UNIVERSITY OF BIRMINGHAM

University of Birmingham
Research at Birmingham

Relaxation dynamics of maximally clustered networks

Klaise, Janis; Johnson, Samuel

[Link to publication on Research at Birmingham portal](#)

Relaxation dynamics of maximally clustered networks

Janis Klaise[1],[*] and Samuel Johnson[2],[†]

[1]*Centre for Complexity Science, University of Warwick, Coventry CV4 7AL, United Kingdom*
[2]*School of Mathematics, University of Birmingham, Birmingham B15 2TT, United Kingdom*

We study the relaxation dynamics of fully clustered networks (maximal number of triangles) to an unclustered state under two different edge dynamics—the double-edge swap, corresponding to degree-preserving randomization of the configuration model, and single edge replacement, corresponding to full randomization of the Erdős-Rényi random graph. We derive expressions for the time evolution of the degree distribution, edge multiplicity distribution and clustering coefficient. We show that under both dynamics networks undergo a continuous phase transition in which a giant connected component is formed. We calculate the position of the phase transition analytically using the Erdős-Rényi phenomenology.

I. INTRODUCTION

Network science has enjoyed unprecedented popularity in the past two decades as a paradigm for studying complex, interacting systems such as the Internet [1,2], the World Wide Web [3], food webs [4], scientific collaboration networks [5], social and biological networks [6–8], contact networks [9,10], and many others [11]. Many of these empirical networks exhibit a high degree of *clustering* or *transitivity*, i.e., a significant number of short, closed loops forming triangles [12,13]. This phenomenon is most commonly quantified by the clustering coefficient, defined as the proportion of connected triads that are also triangles in a network [14].

The classical random network models, the Erdős-Rényi (ER) random graph [15,16] and the configuration model [14,17,18], both suffer from being unable to generate networks with significant values of the clustering coefficient, thus making them unsuitable for modeling many real networks. High values of the clustering coefficient observed in empirical networks have lead to a surge of random network models that are capable of generating significant numbers of triangles [13,19–22]. The relationship between clustering and other network properties has also been studied extensively [23–30]. However, despite the large body of research, the inherent violation of edge independence in highly clustered networks has made it difficult to understand the full implications of clustering. Common issues encountered when dealing with highly clustered networks include difficulties of network sampling [22,31], inability to use edge independence to derive accurate results [26,29], and potentially overstated inferences of causality [28,30]. This points to a need for more fundamental research in clustered networks.

In this paper we explore simple dynamics of highly clustered networks relaxing to an unclustered equilibrium state. Specifically, we study the evolution of the clustering coefficient under two edge rewiring schemes starting with fully clustered,

degree-regular networks, i.e., networks in which all nodes have the same number of neighbors and a maximal number of triangles. We find that under both dynamics whose equilibrium distributions correspond to the Erdős-Rényi random graph and the configuration model, respectively, a giant connected component emerges via a continuous phase transition. We provide an analytical prediction of the critical point for this transition as well as derive time evolution equations for various network properties.

II. METHODS

A. Network metrics

We consider undirected graphs with N nodes and L edges described by a symmetric $N \times N$ adjacency matrix $\mathbf{A}$ with binary edge variables $A_{ij} \in \{0,1\}$ for $i, j \in \{1, \ldots, N\}$ with $A_{ij} = 1, i \neq j$ indicating an edge between nodes i and j so that $L = \sum_{i,j} A_{ij}$. The degree distribution of a network is defined as $p_k = N_k/N$, where N_k is the number of nodes with degree k. We denote the nth moment of the degree distribution by $\langle k^n \rangle$.

We define the *multiplicity* m_{ij} of an edge ij to be the number of triangles it participates in [24]. Similarly to the degree distribution, we define the edge multiplicity (or simply multiplicity) distribution as $q_m = L_m/L$, where L_m is the number of edges with multiplicity m. We denote the nth moment of the multiplicity distribution by $\langle m^n \rangle$.

The clustering coefficient of a network is defined as 3 times the number of triangles divided by the number of connected triples, i.e., $C = 3N_\triangle/N_\wedge$ [14]. This measure of clustering is properly normalized so that $C \in [0,1]$. It also admits a probabilistic interpretation—it is the probability that a randomly chosen triple of nodes is closed.

We can express the clustering coefficient in terms of the degree and multiplicity distributions. For any network we have

$$N_\wedge = \sum_k \binom{k}{2} N_k = N \sum_k \binom{k}{2} p_k = N \frac{\langle k^2 \rangle - \langle k \rangle}{2} \quad (1)$$

[*]J.Klaise@warwick.ac.uk
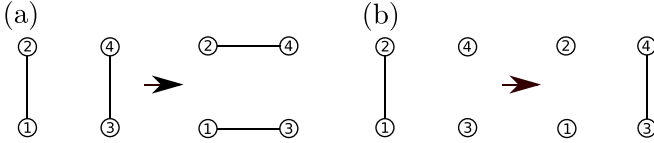[†]s.johnson.4@bham.ac.uk

FIG. 1. (a) Double edge swap or degree-preserving randomization. (b) Edge replacement or full randomization.

and

$$3N_\triangle = \sum_m L_m = L \sum_m m q_m = L \langle m \rangle. \qquad (2)$$

Putting the above results together and noting that in any network $L = N \langle k \rangle / 2$, we obtain the following general expression for the clustering coefficient:

$$C = \frac{\langle k \rangle \langle m \rangle}{\langle k^2 \rangle - \langle k \rangle}. \qquad (3)$$

B. Random network ensembles

We study relaxation dynamics of k-regular networks under edge rewiring in two random network ensembles—the configuration model (CM) and the ER random graph.

