Network Reliability: The effect of local network structure on diffusive processes

Mina Youssef[1,*], **Yasamin Khorramzadeh**[1,2,†], and **Stephen Eubank**[1,2,3,‡]

[1]Network Dynamics and Simulation Science Laboratory, Virginia Bioinformatics Institute, Virginia Tech, Blacksburg, Virginia 24061, USA

[2]Department of Physics, Virginia Tech, Blacksburg, Virginia 24061, USA

[3]Department of Population Health Sciences, Virginia Tech, Blacksburg, Virginia 24061, USA

Abstract

This paper re-introduces the network reliability polynomial – introduced by Moore and Shannon in 1956 – for studying the effect of network structure on the spread of diseases. We exhibit a representation of the polynomial that is well-suited for estimation by distributed simulation. We describe a collection of graphs derived from Erd s-Rényi and scale-free-like random graphs in which we have manipulated assortativity-by-degree and the number of triangles. We evaluate the network reliability for all these graphs under a reliability rule that is related to the expected size of a connected component. Through these extensive simulations, we show that for positively or neutrally assortative graphs, swapping edges to increase the number of triangles does not increase the network reliability. Also, positively assortative graphs are more reliable than neutral or disassortative graphs with the same number of edges. Moreover, we show the combined effect of both assortativity-by-degree and the presence of triangles on the critical point and the size of the smallest subgraph that is reliable.

I. INTRODUCTION

We study the dynamics on a variety of networks for a networked $S – I – R$ model of epidemics, in which each vertex can be in one of the three states *Susceptible*, *Infectious*, or *Recovered* [1, 2]. As is well known, this process is equivalent to bond percolation [3], and thus exhibits a percolation phase transition and associated critical phenomena in an infinite network. The mean field dynamics are also well understood: critical phenomena such as scaling exponents depend only on the degree, also known as the coordination number. Corrections to mean field dynamics [4, 5] have been established that take into account variations in degree from one vertex to another [6, 7]. Often, following [8], the variation is taken to follow a power law distribution. However, the most important variation is not necessarily in degree, but in the number and overlaps of loops of a given length. Both the degree and the distribution of loops are completely determined by the dimension for regular grids, where much of the theory was developed, but not for generic graphs. In this paper, we illustrate how to use the concept of network reliability to elucidate how details of network topology influence the spread of epidemics. There are many structural aspects of contact networks that interact in complicated ways with each other and with the dynamical properties of disease transmission to create population-level dynamics in infectious disease outbreaks. For concreteness, we focus on the effect of degree assortativity and the number of

---
[*]myoussef@vbi.vt.edu
[†]yasi@vbi.vt.edu
[‡]seubank@vbi.vt.edu

triangles. As we show below, the complicated interaction between these structural measures generates a wide range of population-level effects.

We show how to characterize a network by the way its overall attack rate – the mean cumulative fraction of vertices infected before this transient dynamics reaches a fixed point – varies with disease transmissibility. Interventions alter the network structure, changing the overall attack rate. In [9], we found that isolating infected people within a household, i.e. limiting their contacts with other household members, can significantly reduce the population-level attack rate for a wide range of transmissibility. In this case we can characterize the network after intervention as uniformly more resistant to epidemic outbreak than the original network.

The overall attack rate is a special case of the *Network Reliability Polynomial*[10] formalism. This formalism was introduced to analyze specific networks. Hence, one of its strengths for characterizing networks is that it makes no assumptions about regularities or symmetries. We define and provide algorithms for calculating and estimating coefficients of the reliability polynomial, provide illustrative examples on several networks, and show how it can be used to understand complicated phenomena. The novelty in this work is not the concept of reliability itself – the *IEEE Transactions on Reliability* is now in its 61st year – nor is it in the statistical physics of reliability. It is in our suggestions that

1. coefficients of the reliability polynomial are among the best ways to characterize graph structure and

2. network analysis in terms of reliability provides insights into global effects of local structural details that elude other approaches.

Reliability refocuses the question of structural effects from the individual interactions between elements to global dynamical properties, suggesting new methods of analysis. In contrast to methods using coefficients of the characteristic polynomials of adjacency and Laplacian matrices [11, 12] or combinations of centrality measures [13] that describe the graph structure, the coefficients of the reliability polynomial transform all the information in the network adjacency matrix into a form that, by design, reflects dynamical phenomena of interest. Hence it is a unique structural measure that is immediately connected with dynamics. Network reliability is amenable to study from many perspectives, and much is known about the general properties of the reliability polynomial [14].

In contrast, the literature about the relation between dynamics and common graph statistics such as assortativity-by-degree and clustering coefficient is confusing and sometimes inconsistent. For example, consider what is known about the relationship between the spread of $S - I - R$ epidemics and assortativity-by-degree. Assortativity can be defined as a correlation coefficient between the degrees of vertices at each end of an edge. Thus it ranges from highly assortative (+1) through neutrally assortative (near 0) to highly disassortative (−1). The spread of $S - I - R$ epidemics in correlated and uncorrelated networks has been addressed in [15–20]. Given a social network, Nold in [16] grouped individuals based on their number of contacts. Thus, high epidemic prevalence appears in groups with highest number of contacts. In contrast to Nold, Moreno and Pacheco[15] reported that positively assortative networks have fewer large outbreaks than neutrally assortative networks. Moreover, for finite size networks, the epidemic threshold for positively assortative networks is larger than that for neutrally assortative networks, indicating more robustness against the spread of epidemics. Consistent with Newman[17], epidemics persist in positively assortative networks longer than in neutrally assortative networks when the initial infected vertex is the one with the largest node degree. Kiss and Kao[18] showed that epidemics spread faster in positively assortative networks than in disassortative (negatively

assortative) networks. However, this result disagrees with D'Agostino et al.[19], in which it is shown that disassortative networks have a shorter longest time to peak epidemic prevalence than assortative networks. The longest timescale is the inverse of the algebraic connectivity representing the slowest mode of diffusion in the network [21]. Disassortative networks have larger algebraic connectivity than assortative networks. In other words, disassortative networks have shorter longest timescale to the epidemic peak than assortative networks. Thus, epidemics spread faster in disassortative networks than in assortative networks. Finally, the combined impact of both assortativity-by-degree and clustering coefficient on the spread of epidemics is studied in Badham and Stocker[20]. Through extensive simulations on a limited set of networks, the authors found that both the total epidemic size and the average secondary infection size are smaller for highly clustered and/or highly positively assortative networks. However, for smaller values of these properties, the epidemic final size is inconsistent with the increase of either the assortativity value or the clustering coefficient.

Using the network reliability polynomial, we study the effect of both assortativity-by-degree and number of triangles on the spread of epidemics in Erdős-Rényi and scale-free-like random graphs. We evaluate the reliability for graphs we have generated that exhibit a wide range of values for these two parameters. We focus on a particular criterion for reliability: having an expected attack rate of at least 20% of the vertices. On one hand, we found that assortative graphs are more reliable than disassortative graphs; on the other hand, network reliability decreases for highly assortative scale-free-like graphs. In addition, network reliability decreases as the number of triangles increases in the graph. We also study features of the reliability polynomial such as its critical point and the minimum (respectively, maximum) number of edges for which subgraph is guaranteed to be reliable (respectively, unreliable). Finally, we illustrate the effect of graph size on reliability, showing that the transition in network reliability becomes sharper as the graph size increases.

