i=1}^{L} j_i k_i - \left[\sum\limits_{i=1}^{L} j_i\right]^2}{L\sum\limits_{i=1}^{L} j_i^2 - \left[\sum\limits_{i=1}^{L} j_i\right]^2}. \qquad (3)$$

Here L is the number of links in the network and j_i and k_i are the degrees of nodes at each end of link i. If high-degree nodes are linked exclusively to other high degree nodes, then $r \approx 1$. If high degree nodes are exclusively linked to low degree nodes, $r \approx -1$. The assortativity, r also serves as an order parameter for the BRM.

D. Modularity

A number of heuristic methods exist for identifying community structure in complex networks [41,42], each with its own strengths and weaknesses. We shall use a measure proposed by Newman and Girvan [13] called *modularity*. Assume one has partitioned a network into k nonoverlapping communities. Define e_{ij} as the fraction of all edges which connect a node in community i to a node in community j. Thus $a_i = \Sigma_j e_{ij}$ is the fraction of all links which connect to community i. The modularity of the partition is then defined as:

$$Q = \sum_i (e_{ii} - a_i^2), \qquad (4)$$

and the modularity of the network is the maximum of Q over all partitions. Q measures the fraction of 'internal' links, versus the fraction expected for a random network with the same degree sequence. It is large when communities are mostly isolated with few cross links.

The main problem in computing Q for a network is the optimization over all partitions, which is usually done with some heuristics. We use a greedy algorithm introduced by

Newman [43] which starts with each node in its own community (i.e., all communities are of size one). Joining two communities i and j would produce a change δQ_{ij}. After checking all pairs (i,j), the pair with the largest δQ_{ij} is joined. This is repeated until all δQ_{ij} are negative, i.e., until Q is locally maximal. We follow the efficient implementation of this method described by Clauset *et al.* [35] and find that the modularity also serves as an order parameter for the BRM.

E. Exponential network ensembles and network Hamiltonians

Let us assume that $\mathcal{G}$ is a set of graphs. This could be the set of all graphs with fixed number N of nodes, or with fixed N and fixed number of links L, or with fixed N and fixed degree sequence etc. Consider a particular graph $G \in \mathcal{G}$. Following [27,28], a network *Hamiltonian* $H(G)$ is any function defined on $\mathcal{G}$, used to define an exponential ensemble (analogous to a canonical ensemble in statistical mechanics) by assigning a weight

$$P(G) \propto e^{-H(G)} \qquad (5)$$

to any graph, similar to the Boltzmann-Gibbs weight.

Examples of exponential ensembles are the Erdös-Rényi model $G(N,p)$ where $H = -L \ln[p/(1-p)]$ and the Strauss model with

$$H_{\text{Strauss}} = \theta L - \beta n_\Delta. \qquad (6)$$

Here, p [which is not to be confused with $P(G)$] is the probability that a link exists between any two nodes, while θ and β are "fugacities" conjugate to L and n_Δ, respectively.

In the configuration model, $\mathcal{G}$ is the set of all graphs with a fixed degree sequence and $H=0$. Thus all graphs with the same degree sequence have the same weight. In the "triangle conserving" model of Milo *et al.* [17], again $\mathcal{G}$ is the set of graphs with fixed degree sequence, but

$$H_{\text{Milo}} = \beta |n_\Delta - n_{\Delta,0}|. \qquad (7)$$

where $n_{\Delta,0}$ is a prespecified target number of triangles, usually the number found in the empirical network being studied. Finally, in the BRM, $\mathcal{G}$ is again the same but

$$H_{\text{BRM}} = -\beta n_\Delta. \qquad (8)$$

Thus, while larger weights are given in the BRM with $\beta > 0$ to graphs with more triangles (high clustering), the triangle conserving model of [17] assigns the largest weight to graphs with $n_\Delta = n_{\Delta,0}$.

F. Simulations: Rewiring

Numerical simulations of these ensembles are often accomplished using the Markov chain Metropolis-Hastings method [37]. This is particularly easy for models without a fixed degree sequence, such as the Strauss model. There, new configurations are simply generated by randomly adding or removing links. This is not possible for the ensembles with a fixed degree sequence, where the most natural method is *rewiring* [19]. Below we first review the unbiased case cor-

responding to the configuration model, and then discuss the two biased cases H_{Milo} and H_{BRM}.

1. Unbiased rewiring

Starting from a current graph G, a new graph G' is proposed as follows: Two links, which have no node in common, are chosen at random, e.g., $X-Y$ and $W-Z$. Links are then swapped randomly either to $X-W$ and $Y-Z$, or to $X-Z$ and $Y-W$. If this leads to a double link (i.e., one or both of the proposed new links is already present), the new graph G' is discarded and G is kept. Otherwise, G' is accepted. This method conserves the degree sequence, satisfies detailed balance, and is ergodic [19]. Thus, it leads to equidistribution among all graphs sharing the same degree sequence as the original graph. Although no exact results for the speed of equilibration seem to exist, previous experience [19,20] suggests that this unbiased rewiring is fast and can be used efficiently even for large networks.

2. Biased rewiring

For biased rewiring with a Hamiltonian $H(G)$, the proposal stage is the same as for unbiased rewiring. Only the acceptance step has to be modified, according to the standard Metropolis-Hastings procedure [37,44]: If $H(G') \leq H(G)$, then G' is accepted—unless it makes a double link. Otherwise the swap is accepted with probability

$$p = e^{H(G)-H(G')} p \leq 1, \tag{9}$$

unless it makes a double link, and rejected with probability $1-p$.

The precise protocols for simulating the two biased models studied in this paper are different. The BRM starts with the actual network G_0 whose degree sequence we want to use, and first proposes M_0 *unbiased* swaps, with M_0 sufficiently large so that the system ends up in the typical region of the unbiased ensemble. After that we increase β in small steps (typically $\delta\beta=0.002$), starting at $\beta=0$. After each step in β, M_1 swaps are proposed in order to equilibrate approximately, and then at this β an ensemble average is obtained by making m measurements, each separated by M_2 additional proposed swaps. Thus the total number of proposed swaps at each fixed β is $M_1+(m-1)M_2$. Typically, $M_0 \approx 10^6$, $M_1 > 10^5$, $M_2 \approx 10^3 - 10^5$, and $m \approx 500 - 10\,000$. The precise choice of these values depends on the network size, with larger networks needing more rewiring to equilibrate.

Following the m measurements we increase β and repeat the procedure, until a preset maximal value β_{max} is reached. After that, we reverse the sign of $\delta\beta$ and continue with the same parameters M_1, M_2, and m until we reach again $\beta=0$, thereby forming a hysteresis loop. Triangle bias values during the ascending part of the loop are denoted by β^+; those in the descending part are denoted as β^-. In cases where we start from a real world network with $n_{\Delta,0}$ triangles, we choose β_{max} sufficiently large so that $n_\Delta(\beta^+) > n_{\Delta,0}$. Then the range of clustering covered in the hysteresis loop includes the clustering coefficient of the original network, G_0.