The CM [14,17,18] is defined by drawing a valid degree sequence $\mathbf{k} = \{k_i\}_{i=1}^N$ from a degree distribution p_k and producing a network realization uniformly at random from all possible networks with that degree sequence [14,32]. Provided the second moment of the degree distribution remains finite, it can be shown that the clustering coefficient scales as $C \sim 1/N$ so that in the thermodynamic limit ($N \to \infty$) the resulting networks are treelike [14].

The ER random graph [15,16,33] is defined by placing L edges uniformly at random between N nodes [34]. If we require that the mean degree $\langle k \rangle = 2L/N$ be fixed, then the degree distribution of the ER model in the thermodynamic limit is Poisson with mean $\langle k \rangle$ [14]. The ER model is thus a special case of the CM and has the same scaling behavior of the clustering coefficient.

Given that both the CM and ER random graphs are asymptotically triangle free, it is natural to consider them as equilibrium ensembles for relaxation dynamics of highly clustered networks into an unclustered state. To this end we describe two edge rewiring mechanisms that have the CM and the ER random graphs as equilibrium distributions (see Fig. 1 for a graphical demonstration).

1. Double edge swap

The double-edge swap [35,36] is defined by choosing two existing edges in the network at random and rewiring their ends to produce two new edges while deleting the original two. This is also known as *degree-preserving* randomization and so naturally produces network realizations in the CM ensemble with a fixed degree sequence. The double-edge swap defines a Markov chain whose equilibrium distribution is the CM [35].

2. Edge replacement

Alternatively, one can fully randomize a network by picking an edge at random and placing it anywhere in the network
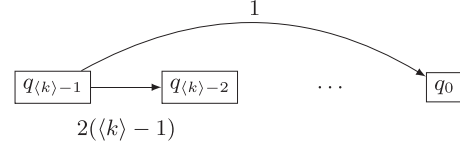

FIG. 2. Transition rates in the multiplicity distribution for a single clique of size $\langle k \rangle + 1$.

where there is no edge already [37,38]. In this scheme the number of edges is preserved but the degrees of the nodes are not. Edge replacement defines a Markov chain whose equilibrium distribution is the ER ensemble.

A double-edge swap or an edge replacement constitutes an *elementary rewiring step*.

III. RESULTS

To assess the evolution of network measures over time, we take into account the network size and the rewiring scheme (either CM or ER) to normalize the number of elementary rewiring steps per number of edges. If r_{CM} and r_{ER} are the number of elementary rewiring steps in the CM and ER ensembles, respectively, then we define the corresponding time variables as

$$\begin{aligned} t_{ER} &= \frac{r_{ER}}{L} \\ t_{CM} &= \frac{2r_{CM}}{L}. \end{aligned} \qquad (4)$$

These definitions have the useful interpretation that when $t_{scheme} = 1$, the rewiring scheme has, on average, modified each edge in the network.

A. Multiplicity distribution

The multiplicity distribution evolves over time as edges are rewired and triangles are destroyed. The initial configuration of a k-regular network is a disjoint union of $N/(\langle k \rangle + 1)$ cliques of size $\langle k \rangle + 1$, which ensures maximal clustering $C = 1$. In other words, at time $t = 0$, the multiplicity distribution is

$$\begin{aligned} q_{\langle k \rangle - 1} &= 1 \\ q_m &= 0 \text{ if } m \neq \langle k \rangle - 1. \end{aligned} \qquad (5)$$

Consider the smallest informative time step $\Delta t_{CM} = 2/L$ or $\Delta t_{ER} = 1/L$ corresponding to exactly one elementary rewiring step. At $t = 0$ a clique of size $\langle k \rangle + 1$ has exactly $\binom{\langle k \rangle + 1}{2}$ edges, all of which have maximal multiplicity $\langle k \rangle - 1$. Rewiring any single edge will destroy $\langle k \rangle - 1$ triangles leading to a decrease of $2(\langle k \rangle - 1) + 1$ edges with maximal multiplicity, one for the rewired edge and an additional two for each destroyed triangle. Assuming that no new triangles are created, the single rewired edges will have multiplicity zero. Figure 2 shows the transition rates in the multiplicity distribution of a single clique.

