USING MARKOV CHAINS TO DETERMINE EXPECTED PROPAGATION TIME FOR PROBABILISTIC ZERO FORCING*

YU CHAN[†], EMELIE CURL[†], JESSE GENESON[†], LESLIE HOGBEN[†‡], KEVIN LIU[†], ISSAC ODEGARD[†], AND MICHAEL ROSS[†]

Abstract. Zero forcing is a coloring game played on a graph where each vertex is initially colored blue or white and the goal is to color all the vertices blue by repeated use of a (deterministic) color change rule starting with as few blue vertices as possible. Probabilistic zero forcing yields a discrete dynamical system governed by a Markov chain. Since in a connected graph any one vertex can eventually color the entire graph blue using probabilistic zero forcing, the expected time to do this is studied. Given a Markov transition matrix for a probabilistic zero forcing process, an exact formula is established for expected propagation time. Markov chains are applied to determine bounds on expected propagation time for various families of graphs.

1. Introduction. A *graph*, which can be used to model relationships between objects, is a pair $G = (V, E)$. The set $E = E(G)$ of *edges* (relationships) consists of 2-element subsets of the set $V = V(G)$ of *vertices* (objects). Two vertices v, w are *adjacent* if $\{v, w\} \in E$. Suppose a graph G is colored so that every vertex is blue or white. Vertices in the graph can change color based on the *zero forcing color change rule*: If a blue vertex v is adjacent to exactly one white vertex w, then the white vertex changes to blue. In this case, we say that v *forces* w and denote this by $v \to w$. A set of vertices S is called a *zero forcing set* if when the vertices in S are colored blue and those in $V \setminus S$ are colored white, repeated application of the color change rule forces all of the vertices to be blue. The *zero forcing number* of a graph G, denoted $Z(G)$, is the minimum cardinality of a zero forcing set [1]. Throughout this paper, a force performed using the zero forcing color change rule is called a *deterministic force*.

Zero forcing was introduced in the study of the control of quantum systems by mathematical physicists who called it the "graph infection number" [3, 4]. Zero forcing was also introduced independently in the study of the minimum rank problem in combinatorial matrix theory to bound the maximum nullity [1]. Zero forcing and its positive semidefinite variant have been used extensively in the study of the minimum rank problem (see [11] and the references therein). Parameters derived from zero forcing have also been studied. Examples include propagation time (e.g., [13, 16]) and throttling (e.g., [5]). Zero forcing also has connections to graph searching [17] and power domination [2].

Two vertices are called *neighbors* if they are adjacent, and the set of neighbors of a vertex v in G is denoted by $N(v)$. The *closed neighborhood* of a vertex v is $N[v] = N(v) \cup \{v\}$. A variant of zero forcing called *probabilistic zero forcing* was introduced by Kang and Yi [15] and is defined as follows: In one *round*, each blue vertex u attempts to *force* (change the color to blue) each of its white neighbors w independently

---

with probability

$$\mathbf{Pr}(u \to w) = \frac{|N[u] \cap B|}{\deg u},$$

where B denotes the set of blue vertices. Because a vertex u attempts to force each of its white neighbors independently, this action is a binomial (or Bernoulli) experiment with probability of success given by the previous formula. This color change rule is known as the *probabilistic color change rule*, and *probabilistic zero forcing* refers to the process of coloring a graph blue by repeated application of the probabilistic color change rule.

The study of probabilistic zero forcing therefore produces a discrete dynamical system that plausibly describes many real world applications. Some of these applications include modeling the spread of a rumor through a social network, the spread of an infectious disease in a population, or the dissemination of a computer virus in a network. In addition, this type of zero forcing offers a new approach to coloring a graph. It should be noted that while for traditional zero forcing, the parameter of primary interest is the minimum number of vertices required to force the entire graph blue, in probabilistic zero forcing one blue vertex per connected component is necessary and sufficient to eventually color an entire graph blue. Therefore, finding a minimum probabilistic zero forcing set is not an interesting problem. However, there are parameters related to probabilistic zero forcing that are of interest.

One such parameter is expected propagation time, which is the focus of this paper. Suppose that G is a connected graph with the vertices in $B \neq \emptyset$ colored blue and all other vertices white. The *probabilistic propagation time* of B, denoted by $\mathrm{pt}_{pzf}(G, B)$, is defined as the random variable equal to the number of the round in which the last white vertex turns blue when applying the probabilistic color change rule [12]. For a connected graph G and a set $B \subseteq V(G)$ of vertices, the *expected propagation time of B* is the expected value of the propagation time of B [12], i.e.,

$$\mathrm{ept}(G, B) = \mathbf{E}[\mathrm{pt}_{pzf}(G, B)].$$

The *expected propagation time* of a connected graph G is the minimum of the expected propagation time of B over all one-vertex sets B of G [12], i.e.,

$$\mathrm{ept}(G) = \min\{\mathrm{ept}(G, \{v\}) : v \in V(G)\}.$$

The use of Markov chains for probabilistic zero forcing was introduced in [15] and studied further in [12]. If M is the $s \times s$ Markov matrix where the first state is one blue vertex and the last state is all vertices blue, then the probability that all vertices are blue after round r is the $(1, s)$-entry of M^r, i.e., $(M^r)_{1s}$. Thus, the probability that the propagation time is r is $\left(M^r - M^{r-1}\right)_{1s}$, and the expected propagation time is [12]

$$\mathrm{ept}(G, B) = \sum_{r=1}^{\infty} r\left(M^r - M^{r-1}\right)_{1s}.$$

In Section 2, we provide an exact method to calculate $\mathrm{ept}(G, B)$. We apply this result to examine the effect of various graph operations, including showing that is possible for each of the common graph operations vertex deletion, edge deletion, edge subdivision, and edge contraction to raise or lower the expected propagation time. In particular, we exhibit arbitrarily large graphs, the tadpole graphs $T_{4,m}$, for which adding an edge increases the expected propagation time, thereby answering a question in [12]; for each tadpole and tadpole with extra edge we obtain an exact formula for the expected propagation time in terms of the order.

This section also includes a characterization of the Markov matrices for the complete graph K_n on n vertices and information on graphs that exhibit interesting properties of the expected propagation time, such as n-Suns. Additional data on the expected propagation times of graphs discussed in Section 2, including all graphs that have at most four vertices and various families of graphs, is presented in the appendices [6, 7]. In particular, the *Sage* code listed in [6] offers the user the ability to compute the expected propagation time directly from the graph provided the graph is small enough that the user's computer can complete the computation (since the program is finding the Markov matrix, it is inefficient of computer time but efficient of human time).

In Section 3, we prove that $\text{ept}(K_n) = \Theta(\log \log n)$, improving the upper bound given in [12], and $\text{ept}(K_{c,n}) = \Theta(\log n)$, where $c \geq 1$ is a fixed integer and $K_{m,n}$ is a complete bipartite graph. We prove that $\text{ept}(G) = O(n)$ for any connected graph G on n vertices. Furthermore, we prove a $\Theta(\log n)$ bound on the expected propagation time of graphs on n vertices obtained by adding a universal vertex to a graph of bounded degree.

We define some additional terms from graph theory and notation that we will use throughout the paper. The *order* of a graph is the number of vertices. The *path* P_n of order n is a graph whose vertices can be listed in the order $v_1, \ldots, v_n$ such that the edges of the graph are $\{v_i, v_{i+1}\}$ for $i = 1, \ldots, n-1$. The *cycle* C_n of order n is a graph whose vertices can be listed in the order $v_1, \ldots, v_n$ such that the edges of the graph are $\{v_i, v_{i+1}\}$ for $i = 1, \ldots, n-1$ and $\{v_1, v_n\}$. The *complete graph* K_n is the graph of order n with all possible edges. The *complete bipartite graph* $K_{m,n}$ is the graph of order $m+n$ whose vertices can be divided into two parts $u_1, \ldots, u_m$ and $v_1, \ldots, v_n$ such that the edges of the graph are $\{u_i, v_j\}$ for $1 \leq i \leq m$ and $1 \leq j \leq n$. As a shorthand, we denote the edge $\{u, v\}$ as uv (since the graphs in the paper are not directed, the same edge could be written as vu). If v is a vertex in G, then $G - v$ denotes the graph obtained from G by removing the vertex v and all edges that contain v. If B is a set of blue vertices in G and v is a white vertex, we use $B \to v$ to denote that some vertex in B forces v.

2. Markov chains for probabilistic zero forcing. In this section, we introduce a method to compute expected propagation time exactly from the Markov transition matrix (see Theorem 2.2). In Section 2.1, we apply this theorem to determine the possible effects of various graph operations on expected propagation time, answering the question of whether adding an edge can raise expected propagation time (cf. [12, Question 2.16]). We then apply Markov chain methods to compute expected propagation time of various families of graphs in Section 2.2.