For the biased model of Milo *et al.* [17] we skip the first stage by setting $M_0=0$, and jump immediately to a value of

β, estimated after preliminary runs, which is large enough to reach $n_{\Delta,0}$ at long times. At this β, we make M_1 swaps to equilibrate, and then make m measurements, each separated by M_2 further swaps. An alternative protocol using multiple annealing periods is discussed in Sec. III D. Averages are taken only over configurations with exactly $n_{\Delta,0}$ triangles. If β is too small, the bias is not large enough to keep n_Δ near $n_{\Delta,0}$, and n_Δ drifts to smaller values. Even if β is sufficiently large in principle, the algorithm slows down if β is near its lower limit, since n_Δ seldom hits its target value. On the other hand, if β is too large, most swaps are rejected, leading to increased relaxation times. Thus choosing an optimal β for a given network is somewhat delicate.

III. RESULTS

We explored the behavior of the BRM for three different network classes: k–regular networks, where every node in the network has degree k; Poisson degree distributions as in Erdös-Rényi networks; and typical fat-tailed distributions as in most empirical networks. Although we studied many more cases (Erdös-Rényi networks with different connectivities and sizes and several different protein-protein interaction networks), we present here only results for k-regular networks with different k, for one Erdös-Rényi network, and for two empirical networks with fat-tailed degree distributions: A high-energy physics collaboration network (HEP) [45] and a protein-protein interaction network for yeast (*S. cerevisiae*) [46]. Except for k–regular networks, multiple discontinuous phase transitions appear, while only a single discontinuous phase transition arises for all k-regular networks with $k > 2$.

A. *k*-regular networks: analytical and numerical simulation results

1. Numerical simulations of k-regular graphs

For each k, the configuration with maximal n_Δ is a disjoint set of $(k+1)$-cliques, where the graph decomposes into disjoint completely connected components of $k+1$ nodes. When N is divisible by $k+1$, this gives

$$n_\Delta^{(k,\text{max})} = \frac{N}{k+1}\binom{k+1}{3}. \tag{10}$$

For $k=2$, $n_\Delta^{(k,\text{max})}$ is reached in a smooth way for increasing β. However, for each $k \geq 3$, and for sufficiently large N, n_Δ first increases proportional to $\exp(\beta)$. Then the increase accelerates and finally n_Δ jumps to a value very close to $n_\Delta^{(k,\text{max})}$. This is illustrated for $k=3$ in Fig. 1, where we plot hysteresis curves for n_Δ vs β. From this and from similar plots for different k we observe the following features:

(i) For small β, all curves are roughly described by

$$n_\Delta \approx \frac{(k-1)^3}{6} e^\beta \tag{11}$$

as indicated by the straight line in Fig. 1. This approximation seems to become exact as $N \to \infty$. Note that Eq. (11) is the same as the corresponding equation for the number of triangles in large exponential random graphs [29,30] with
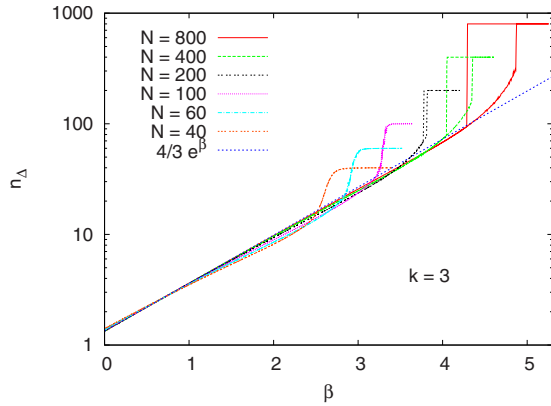
FIG. 1. (Color online) Average number of triangles for regular networks with $k=3$ vs β. All curves depict full hysteresis cycles, with $M_1=2\times10^5$ initial swaps after each increment of β, and $M_2=5000$ additional swaps after each of $m=4\times10^4$ measurements at the same value of β. Hysteresis loops are evident for $N\geq200$, but not for $N\leq100$. The straight line corresponds to Eq. (11).

FIG. 2. Average number of triangles for k-regular networks, with $k=2, 3, 5, 10,$ and 16, vs the triangle bias or conjugate fugacity β. Network size is $N=400$ for all curves. In these simulations β was slowly increased, until a jump in n_Δ was seen (for $k\geq3$). The straight line shows the theoretical prediction for $k=2$: $n_\Delta=\frac{1}{6}e^\beta$. The inset shows n_Δ/e^β for $k=2$. All curves for $k>2$ exhibit hysteresis at this value of N but, for clarity, only the curves for increasing β are shown.

$(k-1)$ replaced by $\langle k\rangle$. This implies that n_Δ is independent of N, and the clustering coefficient is proportional to $1/N$.

(ii) While the curves are smooth and do not show hysteresis for small N, they both jump and exhibit hysteresis above a k-dependent value of N. This is our best indication that the phenomenon is a first order phase transition in the thermodynamic limit, similar to the one in the Strauss model. Above the jump, the curves saturate, within the resolution of the plot, the bound given in Eq. (10).

(iii) The jumps occur when n_Δ becomes extensive. Hence the value β^* where the jump occurs increase approximately logarithmically with N. Although size dependent transition points are not very common, there are some well known examples. These include models with long range or mean field interactions, where the number of interaction terms increases faster than N. In the present case the reason for the logarithmic increase is that k-regular graphs become sparser as N increases. Thus the density of triangles decreases, and in a Markov chain MC method there are increasingly more proposed moves which destroy triangles than moves which create them. To compensate for this, $\exp(\beta^*)$ has to increase $\propto N$.

In Fig. 2, we show the average number of triangles as a function of β for k-regular networks, for $k=2, 3, 5, 10,$ and 16, with $N=400$ nodes. For each network we perform $M_0=4\times10^6$ initial swaps. Each curve corresponds to $M_1=4\times10^6$ swaps after each increase in β, and $M_2=2\times10^5$ additional swaps after each of $m\geq5000$ measurements at the same value of β. For clarity we show only values for increasing β, although strong hysteresis occurs for all $k\geq3$ when $N=400$.

