Which Networks Are Least Susceptible to Cascading Failures?

Larry Blume [*]   David Easley [†]   Jon Kleinberg [‡]   Robert Kleinberg [§]   Éva Tardos [¶]

July 2011

Abstract. The resilience of networks to various types of failures is an undercurrent in many parts of graph theory and network algorithms. In this paper we study the resilience of networks in the presence of *cascading failures* — failures that spread from one node to another across the network structure. One finds such cascading processes at work in the kind of contagious failures that spread among financial institutions during a financial crisis, through nodes of a power grid or communication network during a widespread outage, or through a human population during the outbreak of an epidemic disease.

A widely studied model of cascades in networks assumes that each node v of the network has a threshold $\ell(v)$, and fails if it has at least $\ell(v)$ failed neighbors. We assume that each node selects a threshold $\ell(v)$ independently using a probability distribution μ. Our work centers on a parameter that we call the μ-risk of a graph: the maximum failure probability of any node in the graph, in this threshold cascade model parameterized by threshold distribution μ. This defines a very broad class of models; for example, the large literature on edge percolation, in which propagation happens along edges that are included independently at random with some probability p, takes place in a small part of the parameter space of threshold cascade models, and one where the distribution μ is monotonically decreasing with the threshold. In contrast we want to study the whole space, including threshold distributions with qualitatively different behavior, such as those that are sharply increasing.

We develop techniques for relating differences in μ-risk to the structures of the underlying graphs. This is challenging in large part because, despite the simplicity of its formulation, the threshold cascade model has been very hard to analyze for arbitrary graphs G and arbitrary threshold distributions μ. It turns out that when selecting among a set of graphs to minimize the μ-risk, the result depends quite intricately on μ. We develop several techniques for evaluating the μ-risk of d-regular graphs. For $d = 2$ we are able to solve the problem completely: the optimal graph is always a clique (i.e. triangle) or tree (i.e. infinite path), although which graph is better exhibits a surprising non-monotonicity as the threshold parameters vary. When $d > 2$ we present a technique based on power-series expansions of the failure probability that allows us to compare graphs in certain parts of the parameter space, deriving conclusions including the fact that as μ varies, at least three different graphs are optimal among d-regular graphs. In particular, the set of optimal graphs here includes one which is neither a clique nor a tree.

1 Introduction

The resilience of networks to various types of failures is an undercurrent in many parts of graph theory and network algorithms. For example, the definitions of cuts and expansion each capture types of robustness in the presence of worst-case edge or node deletion, while the study of network reliability is based on the question of connectivity in the presence of probabilistic edge failures, among other issues.

In this paper we are interested in the resilience of networks in the presence of *cascading failures* — failures that spread from one node to another across the network structure. One finds such cascading processes at work in the kind of contagious failures that spread among financial institutions during a financial crisis [1], in the breakdowns that spread through nodes of a power grid or communication network during a widespread outage [3], or in the course of an epidemic disease as it spreads through a human population [2].

To represent cascading failures we use the following basic *threshold cascade model*, which has been studied extensively both in the context of failures and also in other settings involving social or biological contagion [6, 8, 9, 10, 11, 12, 13, 14].[1] We are given a graph G, and each node v chooses a *threshold* $\ell(v)$ independently from a distribution μ on the natural numbers, choosing threshold $\ell(v) = j$ with probability $\mu(j)$. The quantity $\ell(v)$ represents the number of failed neighbors that v can withstand before v fails as well — thus we can think of μ as determining the distribution of levels of "health" of the nodes in the population, and hence implicitly controlling the way the failure process spreads on G. To determine the outcome of the failure process, we first declare all nodes with threshold 0 to have failed. We then repeatedly check whether any node v that has not yet failed has at least $\ell(v)$ failed neighbors — if so, we declare v to have failed as well, and we continue iterating. For example, Figure 1 shows the outcome of this process on two different graphs G with particular choices of node thresholds.

For a given node r in G, we define its *failure probability* $f_\mu(G, r)$ to be the probability it fails when node thresholds $\ell(v)$ are drawn independently from μ and then the threshold cascade model is run with these thresholds. Now we let $f_\mu^*(G) = \sup_{r \in V(G)} f_\mu(G, r)$, namely, the *maximum failure probability in G*. We view $f_\mu^*(G)$ as our measure of the resilience of G against cascading failures that operate under the threshold distribution μ; accordingly, we refer to $f_\mu^*(G)$ as the μ-*risk* of G, and we seek graphs of low μ-risk.

A Motivating Contrast: Cliques and Trees. How do different network structures compare in their resilience to a cascading failure? Because the failure probability clearly goes up as we add edges to a given node set, we take the top-level issue of edge density out of consideration by posing this question over the set of all (finite or infinite) connected d-regular graphs, for a fixed choice of d. We use $\mathcal{G}_d$ to denote this set of graphs, and for graphs in $\mathcal{G}_d$ we ask how they compare according to their μ-risk.[2] When we consider $\mathcal{G}_d$, we will also restrict the threshold distributions to the set of all distributions supported on $\{0, 1, 2, \ldots, d\}$, a set which we denote by Γ_d.

As a first concrete example of the kind of results to come, we consider a comparison between two basic d-regular graphs; the analysis justifying this comparison will follow from the framework

---

[1]The threshold cascade model is also related to the *nonlinear voter model* [7], though somewhat different in its specifics.

[2]Unless explicitly noted otherwise, all quantification over graphs in this paper takes place over the set of connected graphs only. This does not come at any real loss of generality, since the μ-risk of a disconnected graph is simply the supremum of the μ-risk in each connected component.
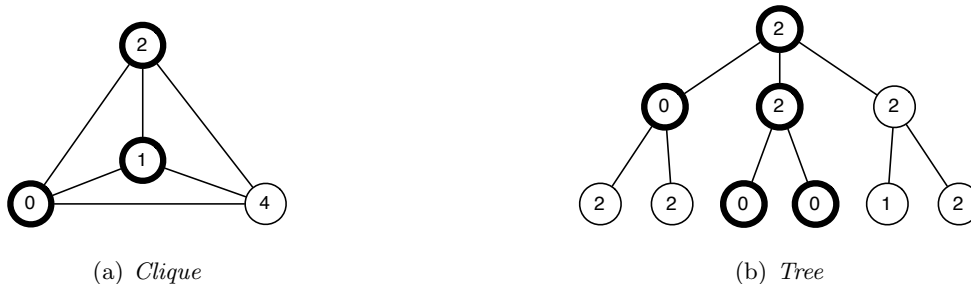
(a) *Clique*   (b) *Tree*

Figure 1: The spread of failures on two graphs according to the threshold cascade model. On each graph, the thresholds are drawn inside the nodes, and the nodes with thick borders are those that fail as a result of the process.

developed in the paper. To begin with, for conjecturing structures that produce low μ-risk, we can draw on intuitions from the motivating domains discussed above. A standard notion in epidemic disease is that it is dangerous to belong to a large connected component, and this suggests the clique K_{d+1} as a resilient network. On the other hand, a principle in financial networks is that it is important to have diversity among one's neighbors — in the present context, a lack of edges among one's neighbors — so that shocks are uncorrelated. This suggests the infinite complete d-ary tree T_d as a resilient network. (By way of illustration, note that if we were to continue the tree in Figure 1(b) indefinitely downward, we would have the complete 3-ary tree T_3.)

An intriguing point, of course, is that these two sources of intuition point in completely opposite directions. But as one consequence of the framework we develop here (in Section 4) we will see that both intuitions are essentially correct — each of K_{d+1} or T_d can be better than the other, for different choices of the threshold distribution. Specifically, we will show that there exist $\mu, \nu \in \Gamma_d$ such that $f_\mu^*(K_{d+1}) < f_\mu^*(T_d)$ and $f_\nu^*(T_d) < f_\nu^*(K_{d+1})$.

In fact, this trade-off between cliques and trees shows up in instructive ways on very simply parametrized subsets of the space Γ_d. For example, suppose we choose a very small value $\varepsilon > 0$, and for a variable x we define $(\mu(0), \mu(1), \mu(2)) = (\varepsilon, x, 1 - \varepsilon - x)$ with $\mu(j) = 0$ for $j > 2$. Then when $x = 1 - \varepsilon$, so that all thresholds are either 0 or 1, a node's failure probability is strictly increasing in the size of the component it belongs to, and so K_{d+1} uniquely minimizes the μ-risk. At the other extreme, when $x = 0$, a short argument shows that K_{d+1} uniquely optimizes the μ-risk here too. But as we prove in Section 5.1, it is possible to choose a value of x strictly between 0 and $1 - \varepsilon$ for which T_d has strictly lower μ-risk than K_{d+1}. Moreover, the value of x where T_d has lower μ-risk accords with the financial intuition about the value of diversity: it occurs when x is very small, but significantly larger than ε, so that thresholds of 1 are much more numerous than thresholds of 0. In this case, failures are still rare, but if a node u has connected neighbors v and w, then there is a non-trivial risk that v will have threshold 0 and w will have threshold 1, at which point v's failure will ricochet off w and bring down u as well, even if u has the maximum (and most likely) threshold of 2. In this region of the space Γ_d of threshold distributions, it is safer to have no links among your neighbors, even at the expense of producing very large connected components.

There is also an important qualitative message underlying this contrast: the question of which graph is more resilient against cascading failures depends sensitively on the way in which failure moves through the graph (via the mixture of thresholds determined by μ).

3

This contrast, and the reasons behind it, suggest that the space Γ_d has a rich structure when viewed in terms of the μ-risk it induces on graphs. Indeed, as we've just seen, even monotonic trade-offs between simple parameters of $\mu \in \Gamma_d$ can produce non-monotonic transitions between graphs — for example, with K_{d+1} first being better, then worse, then better again compared to T_d as we vary x above.

