Fast approximations for the reliability of networks

Jason Brown,[1,*] Theodore Kolokolnikov,[1,†] and Rob Kooij[2,3,‡]

[1] *Department of Mathematics and Statistics, Dalhousie University Halifax, Nova Scotia, B3H3J5, Canada*
[2] *Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, Delft, Netherlands*
[3] *TNO (Netherlands Organization for Applied Scientific Research), The Hague, The Netherlands*

We introduce the following edge-removal process: remove edges at random, one at a time, until the graph becomes disconnected. We show that the expected number of edges thus removed is equal to $(m+1)A$, where m is the number of edges in the graph, and A is the average of the all-terminal reliability polynomial. Based on this process, we propose a Monte-Carlo algorithm to quickly estimate the graph reliability (whose exact computation is generally an NP-hard problem). Moreover, we show that the distribution of the edge-removal process can be used to quickly estimate the reliability polynomial. We then propose increasingly accurate asymptotics for graph reliability based solely on degree distributions of the graph. These asymptotics are tested against several real-world networks and are shown to be accurate for sufficiently dense graphs, whose average degree is above 2. They start to fail for "subway-like" networks that contain many paths of vertices of degree two. Different asymptotics are derived for such networks.

1. INTRODUCTION

For a network (which we assume here is represented by a finite undirected graph, possibly with multiple edges) there are many models of robustness to component failure. The simplest measures include the *minimum degree* (the minimum number of edges whose removal disconnects a vertex from the rest of the graph), the *edge-connectivity* (the minimum number of edges whose removal disconnects the graph), and the *vertex connectivity* (the minimum number of vertices whose removal disconnects the graph or reduces it down to a single vertex). Another indirect measure is the algebraic connectivity of the graph [1] which measures how fast information propagates through the graph [2–5].

However, such static measures are rather coarse and do not take into account that the components of a graph may, at different junctures, fail to be operational. So more nuanced probabilistic models have been described, where the components (vertices and/or edges) are subject to random failures. In the most common of these models, it is the edges that fail independently with the same probability q, while the vertices are always operational, and we ask for the probability that the network is operational, where operational can mean, for example:

- all vertices can communicate (the *all-terminal reliability* $R_G(q)$, or simply $R(q)$), or
- two specific vertices s and t (called the *terminals*) can communicate (*two-terminal reliability* $R_{G,s,t}(q)$), or
- a specific subset of vertices can communicate with one another (*K-terminal reliability*, $R_{G,K}(q)$).

All of these polynomials (in variable q) are intractable, that is, NP-hard (see [6]) Of course, there are other variants in the literature, some for node failures rather than edge failures, some for directed graphs, and some that allow for failure dependencies. A general reference is [6].

In our study of all-terminal reliability, we have been led recently to propose a new variant of network robustness based on an algorithmic procedure. Consider the following **edge deletion process:** delete edges from the graph at random, one at a time. Stop when the graph becomes disconnected. Let s_k denote the probability that under the edge deletion process the process stops exactly after k steps. What is the probability distribution of the s_k's? Define

$$\psi = \sum_{k=0}^{\infty} k s_k \tag{1.1}$$

to be the mean of s_k. What can we say about ψ, the average number of edge deletions to disconnection? How does this measure relate (if at all) to the existing reliability measures? We are interested in the answers to these questions both for abstract graphs, as well as for networks that appear in practice.

We remark that our interest grew out of a common Monte Carlo simulation of all-terminal reliability. Namely, start with a (connected) graph G of order n and size m (that is, with n vertices and m edges), a non-negative integer $k \in \{0, 1, \ldots, m-n+1\}$ and for a large positive integer N, chose a random spanning subgraph of G with $m-k$ edges and determine the proportion r_k of such spanning subgraphs that are connected. This value r_k is an approximation to $R_G(k/m)$. A more efficient algorithm is to chose the k edges sequentially for deletion and stop when the graph becomes disconnected (as further edge deletion cannot subsequently make the graph connected!).

Let us begin with a couple of examples.

* Jason.Brown@dal.ca
† tkolokol@gmail.com, corresponding author
‡ R.E.Kooij@tudelft.nl

Example 1 First of all, consider the cycle graph C_n of order n. Clearly the removal of any single edge does not disconnect the graph, but the removal of any two edges does disconnect the graph. Thus $s_2 = 1$ while $s_k = 0$ for all $k \neq 2$, and hence

$$\psi(C_n) = \sum k s_k = 2. \tag{1.2}$$

Example 2 As another example, consider the graph H of order 3, with vertices x_1, y and x_2, where each x_i is joined to y by l edges. It is not hard to see that

$$s_k = \begin{cases} 0 & k < l \\ \frac{2l\binom{k}{l}}{k\binom{2l}{l}} & l \leq k \leq 2l-1 \\ 0 & k \geq 2l \end{cases}, \tag{1.3}$$

and hence

$$\psi(H) = \sum_{k=l}^{2l-1} \frac{2l\binom{k}{l}}{\binom{2l}{l}} = \frac{2l^2}{l+1}. \tag{1.4}$$

While the definition of the s_k's and ψ do not, on the surface, seem to relate to the all-terminal reliability (the later is a *polynomial* in variable q), we shall see in the next section that there is indeed a deep connection between the two notions.

2. THE CONNECTIONS TO ALL-TERMINAL RELIABILITY

The aim of this section is to quantify the connections between the edge deletion process, in terms of the stochastic variable s_k and the all-terminal reliability. We will establish connections for three metrics related to the all-terminal reliability: the average reliability (AR) of the graph, the full all-terminal reliability polynomial of the graph, and the 99-percentile value of the all-terminal reliability (the choice of the 99-percentile is arbitrary, but is indicative of how reliable a network is expected to be in practice).