Let G be a graph and $B \subset V(G)$ be nonempty. A *simple state for B* is a coloring of the vertices that can be reached by starting with exactly the vertices in B blue, and then applying the probabilistic color change rule iteratively. We normally combine simple states that behave analogously into one *state for B*. For example, in K_n starting with one blue vertex, we use n states, with state k being the condition of having k blue vertices. In most graphs, it matters which vertices are blue, and this is reflected by distinguishing states with the same number of blue vertices but different behavior.

An *ordered state list for B*, denoted by $\mathcal{S} = (S_1, \ldots, S_s)$, is an ordered list of all states for B in which S_1 is the initial state (where exactly the vertices in B are blue), S_s is the final state (where all vertices are blue), and the states $S_k, k = 2, \ldots, s-1$ are in some chosen order. A graph G and an ordered state list $\mathcal{S}$ determine the Markov transition matrix for the process, which is denoted by $M(G, \mathcal{S})$. Reordering the states $S_2, \ldots, S_{s-1}$ results in a Markov transition matrix that is obtained by a permutation similarity of $M(G, \mathcal{S})$. We use $|S_k|$ to denote the number of blue vertices in state S_k, and say $\mathcal{S}$ is *properly ordered* if $|S_i| < |S_j|$

implies $i < j$.

PROPOSITION 2.1. *Let G be a graph and let $B \subset V(G)$ be nonempty. Let $\mathcal{S}$ be an ordered state list for B and let $M(G, \mathcal{S}) = [m_{ij}]$. Then $\mathrm{spec}(M(G, \mathcal{S})) = \{m_{kk} : k = 1, \ldots, s\}$, every eigenvalue is a real number in the interval $[0, 1]$, and 1 is a simple eigenvalue of $M(G, \mathcal{S})$. If $\mathcal{S}$ is a properly ordered state list for B, then $M(G, \mathcal{S})$ is upper triangular.*

Proof. Assume first that $\mathcal{S}$ is properly ordered. If $i \neq j$ and it is possible to go from S_i to S_j in one round, then $|S_i| < |S_j|$ so $i < j$. Thus, $M(G, \mathcal{S})$ is an upper triangular matrix and the eigenvalues are the diagonal entries. The probability m_{kk} of remaining in state S_k is less than one for $k < s$, is equal to one for $k = s$, and all m_{kk} are nonnegative. Thus, one is a simple eigenvalue and is the spectral radius of $M(G, \mathcal{S})$.

Note that a permutation similarity does not change the eigenvalues of $M(G, \mathcal{S})$ or the (unordered) multiset of diagonal entries (although the order of the diagonal entries may change). Thus, the statements about the spectrum are true without the assumption that $\mathcal{S}$ is properly ordered.   □

THEOREM 2.2. *Suppose that G is a graph, $B \subset V(G)$ is nonempty, $\mathcal{S}$ is an ordered state list for B with s states, and $M = M(G, \mathcal{S})$. Then*

$$\mathrm{ept}(G, B) = ((M - \mathbb{1}\mathbf{e}_s{}^T - I)^{-1})_{1s} + 1,$$

where $\mathbb{1} = [1, \ldots, 1]^T$ and $\mathbf{e}_s = [0, \ldots, 0, 1]^T$.

Proof. Define $\tilde{M} = M - \mathbb{1}\mathbf{e}_s{}^T$. Since $M\mathbb{1} = \mathbb{1}$ and $\mathbf{e}_s{}^T M = \mathbf{e}_s{}^T$, $\tilde{M}\mathbb{1} = \mathbf{0}$ and $\mathbf{e}_s^T \tilde{M} = \mathbf{0}^T$. An inductive argument shows that $M^k = \tilde{M}^k + \mathbb{1}\mathbf{e}_s{}^T$ for $k \geq 1$. Furthermore, the spectrum of $\tilde{M}$ is obtained from $\mathrm{spec}(M)$ by replacing eigenvalue 1 with 0 (subtracting $\mathbb{1}\mathbf{e}_s{}^T$ has the effect of deflating M on eigenvalue 1, as is done in the proof of [14, Theorem 8.2.7]). Recall that $\mathrm{ept}(G, B) = \sum_{r=1}^{\infty} r(M^r - M^{r-1})_{1s}$ [12], so we consider $\sum_{r=1}^{\ell} r(M^r - M^{r-1})$ as $\ell \to \infty$.

$$\sum_{r=1}^{\ell} r(M^r - M^{r-1}) = \tilde{M} + \mathbb{1}\mathbf{e}_s{}^T - I + \sum_{r=2}^{\ell} r\left(\tilde{M}^r + \mathbb{1}\mathbf{e}_s{}^T - \left(\tilde{M}^{r-1} + \mathbb{1}\mathbf{e}_s{}^T\right)\right)$$

$$= \mathbb{1}\mathbf{e}_s{}^T + \sum_{r=1}^{\ell} r\tilde{M}^r - \sum_{r=1}^{\ell-1} r\tilde{M}^r - \sum_{r=1}^{\ell-1} \tilde{M}^r - I$$

$$= \mathbb{1}\mathbf{e}_s{}^T + \ell\tilde{M}^\ell - \left(I + \tilde{M} + \cdots + \tilde{M}^{\ell-1}\right)$$

$$= \mathbb{1}\mathbf{e}_s{}^T + \ell\tilde{M}^\ell - (\tilde{M} - I)^{-1}(\tilde{M}^\ell - I).$$

Since the spectral radius is less than one, $\ell\tilde{M}^\ell \to 0$ and $(\tilde{M} - I)^{-1}\tilde{M}^\ell \to 0$ as $\ell \to \infty$. Thus,

$$\lim_{\ell \to \infty} \sum_{r=1}^{\ell} r(M^r - M^{r-1}) = (\tilde{M} - I)^{-1} + \mathbb{1}\mathbf{e}_s{}^T$$

and

$$\mathrm{ept}(G, B) = \left((M - \mathbb{1}\mathbf{e}_s{}^T - I)^{-1}\right)_{1s} + 1. \quad □$$

2.1. Graph operations. In this section, we present examples illustrating the possible effects of graph operations on expected propagation time. Some examples use results from Section 2.2; that section is independent of this one, and the use of such results in this section motivates the choice of some of the families for which expected propagation time is determined there.

Some of the examples presented here will use the values of expected propagation time of cycles and paths computed in [12], so we list them here: $\text{ept}(C_n) = \frac{n}{2} + \frac{1}{3}$ if n is even, $\text{ept}(C_n) = \frac{n}{2} + \frac{1}{2}$ if n is odd, $\text{ept}(P_n) = \frac{n}{2} + \frac{2}{3}$ if n is even, and $\text{ept}(P_n) = \frac{n}{2} + \frac{1}{2}$ if n is odd.

Vertex deletion. As might be expected, deleting a vertex (and all the edges incident with it) can lower the expected propagation time. One such example is deleting an endpoint of a path P_n with n even, because $\text{ept}(P_{n-1}) = \frac{n-1}{2} + \frac{1}{2} = \frac{n}{2}$ since $n-1$ is odd, whereas $\text{ept}(P_n) = \frac{n}{2} + \frac{2}{3}$. It follows from results in [12] that deleting a vertex can raise expected propagation time substantially: The underlying idea is that a universal vertex is a powerful tool for achieving low expected propagation time, and removing it can raise expected propagation time. For example, the *wheel* on n vertices, denoted by W_n, is obtained from C_{n-1} by adding a new vertex adjacent to every vertex in C_{n-1}. Since W_n has a universal vertex, $\text{ept}(W_n) = O(\log n)$ [12], whereas $\text{ept}(C_{n-1})$ is approximately $\frac{n}{2}$.