Our overall plan in this paper is thus to develop techniques for relating differences in μ-risk to the structures of the underlying graphs. This is challenging in large part because, despite the simplicity of its formulation, the threshold cascade model has been very hard to analyze for arbitrary graphs G and arbitrary threshold distributions μ. Existing results have either made the strong assumptions that either μ obeys a *diminishing property* (that threshold probabilities exhibit some form of monotonic decrease in the threshold size) [10, 12] or that the underlying graph G is a tree [8, 14], a lattice [7], or a complete graph [9, 13]. In fact, even the existing techniques developed specifically for cliques and trees do not appear strong enough to identify the contrast discussed above, which emerges from our framework in Section 5.1. And for comparing graphs outside these special cases, very few tools are available; one of our motivating goals is to develop tools of this type.

It is also worth noting that the large literature on edge percolation, in which propagation happens along edges that are included independently at random with some probability p, deals with a particular class of models that — when viewed in terms of thresholds — have the diminishing property discussed above. This includes the large literature on $G_{n,p}$, viewed as random edge sets of the complete graph [5]; the authors' own recent work on network formation in the presence of contagion exclusively used a model based on this type of edge percolation [4]. The point is that for this special case, component size is the dominant effect, and so the graphs of minimum μ-risk are essentially cliques; working in this part of the space thus does not enable one to look at trade-offs between "open" and "closed" neighborhoods as in our motivating discussion of K_{d+1} vs. T_d. (As we will see, the constructions of $\mu \in \Gamma_d$ that favor T_d indeed involve thresholds with a sharply *increasing* property over part of the support set; for certain applications, this increasing property is often viewed as crucial, which accords with the intuition discussed earlier.) Hence we need to look beyond models with an edge percolation structure to see things that even qualitatively resemble the phenomena we are trying to study.

Summary of Results. The contrast between K_{d+1} and T_d establishes that there is no single graph H such that H achieves the minimum μ-risk for all distributions $\mu \in \Gamma_d$. It is thus natural to ask whether K_{d+1} and T_d are sufficient to jointly "cover" the space Γ_d, in the sense that at least one of them is optimal at each $\mu \in \Gamma_d$. More generally, we say that a (finite or infinite) set of graphs $\mathcal{H} = \{H_1, H_2, \ldots\} \subseteq \mathcal{G}_d$ is a *sufficient set* for Γ_d if for each $\mu \in \Gamma_d$, at least one member of $\mathcal{H}$ achieves the minimum μ-risk over all graphs in $\mathcal{G}_d$. In this terminology, our question becomes:

(∗) *Does $\{K_{d+1}, T_d\}$ form a sufficient set for Γ_d?*

One consequence of the results in the paper is a complete answer to Question (∗). We find, in fact, that the answer to this question depends on the value of d.

We begin with a fairly complete analysis of μ-risk for the case of degree $d = 2$, answering Question (∗) affirmatively in this case. While the set of graphs in $\mathcal{G}_2$ is clearly very simple (cycles of each length ≥ 3, and the infinite path), the behavior of μ-risk on $\mathcal{G}_2$ is still rich enough that the non-monotonic phenomenon discussed above takes place even between K_3 and T_2. (Observe that

4

T_2, the infinite 2-ary tree, is better known as the infinite path). We find in fact that at each μ with $0 < \mu(0) < 1$, at least one of K_3 or T_2 achieves strictly lower μ-risk than every other graph in $\mathcal{G}_2 - \{K_3, T_2\}$.

When $d > 2$, the behavior of μ-risk on $\mathcal{G}_d$ becomes much more complicated. Here we establish that for each $d > 2$, the two graphs $\{K_{d+1}, T_d\}$ do not form a sufficient set for Γ_d. We do this by considering a graph that we call the (d-regular) *tree of triangles* Δ_d, consisting essentially of a collection of disjoint triangles attached according to the structure of an infinite regular tree. (Δ_d is specified precisely in Section 5.2, and depicted schematically for the case $d = 3$ in Figure 2). We construct a distribution $\mu \in \Gamma_d$ for which Δ_d has strictly lower μ-risk than both K_{d+1} and T_d. Intuitively, the tree of triangles "interpolates" between the complete neighborhood diversification of T_d and the complete neighborhood closure of K_{d+1}, and hence points toward a further structural dimension to the problem of minimizing μ-risk.

Despite the complex structure of μ-risk when $d > 2$, we have a set of results making it possible to compare the μ-risk of certain specific graphs to the μ-risk of arbitrary graphs. In addition to the comparisons among K_{d+1}, T_d, and Δ_d described above, we establish the following further results for K_{d+1} and T_d. First, as noted above, it is not hard to show that there are distributions $\mu \in \Gamma_d$ for which K_{d+1} has strictly lower μ-risk than any other $G \in \mathcal{G}_d$. A much more intricate argument establishes a type of optimality property for T_d as well: for each graph $G \in \mathcal{G}_d$, we construct a distribution $\mu_G \in \Gamma_d$ for which T_d has strictly lower μ_G-risk than G. This is a broad generalization of the T_d-vs.-K_{d+1} comparison, in that it says that such a comparison is possible for every $G \in \mathcal{G}_d$: in other words, T_d is more resilient than every other connected d-regular graph at some point in Γ_d.

Our analysis in fact establishes a strengthening of this result for T_d — for every finite set $\mathcal{H}$ of connected d-regular graphs, there is a distribution $\mu_{\mathcal{H}} \in \Gamma_d$ on which T_d achieves strictly lower $\mu_{\mathcal{H}}$-risk than each member of $\mathcal{H}$. And this in turn yields a negative answer to a more general version of Question ($*$): When $d > 2$, there is no two-element sufficient set of graphs for Γ_d.

Our results for $d > 2$ are based on a unifying technique, motivated by the construction of the distribution $\mu = (\varepsilon, x, 1 - \varepsilon - x)$ used to compare K_{d+1} and T_d above. The technique is based on using power series approximations to study the μ-risk for μ in the vicinity of particular threshold distributions; roughly speaking, it works as follows. We focus on cases in which the distribution μ concentrates almost all of its probability on a single threshold $\ell_{\max}$ and the remaining probability is divided up over values $j < \ell_{\max}$. The random draw of a threshold from μ in this case can be treated as a small perturbation of the fixed threshold distribution in which every node gets threshold $\ell_{\max}$ and no nodes fail. A given node's failure probability can then be expressed using a power series in the variables $\{\mu(j) \mid j < \ell_{\max}\}$ and the power series coefficients for different graphs provide enough information to compare them according to μ-risk when the probabilities $\{\mu(j) \mid j < \ell_{\max}\}$ are sufficiently close to zero. The computation of the power series coefficients then reduces to a counting problem involving certain partial assignments of thresholds to nodes of G.

In addition to their role in our analyses, we believe that small perturbations of a single fixed threshold are a very natural special case to consider for the threshold cascade model. Specifically, let $\Gamma_d^h(x) \subseteq \Gamma_d$ be the set of distributions in Γ_d such that $\mu(0) > 0$, $\mu(j) < x$ for $j < h$, and $\mu(j) = 0$ for $j > h$. (In other words, most of the probability mass is concentrated on h, and the rest is on values below h.) Threshold distributions in $\Gamma_d^h(x)$ for small $x > 0$ correspond to scenarios in which all nodes begin with a fixed level of "health" h, and then a shock to the system causes a small fraction of nodes to fail, and a small fraction of others to be weakened, with positive thresholds

below h. The study of μ-risk on $\Gamma_d^h(x)$ corresponds simply to the question of which networks are most resilient to the effect of such shocks.

Overall, then, we believe that the techniques developed here suggest avenues for further progress on a set of basic questions involving the threshold cascade model, including sharper comparisons of the μ-risk between different graphs, and how these comparisons depend both on μ and on the underlying graph structure.

2   Definition of the model

In the threshold cascade model, there is a graph G (possibly infinite) in which each node v randomly samples a label $\ell(v) \in \mathbb{N}$. Given a labeling ℓ of graph G, we define a subset $S \subseteq V(G)$ to be *failure-stable* if every node $v \notin S$ has strictly fewer than $\ell(v)$ neighbors in S. We define the set of failed nodes $\Phi(G, \ell)$ to be the intersection of all failure-stable node sets.

Given a graph G with root vertex r, and a distribution μ on node labels, we define the *root failure probability* to be the probability that $r \in \Phi(G, \ell)$ when ℓ is randomly sampled by assigning each node an independent label with distribution μ. We denote the root failure probability by $f_\mu(G, r)$.

It is not hard to see that this definition of $\Phi(G, \ell)$ is equivalent to the one we used in the introduction as stated by the following lemma.

Lemma 2.1. *The set $\Phi(G, \ell)$ is failure-stable. It is also equal to the union of the infinite sequence of sets $\Phi_0(G, \ell) \subseteq \Phi_1(G, \ell) \subseteq \cdots$ defined inductively by specifying that $\Phi_0(G, \ell) = \{v \mid \ell(v) = 0\}$ and $\Phi_{i+1}(G, \ell) = \{v \mid \Phi_i(G, \ell)$ contains at least $\ell(v)$ neighbors of $v\}$. It is also equal to the set of all nodes $v \in V(G)$ such that $v \in \Phi(G_0, \ell)$ for some finite subgraph $G_0 \subseteq G$.*

Proof. It is easy to see that the intersection of failure-stable sets is failure-stable, hence $\Phi(G, \ell)$ is failure-stable. The containment $\bigcup_i \Phi_i(G, \ell) \subseteq \Phi(G, \ell)$ is obvious from the definition of $\Phi_i(G, \ell)$. To establish the reverse containment, it suffices to show that $\bigcup_i \Phi_i(G, \ell)$ is failure-stable. This holds because any v having $\ell(v)$ or more neighbors in $\bigcup_i \Phi_i(G, \ell)$ must also have that same number of neighbors in $\Phi_k(G, \ell)$ for some sufficiently large k; it then follows that $v \in \Phi_{k+1}(G, \ell)$.