We begin by enumerating a number of useful forms of the all-terminal reliability polynomial. For a graph G of order n and size m (here and elsewhere, we shall always assume that any graph under question is connected), its all-terminal reliability has the following forms (each is an expansion of the polynomial under different bases for the underlying vector space of polynomials, see [6]):

$$\begin{aligned} R_G(q) &= \sum_{i=0}^{m} N_i q^{m-i}(1-q)^i \\ &= \sum_{i=0}^{m} F_i q^i (1-q)^{m-i} \\ &= 1 - \sum_{i=0}^{m} C_i q^i (1-q)^{m-i} \end{aligned} \tag{2.5}$$

We refer to these as the N-, F-, and *C-forms* of the reliability polynomial. The interpretation of the coefficients is as follows:

- N_i counts the number of spanning connected subgraphs with i edges.
- F_i counts the number of spanning connected subgraphs with $m - i$ edges.
- C_i counts the number of *cut sets* with i edges, that is, the number of subsets of i edges whose removal leaves the graph disconnected.

Of these three forms, the first two have attracted the most attention in the literature. However, we will see that the connection between the s_k's and all-terminal reliability is most easily drawn with the C-form.

2.1. Connection with the average reliability of a graph

Theorem 2.1. *Let G be a graph having m edges. Remove edges from G at random and one at a time, until the graph becomes disconnected. Let ψ be the average number of edges thus removed. Then*

$$\psi = (m+1)\int_0^1 R(q)dq. \tag{2.6}$$

Theorem 2.1 states that ψ is $m+1$ times the average of the all-terminal reliability polynomial over the interval $[0, 1]$; the latter is known as the *average reliability* (AR) of the graph, and has been studied in [7]. The average reliability of a graph was proposed as a single numerical measure of the robustness of a graph, and allows one to search for the existence of an optimal graph with respect to reliability (even when a graph optimal for all $q \in [0, 1]$ need not exist [8]) among all graphs with a given fixed number of of vertices n and a fixed number of edges m. We remark that the average reliability (and hence ψ) is also known to be intractable [7].

Proof of Theorem 2.1 Define $c_k = \sum_{i=0}^{k} s_i$. Then c_k is the probability that removing k random edges disconnects G. Furthermore, $s_k = c_k - c_{k-1}$ and $c_k = \frac{C_k}{\binom{m}{k}}$, where, as in the C-form, C_k is the number of edge subsets of size k whose deletion disconnects the graph. To see the latter, note that if a subset S of k edges has the property that $G - S$ is disconnected, then this is true under *any* edge ordering of S, so that

$$c_k = \frac{C_k \cdot k!}{m(m-1)\cdots(m-k+1)} = \frac{C_k}{\binom{m}{k}}.$$

We now write the mean ψ as:

$$\begin{aligned}\psi &= \sum_{0}^{m} k s_k \\ &= c_1 - c_0 + 2\left(c_2 - c_1\right) + \cdots + m\left(c_m - c_{m-1}\right) \\ &= -c_0 - c_1 - c_2 - c_3 - \cdots - c_{m-1} + m c_m \\ &= -\sum_{k=0}^{m-1} \frac{C_k}{\binom{m}{k}} + m \\ &= m + 1 - \sum_{k=0}^{m} \frac{C_k}{\binom{m}{k}}\end{aligned}$$

On the other hand, from the C-form of the reliability,

$$R(q) = 1 - \sum_{k=0}^{m} C_k q^k \left(1-q\right)^{m-k},$$

so that, from the above formula for ψ,

$$\begin{aligned}\int_0^1 R(q)dq &= 1 - \sum_{k=0}^{m} \int_0^1 C_k q^k \left(1-q\right)^{m-k} dq \\ &= 1 - \frac{1}{m+1} \sum_{k=0}^{m} \frac{C_k}{\binom{m}{k}} \\ &= \frac{\psi}{m+1},\end{aligned}$$

where we used the fact (see, for example, [9]) that for $i \in \{0, 1, \ldots, m\}$, the integral of the *Bernstein basis polynomial* [10] is $\int_0^1 \binom{m}{i} q^i (1-q)^{m-i} dq = 1/(m+1)$.□

To illustrate the theorem, we again consider the two examples from Section 1.

Example 1. For a cycle C_m we have $R(q) = (1-q)^m + mq(1-q)^{m-1}$, and direct integration yields $\int_0^1 R(q)dq = \frac{2}{m+1}$. On the other hand, removing any two edges (and exactly two edges) results in disconnection, so that $\psi = 2$, consistent with (2.6).

Example 2. For the graph H introduced in Section 1, the reliability polynomial satisfies $R(q) = (1-q^l)^2$ while $m = 2l$. We have

$$(m+1) \int_0^1 R(q)dq = (2l+1) \int_0^1 \left(1-q^l\right)^2 dq = \frac{2l^2}{l+1},$$

consistent with the direct computations that $\psi = \frac{2l^2}{l+1}$ in (1.4).

Theorem 2.1 gives a **practical** way to approximate ψ for large graphs using Monte Carlo simulation without computing the reliability polynomial explicitly, which in general is intractable.

