6-1-2022

Expected propagation time for probabilistic zero forcing

Jesse Geneson
San Jose State University, jesse.geneson@sjsu.edu

Leslie Hogben
Iowa State University

AUSTRALASIAN JOURNAL OF COMBINATORICS
Volume **83(3)** (2022), Pages 397–417

Expected propagation time for probabilistic zero forcing

Jesse Geneson*

Department of Mathematics and Statistics
San José State University
1 Washington Sq., San José, CA 95192
U.S.A.
geneson@gmail.com

Leslie Hogben†

Department of Mathematics
Iowa State University
Ames, IA 50011
U.S.A.
hogben@aimath.org

Abstract

Zero forcing is a coloring process on a graph that was introduced more than fifteen years ago in several different applications. The goal is to color all the vertices blue by repeated use of a (deterministic) color change rule. Probabilistic zero forcing was introduced by Kang and Yi in [*Bull. Inst. Combin. Appl.* 67 (2013), 9–16] and yields a discrete dynamical system, which is a better model for some applications. 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 a natural parameter to study. We determine expected propagation time exactly for paths and cycles, establish the asymptotic value for stars, and present asymptotic upper and lower bounds for any graph in terms of its radius and order. We apply these results to obtain values and bounds on ℓ-round probabilistic zero forcing and confidence levels for propagation time.

* Also at: Department of Mathematics, Iowa State University, Ames, IA 50011, U.S.A.
† Also at: American Institute of Mathematics, 600 E. Brokaw Road, San José, CA 95112, U.S.A.

ISSN: 2202-3518

1 Introduction

Zero forcing is a coloring process on a graph that was introduced independently in the study of control of quantum systems in mathematical physics [8] and the study of the maximum nullity problem in combinatorial matrix theory [3]. It was later observed to have connections with graph searching [29] and power domination [6]. As noted in [24], zero forcing is an example of a cellular automaton; such processes are common in mathematics and computer science. Variants of zero forcing, such as positive semidefinite zero forcing related to the maximum positive semidefinite nullity problem [5] and k-forcing [4], have also been studied (see also [17]). The initial focus of much work on zero forcing and variants was on the minimum number of vertices that needed to be colored initially in order to color the entire graph through the propagation process (see, for example, [16] and the references therein). Computation of the zero forcing number is NP-hard [1] and various papers giving bounds on the zero forcing number in terms of other graph parameters have appeared (see, for example, [12, 13]). More recently, derived parameters such as propagation time [11, 22] and throttling [9] have generated substantial interest and established connections between zero forcing and other graph searching parameters such as Cops and Robbers (see [7]). Zero forcing, its variants, propagation time, and throttling are one of the main subjects of the forthcoming book [23].

Zero forcing on a graph G is described by the following (standard) zero forcing color change rule: Given a set B of vertices of G that are colored blue with the remaining vertices colored white, a blue vertex u can change the color of (force) a white vertex w to blue if w is the only white neighbor of u; this is denoted by $u \to w$. A force performed using the zero forcing color change rule is also called a *deterministic force*. A zero forcing set of G is a set $Z \subseteq V(G)$ of vertices such that when the vertices of Z are colored blue and the remaining vertices are colored white, every vertex can eventually be colored blue by repeated applications of the color change rule. The zero forcing number of G, $\mathrm{Z}(G)$, is the minimum cardinality of a zero forcing set.

Probabilistic zero forcing was introduced by Kang and Yi in [25, Definition 1.1]. Given a set B of currently blue vertices, in one round each blue vertex $u \in B$ *fires* at, i.e., attempts to force (change the color to blue), each of its white neighbors $w \in V(G) \setminus B$ independently with probability

$$\mathbf{Pr}(u \to w) = \frac{|N[u] \cap B|}{\deg u}. \tag{1}$$

The coloring rule just described is the *probabilistic color change rule* (in [25] $\mathbf{Pr}(u \to w)$ is denoted by $F(u \to w)$). *Probabilistic zero forcing* refers to the process of coloring a graph blue by repeatedly applying the probabilistic color change rule. As noted in [25], the definition of probability of a force in (1) has the property that a deterministic force will be performed with probability one.

The probabilistic color change rule produces a discrete dynamical system that is of mathematical interest and that plausibly describes many applications. The evolution of this discrete dynamical system is a Markov process, as Kang and Yi note in [25]. Zero forcing is sometimes used to model rumor spreading in social networks, and given human nature a probabilistic model is reasonable to study. A probabilistic model is also of interest for the spread of infection among a population, or the spread of a computer virus in a network. A related model from the literature is the *push model*, also called *randomized rumor spreading*, where an initial vertex in the graph is given a rumor, and then in each round, every vertex that knows the rumor selects one of its neighbors uniformly at random and passes the rumor to them [27]. For any regular graph on n vertices, it is known that randomized rumor spreading takes $\Omega(\log n)$ rounds to complete [15], and that the complete graph on n vertices has the fastest completion time among regular graphs [18].

Probabilistic zero forcing also offers a new perspective on forcing and propagation, and it is necessary to revise the parameters of interest. For zero forcing, determining the minimum number of vertices needed to color the graph blue is a main research question (on a connected graph of order at least three that is not a path, no one vertex is a zero forcing set). However, in probabilistic zero forcing *any* one vertex in a connected graph can eventually force the entire graph blue. Since a minimum zero forcing set is not of interest, we consider other parameters for the study for probabilistic zero forcing.

A natural object of study is the expected number of rounds needed to color all vertices blue with a given starting set of vertices, especially starting with a single vertex; this is the parameter studied in Section 2. In the deterministic case, the *propagation time* $\mathrm{pt}(G, Z)$ of a zero forcing set Z for G is the number of time steps needed to color all vertices blue, performing independent forces simultaneously at each time step [22]. The *propagation time* $\mathrm{pt}(G)$ of a graph G is the minimum of $\mathrm{pt}(G, Z)$ over all minimum zero forcing sets Z.

We can recast the definition of zero forcing in parallel with that of probabilistic zero forcing, which is particularly useful for defining a time step in the study of propagation time. Given a set B of currently blue vertices, in one time step (analogous to a round) each blue vertex $u \in B$ fires at, i.e., attempts to force (change the color to blue), each of its white neighbors $w \in V(G) \setminus B$ independently with probability

$$\mathbf{Pr}(u \to w) = \begin{cases} 1 & \text{if } w \text{ is the only white neighbor of } u, \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

The *probabilistic propagation time* of a nonempty set Z of vertices of a connected graph G, $\mathrm{pt}_{\mathrm{pzf}}(G, Z)$, is a random variable that reflects the time (number of the round) at which the last white vertex turns blue when applying a probabilistic zero forcing process starting with the set Z blue (if G is not connected, we assume Z contains at least one vertex from each connected component of G). For a graph G of order n and a set $Z \subseteq V(G)$ of vertices, the *expected propagation time of Z for G* is the expected value of the propagation time of Z, i.e.,

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

The *expected propagation time* of a connected graph G is the minimum of the expected propagation time of Z for G over all one-vertex sets Z, i.e.,

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

In Section 2 we determine $\operatorname{ept}(G)$ exactly when $G = P_n$ (a path on n vertices) or $G = C_n$ (a cycle on n vertices). We also determine asymptotic upper and lower bounds on expected propagation time for all connected graphs and apply them to additional families of graphs.

While we consider expected propagation time the most natural object of study for probabilistic zero forcing, there are other parameters of interest. Aazami introduced the study of *ℓ-round zero forcing* for (deterministic) zero forcing in [1, 2]; an ℓ-round zero forcing set can (deterministically) force the entire graph blue in at most ℓ time steps. In Section 3 we define the ℓ-round probability as the maximum possible probability of all vertices being colored blue in ℓ rounds starting with a single vertex of the graph; this parameter is a possible alternative to the parameter $P_B(G)$ introduced by Kang and Yi in [25] (see Section 3.2 for the definition of $P_B(G)$). In Section 4 we approach propagation time from the perspective of confidence levels, i.e., what is the minimum number of rounds needed to ensure the entire graph will be blue with probability at least α (e.g., $\alpha = .95$ for a 95% confidence level). Many of the results on expected propagation time can be applied to obtain results for these parameters.