If G_0 is any finite subgraph of G and $v \in \Phi(G_0, \ell)$, then $v \in \Phi(G, \ell)$ since $\Phi(G_0, \ell) \subseteq \Phi(G, \ell)$. To prove the converse, we will show that the set of all v such that $v \in \Phi(G_0, \ell)$ for some finite G_0 is a failure-stable set. Indeed, suppose that v has neighbors $w_1, \ldots, w_{\ell(v)}$, each belonging to a finite subgraph $G_0(w_i)$ such that $w_i \in \Phi(G_0(w_i), \ell)$. Then $v \in \Phi\left(\cup_{i=1}^{\ell(v)} G_0(w_i), \ell\right)$, as desired. $\square$

3   The case $d = 2$

In this section, we specialize to 2-regular undirected graphs G. For any such graph, one can define a permutation R of the vertex set such that for every $v \in V(G)$, the set of neighbors of v is $\{R(v), R^{-1}(v)\}$. The following algorithm ROOTFAIL processes a labeling ℓ of G and outputs "fail" if and only if the root vertex r belongs to $\Phi(G, \ell)$. The algorithm works as follows. First it inspects the label $\ell(r)$: if this is not equal to 1 or 2, then it halts instantly and outputs "fail" if and only if $\ell(r) = 0$. Otherwise, find the least i such that $R^i(r) \neq 1$ and the least j such that $R^{-j}(r) \neq 1$. Let $\ell^+ = \ell(R^i(r))$, $\ell^- = \ell(R^{-j}(r))$. If i is undefined, then set $i = \infty$ and $\ell^+ = 2$. Similarly, if j is undefined then set $j = \infty$ and $\ell^- = 2$. Now, if $\ell(r) = 1$, output "fail" if and only if $\ell^+ = 0$ or

$\ell^- = 0$. If $\ell(r) = 2$, output "fail" if and only if $\ell^+ = 0$ and $\ell^- = 0$. Define the length of an execution of this algorithm to be equal to $i + j$. (Note that if $i = \infty$ or $j = \infty$, the algorithm ROOTFAIL will not actually halt. For this reason, an actual implementation of ROOTFAIL would have to be more careful to inspect the vertices in interleaved order — $R(r), R^{-1}(r), R^2(r), R^{-2}(r), \ldots$ — until it can prove that the root must fail. Such an implementation is not guaranteed to halt, but when processing any labeling ℓ such that $r \in \Phi(G, \ell)$ it is guaranteed to halt after a finite number of steps and output "fail".)

The key to analyzing the root failure probability in 2-regular graphs is the following observation: there is a probabilistic coupling of the labelings ℓ_P of the infinite path P and the labelings ℓ_C of the n-cycle $C = C_n$, such that for every sample point at which ROOTFAIL(P, ℓ_P) has execution length less than n, ROOTFAIL(C, ℓ_C) also has execution length less than n and the two executions are identical.

We now define some events on the sample space of this coupling. For any k, let $\mathcal{E}_k$ denote the event that ROOTFAIL(P, ℓ_P) has execution length at least k. Let $\mathcal{F}_P$ denote the event that $r \in \Phi(P, \ell_P)$ and let $\mathcal{F}_C$ denote the event that $r \in \Phi(C, \ell_C)$. Since the executions of ROOTFAIL(P, ℓ_P) and ROOTFAIL(C, ℓ_C) are identical on the complement of $\mathcal{E}_n$, we find that

$$\Pr(\mathcal{F}_P) - \Pr(\mathcal{F}_C) = \Pr(\mathcal{E}_n) \cdot [\Pr(\mathcal{F}_P \mid \mathcal{E}_n) - \Pr(\mathcal{F}_C \mid \mathcal{E}_n)].$$

We now proceed to compute each of the conditional probabilities on the right-hand side. Let s, t, u denote the label probabilities $\mu(0), \mu(1), \mu(2)$, respectively. Let $q = \frac{s}{1-t}$, which is the conditional probability that the label of any node is 0, given that its label is not 1. Then we have

$$\Pr(\mathcal{F}_P \mid \mathcal{E}_n) = \frac{t}{t+u}\left(1 - (1-q)^2\right) + \frac{u}{t+u} \cdot q^2.$$

The first term on the right accounts for the case that $\ell(r) = 1$ and the second term accounts for the case that $\ell(r) = 2$. After some manipulation — pulling out $\left(\frac{t}{t+u}\right) q$ from the first term and $\left(\frac{u}{t+u}\right) q$ from the second one — we obtain the formula $\Pr(\mathcal{F}_P \mid \mathcal{E}_n) = q + \frac{t-u}{t+u}\left(q - q^2\right)$. To compute $\Pr(\mathcal{F}_C \mid \mathcal{E}_n)$, note that when $\mathcal{E}_n$ occurs, the root's label is either 1 or 2, and at most one of the remaining labels is not equal to 1. Furthermore, in any such labeling of C, the root fails if and only if one of the other $n-1$ nodes has label 0. Thus,

$$\Pr(\mathcal{E}_n) = (t+u)[t^{n-1} + (n-1)(1-t)t^{n-2}]$$

$$\Pr(\mathcal{E}_n \cap \mathcal{F}_C) = (t+u)(n-1)st^{n-2}$$

$$\Pr(\mathcal{F}_C \mid \mathcal{E}_n) = \frac{(n-1)s}{t + (n-1)(1-t)} = q\left(1 - \frac{t}{t + (n-1)(1-t)}\right)$$

$$\Pr(\mathcal{F}_P \mid \mathcal{E}_n) - \Pr(\mathcal{F}_C \mid \mathcal{E}_n) = \frac{t-u}{t+u}\left(q - q^2\right) + \frac{qt}{t + (n-1)(1-t)}$$

$$\Pr(\mathcal{F}_P) - \Pr(\mathcal{F}_C) = \Pr(\mathcal{E}_n) \cdot \left[\frac{t-u}{t+u}\left(q - q^2\right) + \frac{qt}{t + (n-1)(1-t)}\right].$$

On the last line, both factors are decreasing functions of n. Consequently, when they are both positive, their product is a decreasing function of n. In other words, if an n-cycle is better than an infinite path, then an $(n-1)$-cycle is better still.

We have thus proved the following.

Theorem 3.1. *For each $\mu \in \Gamma_2$, at least one of the 3-cycle or the infinite path has minimum μ-risk over all graphs in $\mathcal{G}_2$.*

4   Computing Failure Probabilities Via Power Series

When $d > 2$, the method of the preceding section does not appear to be applicable. In effect, since the breadth-first search of such a graph builds a tree which, at any stage of the search, may have more than two leaves (in fact, an unbounded number of them) there are many more opportunities for correlation as different leaves of the tree are discovered to refer to the same node of G. For this reason, an analysis along the lines of Section 3 seems hopeless. Instead we specialize to cases in which the distribution μ concentrates almost all of its probability on a single label $\ell_{\max}$ and the remaining probability is divided up over labels $j < \ell_{\max}$. We then express the μ-risk as a power series in the probabilities $\{\mu(j) \mid j < \ell_{\max}\}$, which allows us to compare different graphs according to their low-degree power series coefficients.

4.1   Definitions

We now present the definitions that we need, followed by a description of the power series for the root failure probability and its convergence properties. Throughout this section, we will illustrate the definition on a very simple graph: a 3-node path, with the root r placed at the middle node, and we let v and w be the two other (leaf) nodes of the path.

Throughout this section and the following ones, we will assume that labels take values in the set $\{0, \ldots, \ell_{\max}\}$ for some fixed positive integer $\ell_{\max}$. For purposes of our example, we assume that $\ell_{\max}$, where most of the probability is concentrated, is equal to 2: $\mu(0) = s$ and $\mu(1) = t$ are small positive numbers, and $\mu(2) = 1 - s - t$ is close to 1.

We will compute failure probabilities by working with partial node labelings λ, in which labels are assigned to only some of the nodes, i.e., a partial function λ from $V(G)$ to $\{0, \ldots, \ell_{\max}\}$. Its *domain of definition*, $\mathrm{Dom}(\lambda)$, is the set of all $v \in V(G)$ such that $\lambda(v)$ is defined; when $\mathrm{Dom}(\lambda) = V(G)$ we refer to λ as a *full labeling* or simply a *labeling*.

We say that a partial labeling λ is an *explanation of root failure (ERF)* if the root fails in every full labeling of G that agrees with λ on $\mathrm{Dom}(\lambda)$. We say that λ is a *minimal explanation of root failure (MERF)* if it is an ERF, and every proper sublabeling of λ is not an ERF. Note that $\mathrm{Dom}(\lambda)$ is a finite set whenever λ is a MERF, by Lemma 2.1.

Thus, on the three-node path with r in the middle, there are four MERFs: **(a)** assigning 0 to r; **(b)** assigning 1 to r and 0 to v; **(c)** assigning 1 to r and 0 to w; and **(d)** assigning 0 to v and w. We can think of partial labelings as events in the full sample space of labelings, and (a)-(d) are thus four events that cover the event that r fails. Hence the probability r fails is bounded above by the sum of the probabilities of these four events, which is $s + 2st + s^2$.