Edge deletion. As discussed in [12], there are many graphs for which deleting an edge can raise the expected propagation time, or equivalently, where adding an edge can lower the expected propagation time; one such example is deleting an edge from the cycle C_n to produce the path P_n for even $n \geq 4$, which raises the expected propagation time by $\frac{1}{3}$. It was asked in [12] whether deleting an edge can lower the expected propagation time. We applied Theorem 2.2 to determine the expected propagation time of all graphs of order four; this information is available in Appendix 1 [6]. Examination of the data provides an example where deleting an edge lowers the expected propagation time, as illustrated in Figure 2.1.
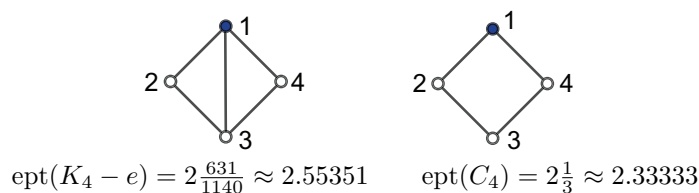

$$\text{ept}(K_4 - e) = 2\tfrac{631}{1140} \approx 2.55351 \qquad \text{ept}(C_4) = 2\tfrac{1}{3} \approx 2.33333$$

FIGURE 2.1. *A graph for which deleting an edge lowers the expected propagation time.*

This idea is generalized in Theorem 2.7, where expected propagation times are determined exactly for of an infinite family of graphs for which deleting an edge decreases expected propagation time. The *tadpole graph* $T_{4,m}$ is constructed from C_4 with vertices p_1, c_2, c_3, c_4 labeled cyclically and P_m with vertices $p_1, \ldots, p_m$ labeled in path order as $T_{4,m} = C_4 \cup P_m$. Form $T'_{4,m}$ by adding the edge $c_2 c_4$ to $T_{4,m}$. See Figure 2.2.
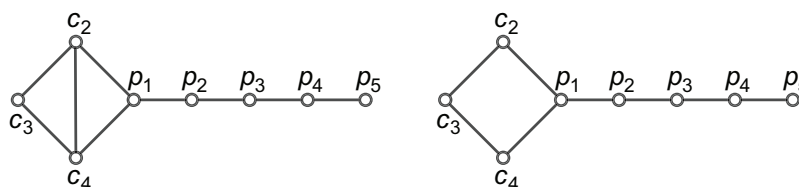

FIGURE 2.2. *The graphs $T'_{4,5}$ and $T_{4,5}$.*

Edge subdivision and contraction. If $e = uv$ is an edge of G, the *subdivision of e*, denoted by G_e, is the graph obtained from G by adding a new vertex w adjacent to u and v and deleting the edge uv. It is easy to find examples where subdividing an edge raises the expected propagation time. For example, subdividing an edge in a path P_n creates P_{n+1}, so raises the expected propagation time for odd n.

It is also possible for edge subdivision to lower expected propagation time. The *double star $DS(\ell, \ell')$* with $\ell, \ell' \geq 2$ is a tree of order $\ell + \ell' + 2$ that has two adjacent vertices u and u' of degrees $\ell + 1$ and $\ell' + 1$, called the *centers*, and every other vertex is a *leaf* (i.e., has degree one). The double star $DS(3, 3)$ and its subdivision $DS(3, 3)_{uu'}$ are shown in Figure 2.3.
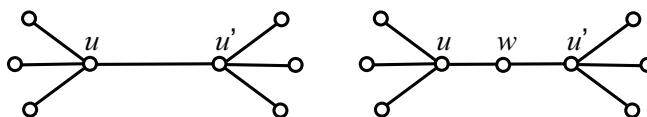

FIGURE 2.3. *The double star $DS(3, 3)$ and the subdivision $DS(3, 3)_{uu'}$.*

By constructing the Markov matrices of $DS(\ell, \ell)$ (starting with a center and starting with a leaf) and of $DS(\ell, \ell)_{uu'}$ (starting with the new vertex w), and then applying Theorem 2.2, we see that $\text{ept}(DS(3, 3)) = 4.71003$ and $\text{ept}(DS(3, 3)_{uu'}) = 4.36807$. For $2 \leq \ell \leq 9$, starting at a center does better for $DS(\ell, \ell)$, but for $\ell \geq 10$, starting at a leaf does better, and raising the number of leaves increases the difference between $\text{ept}(DS(\ell, \ell))$ and $\text{ept}(DS(\ell, \ell)_{uu'})$ – see Table 2.5 in Appendix 2 [7]. Intuitively, the increasing gap between $\text{ept}(DS(\ell, \ell))$ and $\text{ept}(DS(\ell, \ell)_{uu'})$ occurs because at least one set of leaves cannot start being colored until both centers are blue. In the original graph, the expected time until both centers are blue increases with the number of leaves, whereas in the subdivided graph starting with the new vertex w blue, the expected time until both centers is blue is two, i.e., the time to color a path on three vertices.

If $e = uv$ is an edge of G, the *contraction of G by e*, denoted by G/e, is the graph obtained from G by identifying vertices u and v and deleting any loops or duplicate copies of an edge that arise in this process. For $e = uv$ and new vertex w in G_e, contracting edge uw in G_e produces G. Thus, the fact that edge subdivision can raise or lower expected propagation time implies edge contraction can lower or raise expected propagation time.

The examples where $\text{ept}(G - v) - \text{ept}(G) \geq 0$, $\text{ept}(G) - \text{ept}(G_e) \geq 0$, and $\text{ept}(G/e) - \text{ept}(G) \geq 0$ suggest that these gaps grow arbitrarily large as the order of the graph increases, but this phenomenon is not present for other cases. This raises the questions of whether there exist constant upper bounds on $\text{ept}(G) - \text{ept}(G - v)$, $\text{ept}(G) - \text{ept}(G - e)$, $\text{ept}(G - e) - \text{ept}(G)$, $\text{ept}(G_e) - \text{ept}(G)$, and $\text{ept}(G) - \text{ept}(G/e)$; these are suitable subjects for future work.

2.2. Applications of Markov matrices to specific families. In this section, we determine Markov matrices for various families of graphs and then apply Theorem 2.2 to determine expected propagation times.

The complete graph. Let $K_n = (V, E)$ be the complete graph on n vertices. Let B be the set of currently blue vertices and let $b = |B| < n$. Consequently, the number of currently white vertices is equal to $n - b$. For any $v \in B$ and $w \in V \setminus B$, $\mathbf{Pr}(v \to w) = \frac{b}{n-1}$ and $\mathbf{Pr}(v \not\to w) = 1 - \frac{b}{n-1}$. At each given time step, for any given $w \in V \setminus B$, each $v \in B$ will independently attempt to force it. If at least one $v \in B$ is successful, then w is forced. So for any $w \in V \setminus B$ and any integer k such that $0 \leq k \leq n - b$,

$\mathbf{Pr}(\forall v \in B, v \not\to w) = \left(1 - \frac{b}{n-1}\right)^b$ and $\mathbf{Pr}(B \to w) = 1 - \left(1 - \frac{b}{n-1}\right)^b$. Thus, for $b < n - 1$,

$$(2.1) \quad \mathbf{Pr}(\text{exactly } k \text{ white vertices are forced}) = \binom{n-b}{k} \left(1 - \left(\frac{n-1-b}{n-1}\right)^b\right)^k \left(\left(\frac{n-1-b}{n-1}\right)^b\right)^{n-b-k}.$$

For $b = n - 1$, the process is deterministic (note that (2.1) remains valid with $0^0 = 1$). The next theorem follows from the previous statements and (2.1).

THEOREM 2.3. *Let* $\mathcal{S} = (S_1, \dots, S_n)$ *be the ordered state list where* S_k *is the state of having* k *blue vertices in* K_n. *The matrix* $M(K_n, \mathcal{S}) = [m_{ij}]$ *has*

$$m_{ij} = \begin{cases} \binom{n-i}{j-i} \left(1 - \left(\frac{n-1-i}{n-1}\right)^i\right)^{j-i} \left(\left(\frac{n-1-i}{n-1}\right)^i\right)^{n-j} & \text{if } i \leq \min(n-2, j), \\ 1 & \text{if } j = n \text{ and } i = n-1 \text{ or } n, \\ 0 & \text{if } i > j \text{ or } i = j = n-1. \end{cases}$$

Furthermore,

$$\text{spec}(M(K_n, \mathcal{S})) = \left\{0, 1, \left(\frac{n-1-i}{n-1}\right)^{i(n-i)} : i \in \{1, \dots, n-2\}\right\}.$$

The complete bipartite graph. Using a similar process, we can construct a Markov matrix for the complete bipartite graph $K_{m,n}$. Partition $K_{m,n}$ into its partite vertex sets R and R'. We denote each state (a, b), where a and b denote the number of blue vertices in R and R', respectively. In this state, a blue vertices independently attempt to force $n - b \geq 1$ white vertices, each with probability $\frac{b+1}{n}$, and b blue vertices independently attempt to force $m - a \geq 1$ white vertices, each with probability $\frac{a+1}{m}$.

PROPOSITION 2.4. *Given initial state* (a, b), *the probability of forcing exactly* k *vertices in* R *and* ℓ *vertices in* R' *is*

$$\binom{n-b}{\ell} \left(1 - \left(1 - \frac{b+1}{n}\right)^a\right)^\ell \left(1 - \frac{b+1}{n}\right)^{a(n-b-\ell)} \binom{m-a}{k} \left(1 - \left(1 - \frac{a+1}{m}\right)^b\right)^k \left(1 - \frac{a+1}{m}\right)^{b(m-a-k)}$$

where we define $0^0 = 1$.