The outline of this paper is as follows: First, we reintroduce network reliability in terms of reliability rules and reliability polynomials. Then we discuss the estimation of reliability coefficients. We describe an *in silico* laboratory of networks with a range of carefully controlled topological properties. We characterize these networks' reliability in terms of critical points and other features, elucidating the relationship between network reliability and common graph statistics as a function of network size. Finally, we indicate some intriguing open research problems.

II. NETWORK RELIABILITY

See Colbourn [14] for a comprehensive introduction to notions of reliability. Basically, a graph's reliability is the probability that it satisfies a predefined criterion or rule, called a *reliability rule*, even when some of its edges fail. The reliability rule can be designed to capture dynamical phenomena of interest in the graph. Reliability can be evaluated as the weighted fraction of subgraphs of the base graph that meet the criterion, where each subgraph's contribution is weighted by the probability that random edge failures produce that subgraph. Since the probability of producing a particular subgraph depends on its size, the overall graph reliability depends on the number and the size of reliable subgraphs.

More formally, consider a graph $G(V, E)$ with V vertices and E weighted edges. The edges may be directed or undirected, and there may be multiple edges between two vertices. Let the set $\mathcal{S}$ be the set of all subgraphs of G generated by including each edge (i, j) independently with probability $x_{i,j}$. There are 2^E elements of this set. Now consider a binary function $r : \mathcal{S} \rightarrow \{0, 1\}$, the *reliability rule*. If $r(s) = 1$, we say that subgraph s is *accepted* or *reliable*. We define the *reliability $R(G, r, \{x\})$* of a base graph G with respect to the

reliability rule *r* for edge weights {*x*} as the probability that a randomly chosen subgraph *s* is reliable. In other words, a network is reliable to the extent that it remains functional under random removal, i.e. failure, of edges:

$$R(G, r, x) \equiv \sum_{s \in \mathscr{S}} r(s) p_x(s). \quad (1)$$

We will explicitly include the dependence on the graph *G* and the rule *r* in notation such as *R*(*G, r, x*) only when we wish to distinguish the reliability of two different graphs or two different rules.

A. Reliability rules

There are many useful reliability rules, for example:

1. *two terminal:* a subgraph is accepted if it contains at least one directed path from a distinguished vertex *S* (the *source*) to another distinguished vertex *T* (the *terminus*);

2. *at-least-n-terminal:* a subgraph is accepted if it contains at least one connected component of size *n* or greater;

3. *all-terminal:* a subgraph is accepted if it is connected and contains every vertex of the base graph;

4. *attack rate (AR)-α:* a subgraph is accepted if the mean component size across all *vertices* is greater than or equal to *aV*. Note that this is different from the mean component size taken across all *components*. In fact, it is the sum taken across all *components* of the *squared* component size divided by *V*.

For graphs with directed edges, the notion of "connected" can be generalized appropriately. We primarily use the AR-*α* rule in this paper because of its epidemiological relevance. As its name suggests, it gives the probability that the cumulative fraction of vertices infected (sometimes called the "wet set" in a percolation setting) exceeds *α*, averaged over all possible initial conditions in which a single vertex is infected. This rule, along with many other commonly used rules, has the useful property of *coherence*, i.e. adding an edge to a reliable subgraph does not make it unreliable.

B. Reliability polynomials

The reliability defined in Equation 1 depends on the probability of obtaining each particular subgraph when edges are selected *independently* at random. It is this independence of selecting edges that makes reliability such a powerful tool. For instance, the probability of selecting any particular subgraph is simply the product of the probability of selecting each of its edges and not selecting each edge that doesn't appear. As we show below, when the edges are homogeneously weighted with, say, uniform probability of selection *x*, this reduces to a homogeneous polynomial in *x* and (1−*x*) of degree *E*. The case of a few different weights can either be treated by considering a multivariate polynomial in the weights, or by adding multiple edges between vertices; for many different weights, this becomes intractable. We restrict ourselves to the homogeneously weighted case in this paper.

To rewrite Equation 1 in polynomial form, first partition the set of subgraphs $\mathscr{S}$ into subsets $\mathscr{S}_k$ in which each subgraph has exactly $k \leq E$ edges. Each subgraph with *k* edges appears with probability $p = x^k(1-x)^{E-k}$. If $\mathscr{R}$ denotes the set of all reliable subgraphs, then $R_k \equiv |\mathscr{R} \cap \mathscr{S}_k|$ is the number of subgraphs with exactly *k* edges that are accepted by rule *r*. Then the total

contribution of sub-graphs in s_k to $R(x)$ is $R_k x^k (1-x)^{E-k}$. Summing these contributions over all k gives the *reliability polynomial* (for rule r and graph G):

$$R(x) = \sum_{k=0}^{E} R_k x^k (1-x)^{E-k}. \quad (2)$$

Figure 1 shows the network reliability $R(x)$ for several Erdős-Rényi GNM graphs that have been rewired to have positive assortativity-by-degree. General properties to note are: R_k is a non-negative integer in the range $[0, \binom{E}{k}]$; $R(0) = 0$, $R(1) = 1$ for non-trivial networks; and for a coherent rule, R is monotonic non-decreasing.

We can rewrite R_k as a product of two factors, taking

$$R_k = P_k \binom{E}{k} \quad (3)$$

as a definition of P_k. This decomposition splits R_k into what we might call an *entropic* or *combinatorial* factor $\binom{E}{k}$ and a *structural* factor P_k. The entropic factor simply makes explicit the sharp peak in the number of possible subgraphs with k edges around a small region centered at $k = Ex$, i.e. the size of the space from which equi-probable system configurations can be drawn. The factor P_k is structural in the sense that it encodes all the information about the specific graph G that is needed to determine its reliability.

The interpretation of P_k is clear – it is the fraction of possible subgraphs with k edges that are accepted by the reliability criterion. This interpretation suggests a simple estimation procedure for P_k: select a sample of subgraphs with k edges, evaluate the reliability criterion for each, and let the estimated P_k be the fraction of the sampled subgraphs that are reliable. Given a graph in memory, the computational complexity of selecting a subgraph is proportional to k (not E), the number of samples selected, and the complexity of evaluating the criterion. (The complexity of the criterion itself should not be overlooked. For most reliability rules discussed here, it can be evaluated by partitioning the selected subgraph into connected components.) Moreover, since each subgraph can be chosen and its reliability evaluated independently, the algorithm can be distributed easily onto massively parallel distributed machines.

C. Alternative expressions for *R(x)*

There are many possible complete sets of basis functions for polynomials on the unit interval in general, and hence for the reliability polynomial in particular. We find two to be particularly useful, even though they are not orthogonal bases:

1.
the set of E functions $\binom{E}{k} x^k (1-x)^{E-k}$. The coefficients in this basis are the P_k introduced above. Although, as discussed in Colbourn, evaluating these coefficients exactly is computationally hard for many graphs and many reliability rules, the P_k can be estimated to arbitrary precision by a simple, scalable algorithm for any graph. This basis is thus well-suited for computational analysis of particular graphs.

