EXPECTED RANK AND RANDOMNESS IN ROOTED GRAPHS

DAVID EISENSTAT, JENNIFER FEDER, GREG FRANCOS, GARY GORDON, AND AMANDA REDLICH

Abstract. For a rooted graph G, let $EV_b(G;p)$ be the expected number of vertices reachable from the root when each edge has an independent probability p of operating successfully. We determine the expected value of $EV_b(G;p)$ for random trees, and include a connection to unrooted trees. We also consider rooted digraphs, computing the expected value of a random orientation of a rooted graph G in terms of $EV_b(G;p)$. We consider optimal location of the root vertex for the class of grid graphs, and we also briefly discuss a polynomial that incorporates vertex failure.

1. Introduction

We continue the study of the expected number of vertices that remain connected to a distinguished vertex (the *root*) in a rooted graph. When each edge of a rooted graph has an independent probability p of succeeding, we let $EV_b(G;p)$ be the polynomial in p that gives the expected size of the component containing the root (the subscript b denotes *branching* rank). More motivation can be found in [4] and [8], which can be thought of as logical prequels to this work.

When T is a rooted tree, $EV_b(T;p)$ has an especially simple form [1]: $EV_b(T;p) = \sum_{*\neq v\in V} p^{d(*,v)}$, where $d(*,v)$ is the distance from the root vertex $*$ to v. In this case, it is clear that the 'best' configuration is a rooted star (rooted at the central vertex) and the 'worst' configuration is a rooted path (rooted at a leaf). In the former case, $EV_b(T;p) = np$, and in the latter, we have $EV_b(T;p) = \sum_{i=1}^{n} p^i$, where n is the number of edges.

These extreme cases lead naturally to the question of what is the 'typical' value of $EV_b(T;p)$. More precisely, what is the expected number of vertices reachable from the root in a random rooted tree on n vertices? Answering this question occupies Section 2, where we compute an explicit formula for this expected value – essentially an average of averages. These formulas can be evaluated for specified values of p between 0 and 1, but when no information about the distribution of the random variable p is known, a Bayesian approach using $\int_0^1 EV_b(G;p)dp$ as a measure of the expected number of vertices (as a function of n) can be useful. Theorem 2.8 computes this integral for a random rooted tree with n edges: $\int_0^1 EV_b(T;p)dp \approx \frac{n}{\log n}$, a value between the extremes achieved by rooted stars (where $\int_0^1 EV_b(T;p)dp = n/2$) and rooted paths (where $\int_0^1 EV_b(G;p)dp \approx \log n$). Theorem 2.10 gives an interpretation analogous to Theorem 2.6 for a rooted tree polynomial $EV_p(T;p)$ based on a *pruning* rank function.

Key words and phrases. Expected rank, Probabilistic graph.
This work was supported by NSF grant DMS-0243763.

When T is an unrooted tree, we can define an expected rank polynomial $EV_{ur}(T;p)$ via a pruning rank function. This allows us to compute the average number of edges that can be pruned from a random tree, and we also give a connection between the rooted and unrooted pruning expected rank polynomials.

Rooted directed graphs are treated briefly in Section 3. When an undirected graph is randomly oriented, we expect the number of vertices reachable from the root to drop, on average. Corollary 3.2 gives an explicit formula: The expected value of $EV_b(D;p)$ is $EV_b(G;p/2)$, where D is a digraph obtained from the undirected graph G by randomly orienting the edges of G.

For specific classes of graphs, it can be difficult to decide on an optimal location for the root. In general, no such location exists that is optimal for all p. It is possible to construct a graph G in which each of k specified vertices is the optimal location for the root for some value of p between 0 and 1 [10]. Nevertheless, for certain classes of graphs, there is a unique vertex v (up to automorphisms of the graph) in which $EV_b(G_v;p) \geq EV_b(G_w;p)$ for all $0 \leq p \leq 1$ and all vertices w. In Section 4, we consider grids, where the 'obvious' choice for the root is optimal.

Section 5 considers a different question: what happens when the edges are reliable, but the vertices are not? We obtain explicit formulas for the polynomial $EV_v(G;p)$ similar to those in the edge failure case, and in some cases, we obtain much simpler formulas. However, most of the results we present are negative in the sense that they reveal less about the structure of the graph than the edge-failure version of the polynomial. In particular, it does not matter if two edges that are adjacent to the root vertex are joined by an edge. Thus, for instance, the rooted star (rooted at the center) has the same polynomial as a rooted complete graph.

We thank the anonymous referees for several useful suggestions that helped clarify the exposition.

2. Random trees

We begin with some preliminary definitions and results. Let G be a connected rooted graph with edge set E where each edge has the same independent probability p of being operational. For $S \subseteq E$, let $r_b(S)$ be the number of vertices (besides the root) in the component of the subgraph S that contains the root. This rank function is the *branching* rank of the associated greedoid, but we will not need the generality of greedoids in this paper.

Definition 2.1. Let G be a rooted graph. The expected value $EV_b(G;p)$ is

$$EV_b(G;p) = \sum_{S \subseteq E} r_b(S) p^{|S|}(1-p)^{|E-S|}.$$

We give two reformulations of this polynomial, both of which will be important throughout this work. For a vertex $v \in V$, let $Pr(v)$ denote the probability that v remains connected to the root. The following result appears explicitly in [4] and implicitly in [7] and [2, 3].

