Tight Bounds on Probabilistic Zero Forcing on Hypercubes and Grids

Natalie C. Behague

Dept. of Mathematics and Statistics
University of Victoria
Victoria, BC, Canada

nbehague@uvic.ca

Trent G. Marbach   Paweł Prałat

Department of Mathematics
Ryerson University
Toronto, ON, Canada

{trent.marbach,pralat}@ryerson.ca

Abstract

Zero forcing is a deterministic iterative graph colouring process in which vertices are coloured either blue or white, and in every round, any blue vertices that have a single white neighbour force these white vertices to become blue. Here we study probabilistic zero forcing, where blue vertices have a non-zero probability of forcing each white neighbour to become blue. We explore the propagation time for probabilistic zero forcing on hypercubes and grids.

Mathematics Subject Classifications: 05C81, 05C57

1 Introduction

1.1 Definition of Zero Forcing

Zero forcing is an iterative graph colouring procedure which can model certain real-world propagation and search processes such as rumor spreading. Given a graph G and a set of marked, or blue, vertices $Z \subseteq G$, the process of zero forcing involves the application of the *zero forcing colour change rule* in which a blue vertex u forces a non-blue (white) vertex v to become blue if $N(u) \setminus Z = \{v\}$, that is, u forces v to become blue if v is the only white neighbour of u.

We say that Z is a *zero forcing set* if when starting with Z as the set of initially blue vertices, after iteratively applying the zero forcing colour change rule until no more vertices can be forced blue, the entire vertex set of G becomes blue. Note that the order in which forces happen is arbitrary since if u is in a position in which it can force v, this property will not be destroyed if other vertices are turned blue. As a result, we may process vertices sequentially (in any order), or all vertices that are ready to turn blue can

do so simultaneously. The *zero forcing number*, denoted $z(G)$, is the cardinality of the smallest zero forcing set of G.

Zero forcing has sparked a lot of interest recently. Some work has been done on calculating or bounding the zero forcing number for specific structures such as graph products [12], graphs with large girth [8] and random graphs [1, 16], while others have studied variants of zero forcing such as connected zero forcing [4] or positive semi-definite zero forcing [2].

While zero forcing is a relatively new concept (first introduced in [12]), the problem has many applications to other branches of mathematics and physics. For example, zero forcing can give insight into linear and quantum controllability for systems that stem from networks. More precisely, in [5], it was shown that for both classical and quantum control systems that can be modelled by networks, the existence of certain zero forcing sets guarantees controllability of the system.

Another area closely related to zero forcing is power domination [19]. Designed to model the situation where an electric company needs to continually monitor their power network, one method that is used is to place phase measurement units (PMUs) periodically through their network. To reduce the cost associated with this, one asks for the least number of PMUs necessary to observe a specific network fully. To be more specific, given a network modelled with a simple graph G, a PMU placed at a vertex will be able to observe every adjacent vertex. Furthermore, if an observed vertex has exactly one unobserved neighbour, this observed vertex can observe this neighbour. In this way, power domination involves an observation rule compatible with the zero forcing colour change rule.

In the present paper we are concerned with a parameter associated with zero forcing known as the *propagation time*, which is the fewest number of rounds necessary for a zero forcing set of size $z(G)$ to turn the entire graph blue. More formally, given a graph G and a zero forcing set Z, we generate a finite sequence of sets $Z_0 \subsetneq Z_1 \subsetneq \cdots \subsetneq Z_t$, where $Z_0 = Z$, $Z_t = V(G)$, and given Z_i, we define $Z_{i+1} = Z_i \cup Y_i$, where $Y_i \subseteq V(G) \setminus Z_i$ is the set of white vertices that can be forced in the next round if Z_i is the set of blue vertices in the current round. Then the propagation time of Z, denoted $pt(G, Z)$, is defined to be t. The propagation time of the graph G is then given by $pt(G) = \min_Z pt(G, Z)$, where the minimum is taken over all zero forcing sets Z of cardinality $z(G)$. The propagation time for zero forcing has been studied in [13].

1.2 Definition of Probabilistic Zero Forcing

Zero forcing was initially formulated to bound a problem in linear algebra known as the min-rank problem [12]. In addition to this application to mathematics, zero forcing also models many real-world propagation processes. One specific application of zero forcing could be to rumor spreading, but the deterministic nature of zero forcing may not fit the chaotic nature of real-life situations. As such, probabilistic zero forcing has also been proposed and studied where blue vertices have a non-zero probability of forcing white neighbours, even if there is more than one white neighbour. More specifically,

given a graph G, a set of blue vertices Z, and vertices $u \in Z$ and $v \in V(G) \setminus Z$ such that $uv \in E(G)$, in a given time step, vertex u will force vertex v to become blue with probability

$$\mathbb{P}(u \text{ forces } v) = \frac{|N[u] \cap Z|}{\deg(u)},$$

where $N[u]$ is the closed neighbourhood of u.