2. the set of E functions x^k. The coefficients in this basis, which we denote by N_k, can obviously be drives from the P_k by expanding the binomial $(1-x)^{E-k}$, but as we show in a companion manuscript, [22] they also have an important physical interpretation in terms of the number and overlaps of what we call *structural motifs*. This basis is well-suited for reasoning about graph structure in general.

III. IMPORTANT FEATURES OF NETWORK RELIABILITY

A. Minimum and maximum number of edges of reliable subgraphs

In Figure 1, note that $R(x)$ is negligible for $x < 0.1$ and is near unity for $x > 0.25$, i.e. for subgraphs with fewer than $k = 0.1E$ edges or more than $0.25E$ edges, respectively. This is a common feature of network reliability for many different rules, related to max flow / min cut theorems. Let $k_{min} + 1$ represent the minimum number of edges for any subgraph to be reliable, so that $P_{k_{min}} = 0$ and $P_{k_{min}+1} > 0$. Similarly, let k_{max} represent the minimum number of edges that are necessary for every subgraph to be reliable, so that $P_{k_{max}-1} < 1$ and $P_{k_{max}} = 1$. It is always true that $k_{min} < k_{max}$ since the reliability rule is coherent, i.e. adding an edge to a reliable subgraph does not make it unreliable. Thus we can write the probability P_k as follows:

$$P_k = \begin{cases} 0, & k \leq k_{min} \\ 0 < P_k < 1, & k_{min} < k < k_{max} \\ 1, & k_{max} \leq k \leq E. \end{cases} \quad (4)$$

B. Average reliability

The average reliability $\langle R(x) \rangle$ gives the expected outcome for a disease with unknown transmissibility [23, 24]. The transformation between reliability $R(x)$, viewed as a function of x, and its coefficients P_k, viewed as a function of k, has the following nice property: the average value of $R(x)$ is equal to the average value of P_k. To demonstrate this, first note that the Euler Beta function:

$$B(k+1, E-k+1) \equiv \int_0^1 x^k (1-x)^{E-k} dx \quad (5)$$

has the solution

$$B(k+1, E-k+1) = \frac{k!(E-k)!}{(E-1)!}$$
$$= \frac{1}{(E+1)\begin{pmatrix} E \\ k \end{pmatrix}}. \quad (6)$$

Then interchanging the sum and the integral and integrating by parts yields

$$\langle R(x) \rangle = \int_{x=0}^{x=1} \sum_{k=0}^{E} R_k x^k (1-x)^{E-k} dx$$
$$= \sum_{k=0}^{E} P_k \begin{pmatrix} E \\ k \end{pmatrix} \frac{k!(E-k)!}{(1+E)!} \quad (7)$$
$$= (1+E)^{-1} \sum_k^E P_k$$
$$\equiv \langle P_k \rangle$$

Thus, the average of the reliability coefficients represents the average probability of selecting a reliable subgraph.

C. Critical point

Equations 1 or 2 define a partition function for the system: the weighted sum over all configurations (subgraphs) of the reliability of the configuration, weighted by its probability. Thus the reliability can be viewed as an order parameter for the system. In the thermodynamic limit, i.e. for an infinite graph, we expect that the derivative of the reliability with respect to x will diverge at a *critical point* x_c of a phase transition, for example, the percolation phase transition for the all-terminal reliability rule. For finite graphs, we take the value of x for which the derivative of the reliability attains its maximum as defining the critical point x_c. The first derivative of reliability is the probability that the reliable subgraph starts to percolate if the probability of choosing an edge increases from x to $x + dx$ [25]. From Equations 2 and 3, we find that the first derivative of the reliability can also be written as a homogenous polynomial in x and $(1 - x)$, where the coefficients are finite differences of the P_k:

$$\frac{dR(x)}{dx} = \sum_{k=0}^{E-1} [kP_{k+1} - (k+1)P_k] \begin{pmatrix} E \\ k+1 \end{pmatrix} x^k (1-x)^{E-1-k} \quad (8)$$

IV. A LABORATORY FOR STUDYING GRAPHS

We have constructed a set of graphs of three different sizes with several different degree distributions but similar mean degree and carefully controlled ranges of assortativity-by-degree and number of triangles. (For convenience below, we use the general term "assortativity" to mean specifically assortativity-by-degree.) These graphs form an *in silico* laboratory for studying structural effects in graphs. This laboratory, along with software for evaluating network reliability, will be made accessible to the public via the Cyber-Infrastructure for Network Science (CINET) web site http://ndssl.vbi.vt.edu/cinet.

Beginning with a single randomly generated graph instance for each of two degree distributions, we apply assortativity and triangle "raising and lowering" operators $A^\pm$ and $T^\pm$ defined as follows:

- The A^+ and A^- operators increase or decrease, respectively, a graph's assortativity-by-degree.

- The T^+ and T^- operators increase or decrease, respectively, the number of triangles in a graph while leaving its assortativity-by-degree invariant.

The first graph is an Erd s-Rényi random graph (or *GNM* for short) – i.e. one generated by choosing E edges uniformly at random from among V vertices, with $V = 341$ and $E = 992$ in this case. The reason for choosing these values of V and E will become clear below. The degree distribution of this graph is as follows, illustrated in Figure 2a: (1, 9), (2, 7), (3, 33), (4, 58), (5, 54), (6, 53), (7, 57), (8, 31), (918), (10, 8), (11, 7), (12, 3), (13, 2), (14, 1)).

We accepted the first generated *GNM* that was also connected. We claim that this bias toward connectivity has not produced an atypical degree distribution. In the limit as $E \to \infty$ with fixed E/V, the expected degree distribution becomes Poisson, as is well known. Note, however, that the expected degree distribution of *connected* $G(V, E)$ is slightly different from that of *all* $G(V, E)$, since it is less likely that a graph with many vertices of low degree is connected. Consider, for example, that a connected graph cannot have any vertices with degree 0. Selecting any vertex as part of an edge is a Bernoulli process with probability $\frac{2}{V}$.

Hence across all $G(V, E)$, the probability of observing a vertex with degree d is

$p(d) = \begin{pmatrix} E \\ d \end{pmatrix} (\frac{2}{V})^d (1 - \frac{2}{V})^{E-d}$. Thus, roughly 37% of all $G(V = 341, E = 992)$ will have no vertices with degree 0. While this condition alone is not a guarantee of connectedness, it indicates that the degree distribution for our sample graph is not atypical.

Because of the recent interest in scale free graphs, we also considered a "scale-free-like" (*SF L*) graph. The degree distribution of this graph is as follows, illustrated in Figure 2b: (4, 256), (8, 64), (16, 16), (32, 4), (64, 1), with, therefore $V = 341$ vertices and $E = 992$ edges. We consider it scale-free-like because the frequency of finding a vertex with degree d, for those degrees that are present, scales as d^{-2}. We have not included vertices of degree 1 or 2 in this graph because they are less interesting dynamically than those of higher degree.