2.2. Connection with the all-terminal reliability polynomial

We will now show an even deeper connection between the edge-deletion process and the all-terminal reliability polynomial. From the C-form of reliability, we can write

$$1 - R(q) = \sum_{k=0}^{m} c_k \binom{m}{k} q^k \left(1-q\right)^{m-k}. \qquad (2.7)$$

Note that $\binom{m}{k} q^k \left(1-q\right)^{m-k}$ $(k = 0, 1, \ldots, m)$ is a binomial distribution. In the limit of large m, it converges to a normal distribution of mean $\mu = qm$ and standard deviation $\sigma = \sqrt{mq(1-q)}$, so that $\binom{m}{k} q^k \left(1-q\right)^{m-k} \sim \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{(\mu-k)^2}{2\sigma^2}\right)$. Recall that for any continuous function f_k, Laplace's method gives the asymptotics

$$\int_0^{\infty} f_k \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{\left(\mu-k\right)^2}{2\sigma^2}\right) dk \sim f_\mu$$

assuming $\sigma \ll \mu$ (see for example [11]). Approximating the sum in (2.7) by an integral then yields

$$\begin{aligned}&\sum_{k=0}^{m} c_k \binom{m}{k} q^k \left(1-q\right)^{m-k} \sim \\ &\int_0^{\infty} c_k \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{\left(\mu-k\right)^2}{2\sigma^2}\right) dk \sim c_\mu\end{aligned}$$

so that $1 - R(q) \sim c_\mu$, $q = \mu/m$. Replacing μ by k so that $q = k/m$, we obtain

$$c_k \sim 1 - R(k/m). \qquad (2.8)$$

This means that the complement of the reliability polynomial (i.e. $1-R(q)$), is approximated by the normalized histogram of the c_k's. To be more precise, assume we run the edge deletion process M times and denote the number of times the process stops after exactly k steps by S_k. Then we have $s_k = \lim_{M\to\infty} \frac{S_k}{M}$ and $c_k = \lim_{M\to\infty} \frac{1}{M} \sum_{i=0}^{k} S_k$.

As a result, it follows from Eq.(2.8) that for large m, s_k approximates minus the reliability polynomial evaluated at k/m, i.e.

$$s_k \sim -R'\left(\frac{k}{m}\right). \qquad (2.9)$$

We now show that for sufficiently small values of k, c_k can underestimate the values of $1-R(k/m)$ while for sufficiently large values of k overestimation can occur. Let $\lambda(G)$ denote the edge connectivity of a graph G. Then, for graphs with $\lambda > 1$ then for any $1 \leq k < \lambda$ it holds that $c_k = 0$, while $1 - R(k/m) > 0$ although, in general, the value is very small. On the other hand, for $k > m - n + 1$, the edge deletion process always disconnects the network hence for those k it holds that $c_k = 1$, while $1-R(k/m)<1$ although, in general, the difference is very small.

Finally, we can also use Eq.(2.8) to approximate the 99%-percentile of $R(q)$. First determine the largest value of k which satisfies $c_k \leq 0.01$ and denote this value by $k_{99\%}$. Then $q_{99\%}$ can be approximated by

$$q_{99\%} = \frac{k_{99\%}}{m}. \tag{2.10}$$

2.3. Validation on a synthetic and a real-world network

First we validate the results obtained in this section on a so-called crown graph. This graph is constructed by taking a complete bipartite graph $K_{N,2}$ and joining the two nodes in the independent set of size 2, see Fig.1. This graph has $N+2$ nodes and $2N+1$ links, and will be denoted as Cr_{N+2}.

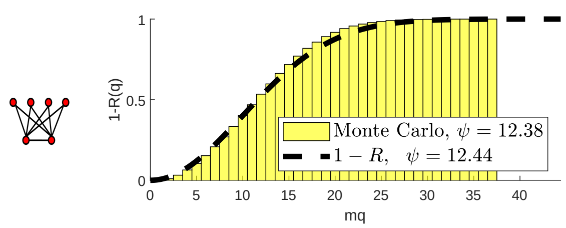

FIG. 1. Left: The crown graph Cr_6. Right: Reliability polynomial: exact (dashed line) compared with Monte Carlo simulation (yellow histogram), for Cr_{50} using $M = 1000$ simulations

For the crown graph Cr_{N+2}, according to [12], we have

$$R(q) = (1-q)^N((1+q)^N - 2^N q^{N+1}) \tag{2.11}$$

The explicit expression (2.11) allows us to obtain a closed-form expression for ψ through (2.6):

$$\psi = \frac{\Gamma(N+2)}{\Gamma(N+3/2)}\left(1-2^{-N-1}\right)\sqrt{\pi}, \tag{2.12}$$

where Γ is the Gamma function.

We now use (2.11) and (2.12) to validate Theorem 2.1 by means of Monte Carlo simulation, as shown in Figure 1. Taking $N=48$, formula (2.12) yields $\psi_{exact} = 12.4389$. On the other hand, using $M = 1000$ Monte Carlo simulations we obtained $\psi_{MC} = 12.39$, in excellent agreement with the exact result. Finally, we use (2.11) to determine $q_{99\%}$ numerically by solving $R(q) = 0.99$. With $N = 48$, (2.11) yields $q_{99\%} = 0.014469$, so that removing $k_{99\%} = 1.40$ edges on average results in a 99% probability of still being connected. According to Monte carlo simulation, there is about a 1.3% probability of disconnection after removing 2 edges (whereas removing one edge never disconnects a crown graph). Using linear interpolation, this yields $k_{99\%,MC} \approx 1.78$. Note that the 99% threshold is in the very tail of the distribution, where the Monte Carlo approximation to $R(q)$ is degraded. Using a 90% threshold instead, the exact formula (2.11) yields $k_{90\%,exact} = 4.589$, compared to $k_{90\%,MC} = 5.0364$, a better agreement (9.7% relative error for $k_{90\%}$ instead of 27% for $k_{99\%}$).

M	1000	2000	4000	8000	16000	32000
E_M	0.1706	0.1225	0.0867	0.0620	0.0439	0.0310

TABLE I. RMSE as a functin of the number of simulations M

Due to nature of MC simulation, the accuracy increases with M but relatively slow. To measure the accuracy, let K_j be the number of edges removed before the disconnection during simulation j. Define the Root Mean Square Error (RMSE) as $E_M = \sqrt{\sum_{j=1}^{M}(\psi_{exact} - K_j)^2/M}$. The following table lists E_M as a function of M. As expected, square root scaling $E_M = O(N^{-1/2})$, typical of Monte Carlo simulations, is observed.