A preprint of this paper appeared on the arXiv in 2018 as [21]. As a result of the posting of that preprint, several other researchers have studied expected propagation time, and in [10, 14, 26] they have improved some of our results about expected propagation time as well as introducing new directions. Whenever a result presented here has subsequently been improved, that is noted immediately after we conclude our proof. In addition to improving the bound in Theorem 2.5, Narayanan and Sun establish the bounds $\log_2 \log_2(2n) \leq \operatorname{ept}(G) \leq \frac{n}{2} + O(\log n)$ for all graphs G of order n. English et al. study expected propagation time of random graphs (and improve Corollary 2.6). Markov chain techniques (see Remark 2.12) are the focus of [10] and numerous additional such results appear there. We include the original results and proof here because because in some cases the improved results rely explicitly on the proofs included here and in some cases the methods differ significantly. A question that appeared in the original preprint has subsequently been answered, and thus does not appear here as a question. Question 2.6 in [21] asked whether adding an edge to the graph could raise the expected propagation time. This question is answered in the affirmative in [10], where the exact values of expected propagation time of a tadpole graph and a tadpole graph with extra edge are presented.

We conclude this introduction with some notation that will be used throughout and statements of results from probability theory that will be used repeatedly. The distance between vertices u and v is denoted by $\operatorname{dist}(u, v)$, and $\operatorname{dist}(u, S) = \min_{x\in S} \operatorname{dist}(u, x)$ for $S \subseteq V(G)$. The *eccentricity* of a set $U \subseteq V(G)$ is defined by $\operatorname{ecc}(U) = \max_{v\in V(G)} \operatorname{dist}(U, v)$. The *radius* of G is $\operatorname{rad}(G) = \min_{u\in V(G)} \operatorname{ecc}(\{u\})$.

For functions $f(n)$ and $g(n)$ from the nonnegative or positive integers to the real numbers, asymptotic bounds are defined as follows: $f(n) = o(g(n))$ if $\lim_{n\to\infty} \frac{f(n)}{g(n)} =$

0, $f(n) = O(g(n))$ if there exists $c > 0$ such that $f(n) \leq cg(n)$ for all n sufficiently large, $f(n) = \omega(g(n))$ if $g(n) = o(f(n))$, $f(n) = \Omega(g(n))$ if $g(n) = O(f(n))$, and $f(n) = \Theta(g(n))$ if $f(n) = O(g(n))$ and $f(n) = \Omega(g(n))$.

Theorem 1.1 (Markov's inequality)**.** *Let X be a nonnegative random variable. For any constant $a > 0$,*

$$\mathbf{Pr}(X \geq a) \leq \frac{\mathbf{E}[X]}{a}.$$

Theorem 1.2 (Chebyshev's inequality)**.** *Let X be a random variable. For any constant $c > 0$,*

$$\mathbf{Pr}(|X - \mathbf{E}[X]| \geq c) \leq \frac{\mathbf{Var}(X)}{c^2}.$$

Observation 1.3. *If the probability of an event is p, then the expected trial of the event's first occurrence in repeated trials is $\frac{1}{p}$.*

2 Expected propagation time

In this section we determine the expected propagation time of cycles and paths exactly, develop several tools for bounding the expected propagation time, and apply these tools to obtain bounds on the expected propagation time of several additional families of graphs.

Proposition 2.1. *For a cycle of order $n > 2$,*

$$\operatorname{ept}(C_n) = \begin{cases} \frac{n}{2} + \frac{1}{3} & \text{if } n \text{ is even,} \\ \frac{n}{2} + \frac{1}{2} & \text{if } n \text{ is odd.} \end{cases}$$

Proof. Since C_n is vertex transitive, it does not matter which vertex is chosen as the blue vertex. Observe that $\operatorname{ept}(C_n)$ is the sum of the expected number of rounds until the first successful probabilistic force plus the number of rounds for the remainder of the vertices to be deterministically forced blue, since the process becomes deterministic as soon as there are at least two adjacent blue vertices. Since the probability that one blue vertex forces at least one of its white neighbors in any round is $\frac{3}{4}$, the expectation for the first force is $\frac{4}{3}$ by Observation 1.3. The number of rounds r needed to deterministically force all remaining vertices once two or three consecutive vertices are blue is the maximum of the distance $\operatorname{dist}(w, B)$ of a white vertex w to the set B of (two or three) blue vertices. For n even, $r = \frac{n-2}{2}$ (regardless of whether there are two or three blue vertices), so $\operatorname{ept}(C_n) = \frac{4}{3} + \frac{n-2}{2} = \frac{n}{2} + \frac{1}{3}$.

Now assume n is odd. The case of three blue vertices must be distinguished from two blue vertices because it affects $r = \max_{w \in V(G) \setminus B} \operatorname{dist}(w, B)$. When there are three consecutive blue vertices, $r = \frac{n-3}{2}$, whereas with two adjacent blue vertices $r = \lceil \frac{n-2}{2} \rceil = \frac{n-1}{2}$. Assuming that the first force has taken place in the prior round, the probability of exactly three blue vertices (two forces occurred) is $\frac{1}{3}$ and the probability of exactly two blue vertices (one force occurred) is $\frac{2}{3}$. Thus $\operatorname{ept}(C_n) = \frac{4}{3} + \frac{1}{3}(\frac{n-3}{2}) + \frac{2}{3}(\frac{n-1}{2}) = \frac{n}{2} + \frac{1}{2}$. □

Proposition 2.2. *For a path of order $n > 2$,*

$$\operatorname{ept}(P_n) = \begin{cases} \frac{n}{2} + \frac{2}{3} & \text{if } n \text{ is even,} \\ \frac{n}{2} + \frac{1}{2} & \text{if } n \text{ is odd.} \end{cases}$$

Proof. Number the vertices of P_n in order from 1 to n. We begin by considering the case that a center vertex $u = \lceil \frac{n}{2} \rceil$ is the blue vertex. For odd n, the situation is the same as that of a cycle (see Lemma 2.1) and $\operatorname{ept}(P_n) = \frac{4}{3} + \frac{1}{3}(\frac{n-3}{2}) + \frac{2}{3}(\frac{n-1}{2}) = \frac{n}{2} + \frac{1}{2}$.

Now assume n is even, which means the distance from $u = \frac{n}{2}$ to n is $\frac{n}{2}$ whereas the distance from u to 1 is $\frac{n}{2} - 1$. Assuming that at least one force takes place, the probability of $\frac{n}{2} + 1$ being forced (with or without $\frac{n}{2} - 1$ being forced) is $\frac{2}{3}$ and the probability of only $\frac{n}{2} - 1$ being forced is $\frac{1}{3}$. Thus $\operatorname{ept}(P_n) = \frac{4}{3} + \frac{2}{3}(\frac{n}{2} - 1) + \frac{1}{3}(\frac{n}{2}) = \frac{n}{2} + \frac{2}{3}$.

Finally, consider the case of $\operatorname{ept}(P_n, \{v\})$ when v is not a center vertex. The computation is analogous to that for the even case, but with the maximum distance being greater than $\frac{n}{2}$, which results in a greater expected propagation time. □

Next we prove a main lemma that provides an upper bound for expected propagation time for the neighborhood of a vertex based on its degree.

Lemma 2.3. *Let G be a graph. Then for any vertex v of G,*

$$\operatorname{ept}(G[N[v]]) = O(\log \deg v).$$

Proof. Let $d = \deg v$. Let G' be the graph obtained from $G[N[v]]$ by removing all edges except for those incident to v; note that the order of G' is $d + 1$. Since the degree of v is the same in both G and G' and additional forcing of vertices in $G[N[v]]$ may be possible in G, $\operatorname{ept}(G[N[v]]) \leq \operatorname{ept}(G', \{v\})$. Thus it suffices to prove that $\operatorname{ept}(G', \{v\}) = O(\log d)$. Since asymptotic bounds are for sufficiently large values, we assume $d \geq 272$. We establish the following three claims, using b to denote the number of currently blue vertices in G' and $w = d + 1 - b$ to denote the number of currently white vertices.

(C1) For $1 \leq b \leq 36$, the probability of at least one new blue vertex in G' in one round is at least $\frac{1}{2}$.