For $n \geq 3$, we apply the previous proposition to determine the two possible Markov matrices for the star $K_{1,n-1}$, starting at the center (the vertex of degree greater than one) or at a leaf.

COROLLARY 2.5. *Let* $\mathcal{S} = (S_0, S_1, \dots, S_{n-1})$ *be an ordered state list for one vertex of* $K_{1,n-1}$ *where for* $k = 1, \dots, n-1$, S_k *is the state with* k *blue leaves and the center blue. When starting with a blue leaf,* S_0 *is the state with one blue leaf and the center white. When starting with the center blue,* S_0 *is the state with no blue leaves. The matrix* $M(T_n, \mathcal{S}) = [m_{ij}]$ *is a* $n \times n$ *matrix. Index the entries by* $0, \dots, n-1$ *with index* k *corresponding to state* S_k.

When starting with a blue center, the only nonzero entries of $M(K_{1,n-1}, \mathcal{S})$ *are as follows: For* $k = 0, \dots, n-3$ *and* $r = 0, \dots, n-1-k$,

$$m_{k,k+r} = \binom{n-1-k}{r} \left(\frac{k+1}{n-1}\right)^r \left(1 - \frac{k+1}{n-1}\right)^{n-1-k-r},$$

and $m_{n-2,n-1} = m_{n-1,n-1} = 1$.

When starting with a blue leaf, rows $1, \ldots, n-1$ of $M(K_{1,n-1}, \mathcal{S})$ are the same as when starting with the center, and the only nonzero entry in row 0 is $m_{0,1} = 1$.

Implementing Proposition 2.4 to construct the Markov matrix and then applying Theorem 2.2 in *Sage* code for various values of m and n yielded the data in Table 2.4 in Appendix 2 [7]. Examination of that data shows that $K_{2,3}$ is an outlier in the sense that $m = 2, n = 3$ is the only pair of values (up to $n = 10$) for which $\mathrm{ept}(K_{m,n})$ is not achieved by choosing a vertex in the larger partite set.

QUESTION 2.6. *Let $K_{m,n}$ have partite vertex sets R and R' of orders $m \leq n$ respectively, and let $u \in R$ and $v \in R'$. Is there any other pair $(m,n) \neq (2,3)$ such that $\mathrm{ept}(K_{m,n}, \{u\}) < \mathrm{ept}(K_{m,n}, \{v\})$?*

The tadpole graph and the tadpole with an extra edge. The tadpole graph $T_{4,5}$ and the tadpole with an extra edge $T'_{4,5}$ are illustrated in Figure 2.2. In the next theorem we compute the exact values of $\mathrm{ept}(T_{4,m})$ and $\mathrm{ept}(T'_{4,m})$ for $m \geq 5$. The only other infinite families of connected graphs for which the exact values of the expected propagation time are known are cycles and paths.

THEOREM 2.7. *For infinitely many positive integers n, there exist graphs on n vertices such that deleting an edge strictly decreases the expected propagation time. For $n \geq 5$, $\mathrm{ept}(T_{4,m}) = \frac{m-1}{2} + \frac{451}{216} = \frac{m-1}{2} + \frac{1353}{648}$ and $\mathrm{ept}(T'_{4,m}) = \frac{m-1}{2} + \frac{1429}{648}$ when m is odd, and $\mathrm{ept}(T_{4,m}) = \frac{m}{2} + \frac{3331}{1944} = \frac{m}{2} + \frac{9993}{5832}$ and $\mathrm{ept}(T'_{4,m}) = \frac{m}{2} + \frac{10357}{5832}$ when m is even.*

Proof. Suppose that $m \geq 5$. For ease of exposition, we assume that the path is horizontal and to the right of the cycle in $T_{4,m}$ and $T'_{4,m}$, as in Figure 2.2. First we note that $\mathrm{ept}(T_{4,2}, \{p_1, p_2\}) = \frac{17}{8}$, while $\mathrm{ept}(T'_{4,2}, \{p_1, p_2\}) = \frac{55}{24}$ (this can be verified by constructing Markov matrices and applying Theorem 2.2). We define the events E_0 and E_1 as follows: E_0 is the event that after the first force has occurred, in every round in which a non-deterministic force is attempted there is a successful non-deterministic force. E_1 is the event that after the first force has occurred, in every round but one in which a non-deterministic force is attempted there is a successful non-deterministic force. We break the proof into two cases depending on the parity of m.

Suppose that $m = 2k$ for some positive integer k: First we show that $\mathrm{ept}(T_{4,m}, \{p_k\}) = \frac{4}{3} + \frac{2}{3}(k-2) + \frac{1}{3}(k-1) + \frac{17}{8} = k + \frac{43}{24} = k + \frac{10449}{5832}$ and $\mathrm{ept}(T'_{4,m}, \{p_k\}) = \frac{4}{3} + \frac{2}{3}(k-2) + \frac{1}{3}(k-1) + \frac{55}{24} = k + \frac{47}{24} = k + \frac{11421}{5832}$. In each case the stated value is the expected time for the vertices to the left of p_k to turn blue, consisting of the expected time for the first force, the time after that to deterministically force p_1, and $\mathrm{ept}(T_{4,2}, \{p_1, p_2\})$ (respectively, $\mathrm{ept}(T'_{4,2}, \{p_1, p_2\})$). The vertices on the right can be ignored because once the first force happens, the time for the vertices to the right of p_k to turn blue is less than or equal to the least possible time for the last vertex on the left of p_k to turn blue.

Any vertex other than p_k and p_{k-1} has a vertex with distance at least $k+2$ from it in both $T_{4,m}$ and $T'_{4,m}$, which exceeds both $\mathrm{ept}(T_{4,m}, \{p_k\})$ and $\mathrm{ept}(T'_{4,m}, \{p_k\})$, so it suffices to compute $\mathrm{ept}(T_{4,m}, \{p_{k-1}\})$ and $\mathrm{ept}(T'_{4,m}, \{p_{k-1}\})$. We split into three cases depending on which vertices are forced in the round where the first force occurs, each of which has probability $\frac{1}{3}$.

For the first case, suppose that only p_k, i.e., the vertex to the right of p_{k-1} gets colored blue in the round with the first force. Then the propagation time for the vertices to the right of p_{k-1} is at most the propagation time for the vertices to the left of p_{k-1}, so the expected propagation time in this case is $\frac{4}{3} + (k-2) + \frac{17}{8} = k + \frac{35}{24}$ for $T_{4,m}$ and $\frac{4}{3} + (k-2) + \frac{55}{24} = k + \frac{39}{24}$ for $T'_{4,m}$.

For the second case, suppose that both p_k and p_{k-2} get colored blue on the first force. Then the

Y. Chan, E. Curl, J. Geneson, L. Hogben, K. Liu, I. Odegard, and M.S. Ross   326

propagation time for the vertices to the right of p_{k-1} is at most the propagation time for the vertices to the left of p_{k-1}, unless E_0 occurs, in which case the propagation time for the vertices to the left of p_{k-1} is one less than the propagation time for the vertices to the right of p_{k-1}. Since $\mathbf{Pr}(E_0)$ is $\frac{8}{9}$ for $T_{4,m}$ and $\frac{4}{9} + \frac{4}{9} \cdot \frac{2}{3} = \frac{20}{27}$ for $T'_{4,m}$, the expected propagation time in this case is $\frac{4}{3} + (k-3) + \frac{17}{8} + \frac{8}{9} = k + \frac{291}{216}$ for $T_{4,m}$ and $\frac{4}{3} + (k-3) + \frac{55}{24} + \frac{20}{27} = k + \frac{295}{216}$ for $T'_{4,m}$.