Proposition 2.2 (Proposition 2.7 of [4]). *Let G be a rooted graph. Then*

$$EV_b(G;p) = \sum_{* \neq v \in V} Pr(v).$$

We will also need the following deletion-contraction expansion for the polynomial.

Proposition 2.3 (Deletion-Contraction: Proposition 2.3 of [4])**.** *Let G be a rooted graph and let e ($\neq$ loop) be an edge adjacent to the root. Then*

$$EV_b(G;p) = (1-p)\cdot EV_b(G-e;p) + p\cdot EV_b(G/e;p) + p.$$

Now we shift our attention to rooted trees. In many applications, it is important to visit each vertex of a rooted tree in a specified order. We will use *preorder* labelings of rooted trees here.

Definition 2.4. A rooted tree T with labels $1,\dots,n$ is *preorder-labeled* if the root is labeled 1 and every other vertex has a label greater than its parent's.

These labelings are exactly those that arise from a preorder traversal of T. This gives a simple recursive structure as follows: to generate a random preorder-labeled tree on n vertices, generate one on $n-1$ vertices and then add a new leaf vertex with a randomly chosen parent.

We will consider two separate expected rank polynomials for rooted trees - one based on the rank function r_b given above (the *branching* rank) and one based on a complementary rank function (the *pruning* rank r_p).

2.1. **Branching rank.** When T is a rooted tree, the polynomial $EV_b(T;p)$ has an especially simple form. We restate this result, which we will use throughout this section.

Proposition 2.5 (Corollary 2.6 of [1])**.** *Let T be a rooted tree. Then*

$$EV_b(T;p) = \sum_{*\neq v\in V} p^{d(*,v)}.$$

What is the expected rank polynomial for a random rooted tree? We answer this question by computing an 'average of averages.' We let $B_n(p)$ be the expected value of $1+EV_b(T;p)$ when T has n vertices.

Theorem 2.6. *Let T be a random preorder-labeled rooted tree with n vertices, including the root. Then the expected value of $EV_b(T;p)$ is given by $B_n(p)-1 = -1+\prod_{k=1}^{n-1}(1+p/k)$.*

Proof. Clearly, $B_1(p)=1$. By Proposition 2.5, given a tree T, $1+EV_b(T;p) = 1+n_1p+n_2p^2+\dots$ can be viewed as a generating function where n_k is the number of vertices at depth k. In this light, each coefficient n_k in $B_n(p)$ is the expected number of vertices at depth k.

When $n>1$, T is determined by two independent choices: the preorder-labeled rooted subtree of size $n-1$ and the parent of vertex n. Thus $B_n(p) = B_{n-1}(p)+f(p)$ where $f(p)$ is the generating function for the distribution of the depths of vertex n. Since the depth of n is the depth of its parent plus one, the distribution of depths of n is exactly that of vertices $1,\dots,n-1$ shifted and rescaled. Multiplying by p accomplishes the shift; going from $n-1$ vertices to 1 requires division by $n-1$. As a result, $f(p) = pB_{n-1}(p)/(n-1)$.

This yields the recurrence:

$$B_1(p) = 1 \qquad\qquad B_n(p) = B_{n-1}(p) + p\frac{B_{n-1}(p)}{n-1}$$

Dividing both sides of the righthand equation by $B_{n-1}(p)$ yields

$$\frac{B_n(p)}{B_{n-1}(p)} = 1 + \frac{p}{n-1}$$

from which it is easily seen that $B_n(p) = -1 + \prod_{k=1}^{n-1}(1 + p/k)$.

$\square$

Note that $\prod_{k=1}^{n-1}(1 + p/k) = \prod_{k=1}^{n-1}(p + k)/k$, so $B_n(p) = \binom{n-1+p}{n-1} = \binom{n-1+p}{p}$, where the domain of the binomial coefficient has been extended via the gamma function.

Example 2.7. We illustrate Theorem 2.6 when $n = 4$. In this case, there are 6 labeled trees, as seen in Figure 1. These trees fall into 4 isomorphism classes. These trees have the following expected value polynomials, where $k = 2, 3$ or 4:

$$EV_b(T_1; p) = p + p^2 + p^3 \tag{1}$$

$$EV_b(T_k; p) = 2p + p^2 \tag{2}$$

$$EV_b(T_5; p) = p + 2p^2 \tag{3}$$

$$EV_b(T_6; p) = 3p \tag{4}$$

This gives an expected value of $\left(11p + 6p^2 + p^3\right)/6 = (1+p)\left(1 + \frac{p}{2}\right)\left(1 + \frac{p}{3}\right) - 1$, as required.