(C2) For $36 \leq b \leq \frac{d}{2}$, the probability of at least $\frac{b}{4}$ new blue vertices in G' in one round is at least $\frac{1}{2}$.

(C3) For $1 \leq w \leq \frac{d}{2}$, the probability of at least $\frac{w}{4}$ new blue vertices in G' in one round is at least $\frac{16}{17}$.

Once the three claims have been established, by Observation 1.3 the expected number of rounds to satisfy the condition for one new blue vertex, at least $\frac{b}{4}$ new blue vertices, or at least $\frac{w}{4}$ new blue vertices, is at most 2, 2, or $\frac{17}{16}$, respectively. Thus the expected number of rounds to go from 1 to at least 36 blue vertices is $O(1)$. Starting with between 36 and $\frac{d}{2}$ blue vertices, the expected number of rounds until the number of blue vertices goes up by 25% is at most 2. Thus the expected number of rounds until the number of blue vertices is at least $\left(\frac{5}{4}\right)^r 36$ is at most $2r$, and the expected number of rounds to go from at least 36 blue vertices to at least $\frac{d}{2}+1$ blue vertices is $O(\log d)$. Starting with at least $\frac{d}{2}+1$ blue vertices, or at most $\frac{d}{2}$ white vertices, the expected number of rounds until the number of white vertices decreases by 25% is $\frac{17}{16}$. Thus the expected number of rounds until the number of white vertices is at most $\left(\frac{3}{4}\right)^r\left(\frac{d}{2}\right)$ is at most $\frac{17}{16}r$, and the expected number of rounds to go from at least $\frac{d}{2}+1$ blue vertices to all blue vertices is $O(\log d)$. Thus $\operatorname{ept}(G', \{v\}) = O(\log d)$.

For (C1), note that when there are b blue vertices, the probability that at least one additional vertex gets colored blue in the current round is

$$1-\left(1-\frac{b}{d}\right)^{d+1-b} \geq 1-\left(1-\frac{b}{d}\right)^{d-b} \geq 1-\frac{1}{e^{b(d-b)/d}} \geq \frac{1}{2}$$

for $b \leq 36$.

For (C2), let $p(b)$ be the probability that the number of vertices forced in the current round is at least $\frac{b}{4}$, given that there are currently b blue vertices and $36 \leq b \leq \frac{d}{2}$. For each white vertex $v_1, \ldots, v_{d+1-b}$, define X_i to be 1 if v_i is colored blue in this round and 0 otherwise. Let $X = \sum_{i=1}^{d+1-b} X_i$. Since the X_i's are independent identically distributed (i.i.d.) random variables with $\mathbf{E}[X_i] = \frac{b}{d}$ and $\mathbf{Var}[X_i] = \frac{b(d-b)}{d^2}$, we have $\mathbf{E}[X] = \frac{b(d+1-b)}{d} > \frac{b}{2}$ and $\mathbf{Var}[X] = \frac{b(d-b)(d+1-b)}{d^2} \leq b$. For $36 \leq b \leq \frac{d}{2}$,

$$\begin{aligned}
1-p(b) = \mathbf{Pr}\left(X < \frac{b}{4}\right) & \\
&\leq \mathbf{Pr}\left(X \leq \frac{b}{4}\right)\\
&= \mathbf{Pr}\left(\frac{b}{2} - X \geq \frac{b}{4}\right)\\
&\leq \mathbf{Pr}\left(\mathbf{E}[X] - X \geq \frac{b}{4}\right)\\
&\leq \mathbf{Pr}\left(|X-\mathbf{E}[X]| \geq \frac{b}{4}\right) \leq \frac{\mathbf{Var}[X]}{(\frac{b}{4})^2} \leq \frac{16}{b} < \frac{1}{2}
\end{aligned}$$

where $\mathbf{Pr}\left(|X-\mathbf{E}[X]| \geq \frac{b}{4}\right) \leq \frac{\mathbf{Var}[X]}{(\frac{b}{4})^2}$ is Chebyshev's inequality and the other equalities and inequalities follow from the definitions, the values above, or algebraic manipulation. Thus in this case $p(b) > \frac{1}{2}$.

For (C3), let $q(w)$ be the probability that the number of new blue vertices in the current round is at least $\frac{w}{4}$, given that there are currently $w \leq \frac{d}{2}$ white vertices in G'. For each white vertex $v_1, \ldots, v_w$, define Y_i to be 1 if v_i is colored blue and 0 otherwise. Let $Y = \sum_{i=1}^{w} Y_i$. Since the Y_i's are i.i.d. with $\mathbf{E}[Y_i] = \frac{d+1-w}{d}$ and $\mathbf{Var}[Y_i] = \frac{(d+1-w)(w-1)}{d}$, we have $\mathbf{E}[Y] = \frac{(d+1-w)w}{d} \geq \frac{w}{2}$ and $\mathbf{Var}[Y] = \frac{(d+1-w)(w-1)w}{d^2}$.

Since $d \geq 272$,

$$1 - q(w) \leq \mathbf{Pr}\Big(Y \leq \frac{w}{4}\Big) \;=\; \mathbf{Pr}\Big(\frac{w}{2} - Y \geq \frac{w}{4}\Big) \;\leq\; \mathbf{Pr}\Big(\mathbf{E}[Y] - Y \geq \frac{w}{4}\Big)$$
$$\leq\; \mathbf{Pr}\Big(|Y - \mathbf{E}[Y]| \geq \frac{w}{4}\Big) \leq \frac{\mathbf{Var}[Y]}{(\frac{w}{4})^2} \leq \frac{16}{d} \leq \frac{1}{17}. \qquad \square$$

The next lemma will be used to establish a general upper bound.

Lemma 2.4. *Let G be a graph and suppose that the vertices $v_1, \dots, v_b$ each have degree at most k and all are colored blue. Then the expected number of rounds until all of their neighbors are colored blue is $O(\log b \log k)$.*

Proof. In Lemma 2.3, we proved that if v is a vertex with k neighbors, then once v turns blue, the expected number of rounds before all of the neighbors of v are blue is at most $c \log k$ for some constant c.

Define a *block* as a consecutive sequence of $2c \log k$ rounds. By Markov's inequality the probability that all of the neighbors of v get colored within one block after v is colored is at least $\frac{1}{2}$. We bound the expected number of blocks for all neighbors of the vertices $v_1, \dots, v_b$ to be successfully colored from the first time at which all of $v_1, \dots, v_b$ have been colored blue.

Let X_i be the random variable for the number of blocks that it takes for all the neighbors of v_i to be colored blue, and define $X = \max(X_i : 1 \leq i \leq b)$. If $F(x) = \min(\mathbf{Pr}(X_i \leq x) : 1 \leq i \leq b)$, then observe that $\mathbf{Pr}(X \leq x) \geq F(x)^b$.

Note that $F(x) \geq 1 - \left(\frac{1}{2}\right)^{\lfloor x \rfloor}$, so

$$\mathbf{E}[X] \leq \int_0^\infty \left(1 - \left(1 - \left(\frac{1}{2}\right)^{\lfloor x \rfloor}\right)^b\right) dx \;=\; \sum_{n=0}^{\infty} \left(1 - \left(1 - \left(\frac{1}{2}\right)^n\right)^b\right)$$
$$\leq\; \sum_{n=0}^{\infty} \min\left(1, b\left(\frac{1}{2}\right)^n\right) \leq \lfloor \log_2 b \rfloor + 3.$$

Thus if the vertices $v_1, \dots, v_b$ all are colored blue, then the expected number of rounds for all neighbors of $v_1, \dots, v_b$ to get colored is $O(\log b \log k)$. $\square$

Since a vertex at a distance r from the one initially blue vertex cannot be reached in fewer than r rounds, it is natural to develop general bounds that apply to all graphs in terms of both the radius, $\mathrm{rad}(G)$, and the order of G.

Theorem 2.5. *For all connected graphs G of order n,*

$$\mathrm{rad}(G) \leq \mathrm{ept}(G) = O(\mathrm{rad}(G)(\log n)^2)$$

and the lower bound is asymptotically tight.

Proof. The lower bound of $\mathrm{rad}(G)$ is immediate because the vertices colored in round i can be distance at most i from the one vertex that was colored blue initially. The path and cycle show that the lower bound is asymptotically tight (see Propositions 2.2 and 2.1).