We now make the *ansatz* that this is the main way the multiplicity distribution changes over time—multiplicity is predominantly decreased by rewiring single edges from cliques and all such rewirings are independent. In this case, we can write down the full transition rate diagram between multiplicity classes as shown in Fig. 3. This gives the following time
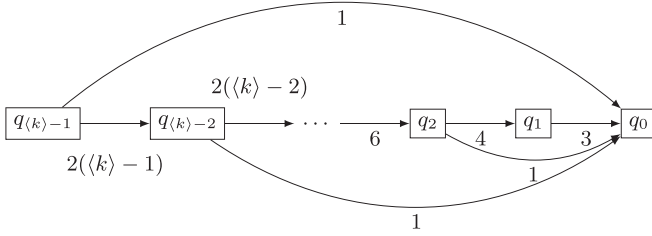
FIG. 3. Transition rates in the full multiplicity distribution.

evolution equations for q_m:

$$\frac{dq_m}{dt} = -(2m+1)q_m + 2(m+1)q_{m+1}, \quad \begin{array}{l} \text{for } m = \\ \langle k \rangle - 1, \ldots, 1 \end{array}$$

$$\frac{dq_0}{dt} = 3q_1 + \sum_{m=2}^{\langle k \rangle - 1} q_m. \tag{6}$$

Figure 4 shows the numerical solution of this system of ordinary differential equations (ODEs) which is in excellent agreement with simulation results. The calculations are valid both in the ER and the CM cases.

1. Average multiplicity

Using the time evolution equations for the multiplicity distribution, we can derive exact expressions of its moments. Specifically, we are interested in the average multiplicity $\langle m \rangle$ as it features in the expression for the clustering coefficient. We have

$$\frac{d\langle m \rangle}{dt} = \sum_{m=1}^{\langle k \rangle - 1} m \frac{dq_m}{dt}. \tag{7}$$

Inserting Eq. 6, we obtain the simple expression

$$\frac{d\langle m \rangle}{dt} = -3\langle m \rangle. \tag{8}$$
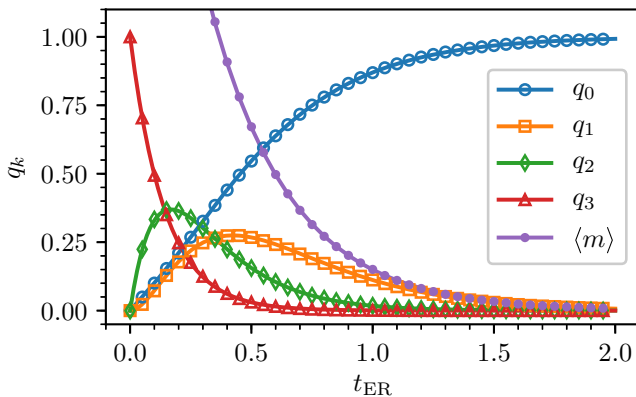

FIG. 4. Evolution of the multiplicity distribution in an ER network with average degree $\langle k \rangle = 4$ and $N = 10^5$. The solid lines are numerical solutions of Eq. 6 while the markers are simulation results. The purple line with filled circles indicates the average multiplicity $\langle m \rangle$.

Using the initial condition $\langle m \rangle(0) = \langle k \rangle - 1$, this has the solution

$$\langle m \rangle = (\langle k \rangle - 1)e^{-3t}. \tag{9}$$

Figure 4 shows the analytic solution of the average multiplicity, which is in perfect agreement with simulation results.

B. Degree distribution

In the case of the ER model, the degree distribution is also changing over time. Consider the degree distribution $p_k(t)$ as a function of time and a time step Δt_{ER}. We can calculate the rate at which $p_k(t)$ changes.

An edge replacement event in the ER model consists of two steps. First, a random edge is selected. Second, a random pair of nodes that are not linked by an edge (let us call this pair a *nonedge*) is selected and the edge selected in the first step is deleted while the nonedge becomes an edge.

When a random edge is selected, p_k can decrease if at least one end of the edge has degree k. Alternatively, p_k can increase if at least one end of the edge has degree $k + 1$. The probability of reaching a node of degree k by following a randomly chosen edge is given by the so-called excess degree distribution [14] which reads $s_k = kp_k/\langle k \rangle$. Given this and the fact that a randomly chosen edge can have 0, 1, or 2 nodes of degree k, we can calculate the expected number of nodes of degree k at the ends of a random edge:

$$\mathbb{E}(k \to k-1) = 2s_k^2 + 2s_k(1 - s_k) = 2s_k = 2\frac{kp_k}{\langle k \rangle}. \tag{10}$$

This is the expected number of nodes whose degree would decrease from k to $k - 1$ during a single edge selection step. Note that at the beginning of the process the degree distribution is regular so $\mathbb{E}(k \to k-1) = 2$ as expected.

Similarly, the expected number of nodes whose degree would decrease from $k + 1$ to k leading to an increase in p_k is

$$\mathbb{E}(k+1 \to k) = 2s_{k+1} = 2\frac{(k+1)p_{k+1}}{\langle k \rangle}. \tag{11}$$

Now consider the second step in the edge replacement event, the selection of a nonedge. When a random nonedge is selected, p_k can also change in two ways. It can increase if at least one of the selected nodes has degree $k - 1$ and it can decrease if at least one of the nodes has degree k. The calculation of the expected number of nodes changed as a result of this is similar to the previous case, but we must consider the distribution of *nondegrees* instead. To this end we study the graph complement of the original network defined as a network in which two nodes are linked if and only if they are not linked in the original network. From here on we denote by an overbar quantities in the graph complement.