In a given round, each blue vertex will attempt to force each white neighbour independently. If this happens, we may say that the edge uv is forced. A vertex becomes blue as long as it is forced by at least one blue neighbour, or in other words if at least one edge incident with it is forced. Note that if v is the only white neighbour of u, then with probability 1, u forces v, so given an initial set of blue vertices, the set of vertices forced via probabilistic zero forcing is always a superset of the set of vertices forced by traditional zero forcing. In this sense, probabilistic zero forcing and traditional zero forcing can be coupled. In the context of rumor spreading, the probabilistic colour change rule captures the idea that someone is more likely to spread a rumor if many of their friends have already heard the rumor.

Under probabilistic zero forcing, given a connected graph, it is clear that starting with any non-empty subset of blue vertices will with probability 1 eventually turn the entire graph blue, so the zero forcing number of a graph is not an interesting parameter to study for probabilistic zero forcing. Initially in [17], the authors studied a parameter that quantifies how likely it is for a subset of vertices to become a traditional zero forcing set the first time-step that it theoretically could under probabilistic zero forcing.

Instead, in this paper, we will be concerned with a parameter that generalizes the zero forcing propagation time. This generalization was first introduced in [11]. Given a graph G, and a set $Z \subseteq V(G)$, let $pt_{pzf}(G, Z)$ be the random variable that outputs the propagation time when probabilistic zero forcing is run with the initial blue set Z. For ease of notation, we will write $pt_{pzf}(G, v) = pt_{pzf}(G, \{v\})$. The propagation time for the graph G will be defined as the random variable $pt_{pzf}(G) = \min_{v \in V(G)} pt_{pzf}(G, v)$. More specifically, $pt_{pzf}(G)$ is a random variable for the experiment in which n iterations of probabilistic zero forcing are performed independently, one for each vertex of G, then the minimum is taken over the propagation times for these n independent iterations.

1.3 Asymptotic Notation

Our results are asymptotic in nature, that is, we will assume that $n \to \infty$. Formally, we consider a sequence of graphs $G_n = (V_n, E_n)$ (for example, G_n is an n-dimensional hypercube or n by n grid) and we are interested in events that hold *asymptotically almost surely* (*a.a.s.*), that is, events that hold with probability tending to 1 as $n \to \infty$.

Given two functions $f = f(n)$ and $g = g(n)$, we will write $f(n) = O(g(n))$ if there exists an absolute constant $c \in \mathbb{R}_+$ such that $|f(n)| \leqslant c|g(n)|$ for all n, $f(n) = \Omega(g(n))$ if $g(n) = O(f(n))$, $f(n) = \Theta(g(n))$ if $f(n) = O(g(n))$ and $f(n) = \Omega(g(n))$, and we write $f(n) = o(g(n))$ or $f(n) \ll g(n)$ if $\lim_{n \to \infty} f(n)/g(n) = 0$. In addition, we write $f(n) \gg g(n)$ if $g(n) = o(f(n))$ and we write $f(n) \sim g(n)$ if $f(n) = (1 + o(1))g(n)$, that

is, $\lim_{n \to \infty} f(n)/g(n) = 1$.

Finally, as typical in the field of random graphs, for expressions that clearly have to be an integer, we round up or down but do not specify which: the choice of which does not affect the argument.

1.4 Results on Probabilistic Zero Forcing

In [11], the authors studied probabilistic zero forcing, and more specifically the expected propagation time for many specific structures. A summary of this work is provided in the following theorem.

Theorem 1 ([11]). *Let* $n > 2$. *Then*

- $\min_{v \in V(P_n)} \mathbb{E}(pt_{pzf}(P_n, v)) = \begin{cases} n/2 + 2/3 & \textit{if } n \textit{ is even} \\ n/2 + 1/2 & \textit{if } n \textit{ is odd,} \end{cases}$

- $\min_{v \in V(C_n)} \mathbb{E}(pt_{pzf}(C_n, v)) = \begin{cases} n/2 + 1/3 & \textit{if } n \textit{ is even} \\ n/2 + 1/2 & \textit{if } n \textit{ is odd,} \end{cases}$

- $\min_{v \in V(K_{1,n})} \mathbb{E}(pt_{pzf}(K_{1,n}, v)) = \Theta(\log n)$,

- $\Omega(\log \log n) = \min_{v \in V(K_n)} \mathbb{E}(pt_{pzf}(K_n, v)) = O(\log n)$.