For the upper bound, initially color a center vertex of G blue. At an arbitrary step of the coloring process, suppose that there are $b \leq n$ blue vertices $v_1, \ldots, v_b$ that have at least one white neighbor. In Lemma 2.4, we proved that if $v_1, \ldots, v_b$ are vertices each with at most k neighbors, then there exists a constant c such that once $v_1, \ldots, v_b$ are all blue, the expected number of rounds before all of their neighbors are blue is at most $c \log k \log b$.

Thus after the round during which the last of $v_1, \ldots, v_b$ is colored blue, the expected number of rounds for all neighbors of $v_1, \ldots, v_b$ to get colored blue is $O((\log n)^2)$. Since all vertices in G are within distance $\mathrm{rad}(G)$ of the initial blue vertex, the expected number of rounds until every vertex in G is blue is $O(\mathrm{rad}(G)(\log n)^2)$. □

The upper bound in Theorem 2.5 is improved in Theorem 3.1 of [26] to $\mathrm{ept}(G) = O\left(\mathrm{rad}(G) \log\left(\frac{n}{\mathrm{rad}(G)}\right)\right)$, and in [26, Theorem 3.6] it is shown that this upper bound is tight.

The Erdős-Rényi random graph with edge probability p is denoted by $G(n,p)$. For a fixed probability p or a probability bounded by a function of n, we say $G(n,p)$ has some property P with *with high probability* if the probability that $G(n,p)$ has the property P goes to one as n goes to infinity.

Corollary 2.6. *For $p \geq (\sqrt{2} + o(1))\frac{\sqrt{\log n}}{\sqrt{n}}$, with high probability* $\mathrm{ept}(G(n,p)) = O((\log n)^2)$.

Proof. For $p \geq (\sqrt{2} + o(1))\frac{\sqrt{\log n}}{\sqrt{n}}$, the random graph $G(n,p)$ has diameter at most 2 with high probability (see, for example, [19, Theorem 7.1]). Thus $G(n,p)$ has radius at most 2 with high probability. Thus the result follows from Theorem 2.5. □

Corollary 2.6 is improved in Theorems 3.1 and 4.1 of [14], showing for fixed probability $0 < p < 1$, $\mathrm{ept}(G(n,p)) = (1 + o(1)) \log_2 \log_2 n$ with high probability. These theorems are individually stronger, e.g., [14, Theorem 3.1] shows that $\mathrm{ept}(G(n,p)) = (1 + o(1)) \log_2 \log_2 n + (1 + o(1)\frac{1}{p}$ with high probability whenever $p = \omega\left(\frac{\log n}{n}\right)$.

Next we apply previous results to graphs with certain properties and families of graphs. The next result is immediate from Lemma 2.3.

Corollary 2.7. *If a graph G of order n has a universal vertex, then* $\mathrm{ept}(G) = O(\log n)$.

Corollary 2.7 is used in the proof Theorem 3.2 of [10] to establish the following result: Let c be a fixed positive integer. Then $\mathrm{ept}(G) = \Theta(\log n)$ for a graph G of order n with a universal vertex u whose deletion leaves a graph having maximum degree at most c.

Theorem 2.8. *For the star on $n+1$ vertices,* $\mathrm{ept}(K_{1,n}) = \Theta(\log n)$.

Proof. The upper bound follows from Lemma 2.3. For the lower bound, let $h(b)$ be the probability that the number of new blue vertices colored in the current round is at most $4b$, given that there are currently b blue vertices in $K_{1,n}$ and the center vertex is blue. Using the same setup with the random variables X_i for each $i = 1, \dots, n+1-b$ and $X = \sum_{i=1}^{n+1-b} X_i$ as in Lemma 2.3, $\mathbf{E}[X] = \frac{b(n+1-b)}{n} \le b$ and $\mathbf{Var}[X] = \frac{b(n-b)(n+1-b)}{n^2} \le b$. We again use Chebyshev's inequality to show that $h(b) = 1 - O(\sqrt{n}^{-1})$ for $\sqrt{n} \le b \le \frac{n}{2}$:

$$\begin{aligned} 1 - h(b) \le \mathbf{Pr}(X - b \ge 3b) &\le \mathbf{Pr}(|X - \mathbf{E}[X]| \ge 3b) \\ &\le \frac{\mathbf{Var}[X]}{(3b)^2} \le \frac{1}{9b} \le \frac{1}{9\sqrt{n}} = O\left(\frac{1}{\sqrt{n}}\right). \end{aligned}$$

Since starting with $\sqrt{n} \le b \le \frac{n}{2}$ blue vertices and coloring at most $4b$ additional vertices blue means there are at most $5b$ blue vertices after the round, the probability that there are at most $5^r b$ blue vertices after r rounds is at least $\left(1 - O(\frac{1}{\sqrt{n}})\right)^r$. Thus going from $b \le \sqrt{n}$ blue vertices to at least $\frac{n}{2}$ blue vertices requires that $5^r\sqrt{n} \ge \frac{n}{2}$, or $r \ge \log_5\left(\frac{\sqrt{n}}{2}\right)$. Thus with probability at least $(1 - O(\sqrt{n}^{-1}))^{\log_5(\sqrt{n}/2)} = 1 - o(1)$, it takes at least $\log_5\left(\frac{\sqrt{n}}{2}\right)$ rounds for the number of blue vertices to increase from at most $\sqrt{n}$ to at least $\frac{n}{2}$. Since $\sqrt{n} \ge 2$ for $n > 3$, we have covered the case in which the first blue vertex is a leaf rather than the center, because in that case the expected propagation time is one more than the expected propagation time starting with two blue vertices, one of which is the center. Thus $\mathrm{ept}(K_{1,n}) = \Omega(\log n)$. $\square$

Theorem 2.8 is extended in Theorem 3.3 of [10] to $\mathrm{ept}(K_{c,n}) = \Theta(\log n)$ for a fixed positive integer c.

Proposition 2.9. *For the complete graph on n vertices,* $\mathrm{ept}(K_n) = \Omega(\log\log n)$.

Proof. Let $\hat{h}(b)$ be the probability that the number of additional blue vertices colored in the current round is at most $4b^2$, given that there are currently b blue vertices. For each white vertex $v_1, \dots, v_{n-b}$, define X_i to be 1 if v_i gets colored blue and 0 otherwise and $X = \sum_{i=1}^{n-b} X_i$. Since the X_i's are i.i.d. with $\mathbf{E}[X_i] = 1 - \left(1 - \frac{b}{n-1}\right)^b$ and $\mathbf{Var}[X_i] = \left(1 - \left(1 - \frac{b}{n-1}\right)^b\right)\left(1 - \frac{b}{n-1}\right)^b$, we have $\mathbf{E}[X] = \left(1 - \left(1 - \frac{b}{n-1}\right)^b\right)(n-b)$, and furthermore $\mathbf{Var}[X] = \left(1 - \left(1 - \frac{b}{n-1}\right)^b\right)\left(1 - \frac{b}{n-1}\right)^b (n-b) \le \mathbf{E}[X] \le b^2$ by Bernoulli's inequality. By algebraic manipulation and Chebyshev's inequality,

$$1 - \hat{h}(b) \le \mathbf{Pr}\big(X - b^2 \ge 3b^2\big) \le \mathbf{Pr}\big(|X - \mathbf{E}[X]| \ge 3b^2\big) \le \frac{\mathbf{Var}[X]}{(3b^2)^2} \le \frac{1}{9b^2} = O\left(\frac{1}{b^2}\right).$$

Since starting with $\log n \le b \le n$ blue vertices and coloring at most $4b^2$ additional vertices blue means there are at most $5b^2 \le b^3$ blue vertices after the round for $b \ge 5$, the probability that there are at most $b^{(3^r)}$ blue vertices after r rounds is at least $\left(1 - O(\frac{1}{(\log n)^2})\right)^r$. Thus with probability at least $1 - o(1)$, going from $b \le \log n$ blue vertices to n blue vertices requires that $(\log n)^{(3^r)} \ge n$, or $r \ge \log_3\left(\frac{\log n}{\log \log n}\right)$. Thus with probability at least $1 - o(1)$, it takes $\Omega(\log \log n)$ rounds for the number of blue vertices to increase from at most $\log n$ to exactly n. Therefore $\text{ept}(K_n) = \Omega(\log \log n)$. □

Proposition 2.9 is extended in Theorem 3.1 of [10] to establish that $\text{ept}(K_n) = \Theta(\log \log n)$. Note that this is much faster than the $\Theta(\log n)$ completion time of randomized rumor spreading for K_n [18].