For the third case, suppose that only p_{k-2}, i.e., the vertex to the left of p_{k-1}, gets colored blue in the round with the first force. Then the propagation time for the vertices to the right of p_{k-1} is at most the propagation time for the vertices to the left of p_{k-1}, unless E_0 or E_1 occurs, in which case the propagation time for the vertices to the left of p_{k-1} is two or one less than the propagation time for the vertices to the right of p_{k-1}. In both $T_{4,m}$ and $T'_{4,m}$, $\mathbf{Pr}(E_0)$ is the same as in the last paragraph. Moreover, $\mathbf{Pr}(E_1)$ is $\frac{1}{9} \cdot \frac{8}{9}$ for $T_{4,m}$ and $\frac{1}{9}\left(\frac{20}{27}\right) + \frac{4}{9} \cdot \frac{1}{3}$ for $T'_{4,m}$. Thus, the expected propagation time in this case is $\frac{4}{3} + (k-3) + \frac{17}{8} + \frac{8}{9} \cdot 2 + \left(\frac{1}{9} \cdot \frac{8}{9}\right) \cdot 1 = k + \frac{1513}{648} = k + \frac{13617}{5832}$ for $T_{4,m}$ and $\frac{4}{3} + (k-3) + \frac{55}{24} + \frac{20}{27} \cdot 2 + \left(\frac{1}{9}\left(\frac{20}{27}\right) + \frac{4}{9} \cdot \frac{1}{3}\right) \cdot 1 = k + \frac{13629}{5832}$ for $T'_{4,m}$.

Observe that in each of the three cases, the expected propagation time for $T_{4,m}$ is less than the expected propagation time for $T'_{4,m}$. We determine $\mathrm{ept}(T_{4,m}, \{p_{k-1}\})$ and $\mathrm{ept}(T'_{4,m}, \{p_{k-1}\})$ by averaging over the cases: $\mathrm{ept}(T_{4,m}, \{p_{k-1}\}) = k + \frac{3331}{1944} = k + \frac{9993}{5832}$ and $\mathrm{ept}(T'_{4,m}, \{p_{k-1}\}) = k + \frac{10357}{5832}$. Therefore, $\mathrm{ept}(T_{4,m}) = k + \frac{3331}{1944}$ and $\mathrm{ept}(T'_{4,m}) = k + \frac{10357}{5832}$.

Suppose that $m = 2k + 1$ for some positive integer k: This proof is similar to the last proof. Again, we first calculate $\mathrm{ept}(T_{4,m}, \{p_k\})$ and $\mathrm{ept}(T'_{4,m}, \{p_k\})$. Like the proof for $m = 2k$ using p_{k-1}, we split the analysis into three cases depending on what happens in the round where the first force occurs. For both cases where the vertex to the right of p_k gets colored blue in the round with the first force, the propagation time for the vertices to the right of p_k is at most the propagation time for the vertices to the left of p_k. If both vertices adjacent to p_k are colored on the first force, the expected propagation time is $\frac{4}{3} + (k-2) + \frac{17}{8} = k + \frac{35}{24}$ for $T_{4,m}$ and $\frac{4}{3} + (k-2) + \frac{55}{24} = k + \frac{39}{24}$ for $T'_{4,m}$. If only the vertex to the right of p_k gets colored on the first successful force, the expected propagation time is $\frac{4}{3} + (k-1) + \frac{17}{8} = k + \frac{59}{24}$ for $T_{4,m}$ and $\frac{4}{3} + (k-1) + \frac{55}{24} = k + \frac{63}{24}$ for $T'_{4,m}$.

For the case where only the vertex to the left of p_k gets colored on the first successful force, the propagation time for the vertices to the right of p_k is at most the propagation time for the vertices to the left of p_k, unless E_0 occurs. Like the second case of the proof for $m = 2k$, $\mathbf{Pr}(E_0)$ is $\frac{8}{9}$ for $T_{4,m}$ and $\frac{4}{9} + \frac{4}{9} \cdot \frac{2}{3} = \frac{20}{27}$ for $T'_{4,m}$. Thus, the expected propagation time in this case is $\frac{4}{3} + (k-2) + \frac{17}{8} + \frac{8}{9} = k + \frac{169}{72} = k + \frac{507}{216}$ for $T_{4,m}$ and $\frac{4}{3} + (k-2) + \frac{55}{24} + \frac{20}{27} = k + \frac{511}{216}$ for $T'_{4,m}$.

Again, the expected propagation time for $T_{4,m}$ is less than the expected propagation time for $T'_{4,m}$ in each of the three cases, and we determine $\mathrm{ept}(T_{4,m}, \{p_k\})$ and $\mathrm{ept}(T'_{4,m}, \{p_k\})$ by averaging over the cases: $\mathrm{ept}(T_{4,m}, \{p_k\}) = k + \frac{451}{216} < k + 2.1$ and $\mathrm{ept}(T'_{4,m}, \{p_k\}) = k + \frac{1429}{648} < k + 2.21$. Any vertex besides p_k has a vertex with distance at least $k+2$ from it in both $T_{4,m}$ and $T'_{4,m}$, and the probability of failure on the first turn of the coloring process is at least $\frac{1}{4}$ except when p_m is the initial blue vertex, so $\mathrm{ept}(T_{4,m}, \{v\}) \geq k + 2.25$ and $\mathrm{ept}(T'_{4,m}, \{v\}) \geq k + 2.25$ for any $v \neq p_k$. Thus, $\mathrm{ept}(T_{4,m}) = k + \frac{451}{216} = k + \frac{1353}{648}$ and $\mathrm{ept}(T'_{4,m}) = k + \frac{1429}{648}$. $\square$

The sun and comb graphs. Let the n-Sun be obtained from the n-cycle C_n by adding a single leaf to each vertex. There are two distinct choices for the initial blue vertex, a vertex on the cycle or a leaf, and once forcing starts all blue vertices on the cycle are necessarily consecutive and the analysis is essentially the same. Finding the expected propagation time of the n-Sun is equivalent to finding the expected propagation time of the embedded cycle C_n, and then adding 1 to color all remaining leaves. Once the propagation

$$\begin{bmatrix} \text{IL} \\ \text{AS} \end{bmatrix}$$

process has produced at least two blue cycle vertices, an *outer cycle vertex* is a blue vertex on the cycle that has a white cycle neighbor, whereas a blue cycle vertex both of whose cycle neighbors are blue is called an *inner cycle vertex*. The leaf neighbor of an outer cycle vertex is called an *outer leaf*. The probability of forcing along the cycle is affected by whether each of the two outer leaves is colored blue; the leaf neighbors of inner cycle vertices have no effect on the cycle propagation. Thus, states are determined by the number of blue vertices in the cycle and how many of the outer leaves are blue.

There are two states involving one blue vertex on the cycle, without or with the adjacent leaf, which we denote 1 and $1L$, respectively. Next, denote the intermediate states (c, ℓ), where $2 \leq c \leq n-2$ and $0 \leq \ell \leq 2$. Here, c indicates the number of blue vertices on the cycle and ℓ indicates the number of outer leaves forced; notice that intermediate states with the same value of ℓ behave similarly to one another. Denote the last four states $(n-1, \ell)$ and (n), where $n-1$ or n of the cycle vertices are blue, respectively, and $0 \leq \ell \leq 2$. All possible outcomes and probabilities for these states starting with a blue vertex on the cycle are given in Table 2.1. Note that we leave out the fully propagated state. Instead, we add 1 to the propagation time found from the Markov matrix to account for the round needed to force all remaining leaves after reaching state (n).

State at time t	State at time $t+1$	Prob.	State at time t	State at time $t+1$	Prob.
1	1	$\frac{8}{27}$	$(c, 0)$	$(c+2, 0)$	$\frac{4}{9}$
1	1L	$\frac{4}{27}$	$(c, 1)$	$(c+1, 0)$	$\frac{1}{9}$
1	$(2, 0)$	$\frac{8}{27}$	$(c, 1)$	$(c+1, 1)$	$\frac{2}{9}$
1	$(2, 1)$	$\frac{4}{27}$	$(c, 1)$	$(c+2, 0)$	$\frac{2}{3}$
1	$(3, 0)$	$\frac{1}{9}$	$(c, 2)$	$(c+2, 0)$	1
1L	1L	$\frac{1}{9}$	$(n-1, 0)$	$(n-1, 0)$	$\frac{1}{81}$
1L	$(2, 1)$	$\frac{4}{9}$	$(n-1, 0)$	$(n-1, 1)$	$\frac{4}{81}$
1L	$(3, 0)$	$\frac{4}{9}$	$(n-1, 0)$	$(n-1, 2)$	$\frac{4}{81}$
$(c, 0)$	$(c, 0)$	$\frac{1}{81}$	$(n-1, 0)$	(n)	$\frac{8}{9}$
$(c, 0)$	$(c, 1)$	$\frac{4}{81}$	$(n-1, 1)$	(n)	1
$(c, 0)$	$(c, 2)$	$\frac{4}{81}$	$(n-1, 2)$	(n)	1
$(c, 0)$	$(c+1, 0)$	$\frac{4}{27}$	(n)	(n)	1
$(c, 0)$	$(c+1, 1)$	$\frac{8}{27}$			

TABLE 2.1
Transition probabilities for the ordered state list $\mathcal{S} = \{1, 1L, \ldots, (c, 0), (c, 1), (c, 2), \ldots, (n-1, 0), (n-1, 1), (n-1, 2), (n)\}$ of the n-Sun as defined above with the initial blue vertex on the cycle.