Obviously, the mean degree for the two graphs is the same. This portion of the CINET graph library includes graphs with several other topologies and degree distributions, e.g. regular grids, of nearly the same size and mean degree.

A. Choosing a range of assortativities

We use the definition of assortativity presented in Newman [26]. We repeatedly apply A^+ and A^- to the instances of *GNM* and *SFL* graphs. The operators A^+ and A^- are described in the Appendix. Applied to GNM, this creates graphs with assortativities in the range [−0.950, 0.979], nearly the full possible range; for SFL, in the range [−0.268, 0.248], only about one quarter of the possible range and apparently in agreement with an estimate by Newman.

We selected for further study only those graphs with assortativities spaced at intervals of approximately 0.05: 41 GNM graphs and 11 SFL graphs.

B. Choosing the number of triangles

For each value of assortativity, for each degree distribution, we repeatedly apply T^+ and T^- as shown in the Appendix. The possible range of the number of triangles varies significantly across assortativities and across degree distributions, and is illustrated in Figure 3. We chose to study graphs containing approximately multiples of 50 triangles. Figure 3 shows the locations of all 300+ graphs included in this study in the assortativity-triangles plane. Clearly, the total number of edges, the degree distribution, and the assortativity place complicated constraints on the total number of triangles in the graph.

Assortativity is defined as a Pearson correlation coefficient, and is thus normalized to lie in the interval [−1, 1]. The clustering coefficient for a given vertex i – and its mean value across all vertices – can similarly be normalized to lie in the interval [0, 1] by dividing the number of triangles including i by the maximum possible number of triangles that could include it, $\begin{pmatrix} d_i \\ 2 \end{pmatrix}$. However, the value of the clustering coefficient for a graph with a given number of triangles depends on how those triangles are distributed across vertices of different degrees. Since we explicitly manipulate the assortativity, this distribution changes dramatically. For example, all else being equal, it is more likely to find a triangle including two vertices of high degree, given that the two are both neighbors of a third. But all else is not equal – if the graph is assortative, it is even more likely than if it is disassortative. For these reasons, in this paper we restrict ourselves to studying the number of triangles directly rather than any normalized version such as the clustering coefficient.

C. Choosing the number of vertices

To study finite size scaling, we constructed graphs with $2V$ and $4V$ vertices. Since the model used to create the original graphs is specific to the number of vertices, there is some latitude in specifying what it means to scale the number of vertices while maintaining the "same" structure. Specifically, we maintained the edge density (the ratio between number of edges and number of vertices) and the node degree distribution.

V. NUMERICAL EVALUATION OF RELIABILITY

We evaluated the network reliability for the AR-a reliability rule on all the graphs described in the previous section. Recall that AR-a gives the probability that the cumulative fraction of vertices infected exceeds a, averaged over all possible initial conditions in which a single vertex is infected. For relevance to the spread of epidemics, we chose a equal 0.2.

A. Erdős-Rényi graphs

1. Evaluation of k_min and k_max—Figure 4 shows the minimum and maximum number of edges (k_{min} and k_{max}) needed to obtain reliable sub-graphs for GNM graphs. We observe that, in general, both k_{min} and k_{max} decrease as the assortativity increases. Because $k_{min} + 1$ is the minimum number of edges needed to obtain a connected component containing 20% of the vertices, it represents the edge density of reliable subgraphs. Consequently, the edge density of reliable subgraphs is lower for assortative graphs than for neutral and disassortative graphs. As mentioned in [26], high degree vertices in assortative networks tend to form cliques, which are also called *core groups* in the epidemiological literature. The edge density within the clique is higher than that of the network as a whole. Therefore, a reliable subgraph will first appear with fewer edges within the clique. In disassortative networks, edges tend to connect vertices with dissimilar node degrees. Thus, a reliable subgraph from a disassortative network will first appear with more edges. In other words, reliable subgraphs in assortative networks have lower edge density than reliable subgraphs in disassortative networks as shown in Figure 4. We also observe that the number of triangles has more effect on k_{max} than on k_{min}. The number of edges k_{max} increases slightly as the number of triangles increases.

2. Evaluation of critical point and the maximum derivative of reliability—Figure 5 shows the critical point x_c for disassortative, neutral and assortative GNM networks. The critical point decreases as assortativity increases; however, the critical point increases as the number of triangles increases. More edges are required to obtain reliable subgraphs from highly clustered graphs. In addition, reliable subgraphs that appear in disassortative networks are more dense than reliable subgraphs from assortative networks. We also report the maximum derivative of $R(x)$ with respect to x for GNM graphs in Figure 7. Clearly, a small change in x, i.e. $x + dx$, increases the network reliability of graphs with fewer triangles more than that of graphs with more triangles. The influence of assortativity on the maximum derivative of $R(x)$ is more noticeable in assortative graphs than in disassortative graphs.

3. Evaluation of average reliability—Figures 5 shows the average reliability for GNM networks with negative and neutral or positive assortativity, respectively. We first analyze the influence of assortativity and triangles on the reliability independently.

- Effect of triangles on reliability: Network reliability decreases as the number of triangles increases for any assortativity. For the AR-a reliability rule, creating a triangle in an unreliable subgraph by adding a new edge does not make the subgraph reliable since the newly added edge does not increase number of vertices in any connected component. However, if the newly added edge connects a vertex

that belongs in one component with a vertex in another component, the probability that the overall subgraph is reliable increases.

- Effect of assortativity on reliability: The more assortative the network is, the more reliable the network is. We know that reliable subgraphs have lower edge density for assortative graphs than for disassortative graphs i.e. $k_{min}^{assort} < k_{min}^{disassort}$ and $k_{max}^{assort} < k_{max}^{disassort}$. Thus, $\langle P_k^{assort} \rangle > \langle P_k^{disassort} \rangle$. Consequently, using Eq. 7, $\langle R(x)^{assort} \rangle$ is larger than $\langle R(x)^{disassort} \rangle$.

In contrast to [26], assortative graphs do not always have many cliques. Therefore, we analyze the combined effect of the number of triangles and assortativity on the reliability using six distinct combinations of graph properties:

1) Assortative graphs with few triangles: High degree vertices have high degree neighbors. However, these vertices are not interconnected and hence do not form cliques. Therefore, reliable subgraphs are weakly locally connected. It is hard for a reliable subgraph to percolate among only high-degree vertices because the edge density is lower for the subgraph containing high degree vertices than for the graph as a whole. Therefore, reliable subgraphs expand across not only high degree vertices but also low degree vertices. Due to the assortative property, the majority of vertices will have high degree. Thus, only a small number of edges is required for a reliable subgraph to appear.

2) Assortative graphs with many triangles: The majority of edges are used to create triangles among vertices with similar node degrees. In other words, vertices with similar degrees form weakly interconnected cliques. Reliable subgraphs appear in cliques with high degree vertices due to their large edge density. Because the cliques are highly locally connected, the number of edges in a reliable subgraph is larger for assortative graphs with large number of triangles than for assortative graphs with small number of triangles. In addition, because cliques are only weakly interconnected, it is hard for a reliable subgraph to expand outside the clique.