It is easy to see that the degrees of nodes in the complement are given by $\bar{k} = N - 1 - k$ where k is the degree of a node in the original network and we have $p_{\bar{k}} = p_k$. Thus, the nonedges are selected proportionally to $\bar{k}$ not k as in the case of edge selection so we must work with the excess nondegree distribution given by $s_{\bar{k}} = \bar{k}p_k/\langle \bar{k} \rangle$. Note that the mean nondegree is given by

$$\langle \bar{k} \rangle = \sum_k \bar{k}p_k = N - 1 - \langle k \rangle. \tag{12}$$
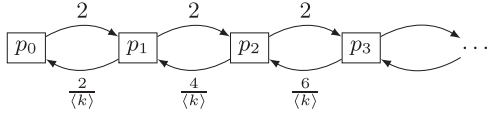
FIG. 5. Transition rates in the degree distribution under the ER model.

As in the case of edge selection, the expected number of nodes whose degree would increase from k to $k + 1$, thus reducing p_k during a single nonedge selection step is

$$\mathbb{E}(k \to k + 1) = 2q_{\bar{k}} = 2\frac{\bar{k}p_k}{\langle \bar{k} \rangle} = \frac{2(N - 1 - k)}{N - 1 - \langle k \rangle}p_k. \quad (13)$$

When N is large we can approximate this by

$$\mathbb{E}(k \to k + 1) \simeq 2p_k. \quad (14)$$

Similarly, p_k can increase if we select a nonedge with at least one node with degree $k - 1$. The expected number of such nodes in a single nonedge selection is

$$\mathbb{E}(k - 1 \to k) \simeq 2p_{k-1}. \quad (15)$$

Figure 5 describes pictorially the transition rates between degree classes as derived here. This allows us to write down the time evolution equations for p_k:

$$\frac{dp_k}{dt} = 2p_{k-1} - 2\left(1 + \frac{k}{\langle k \rangle}\right)p_k + 2\frac{k+1}{\langle k \rangle}p_{k+1}, \quad (16)$$

for $k = 0, 1, \ldots$. This system of ODEs is not closed, so in order to solve it numerically, we must truncate the system at some p_{k^*} setting $p_k = 0$ for all $k > k^*$. The value of k^* should be set high enough so the probability mass unaccounted for is minimal for accurate predictions. We test our predictions by numerically solving the ODEs for a network with average degree $\langle k \rangle = 2$ and setting the cutoff $k^* = 8$. The results are shown in Fig. 6. The numerical solution of the ODE system is in excellent agreement with simulation results. We also note that the cutoff is appropriate for this level of approximation as the
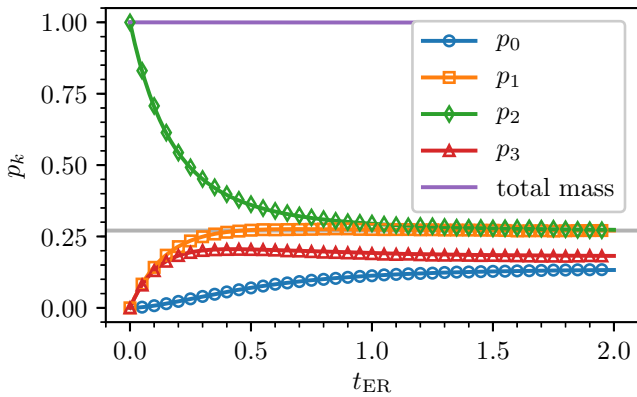

FIG. 6. Evolution of the degree distribution in an ER network with average degree $\langle k \rangle = 2$ and $N = 10^5 - 1$. The solid lines are numerical solutions of Eq. 16 while the markers are simulation results. The gray line indicates the equilibrium value of p_2 in an ER ensemble. The purple line indicates the total probability mass in the system accounted for by truncating the ODE system at $k^* = 8$.

total probability mass does not diverge from unity noticeably over the time period considered.

1. Second moment of the degree distribution

Using the time evolution equations for the degree distribution, we can derive exact expressions of its moments. Specifically, we are interested in the second moment $\langle k^2 \rangle$. We have

$$\frac{d\langle k^2 \rangle}{dt} = \sum_k k^2 \frac{dp_k}{dt}. \quad (17)$$

Inserting Eq. 16 we obtain the simple expression

$$\frac{d\langle k^2 \rangle}{dt} = -4\frac{\langle k^2 \rangle}{\langle k \rangle} + 4\langle k \rangle + 4. \quad (18)$$

Using the initial condition $\langle k^2 \rangle(0) = \langle k \rangle^2$ and recalling that the average degree $\langle k \rangle$ is constant, this has solution

$$\langle k^2 \rangle = \langle k \rangle(\langle k \rangle + 1 - e^{-\frac{4t}{\langle k \rangle}}). \quad (19)$$

C. Clustering coefficient

Putting together the results for the multiplicity and degree distributions, and using Eq. 3, we obtain exact expressions for the clustering coefficient as a function of time in both the CM and ER ensembles:

$$C_{\text{CM}} = e^{-3t}$$
$$C_{\text{ER}} = \frac{(\langle k \rangle - 1)e^{-3t}}{\langle k \rangle - e^{\frac{-4t}{\langle k \rangle}}}. \quad (20)$$

We note that in the CM ensemble, the clustering coefficient has no dependence on the average degree while this is not the case for the ER ensemble. This is because the number of connected triples $N_\wedge$ in the CM ensemble is constant by virtue of having a fixed degree sequence while it is dependent on the evolving degree sequence in the ER ensemble.