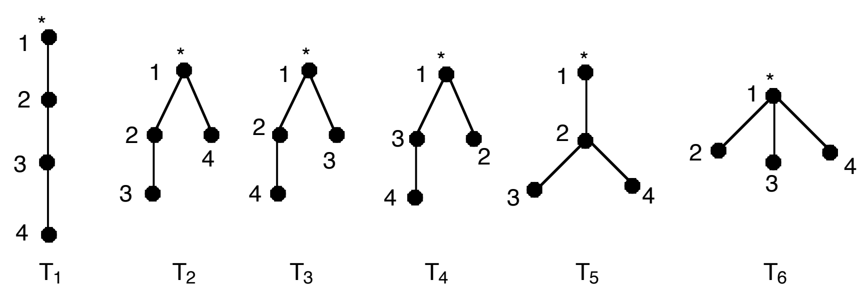

FIGURE 1. The 6 pre-ordered labeled trees on 4 vertices.

When no prior knowledge of the value of p is given, we can estimate the expected number of vertices which remain connected to the root with an integral: $\int_0^1 EV_b(T; p)dp$. This Bayesian approach is introduced in [1], and provides a simple invariant for comparing different rooted trees. For a rooted tree on n vertices, $H_n \le \int_0^1 EV_b(T; p)dp \le n/2$, where H_n is the n^{th} harmonic number, and these bounds are sharp: The lower bound is achieved by a rooted path and the upper bound is achieved by a rooted star (Proposition 4.2 of [1]).

Section 5.3 of [1] asks for the expected value of $\int_0^1 EV_b(T; p)dp$ when T is a random tree. We answer this question now.

Theorem 2.8. *For fixed p, let $B_n(p)$ be the expected value of $1 + EV_b(T; p)$. Then $B_n(p) = \Theta(n^p)$ and $\int_0^1 B_n(p)dp = \Theta(n/\log n)$.*

Proof. From the previous theorem, we have $B_n(p) = \prod_{k=1}^{n-1}(1+p/k)$. Taking the logarithm of both sides, this becomes

$$\log B_n = \sum_{k=1}^{n-1} \log(1+\frac{p}{k})$$

We bound $\log B_n$ with Taylor expansions and simplify:

$$\begin{aligned}
\sum_{k=1}^{n-1}\left[\frac{p}{k} - \frac{(p/k)^2}{2}\right] &\leq \log B_n \leq \sum_{k=1}^{n-1}\frac{p}{k} \\
\Rightarrow p\sum_{k=1}^{n-1}\frac{1}{k} - \frac{p^2}{2}\sum_{k=1}^{n-1}\frac{1}{k^2} &\leq \log B_n \leq p\sum_{k=1}^{n-1}\frac{1}{k} \\
\Rightarrow pH_{n-1} - \frac{p^2}{2}\sum_{k=1}^{n-1}\frac{1}{k^2} &\leq \log B_n \leq pH_{n-1}
\end{aligned}$$

where $H_n = \sum_{k=1}^{n}\frac{1}{k}$ is the n^{th} harmonic number.

Now, since $\sum_{k=1}^{n-1} 1/k^2 = O(1)$ and $\log n \leq H_n \leq \log n+1$, there exists a constant c independent of n such that

$$-c + p\log(n-1) \leq \log B_n(p) \leq c + p\log(n-1)$$

Exponentiating yields

$$e^{-c}(n-1)^p \leq B_n(p) \leq e^c(n-1)^p$$

and $B_n(p) = \Theta(n^p)$. Finally, integrating gives $\int_0^1 B_n(p)dp = \Theta(n/\log n)$. □

2.2. **Pruning rank.** The definition of $EV_b(T;p)$ depends on the branching rank function r_b. In this section, we use the *pruning* rank function r_p to define a new expected rank polynomial $EV_p(T;p)$ for rooted trees. Let T be a rooted tree with edge-set E and let $S \subseteq E$. Then define the pruning rank $r_p(S)$ by

$$r_p(S) = \max_{A\subseteq S}\{|A| : E - A \text{ is a rooted subtree}\}.$$

Equivalently, we can compute the pruning rank $r_p(S)$ by successively removing leaf-edges from S, keeping in mind that edges of S which are not leaves of T may become leaves during this process, and using the convention that the root is never considered a leaf.

We denote the polynomial derived from this pruning rank by $EV_p(T;p)$. For $e \in E$ incident to a vertex v, let $c_e(v)$ be the number of vertices in the component of $T-e$ that contains v. The next result is analogous to Proposition 2.5. The proof follows from Proposition 3.1 of [9], which uses linearity of expected value to express the polynomial as a sum over E.

Proposition 2.9. *Let T be a rooted tree, and let vertices u and v be incident to the edge e with v further from the root than u. Then*

$$EV_p(T;p) = \sum_{e\in E} p^{c_e(v)}.$$

What is the expected value polynomial of a random rooted tree in this context? The next theorem, which is analogous to Theorem 2.6, computes this as another average of averages. We let $P_n(p)$ denote the expected value of $EV_p(T;p)$, where T ranges over all rooted trees on n vertices.

Theorem 2.10. *Let T be a random preorder-labeled tree with n vertices. Then the expected value of $EV_p(T;p)$ is $P_n(p) = \sum_{k=1}^{n-1} \frac{n}{k(k+1)} p^k$.*

Proof. We again interpret $EV_p(T;p) = n_1 p + n_2 p^2 + \dots$ as a generating function. Here, n_k is the number of vertices other than the root with k (improper) descendants, since we can identify each edge with the lower of its endpoints and think of pruning vertices (this follows from Proposition 2.9).