Next we validate our results on a real-life network, namely the DFN communication network from the Internet Topology Zoo [13]. This network has $n = 58$ nodes and $m = 87$ edges. This network is small enough to determine its reliability polynomial exactly. We used the `ReliabilityPolynomial` command from Maple's GraphTheory package to do so. Figure 2 shows the comparison between the normalized histogram for the c_k's (computed using 1000 Monte Carlo simulations), and $1-R(q)$. Visually we see a good fit. In addition, $\psi_{MC} = 6.83$ while $\psi_{exact} = 6.896$, so the relative error is only about 1%. Similarly we obtain $k_{99\%,MC} = 0.1252$ while $k_{99\%,exact} = 0.1190$, a relative error of about 5%.

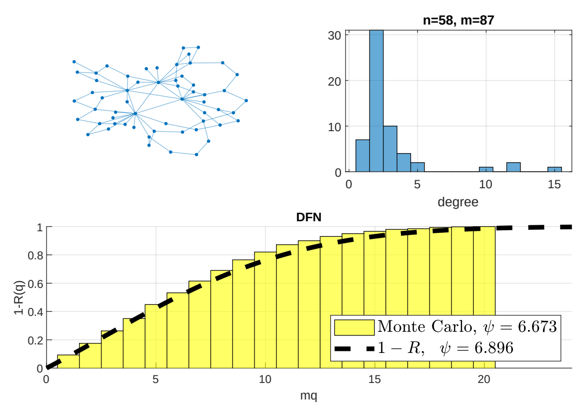

FIG. 2. Top: The DFN network (consisting of 58 nodes and 87 edges) and its degree histogram. Reliability polynomial (dashed line is $1-R(q)$ computed exactly) compared with the CDF for the s_k's obtained through Monte Carlo simulations (yellow histogram).

3. ANALYTIC APPROXIMATIONS FOR RELIABILITY POLYNOMIALS

So far we have compared the results obtained from the edge deletion process for graphs for which an explicit expression for its reliability polynomial is available. However, in general such explicit expressions are intractable. In this section we will introduce two closed-form approximations for $R(q)$, which allow us to compare the results of the edge deletion simulation with the performance metrics ψ and the 99%-percentile.

3.1. First and second order approximations

The probability of a single vertex being isolated is q^d, where d denotes the degree of the vertex. When q is small and for a sufficiently dense graph, the probability of having no isolated vertices asymptotically can be approximated by

$$R_1(q) := \prod_{j=1}^{n} \left(1 - q^{d_j}\right). \tag{3.13}$$

This heuristic ignores any inter-dependence of the vertices, but works well when the graph is sufficiently "dense". Note that (3.13) only depends on the degree sequence of the graph and not on its finer structure. This is a property exhibited by the so-called "random configuration model", see e.g. [14–16] and references therein. For sufficiently dense graphs, we use the heuristic that the probability of being disconnected approaches the probability of having an isolated vertex. While the rigorous justification of this heuristic remains an open problem, it is similar to the classical results for the Erdős-Rényi random graph model [17]. We will see below that R_1 provides a good approximation to R for many realistic networks as well as for random regular graphs of degree 3 or more.

A more accurate formula for $R(q)$ also incorporates the chance of having no isolated two-vertex subgraphs. Given an edge, the chances that it is disconnected from the graph is $q^{a-2}(1-q)$, where a is the sum of the degrees of the vertices adjoining this edge. This leads to the following, more accurate asymptotics: $R \sim R_2$, where

$$R_2(q) := \prod_{j=1}^{n} \left(1 - q^{d_j}\right) \prod_{j=1}^{m} \left(1 - q^{a_j-2}(1-q)\right), \tag{3.14}$$

where a_j is the sum of degrees of the two vertices adjoining an edge j. Higher-order asymptotics can be written by considering isolated graphs of 3 or more vertices. However we will show in the next section that for most practical examples we considered, R_2 (and even R_1) provides high accuracy.

The asymptotic formulae (3.13, 3.14) can also be used to estimate ψ in (2.6), the expected number of edges that need to be deleted in order to disconnect the network. We have the following, increasingly accurate approximations for ψ:

$$\psi_1 := (m+1) \int_0^1 R_1(q)dq; \tag{3.15}$$

$$\psi_2 := (m+1) \int_0^1 R_2(q)dq; \tag{3.16}$$

4. REGULAR GRAPHS

In this subsection we consider the special case of d-regular graphs, in particular for large n. For this case, the approximation Eq.(3.13) becomes

$$R_1(q) = (1-q^d)^n \tag{4.17}$$

whereas (3.14) yields

$$R_2(q) = (1-q^d)^n (1 - q^{2d-2}(1-q))^{nd/2}. \tag{4.18}$$

From (4.17) we estimate:

$$q_{99\%} \sim (1 - (0.99)^{1/n})^{1/d} \tag{4.19}$$

To compute $\int_0^1 R_1(q)dq$,we estimate $(1-q^d)^n \sim \exp\left(-nq^d\right)$, $n \gg 1$, so that

$$\int_0^1 R_1(q)dq \sim \frac{1}{d}\Gamma(1/d)n^{-1/d}, \quad n \gg 1. \tag{4.20}$$

Substituting $m = \frac{nd}{2}$ for d-regular graphs into (2.6) yields $\psi \sim \psi_1$ for $n \gg 1$ where:

$$\psi_1 = \frac{n^{1-1/d}}{2}\Gamma\left(\frac{1}{d}\right), \quad n \gg 1 \tag{4.21}$$

For example for a 3-regular graph, this estimate yields $\psi_1 = 1.3394n^{2/3}$.