A *spider* is a tree with exactly one vertex of degree at least three, which is called the *body vertex*. The *legs* are the paths that result from deleting the body vertex. The number of legs is the degree of the body vertex.

Proposition 2.10. *Let G be a spider with k legs. Then* $\text{ept}(G) = \text{rad}(G) + O(\log k)$.

Proof. By Theorem 2.5, $\text{ept}(G) \ge \text{rad}(G)$. For the upper bound, we initially color a center vertex v of G blue. Let u be the body vertex of G. If $v \ne u$, then the expected time of first force is $\frac{4}{3}$. After the first force the process becomes deterministic until u is colored blue. By Lemma 2.3, the number of rounds after u is colored blue for all of u's neighbors to get colored blue is $O(\log k)$. Then the process becomes deterministic until the graph is all blue. This proves the upper bound since all vertices of G are within $\text{rad}(G)$ of v. □

Recall that a *full k-ary tree of height h*, denoted by $T_{k,h}$, is constructed from a root by performing h steps in which k leaves are appended to each vertex of degree at most one. Observe that the order of $T_{k,h}$ is $n = \frac{k^{h+1}-1}{k-1}$, so $h = \log_k((k-1)n+1) - 1$.

Proposition 2.11. *Let $T_{k,h}$ be a full k-ary tree of order n. Then* $\log_k((k-1)n+1) - 1 \le \text{ept}(T_{k,h}) = O((\log n)^2)$, *where the constant in the upper bound depends on k.*

Proof. The lower bound follows from Theorem 2.5 and $\text{rad}(T_{k,h}) = h = \log_k((k-1)n+1) - 1$. For the upper bound, we initially color the root vertex v of G blue. The expected number of rounds for all of the neighbors of v to be colored blue is $O(\log k)$ by Lemma 2.3.

Suppose that at some stage of the coloring process, all of the k^t vertices $v_1, \ldots, v_{k^t}$ in level t of the k-ary tree have been colored blue. By Lemma 2.4, the expected number of rounds for all neighbors of $v_1, \ldots, v_{k^t}$ to get colored after $v_1, \ldots, v_{k^t}$ have been colored is $O(t)$, where the constant in the bound depends on k.

Since $T_{k,h}$ has $h = \Theta(\log n)$ levels (where the constants in the bound depend on k), the expected number of rounds for every vertex in G to be colored is $\sum_{t=1}^{O(\log n)} O(t) = O((\log n)^2)$. □

Finally, note that a probabilistic zero forcing process is a Markov chain. For a discussion of the construction of the Markov transition matrix, see [25] (which introduced the Markov chain approach), [10] (which builds on results presented here), or the proof of Lemma 2.13 for an example of such a matrix. In the next remark, we explain how to use the Markov transition matrix M to compute the expected propagation time $\operatorname{ept}(G, Z)$.

Remark 2.12. Given a graph G and an initial set of blue vertices Z, each possible set of blue vertices in the probabilistic zero forcing process is a *state*. Suppose there are m states, in state 1 the blue vertices are exactly those in Z, and in state m all vertices are blue. Let M denote the $m \times m$ Markov transition matrix defined by these states and $\mathbf{q} = [1, 0, \ldots, 0]$ (since state 1 has exactly the vertices of Z blue). Then the probability that all vertices are blue after round r is $(\mathbf{q}M^r)_m = (M^r)_{1m}$, so

$$\operatorname{ept}(G, Z) = \sum_{r=1}^{\infty} r \left((M^r)_{1m} - \left(M^{r-1}\right)_{1m} \right).$$

The method described in Remark 2.12 is particularly useful for specific small cases, as in the next lemma. Recall that it was shown in Lemma 2.3 that $\operatorname{ept}(G[N[v]]) = O(\log \deg v)$ by considering a spanning star of $G[N[v]]$ with center v.

Lemma 2.13. *Let G be a graph and let v be a vertex of G. Then*

$$\operatorname{ept}(G[N[v]]) \leq \begin{cases} 1 & \textit{if } \deg v = 1, \\ 2 & \textit{if } \deg v = 2, \\ 2.76316 & \textit{if } \deg v = 3, \\ 3.34171 & \textit{if } \deg v = 4. \end{cases}$$

Proof. Let $d = \deg v$. Let G' be the graph obtained from $G[N[v]]$ by removing all edges except for those incident to v, so G' is a star. As in the proof of Lemma 2.3, it suffices to determine $\operatorname{ept}(G', \{v\})$ for $d = 1, 2, 3, 4$.

The case $\deg v = 1$ is deterministic zero forcing. For $\deg v = 2$, $\operatorname{ept}(G', \{v\}) = \operatorname{ept}(P_3) = 2$ by Proposition 2.2. The remaining probabilities are computed using the Markov transition matrices M_d for $d = 3, 4$ shown below (see [20] for the computational details). These transition matrices were defined by using the states consisting of $0, 1, \ldots, d$ blue leaves.

$$M_3 = \begin{bmatrix} \frac{8}{27} & \frac{4}{9} & \frac{2}{9} & \frac{1}{27} \\ 0 & \frac{1}{9} & \frac{4}{9} & \frac{4}{9} \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad M_4 = \begin{bmatrix} \frac{81}{256} & \frac{27}{64} & \frac{27}{128} & \frac{3}{64} & \frac{1}{256} \\ 0 & \frac{1}{8} & \frac{3}{8} & \frac{3}{8} & \frac{1}{8} \\ 0 & 0 & \frac{1}{16} & \frac{3}{8} & \frac{9}{16} \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \qquad \square$$

The use of Markov transition matrices to determine expected propagation times was subsequently developed more fully in [10].

3 ℓ-round probability

While we believe that expected propagation time is the most interesting parameter associated with probabilistic zero forcing, several related parameters are also of interest. Some results for the parameters discussed in this section and in Section 4 can be obtained by applying results for expected propagation time. In this section we build on Aazami's definition of ℓ-round (deterministic) zero forcing to define the ℓ-round probability for probabilistic zero forcing. An *ℓ-round zero forcing set* [1, 2] is a set of vertices that is able to color the entire graph blue in at most ℓ time steps. The *ℓ-round propagation problem* is to determine the minimum size of an ℓ-round zero forcing set. A dynamic programming algorithm is provided in [1] to optimally solve the ℓ-round propagation problem in polynomial time for graphs of bounded tree-width.

Given a connected graph G, a set B of blue vertices, and a positive integer ℓ, the *ℓ-round probability of* B, $P_B^{(\ell)}(G)$, is the probability that all vertices of G are blue after ℓ rounds of probabilistic zero forcing starting with exactly the vertices of B blue. The *ℓ-round probability of* G, $P^{(\ell)}(G)$, is the maximum of $P_B^{(\ell)}(G)$ over all one vertex sets B. We begin by determining the ℓ-round probability for cycles and paths.

Proposition 3.1. *For a cycle of order $n > 2$, $P^{(\ell)}(C_n) = 0$ for $\ell < \left\lfloor \frac{n}{2} \right\rfloor$, and for $\ell \geq \left\lfloor \frac{n}{2} \right\rfloor$,*

$$P^{(\ell)}(C_n) = \begin{cases} 1 - \left(\frac{1}{4}\right)^{\ell - n/2 + 1} & \text{if } n \text{ is even,} \\ 1 - \frac{3}{4}\left(\frac{1}{4}\right)^{\ell - (n-1)/2} & \text{if } n \text{ is odd.} \end{cases}$$

Proof. The probability that the first force occurs in the k^{th} round is $\left(\frac{1}{4}\right)^{k-1}\left(\frac{3}{4}\right)$. The probability that two forces occur in the first round that has a force is $\frac{1}{3}$, and the probability that only one force occurs on the first round that has a force is $\frac{2}{3}$.