Now write $P_n(p) = a_{n,1}p + a_{n,2}p^2 + \dots$ Then the coefficient $a_{n,k}$ is the expected number of vertices with k descendants. In order to compute $a_{n,k}$, we develop and solve a binomial-like recurrence.

We first find a formula for $a_{n,1}$, the expected number of leaves. When adding a new vertex to a tree with $n-1$ vertices, each leaf in the subtree will cease to be a leaf with probability $1/(n-1)$; the new vertex will always be a leaf. This yields the following recurrence:

$$a_{1,1} = 1 \qquad\qquad a_{n,1} = \left(1 - \frac{1}{n-1}\right) a_{n-1,1} + 1$$

Multiplying through by $n-1$ yields:

$$0a_{1,1} = 0 \qquad\qquad (n-1)a_{n,1} = (n-2)a_{n-1,1} + (n-1)$$

Hence $(n-1)a_{n,1} = n(n-1)/2$; thus, for all $n > 1$, $a_{n,1} = n/2$

We next compute $a_{n,n-1}$ for $n > 1$. Only vertex 2 can have $n-1$ descendants, and it has this property exactly when the root, vertex 1, has only one child. Thus $a_{n,n-1}$ is the probability that vertex 2 is the only vertex with 1 as its parent. In adding a vertex to an $(n-1)$-vertex tree, the root will be chosen as a parent with probability $1/(n-1)$, so

$$a_{2,1} = 1 \qquad\qquad a_{n,n-1} = \left(1 - \frac{1}{n-1}\right) a_{n-1,n-2}$$

This gives $a_{n,n-1} = 1/(n-1)$.

Finally, when $n-1 > k > 1$, we can obtain a formula for $a_{n,k}$ as follows. In adding a vertex to an $(n-1)$-vertex preorder-labeled tree, the existing vertices with k descendants will continue to have k descendants with probability $1 - k/(n-1)$. The existing vertices with $k-1$ descendants will gain another descendant with probability $(k-1)/(n-1)$. Thus, $a_{n,k}$ satisfies the following formulas:

$$a_{n,1} = \frac{n}{2} \qquad\qquad a_{n,n-1} = \frac{1}{n-1}$$
$$a_{n,k} = \left(1 - \frac{k}{n-1}\right) a_{n-1,k} + \frac{k-1}{n-1} a_{n-1,k-1}$$

It is not hard to verify that the solution is $a_{n,k} = \frac{n}{k(k+1)}$.

□

2.3. **Unrooted trees.** When T is an unrooted tree, it is possible to define a rank function on subsets of edges of T based on edge pruning. This rank function endows the tree with an *antimatroid* structure; see [9] for a study of the expected rank polynomial in this context.

Let T be an unrooted tree. Then we define the expected rank polynomial $EV_{ur}(T;p)$ as before:

$$EV_{ur}(T;p) = \sum_{S \subseteq E} r_p(S) p^{|S|}(1-p)^{|E-S|},$$

where $r_p(S) = \max_{A \subseteq S}\{|A| : E - A \text{ is a subtree}\}$.

We will also need the following result, which is similar to Proposition 2.9. For an edge e of the unrooted tree T that is incident to vertex v, recall that $c_e(v)$ equals the number of vertices in the component of $T - e$ containing the vertex v.

Proposition 2.11 (Corollary 3.6 of [1])**.** *Let T be an unrooted tree with edges $E(T)$, where $|E(T)| = n$, and assume the edge e is incident to the vertices u and v. Then*

$$EV_{ur}(T;p) = \sum_{e \in E(T)} \left(p^{c_e(u)} + p^{c_e(v)} - p^n \right).$$

If T has n edges, then we have extreme values for $EV_{ur}(T;p)$ of np when T is a star, and $2p + 2p^2 + \cdots + 2p^{n-1} - (n-2)p^n$ when T is a path. As with the pruning rooted case, let $P_n(p)$ be the expected value of $EV_{ur}(T;p)$ when T has n vertices.

Proposition 2.12. *When $p > 0$ is fixed, $P_n(p) = \Theta(n)$.*

Proof. Note that the formula for $P_n(p)$ from Theorem 2.10 provides a lower bound for $P_n(p)$ in this context, since the pruning rank of any subset $S \subseteq E$ in a rooted tree can only increase when the root is ignored and we consider the pruning rank of the corresponding unrooted tree. Thus, $P_n(p)$ is bounded below by $(n/2)p$, and (trivially) above by $n-1$.

□

An interpretation for this result is that, for a fixed value of p, we expect to be able to prune a positive fraction of the edges from a random tree, independent of the size of the tree. The fraction depends only on the probability p of edge success.

The close relation between the pruning rank functions for rooted and unrooted trees explains the similarity between the formulas of Propositions 2.9 and 2.11. We exploit that connection in the next proposition. For a given (unrooted) tree T, let $f(T;p)$ be the expected value of the rooted pruning polynomial $EV_p(T;p)$, where this average is computed over the collection of all rooted trees obtained from T by placing the root at each vertex of T, in turn.