Regular networks with $k=2$ exhibit no hysteresis, and no indication of any phase transition. As illustrated in the inset of Fig. 2, for all values of β, n_Δ is very well described by Eq. (11) for $k=2$ up to the point where it reaches the bound in Eq. (10). Close to that point there is a small bump in the curve shown in the inset that is explained next.
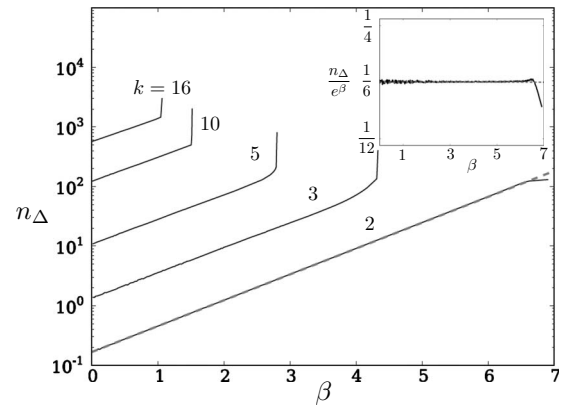
2. $k=2$ analytical results

We now derive analytically Eq. (11) for $k=2$, and show how this equation become exact for $N\to\infty$. A regular $k=2$ network consists of N nodes and N links arranged in a set of disjoint simple loops. Triangles are the smallest possible loops, since self-links and double links are not allowed. For large N and small β nearly all loops are large. So the number of loops of length <7 is of order $1/N$ and can be neglected for $N\to\infty$ for finite β. Since the biased rewiring process favors the creation of loops of length three, we must keep track of their number to determine when the process has reached equilibrium.

Consider now a network of size N with n_Δ triangles and bias β. The rewiring process reaches equilibrium when the probability to destroy a triangle is equal to the probability to create a new one. First we calculate the probability to select a swap that destroys a triangle. The total number of ways to choose a pair of links and perform a swap is $\mathcal{N}=\frac{N(N-1)}{2}\times2$, where $\frac{N(N-1)}{2}$ gives the number of distinct pairs of links and the extra factor of two accounts for the two possible ways to swap links. To destroy a triangle, one of the links must be chosen from it, and the other from a larger loop. The chance that both links are chosen from triangles, which would lead to the destruction of both, can be neglected. There are $3n_\Delta$ possible links in triangles to choose from, and $(N-3n_\Delta)$ links in larger loops. Thus the probability of choosing a swap which would destroy a triangle is

$$p_{\Delta-}=\frac{3n_\Delta(N-3n_\Delta)\times2}{\mathcal{N}}=\frac{6n_\Delta}{N}\times[1+O(N^{-1})],\quad(12)$$

where the factor of two in the numerator corresponds to the fact that both possible swaps destroy a triangle and the correction term also takes into account the neglected loops of lengths 4, 5, and 6.

TABLE I. The number of nodes (N), the number of links (L), the average degree $\langle k \rangle$, the clustering coefficient (C), the assortativity (r), the modularity (Q), and the number of connected triples (N_3) for each of the networks discussed in Sec. III B. ER stands for Erdös-Rényi. HEP is the scientific collaboration network in [45], and Yeast is the protein-protein interaction network in [46].

Network properties						
Network	N	$\langle k \rangle$	C	r	Q	N_3
ER	800	5.0	0.002	−0.0004	0.196	10560
HEP	7610	4.1	0.33	0.29	0.397	121083
Yeast	1373	10.0	0.58	0.58	0.380	104970

To add a triangle to the network, two links must be chosen from the same long loop. They must be separated by exactly two links. Since ℓ such pairs occur in a loop of length ℓ, the total number of such pairs in the network is N. This neglects terms of $\mathcal{O}(1)$, corresponding to the triangles and other loops shorter than $\ell = 7$. Hence the probability to add a triangle

$$p_{\Delta+} = \frac{N}{\mathcal{N}} = N^{-1} \times [1 + O(N^{-1})], \quad (13)$$

where there is no factor of two in the numerator because only one of the two possible swaps creates a triangle. Equilibration is achieved when

$$p_{\Delta+} = e^{-\beta} p_{\Delta-}, \quad (14)$$

giving

$$n_\Delta = \frac{e^\beta}{6} \quad (15)$$

up to correction terms of order $1/N$. Indeed Eq. (15) is just Eq. (11) for $k = 2$.

The simple exponential dependence of n_Δ on β occurs because swaps create/destroy triangles independently and one at a time—except in the rare case of breaking up a loop of length six. For networks with nodes of degree greater than two this still holds for small β. As β increases, nodes cluster together more densely, allowing each link to participate in many triangles. This cooperative behavior is not captured in Eq. (11). The presence of triangles helps other triangles to form and makes it harder for them to be removed. Such cooperativity explains intuitively the existence of first order phase transitions for $k \geq 3$ but not for $k = 2$, where these effects are not possible.

Indeed, for n_Δ very close to its limit $n_\Delta^{(2,\text{max})}$ *some* cooperativity appears even for $k = 2$. The configuration with $n_\Delta = n_\Delta^{(2,\text{max})}$ can be changed only by breaking up *two* triangles and joining their links in a loop of length six. When n_Δ is close to $n_\Delta^{(2,\text{max})}$, link swaps which involve two triangles become increasingly prevalent. The tendency to create and destroy triangles two at a time introduces weak cooperativity, which is only noticable when $n_\Delta^{(2,\text{max})} - n_\Delta = O(1)$. It is not strong enough to lead to a phase transition, but it explains the small bump seen in the inset of Fig. 2.

B. Networks with nontrivial degree sequences

We explored the behavior of our BRM for various networks. These included ER graphs with different sizes and different connectivities and several real-world networks. The latter typically show more or less fat tails. In order to find any dependence on the fatness, we also changed some of the sequences manually in order to reduce or enhance the tails. We found no significant systematic effects beyond those visible already from the following three typical networks, and restrict our discussion to these: an ER graph [15], a HEP [45], and a yeast protein binding network (Yeast) [46]. Table I collects some of their properties.

Figures 3–5 show the clustering coefficient $C = 3n_\Delta/N_3$. for these three networks. In each case $M_0 = 10^6$, $M_1 = 1.5 \times 10^5$, $M_2 = 50000$, and $m = 500$. In each of them a full hysteresis cycle is shown, with the lower curves (labeled β^+) corresponding to increasing and the upper curves (β^-) corresponding to decreasing β. In Figs. 4 and 5 the dotted line shows the number of triangles in the empirical networks. These fall well within the hysteresis loops.