Next we compute the asymptotics to two orders. Applying Laplace's method, we estimate:

$$\begin{aligned}
\int_0^1 R_2(q)dq &\sim \int_0^1 e^{-nq^d} e^{-n\frac{d}{2}q^{2d-2}(1-q)} dq \\
&= \frac{1}{d}n^{-1/d} \int_0^n e^{-x} e^{-\frac{d}{2}x^{2-2/d}n^{-1+2/d}\left(1-x^{1/d}n^{-1/d}\right)} x^{1/d-1}dx \\
&\sim \frac{1}{d}n^{-1/d} \int_0^\infty e^{-x}\left(1 - \frac{d}{2}x^{2-2/d}n^{-1+2/d}\right) x^{1/d-1}dx \\
&\sim \frac{1}{d}n^{-1/d}\left(\Gamma(1/d) - \frac{d}{2}n^{-1+2/d}\Gamma(2-1/d)\right)
\end{aligned}$$

so to two orders in n, we obtain $\psi \sim \psi_2$ where

$$\psi_2 = \frac{1}{2}n^{1-1/d}\Gamma(1/d) - \frac{d}{4}n^{1/d}\Gamma(2-1/d). \tag{4.22}$$

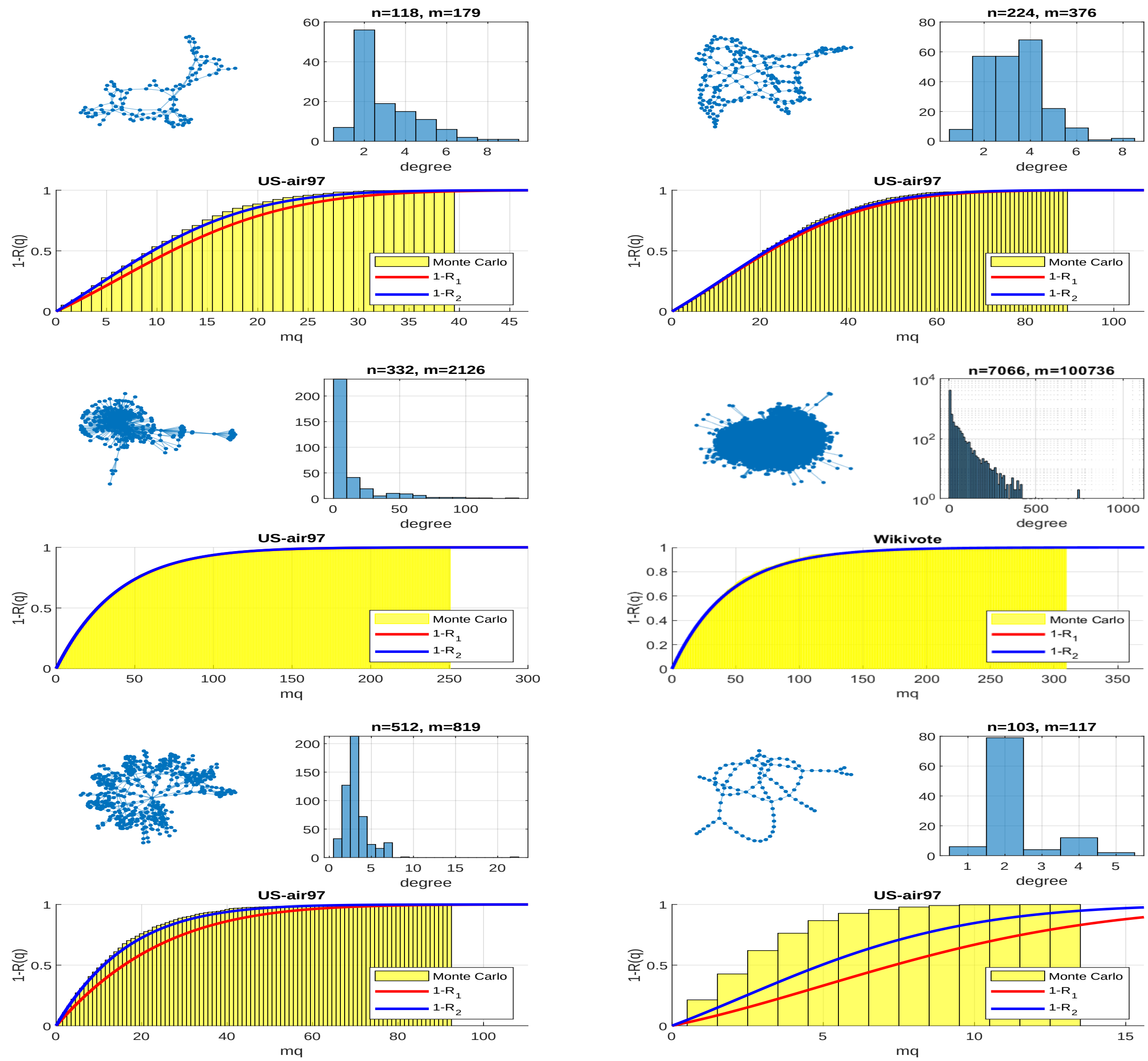

FIG. 3. Comparison of Monte Carlo (using 1000 simulations) and asymptotic reliability computations for some real-world networks. Figures show the full distribution s_k obtained using Monte Carlo simulations, and visual comparison to asymptotics (3.13), (3.14). The table shows numerical values for the averages as well as 99% threshold values.

5. VALIDATION OF THE APPROXIMATIONS FOR THE RELIABILITY POLYNOMIALS

In this section we compare the outcomes of the edge deletion process with the approximations Eqs.(3.13)-(3.14) for the reliability polynomial.