We can modify the above process for expected propagation time starting at a leaf rather than on the cycle. If the initial blue vertex is a leaf, the first step is deterministic, yielding state $1L$. Afterwards, the states and probabilities proceed as before. Thus, we simply need to construct the list of states starting at $1L$ instead of 1, and after finding the expected propagation time from the Markov matrix, add 2 to account for the first and last deterministic steps.

Using Theorem 2.2 and adding 1 for the final round, we can obtain exact values for $\mathrm{ept}(n\text{-Sun})$. Decimal approximations of these values with the initial vertex on the cycle are given in Table 2.2 (the *Sage* code that produced it can be found in [7]). This table also lists the differences in expected propagation time for consecutive n, i.e., $\Delta \mathrm{ept}(n\text{-Sun}) = \mathrm{ept}(n\text{-Sun}) - \mathrm{ept}((n-1)\text{-Sun})$. The clear trend that $\Delta \mathrm{ept}(n\text{-Sun}) \to 0.6875$ as n becomes large leads to the next conjecture.

n	ept(n-Sun)	Δ ept(n-Sun)	n	ept(n-Sun)	Δ ept(n-Sun)
5	4.77765692007797	0.729718323586744	25	18.5143540671558	0.687500000595474
6	5.44614700021659	0.668490080138619	26	19.2018540669176	0.687499999761808
7	6.14172265492263	0.695575654706038	27	19.8893540670129	0.687500000095277
8	6.82588757375988	0.684164918837255	28	20.5768540669748	0.687499999961890
9	7.51474489939839	0.688857325638504	29	21.2643540669900	0.687500000015245
10	8.20169679288223	0.686951893483841	30	21.9518540669839	0.687499999993904
11	8.88941718576886	0.687720392886632	31	22.6393540669863	0.687500000002437
12	9.57682877299639	0.687411587227531	32	23.3268540669854	0.687499999999023
13	10.2643641949093	0.687535421912946	33	24.0143540669858	0.687500000000391
14	10.9518500135211	0.687485818611719	34	24.7018540669856	0.687499999999844
15	11.6393556888815	0.687505675360446	35	25.3893540669857	0.687500000000064
16	12.3268534181140	0.687497729232458	36	26.0768540669856	0.687499999999975
17	13.0143543265595	0.687500908445543	37	26.7643540669856	0.687500000000011
18	13.7018539631505	0.687499636590999	38	27.4518540669856	0.687499999999996
19	14.3893541085209	0.687500145370441	39	28.1393540669856	0.687500000000000
20	15.0768540503712	0.687499941850303	40	28.8268540669856	0.687500000000000
21	15.7643540736315	0.687500023260217	41	29.5143540669856	0.687500000000000
22	16.4518540643273	0.687499990695837	42	30.2018540669856	0.687500000000000
23	17.1393540680490	0.687500003721681	43	30.8893540669856	0.687500000000000
24	17.8268540665603	0.687499998511324	44	31.5768540669856	0.687500000000000
25	18.5143540671558	0.687500000595474	45	32.2643540669856	0.687500000000000

TABLE 2.2
Expected propagation times for the n-Sun, and differences Δ ept(n-Sun) = ept(n-Sun) − ept(($n − 1$)-Sun) for $n = 5, \ldots, 45$.

CONJECTURE 2.8. $\lim_{n \to \infty} (\text{ept}(n\text{-Sun}) - \text{ept}((n-1)\text{-Sun})) = \frac{11}{16} = 0.6875$.

In general, starting with a blue leaf yields a greater expected propagation time than starting with a blue cycle vertex, although propagation starting at a leaf still suggests the aforementioned limit of $\frac{11}{16}$.

We can use a similar process to construct the Markov matrix for the n-Comb, which is obtained from the path P_n by adding a leaf to each vertex. As the initial blue vertex, choose $v = \lfloor \frac{n+1}{2} \rfloor$ on the embedded path, which is the center vertex for odd n and the left center vertex for even n. For the comb, we need to track both the number of vertices forced to the left and to the right of the initial vertex, along with whether or not the outer leaves are blue. The details, which are similar to the n-Sun but messier, are given in Appendix 2 [7], along with data.

3. Asymptotic bounds for probabilistic zero forcing. In this section, we prove asymptotically tight bounds up to a constant factor on several families of graphs, including some that were partially bounded in [12]. We prove that ept(K_n) = $\Theta(\log \log n)$. Next we generalize the bound ept($K_{1,n}$) = $\Theta(\log n)$ from [12] by proving that ept($K_{c,n}$) = $\Theta(\log n)$ for constant c, where the bound depends on c. Generalizing the same bound in a different direction, we show $\Theta(\log n)$ bounds on graphs obtained by adding a universal vertex to a graph of maximum degree at most c (a universal vertex is adjacent to every other vertex). Finally, we prove that ept(G) = $O(n)$ for all connected graphs G of order n.

Geneson and Hogben [12] proved that $\mathrm{ept}(K_n) = \Omega(\log \log n)$. In the next result, we show that bound is tight by proving that $\mathrm{ept}(K_n) = O(\log \log n)$. The method of proof is similar to that used in the proof in [12] that $\mathrm{ept}(K_{1,n}) = O(\log n)$.

THEOREM 3.1. *For positive integers n, $\mathrm{ept}(K_n) = \Theta(\log \log(n))$.*

Proof. Let K_n be the complete graph on n vertices for $n \geq 5$. Let b be the number of currently blue vertices and $w = n - b$ be the number of currently white vertices. For each white vertex $v_1, \ldots, v_w$, define the indicator random variable X_i to be 1 if v_i is colored blue in the current round and 0 otherwise, and define $X = \sum_{i=1}^{w} X_i$. Since the X_i's are i.i.d., we have that $\mathbf{E}[X] = w\mathbf{E}[X_i] = w \left(1 - \left(1 - \frac{b}{n-1} \right)^b \right)$ and $\mathbf{Var}[X] = w \left(1 - \left(1 - \frac{b}{n-1} \right)^b \right) \left(1 - \frac{b}{n-1} \right)^b$. Since $\left(1 - \left(1 - \frac{b}{n-1} \right)^b \right) \leq \left(1 - \left(1 - \frac{b^2}{n-1} \right) \right) = \frac{b^2}{n-1}$ by Bernoulli's inequality, $\mathbf{Var}[X] \leq \frac{w}{n-1} b^2 \left(1 - \frac{b}{n-1} \right)^b \leq b^2$.

For $1 \leq b \leq \frac{\sqrt{n}}{\log n}$, we first use binomial expansion on $\mathbf{E}[X]$ to obtain $\mathbf{E}[X] > \frac{wb^2}{n-1} - \sum_{k=1}^{\lfloor b/2 \rfloor} w \binom{b}{2k} \left(\frac{b}{n-1} \right)^{2k}$. For each term in the summation,

$$w \binom{b}{2k} \left(\frac{b}{n-1} \right)^{2k} \leq (n-1) \cdot \frac{b^{2k}}{(2k)!} \cdot \frac{b^{2k}}{(n-1)^{2k}} = \frac{b^2}{(2k)!} \cdot \frac{b^{4k-2}}{(n-1)^{2k-1}}.$$

Since $b = o(\sqrt{n})$, we conclude $b^{4k-2} = o(\sqrt{n}^{4k-2}) = o(n^{2k-1})$, and using this, we find

$$\frac{b^2}{(2k)!} \cdot \frac{b^{4k-2}}{(n-1)^{2k-1}} = \frac{b^2}{(2k)!} \cdot o(1) = \frac{o(b^2)}{(2k)!}.$$

Since $\sum_{k=1}^{\infty} \frac{1}{(2k)!}$ converges, this implies

$$\sum_{k=1}^{\lfloor b/2 \rfloor} w \binom{b}{2k} \left(\frac{b}{n-1} \right)^{2k} = \sum_{k=1}^{\lfloor b/2 \rfloor} \frac{o(b^2)}{(2k)!} = o(b^2).$$

For $b \leq \frac{\sqrt{n}}{\log n}$, we have $w \geq n - \frac{\sqrt{n}}{\log n}$, so $\frac{w}{n-1} = 1 - o(1)$. We conclude that

$$\mathbf{E}[X] > \frac{wb^2}{n-1} - \sum_{k=1}^{\lfloor b/2 \rfloor} w \binom{b}{2k} \left(\frac{b}{n-1} \right)^{2k} = b^2 - o(b^2).$$