For small values of β^+ all three figures exhibit exponential increase in the clustering coefficient similar to that observed in k-regular networks. At different values of β, however, there is a sudden, dramatic increase in C, which does *not* lead to saturation as it did for k-regular networks. The first phase transition is followed by a series of further transitions
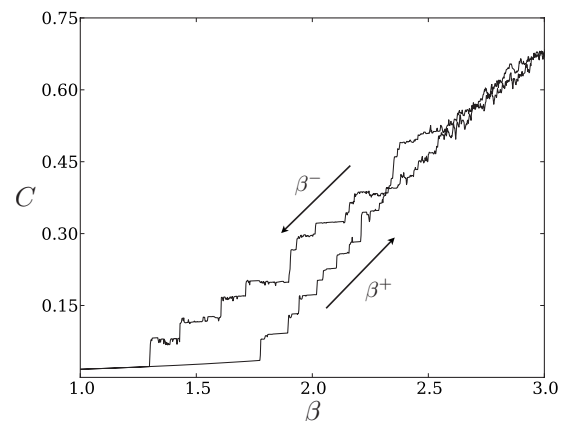

FIG. 3. Clustering coefficient C in the BRM for an ER graph with $N = 800$ nodes and $\langle k \rangle = 5$, plotted against the triangle bias β. The lower curve corresponds to slowly increasing β, the upper to decreasing β.
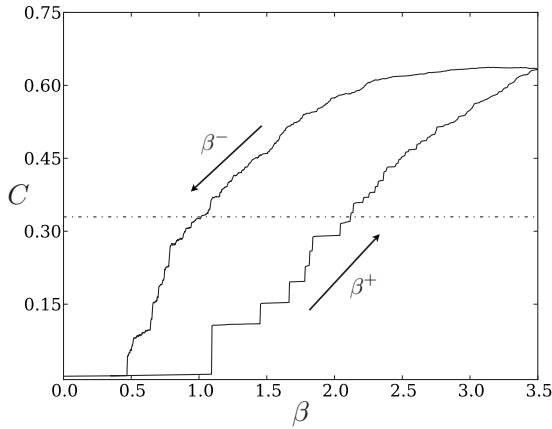
FIG. 4. Similar to Fig. 3, but for the BRM of the HEP network. The dotted line indicates the clustering coefficient in the real network, which falls well within the hysteresis loop.

through which the network becomes more and more clustered. Many of the jumps are comparable in absolute magnitude to the first jump. Although the rough positions of the jumps depend only on the degree sequences, their precise positions and heights change slightly with the random number sequences used and with the speed with which β is ramped up and down. Thus the precise sequence of jumps presumably has no deep significance, but the existence of jumps and the general appearance of the hysteresis loops seem to be a universal feature found in *all* cases.

Associated with the jumps in C are jumps in all other network characteristics we looked at; see Fig. 6. The jumps in the other quantities always occur at *exactly* the same positions as those in C or, equivalently, n_Δ. Obviously, at each jump a significant restructuring of the network occurs, which affects all measurable quantities. Speculations as to how these reorganizations can best be described and what is their most "natural" driving mechanism are put forward in Sec. III C.

In the downward branch of the hysteresis loop, as β^- decreases toward zero, the clustering coefficient remains high for a long time, forming a significant hysteresis loop. This loop suggests that all jumps should be seen as discon-
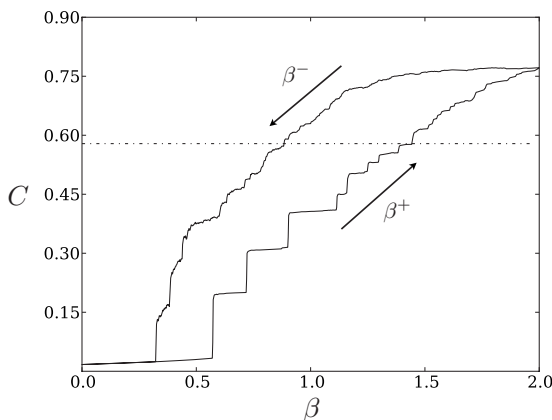

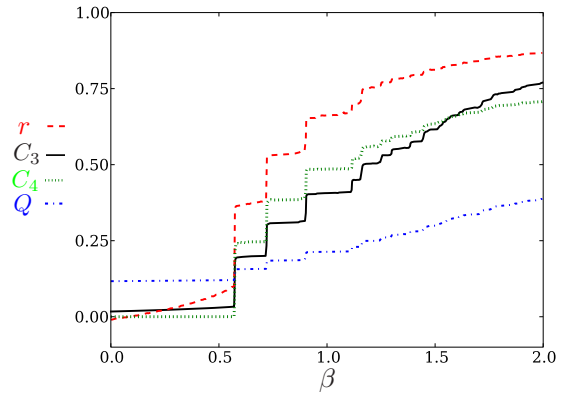FIG. 5. Same as Fig. 4, but for the BRM of the Yeast network.

FIG. 6. (Color online) Four network characteristics: modularity (Q), clustering coefficient (C_3), four-clique clustering coefficient (C_4), and assortativity (r) vs β for the BRM of the Yeast network. These data are drawn from the same simulation as in Fig. 5, but for clarity only the results for increasing values of β are shown. All quantities exhibit simultaneous jumps and hence serve as equivalent order parameters for the transitions.

tinuous, first order phase transitions. Since all studied systems are finite, these hysteresis loops disappear for infinitely slow ramping of β. But the sampling shown involved $>2.5 \times 10^6$ attempted swaps for each value of β, and no systematic change in the hysteresis was seen when compared to sweeps that were twice as fast.

In Fig. 7, we plot the assortativity against the number of triangles for the *same* values of $\beta^\pm$, normalizing both quantities to the unit interval. We guessed that in this way we might find universal curves which are the same for β^+ and β^-, and maybe even across different networks. Indeed, we see a quite remarkable data collapse. It is certainly not perfect, but definitely better than pure chance. It suggests that the triangle bias in the BRM leads to networks where the two characteristics $n_\Delta/n_{\Delta,\max}$ and $(r-r_{\min})/(r_{\max}-r_{\min})$ are strongly—but nonlinearly correlated. Roughly similar collapse was found for other quantities such as the modularity. This indicates a potential scaling relationship between these order parameters in the BRM. For the two empirical net-
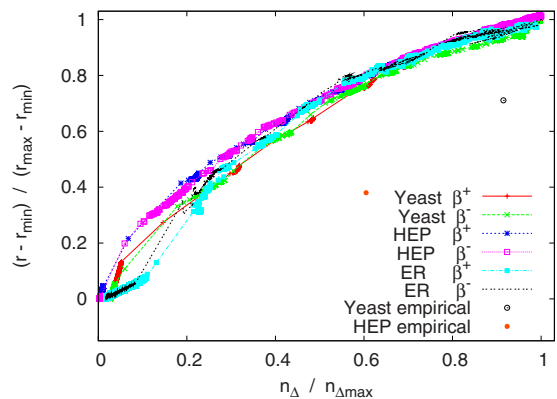

FIG. 7. (Color online) Values of the rescaled characteristics $n_\Delta/n_{\Delta,\max}$ and $(r-r_{\min})/(r_{\max}-r_{\min})$, measured at the same values of $\beta^\pm$, and plotted against each other. The points represent the values for the real HEP and Yeast networks.

works, we show also the real values of these characteristics. They fall far from the common curve, indicating that these networks are not typical for the BRM at any value of β.

Among the networks studied here, the ER network is closest to a k-regular network, so it could be expected to show behavior closest to that studied in the last subsection. But this is not evident Figs. 3–5. However, hysteresis appears to be more closely tied to individual jumps for the ER network, while it is more global (and thus, also more important overall) for HEP and Yeast networks.