To get the precise failure probability of r, we need to incorporate inclusion-exclusion terms arising from overlaps in these four MERFs. In our example, there are two distinct labelings that correspond to such overlaps:

(i) assigning 0 to all three nodes: this arises when events (a) and (d) both occur, so it contributes $-s^3$ to the probability.

(ii) assigning 1 to r and 0 to both v and w: this arises when any two out of (b), (c), and (d) occur, and also when all three occur. By the inclusion-exclusion formula, this contributes

$-3s^2t + s^2t = -2s^2t$ to the probability, with the first term coming from two-way overlaps and the second term coming from the three-way overlap.

Putting all this together, we get the root failure probability for the small example: $s + 2st + s^2 - s^3 - 2s^2t$.

MERFS give rise to such overlaps when they are compatible. We sat that two partial labelings λ_1, λ_2 are *compatible* if $\lambda_1(v) = \lambda_2(v)$ for every $v \in \text{Dom}(\lambda_1) \cap \text{Dom}(\lambda_2)$. The *union* of two compatible partial labelings λ_1, λ_2 is the unique partial function λ such that

$$\{(v, \lambda(v)) \mid v \in \text{Dom}(\lambda)\} = \{(v, \lambda_1(v)) \mid v \in \text{Dom}(\lambda_1)\} \cup \{(v, \lambda(v)) \mid v \in \text{Dom}(\lambda_2)\}.$$

For notational reasons, it will be convenient to make the union operation into a binary operation that is defined for any pair of partial labelings, not only for compatible pairs. To do so, we define the set Λ to be a set consisting of all partial labelings, together with one special element denoted $\perp$ that is interpreted to be incompatible with every element of Λ, including itself. We extend the union operation $\cup$ to a binary operation on Λ by specifying that $\lambda_1 \cup \lambda_2 = \perp$ when λ_1 and λ_2 are incompatible. For a partial labeling λ, we define $\mathcal{E}(\lambda)$ to be the set of all full labelings that extend λ; note that $\mathcal{E}(\perp) = \emptyset$, and that for every two partial labelings λ_1, λ_2 we have the relation $\mathcal{E}(\lambda_1) \cap \mathcal{E}(\lambda_2) = \mathcal{E}(\lambda_1 \cup \lambda_2)$.

For the inclusion-exclusion formula, we'll need to think about finite unions of MERFs which we'll call *UMERFs*. For graph G with root vertex r, we will denote the set of all MERFs by $\mathbb{M}(G, r)$ and the set of all UMERFs by $\mathbb{U}(G, r)$. We will sometimes abbreviate these to $\mathbb{M}, \mathbb{U}$ when the identity of the graph and root vertex are obvious from context.

We can now describe the plan for arbitrary graphs, including infinite ones, when $\mu(j) = s_j$ are small numbers for $j < \ell_{\max}$, and $\mu(\ell_{\max}) = 1 - \sum_{j=0}^{\ell_{\max}-1} s_j$. We first show that when $\ell_{\max} > d/2$, for any vector of natural numbers $\mathbf{i} = (i_0, i_1, \ldots, i_{\ell_{\max}-1})$, there are only finitely many MERFs that assign i_k nodes a label of k, for $k = 0, \ldots, \ell_{\max} - 1$. Moreover, we can write the root's failure probability as a multivariate power series of the form $\sum_{\mathbf{i}} a_{\mathbf{i}} s_0^{i_0} s_1^{i_1} \cdots s_{\ell_{\max}-1}^{i_{\ell_{\max}-1}}$, and this power series has a positive radius of convergence. We use this to compare failure probabilities in different graphs by enumerating a finite set of terms in the power series until we identify a difference between them.

4.2   A power series for computing the root failure probability

We make the set of all labelings ℓ into a probability space by declaring the labels $\{\ell(v) \mid v \in V(G)\}$ to be independent random variables with common distribution μ. The measurable sets in this probability space are the σ-field generated by the sets $\mathcal{E}(\lambda)$, where λ ranges over all partial labelings of G.

By Lemma 2.1, whenever the root fails there is a MERF that explains the failure, i.e. the event $r \in \Phi(G, \ell)$ is the union of the events $\mathcal{E}(\lambda)$ for $\lambda \in \mathbb{M}$. Since $\mathbb{M}$ is a countable set, we can choose an arbitrary one-to-one correspondence $m : \mathbb{N} \to \mathbb{M}$. Then

$$\Pr(r \in \Phi(G, \ell)) = \Pr\left(\bigcup_{i=1}^{\infty} \mathcal{E}(m(i))\right) = \lim_{n \to \infty} \Pr\left(\bigcup_{i=1}^{n} \mathcal{E}(m(i))\right). \tag{1}$$

Each of the probabilities on the right-hand side can be expanded using the inclusion-exclusion

9

formula:

$$\Pr\left(\bigcup_{i=1}^{n}\mathcal{E}(m(i))\right) = \sum_{k=1}^{n}(-1)^{k+1}\sum_{1\leq i_1<\cdots<i_k\leq n}\Pr\left(\mathcal{E}(m(i_1))\cap\cdots\cap\mathcal{E}(m(i_k))\right)$$

$$= \sum_{k=1}^{n}(-1)^{k+1}\sum_{1\leq i_1<\cdots<i_k\leq n}\Pr\left(\mathcal{E}(m(i_1)\cup\cdots\cup m(i_k))\right). \tag{2}$$

The right-hand side of (2) is easy to evaluate: using variables s_i $(i = 0, \ldots, \ell_{\max})$ to denote the values $s_i = \mu(i)$, the probability of the event $\mathcal{E}(\lambda)$ for any partial labeling is given by

$$\Pr(\mathcal{E}(\lambda)) = \prod_{v\in\text{Dom}(\lambda)} s_{\lambda(v)} \stackrel{\Delta}{=} s_\lambda, \tag{3}$$

where this is taken as the definition of s_λ.

Combining (2) and (3), and regrouping the terms we get the following lemma.

Lemma 4.1.

$$\Pr\left(\bigcup_{i=1}^{n}\mathcal{E}(m(i))\right) = \sum_{\lambda\in\mathbb{U}}\sum_{k=1}^{n}(-1)^{k+1}a_\lambda^{k,n}s_\lambda. \tag{4}$$

Here, $a_\lambda^{k,n}$ for a UMERF λ and integers $1 \leq k \leq n$, is defined to be the number of distinct k-tuples $(i_1, \ldots, i_k)$ such that $1 \leq i_1 < \cdots < i_k \leq n$ and $\lambda = m(i_1) \cup \cdots \cup m(i_k)$.

4.3   Convergence of the power series

To take the limit as $n \to \infty$ and obtain a well-defined power series, it is necessary to have a finiteness theorem that justifies that the coefficient of s_λ eventually stabilizes as n grows. In fact, in order for the power series to have positive radius of convergence the coefficients must grow no faster than exponentially. Proving such bounds requires bounding the number of UMERFs of a given size. In general this is not possible: for some graphs and some settings of the parameter $\ell_{\max}$, the number of UMERFs of a specified size is not even finite. As a simple example, consider an infinite path and $\ell_{\max} = 1$; there are infinitely many MERFs λ consisting of a single node labeled with 0. This example generalizes to any positive even degree d: the graph G is formed from an infinite sequence of independent sets of size $d/2$, with every two consecutive such independent sets being joined by a complete bipartite graph. When $\ell_{\max} = d/2$, there are infinitely many MERFs obtained by taking one of the independent sets in the sequence and labeling all of its nodes with 0. Each of these MERFs λ has $i(\lambda) = (d/2, 0, \ldots, 0)$.

The remainder of this section is devoted to specifying some sufficient conditions under which the right-hand side of Equation (4) can be rewritten as a power series with positive radius of convergence. For any partial labeling λ, we define its *size* $|\lambda| = |\text{Dom}(\lambda)|$ to be the number of nodes it labels. We begin by identifying some sufficient conditions under which we can assert that for every partial labeling λ, the number of nodes that are guaranteed to fail in every labeling extending λ is at most $O(|\lambda|)$.

Lemma 4.2. *Suppose we are given a graph G, a default threshold $\ell_{\max}$, and a partial labeling λ. Let $\overline{\lambda}$ be the full labeling that extends λ by assigning label $\ell_{\max}$ to each node not labeled by λ, and let $F = \Phi\left(G, \overline{\lambda}\right)$.*

10

1. *If G is d-regular and $d < 2\ell_{\max}$ then $|F| \leq (d+1)|\lambda|$.*

2. *Suppose that for every node v of G, every connected component of $G \setminus \{v\}$ contains strictly fewer than $\ell_{\max}$ neighbors of v. Then $|F| < 2|\lambda|$.*

Proof. Arrange the elements of F into a sequence $v_1, v_2, \ldots$ such that each of the sets $\Phi_i\left(G, \overline{\lambda}\right)$ is an initial segment of the sequence. Thus, each $v \in F$ has at least $\overline{\lambda}(v)$ neighbors that precede it in the sequence. We can think of the sequence $v_1, v_2, \ldots$ as specifying a possible order in which the nodes of F failed in an execution of the threshold cascade model. To prove both parts of the lemma we will define a potential function that maps vertex sets to non-negative integers, then evaluate the potential function on each initial segment of the sequence, and consider how the value of the potential function changes every time a new node fails (i.e., is added to the initial segment). We will use two different potential functions corresponding to the two parts.