5.1. Validation on real-world networks

In this subsection we consider a number of real-world networks, taken from the Internet Topology Zoo [13] and the Network Repository [18]. We apply Theorem 2.1 and asymptotics (3.15) and (3.16) to several real-world networks; the results are presented in Figure 3 and Table II. The column ψ_{MC} is computed using the Monte Carlo method, averaged over 1000 simulations.

Columns ψ_1 and ψ_2 are asymptotic estimates as given by (3.16). In addition, we include the reliability measure $q_{99\%}$. Despite the diversity of networks presented in Figure 3, the asymptotics agree very well with Monte Carlo simulations for all networks shown except “Singapore”, which represents the subway network in Singapore. The agreement breaks down for such network because it consists of many “strings”: paths where each vertex has degree at most two. We will discuss how to improve asymptotics of such “subway” networks in Section 6.

It is interesting to note that the approximations for

Network	n	avg deg	deg std	ψ_{MC}	ψ_1	ψ_2	$q_{99\%}$ (MC)	99% (R_1)	99% (R_2)
Hagy	37	10.81	6.94	91.171	93.096	92.856	0.0928	0.0913	0.0913
ISNA	60	3.13	2.80	3.62	4.18812	3.98228	0.000404	0.000480	0.000459
Illinois	70	7.82	3.43	92.578	94.5861	93.8267	0.0636	0.0553	0.0553
Singapore	103	2.27	0.818	3.0090	8.1923	5.7095	0.00033	0.000901	0.0016
IEEE118	118	3.03	1.574	11.045	13.02	11.056	0.000868	0.00145	0.00110
Berlin	224	3.35	1.27	24.017	25.1483	24.1642	0.00115	0.00124	0.00124
US-air97	332	12.807	20.134	35.335	37.04	36.882	0.00014692	0.00018276	0.00018276
s838	512	3.12	1.63	15.39	20.5658	15.0196	0.000221	0.000306	0.000208
wikivote	7066	28.513	57.733	43.61	44.36	44.17	4.72E-06	4.45E-06	4.44E-06
ia-email-EU-def	32430	3.35	18.18	2.178	2.1407	2.13919	4.103E-07	4.03E-07	4.02E-07

TABLE II. Comparison of Monte Carlo and asymptotic estimates for selected real-world networks (see also Figure 3

$q_{99\%}$ using either R_1 or R_2 give identical or nearly identical results (to all digits shown) in about half of the cases.

5.2. Validation on regular graphs

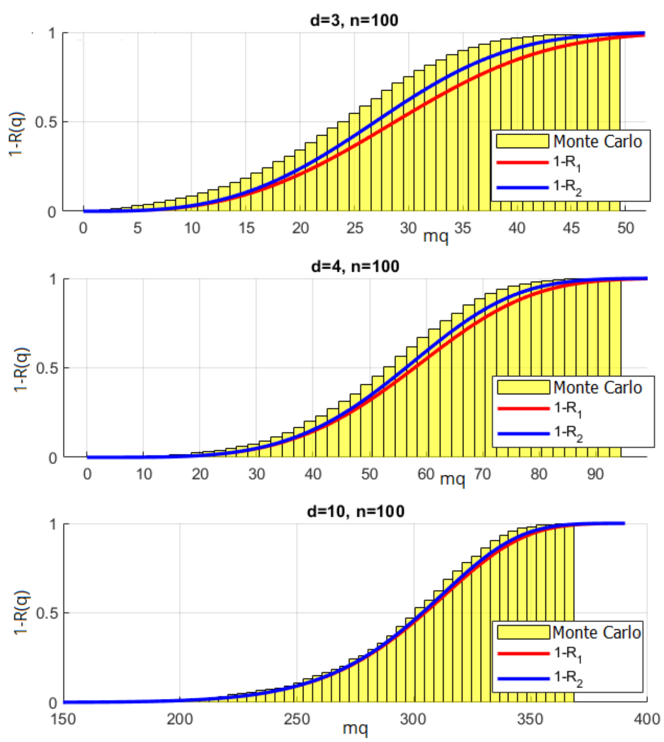

FIG. 4. Comparison of Monte Carlo and asymptotic reliability computations for random d-regular graphs, with d as indicated and $n = 100$.

The following table shows the comparison between asymptotic and Monte Carlo simulations for random d-regular graphs (constructed using the configuration model [16]) with $n = 100$ and d as indicated:

d	ψ	ψ_1	ψ_2	$\%err_1$	$\%err_2$	$q_{99\%,MC}$	$q_{99\%,as}$
3	23.5	28.856	25.71	18%	8%	0.015	0.046
4	53.3	57.32	54.41	7%	2%	0.074	0.100
10	296	300.13	296.32	1.7%	$< 1\%$	0.3870	0.3983

TABLE III. Asymptotics vs. Monte Carlo results for regular graphs

Here, ψ and $q_{99,MC}$ are computed using the Monte Carlo method using 10000 realizations and is believed to be accurate to within less than 1% (as validated by averaging over several random subsets of simulations of size 5000). The relative errors $err_1 = 1 - \psi/\psi_1$ and $err_2 = 1 - \psi/\psi_2$, are also shown. Note that for $d = 3$, ψ_1 captures about 82% of cases, whereas getting either an isolated vertex or an isolated 2-graph captures 92% of all cases (the remaining cases correspond to getting an isolated 3-graph and higher). As the graph density is increased, disconnection due to isolated vertices captures more and more cases; e.g. it captures >98% of cases when $d = 10$.