For the HEP and Yeast networks, we can compare the clustering of the BRM ensemble to that in the real empirical networks. The latter numbers are shown as dashed lines in Figs. 4 and 5. In both cases, the line intersects the hysteresis loop where it is very broad. This means that a large value of β^+ is required to reach the real network's level of clustering when the bias is increased, whereas a much lower value β^- must be reached before these triangles can be rewired out of the network again. This gap between β^+ and β^- at fixed C has important implications for the triangle "conserving" null model of [17], as we discuss later.

C. Clique adjacency plots and clustering cores

Up to now we have not given any plausible arguments for why clustering increases via a sequence of jumps rather than in a single jump or in a continuous manner. One might expect *a priori* that each jump is related to the breakup of a connected component into disconnected subgraphs, just as the phase transition in regular graphs is associated with such a breakup. Indeed this type of percolation phase transition was seen by Volz [24] in different networks as a function of a tunable clustering. However, by counting the number of disconnected components for the first few jumps we ruled out this scenario as a generic cause for the jumps, although we do expect the networks to break up when β becomes very large.

Instead, we now argue that each jump is associated with the sudden formation of a highly connected cluster of higher than average degree nodes, where each link participates in many triangles. The first jump encountered on ramping up β occurs when some of the strongest hubs link among themselves, forming a highly connected cluster. Subsequent jumps indicate the formation of other clusters with high intra—but low interconnections. Hence the clusters correspond to communities according to a standard definition [35,47]. What distinguishes this picture of modularity is that it automatically leads to large assortativity: Since high-degree nodes form the first cluster(s), there is a strong tendency for clusters to contain nodes with similar degrees. For previous discussions of how clustering of nodes depends on their degree, see e.g., [36,48,49]). The BRM demonstrates how a bias for transitive relations or triangles can give rise to community structure *de novo*, whereas in other models, community structure must be put in by hand [2,50].

We refer to the clusters of tightly connected nodes created by the BRM as *cluster cores*. To visualize them, we use q-*clique adjacency plots* (qCAPs). A q-clique adjacency plot is based on an integer-valued $N \times N$ matrix T_{ij}^q called the

q-clique adjacency matrix. When no link exists between nodes i and j, $T_{ij}^q = 0$. Otherwise, the matrix element is equal to the number of q-cliques in which the link between i and j participates. For example, if $q=3$, $T_{ij}^{q=3}$ is nonzero only when i and j are connected. It counts the number of common neighbors. Hence, $T_{ij}^{q=3} = m_{ij}$ where m_{ij} is the edge multiplicity in [25,36]. However, instead of looking at coarse grained quantities, which lump together all node pairs with degree k and k' sharing an edge, we examine microscopically each individual link. $T_{ij}^{q=3}$ can be considered a proximity measure for nodes: linked nodes with many common neighbors are likely to belong to the same community. Similar proximity measures between nodes which depend on the similarity of their neighborhoods have been used in [50–52]. To visualize T_{ij}, we first rank the nodes and then plot for each pair of ranked nodes a pixel with corresponding color or gray scale indicating the value of T_{ij}^q. Possible ranking schemes are by degree, by the number of q-cliques attached to the node, or by achieving the simplest looking block-diagonalized q-clique adjacency plot.

Examples for the Yeast degree sequence are given in Fig. 8. The four rows, descending from the top, show the 3CAPs for typical members of the BRM ensemble before the first jump and after the first, second, and third jumps. The plots in the left column rank nodes according to their degree. The plots on the right show the same matrices after "diagonalization," with nodes forming the first cluster placed in the top ranks, followed by the nodes forming the second cluster, and then nodes forming the third cluster. This diagonalization can be performed using a community detection algorithm [35], followed by some heuristics to order nodes within each community so that the centers of the blocks have the highest values of T_{ij}^q. The community detection algorithm occasionally includes a few nodes with low degree in the cluster core community; these nodes are not shown here. But the relevant parts of the 3CAPs are shown: nodes with lower ranks do not play any substantial role except for very large values of β. We notice several features:

(i) Not all high-degree nodes participate in the first cluster cores. Obviously, the selection of participating nodes is to some degree random. When sufficiently many links are established the cores freeze and cannot be changed easily. This agrees with our previous observation that the positions of the jumps change unsystematically with details like the random number sequence or the speed with which β is ramped.

(ii) Cluster cores that have been formed once are not modified when β is further increased. Again this indicates that existing cores are essentially frozen.

(iii) Cluster cores corresponding to different jumps do not overlap.

All three points are in perfect agreement with our previous finding that hysteresis effects are strong and that once these structure appear, they are preserved when β is increased further.

From other examples (and from later jumps for the same Yeast BRM) we know that the last two items in the list are not strictly correct in general, although changes of cores and overlap with previous cores do not occur often. Thus the results in Fig. 8 are too extreme to be typical. When a cluster core is formed, most of the links connected to these nodes
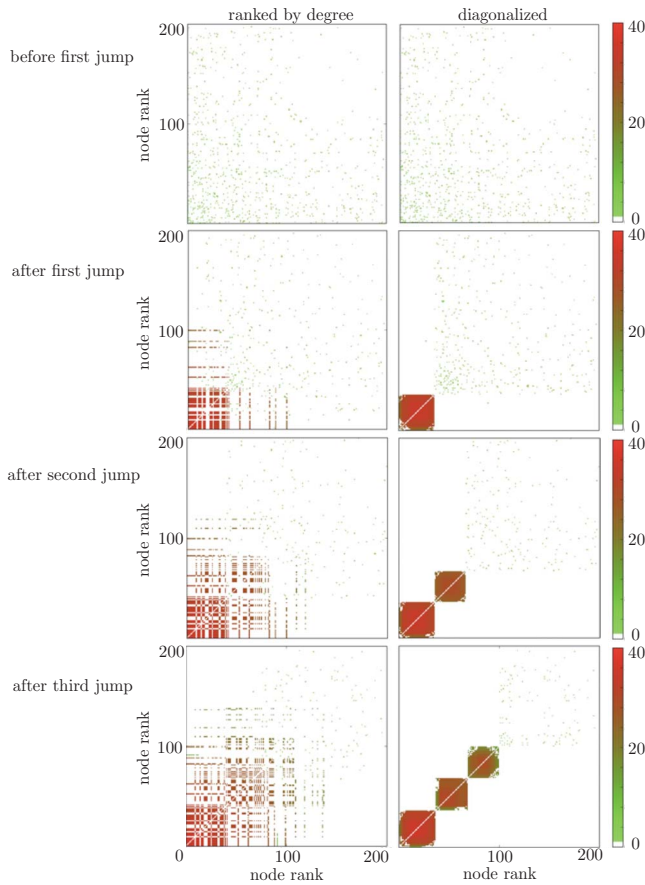
FIG. 8. (Color online) Relevant parts of three-clique, or triangle adjacency plots (3CAPs) for the Yeast network. The color of each point indicates the number of triangles in which the link participates. Each pair of plots shows (from top to bottom) the 3CAP for a typical member of the ensemble shortly before the first jump seen in Fig. 5, shortly after it, shortly after the second jump, and shortly after the third jump. The plots on the left hand side show the 3CAP with the nodes ranked in order of their respective degree. In the "diagonalized" plots we rearranged the ranking so that nodes which participate in the three clusters formed by each jump are ranked together, at the head of the list. To do this we use the community detection algorithm of [35] (see text). The rest of the nodes are ranked by degree.

are used to connect to other members of the cluster core, and the few links left over do not have a substantial effect on the further evolution of the core.