3) Neutral graphs with few triangles: With equal probability, a randomly selected edge connects vertices with similar degrees or vertices with different degrees. High degree vertices are weakly connected and the subgraph containing them has low edge density. Being neutral and having few triangles in the graph, a reliable subgraph expands across vertices with a wide range of degrees. Therefore, many edges are required to increase the edge density of a subgraph to become reliable. Thus, a reliable subgraph requires more edges for neutral graphs than for assortative graphs, if they both have few triangles.

4) Neutral graphs with many triangles: Many triangles exist in the graph without composing cliques. Because the graph is neutral and because triangles do not increase the reliability of graphs, the number of edges needed for a subgraph to be reliable and to expand across the graph is larger for graphs with many triangles than for graphs with few triangles.

5) Disassortative graphs with few triangles: Vertices with different node degrees are connected but do not form cliques. Thus, subgraphs with larger edge density than that of the graph as a whole exist with many edges and vertices. Consequently, in contrast to reliable subgraphs that appear with fewer edges in assortative graphs with few triangles, reliable subgraphs appear with many edges from high density subgraphs.

6) Disassortative graphs with many triangles: Vertices with different node degrees are connected together forming triangles. As discussed above, in finite graphs, triangles do not increase the reliability of graphs.

B. Scale-free-like graphs

Results obtained from SFL graphs are in agreement with results from GNM graphs except for assortativity $A > 0.1$. For assortativity increases above 0.1, k_{min} and k_{max} increase as shown in Figure 4, $\langle R(x) \rangle$ decreases and x_c increases as shown in Figure 6, and the derivative of reliability at critical point decreases as shown in Figure 7. Thus, the edge density of reliable subgraphs is larger for SFL graphs with assortativity $A > 0.1$ than for neutral SFL graphs. To understand this phenomenon, note that SFL graphs with near-maximal assortativity tend to have large number of triangles, because vertices with similar degrees create cliques. These cliques represent communities with vertices that are strongly connected, while different communities are weakly interconnected. Thus, the number of communities decreases [27] and approaches the number of distinct degree values as assortativity increases for highly assortative SFL graphs. Due to the degree distribution of SFL graphs, the majority of lowest degree vertices belong to a single community. The edge density is lower for this community than for the graph as a whole. Conversely, the communities of high degree vertices contain only a few vertices. Therefore, for reliable subgraphs to appear in communities with high edge density, the reliable subgraphs have to extend across different communities that are weakly interconnected. Consequently, a large number of edges is required to obtain reliable subgraphs from highly assortative SFL graphs. This result causes the critical point to increase with assortativity leading to a decrease in the average reliability.

C. Network reliability and scaling

We study the effect of graph size by evaluating the reliability on GNM graphs with fixed average node degrees and sizes V, $2V$ and $4V$. Three different assortativity values are used, while the number of triangles is held constant at 100. The results are summarized in Table I. Let k' be the normalized number of edges with respect to the total number of edges in the graph, e.g. $k'_{min} = \frac{k_{min}}{4E}$ for graphs with $4V$ vertices and $4E$ edges. We observe that the average reliability, k'_{min} and maximum derivative increase as the graph size increases, while k'_{max}, x_c and $k'_{max} - k'_{min}$ decrease as the graph size increases. In addition, results show that the derivative of the reliability with respect to x diverges for larger graph sizes. In other words, the transition from $R(x) = 0$ to $R(x) = 1$ becomes sharper for large graphs than for small graphs. Consequently, at the thermodynamic limit, $k'_{max} - k'_{min}$ converges to 0 i.e. k'_{min} and k'_{max} reach their convergence value k'_{therm}. Thus, $\langle R(x) \rangle$ and x_c converge to $1 - k'_{therm}$ and k'_{therm}, respectively. Therefore, network reliability moves toward a sharp transition for infinite size systems, reflecting a first order phase transition from a region of unreliable subgraphs on one side to a region with only reliable subgraphs on the other side.

VI. CONCLUSION AND FUTURE WORK

The classical concept of network reliability provides a rich theoretical basis, supported by computational estimation procedures, to study the effect of structural properties on diffusion dynamics. We have highlighted various features of reliability that provide useful characterizations of graph structure, e.g. the minimum and maximum number of edges needed to obtain reliable subgraphs, the average reliability and the critical point. We have created and made widely available a library of graphs with carefully controlled structural properties, i.e. assortativity-by-degree and triangles.

Simulation results for Erdős-Rényi and scale-free-like random graphs in this library reveal that increasing the assortativity and number of triangles has opposite effects on the probability that an epidemic outbreak will achieve an average attack rate of 20%. We found

that the required number of edges decreases as the degree assortativity increases; however, the required number of edges increases as the number of triangles increases. In addition, average network reliability increases as the degree assortativity increases but decreases as the number of triangles increases. Moreover, the critical point decreases and the derivative of reliability at critical point diverges as the degree assortativity increases, while the opposite is true for increasing number of triangles. In contrast to assortative GNM graphs, network reliability decreases as assortativity increases for assortative SFL graphs. Furthermore, we have demonstrated that the transition from unreliable subgraphs to reliable subgraphs behaves as expected.

Obviously, there are many avenues for future work in this area, such as studying the relationship between reliability and other common graph statistics. In a companion paper, we show the relationship between network reliability and statistical physics and we demonstrate the power of reliability for reasoning about graph structure using the overlaps of structural motifs. We also introduce a new measure of centrality in [28]– similar to betweenness but more closely tailored to specific dynamics – and use it to compare graphs. To extend the application of network reliability to epidemiology, we will use reliability to characterize large, realistic social networks and the effect of changes brought about by outbreak control interventions. We will also study an extension of the reliability polynomial to the case in which the rule allows continuous rather than binary outcomes.

Acknowledgments

References

1. Eubank S, Guclu H, Kumar VSA, Marathe M, Srinivasan A, Toroczkai Z, Wang N. Nature. 2004; 429(6988):180. [PubMed: 15141212]

2. Eubank S, Barrett C, Beckman R, Bisset K, Durbeck L, Kuhlman C, Lewis B, Marathe A, Marathe M, Stretz P. Journal of Biological Dynamics. 2010; 4(5):446. [PubMed: 20953340]

3. Grassberger P. Mathematical Biosciences. 1983; 157

4. Kermack W, McKendrick A. Proceedings of Royal Society. 1927; 115:700.

5. Anderson, R.; May, R. Infectious Diseases in Humans. Oxford University Press; 1992.

6. Moreno Y, Pastor-Satorras R, Vespignani A. European Physical Journal B. 2002; 26:521.

7. Youssef M, Scoglio C. J Theor Biol. 2011; 21:136. [PubMed: 21663750]

8. Barabasi AL, Albert R. Science. 1999; 286:509. [PubMed: 10521342]

9. Marathe A, Lewis B, Chen J, Eubank S. Plos One. 2011; 6(8)

10. Moore E, Shannon C. Journal of the Franklin Institute. 1956; 262:191.

11. Cvetkovic, D.; Doob, M.; Sachs, H. Spectra of Graphs, Theory and Applications. 3. Johan Ambrosius Barth Verlag; 1995.