First assume n is even. Then there will be $\frac{n}{2} - 1$ rounds after the first force, regardless of how many forces occur on the first round that has a force (call this round k). Thus, the process takes $t = \frac{n}{2} - 1 + k$ rounds with probability $\left(\frac{1}{4}\right)^{k-1}\left(\frac{3}{4}\right) = \frac{3}{4}\left(\frac{1}{4}\right)^{t-n/2}$. This implies that $P^{(\ell)}(C_n) = 0$ for $\ell < \frac{n}{2}$ and

$$P^{(\ell)}(C_n) = 1 - \sum_{t=\ell+1}^{\infty} \frac{3}{4}\left(\frac{1}{4}\right)^{t-n/2} = 1 - \left(\frac{1}{4}\right)^{\ell-n/2+1}$$

for every $\ell \geq \frac{n}{2}$.

Now assume n is odd. Then there will be $\frac{n-1}{2}$ rounds after the first force if there is only one force on the first round that has a force, and there will be $\frac{n-3}{2}$ rounds after the first force if there are two forces on the first round that has a force. The probability of two forces in the first round is $\frac{1}{3}\left(\frac{3}{4}\right) = \frac{1}{4}$, and $t = 1 + \frac{n-3}{2} = \frac{n-1}{2}$ rounds are needed to color all vertices blue. For each $t \geq \frac{n-1}{2} + 1$, there are two ways to achieve the last vertex turning blue in round t: Only one force happens in round $t - \frac{n-1}{2}$ (and no forces earlier) with probability $\frac{2}{3}\left(\frac{1}{4}\right)^{t-(n-1)/2-1}\left(\frac{3}{4}\right)$. Two forces happen in round $t - \frac{n-1}{2} + 1$ (and no forces earlier) with probability $\frac{1}{3}\left(\frac{1}{4}\right)^{t-(n-1)/2}\left(\frac{3}{4}\right)$. So the probability of the last vertex turning blue in round $t \geq \frac{n-1}{2}+1$ is $\frac{9}{16}\left(\frac{1}{4}\right)^{t-(n+1)/2}$. This implies that $P^{(\ell)}(C_n) = 0$ for $\ell < \frac{n-1}{2}$ and

$$P^{(\ell)}(C_n) = \frac{1}{4} + \frac{3}{4}\left(1 - \left(\frac{1}{4}\right)^{\ell-(n-1)/2}\right) = 1 - \frac{3}{4}\left(\frac{1}{4}\right)^{\ell-(n-1)/2}$$

for every $\ell \geq \frac{n-1}{2}$. □

Proposition 3.2. *For a path of order $n > 2$, $P^{(\ell)}(P_n) = 0$ for $\ell < \left\lfloor \frac{n}{2} \right\rfloor$, and for $\ell \geq \left\lfloor \frac{n}{2} \right\rfloor$,*

$$P^{(\ell)}(P_n) = \begin{cases} 1 - \frac{1}{2}\left(\frac{1}{4}\right)^{\ell-n/2} & \text{if } n \text{ is even,} \\ 1 - \frac{3}{4}\left(\frac{1}{4}\right)^{\ell-(n-1)/2} & \text{if } n \text{ is odd.} \end{cases}$$

Proof. As in the proof of Proposition 2.2, it suffices to start the forcing at a central vertex, since the forcing process becomes deterministic after the first force. The probability of the first force occurring on the k^{th} round and the probabilities of one or two forces in the round with the first force are the same as for a cycle. If n is odd, then the situation is the same as for a cycle, so $P^{(\ell)}(P_n) = 0$ for $\ell < \frac{n-1}{2}$ and $P^{\ell}(P_n) = 1 - \frac{3}{4}\left(\frac{1}{4}\right)^{\ell-(n-1)/2}$ for every $\ell \geq \frac{n-1}{2}$.

If n is even, then as in the proof of Proposition 2.2, we need to distinguish whether the neighbor in the longer direction is forced: There will be $\frac{n}{2} - 1$ rounds after the first force with probability $\frac{2}{3}$, and there will be $\frac{n}{2}$ rounds after the first force with probability $\frac{1}{3}$. Thus if n is even, then the process takes $t = \frac{n}{2}$ rounds with probability $\left(\frac{3}{4}\right)\left(\frac{2}{3}\right) = \frac{1}{2}$ and $t \geq \frac{n}{2}+1$ rounds with probability $\left(\frac{2}{3}\right)\left(\frac{1}{4}\right)^{t-n/2}\left(\frac{3}{4}\right) + \left(\frac{1}{3}\right)\left(\frac{1}{4}\right)^{t-n/2-1}\left(\frac{3}{4}\right) = \left(\frac{3}{8}\right)\left(\frac{1}{4}\right)^{t-n/2-1}$. This implies that $P^{(\ell)}(P_n) = 0$ for $\ell < \frac{n}{2}$ and $P^{(\ell)}(P_n) = \frac{1}{2} + \frac{1}{2}(1 - \left(\frac{1}{4}\right))^{\ell-n/2})$ for every $\ell \geq \frac{n}{2}$. □

3.1 Applications of expected propagation time to ℓ-round probability

We can obtain numerous corollaries about $P^{(\ell)}(G)$ from prior results about expected propagation time and the next two lemmas. Lemma 3.3 is an application of Markov's inequality and Lemma 3.4 follows from Lemma 3.3.

Lemma 3.3. *If G is a connected graph and $\ell > \operatorname{ept}(G)$, then $P^{(\ell)}(G) \geq 1 - \frac{\operatorname{ept}(G)}{\ell}$.*

Lemma 3.4. *For all connected graphs G, $\operatorname{ept}(G) = O(f(n))$ implies $P^{(\ell)}(G) = 1 - o(1)$ for $\ell = \omega(f(n))$.*

Corollary 3.5. *$P^{(\ell)}(G) = 1 - o(1)$ for $\ell = \omega(\log n)$ for every graph G with a universal vertex.*

We obtain the next table of bounds on $P^{(\ell)}(G)$ for various graph classes G. The statements in rows 2, 4, 6, 7, and 8 follow from Lemma 3.4 together with Theorem 2.8, [10, Theorem 3.1], Propositions 2.10 and 2.11, and Corollary 2.6. The third and fifth rows use the same methods as in Theorem 2.8 and Proposition 2.9.

An interesting open problem would be to see if more sophisticated techniques than Markov's inequality, such as Chebyshev's inequality, could be used to sharpen any of the following bounds.

graph class G	ℓ	$P^{(\ell)}(G)$
$K_{1,n}$	$\omega(\log n)$	$1-o(1)$
$K_{1,n}$	$o(\log n)$	$o(1)$
K_n	$\omega(\log\log n)$	$1-o(1)$
K_n	$o(\log\log n)$	$o(1)$
spider S with k legs	$\omega(\mathrm{rad}(S)+\log k)$	$1-o(1)$
full k-ary tree T	$\omega((\log n)^2)$	$1-o(1)$
$G(n,p)$ for fixed $0<p<1$	$\omega((\log n)^2)$	$1-o(1)$

Table 3.1: Bounds on $P^{(\ell)}(G)$ for classes of graphs G.

Corollary 3.6. *$P^{(\ell)}(G) = 0$ for $\ell < rad(G)$ and $P^{(\ell)}(G) = 1 - o(1)$ for all $\ell = \omega(\mathrm{rad}(G)(\log n)^2)$ for every connected graph G.*

Proof. The first statement is true since the number of steps in the coloring process cannot be less than $\mathrm{rad}(G)$ if we start with only one blue vertex. The second statement follows from Lemma 3.4 and Theorem 2.5. □

3.2 Discussion of Kang and Yi's $P_B(G)$ and its properties

After defining probabilistic zero forcing in [25], Kang and Yi define the probability $P_B(G)$ as follows: Let k_o be the first round in which it is possible to have a deterministic zero forcing set colored blue, starting with exactly the vertices in B colored blue. Define $P_B(G)$ to be the probability that a deterministic zero forcing set has been colored blue in round k_o.