We find that as β^- decreases, the clustering cores persist well below the value of β^+ at which they were created (not shown here). This provides further evidence that once a link participates in a large number of triangles, it is stable and unlikely to be removed.

These adjacency plots are also useful for analyzing empirical networks, independent of any rewiring procedure, to help visualize community structure. While nodes in different communities often are linked, these links between communities usually take part in fewer triangles than links within communities. Thus simply replacing the standard adjacency matrix by the three-clique adjacency matrix should help discover and highlight community structure [50–52].
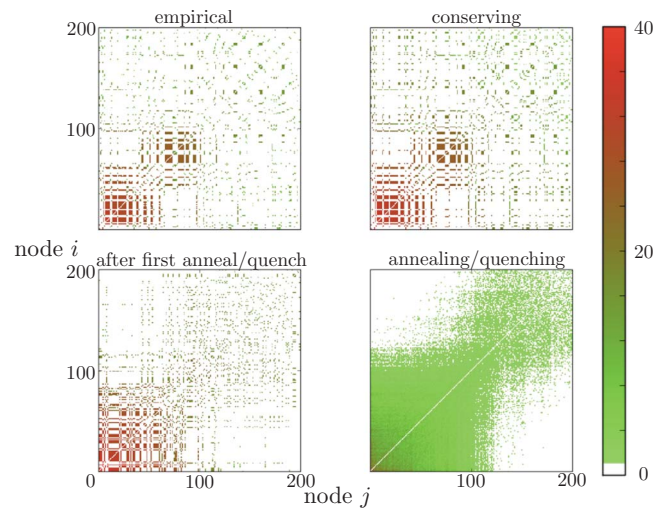

FIG. 9. (Color online) Parts of 3CAPs for the real yeast protein-protein interaction network of [46], for a typical network of the "triangle conserving" ensemble with no annealing, for a network obtained after an "annealing" period with $\beta=0$ and a subsequent quench with $\beta\neq0$ using "triangle conserving" rewirings [17], and for an ensemble obtained by 500 such annealing/quenching alternations.

In the top left panel of Figs. 9 and 10 we show parts of the 3CAPs for the yeast protein-protein interaction and HEP networks respectively. In both cases, nodes are ranked by degree. We see that the triangles tend to be formed between strong hubs. But clustering in the real networks does not strictly follow the node degree, in the sense that some of the strongest hubs are not members of prominent clusters. This shows that real networks often have features which are not encoded in their degree sequence, and that a null model entirely based on the latter will probably fail to reproduce these features. We see also that links typically participate in *many* triangles, if they participate in at least one. This is in contrast to a recently proposed clustering model, which assumes that each link can only participate in a single triangle [53].
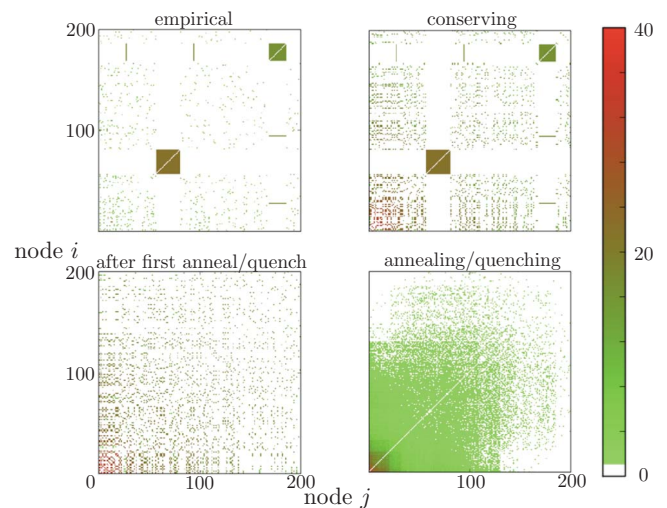

FIG. 10. (Color online) Analogous to Fig. 9 using the high-energy physics collaboration network instead [45].

D. Triangle conserving null models

In the previous subsection we looked at the case where the bias is "unidirectional." In contrast, Milo *et al.* [17] considered the case where the bias tends to increase the number of triangles when it is below a target number $n_{\Delta,0}$, but pushes it *down* when it is above [54]. In this way neither jumps nor any hysteresis is encountered. But that does not mean that the method is not plagued by the same underlying problem–extremely sluggish dynamics and broken ergodicity on realistic time scales.

In the most straightforward implementation of triangle conserving rewiring with the Hamiltonian H_{Milo} [17], one first estimates via preliminary runs a value of β which is sufficiently large so that n_Δ fluctuates around $n_{\Delta,0}$. Then one starts with the original network and rewires it using this β, *without first "annealing" the network at $\beta=0$*. The effect of this is seen in the top right panels of Figs. 9 and 10. In both cases, the three-CAPs shown were obtained after more than 10^9 attempted swaps. At $\beta=0$, this number is more than enough to equilibrate the ensemble. For the large values of β needed ($\beta=1.5$ for Yeast, and $\beta=2.4$ for HEP), some changes to the initial configurations are seen but the original structures remain intact. This is particularly true for the strongest clusters existing in the real networks. Triangles not taking part in these clusters change more rapidly, but are also less important.

Thus we demonstrate a pitfall inherent in triangle conserving rewiring: when the bias is strong enough to keep the number of triangles in the network around the desired target number, the bias is also so large that links within the cluster cores are hardly ever randomized.

As a way out of this dilemma, we can alternate epochs of triangle conserving swaps with "annealing periods" where $\beta=0$. This guarantees that memory is wiped out during each annealing period (see the lower left panels in Figs. 9 and 10), and each "quenching epoch" contributes one independent configuration to the ensemble. After many such cycles we can obtain an ensemble, which looks more evenly sampled—as indicated in the lower right panels in Figs. 9 and 10. However, even then we cannot be sure that it accurately represents the equilibrium ensemble for the Hamiltonian H_{Milo}. Apart from the last caveat, the method would presumably be too slow for practical applications where high accuracy and precise variances of ensemble observables are needed, since one needs one entire cycle per data point. But it can be useful when it is sufficient to estimate fluctuations roughly, and when high precision is not an issue. To illustrate this, Fig. 11 shows results for the HEP network where 200 anneal/quench cycles were made for two different values of β ($\beta=2.3$ and $\beta=2.5$). In each cycle, the quenching epoch terminated when the number of triangles reached the value of the real network, and the values of r and of the number of four-cliques were recorded. Figure 11 shows that these values scatter considerably, but are in any case far from the values for the real network. Thus the ensemble is a poor model for the real HEP network. Figure 11 also shows that r and $n_{4-\text{clique}}$ depend slightly on β (as was expected), but not so much as to invalidate the above conclusion.