12. Van Mieghem, P. Graph Spectra for Complex Networks. Cambridge University Press; 2012.

13. Trajanovski S, Martin-Hernandez J, Winterbach W, Van Mieghem P. Journal of Complex Networks. 2013; 1:44.

14. Colbourn, CJ. The Combinatorics of Network Reliability. Oxford University Press; 1987.

15. Moreno JGY, Pacheco A. Phys Rev E. 2003; 68

16. Nold A. Mathematical biosciences. 1980; 52:227.

17. Newman MEJ. Phys Rev Lett. 2002; 89

18. Kiss DGI, Kao R. J R Soc Interface. 2008; 791

19. Agostino GD, Scala A, Zlatic V, Caldarelli G. EPL. 2012; 97

20. Badham J, Stocker R. J Theor Pop Biol. 2010; 77:71.

21. Almendral JA, Daz-Guilera A. New Journal of Physics. 2007; 9

22. Khorramzadeh Y, Youssef M, Eubank S. 2013 in preparation.

23. Youssef M, Kooij R, Scoglio C. Journal of Computational Science. 2011; 2:286.

24. Van Mieghem P. Journal of Computer Communications. 2012; 35:1494.

25. Stauffer, D.; Aharony, A. Introduction to percolation theory. 2. CRC Press; 1991.

26. Newman MEJ. Physical Review E. 2003; 286

27. Van Mieghem P, Ge X, Schumm P, Trajanovski S, Wang H. Physical Review E. 2010; 82

28. Eubank, S.; Youssef, M.; Khorramzadeh, Y. Proceedings of the 2nd International Workshop on Complex Networks and their Applications; Tokyo, Japan. December 2–5, 2013; 2013. (in press)

29. Van Mieghem P, Wang H, Ge X, Tang S, Kuipers FA. The European Physical Journal B. 2010; 76:643.

30. Winterbach W, de Ridder D, Wang H, Reinders M, Van Mieghem P. The European Physical Journal B. 2012; 85

Appendix A: Implementing A+, A−, T+, and T−

We implement the $A^{\pm}$ and $T^{\pm}$ operators using edge swapping, or graph rewiring. We choose candidate edges uniformly at random from among all the edges in the graph and, if the candidates meet certain constraints, swap them. Figure 8 illustrates the swaps involved.

Specifically, given a graph G defined by edge set ε, which includes edges (i, j) and (k, l), the operator $A^{+}_{ij,kl}$ (resp. $A^{-}_{ij,kl}$) returns either the same graph new graph G or a G' with the edge set $\varepsilon - \{(i, j), (k, l)\} + \{(i, k), (j, l)\}$, whichever increases (resp. decreases) the assortativity. That is, $A^{+}_{ij,kl}(G) = \arg\max_{G,G'} a(g)$. We check the constraints that i, j, k, l are all distinct, that the edges (i, k) and (j, l) do not already exist, and that the graph G' remains connected. (This last constraint can be checked by ensuring that the pairs of vertices originally connected by edges are in the same component of G'.). Since this edge swap does not change the degree of the affected vertices, the direction of the change in assortativity is easily computed by comparing the values $d_i d_j + d_k d_l$ and $d_i d_k + d_j d_l$.

The triangle operators must satisfy more constraints, both because they are intended to maintain the assortativity invariant and because triangles are less local than edges. In this case, we randomly choose a vertex A, and randomly pick two of its neighbors, B and C, that are not connected by an edge. As illustrated in the right panel of Figure 8, we find a neighbor D of B that is not A or C, has the same degree as C, and has no neighbors in common with B. I.e. the edge (B, D) is not a part of any triangle. We repeat this, replacing vertex B with C to find E, with the additional constraint that E is not a neighbor of D. Then we swap edges (B, D) and (C, E) for edges (B, C) and (D, E). By construction, this does not change the assortativity, but it creates at least one more triangle than was present before, namely (A, B, C).