A graph G together with an assignment of one of the colors blue and white to each vertex of G is called a *colored graph*. We use G_B to denote the colored graph with underlying graph G and set of blue vertices B. For $B \subseteq V(G)$, S_B^k denotes the set of colored graphs that are possible (i.e., have positive probability) after the kth round starting with G_B; note that $S_B^0 = \{G_B\}$. For $R^k \subseteq S_B^k$, $P^{(k)}(R^k)$ is the probability that after round k the result is one of the colored graphs in R^k. Let $T_B^k = \{G_Z \in S_B^k : Z \text{ is a (deterministic) zero forcing set for } G\}$. Then $P_B(G) = P^{(k_0)}(T_B^{k_0})$ where k_0 is the least k that $T_B^k \neq \emptyset$ [25] (and $P_\emptyset(G) = 0$).

In [25, page 13], Kang and Yi claim the following three properties are clear for $P_B(G)$:

1. $P_\emptyset(G) = 0$.
2. If Z is a zero forcing set for G, then $P_Z(G) = 1$.
3. If $A \subseteq B \subseteq V(G)$, then $P_A(G) \leq P_B(G)$.

The first of these properties is by definition (this was not explicit in the definition in [25] but is clearly what is intended), and the second follows from the fact that $T_Z^0 = \{G_Z\}$ when Z is a zero forcing set, and thus $P^{(0)}(T_Z^0) = 1$. However, the third property, $A \subseteq B \Rightarrow P_A(G) \leq P_B(G)$, is not true, as shown in Example 3.7. We show that ℓ-round probability of B has the first and third properties desired by Kang and Yi and a modified form of the second (see Proposition 3.8 below).

The problem with the third property for $P_B(G)$ is that the definition depends on the round in which it is first possible to have a zero forcing set colored blue. With a larger B, this may occur in an earlier round but with lower probability. This is illustrated in the next example.

G_A G_B

Figure 3.1: Two colorings of a graph G with $A \subset B$ and $P_A(G) > P_B(G)$.

Example 3.7. Let G be the graph shown in Figure 3.1 with two colorings $A = \{1\}$ and $B = \{1, 3\}$. Any zero forcing set for G must contain at least one of 4 and 5, and this is sufficient to guarantee a zero forcing set is blue given that vertex 1 is blue.

For G_A, it takes at least three rounds to reach vertex 4 or 5 and it is possible to color 4 or 5 blue in the third round. Thus $P_A(G) = P^{(3)}(T_A^3)$ where T_A^3 is the set of all colored graphs attainable from G_A in three rounds that have at least one of 4 or 5 blue. Vertex 2 is forced in the first round and vertex 3 is colored blue in the second round, so $T_A^3 = \{G_{\{1,2,3,4\}}, G_{\{1,2,3,5\}}, G_{\{1,2,3,4,5\}}\}$. The probability $3 \to 4$ (or $3 \to 5$) in the third round is $\frac{2}{3}$, so the probability of at least one of 4 and 5 being colored blue is $1 - \left(\frac{1}{3}\right)\left(\frac{1}{3}\right) = \frac{8}{9} = P^{(3)}(T_A^3) = P_A(G)$.

For G_B, it is possible to color at least one of 4 or 5 in round one. The probability of $3 \to 4$ (or $3 \to 5$) is $\frac{1}{3}$, so the probability of at least one being forced in round one is $1 - \left(\frac{2}{3}\right)\left(\frac{2}{3}\right) = \frac{5}{9} = P^{(1)}(T_B^1) = P_B(G)$.

Thus
$$P_A(G) = \frac{8}{9} > \frac{5}{9} = P_B(G).$$

Property (ii) of Proposition 3.8 has the added stipulation that $\ell \geq \mathrm{pt}(G, Z)$, which is necessary and seems reasonable given the definition of ℓ-round probability. In the definition of $P_B(G)$, the probability is measured after the first round in which a zero forcing set can be colored blue, and this is incompatible with the third property. Note a major difference between $P_B(G)$ and $P_B^{(\ell)}(G)$ is that $P_B(G)$ requires only a deterministic zero forcing set to be colored blue, while $P_B^{(\ell)}(G)$ requires the whole vertex set to be colored. We believe the definition of $P_B^{(\ell)}(G)$ is more natural for the application to rumor spreading, since the push model of randomized rumor spreading also requires that the whole vertex set be colored [27].

Proposition 3.8. *Let G be a graph and ℓ be a positive integer. Then*

(i) $P_{\emptyset}^{(\ell)}(G) = 0$.

(ii) *If Z is a zero forcing set for G and $\ell \geq \text{pt}(G, Z)$, then $P_Z^{(\ell)}(G) = 1$.*

(iii) *If $A \subseteq B \subseteq V(G)$, then $P_A^{(\ell)}(G) \leq P_B^{(\ell)}(G)$.*

Proof. (i): Since only blue vertices can force, $P_{\emptyset}^{(\ell)}(G) = 0$. For (ii), let Z be a zero forcing set. Using only zero forcing will color the entire graph blue in $\text{pt}(G, Z)$ time steps (rounds). Since probabilistic zero forcing is not slower, $P_Z^{(\ell)}(G) = 1$. For (iii), a long proof involving multidimensional rectangles was presented in [21]. However, Lemma 2.2 of the later paper [14] used a much simpler coupling argument to prove the same result (without relying on any results in [21]). □

4 Confidence propagation time

In this section, we consider another perspective on probabilistic zero forcing to which we can apply results on expected propagation time. For any connected graph G, define $\text{pt}_{\text{pzf}}(G, Z, \alpha)$ to be the least number of rounds t such that the probability that all the vertices are blue after round t is greater than or equal to α, assuming that the vertices in Z are colored blue initially. This can be thought of as the the time at which you have *alpha*-confidence that the graph is all blue when starting with Z, and is called the *α-confidence propagation time.* Define $\text{pt}_{\text{pzf}}(G, \alpha) = \min_{v \in V(G)} \text{pt}_{\text{pzf}}(G, \{v\}, \alpha)$.

Confidence propagation time can be determined immediately from ℓ-round probability when this is known, as for cycles and paths (Propositions 3.1 and 3.2). More specifically, for each line of the following propositions with a number of rounds x and a probability range $p_0 < \alpha \leq p_1$, the p_0 values are obtained by using $\ell = x - 1$ in Propositions 3.1 and 3.2, and the p_1 values are obtained by using $\ell = x$.

Corollary 4.1. *For a cycle of order n,*

$$\text{pt}_{\text{pzf}}(C_n, \alpha) = \begin{cases} \frac{n}{2} & \text{if } n \text{ is even and } 0 < \alpha \leq \frac{3}{4}, \\ \frac{n}{2} + x & \text{if } n \text{ is even and } 1 - (\frac{1}{4})^x < \alpha \leq 1 - (\frac{1}{4})^{x+1}, \\ \frac{n-1}{2} & \text{if } n \text{ is odd and } 0 < \alpha \leq \frac{1}{4}, \\ \frac{n-1}{2} + x & \text{if } n \text{ is odd and } 1 - \frac{3}{4}(\frac{1}{4})^{x-1} < \alpha \leq 1 - \frac{3}{4}(\frac{1}{4})^x. \end{cases}$$

Corollary 4.2. *For a path of order n,*

$$\text{pt}_{\text{pzf}}(P_n, \alpha) = \begin{cases} \frac{n}{2} & \text{if } n \text{ is even and } 0 < \alpha \leq \frac{1}{2}, \\ \frac{n}{2} + x & \text{if } n \text{ is even and } 1 - \frac{1}{2}(\frac{1}{4})^{x-1} < \alpha \leq 1 - \frac{1}{2}(\frac{1}{4})^x, \\ \frac{n-1}{2} & \text{if } n \text{ is odd and } 0 < \alpha \leq \frac{1}{4}, \\ \frac{n-1}{2} + x & \text{if } n \text{ is odd and } 1 - \frac{3}{4}(\frac{1}{4})^{x-1} < \alpha \leq 1 - \frac{3}{4}(\frac{1}{4})^x. \end{cases}$$

The next lemma is analogous to Lemma 3.3 in the last section.

Lemma 4.3. *If G is a connected graph, then* $\mathrm{pt}_{\mathrm{pzf}}(G,\alpha) \le \frac{\mathrm{ept}(G)}{1-\alpha}$.