D. Giant connected component

We find that under both rewiring schemes there is an emergence of global connectivity via the appearance of a giant connected component (GCC) at some critical time t^c (equivalently, critical clustering coefficient C^c). We confirm from simulation results that a GCC emerges in a continuous phase transition Figs. (7 and 8 for the CM and Figs. 9 and 10 for the ER ensembles). Note that the large fluctuations in the 2-regular case is due to the fact that 2-regular networks are exactly at the poing of criticality in the unclustered CM case ($C = 0$). This phenomenon has been studied in the context of reversible polymerization of rings [39].

We can calculate the critical point analytically by using the known result that a GCC in an ER random graph emerges when $\langle k \rangle = 1$ [14]. We conjecture that a GCC induced by edge rewiring emerges when the average number of external edges between the original $N/(\langle k \rangle + 1)$ cliques of size $\langle k \rangle + 1$ exceeds 1. If this is the case, then the critical number of elementary rewiring steps is

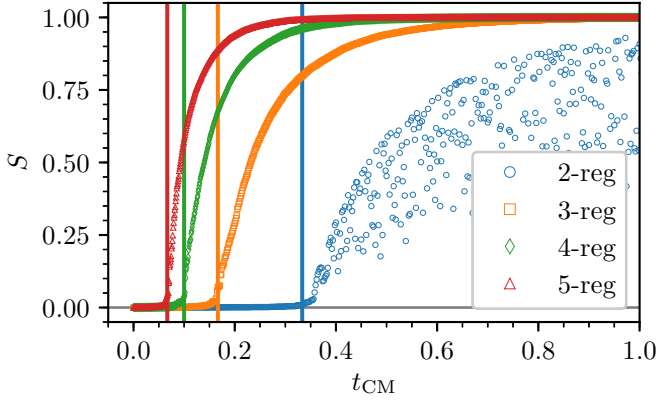$$r^c = \frac{N}{2(\langle k \rangle + 1)}. \quad (21)$$

FIG. 7. Proportion of nodes S in the giant connected component as a function of time t_{CM} for a few select k-regular networks. We observe a continuous phase transition at a critical point t_{CM}^c which depends on the average degree of the network as explained in the main text. Vertical lines correspond to the analytically calculated critical points.

FIG. 9. Proportion of nodes S in the giant connected component as a function of time t_{ER} for a few select mean degree $\langle k \rangle$ networks. We observe a continuous phase transition at a critical point t_{ER}^c which depends on the average degree of the network as explained in the main text. Vertical lines correspond to the analytically calculated critical points.

Expressing this in terms of the time variable, we obtain the critical time for both the CM and the ER rewiring schemes:

$$
\begin{aligned}
t_{CM}^c &= \frac{2}{\langle k \rangle (\langle k \rangle + 1)} \\
t_{ER}^c &= \frac{1}{\langle k \rangle (\langle k \rangle + 1)}.
\end{aligned}
\tag{22}
$$

Note that these differ by a factor of two. This is because in the CM rewiring scheme, even though every elementary rewiring step involves two edges, the two rewirings are not independent—during one rewiring step it is possible to connect at most two disconnected components.
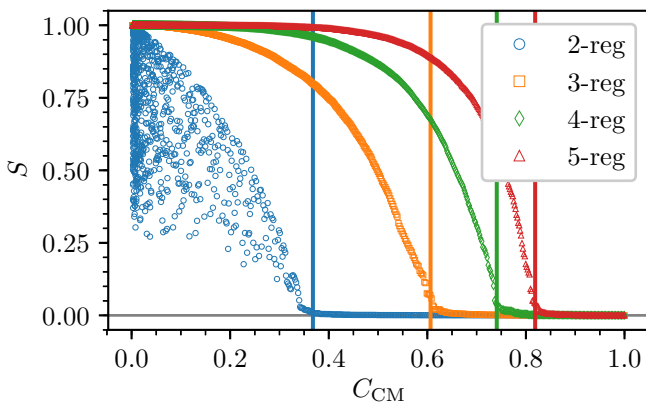
Expressed in terms of the clustering coefficient, the critical thresholds read:

$$
\begin{aligned}
C_{CM}^c &= e^{-6/\langle k \rangle (\langle k \rangle + 1)} \\
C_{ER}^c &= \frac{(\langle k \rangle - 1) e^{-3/\langle k \rangle (\langle k \rangle + 1)}}{\langle k \rangle - e^{-4/\langle k \rangle^2 (\langle k \rangle + 1)}}.
\end{aligned}
\tag{23}
$$

Figures 7 and 8 confirm that these are in excellent agreement with simulations in the CM case and Figs. 9 and 10 confirm a good agreement in the ER case which improves as the mean degree increases.