For Part 1 define $\varphi(S)$, for any vertex set S, to be the number of edges of G having one endpoint in S and the other in its complement. Each time a new node v_k fails, it increases the value of φ by at most d since it has only d neighbors. Furthermore, if $v_k \notin \mathrm{Dom}(\lambda)$ then v_k has at least $\ell_{\max}$ neighbors that precede it in the sequence and at most $d - \ell_{\max}$ that succeed it. Thus, the net change in φ is bounded above by $(d - \ell_{\max}) - \ell_{\max}$, which is at most -1 by our assumption that $d < 2\ell_{\max}$. The potential function φ thus starts at 0, increases by at most $d|\lambda|$ over the whole sequence of failures, and is never negative; hence there can be at most $d|\lambda|$ steps of the sequence when it strictly decreases, and therefore at most $d|\lambda|$ nodes in $F \setminus \mathrm{Dom}(\lambda)$. Consequently $|F| \leq (d+1)|\lambda|$.

For Part 2, we instead use the potential function $\psi(S)$ defined as the number of connected components in the induced subgraph $G[S]$. Each time a new node v_k fails, it increases ψ by at most 1. Now consider how ψ changes when a node $w \notin \mathrm{Dom}(\lambda)$ fails. Since $\overline{\lambda}(w) = \ell_{\max}$, we know that w has at least $\ell_{\max}$ neighbors that precede it in the sequence. By our assumption on the structure of G, at least two of these neighbors belong to different connected components of $G \setminus \{w\}$. These components merge together when w fails, causing ψ to decrease by at least 1. Since the initial value of ψ is 0 and its final value is strictly positive, and it increases by at most 1 in each step, we know that the number of steps in which ψ increases must be greater than the number of steps in which it decreases. Hence, $|F \setminus \mathrm{Dom}(\lambda)| < |\lambda|$, implying $|F| < 2|\lambda|$ as claimed. $\square$

The next lemma provides a simple method for bounding the number of UMERFs of size z by an exponential function of z.

Lemma 4.3. *Suppose, for a given graph G and default threshold $\ell_{\max}$, that there exists a constant c such that every partial labeling λ satisfies $|\Phi\left(G, \overline{\lambda}\right)| \leq c|\lambda|$. Then for every z, the number of UMERFs of size z is at most $(d+1)^{3cz}$. In particular, this upper bound is at most $(d+1)^{3(d+1)z}$ whenever one of the sufficient conditions in Lemma 4.2 holds.*

Proof. Let λ be a partial labeling and let $F = \Phi\left(G, \overline{\lambda}\right)$. If λ is a MERF, then F induces a connected subset of G, since otherwise we could remove the labels provided by λ in any component of $G[F]$ not containing the root r and arrive at a proper sublabeling of λ that is also an ERF. This implies that if λ is a UMERF, the set $G[F]$ must also be connected, since it is the union of a finite set of connected graphs all containing a common node r. We can describe any such F uniquely by specifying the sequence of edge labels (each indexed from 1 to d) that are taken in the $\leq 2|F|$ steps of a depth-first search traversal of $G[F]$ starting from r. Hence there are at most $d^{2|F|} \leq d^{2c|\lambda|}$ such sets. As each UMERF of size $|\lambda|$ is uniquely associated with such a set F together with a labeling

of its nodes, we obtain an upper bound of $d^{2cz} \cdot (1 + \ell_{\max})^{cz}$ on the number of UMERFs of size z. The lemma follows because $1 + \ell_{\max} \leq d + 1$. $\qquad \square$

Assume for the remainder of this section that G and $\ell_{\max}$ satisfy one of the two sufficient conditions in Lemma 4.2; thus, the hypothesis of Lemma 4.3 holds with $c = d + 1$. The conclusion of Lemma 4.3 is already enough for us to be able to express the series on the right-hand side of Equation (4) via a more useful indexing. First, for any UMERF λ, let $i(\lambda)$ denote the vector of natural numbers $\mathbf{i} = (i_0, i_1, \ldots, i_{\ell_{\max}})$ such that λ assigns exactly i_k nodes a label of k. The corresponding event $\mathcal{E}(\lambda)$ has probability $s_\lambda = s_0^{i_0} s_1^{i_1} \cdots s_{\ell_{\max}}^{i_{\ell_{\max}}}$, a quantity we will abbreviate as $s^{\mathbf{i}}$.

For any vector of natural numbers $\mathbf{i} = (i_0, i_1, \ldots, i_{\ell_{\max}})$, let $|\mathbf{i}| = \sum_{k=0}^{\ell_{\max}} i_k$; the number of UMERFs λ with $i(\lambda) = \mathbf{i}$ is bounded by the expression in Lemma 4.3, with $z = |\mathbf{i}|$ and $c = d + 1$. Moreover, any MERF λ' that appears in a union of MERFs forming λ must have a vector $i(\lambda')$ that is coordinate-wise dominated by $i(\lambda)$, and hence Lemma 4.3 implies that only a finite set of MERFs can appear in unions that form λ. It follows that the sequence of coefficients $a_\lambda^{k,n}$ eventually stabilizes as $n \to \infty$ — that is, for every λ, k there is an integer a_λ^k and a threshold n_0 such that $a_\lambda^{k,n} = a_\lambda^k$ for all $n \geq n_0$.

Thus we can group together all UMERFs λ with $i(\lambda) = \mathbf{i}$ and write

$$\Pr\left(\bigcup_{i=1}^{\infty} \mathcal{E}(m(i)) \right) = \sum_{\mathbf{i}} \sum_{\substack{\lambda \in \mathbb{U} \\ i(\lambda) = \mathbf{i}}} \sum_k (-1)^{k+1} a_\lambda^k s^{\mathbf{i}} = \sum_{\mathbf{i}} a_{\mathbf{i}} s^{\mathbf{i}}, \tag{5}$$

where the right-hand side should be taken as the definition of $a_{\mathbf{i}}$, and the grouping by $\mathbf{i}$ in the sum on the right-hand side is justified by the fact that in the preceding triple summation, the sums over λ and k range over finite sets.

If we can show that $a_{\mathbf{i}}$ depends only exponentially on $|\mathbf{i}|$, this will establish that the power series has a positive radius of convergence. We observe that if the third summation weren't present in Equation (5), and instead we only were summing over $k = 1$ (corresponding to MERFs), then such an exponential upper bound would follow directly from Lemma 4.3. It follows that to show an exponential upper bound on $|a_{\mathbf{i}}|$, it is sufficient, for each fixed UMERF λ with $i(\lambda) = \mathbf{i}$, to show that $|\sum_k (-1)^{k+1} a_\lambda^k|$ is bounded above by an exponential function of $|\mathbf{i}|$.

To do this, we consider the (potentially very large) set of all MERFs $\lambda_1, \ldots, \lambda_m$ that can appear in a union forming λ. Let $\mathrm{Dom}(\lambda) = D$, with $|D| = n$, and $\mathrm{Dom}(\lambda_j) = D_j$. For each subset of k of these MERFs whose union equals D, we get a term $(-1)^{k+1}$ in the sum we are bounding. We would like to show that the absolute sum of all these terms is bounded above by an exponential function of n, but since there could be many more than this many terms in the sum, we need an argument that actually exploits the cancellation among terms of the form $(-1)^{k+1}$, rather than naïvely treating each as potentially having the same sign.

The upper bound we need follows from our next lemma.

Lemma 4.4. *Let D be an n-element set, and let $D_1, \ldots, D_m$ be (not necessarily distinct) subsets of D. Let $\mathcal{C}$ be the collection of all subsets $J \subseteq \{1, \ldots, m\}$ for which $\bigcup_{j \in J} D_j = D$. Then*

$$\left| \sum_{J \in \mathcal{C}} (-1)^{|J|} \right| \leq 2^n.$$

(The crucial point is that the right-hand side is independent of m.)

12

Proof. We prove this by induction on n, with the case of $n = 1$ being easy. For $n > 1$, choose any element $x \in D$ and let $D' = D - x$. We define $\mathcal{C}_0$ to be the collection of all $J \subseteq \{1, \ldots, m\}$ for which $\bigcup_{j \in J} D_j \supseteq D'$. and $\mathcal{C}_1$ to be the collection of all $J \subseteq \{1, \ldots, m\}$ for which $\bigcup_{j \in J} D_j = D'$. Now, by the induction hypothesis applied to the sets D' and $\{D_j - x : j = 1, 2, \ldots, m\}$, we have $\left| \sum_{J \in \mathcal{C}_0} (-1)^{|J|} \right| \leq 2^{n-1}$. By the induction hypothesis applied to the sets D' and $\{D_j : x \notin D_j\}$, we have $\left| \sum_{J \in \mathcal{C}_1} (-1)^{|J|} \right| \leq 2^{n-1}$. Finally, $\mathcal{C}_1 \subseteq \mathcal{C}_0$ and $J \in \mathcal{C}$ if and only if $J \in \mathcal{C}_0 - \mathcal{C}_1$, so we have $\sum_{J \in \mathcal{C}} (-1)^{|J|} = \sum_{J \in \mathcal{C}_0} (-1)^{|J|} - \sum_{J \in \mathcal{C}_1} (-1)^{|J|}$, from which it follows that $\left| \sum_{J \in \mathcal{C}} (-1)^{|J|} \right| \leq 2^n$. $\square$

Putting these bounds together, we see that $|a_{\mathbf{i}}|$ is bounded above by an exponential function of $|\mathbf{i}|$, and hence:

Theorem 4.5. *If $d < 2\ell_{\max}$, the power series in Equation (5) has a positive radius of convergence. The power series also has a positive radius of convergence if for every node v, every connected component of $G \setminus \{v\}$ contains strictly fewer than $\ell_{\max}$ neighbors of v.*

5   Comparing Cliques, Trees, and Trees of Triangles

5.1   Comparing T_d to K_{d+1}

In the introduction, we noted that it is easy to identify two distinct settings of the parameters for μ for which K_{d+1} has uniquely optimal μ-risk among connected d-regular graphs. First, when $\ell_{\max} = 1$, the probability the root fails is monotonic in the size of the connected component that contains it, and K_{d+1} uniquely minimizes this for connected d-regular graphs. But K_{d+1} is also uniquely optimal for larger values of $\ell_{\max} \leq d$, when μ assigns every label to be either 0 or $\ell_{\max}$. Indeed, in this case, the only way the root can fail in K_{d+1} is if at least $\ell_{\max}$ of its neighbors fail. This event also causes the root to fail in any connected d-regular graph G, but when $G \neq K_{d+1}$ there are other positive-probability events that also cause the root to fail, so again K_{d+1} is uniquely optimal.