Proposition 2.13. *Let T be a tree with n edges. Then*

$$EV_{ur}(T;p) = f(T;p) + p^{n+1} f(T;p^{-1}) - np^n.$$

Proof. Let e be an edge of the unrooted tree T. By Proposition 2.11, e contributes $p^{c_e(u)} + p^{c_e(v)} - p^n$ to $EV_{ur}(T;p)$.

By Proposition 2.9, for a rooted tree obtained from T by placing the root at some vertex of T, e will contribute $p^{c_e(u)}$ or $p^{c_e(v)}$ to $EV_p(T;p)$, resp., depending on whether the root is closer to v or u, resp. Thus, the edge e contributes the term $\left(c_e(u)p^{c_e(v)} + c_e(v)p^{c_e(u)}\right)/(n+1)$ to $f(T;p)$. Since $c_e(u) + c_e(v) = n+1$

for any edge e, we also get a contribution of $\left(c_e(u)p^{c_e(u)} + c_e(v)p^{c_e(v)}\right)/(n+1)$ to $p^{n+1}f(T; p^{-1})$.

Thus, the total contribution an edge e makes to the RHS of the formula is $p^{c_e(u)} + p^{c_e(v)}$. Subtracting np^n and comparing the results with Propositions 2.9 and 2.11 completes the proof.

□

As an example of Proposition 2.13, let T be the unrooted tree of Figure 2. Let T_i (for $1 \le i \le 5$) denote the rooted tree obtained from T by placing the root at vertex i. Then $EV_{ur}(T; p) = 3p + p^2 + p^3 - p^4$, and

$$\begin{aligned} EV_p(T_1; p) &= EV_p(T_2; p) = 2p + p^2 + p^4 \\ EV_p(T_3; p) &= 3p + p^2 \\ EV_p(T_4; p) &= 3p + p^3 \\ EV_p(T_5; p) &= 2p + p^3 + p^4 \end{aligned}$$

The reader can check that $\left(\sum_{i=1}^{5} EV_p(T_i; p) + p^5 \sum_{i=1}^{5} EV_p(T_i; p^{-1})\right)/5 - 4p^4 = 3p + p^2 + p^3 - p^4 = EV_{ur}(T; p)$.

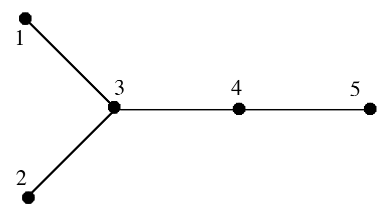

FIGURE 2. Tree for Proposition 2.13.

3. Rooted Digraphs

When D is a rooted directed graph, we can define an expected rank polynomial based on branching rank in the digraph D. $EV_b(D; p)$ still gives the expected number of vertices reachable from the root, where directions of edges must be respected.

If D is a rooted digraph obtained from a rooted graph by orienting the edges of G in some manner, then it is clear that $EV_b(D; p) \le EV_b(G; p)$ for all p (since any orientation can only decrease the probability that a given vertex is reachable). In this section, we answer the question of how much this expectation drops, on average.

We begin with a result that essentially allows edges of a rooted graph to fail in two stages.

Proposition 3.1. *Let G be a rooted graph and G' be a random graph derived from G by deleting edges with probability $1 - p'$. Then $EV_b(G'; p) = EV_b(G; pp')$.*

Proof. Let m be the number of edges in G. We induct on m. The case where $m = 0$ is immediate. The case $m > 0$ and G has no edges incident to the root is also immediate. Otherwise, let e be an edge incident to the root. With probability $1-p'$, e has been deleted in G' and $EV_b(G';p) = EV_b(G-e;pp')$. If e hasn't been deleted, we have

$$\begin{aligned} EV_b(G';p) &= (1+EV_b(G'/e;p))p + (1-p)EV_b(G'-e;p) \\ &= (1+EV_b(G/e;pp'))p + (1-p)EV_b(G-e;pp') \end{aligned}$$

where the final step follows by the induction hypothesis. Putting it all together,

$$\begin{aligned} EV_b(G';p) &= p'[p(1+EV_b(G/e;pp')) + (1-p)EV_b(G-e;pp')] \\ &\quad + (1-p')EV_b(G-e;pp') \\ &= pp'(1+EV_b(G/e;pp')) + (1-pp')EV_b(G-e;pp') \\ &= EV_b(G;pp') \end{aligned}$$

which is the desired result.

□

As an alternative to the inductive proof given, note that independence guarantees that each edge succeeds with probability pp'. The proposition then follows immediately. The inductive proof illustrates the utility of deletion-contraction arguments, however.

The next result concludes our treatment of randomness. We compute the average expected rank polynomial of a rooted directed graph obtained from a fixed rooted graph, averaged over all orientations of G.