What is the cause of the discrepancy of the analytical result for the critical point and the numerical simulations, particularly for low mean degree ER networks? We conjecture that this is due to some edges being rewired multiple times while others are not rewired at all. This would have the effect of increasing the critical time because we have to wait slightly

FIG. 8. Proportion of nodes S in the giant connected component as a function of clustering C_{CM} for a few select k-regular networks under the CM rewiring scheme. We observe a continuous phase transition at a critical point C_{CM}^c which depends on the average degree of the network as explained in the main text. Vertical lines correspond to the analytically calculated critical points.
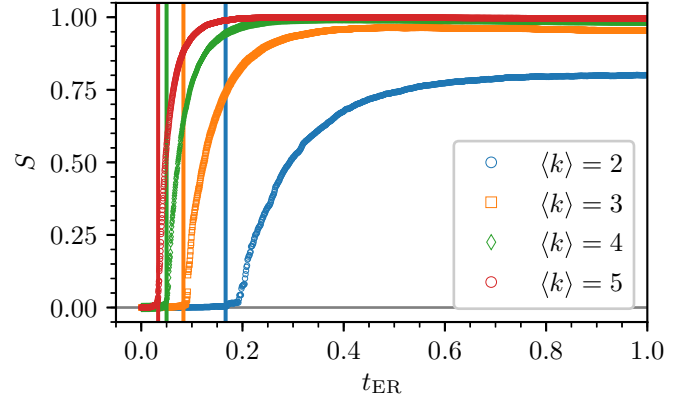

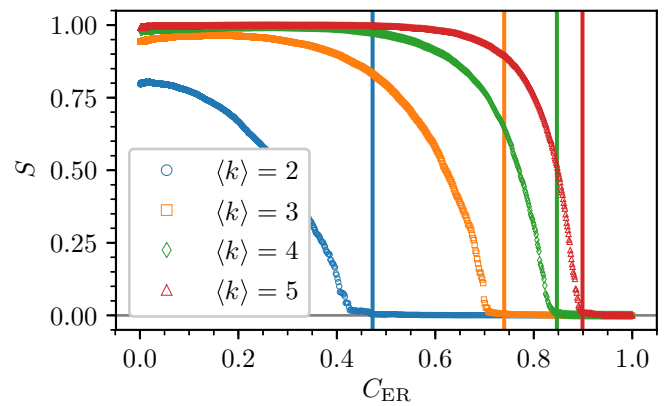FIG. 10. Proportion of nodes S in the giant connected component as a function of clustering C_{ER} for a few select mean degree $\langle k \rangle$ networks under the ER rewiring scheme. We observe a continuous phase transition at a critical point C_{ER}^c which depends on the average degree of the network as explained in the main text. Vertical lines correspond to the analytically calculated critical points.

longer until the average number of rewired edges *discouting edges rewired multiple times* reaches the point where long range connectedness emerges. Figure 9 seems to confirm this to be the case. Let us calculate this revised critical time in the ER case.

During an edge replacement step, the probability of any edge being chosen for rewiring is $1/L$. So after r rewiring events the probability that a specific edge has not been rewired is

$$\mathbb{P}(\text{not rewired}) = \left(1 - \frac{1}{L}\right)^r. \qquad (24)$$

Substituting $r = Lt$ since we are in the ER case and taking the limit as $L \to \infty$, we get

$$\mathbb{P}(\text{not rewired}) = e^{-t}. \qquad (25)$$

The new revised time for the emergence of the GC, call it t^{r}, is then the time at which point this probability drops below a certain threshold. What is this threshold? It should be when the proportion of edges that have been rewired gives rise to a GCC which is precisely given by t^{c}. We can then write

$$e^{-t^{\text{r}}} = 1 - t^{\text{c}}. \qquad (26)$$

Note that by Taylor expansion we have $t^{\text{r}} \simeq t^{\text{c}}$ if this time is small as in the case when the average degree $\langle k \rangle \to \infty$. This explains why the t^{c} value becomes a better predictor for the critical threshold as the mean degree increases as seen in Fig. 9.

The revised critical point in the ER case is thus

$$t^{\text{r}} = -\log(1 - t^{\text{c}}) = \log\left[\frac{\langle k \rangle(\langle k \rangle + 1)}{\langle k \rangle(\langle k \rangle + 1) - 1}\right]. \qquad (27)$$

Figure 11 confirms that t^{r} is a better predictor of the location of the phase transition. The difference between t^{c} and t^{r} becomes negligible as the mean degree increases.