In [6], the authors used tools developed for Markov chains to analyze the expected propagation time for many small graphs. The authors also showed, in addition to other things, that $\min_{v \in V(K_n)} \mathbb{E}(pt_{pzf}(K_n, v)) = \Theta(\log \log n)$ and for any connected graph G, $\min_{v \in V(G)} \mathbb{E}(pt_{pzf}(G, v)) = O(n)$. This result was then improved in [18], where the authors showed that

$$\log_2 \log_2(n) \leqslant \min_{v \in V(G)} \mathbb{E}(pt_{pzf}(G, v)) \leqslant \frac{n}{2} + o(n)$$

for general connected graphs G. In the same paper, the authors also showed that

$$\min_{v \in V(G)} \mathbb{E}(pt_{pzf}(G, v)) = O(r \log(n/r)), \tag{1}$$

where $r \geqslant 1$ denotes the radius of the connected graph G. Moreover, they provided a class of graphs for which the bounds of the theorem are tight.

In addition to the results mentioned above, the authors of [11] also considered the *binomial random graph* $\mathcal{G}(n, p)$, proving the following theorem.

Theorem 2 ([11]). *Let* $0 < p < 1$ *be constant. Then a.a.s. we have that*

$$\min_{v \in V(\mathcal{G}(n,p))} \mathbb{E}(pt_{pzf}(\mathcal{G}(n, p), v)) = O((\log n)^2).$$

However, the authors in [18] conjectured that for the random graph, a.a.s.

$$\min_{v \in V(\mathcal{G}(n,p))} \mathbb{E}(pt_{pzf}(\mathcal{G}(n,p), v)) = (1 + o(1)) \log \log n.$$

Of course, Theorem 2 can be improved immediately via (1) and the fact that for $0 < p < 1$ constant, the radius of $\mathcal{G}(n,p)$ is a.a.s. 2 (see e.g. [10]), but even with this improvement, the bound is still far from the conjectured value. In [9], the authors explored probabilistic zero forcing on $\mathcal{G}(n,p)$ in more detail and, in particular, proved the above conjecture. Their results can be summarized in the following theorem that shows that probabilistic zero forcing occurs much faster in $\mathcal{G}(n,p)$ than in a general graph G, as evidenced by the bounds of (1).

Theorem 3 ([9]). *Let* $v \in V(\mathcal{G}(n,p))$ *be any vertex of* $\mathcal{G}(n,p)$.
If $p = \log^{-o(1)} n$ *(in particular, if* p *is a constant), then a.a.s.*

$$pt_{pzf}(\mathcal{G}(n,p), v) \sim \log_2 \log_2 n.$$

On the other hand, if $\log n / n \ll p = \log^{-\Omega(1)} n$*, then a.a.s.*

$$pt_{pzf}(\mathcal{G}(n,p), v) = \Theta(\log(1/p)).$$

The results of most interest to us are from [14]. The authors of that paper are concerned with hypercube graphs Q_n and grid graphs $G_{m \times n}$, and prove the following result.

Theorem 4 ([14]). *The following bounds hold:*

- $\min_{v \in V(Q_n)} \mathbb{E}(pt_{pzf}(Q_n, v)) = O(n \log n)$,

- $(1/2 + o(1))(m + n) \leqslant \min_{v \in V(G_{m \times n})} \mathbb{E}(pt_{pzf}(G_{m \times n}, v)) \leqslant (4 + o(1))(m + n)$.

1.5 Our Results

In this paper, we improve both results from [14]. Instead of considering the expectation of the propagation time, we will calculate bounds on the propagation time that a.a.s. hold. With some more work one can use our approaches to show that the same bounds hold for expectation, hence improving the results from [14], but we are more interested in statements that hold a.a.s. The lower bounds for both families of graphs are trivial so the improvement is for the upper bounds.

Theorem 5. *The following bounds hold a.a.s.:*

(a) $pt_{pzf}(Q_n) \sim n$,

(b) $(1 + o(1))(m + n)/2 \leqslant pt_{pzf}(G_{m \times n}) \leqslant (1 + 10^{-7} + o(1))(m + n)/2$.

The bounds for Q_n are asymptotically tight. Unfortunately, we could not prove the matching upper bound for $G_{m \times n}$ but our upper bound is very tight, and so it supports the conjecture that a.a.s. $pt_{pzf}(G_{n \times n}) \sim n$. This conjecture is supported by independent simulations performed in [14] as well as our own.

2 Preliminaries

2.1 Chernoff inequality

Let us first state a specific instance of Chernoff's bound that we will find useful. Let $X \in \mathrm{Bin}(n, p)$ be a random variable with the binomial distribution with parameters n and p. Then, a consequence of *Chernoff's bound* (see e.g. [15, Corollary 2.3]) is that

$$\mathbb{P}(|X - \mathbb{E}X| \geqslant \varepsilon \mathbb{E}X) \leqslant 2 \exp\left(-\frac{\varepsilon^2 \mathbb{E}X}{3}\right) \tag{2}$$

for $0 < \varepsilon < 3/2$. Moreover, let us mention that the bound holds for the general case in which $X = \sum_{i=1}^{n} X_i$ and $X_i \in \mathrm{Bernoulli}(p_i)$ with (possibly) different p_i (again, e.g. see [15] for more details).