As a first application of our power-series technique, we now show that there are parameter settings for which T_d has lower root failure probability than K_{d+1}. For this comparison, we consider μ such that $\ell_{\max} = 2$, and label 0 has probability s, while label 1 has probability t, where s and t are small quantities that will be defined precisely later. Observe that when $\ell_{\max} = 2$, T_d satisfies the hypothesis of Lemma 4.2, Part 2, and hence its power series has a positive radius of convergence. The power series for K_{d+1} is actually a polynomial in s and t, since K_{d+1} is a finite graph, so its radius of convergence is infinite.

Let us work out some of the low-degree terms for T_d and for K_{d+1}. For T_d, the coefficient on the term s is 1, corresponding to the MERF in which the root gets labeled 0. The coefficient on the term st is d, corresponding to MERFs in which the root gets labeled 1 and any one of the root's d neighbors gets labeled 0. There are no inclusion-exclusion corrections contributing to either of these coefficients.

For K_{d+1}, the coefficient on the term s is 1, as in T_d, corresponding to the root getting labeled 0. However, the coefficient on the term st is d^2: there are d MERFs in which the root gets labeled 1 and any one of the root's d neighbors gets labeled 0; there are also $d(d-1)$ more MERFs in which one neighbor of the root gets labeled 0 and another gets labeled 1.

Now, suppose we set $s = t^3$. Then the power series for the root failure probability in T_d is $t^3 + dt^4 + O(t^5)$, whereas the power series for the root failure probability in K_4 is $t^3 + d^2 t^4 + O(t^5)$.
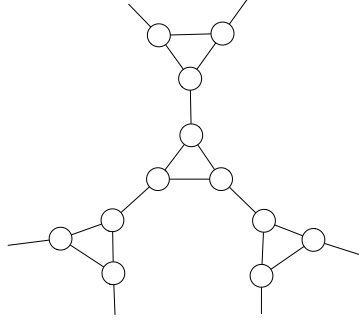
13

Figure 2: The tree of triangles Δ_d for $d = 3$.

The $O(t^5)$ estimate of the error term is valid inside the power series' radius of convergence. Hence, for t sufficiently small and $s = t^3$, we find that $f_\mu^*(T_d) < f_\mu^*(K_{d+1})$.

We have thus shown

Theorem 5.1. *For each $d \geq 3$, there exists a $\mu \in \Gamma_d$ for which T_d has strictly lower μ-risk than K_{d+1}.*

5.2 Comparing Δ_d to K_{d+1} and T_d

We now show that when $d > 2$, the graphs $\{K_{d+1}, T_d\}$ do not form a sufficient set for Γ_d. We do this by establishing the following theorem.

Theorem 5.2. *For each $d \geq 3$, there exists a $\mu \in \Gamma_d$ for which the d-regular tree of triangles Δ_d has strictly lower μ-risk than either T_d or K_{d+1}.*

The d-regular tree of triangles Δ_d is a graph consisting of a collection of disjoint triangles connected according to the structure of T_{3d-6}: it is the graph obtained by starting from T_{3d-6} and replacing each node u (with neighbors $v_1, \ldots, v_{3d-6}$ in T_{3d-6}) by three nodes $\{u_1, u_2, u_3\}$. These three nodes u_1, u_2, u_3 are mutually connected into a triangle, and u_i is also connected to one node in each of the triangles that replaces v_j, for $j = (i-1)(d-2) + 1, \ldots, i(d-2)$. We draw a small portion of Δ_d's repeating structure, in the case $d = 3$, in Figure 2.

We construct the distribution μ in Theorem 5.2 from a small perturbation of the fixed threshold $\ell_{\max} = 3$. To analyze the root failure probability in Δ_d in this case, we first observe that its power series has a positive radius of convergence for all $d \geq 3$, since Δ_d satisfies the hypothesis of Lemma 4.2, Part 2. (A connected component of $\Delta_d \setminus \{v\}$ can contain at most 2 neighbors of v.) Thus, we can compare the root failure probabilities in Δ_d, K_{d+1}, and T_d by comparing low-degree terms in their power series, as we did when we compared K_{d+1} with T_d in Section 5.1. Because the calculations are somewhat lengthy, we present them in Appendix A rather than in the main text.

6 Comparing T_d to an arbitrary d-regular graph

In Section 5.1 we compared $f_\mu^*(T_d)$ with $f_\mu^*(K_{d+1})$, for $d \geq 3$, when μ is a small perturbation of $\ell_{\max} = 2$ — that is, when $(\mu(0), \mu(1), \mu(2)) = (s, t, 1 - s - t)$. We saw that the tree has strictly

14

lower μ-risk than the clique when t is sufficiently small and s is sufficiently small relative to t. Generalizing this, the same power-series technique can be used to show that for *any* connected d-regular graph other than T_d, one can find a setting of $s, t > 0$ such that $f_\mu^*(T_d) < f_\mu^*(G)$. This will establish the following theorem, the proof of which is the main focus of the present section.

Theorem 6.1. *For each $d \geq 3$ and each graph $G \in \mathcal{G}_d$, there exists a $\mu_G \in \Gamma_d$ for which T_d has strictly lower μ_G-risk than G.*

Most of the proof applies to all values of $d \geq 3$. At the end of the analysis, we separately handle the cases of $d = 3$ and $d > 3$. Focusing on $d = 3$ first allows us to use the condition that $d < 2\ell_{\max} = 4$ and hence ensure that the root failure probability in the graph G has a power series expansion with a positive radius of convergence. After analyzing the case of $d = 3$, we extend the proof to $d > 3$; this still depends on evaluating power series coefficients but requires some new techniques to handle the potential non-convergence of the power series for G. But to begin with, we allow $d \geq 3$ to be arbitrary.

Since G is a connected graph that is not a tree, it has finite girth L. Let r be a node of G that belongs to an L-cycle, and let r' be an arbitrary node of $T = T_d$. Applying the results of Section 4, we will be bounding the probabilities $f_\mu(G, r)$ and $f_\mu(T, r')$ using sums of monomials s_λ indexed by UMERFs λ. Any such monomial $s_\lambda = s^i t^j$ has $i \geq 1$: all MERFs have at least one threshold-zero node, since otherwise the failed set is empty. We will be setting $s = t^L \ll 1$, so that all the monomials whose magnitude is greater than t^{2L-1} are of the form st^j ($0 \leq j \leq L - 2$). Focusing, therefore, on UMERFs λ having $i(\lambda) = (1, j)$, we will establish the facts summarized in the following lemma.

Lemma 6.2. *Let G be any d-regular graph of girth L.*

(1) If λ is any UMERF in G such that $i(\lambda) = (1, j)$, where $0 \leq j \leq L - 2$, then λ is a MERF.

(2) When $0 \leq j < L - 2$, there is a one-to-one correspondence between MERFs λ such that $i(\lambda) = (1, j)$ in G and in $T = T_d$.

(3) When $j = L - 2$, G has strictly more MERFs with $i(\lambda) = (1, j)$ than does T.

Proof. Let λ be a UMERF such that $i(\lambda) = (1, j)$, $0 \leq j \leq L - 2$, and let v be the unique node in $\text{Dom}(\lambda)$ such that $\ell(v) = 0$. If we extend λ to a labeling ℓ by assigning threshold 2 to every node not in $\text{Dom}(\lambda)$, then the failed set $\Phi(G, \ell)$ is, by Lemma 2.1, the union of an increasing sequence of sets $\Phi_0(G, \ell) \subseteq \Phi_1(G, \ell) \subseteq \cdots$. Each of these induces a connected subgraph containing v, since the initial set $\Phi_0(G, \ell)$ is the singleton $\{v\}$, and a node belonging to $\Phi_i(G, \ell)$ must have at least one neighbor in $\Phi_{i-1}(G, \ell)$. Now consider the smallest i (if any) such that $\Phi_i(G, \ell)$ contains a node $w \notin \text{Dom}(\lambda)$. As $\ell(w) = 2$, the node w must have two neighbors $x, y \in \Phi_{i-1}(G, \ell)$. Combining the edges $(w, x), (w, y)$ with a path from x to y in $\Phi_{i-1}(G, \ell)$ we obtain a cycle in G whose vertex set belongs to $\{w\} \cup \text{Dom}(\lambda)$, a set of cardinality $j + 2$.

When $j < L - 2$, this contradicts our assumption that G has girth L, and thus we may conclude that $\Phi(G, \ell) = \text{Dom}(\lambda)$. In this case, the induced subgraph on $\text{Dom}(\lambda)$ is a tree containing v and r, and in fact it must be a path connecting v to r. (Any threshold-1 nodes in $\text{Dom}(\lambda)$ lying outside this path cannot belong to any MERF, contradicting our assumption that λ is a UMERF.) It follows that λ is a MERF, and that the number of such MERFs, for a specified value of j, is equal to the number of j-hop simple paths in G terminating at r. Our assumption that G is d-regular,

with girth greater than $j + 2$, implies that *every* non-backtracking walk of j hops terminating at r is a simple path, and that the vertex sets of all these simple paths are distinct. Consequently, in both G and T the number of MERFs λ such that $i(\lambda) = (1, j)$ is equal to the number of j-hop non-backtracking walks in a d-regular graph, i.e. $d(d-1)^{j-1}$.