Another aspect that could influence the position of the critical point is the possibility of having rewired more edges than needed to connect previously disconnected components.
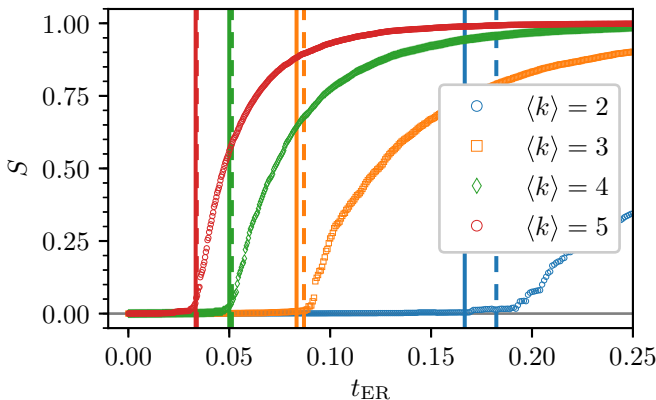

FIG. 11. Proportion of nodes in the giant connected component S as a function of time t_{ER} for a few select mean degree $\langle k \rangle$ networks. Solid vertical lines correspond to the critical time t_{ER}^{c} while dashed vertical lines correspond to the revised critical time t_{ER}^{r}.

We show in the Appendix that this should have no bearing on the critical point in large networks.

IV. DISCUSSION

In this paper we have studied the evolution of highly clustered networks under random edge rewiring dynamics. Our main result is showing the existence of a phase transition in which a giant connected component emerges. Del Genio and House [23] showed that equilibrium ensembles of degree-regular networks with prescribed clustering always admit a giant connected component. As a consequence, spreading processes such as infectious diseases in contact networks could always become endemic regardless of the level of clustering. By contrast, our work implies that this need not be the case in nonequilibrium systems. Depending on the precise mechanisms of time evolution of real networks and the level of clustering maintained, a giant connected component facilitating spreading processes may or may not exist. We have studied a model in which highly clustered populations undergo fully random connectivity changes and even in this simple scenario we observe two different modes of global connectivity.

Another interesting aspect of our work is from the perspective of statistical mechanics. A maximally clustered network is essentially the lowest entropy microstate in the context of the random network ensembles studied here. This is because such a network, under relabelling of nodes, is unique and least likely to be produced by chance at equilibrium. By using this configuration as a starting state for network dynamics, we have shown that the emergence of global connectivity is effectively delayed. This raises several other research questions, for example, is random rewiring the most or least effective method of delaying the onset of global connectivity? It is probable that more sophisticated rewiring methods involving choice, such as those studied in explosive percolation [40], would lead to different critical thresholds. We have also limited ourselves to studying rewiring that consistently destroys triangles, but what about rewiring with a view to increasing the number of triangles? A number of greedy as well as equilibrium algorithms exist and are widely applied to model highly clustered networks [28,41], but it is unclear how they cover the space of all networks and can lead to interesting behavior such as hysteresis loops [28]. Finally, there is a growing literature on generalized network structures such as simplicial complexes which allow for more general types of connections between nodes [42–45]. Simplicial complexes arise as natural generalizations of the static configuration model [42,45] and growing network models [43,44] for studying systems in which many-body rather than pairwise interactions are the norm. The methods developed in this paper could be extended to this framework by studying the effect of random rewiring of not just links but also triangles and higher-order cliques. Indeed, clustering in network structures still leaves much to be explored.

APPENDIX: EXTRANEOUS EDGES

Another mechanism that could change the location of the critical point t^c is the number of extraneous edges between already connected components. A GCC is formed when there are enough external edges between the initial cliques. Only one external edge is needed to connect two cliques, but there are multiple ways to do it and sometimes multiple edges end up linking together the same cliques. For example, we need only one edge two join two disconnected triangles, but there are a total of nine ways to do it; moreover, there is no guarantee that we will not end up with multiple edges between these triangles.

More generally, let the average degree $\langle k \rangle$ be fixed, and then at $t = 0$ there are $n = N/(\langle k \rangle + 1)$ cliques of size $\langle k \rangle + 1$. Any two cliques can therefore be connected in $(\langle k \rangle + 1)^2$ ways.

Suppose we never want to make more than one external edge to connect disconnected components. Then at $t = 0$ the number of choices for placing an external edge is given by

$$(\langle k \rangle + 1)\binom{n}{2} = \frac{N(N - \langle k \rangle - 1)}{2}. \qquad (A1)$$

After each rewiring event, the number of choices decreases by $(\langle k \rangle + 1)^2$, so after $r - 1$ rewires, the probability of placing an extraneous edge on the next rewire, r, is

$$\mathbb{P}\left(\text{extra edge on step } r\right) = \frac{2(\langle k \rangle + 1)^2 r}{N(N - \langle k \rangle - 1)}. \qquad (A2)$$

Thus, the expected number of extraneous edges after r rewiring events is

$$\mathbb{E}\left(\text{extra edges by step } r\right) = \sum_{r'=0}^{r} \frac{2(\langle k \rangle + 1)^2 r'}{N(N - \langle k \rangle - 1)}. \qquad (A3)$$

In particular, setting $r = r^c = N/2(\langle k \rangle + 1)$ we get

$$\mathbb{E}\left(\text{extra edges by step } r^c\right)$$
$$= \frac{2(\langle k \rangle + 1)^2}{N(N - \langle k \rangle - 1)} \frac{N}{4(\langle k \rangle + 1)}\left[\frac{N}{2(\langle k \rangle + 1)} + 1\right]$$
$$= \frac{N + 2\langle k \rangle + 2}{4(N - \langle k \rangle - 1)}. \qquad (A4)$$

Taking the limit $N \to \infty$, we get

$$\mathbb{E}\left(\text{extra edges by step } r_c\right) \simeq \tfrac{1}{4}, \qquad (A5)$$

which is fixed and independent of network size. Therefore, the formation of extraneous edges does not affect the location of the critical point in the large network limit.