Reference [38] proposed "*dK* randomized ensembles" as null models that progressively converge to the actual net-
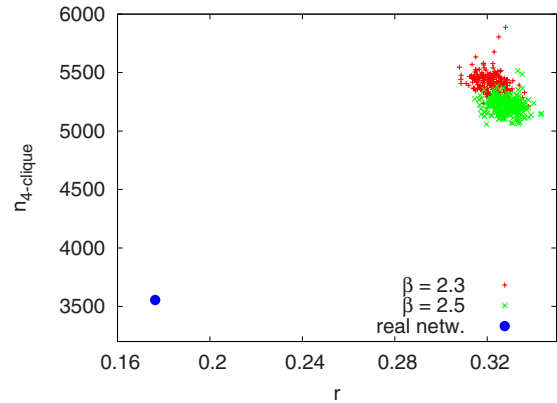

FIG. 11. (Color online) Values of the assortativity r and of the number of four-cliques in the real high-energy physics network and in 400 members of the triangle-conserving ensemble. The 400 realizations were obtained by 200 anneal/quench cycles with $\beta=2.3$ and 200 cycles with $\beta=2.5$, as described in the text.

work as d is increased. In the $3K$ ensemble, the number of triangles and wedges for each triple of degrees k, k', and k'' are fixed. As a result, in any configuration the number of allowed link swaps is much smaller than in triangle conserving rewiring. Based on our results for the latter ensemble, we would suggest that the time required to find any configurations significantly different from the original one (if they exist) are inaccessible with known algorithms. Hence, one cannot be sure of their claim that the $3K$ ensembles reproduce all the local and global properties for most undirected networks.

IV. CONCLUSION

In highly clustered networks—including most real world networks–clustering is concentrated among the highest degree nodes. The Strauss model correctly points to an important feature: clustering is cooperative. Once many triangles appear in a certain part of the network, they actuate the creation of more triangles in the region leading to a positive feedback loop. Thus, clustering cannot be smoothly and evenly introduced into a network; it is often driven by densely interconnected, high-degree regions. In triangle biased methods these high-degree regions emerge suddenly and thereafter prove resistant to subsequent randomization by rewiring.

The biased rewiring model put forward in this paper samples the ensemble of networks with an exponential weight, similar to the Strauss model. The density of triangles is controlled by a conjugate fugacity or triangle bias β. However, we impose a fixed degree sequence and thus avoid the catastrophic increase of connectivity at the phase transition as in the Strauss model. Yet there the first-order transitions remain.

For regular networks, i.e., those where every node has the same degree, there is a single phase transition. In the phase with strong clustering (large fugacity), the dominant configurations are a collection of disjoint k-cliques for $k>2$. If the degree sequence is inhomogeneous, the formation of *cluster*

cores no longer happens at the same β for different parts of the network. Thus the single phase transition is replaced by a sequence of discrete, discontinuous jumps, which resemble both first order transitions and Barkhausen jumps. As in the original Barkhausen phenomenon, frozen randomness is crucial for the multiplicity of jumps. There, each jump corresponds to a *flip* of a spin cluster *already defined* by the randomness—at least at zero temperature [33,34]. In the present case, however, each jump corresponds to the *creation* of a cluster whose detailed properties are not only determined by the quenched randomness (the degree sequence), but also depend on the "thermal" (nonquenched) noise.

As in any first-order phase transition, our model shows strong hysteresis. Cluster cores, once formed, are extremely stable and cannot be broken up easily. They form virtually permanent communities where the intracommunity links far outnumber the intercommunity ones. Indeed a standard community detection algorithm [35] detects the cluster cores together with other communities, which concentrate fewer triangles together.

This hysteresis limits the usefulness of triangle biased or triangle conserving rewiring as a null model. Even though the phase transitions do not appear explicitly in the null model of [17], the underlying dynamics are the same. Due to the very long time scales involved, Monte Carlo methods cannot sample evenly from these ensembles, but remain localized near the starting configuration. Care should be taken to demonstrate that results found using them are broadly consistent across various sampling procedures. This problem seems to us to be even more acute for the dK randomized ensembles in [38]. In the $3K$ ensembles not only is the total number of triangles conserved but the number of triangles and wedges with any triple of degrees (k, k', k'') is also fixed.

The spontaneous emergence of cluster cores in the BRM suggests that triangle bias can give rise to community structure in networks, without the need to define communities *a priori*–thanks to the cooperativity of triangle formation. Together with jumps in the number of triangles, or the clustering coefficient, simultaneous jumps in many other network properties appear at the same control parameter. Both assortativity and modularity emerge as a consequence of the clustering bias. In particular, we found jumps in the number of k-cliques with $k > 3$, in the assortativity, and in the modularity. Hence, all these quantities serve as order parameters associated with phase transitions in the BRM. This immediately raises the question whether the model can be generalized so that a different, independent fugacity is associated to each order parameter. For assortativity, this was proposed some time ago by Newman [12]. In the present notation, biased rewiring models with and without a target triangle number $n_{\Delta,0}$ and target assortativity r_0 are given by the Hamiltonians

$$H_{\text{Milo}}(G; \beta, \gamma) = \beta |n_\Delta(G) - n_{\Delta,0}| + \gamma |r(G) - r_0| \quad (16)$$

and

$$H_{\text{BRM}}(G; \beta, \gamma) = -\beta n_\Delta(G) - \gamma r(G), \quad (17)$$

respectively, where γ is the conjugate fugacity associated to the assortativity. It is an open question whether such a model might lead to less extreme clustering and thus might be more realistic. First simulations [55] indicate that driving assortativity leads to smooth increases in all other quantities, without jumps. The reason seems to be that the basic mechanism leading to increased assortativity—the replacement of existing links by links between similar nodes—is not cooperative, and doesn't lead to a positive feedback loop. But further studies are needed to test this conjecture.