The last two columns in Table III show the 99-percentile, computed using Monte Carlo simulation ($q_{99\%,MC}$) and the asymptotic approximation in formula (4.19) ($q_{99\%,as}$). We observe an increasing accuracy of the asymptotics with increased d. The agreement for $q_{99\%}$ is rather poor when $d = 3$ and $n = 100$ because in that case $mq_{99\%,as} \approx 2$, which is rather small (at the very tail of the distribution). The agreement is much better when $d = 10$.

6. SUBWAY-LIKE NETWORKS

As seen in Figure 3, the asymptotics (3.13) break down for the subway network of Singapore. It is characterized by many paths made of consecutive vertices of degree-two, interspersed with a few transfer stations that have higher degree. We will refer to these types of networks as the "subway-type" network [19]. For this type of network, the breakdown of connectivity is most likely to happen because one of the "paths" gets disconnected, rather than getting an isolated vertex. This is the reason why R_1 or R_2 fail to estimate reliability in this case.

To obtain a better estimate of ψ for "subway-type" networks, we first process the graph by removing the "one-shell" from the graph as follows. Remove all vertices of degree one and its associated edge from the graph. Repeat, until there are no more vertices of degree one. If the reliability polynomial of the graph with the one-shell removed is $R(q)$, then the reliability polynomial of the original graph is $(1-q)^K R(q)$, where K is the number of edges/vertices in the one-shell that were removed.

In what follows, we shall assume without loss of generality that the one-shell has been removed to simplify the computation. We first compute the probability that in a single path consisting of l edges, some vertex cannot communicate with either of the end terminals. This probability corresponds to failure of at least two edges along this path. The probability of such a failure is thus given by $1 - p^l - lp^{l-1}q$ where $p = 1 - q$. This yields the following estimate for the reliability polynomial $R \sim R_{shell}$ for subway-type networks:

$$R_{shell}(q) = \prod_j ((1-q)^{l_j} + (1-q)^{l_j - 1} q l_j) \tag{6.23}$$

where the product is taken over all the paths (consisting of vertices of degree 2 inside the graph, and with l_j being the number of edges in such a path.

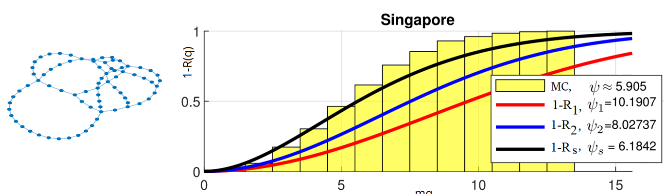

FIG. 5. Left: Subway-type network "Singapore", with one-shell removed. Right: Comparision of Monte Carlo simulation and approximations to the reliability polynomials (3.13), (3.14) and (6.23), as well as their averages.

The corresponding estimate for ψ becomes

$$\psi_s = (m+1) \int_0^1 R_{shell}(q) dq. \tag{6.24}$$

The function $R_{shell}(q)$ and the corresponding ψ_s is shown in Figure 5. As can be seen, ψ_s does a much better job for subway-type graphs than either ψ_1 or ψ_2.

Finally, the two asymptotics (isolated vertex, path failure) can be combined for an even better approximation, but we do not pursue this further here.

7. DISCUSSION

Graph connectivity is an important measure in network theory. In this work, we have presented a simple Monte Carlo algorithm which consists of removing edges at random until the graph becomes disconnected. As the number of simulations increases, this method recovers *exactly* the average of the reliability polynomial. It can also be used to estimate the full reliability polynomial, whose exact computation is an NP-complete problem. We presented simple asymptotic estimates for reliability polynomial of sufficiently dense graphs, based on the probability of getting an isolated vertex or a two-vertex graph. For sparse "subway-type" networks, we presented a different estimate based on the number of loops in the graph. All of these estimates have been shown to work for many real-world networks as well as random regular graphs.

We end this paper with some open questions and conjectures.

Open question 1: Choose a random graph consisting of n_1 vertices of degree 2, and n_2 vertices of degree 3. Suppose that $n_1 \gg n_2$, in which case the graph is subway-like. What are the asymptotics of AR in this case?

Open question 2 What is the average reliability of Erdős-Rényi graphs? The degree distribution is Poisson in this case.

More generally, what is the "best" degree distribution if we wish to optimize AR?

Open question 3. Among all graphs on fixed number of n vertices and m edges, what is the graph that maximizes AR?

Consider the case of $n = 12$ and $m = 18$ (which includes all cubic graphs of 12 vertices). We used Brendon McKay program Nauty [20] to generate a total of about 2×10^7 such graphs. Further restricting to only graphs whose vertices have a minimum degree of 2 yields about 2 million graphs. By contrast, there are only 87 cubic graphs on 12 vertices. Figure 6 shows the plot of AR versus the algebraic connectivity (AC) for this collection of graphs.

The unique maximizer of AR=0.350925 has AC=1.467911, which is the second-highest AC. The unique maximizer of AC=1.438447 has AR=0.350792, which is the second-highest AR. Among high-AR graphs, the first 28 are cubic, having AR from 0.3509 to 0.3429. The highest non-cubic graph has AR of 0.34121, and has the girth of 5. By contrast, there are many non-cubic graphs with high AC; among the 30 graphs having the highest AC, only 2 are cubic. In terms of girth, there are a total of 7 girth-5 graphs, two of which are cubic.