When $j = L - 2$, the set $\Phi(G, \lambda)$ induces either a tree (in which case, by the same reasoning as before, it must be a j-hop path from v to r) or an L-cycle containing a single node w such that $\ell(w) = 2$, a single node v with $\ell(v) = 0$, and all other nodes having label 1. In the latter case, our assumption that every node in $\mathrm{Dom}(\lambda)$ belongs to a MERF implies either that $w = r$, or that v, w, r occur consecutively on the cycle C and that $C \setminus \{w\}$ is a j-hop path from v to r. Finally, it is easy to see that in all of these cases, λ is a MERF. We have thus shown that every UMERF in G with $i(\lambda) = (1, L-2)$ is a MERF, and that the number of these MERFs is at least $d(d-1)^{L-3} + L - 1$. Here, $d(d-1)^{L-3}$ counts the number of $(L-2)$-hop paths terminating at the root — which is also the coefficient of st^{L-2} in the power series for T_d — and $L - 1$ counts the number of ways of labeling $C \setminus \{r\}$ with a single 0 and $L - 2$ 1's. $\qquad\square$

Proof of Theorem 6.1, $d = 3$ case.   As $d < 2\ell_{\max}$, the power series for $f_\mu(G, r)$ and $f_\mu(T, r')$ converge for sufficiently small s and t. Thus the difference $f_\mu(G, r) - f_\mu(T, r')$, may be expressed as $\sum_{\mathbf{i}=(i,j)}(a_{ij}^G - a_{ij}^T)s^i t^j$ where a_{ij}^G and a_{ij}^T are the power series coefficients in (5) for G and T, respectively. Grouping the terms into those with $Li + j \le 2L - 2$ and those with $Li + j \ge 2L - 1$, we find that the first set of terms includes only pairs (i, j) such that $i = 1$, $0 \le j \le L - 2$, and by Lemma 6.2,

$$\sum_{Li+j \le 2L-2}(a_{ij}^G - a_{ij}^T)s^i t^j = (a_{1,L-2}^G - a_{1,L-2}^T)st^{L-2} \ge (L-1)t^{2L-2}. \tag{6}$$

Recall, from Lemmas 4.3 and 4.4, that the number of UMERFs λ such that $i(\lambda) = (i, j)$ is bounded above by $(d+1)^{3(d+1)(i+j)}$ and that the coefficient $\sum_k (-1)^{k+1}a_\lambda^k$ for each of them is bounded by 2^{i+j} in absolute value. Thus,

$$
\left| \sum_{Li+j \ge 2L-1}(a_{ij}^G - a_{ij}^T)s^i t^j \right| \le \sum_{k=2L-1}^{\infty} \sum_{Li+j=k} 2^{i+j+1}(d+1)^{3(d+1)(i+j)}t^{Li+j}
$$

$$
< \sum_{k=2L-1}^{\infty} k 2^{k+1}(d+1)^{3(d+1)k}t^k
$$

$$
< \sum_{k=2L-1}^{\infty} [4(d+1)^{3(d+1)}t]^k \;\; = \;\; \frac{\left(4(d+1)^{3(d+1)}\right)^{2L-1} t^{2L-1}}{1 - 4(d+1)^{3(d+1)}t},
$$

where the last line is justified as long as the denominator is strictly positive. By choosing t sufficiently small, we can ensure not only that the denominator is strictly positive but that the quantity on the last line is less than t^{2L-2}. Then, the positive $(L-1)t^{2L-2}$ contribution from the low-degree terms in the power series more than offsets the possibly negative contribution from the high-degree terms, and this proves that $f_\mu(G, r) > f_\mu(T, r')$, as claimed.

Proof of Theorem 6.1, $d > 3$ case.   When comparing the infinite d-regular tree T_d against another connected d-regular graph G when $d \ge 2\ell_{\max}$, a tricky issue arises because the power series for G need not converge. Recall, however, that the power series for T_d still converges. Thus, to

compute the root failure probablity $f_\mu(T, r')$ we will continue to use the full power series, and we will continue to denote it by $\sum_{i,j} a_{ij}^T s^i t^j$.

To estimate $f_\mu(G, r)$ we must adopt a different approach. Specifically, we number MERFs $m(1), m(2), \ldots$ in order of increasing $Li + j$, where $(i, j) = \mathbf{i}(m)$. Instead of taking the union of the entire countable sequence of events $\mathcal{E}(m(k))$ ($k = 1, 2, \ldots$) we truncate this sequence at the highest value of n such that $\mathbf{i}(m(n)) = (i, j)$ with $Li + j \leq 2L - 2$. We bound $f_\mu(G, r)$ from below by $\Pr\left(\bigcup_{i=1}^n \mathcal{E}(m(i))\right)$ and evaluate this probability using (4). Denote the resulting polynomial by $\sum_{i,j} a_{ij}^G s^i t^j$. Note that it is a polynomial, not a power series, since we are taking a union of only finitely many events, each described by a MERF.

With this revised interpretation of the coefficients a_{ij}^G, the bound in Equation (6) is once again justified by Lemma 6.2. To bound the remainder term $\left|\sum_{Li+j \geq 2L-1}(a_{ij}^G - a_{ij}^T)s^i t^j\right|$ from above, we use

$$\left|\sum_{Li+j \geq 2L-1}(a_{ij}^G - a_{ij}^T)s^i t^j\right| \leq \left|\sum_{Li+j \geq 2L-1} a_{ij}^G s^i t^j\right| + \left|\sum_{Li+j \geq 2L-1} a_{ij}^T s^i t^j\right| \tag{7}$$

and deal with each term on the right-hand side separately. The second term, involving power series coefficients of T, is dealt with exactly as before, yielding an upper bound of order $O(t^{2L-1})$ for that term as $t \to 0$. The first term is a *polynomial* in s and t, not a power series. Upon substituting $s = t^L$, it becomes a univariate polynomial in t, in which each monomial has an exponent of $2L - 1$ or higher. No matter how large the coefficients of this polynomial may be, they are finite, and so as $t \to 0$ the absolute value of the polynomial is $O(t^{2L-1})$.

Thus, the right-hand side of (7) is $O(t^{2L-1})$ as $t \to 0$, and as before the proof finishes by observing that for small enough t this error term cannot possibly offset the positive $(L-1)t^{2L-2}$ contribution from the low-degree terms. The conclusion is that

$$\Pr_G\left(\bigcup_{i=1}^n \mathcal{E}(m(i))\right) - f_\mu(T, r') \geq (L-1)t^{2L-2} - O(t^{2L-1})$$

as $t \to 0$. Recalling that $f_\mu(G, r) \geq \Pr_G\left(\bigcup_{i=1}^n \mathcal{E}(m(i))\right)$, we conclude that the failure probability of r in G exceeds the failure probability of r' in T.

6.1 A Connection to Sufficient Sets

A strengthening of Theorem 6.1 has a consequence for sufficient sets, as we now discuss. (Recall that a set of graphs $\mathcal{H} \subseteq \mathcal{G}_d$ is a *sufficient set* for Γ_d if for each $\mu \in \Gamma_d$, at least one member of $\mathcal{H}$ achieves the minimum μ-risk over all graphs in $\mathcal{G}_d$.) We first describe the relevant strengthening of the theorem. Notice that the proof of Theorem 6.1 in fact shows something stronger than was claimed. If we have any finite set of graphs $\mathcal{H} \subseteq \mathcal{G}_d$, none of which is T_d, then we can define L to be the maximum girth of any graph in $\mathcal{H}$. Using this value of L, we can define a distribution μ just as before, and the analysis in the proof of Theorem 6.1 then directly establishes the following.

Theorem 6.3. *For every finite set $\mathcal{H}$ of connected d-regular graphs, there is a distribution $\mu_\mathcal{H} \in \Gamma_d$ for which T_d achieves strictly lower $\mu_\mathcal{H}$-risk than each member of $\mathcal{H}$.*

In other words, rather than simply being more resilient than any single other graph G at some point in Γ_d, the tree T_d is in fact simultaneously more resilient than any finite set of other graphs at some point in Γ_d.

From this stronger form of the result, we obtain the following immediate consequence.

Theorem 6.4. *When $d \geq 3$, there is no sufficient set of size 2 for Γ_d.*

Proof. If there were such a set $\mathcal{H} \subseteq \mathcal{G}_d$ of size 2, then it would have to contain K_{d+1}, since K_{d+1} uniquely minimizes the μ-risk for some distributions $\mu \in \Gamma_d$. The other graph in $\mathcal{H}$ can't be T_d, since by Theorem 5.2 there are μ for which Δ_d has strictly lower μ-risk than both K_{d+1} and T_d. But if the other graph in $\mathcal{H}$ were some $G \neq T_d$, then by Theorem 6.3 we could find a μ for which T_d has lower μ-risk than both K_{d+1} and G, and so this is not possible either. $\square$

References

[1] Franklin Allen and Douglas M. Gale. Financial contagion. *Journal of Political Economy*, 108(1):1–33, February 2000.

[2] Roy M. Anderson and Robert M. May. *Infectious Diseases of Humans.* Oxford University Press, 1992.

[3] Chalee Asavathiratham, Sandip Roy, Bernard Lesieutre, and George Verghese. The influence model. *IEEE Control Systems Magazine*, 21(6):52–64, December 2001.