Corollary 3.2. *Let D be a rooted digraph obtained from a rooted graph by randomly orienting the edges of G. Then the expected value of $EV_b(D;p)$ is given by $EV_b(G;p/2)$.*

Proof. By Proposition 2.2, we can express $EV_b(G;p)$ and $EV_b(D;p)$ using paths joining the root to a given vertex v. In G, suppose an edge e is used along the operational path joining $*$ to v. Then, on average, the same path will also use the edge e in a directed path joining $*$ to v in D with probability $1/2$, since e will be oriented 'correctly' along the path with that probability. Now applying Proposition 3.1 with $p' = 1/2$ gives the result. □

For example, let G be the rooted cycle C_3. Then $EV_b(G;p) = 2p + 2p^2 - 2p^3$. Now there are 8 orientations of C_3 which fall into 4 isomorphism classes, pictured in Figure 3. Then $EV_b(D_1;p) = 2p + p^2 - p^3$, $EV_b(D_2;p) = p + p^2$, $EV_b(D_3;p) = p$ and $EV_b(D_4;p) = 0$. Each digraph orientation of C_3 is represented by one of these D_i (for $1 \le i \le 4$), and each D_i corresponds to two orientations of G. Thus, the expected value of a random orientation of G is $2(2p + p^2 - p^3 + p + p^2 + p + 0)/8 = p + p^2/2 - p^3/4 = EV_b(G;p/2)$, which agrees with Corollary 3.2.

4. Grids

In this section, we concentrate on the optimal vertex location for grids. As usual, we assume that each edge of the rooted graph succeeds independently with probability p. Rooted trees are considered in [4] and [5]; the optimal location for the root generally depends on the value of p. In fact, the optimal location can switch arbitrarily often in general – see [10] and [8]. For certain graphs, however, the optimal location is independent of p. In particular, for grids, we can show that the 'obvious' locations for the root are indeed optimal for all values of p.

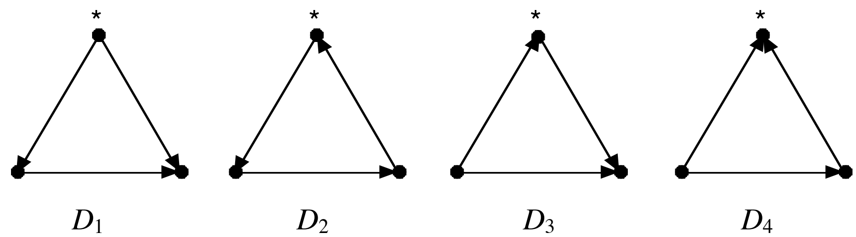

FIGURE 3. The possible orientations of C_3.

A *grid* graph $G_{m,n}$ is the graph whose vertices are ordered pairs of integers (a, b), where $1 \leq a \leq m$ and $1 \leq b \leq n$, with edges joining (a, b) to $(a+1, b)$ and $(a, b+1)$ (provided $a+1$ and $b+1$ are in the valid range). Grids are important in many applications, especially because of the way many cities are designed. Determining the Tutte polynomial of a grid using deletion-contraction is examined in [11]; the authors develop a complicated recursion, which they do not solve explicitly.

Our goal is more modest; we wish to prove the optimal location for the root (from the viewpoint of the expected value polynomial) is a central vertex in the grid, i.e., a vertex of minimum eccentricity. This is 'obvious,' but the proof relies on a few lemmas.

Let P and Q be horizontally adjacent vertices of the grid $G_{m,n}$. We partition the edges of $G_{m,n}$ into two parts; a grid $A = G_{m,d}$ that is symmetric with respect to vertices P and Q, and $B = G_{m,n} - G_{m,d}$.

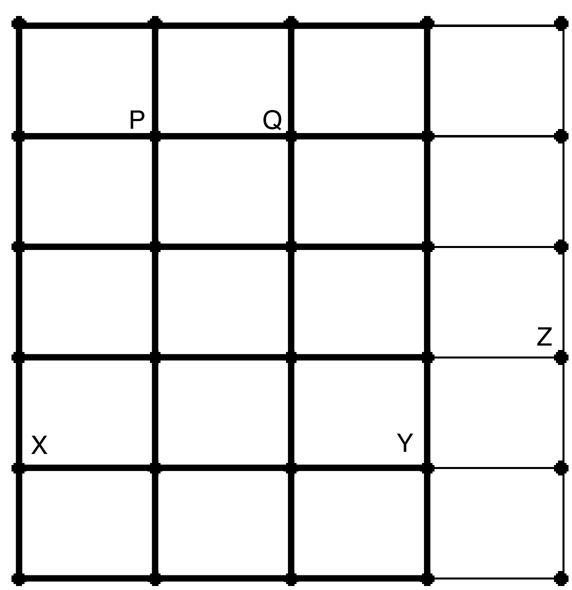

FIGURE 4. The grid $G_{5,6}$ with edges in A in bold.

Let X and Y be vertices of A which are symmetric with respect to the reflection that interchanges P and Q in A, and let Z be any vertex in the grid $G_{m,n}$ which is not incident to any edge of A (see Fig. 4). Then any path joining the vertices P and Z in $G_{m,n}$ must pass through the vertical line containing Q. This observation gives the next lemma. We write $Pr(X,Y)$ for the probability that there is an operational path joining vertices X and Y in the probabilistic graph G.

Lemma 4.1. *Suppose P and Q are adjacent in a grid $G_{m,n}$, with Q more central than P. Let A and B be the partition of the edges of $G_{m,n}$ as above, and let Z be any vertex in the grid $G_{m,n}$ which is not incident to any edge of A. Then $Pr(P,Z) < Pr(Q,Z)$.*

The next lemma considers the probability of reaching vertices in A from our two roots P and Q.