Proof. By Markov's inequality, the probability that G is not all blue by time T is at most $\frac{\mathrm{ept}(G)}{T}$. When $T = \frac{\mathrm{ept}(G)}{1-\alpha}$, this probability is at most $1-\alpha$. □

The upper bounds in the next corollary follow from Lemma 4.3 and results in Section 2 (and in the case of (3), the stronger result that $\mathrm{ept}(K_n) = \Theta(\log\log n)$ established in [10, Theorem 3.1]) . The proofs for the two lower bounds use the same method as in the proofs of Theorem 2.8 and Proposition 2.9.

Corollary 4.4. *For every constant* $0 < \alpha < 1$*:*

(1) $\mathrm{pt}_{\mathrm{pzf}}(G,\alpha) = O(\log n)$ *for every graph G with a universal vertex.*

(2) $\mathrm{pt}_{\mathrm{pzf}}(K_{1,n},\alpha) = \Theta(\log n)$.

(3) $\mathrm{pt}_{\mathrm{pzf}}(K_n,\alpha) = \Theta(\log\log n)$.

(4) $\mathrm{pt}_{\mathrm{pzf}}(T,\alpha) = O((\log n)^2)$ *for every full k-ary tree T, where the constant in the bound depends on k.*

(5) $\mathrm{rad}(G) \le \mathrm{pt}_{\mathrm{pzf}}(G,\alpha) = O(\mathrm{rad}(G)(\log n)^2)$ *for every connected graph G.*

(6) *With high probability,* $\mathrm{pt}_{\mathrm{pzf}}(G(n,p),\alpha) = O((\log n)^2)$ *for all fixed* $0 < p < 1$.

For spiders, we obtain a tighter bound than what is given by Lemma 4.3. This follows from a proof very similar to Proposition 2.10, using Markov's inequality on the sum of the times for the first force to occur and for all neighbors of the body vertex to be colored after the body vertex is colored.

Proposition 4.5. *If G is a spider with k legs, then* $\mathrm{pt}_{\mathrm{pzf}}(G,\alpha) = \mathrm{rad}(G) + O(\log k)$*, where the constant in the $O(\log k)$ depends on α.*

5 Concluding remarks

In [25] Kang and Yi provide key definitions for probabilistic zero forcing: the probability of a force (1) and the concept of a round. Here we use these definitions to begin the study of expected propagation time, ℓ-round probability, and confidence propagation time. Many questions about these parameters remain. Here we list some examples. Can one develop a reasonable approximation algorithm for $\mathrm{ept}(G)$? Which vertices v of G achieve $\mathrm{ept}(G)$? Clearly any vertex achieves $\mathrm{ept}(G)$ in a vertex-transitive graph, and we showed that a center vertex achieves $\mathrm{ept}(G)$ when G is a path. However, it is not the case that a center vertex or (when it exists) a universal vertex must realize $\mathrm{ept}(G)$; it was shown in [10] that the center vertex of $K_{1,3}$ (which is universal) does not achieve $\mathrm{ept}(G)$. What is the expected number of vertices in $G(n,p)$ that achieve $\mathrm{ept}(G)$? The fact that many results from expected propagation time can be applied to obtain results on ℓ-round probability and confidence propagation provides additional evidence that expected propagation time is a parameter worthy of study.

Acknowledgements

The authors thank the referees for their comments, which improved the exposition.

References

[1] A. Aazami, Hardness results and approximation algorithms for some problems on graphs, PhD Thesis, University of Waterloo, 2008. http://hdl.handle.net/10012/4147

[2] A. Aazami, Domination in graphs with bounded propagation: algorithms, formulations and hardness results, *J. Comb. Optim.* 19 (2010), 429–456.

[3] 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 and A. Wangsness), Zero forcing sets and the minimum rank of graphs, *Linear Algebra Appl.* 428 (2008), 1628–1648.

[4] D. Amos, Y. Caro, R. Davila and R. Pepper, Upper bounds on the k-forcing number of a graph, *Discrete Appl. Math.* 181 (2015), 1–10.

[5] F. Barioli, W. Barrett, S. Fallat, H. T. Hall, L. Hogben, B. Shader, P. van den Driessche and H. van der Holst, Zero forcing parameters and minimum rank problems, *Linear Algebra Appl.* 433 (2010), 401–411.

[6] K. F. Benson, D. Ferrero, M. Flagg, V. Furst, L. Hogben, V. Vasilevska and B. Wissman, Zero forcing and power domination for graph products, *Australas. J. Combin.* 70 (2018), 221–235.

[7] J. Breen, B. Brimkov, J. Carlson, L. Hogben, K. E. Perry and C. Reinhart, Throttling for the game of Cops and Robbers on graphs, *Discrete Math.* 341 (2018), 2418–2430.

[8] D. Burgarth and V. Giovannetti, Full control by locally induced relaxation, *Phys. Rev. Lett.* PRL 99 (2007), 100501.

[9] S. Butler and M. Young, Throttling zero forcing propagation speed on graphs, *Australas. J. Combin.* 57 (2013), 65–71.

[10] Y. Chan, E. Curl, J. Geneson, L. Hogben, K. Liu, I. Odegard and M. S. Ross, Using Markov chains to determine expected propagation time for probabilistic zero forcing, *Electron. J. Linear Alg.* 36 (2020), 318–333.

[11] K. Chilakamarri, N. Dean, C. Kang and E. Yi, Iteration Index of a Zero Forcing Set in a Graph, *Bull. Inst. Combin. Appl.* 64 (2012), 57–72.

[12] R. Davila, T. Kalinowski and S. Stephen, A lower bound on the zero forcing number, *Discrete Appl. Math.* 250 (2018), 363–367.

[13] R. Davila and F. Kenter, Bounds for the zero-forcing number of graphs with large girth, *Theory Appl. Graphs* 2, no. 2, (2015), Art. 1, 8 pp.

[14] S. English, C. MacRury and P. Prałatt, Probabilistic zero forcing on random graphs, *European J. Combin.* 91 (2021), 103207, 22 pp.

[15] R. Elsässer and T. Sauerwald, On the runtime and robustness of randomized broadcasting, *Theor. Comp. Sci.* 410 (2009), 3414–3427.

[16] S. Fallat and L. Hogben, Minimum Rank, Maximum Nullity, and Zero Forcing Number of Graphs, in *Handbook of Linear Algebra*, 2nd ed. (Ed. L. Hogben), CRC Press, Boca Raton, 2014.

[17] S. Fallat, S. Severini and M. Young, AIM workshop: Zero forcing and its variants, American Institute of Mathematics, Jan 30-Feb 3, 2017; webpage `https://aimath.org/pastworkshops/zeroforcing.html` contains workshop report and list of open problems.

[18] A. Frieze and G. Grimmett, The shortest-path problem for graphs with random-arc-lengths, *Disc. Appl. Math.* 10 (1985), 57–77.

[19] A. Frieze and M. Karoński, *Introduction to Random Graphs*, Cambridge University Press, Cambridge, 2015.

[20] J. Geneson and L. Hogben, Computations in *Sage* supporting 'Expected propagation time for probabilistic zero forcing,' PDF available at `https://aimath.org/~hogben/GH18%20--%20Sage.pdf`, *Sage* worksheet available at `https://sage.math.iastate.edu/home/pub/92/`.

[21] J. Geneson and L. Hogben, Propagation time for probabilistic zero forcing, available at `https://arxiv.org/abs/1812.10476v1`.

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

[23] L. Hogben, J. C.-H. Lin and B. Shader, *Inverse Problems and Zero Forcing for Graphs*, American Mathematical Society, Mathematical Surveys and Monographs series, 2022 (to appear).

[24] T. Kalinowski, N. Kamčev and B. Sudakov, The zero forcing number of graphs, *SIAM J. Discrete Math.* 33 (2019), 95–115.

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

[26] S. Narayanan and A. Sun, Bounds on expected propagation time of probabilistic zero forcing, *European J. Combin.* 98 (2021), Paper No. 10405, 16 pp.

[27] K. Panagiotou, X. Pérez-Giménez, T. Sauerwald and H. Sun, Randomized Rumor Spreading: the Effect of the Network Topology, *Combin. Prob. Comput.* 24 (2015), 457–479.

[28] N. Warnberg, Positive semidefinite propagation time, *Discrete Appl. Math.* 198 (2016), 274–290.

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

(Received 23 Aug 2021; revised 6 Mar 2022, 31 May 2022)