Since $\mathbf{E}(X) = b^2 - o(b^2)$ and $b^2 = o(n)$, $\mathbf{E}(X) > \frac{5}{6} b^2$ for n sufficiently large. Thus, by Chebyshev's inequality,

$$\mathbf{Pr}(X < \frac{1}{2} b^2) \leq \mathbf{Pr}(|X - E(X)| > \frac{1}{3} b^2) \leq \frac{\mathbf{Var}(X)}{\left(\frac{1}{3} b^2 \right)^2} \leq \frac{9}{b^2} \leq 9/16$$

for $b \geq 4$. Therefore, there exists c such that the expected number of rounds to transition from b blue vertices to at least $\frac{1}{2} b^2$ blue vertices is at most c. To establish an upper bound on the expected number of rounds until there are at least $\frac{\sqrt{n}}{\log n}$ blue vertices, consider $f(x) = 2^{2^x+1}$, which satisfies $f(k+1) = \frac{1}{2} f(k)^2$. If $2^{2^r+1} = \frac{\sqrt{n}}{\log n}$, then $r = \log_2 \left(\log_2 \left(\frac{\sqrt{n}}{\log n} \right) - 1 \right)$. Since the expected time to transition from 1 to 4 blue vertices is bounded by a constant, the total expected time to transition from 1 to $\frac{\sqrt{n}}{\log n}$ blue vertices is at most $cr + O(1) = O(\log \log n)$.

For $\frac{\sqrt{n}}{\log n} \le b \le \sqrt{n} \log n$, Claim (C2) established in the proof of Lemma 2.5 in [12] implies $\mathbf{Pr}\left(X \ge \frac{b}{4}\right) = \Omega(1)$. Thus, there exists a constant D such that the expected number of rounds to transition from b blue vertices to at least $b + \frac{b}{4} = \frac{5}{4}b$ blue vertices is at most D. The expected total rounds to transition from $\frac{\sqrt{n}}{\log n}$ to $\sqrt{n} \log n$ blue vertices is at most Dr, where r is found by solving $\left(\frac{5}{4}\right)^r = \frac{\sqrt{n}\log n}{\sqrt{n}/\log n}$, which gives us $r = 2\log_{5/4}\log n$ and $Dr = O(\log\log n)$.

For $n \ge 5$, $\left(\frac{1}{n}\right)^{\log n} \le \frac{1}{n^{\log 5}} < \frac{1}{n^{1.5}}$. So, for $\sqrt{n}\log n \le b \le n - 2$,

$$\left(1 - \frac{b}{n-1}\right)^b \le \left(1 - \frac{\sqrt{n}\log n}{n}\right)^{\sqrt{n}\log n} < \left(e^{-\log n}\right)^{\log n} < \frac{1}{n^{1.5}}.$$

Note that X ranges from 0 to w, so $w - X$ is nonnegative. This allows us to apply Markov's inequality and linearity of expectation to show

$$\mathbf{Pr}(X < w - \sqrt{w}) = \mathbf{Pr}(w - X > \sqrt{w}) \le \frac{\mathbf{E}[w - X]}{\sqrt{w}} = \sqrt{w}\left(1 - \frac{b}{n-1}\right)^b < \sqrt{w} \cdot \frac{1}{n^{1.5}} < \frac{1}{n}.$$

For the complementary event, we conclude $\mathbf{Pr}(X \ge w - \sqrt{w}) \ge \frac{n-1}{n}$. Then the expected time to transition from w white vertices to at most $\sqrt{w}$ white vertices is at most $\frac{n}{n-1}$. Hence, the expected number of rounds to transition from $w = n - \sqrt{n}\log n$ to 2 white vertices is at most $\frac{n}{n-1} \cdot r$, where r is given by $w^{(1/2)^r} = 2$. Solving this equation, we find $r = \log_2\log_2 w$, implying that $\frac{n}{n-1} \cdot r = \frac{n}{n-1} \cdot \log_2\log_2 w = O(\log\log n)$. Note that for $w \le 2$, the expected time that remains is bounded by a constant. Thus, $\text{ept}(K_n) = \Theta(\log\log(n))$. $\square$

It is known that if a graph G of order n has a universal vertex, then $\text{ept}(G) = O(\log n)$ [12, Corollary 2.6]. In the next result, we use this fact to prove that $\text{ept}(G) = \Theta(\log n)$ for graphs G obtained by adding a universal vertex to a (not necessarily connected) graph of maximum degree at most c.

THEOREM 3.2. *Let c be a fixed positive integer and let $\mathcal{F}_c$ be the family of graphs having maximum degree at most c. Let G be a graph of order n with a universal vertex u such that $G - u \in \mathcal{F}_c$. Then $\text{ept}(G) = \Theta(\log n)$.*

Proof. The upper bound follows from [12, Corollary 2.6]. For the lower bound, we consider two cases, based on the the number $\hat{b}$ of blue vertices when u is colored blue at time t. First, suppose that $\hat{b} \ge \sqrt{n}$. Since the maximum degree is at most c, $\sqrt{n} \le \hat{b} \le 1 + c + c^2 + \cdots + c^t = \frac{c^{t+1}-1}{c-1}$. Thus, $\log_c(\sqrt{n}(c-1)+1) - 1 \le t$, and we have the desired lower bound.

If instead $\hat{b} < \sqrt{n}$, we consider the expected number of rounds to transition from at most $\sqrt{n}$ blue vertices to at least $\frac{n}{2}$ blue vertices. Let X be the random variable for the number of new blue vertices in the current round, and let $g(b) = \mathbf{Pr}(X \le 4b + cb)$, where b is the current number of blue vertices. We will show that $g(b) = 1 - O\left(\frac{1}{\sqrt{n}}\right)$ for $\sqrt{n} \le b \le \frac{n}{2}$. To this end, note that X is at most the sum of the number of vertices forced by u, which we will denote by s, plus the number of vertices forced by vertices other than u, which we will denote by r. Then, $Pr[s \ge 4b] = O\left(\frac{1}{\sqrt{n}}\right)$ by the proof of Theorem 2.7 in [12]. Because the maximum degree is at most c, we also have $r \le cb$. Thus, $1 - g(b) = O\left(\frac{1}{\sqrt{n}}\right)$. From this point, the same steps as in the proof of Theorem 2.7 in [12] show that with probability $1 - o(1)$, the number of rounds to go from at most $\sqrt{n}$ blue vertices to at least $\frac{n}{2}$ blue vertices is $\Omega(\log n)$ (with the constant dependent on c), so $\text{ept}(G) = \Omega(\log n)$. $\square$

The next result builds on ideas in [12].

THEOREM 3.3. *For any positive integers m and n, $\mathrm{ept}(K_{m,n}) = O(\log(m+n))$. For a fixed positive integer c, $\mathrm{ept}(K_{c,n}) = \Theta(\log(n))$.*

Proof. For the upper bounds: It was shown in [12, Lemma 2.5] that $\mathrm{ept}(G[N[v]]) = O(\log \deg v)$ for any vertex v. This implies $\mathrm{ept}(K_{m,n}) = O(\log(m) + \log(n))$. If $m \le n$, then $\log(m) + \log(n) \le 2\log n$, so $\mathrm{ept}(K_{m,n}) = O(\log(n))$, which also implies $\mathrm{ept}(K_{m,n}) = O(\log(m+n))$ (and no assumption $m \le n$ is needed on the latter).

Let c be a fixed positive integer. We consider the lower bound on $\mathrm{ept}(K_{c,n})$. Let R and R' denote the partite sets of orders c and n respectively. We show first that the expected number of rounds to color all vertices in R blue is $O(1)$. Suppose first that the one initial blue vertex is in R. By Claim $(C1)$ established in the proof of Lemma 2.5 in [12], the probability of at least one new blue vertex in a round is at least one half, so the expected time of the first force is at most 2. Once at least one vertex in R' is blue, the expected number of rounds to color R blue is at most $\mathrm{ept}(K_{1,c})$. Thus, the expected number of rounds to color R blue is a constant.