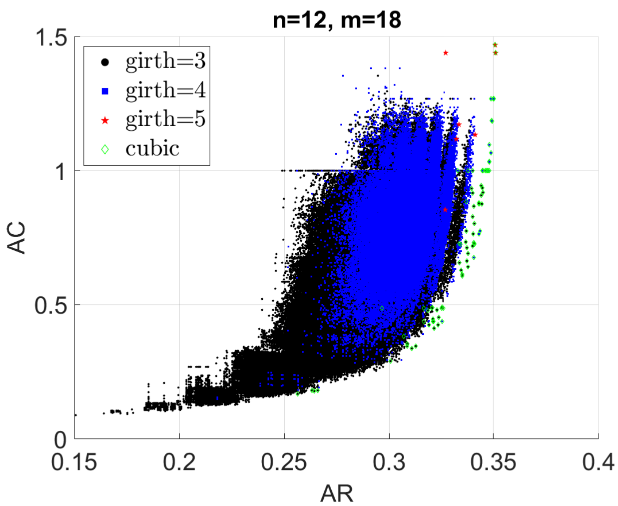

FIG. 6. Algebraic connectivity (AC) versus average reliability (AR) for all connected graphs on 12 vertices and 18 edges with miminum degree 2. There is a total 2189608 such graphs, of which 7 have girth 5, 74021 have girth 4, and the remaining 2115580 graphs have girth 3.

Based on these observations, we propose the following conjecture.

Conjecture 1. Among all the graphs of fixed number of nodes n and fixed integer average degree d, the graph that maximizes AR is a $d-$regular graph.

Finally, let us mention some related problems. In [21, 22], the authors modelled human frailty and death events using complex scale-free networks. In their model, nodes are damaged at random (simulating an ageing process) until too much damage is accumulated in the main nodes. It would be interesting to derive the asymptotics of their model.

Network reliability also has a connection with percolation theory. The percolation process can be viewed as a version of the edge removal process on a grid but with a different termination condition, namely the emergence of a giant component, rather than the network becoming disconnected [23]. Depending on the situation, this may be a better measure of network reliability than a simple disconnection threshold.

Finally, our approach for analysing network reliability can also be applied to the computation of two-terminal reliability [6].

[1] M. Fiedler, Algebraic connectivity of graphs, Czechoslovak Mathematical Journal 23 (2) (1973) 298–305.
[2] A. Ghosh, S. Boyd, Growing well-connected graphs, 2006 45th IEEE Conference on Decision and Control (2006) 6605–6611.
[3] R. Olfati-Saber, R. M. Murray, Consensus problems in networks of agents with switching topology and time-delays, Automatic Control, IEEE Transactions on 49 (9) (2004) 1520–1533.
[4] K. Ogiwara, T. Fukami, N. Takahashi, Maximizing algebraic connectivity in the space of graphs with a fixed number of vertices and edges, IEEE Transactions on Control of Network Systems 4 (2) (2015) 359–368.
[5] T. Kolokolnikov, Maximizing algebraic connectivity for certain families of graphs, Linear Algebra and its Applications 471 (2015) 122–140.
[6] C. J. Colbourn, The combinatorics of network reliability, International Series of Monographs on Computer Science, The Clarendon Press, Oxford University Press, New York, 1987.
[7] J. I. Brown, D. Cox, R. Ehrenborg, The average reliability of a graph, Discrete Applied Mathematics 177 (2014) 19–33.
[8] J. Brown, Maximal intervals of decrease and inflection points for node reliability, Discrete Applied Mathematics 313 (2022) 40–44.
[9] G. G. Lorentz, Bernstein polynomials, Mathematical Expositions, No. 8, University of Toronto Press, Toronto, 1953.
[10] R. T. Farouki, The bernstein polynomial basis: A centennial retrospective, Computer Aided Geometric Design 29 (6) (2012) 379–419.
[11] M. HOLMES, Introduction to perturbation methods, Texts in Applied Mathematics (1995).
[12] W. Zhou, M. Janic, R. Kooij, F. Kuipers, On the availability of networks, in: s.n. (Ed.), Proceedings of Broadband 2007, IET, 2007, pp. 1–6, null ; Conference date: 03-12-2007 Through 06-12-2007.
[13] S. Knight, H. Nguyen, N. Falkner, R. Bowden, M. Roughan, The internet topology zoo, Selected Areas in Communications, IEEE Journal on 29 (9) (2011) 1765 –1775. doi:10.1109/JSAC.2011.111002.
[14] M. Molloy, B. Reed, A critical point for random graphs with a given degree sequence, Random structures & algorithms 6 (2-3) (1995) 161–180.
[15] L. Federico, R. Van der Hofstad, Critical window for connectivity in the configuration model, Combinatorics, Probability and Computing 26 (5) (2017) 660–680.
[16] T. Luczak, Sparse random graphs with a given degree sequence, in: Proceedings of the Symposium on Random Graphs, Poznan, 1989, pp. 165–182.
[17] P. Erdős, A. Rényi, On random graphs I., Publ. Math. Debrecen 6 (1959) 290–297.
[18] R. A. Rossi, N. K. Ahmed, The network data repository with interactive graph analytics and visualization, in: AAAI, 2015.
URL http://networkrepository.com
[19] S. Derrible, C. Kennedy, The complexity and robustness of metro networks, Physica A: Statistical Mechanics and its Applications 389 (17) (2010) 3678–3691. doi:https://doi.org/10.1016/j.physa.2010.04.008.
URL https://www.sciencedirect.com/science/article/pii/S0378437110003262
[20] B. D. McKay, Nauty user's guide (version 2.4), Computer

Science Dept., Australian National University (2007) 225–239.

[21] S. G. Farrell, A. B. Mitnitski, K. Rockwood, A. D. Rutenberg, Network model of human aging: frailty limits and information measures, Physical Review E 94 (5) (2016) 052409.

[22] A. Mitnitski, A. Rutenberg, S. Farrell, K. Rockwood, Aging, frailty and complex networks, Biogerontology 18 (4) (2017) 433–446.

[23] S. R. Broadbent, J. M. Hammersley, Percolation processes: I. crystals and mazes, in: Mathematical proceedings of the Cambridge philosophical society, Vol. 53, Cambridge University Press, 1957, pp. 629–641.