As Newman remarked in [53], clustering in networks "has proved difficult to model mathematically." In that paper, he introduced a model where each link can participate in one triangle at most. In this way, the phase transitions seen in the Strauss model and in the present model are avoided. However, in the real-world networks studied here we found that the number of triangles in which a link participates is broadly distributed, suggesting that the model [53] may not be realistic for networks with significant clustering. Indeed, for each graph, specifying for each link the number of triangles (or more generally q-cliques) in which it participates adds valuable information to the adjacency matrix (which just specifies whether the link exists or not). The resulting "three-clique adjacency plots" revealed community structure that would not have been easy to visualize otherwise. These adjacency plots may turn out to be useful in many other contexts. For instance, it would be interesting to analyze other models for complex networks with significant clustering [3,21–25] and compare them to real-world networks using our adjacency plots. Indeed the quest for understanding clustering in complex networks is not yet finished.

[1] R. Guimera, L. Danon, A. Diaz-Guilera, F. Giralt, and A. Arenas, Phys. Rev. E **68**, 065103(R) (2003).

[2] M. E. J. Newman and J. Park, Phys. Rev. E **68**, 036122 (2003).

[3] J. Davidsen, H. Ebel, and S. Bornholdt, Phys. Rev. Lett. **88**, 128701 (2002).

[4] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener, Comput. Netw. **33**, 309 (2000).

[5] M. Baiesi and M. Paczuski, Phys. Rev. E **69**, 066106 (2004).

[6] J. Davidsen, P. Grassberger, and M. Paczuski, Geophys. Res. Lett. **33**, L11304 (2006).

[7] D. Hughes, M. Paczuski, R. O. Dendy, P. Helander, and K. G. McClements, Phys. Rev. Lett. **90**, 131101 (2003).

[8] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A. Barabási, Nature (London) **407**, 651 (2000).

[9] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Phys. Rev. E **64**, 026118 (2001).

[10] M. E. J. Newman, Phys. Rev. E **68**, 026121 (2003).

[11] S. Wasserman and K. Faust, *Social Network Analysis: Methods*

and Applications (Cambridge University Press, Cambridge, England, 1994).

[12] M. E. J. Newman, Phys. Rev. Lett. **89**, 208701 (2002).

[13] M. E. J. Newman and M. Girvan, Phys. Rev. E **69**, 026113 (2004).

[14] J. G. Foster, D. V. Foster, P. Grassberger, and M. Paczuski, Phys. Rev. E **76**, 046112 (2007).

[15] P. Erdos and A. Renyi, Publ. Math. (Debrecen) **6**, 156 (1959).

[16] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).

[17] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon, Science **298**, 824 (2002).

[18] M. Molloy and B. Reed, Random Struct. Algorithms **6**, 161 (1995).

[19] S. Maslov and K. Sneppen, Science **296**, 910 (2002).

[20] K. Baskerville, P. Grassberger, and M. Paczuski, Phys. Rev. E **76**, 036107 (2007).

[21] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, Phys. Rev. E **63**, 062101 (2001).

[22] P. Holme and B. J. Kim, Phys. Rev. E **65**, 026107 (2002).

[23] A. Barrat and R. Pastor-Satorras, Phys. Rev. E **71**, 036127 (2005).

[24] E. Volz, Phys. Rev. E **70**, 056115 (2004).

[25] M. A. Serrano and M. Boguñá, Phys. Rev. E **72**, 036133 (2005).

[26] B. Bollobas, *Random Graphs* (Cambridge University Press, Cambridge, England, 2001).

[27] J. Berg and M. Lässig, Phys. Rev. Lett. **89**, 228701 (2002).

[28] J. Park and M. E. J. Newman, Phys. Rev. E **70**, 066117 (2004).

[29] Z. Burda, J. Jurkiewicz, and A. Krzywicki, Phys. Rev. E **69**, 026106 (2004).

[30] J. Park and M. E. J. Newman, Phys. Rev. E **72**, 026136 (2005).

[31] This transition can also be seen as a first order *percolation* transition, since a giant percolating cluster is formed when the triangle fugacity is increased through the critical point. It is, however, substantially different from 'explosive percolation' in Achlioptas processes [54], which is also a first order percolation transition. While the Strauss model is a genuine thermodynamic model with Hamiltonian structure, and the phase transition happens as a true control parameter is increased, explosive percolation is a strictly nonequilibrium process where the control is done via a *density* of established bonds. One could also try to control the bond density in the Strauss model. This would lead to a continuous transition with phase coexistence, as in any system that undergoes a thermodynamic first order transition.

[32] H. D. Young, R. A. Freedman, T. R. Sandin, and A. L. Ford, *Sears and Zemansky's University Physics*, 10th ed. (Addison-Wesley, Reading, MA, 1999).

[33] O. Perković, K. Dahmen, and J. P. Sethna, Phys. Rev. Lett. **75**, 4528 (1995).

[34] P. Cizeau, S. Zapperi, G. Durin, and H. E. Stanley, Phys. Rev. Lett. **79**, 4669 (1997).

[35] A. Clauset, M. E. J. Newman, and C. Moore, Phys. Rev. E **70**, 066111 (2004).

[36] M. A. Serrano and M. Boguñá, Phys. Rev. E **74**, 056114 (2006).

[37] M. E. J. Newman and G. T. Barkema, *Monte Carlo Methods in Statistical Physics* (Oxford University Press, New York, 1999).

[38] A. Jamakovic, P. Mahadevan, A. Vahdat, M. Boguñá, and D. Krioukov, e-print arXiv:0908.1143.

[39] D. J. Watts and S. H. Strogatz, Nature (London) **393**, 440 (1998).

[40] A. Barrat and M. Weigt, Eur. Phys. J. B **13**, 547 (2000).

[41] C. E. Porter, Journal of Computer-Mediated Communication **10**, 00 (2004).

[42] S. Fortunato, Phys. Rep. **486**, 75 (2010).

[43] M. E. J. Newman, Phys. Rev. E **69**, 066133 (2004).

[44] W. K. Hastings, Biometrika **57**, 97 (1970).

[45] M. E. J. Newman, Proc. Natl. Acad. Sci. U.S.A. **98**, 404 (2001).

[46] A. Gavin *et al.*, Nature (London) **415**, 141 (2002).

[47] M. Newman, Proc. Natl. Acad. Sci. U.S.A. **103**, 8577 (2006).

[48] E. Ravasz and A. L. Barabási, Phys. Rev. E **67**, 026112 (2003).

[49] S. N. Soffer and A. Vázquez, Phys. Rev. E **71**, 057101 (2005).

[50] E. Ravasz, A. L. Somera, D. A. Mongru, Z. N. Oltvai, and A.-L. Barabási, Science **297**, 1551 (2002).

[51] E. A. Leicht, P. Holme, and M. E. J. Newman, Phys. Rev. E **73**, 026120 (2006).

[52] Y. Ahn, J. Bagrow, and S. Lehmann, arXiv:0903.3178 (unpublished).

[53] M. E. J. Newman, Phys. Rev. Lett. **103**, 058701 (2009).

[54] D. Achlioptas, R. D'Souza, and J. Spencer, Science **323**, 1453 (2009).

[55] D. V. Foster *et al.* (unpublished).