The T^{-} operator accomplishes the swap from the bottom of the right panel of Figure 8 to the top of the right panel. We first find vertices (A, B, C) that form a triangle. Then we find an edge (D, E), such that 1) D and E are both different from A, 2) D and E are not neighbors of A, B, or C, 3) E has the same degree as B, 4) D has the same degree as C, 5) B and D have no common neighbor and 6) C and E have no common neighbor. Then, as usual, we swap edges and test for connectivity in the new graph.
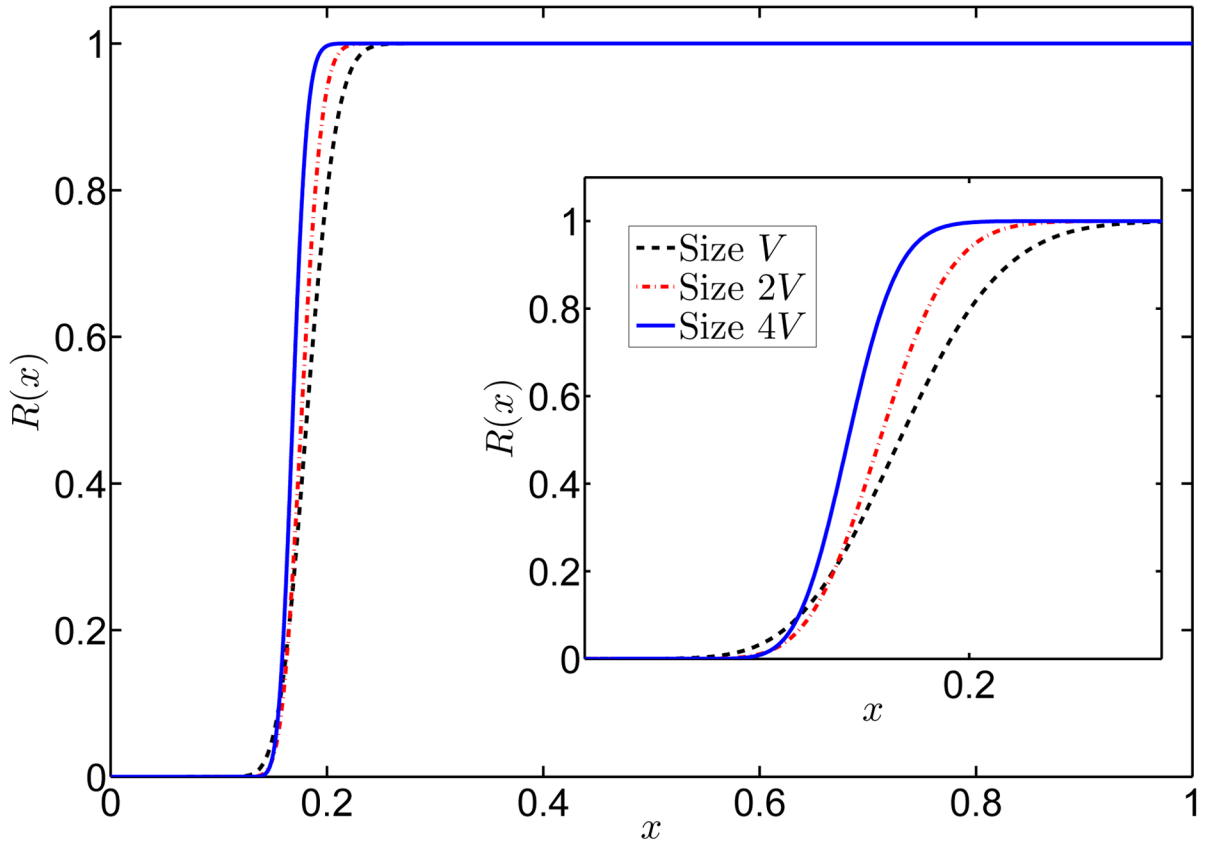
FIG. 1.
(Online color) Network reliability $R(x)$ for assortative Erd s-Rényi GNM graphs with sizes V, $2V$ and $4V$. The inset shows the transition from $R(x) = 0$ to $R(x) = 1$.
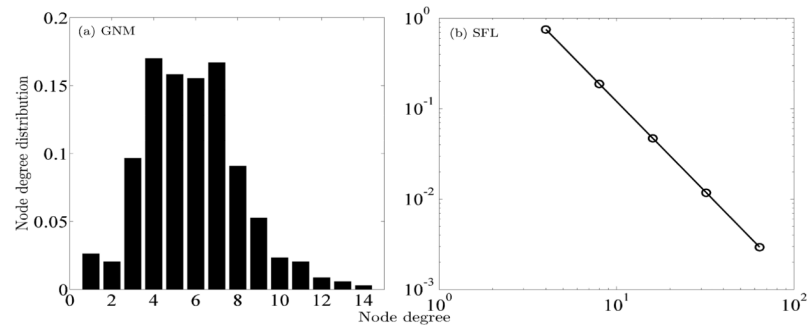
FIG. 2.
Degree distributions for the Erd s-Rényi graphs, GNM (left panel); and the scale-free like graphs, SFL (right panel). Note the logarithimic axes for the SFL graphs.
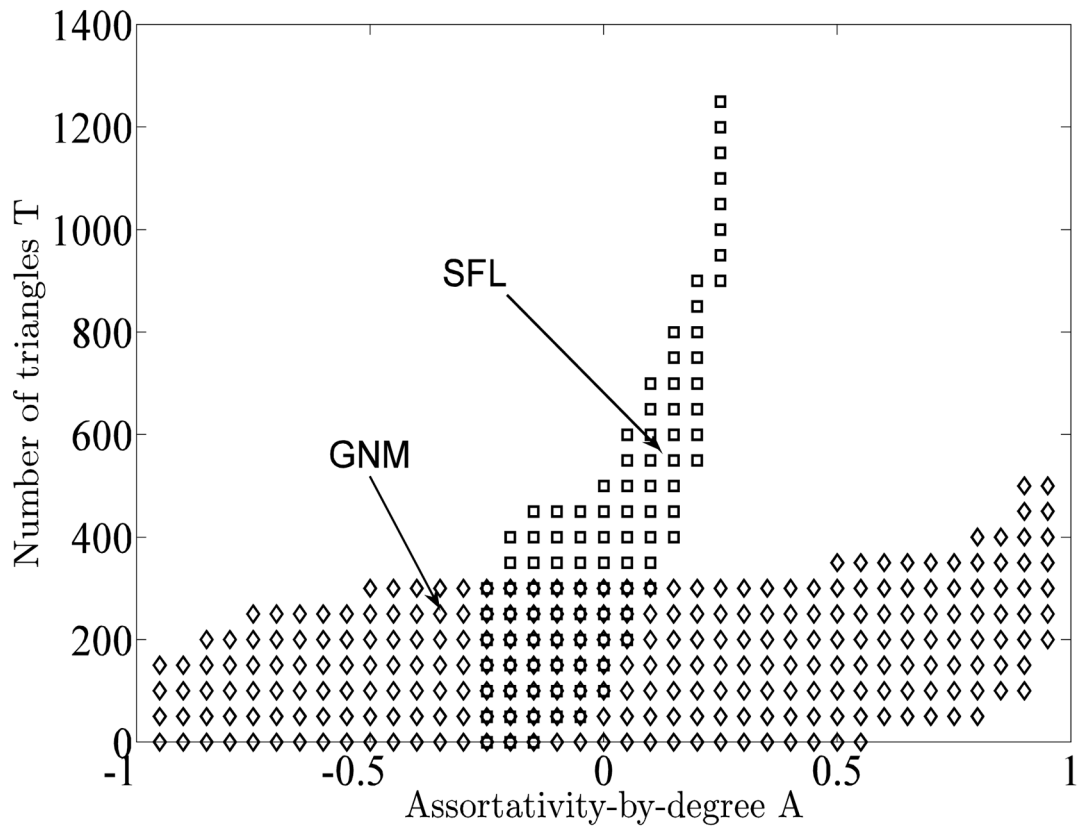
FIG. 3.
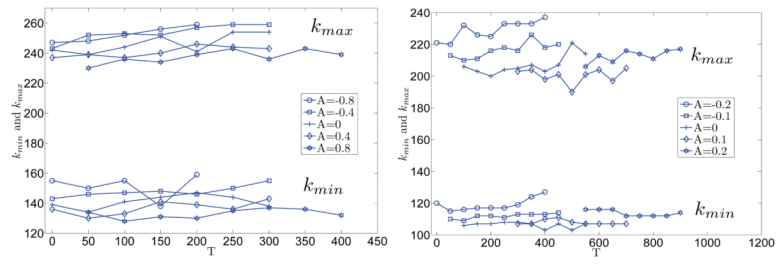Accessible ranges of assortativity and number of triangles for GNM and SFL graphs.