So suppose that all the vertices in R are blue and let b denote the current number of blue vertices. For each white vertex $v_1, \dots, v_{n+c-b} \in R'$, let X_i be the indicator random variable that v_i is colored blue in the current round. Let $X = \sum_{i=1}^{n+c-b} X_i$, and

$$\mathbf{Pr}(R \to v_i) = 1 - \mathbf{Pr}(\forall u \in R, u \nrightarrow v_i) = 1 - (1 - \mathbf{Pr}(u \to v_i))^c = 1 - \left(1 - \frac{1+b-c}{n}\right)^c.$$

Using Bernoulli's inequality for the first inequality below, we have

$$
\begin{aligned}
\mathbf{E}[X] &= \sum_{i=1}^{n+c-b} \mathbf{E}[X_i] \\
&= (n+c-b)\left(1 - \left(1 - \frac{1+b-c}{n}\right)^c\right) \\
&\le (n+c-b)\left(1 - \left(1 + c\left(-\frac{1+b-c}{n}\right)\right)\right) \\
&= \frac{(n+c-b)(1+b-c)c}{n} \\
&\le cb.
\end{aligned}
$$

Since the X_i are i.i.d. and $X_i^2 = X_i$,

$$\mathbf{Var}[X] = (n+c-b)\left(1 - \left(1 - \frac{1+b-c}{n}\right)^c\right)\left(1 - \frac{1+b-c}{n}\right)^c \le \mathbf{E}[X] \le cb.$$

Consider the case in which $\sqrt{n} \le b \le \frac{n}{2}$, and define $h(b)$ to be the probability that the number of new blue vertices in the current round is at most $2cb$. Then Chebyshev's inequality justifies the third inequality below:

$$1 - h(b) \le \mathbf{Pr}(X - cb \ge cb) \le \mathbf{Pr}(|X - \mathbf{E}[X]| \ge cb) \le \frac{\mathbf{Var}[X]}{(cb)^2} \le \frac{1}{c\sqrt{n}} = O\left(\frac{1}{c\sqrt{n}}\right).$$

Starting with $\sqrt{n} \le b \le \frac{n}{2}$ blue vertices and coloring at most $2cb$ additional blue vertices per round implies that the probability that there are at most $(3c)^r b$ blue vertices after r rounds is at least $(h(b))^r = \left(1 - O\left(\frac{1}{c\sqrt{n}}\right)\right)^r$. Thus, going from at most $\sqrt{n}$ blue vertices to at least $\frac{n}{2}$ blue vertices requires that

$(3c)^r \sqrt{n} \geq \frac{n}{2}$, or $r \geq \log_{3c}\left(\frac{\sqrt{n}}{2}\right)$. Hence, the probability is at least $\left(1 - O\left(\frac{1}{c\sqrt{n}}\right)\right)^{\log_{3c}\left(\sqrt{n}/2\right)} = 1 - o(1)$ that it takes at least $\log_{3c}\left(\frac{\sqrt{n}}{2}\right)$ rounds for the number of blue vertices to increase from at most $\sqrt{n}$ to at least $\frac{n}{2}$. So $\mathrm{ept}(K_{c,n}) = \Omega(\log(n))$. ∎

It is shown in [12] that $\mathrm{ept}(G) = O(\mathrm{rad}(G)(\log n)^2)$ for connected graphs G of order n, where $\mathrm{rad}(G) = \min_{u \in V(G)} \max_{v \in V(G)} \mathrm{dist}(u, v)$ is the *radius* of G. The next result implies that $\mathrm{ept}(G) = O(n)$ for connected graphs G of order n.

THEOREM 3.4. *Let G be a connected graph of order n. Then $\mathrm{ept}(G, S) \leq \frac{e}{e-1}(n - |S|)$ for any set S of vertices of G.*

Proof. We prove this by reverse strong induction on $k = |S|$. It is immediate for $k = n$. Now fix some $k < n$ and suppose that the theorem is true for any $i > k$. Let S be an initial set of blue vertices. Since G is connected, there exists some $b \in S$ with at least one white neighbor. Let $d = \deg(b)$, so $d - j + 1$ of the neighbors are white for some integer j with $1 \leq j \leq d$.

Suppose that there have been no forces yet in the graph. The probability that b does not force any of its white neighbors in the current round is at most

$$\left(1 - \frac{j}{d}\right)^{d-j+1} = \left(1 - \frac{j}{d}\right)^{(d/j)(j(d-j+1)/d)} \leq \frac{1}{e}^{j(d-j+1)/d} \leq \frac{1}{e},$$

where the first inequality follows from the fact that $(1 - \frac{1}{x})^x \leq \frac{1}{e}$ for $x \geq 1$ and the last inequality follows from the fact that $\frac{j(d-j+1)}{d}$ is minimized at $j = 1$ and $j = d$ for all real $j \in [1, d]$.

If there have not been any forces yet, the probability of a force in the current round is at least $\frac{e-1}{e}$, so the expected number of rounds until the first force is at most $c = \frac{e}{e-1}$. After the first force, there are at least $k + 1$ blue vertices. Therefore, $\mathrm{ept}(G, S) \leq c + c(n - k - 1) \leq c(n - k)$ by the induction hypothesis. ∎

COROLLARY 3.5. *If G is a connected graph on n vertices, then $\mathrm{ept}(G) = O(n)$.*

REFERENCES

[1] AIM Minimum Rank – Special Graphs Work Group (F. Barioli, W. Barrett, S. Butler, S.M. Cioaba, D. Cvetković, S.M. Fallat, C. Godsil, W. Haemers, L. Hogben, R. Mikkelson, S. Narayan, O. Pryporova, I. Sciriha, W. So, D. Stevanović, H. van der Holst, K. Vander Meulen, A. Wangsness). Zero forcing sets and the minimum rank of graphs. *Linear Algebra Appl.*, 428:1628–1648, 2008.

[2] K.F. Benson, D. Ferrero, M. Flagg, V. Furst, L. Hogben, V. Vasilevska, and B. Wissman. Zero forcing and power domination for graph products. *Australasian J . Combinatorics*, 70:221–235, 2018.

[3] D. Burgarth and V. Giovannetti. Full control by locally induced relaxation. *Phys. Rev. Lett.*, 99:100501, 2007.

[4] D. Burgarth and K. Maruyama. Indirect Hamiltonian identification through a small gateway. *New J. Phys.*, 11:103019, 2009.

[5] S. Butler and M. Young. Throttling zero forcing propagation speed on graphs. *Australasian J . Combinatorics*, 57:65–71, 2013.

[6] Y. Chan, E. Curl, J. Geneson, L. Hogben, K. Liu, I. Odegard, and M.S. Ross. Appendix 1: Expected propagation time for graphs of orders 4, 5, 6, and 7. Available at https://aimath.org/~hogben/Appendix1.pdf. *Sage* worksheet published at https://sage.math.iastate.edu/home/pub/118/.

[7] Y. Chan, E. Curl, J. Geneson, L. Hogben, K. Liu, I. Odegard, and M.S. Ross. Appendix 2: Expected propagation time data for various families of graphs. Available at https://aimath.org/~hogben/Appendix2.pdf. *Sage* worksheet for expected propagation time of complete graphs, published at https://sage.math.iastate.edu/home/pub/123/. *Sage* worksheet for expected propagation time of complete bipartite graphs, published at https://sage.math.iastate.edu/home/pub/122/. *Sage* worksheet for expected propagation time of double stars, published at https://sage.math.iastate.edu/home/pub/124/. *Sage* worksheet for expected propagation time of combs and suns, published at https://sage.math.iastate.edu/home/pub/119/.

[8] K. Chilakamarri, N. Dean, C.X. Kang, and E. Yi. Iteration index of a zero forcing set in a graph. *Bull. Inst. Combin. Appl.*, 64:57–72, 2012.

[9] R Diestel. *Graph Theory*, fourth edition. Graduate Texts in Mathematics, Springer, Heidelberg, 2010.

[10] S.M. Fallat and L. Hogben. The minimum rank of symmetric matrices described by a graph: A survey. *Linear Algebra Appl.*, 426:558–582, 2007.

[11] S. Fallat and L. Hogben. Minimum rank, maximum nullity, and zero forcing number of graphs. In: L. Hogben (editor), *Handbook of Linear Algebra*, second edition, CRC Press, Boca Raton, 2014.

[12] J. Geneson and L. Hogben. Propagation time for probabilistic zero forcing. Available at https://arxiv.org/abs/1812.10476.

[13] L. Hogben, M. Huynh, N. Kingsley, S. Meyer, S. Walker, and M. Young. Propagation time for zero forcing on a graph. *Discrete Appl. Math.*, 160:1994–2005, 2012.

[14] R.A. Horn and C.R. Johnson. *Matrix Analysis*, second edition. Cambridge University Press, New York, 2013.

[15] C.X. Kang and E. Yi. Probabilistic zero forcing in graphs. *Bull. Inst. Combin. Appl.*, 67:9–16, 2013.

[16] N. Warnberg. Positive semidefinite propagation time. *Discrete Appl. Math.*, 198:274–290, 2016.

[17] B. Yang. Fast-mixed searching and related problems on graphs. *Theoret. Comput. Sci.*, 507:100–113, 2013.