Lemma 4.2. *Suppose P and Q are adjacent in a grid $G_{m,n}$, with Q more central than P. Let A and B be the partition of the edges of $G_{m,n}$ as above, and let X and Y be vertices of A which are symmetric with respect to the reflection that interchanges P and Q in A. Then $Pr(P,X) + Pr(P,Y) < Pr(Q,X) + Pr(Q,Y)$.*

Proof. Let $Pr_A(C,D)$ denote the probability that vertices C and D can communicate using only edges of A, and let $Pr_B(C,D)$ be the probability that C and D communicate using some edges of B. Then $Pr(C,D) = Pr_A(C,D) + Pr_B(C,D)$.

Thus $Pr(Q,X)+Pr(Q,Y)-Pr(P,X)-Pr(P,Y) = Pr_A(Q,X)+Pr_A(Q,Y)-Pr_A(P,X)-Pr_A(P,Y)+Pr_B(Q,X)+Pr_B(Q,Y)-Pr_B(P,X)-Pr_B(P,Y)$. Now, by symmetry, we have $Pr_A(P,X) = Pr_A(Q,Y)$ and $Pr_A(P,Y) = Pr_A(Q,X)$. This gives $Pr(Q,X) + Pr(Q,Y) - Pr(P,X) - Pr(P,Y) = Pr_B(Q,X) + Pr_B(Q,Y) - Pr_B(P,X) - Pr_B(P,Y)$.

It remains to show $Pr_B(Q,X) > Pr_B(P,X)$ and $Pr_B(Q,Y) > Pr_B(P,Y)$ for all values of p between 0 and 1. Now any successful path joining vertices P and X which uses edges of B must reach a vertex Z not incident to any edge of A. By Lemma 4.1, $Pr(P,Z) < Pr(Q,Z)$ for all $0 < p < 1$. Since any path can be decomposed into that portion reaching such a vertex Z and the remainder of the path, we have $Pr_B(Q,X) > Pr_B(P,X)$. The same argument also shows $Pr_B(Q,Y) > Pr_B(P,Y)$, and this completes the proof.

□

The proof of the next lemma follows immediately from Proposition 2.2 and Lemmas 4.1 and 4.2. We write $EV_X(G_{m,n};p)$ for the expected value polynomial when the grid is rooted at the vertex X. (We also write EV instead of EV_b throughout this section – all of the calculations use branching rank.)

Lemma 4.3. *Let P and Q be adjacent vertices in $G_{m,n}$, with Q more central than P. Then $EV_P(G_{m,n};p) < EV_Q(G_{m,n};p)$ for all $0 < p < 1$.*

We can now use Lemma 4.3 repeatedly to prove our main result on grids. Note that $G_{m,n}$ may have a unique central vertex (if both m and n are odd), two central vertices (if precisely one of m and n is odd) or four central vertices (if both m and n are even).

Proposition 4.4. *Let $G_{m,n}$ be a grid with central vertex Q and non-central vertex P. Then $EV_P(G_{m,n};p) < EV_Q(G_{m,n};p)$ for all $0 < p < 1$.*

It should be possible to extend this argument to higher dimensional grid graphs; in particular, we again can partition the edges into a symmetric piece A and the rest of the edges B. Then lemmas analogous to 4.1, 4.2 and 4.3 should remain valid. We leave consideration of higher dimensions to the interested reader.

5. Vertex failure

The standard assumptions of reliability theory focus on edge failure. There are theoretical and practical reasons for this [6], but it is possible to incorporate vertex failure into our model of reliability. We present a brief overview here; many formulas are quite similar to the edge failure formulas developed in previous sections, and some formulas are much easier (e.g., complete graphs).

In this section, we assume edges are perfectly reliable, but the failure of a vertex means that no path from the root can pass successfully through that vertex. We also make the usual assumption about independence of vertex failure. (We assume the root vertex is always operational; if the root were allowed to fail, we could easily adjust all of the formulas that follow.)

To obtain a polynomial invariant for rooted graphs that is sensitive to vertex failure, we modify Definition 2.1. As in the case of edge failure, we use branching rank.

Definition 5.1. Let G be a rooted graph with root vertex $*$ and let $V^* = V(G) - \{*\}$ denote the rest of the vertices. Suppose that each $v \in V^*$ has an independent probability p of being operational. Then define

$$EV_v(G;p) = \sum_{S \subseteq V^*} r_b(S) p^{|S|}(1-p)^{|V^*|-|S|},$$

where $r_b(S)$ is the number of vertices in the same component as the root in the induced subgraph on S.

Linearity of expectation can be applied to indicator functions to give an expansion analogous to the one given in Proposition 2.2.

Proposition 5.2. *Let G be a rooted graph and let $Pr_v(v)$ be the probability that the root $*$ and v remain in the same component of G. Then*

$$EV_v(G;p) = \sum_{v \in V^*} Pr_v(v).$$

An immediate consequence of Proposition 5.2 is the following: If a rooted graph has n vertices and the root is adjacent to every vertex in G, then $EV_v(G;p) = (n-1)p$.

The next result is analogous to Proposition 2.3. We omit the straightforward proof.