FIG. 4.
(Online color) k_{min} (bottom) and k_{max} (top) for disassortative, neutral and assortative GNM (left panel) and SFL (right panel) graphs under an AR-α reliability rule with $\alpha=0.2$.
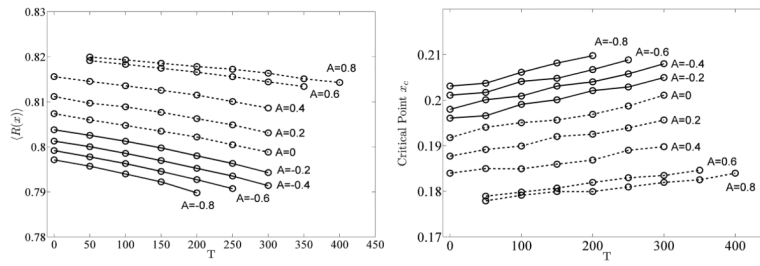
FIG. 5.
The average reliability $\langle R(x) \rangle$ (left panel) and the critical points (right panel) for disassortative, neutral and assortative GNM graphs under an AR-α reliability rule with $\alpha = 0.2$.
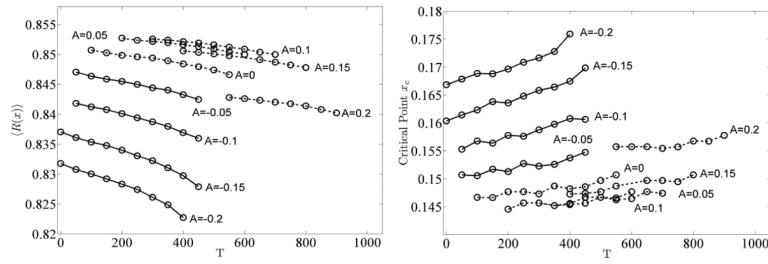
FIG. 6.
The average reliability $\langle R(x) \rangle$ (left panel) and the critical points (right panel) for disassortative, neutral and assortative SFL graphs under an AR-α reliability rule with $\alpha = 0.2$.
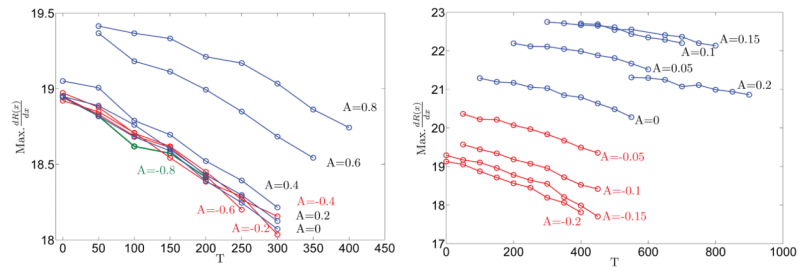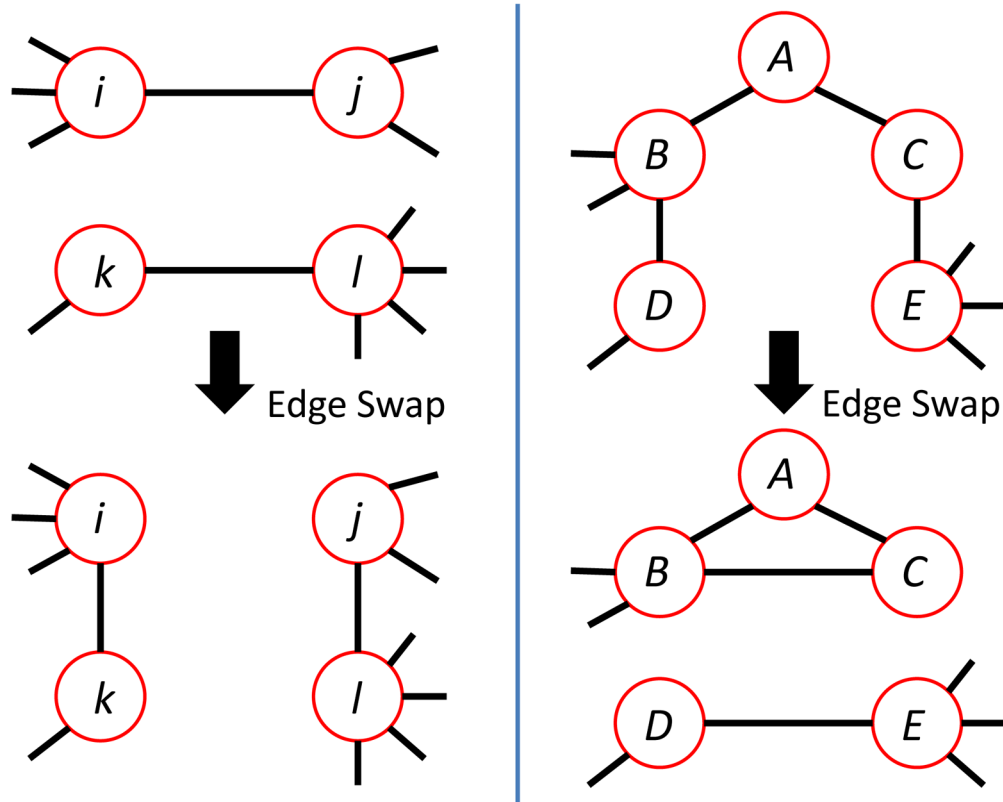
FIG. 7.

(Color online) Peak value of the derivative of $R(x)$ for disassortative, neutral and assortative GNM (left panel) and SFL (right panel) graphs under an AR-α reliability rule with $\alpha = 0.2$.

FIG. 8.

(Color online) A degree distribution-preserving edge swap. The four vertices shown are connected by the edge stubs shown here to the rest of the graph, which is not shown. (left panel) Assortativity-changing edge swap: In this example, $d_i = 4$, $d_j = 3$, $d_k = 2$, and $d_l = 5$. Hence $d_i d_j + d_k d_l = 22$ and $d_i d_k + d_j d_l = 23$ so the graph at the bottom has a higher assortativity than the one at the top [29, 30]. If the graph at the top is connected, the graph at the bottom will also be connected if there is a path from i to j. In that case, this will be an acceptable edge swap for A^+, but not for A^-. (right panel) Assortativity-preserving and triangle-changing edge swap: B and D have no neighbors in common, nor do C and E. In this example, $d_B = d_E = 4$ and $d_C = d_D = 2$. Hence $d_B d_D + d_C d_E = 40 = d_B d_C + d_D d_E$ so the graph at the top has the same assortativity as the one at the bottom. If the graph at the top is connected, the graph at the bottom will also be connected if there is a path from B to D. In that case, the swap from top configuration to bottom configuration will be an acceptable one for T^+; the swap in the opposite direction will be acceptable for T^-, since the swap cannot disconnect the graph. This swap changes the number of triangles in the graph by at least one – more, if B and C have any common neighbors besides A.

TABLE I

Evaluation of average reliability $\langle R(x) \rangle$, k'_{min}, k'_{max}, derivative of reliability at critical point and the critical point x_c for GNM graphs with different graph sizes V, $2V$ and $4V$. Each graph has assortativity $A = -0.85$, 0 and 0.85 and number of triangles $T = 100$.

$A = -0.85$	V	$2V$	$4V$
$\langle R(x) \rangle$	0.7935	0.7950	0.7972
k'_{min}	0.1522	0.1623	0.1767
k'_{max}	0.2550	0.2440	0.2349
$k'_{max} - k'_{min}$	0.1028	0.0817	0.0582
$\max \frac{dR(x)}{dx}$	18.6515	25.4679	38.7695
x_c	0.2066	0.2046	0.2021
$A = 0$	V	$2V$	$4V$
$\langle R(x) \rangle$	0.8049	0.8066	0.8067
k'_{min}	0.1391	0.1563	0.1641
k'_{max}	0.2460	0.2319	0.2228
$k'_{max} - k'_{min}$	0.1069	0.0756	0.0587
$\max \frac{dR(x)}{dx}$	18.7144	25.5718	35.2221
x_c	0.1945	0.1925	0.1928
$A = 0.85$	V	$2V$	$4V$
$\langle R(x) \rangle$	0.8173	0.8228	0.8311
k'_{min}	0.1270	0.1462	0.1447
k'_{max}	0.2429	0.2172	0.2016
$k'_{max} - k'_{min}$	0.1159	0.0710	0.0569
$\max \frac{dR(x)}{dx}$	18.9097	27.5449	38.7695
x_c	0.1804	0.1759	0.1680