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

[2] S. Johnson, J. J. Torres, J. Marro, and M. A. Muñoz, Phys. Rev. Lett. **104**, 108702 (2010).

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

[4] S. Johnson, V. Domínguez-García, L. Donetti, and M. A. Muñoz, Proc. Natl. Acad. Sci. USA **111**, 17923 (2014).

[5] M. E. J. Newman, Proc. Natl. Acad. Sci. USA **98**, 404 (2001).

[6] G. Robins, P. Pattison, Y. Kalish, and D. Lusher, Soc. Netw. **29**, 173 (2007).

[7] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A.-L. Barabasi, Nature **407**, 651 (2000).

[8] E. Bullmore and O. Sporns, Nat. Rev. Neurosci. **10**, 186 (2009).

[9] L. Danon, T. A. House, J. M. Read, and M. J. Keeling, J. R. Soc. Interface **9**, 2826 (2012).

[10] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, Rev. Mod. Phys. **87**, 925 (2015).

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

[12] S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications*, Vol. 8 (Cambridge University Press, Cambridge, 1994).

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

[14] M. Newman, *Networks: An Introduction* (Oxford University Press, New York, 2010).

[15] E. N. Gilbert, Ann. Math. Stat. **30**, 1141 (1959).

[16] P. Erdős and A. Rényi, Publ. Math. Debr. **6**, 290 (1959).

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

[18] B. Bollobás, *Random Graphs*, Cambridge Studies in Advanced Mathematics (Cambridge University Press, Cambridge, 2001).

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

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

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

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

[23] C. I. Del Genio and T. House, Phys. Rev. E **88**, 040801 (2013).

[24] M. Á. Serrano and M. Boguñà, Phys. Rev. E **74**, 056114 (2006).

[25] M. Á. Serrano and M. Boguñà, Phys. Rev. E **74**, 056115 (2006).

[26] M. Á. Serrano and M. Boguñà, Phys. Rev. Lett. **97**, 088701 (2006).

[27] A. Vázquez, R. Dobrin, D. Sergi, J.-P. Eckmann, Z. N. Oltvai, and A.-L. Barabási, Proc. Natl. Acad. Sci. USA **101**, 17940 (2004).

[28] D. Foster, J. Foster, M. Paczuski, and P. Grassberger, Phys. Rev. E **81**, 046115 (2010).

[29] Y. Berchenko, Y. Artzy-Randrup, M. Teicher, and L. Stone, Phys. Rev. Lett. **102**, 138701 (2009).

[30] D. V. Foster, J. G. Foster, P. Grassberger, and M. Paczuski, Phys. Rev. E **84**, 066117 (2011).

[31] S. Horvát, E. Czabarka, and Z. Toroczkai, Phys. Rev. Lett. **114**, 158701 (2015).

[32] C. I. Del Genio, H. Kim, Z. Toroczkai, and K. E. Bassler, PLoS ONE **5**, e10012 (2010).

[33] P. Erdős and A. Rényi, Publ. Math. Inst. Hung. Acad. Sci. **5**, 17 (1960).

[34] Another common definition leading to a slightly different model is to place each of the possible $\binom{N}{2}$ edges with equal probability p, but this does not enforce a fixed number of edges.

[35] A. R. Rao, R. Jana, and S. Bandyopadhyay, Sankhya **58**, 225 (1996).

[36] R. Milo, N. Kashtan, S. Itzkovitz, M. E. J. Newman, and U. Alon, arXiv:cond-mat/0312028 (2003).

[37] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabási, Nature **473**, 167 (2011).

[38] T. Nepusz and T. Vicsek, Nat. Phys. **8**, 568 (2012).

[39] E. Ben-Naim and P. L. Krapivsky, Phys. Rev. E **83**, 061102 (2011).

[40] R. M. D'Souza and J. Nagler, Nat. Phys. **11**, 531 (2015).

[41] M. Ritchie, L. Berthouze, T. House, and I. Z. Kiss, J. Theor. Biol. **348**, 21 (2014).

[42] O. T. Courtney and G. Bianconi, Phys. Rev. E **93**, 062311 (2016).

[43] G. Bianconi and C. Rahmede, Phys. Rev. E **93**, 032315 (2016).

[44] G. Bianconi and C. Rahmede, Sci. Rep. **7**, 41974 (2017).

[45] J.-G. Young, G. Petri, F. Vaccarino, and A. Patania, Phys. Rev. E **96**, 032312 (2017).