Proposition 5.3. *Let G be a rooted graph with root vertex $*$ and let e be an edge with endpoints $*$ and v. Then*

$$EV_v(G;p) = (1-p) \cdot EV_v(G-v;p) + p \cdot EV_v(G/e;p) + p.$$

In each of the terms on the right-hand side of the deletion-contraction recursion, the number of vertices is reduced by 1. Note that if G/e has multiple edges, we can replace it with a simple graph since the edges are assumed to be perfectly reliable. Thus, for example, if $G = K_n$ is a complete graph, we can replace K_n/e by K_{n-1}. Then Proposition 5.3 can be applied to give inductive proofs of the following formulas.

Proposition 5.4. *Let T be a rooted tree with n edges, C_n a rooted n-cycle, and K_n a rooted complete graph on n vertices (including the root). Then*

(1) $EV_v(T;p) = \sum_{v \in V^*} p^{d(*,v)}$,
(2) $EV_v(C_n;p) = 2\sum_{k=1}^{n-2} p^k - (n-3)p^{n-1}$,
(3) $EV_v(K_n;p) = (n-1)p$.

Note that the formulas for rooted trees and cycles are virtually identical to the edge failure case. The formula for the complete graph is much simpler, though (see [2, 3]). Note that the formula for complete graphs follows immediately from the remark following Proposition 5.2.

Proposition 5.4 can be extended to give an explicit formula for complete mutipartite graphs.

Corollary 5.5. *Let $K_{n_1,\ldots,n_k}$ be a multipartite graph with vertex partition $\cup_{i=1}^k V_i$, where $|V_i| = n_i$ for all $1 \leq i \leq k$. Suppose the root vertex $* \in V_1$, and write $N = \sum_{i=1}^k n_i$. Then*

$$EV_v(K_{n_1,\ldots,n_k}) = (N-n_1)p + (n_1 - 1)\left(1 - (1-p)^{N-n_1}\right).$$

Proof. If $v \notin V_1$, then v is adjacent to $*$, so $Pr_v(v) = p$. If $v \in V_1$, then v will be connected to $*$ provided one of the vertices adjacent to $*$ is operational. The formula follows.

□

Note that the formula from Corollary 5.5 reduces to $(n-1)p$ when $n_1 = 1$, i.e., when $*$ is adjacent to every other vertex.

We conclude with a combinatorial result. The derivative of $EV_v(G;p)$ encodes information about the graph (although not as much as the edge-failure polynomial - see [10] or [8]).

Proposition 5.6. *If G is 2-connected, then $EV_v(G;1)' = |V| - 1$.*

The proof of this proposition follows from differentiating the deletion-contraction formula of Proposition 5.3 and using induction. We leave the details to the reader.

As an example, from Proposition 5.4, we have $EV_v(K_n;p) = (n-1)p$, so $EV_v(K_n;1)' = n-1$. For rooted cycles, we get $EV_v(C_n;1)' = 2(1+2+\cdots+(n-2)) - (n-1)(n-3) = n-1$. The necessity of the 2-connectedness condition can be seen, for example, by considering a rooted path, where the result is false.

References

[1] M. Aivaliotis, G. Gordon and W. Graveman, *When bad things happen to good trees*, J. Graph Theory **37** (2001), 79-99.

[2] A. Amin, K. Siegrist and P. Slater, *Pair-connected reliability of a tree and its distance degree sequences*, Cong. Numer. **58** (1987), 29-42.

[3] A. Amin, K. Siegrist and P. Slater, *Exact formulas for reliability measures for various classes of graphs*, Cong. Numer. **58** (1987), 43-52.

[4] A. Bailey, G. Gordon, M. Patton, J. Scancella, *Expected value expansions in rooted graphs*, Discrete Appl. Math. **128** (2003), 555-571.

[5] S. Chaudhary and G. Gordon, *Tutte polynomials for trees*, J. Graph Theory **15** (1991) 317-331.

[6] C. Colbourn, *The combinatorics of network reliability*, Oxford University Press, Oxford, 1987.

[7] C. Colbourn, *Network resilience*, SIAM J. Alg. Disc. Meth. **8** (1987), 404-409.

[8] D. Eisenstat, G. Gordon and A. Redlich, *Combinatorial properties of a rooted graph polynomial*, submitted.

[9] G. Gordon, *Expected rank in antimatroids,* Adv. in Applied Math. **32** (2004), 299-318.
[10] G. Gordon and E. Jager, *Roots for rooted graph polynomials*, (2003) preprint.
[11] M.E. Gegúndez, A. Márquez, and P. Revuelta, *The Tutte Polynomial on the square lattice*, (2001) preprint.

Dept. of Computer Science, Princeton University, 35 Olden Street, Princeton, NJ 08540
Email address: deisenst@cs.princeton.edu

Dept. of Mathematics, Washington Univ., St. Louis, MO 63130
Email address: jmfeder@gmail.com

Dept. of Mathematics 301 Thackeray Hall University of Pittsburgh Pittsburgh, PA 15260
Email address: gpf3@pitt.edu

Dept. of Mathematics, Lafayette College, Easton, PA 18042
Email address: gordong@lafayette.edu

Dept. of Mathematics, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA 02139
Email address: aredlich@uchicago.edu