[4] Larry Blume, David Easley, Jon Kleinberg, Robert Kleinberg, and Éva Tardos. Network formation in the presence of contagious risk. In *Proc. 12th ACM Conference on Electronic Commerce*, 2011.

[5] Bela Bollobás. *Random Graphs.* Cambridge University Press, second edition, 2001.

[6] Damon Centola and Michael Macy. Complex contagions and the weakness of long ties. *American Journal of Sociology*, 113:702–734, 2007.

[7] J. T. Cox and Richard Durrett. Nonlinear voter models. In Richard Durrett and Harry Kesten, editors, *Random Walks, Brownian Motion, and Interacting Particle Systems*, pages 189–202. Birkhauser, 1991.

[8] Peter Dodds and Duncan Watts. Universal behavior in a generalized model of contagion. *Physical Review Letters*, 92(218701), 2004.

[9] Mark Granovetter. Threshold models of collective behavior. *American Journal of Sociology*, 83:1420–1443, 1978.

[10] David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence in a social network. In *Proc. 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 137–146, 2003.

[11] Stephen Morris. Contagion. *Review of Economic Studies*, 67:57–78, 2000.

[12] Elchanan Mossel and Sebastien Roch. On the submodularity of influence in social networks. In *Proc. 39th ACM Symposium on Theory of Computing*, 2007.

[13] Thomas Schelling. *Micromotives and Macrobehavior.* Norton, 1978.

[14] Duncan J. Watts. A simple model of global cascades on random networks. *Proc. Natl. Acad. Sci. USA*, 99(9):5766–5771, April 2002.

A   Power-Series Computations Comparing Δ_d to K_{d+1} and T_d

In this appendix, we provide the details of the power-series computations that establish a distribution $\mu \in \Gamma_d$ for which the d-regular tree of triangles Δ_d has strictly lower μ-risk than both K_{d+1} and T_d.

To find a choice of parameters for which Δ_d has lower root failure probability than either T_d or K_{d+1}, we consider distributions μ for which we have $(\mu(0), \mu(1), \mu(2), \mu(3)) = (s, t, u, 1 - s - t - u)$ for small s, t, and u, and $\mu(j) = 0$ for $j > 3$. This is a small perturbation of a default threshold of $\ell_{\max} = 3$. Recall that in this case, Theorem 4.5 justifies using a power series for the root failure probability in both T_d and Δ_d; the use of a power series for K_{d+1} is justified because K_{d+1} is a finite graph, hence its power series is really a polynomial. We can proceed by comparing the monomials associated with these three graphs. It will turn out to be sufficient to go up to monomials of total degree at most 3 in order to identify choices for s, t, and u defining our distribution μ.

We work out the coefficients on all monomials of total degree up to 3, as follows. We do this tersely, listing the MERFs and UMERFs that account for each. We also omit terms without a factor of s, since every MERF and UMERF must assign at least one label of 0.

For K_{d+1}:

- $1s$: The root gets label 0.

- $0s^2$: No MERF or UMERF can assign only two labels of 0, since no third node will fail.

- dst: The root gets label 1 and one neighbor gets label 0.

- $0su$: No MERF or UMERF can assign just one 0 and one 2, since the node with label 2 will not fail.

- $\binom{d}{3}s^3$: There is a MERF for each 3-tuple of neighbors who are assigned labels of 0.

- $\binom{d}{2}(d-3)s^2 t$: There are $\binom{d}{2}(d-2)$ MERFs contributing to this term: in each, two of the root's neighbors are labeled 0, and a third is labeled 1. There are also $\binom{d}{2}$ UMERFs contributing to this term: in each, two neighbors of the root are labeled 0 and the root is labeled 1. Each of these UMERFs can be written as the union of two MERFs (the root labeled 1 and a neighbor labeled 0), so each contributes -1. Thus the total is $\binom{d}{2}(d-3)$.

- $\binom{d}{2}(d-1)s^2 u$: There are $\binom{d}{2}$ MERFs where the root gets labeled 2 and two neighbors are labeled 0; there are $\binom{d}{2}(d-2)$ more MERFs where two neighbors of the root are labeled 0 and a third is labeled 2.

- $\binom{d}{2}(d-2)st^2$: Two neighbors of the root get label 1 and a third gets label 0.

- $0su^2$: A single 0 cannot cause any nodes of label 2 to fail.

- $d(d-1)^2 stu$: There are $d(d-1)$ MERFs where the root has label 2, one neighbor has label 0, and another has label 1. There are $d(d-1)(d-2)$ more MERFs where the three of the neighbors have labels 0, 1, and 2 in some order.

For T_d, the terms of total degree up to 2, as well as the s^3 and su^2 terms, are the same as for K_{d+1}, and by the same arguments. For the other degree-3 terms:

- $-\binom{d}{2}s^2t$: The $\binom{d}{2}(d-2)$ MERFs from K_{d+1} for this term are not present in T_d, but the $\binom{d}{2}$ UMERFs each contributing -1 are.

- $\binom{d}{2}s^2u$: There are $\binom{d}{2}$ MERFs where the root gets labeled 2 and two neighbors are labeled 0. (The other $\binom{d}{2}(d-2)$ MERFs from K_{d+1}, where two neighbors of the root are labeled 0 and a third is labeled 2, are not present here.)

- $d(d-1)st^2$: A grandchild of the root gets label 0, and the parent of this child together with the root get label 1.

- $0stu$: There is no MERF or UMERF that assigns labels of 0, 1, and 2 to three nodes.

For Δ_d, we think of the root r as having neighbor set $Z = \{v_0, v_1, v_2, \ldots, v_{d-1}\}$, where the only edge among the nodes in Z is between v_0 and v_1. We will refer to neighbors of Z that do not belong to the set $Z \cup \{r\}$ as *depth-two neighbors*. Again the terms of total degree up to 2, as well as the s^3 and su^2 terms, are the same as for K_{d+1}, and by the same arguments. For the other degree-3 terms:

- $(2(d-2) - \binom{d}{2})s^2t$: There are $\binom{d}{2}$ UMERFs each contributing -1 as in K_{d+1} and T_d, and there are also $2(d-2)$ MERFs in which one of v_0 or v_1 is labeled 0, the other is labeled 1, and a neighbor of the root in $Z - \{v_0, v_1\}$ is labeled 0.

- $\binom{d}{2}s^2u$: We get the $\binom{d}{2}$ MERFs that were present in T_d, but not the additional $\binom{d}{2}(d-2)$ that were present in K_{d+1}.

- $(d+1)(d-2)st^2$: There are $(d-2)(d-1)$ MERFs where the root and a node $w \in Z - \{v_0, v_1\}$ are labeled 1, and one of w's depth-two neighbors is labeled 0. There are $2(d-2)$ more MERFs where, for $i \in \{0, 1\}$, the root and v_i are labeled 1, and one of v_i's depth-two neighbors is labeled 0.

- $2stu$: The root is labeled 2, one of v_0 or v_1 is labeled 1, and the other is labeled 0.

For one of these graphs G, let $f_G(s, t, u)$ denote the root failure probability, and let $g_G(s, t, u)$ denote the sum of all terms in the power series of total degree 4 and higher. Then we have

$$
\begin{aligned}
f_{K_{d+1}}(s,t,u) - f_{\Delta_d}(s,t,u) &= \left[\binom{d}{2}(d-3) - \left(2(d-2) - \binom{d}{2}\right)\right]s^2t + \left[\binom{d}{2}(d-1) - \binom{d}{2}\right]s^2u \\
&\quad + \left[\binom{d}{2}(d-2) - (d+1)(d-2)\right]st^2 + \left[d(d-1)^2 - 2\right]stu \\
&\quad + (g_{K_{d+1}}(s,t,u) - g_{\Delta_d}(s,t,u)) \\
&= \left[\binom{d}{2} - 2\right](d-2)s^2t + \left[\binom{d}{2}(d-2)\right]s^2u \\
&\quad + \left[\binom{d}{2} - (d+1)\right](d-2)st^2 + \left[d(d-1)^2 - 2\right]stu \\
&\quad + (g_{K_{d+1}}(s,t,u) - g_{\Delta_d}(s,t,u))
\end{aligned}
$$

and

$$
\begin{aligned}
f_{T_d}(s,t,u) - f_{\Delta_d}(s,t,u) &= -2(d-2)s^2 t + (d(d-1) - (d+1)(d-2))st^2 - 2stu \\
&\quad + (g_{T_d}(s,t,u) - g_{\Delta_d}(s,t,u)) \\
&= (4-2d)s^2 t + 2st^2 - 2stu + (g_{T_d}(s,t,u) - g_{\Delta_d}(s,t,u))
\end{aligned}
$$

Now, we choose $t = ds$ and $u = s$. When $d = 3$, we get

$$
f_{K_{d+1}}(s,t,u) - f_{\Delta_d}(s,t,u) = 27s^3 + \Theta(s^4),
$$

and when $d > 3$ each of the non-zero degree-3 terms of $f_{K_{d+1}}(s,t,u) - f_{\Delta_d}(s,t,u)$ has a positive coefficient, so we get

$$
f_{K_{d+1}}(s,t,u) - f_{\Delta_d}(s,t,u) = c_d s^3 + \Theta(s^4)
$$

for $c_d > 0$. We also have

$$
f_{T_d}(s,t,u) - f_{\Delta_d}(s,t,u) = \left[4d - 2d^2 + 2d^2 - 2d\right]s^3 + \Theta(s^4) = 2ds^3 + \Theta(s^4).
$$

It follows that for $\mu = (s, ds, s, 1 - (d+2)s)$ and s sufficiently small, Δ_d has a smaller root failure probability than either of K_{d+1} or T_d.