2.2 Azuma–Hoeffding inequality

Let $X_0, X_1, \ldots$ be an infinite sequence of random variables that make up a martingale; that is, for any $a \in \mathbb{N}$ we have $\mathbb{E}[X_a | X_{a-1}] = X_{a-1}$. Suppose that there exist constants $c_a > 0$ such that $|X_a - X_{a-1}| \leqslant c_a$ for each $a \leqslant t$. Then, the *Azuma–Hoeffding inequality* implies that for every $b > 0$,

$$\mathbb{P}\Big(\exists i (0 \leqslant i \leqslant t) : |X_i - X_0| \geqslant b\Big) \leqslant 2 \exp\left(-\frac{b^2}{2 \sum_{a=1}^{t} c_a^2}\right). \tag{3}$$

2.3 Chernoff–Heoffding bounds for Markov chains

This result is due to Chung, Lam, Liu and Mitzenmacher [7]. Let M be a discrete time ergodic Markov chain with state space $[n]$ and the stationary distribution π. M may be interpreted as either the chain itself or the corresponding n by n transition matrix. The *total variation distance* between u and w, two distributions over $[n]$, is defined as

$$\|u - w\|_{TV} = \max_{A \subseteq [n]} \left| \sum_{i \in A} u_i - \sum_{i \in A} w_i \right|.$$

For any $\varepsilon > 0$, the *mixing time* of Markov chain M is defined as

$$T(\varepsilon) = \min\left\{t \in \mathbb{N} : \max_{x} \left\|xM^t - \pi\right\|_{TV} \leqslant \varepsilon\right\},$$

where x is an arbitrary initial distribution. We will also need a definition of the π-norm. The *inner product under the π-kernel* is defined as $\langle u, v \rangle_\pi = \sum_{i \in [n]} u_i v_i / \pi_i$. Then, the π-*norm* of u is defined as $\|u\|_\pi = \sqrt{\langle u, u \rangle_\pi}$. Note that, in particular, $\|\pi\|_\pi = 1$.

Now we are ready to state the main result from [7]. Let $T = T(\varepsilon)$ be the ε-mixing time of Markov chain M for $\varepsilon \leqslant 1/8$. Let $(V_1, V_2, \ldots, V_t)$ denote a t-step random walk on M starting from an initial distribution φ on $[n]$. For every $i \in [t]$, let $f_i : [n] \to [0, 1]$ be

a weight function at step i such that the expected weight satisfies $\mathbb{E}_{v \leftarrow \pi}[f_i(v)] = \mu$ for all i. Define the total weight of the walk by $X = \sum_{i=1}^{t} f_i(V_i)$. There exists some constant c (which is independent of μ, δ and ϵ) such that for $0 \leqslant \delta \leqslant 1$,

$$\mathbb{P}\Big((1-\delta)\mu t \leqslant X \leqslant (1+\delta)\mu t\Big) \geqslant 1 - c\,\|\varphi\|_\pi \exp\left(-\delta^2 \mu t/(72T)\right). \tag{4}$$

2.4 Useful Coupling

We will be using coupling to simplify both upper and lower bounds. Having said that, in our case we will only need to prove upper bounds. Indeed, for lower bounds, it might be convenient to make some white vertices blue at some point of the process. Similarly, for upper bounds, one might want to make some blue vertices white. Given a graph G, $S, T \subseteq V(G)$, and $\ell \in \mathbb{N}$, let $A(S, T, \ell)$ be the event that starting with blue set S, after ℓ rounds every vertex in T is blue. The following simple observation was proved in [9].

Lemma 6 ([9]). *For all sets $S_1 \subseteq S_2 \subseteq V(G)$, $T \subseteq V(G)$, and $\ell \in \mathbb{N}$,*

$$\mathbb{P}(A(S_1, T, \ell)) \leqslant \mathbb{P}(A(S_2, T, \ell)).$$

3 Hypercubes

This section is devoted to proving part (a) of Theorem 5. Let us start with a formal definition of the hypercube. The *n-dimensional hypercube* Q_n has vertex set consisting of all binary strings of length n and there is an edge between two vertices if and only if their binary strings differ in exactly one bit. For $0 \leqslant k \leqslant n$, *level k* of the hypercube Q_n is defined to be the set of all vertices whose binary strings contain exactly k ones. Note that each vertex in level k has k neighbours in level $k-1$ and $n-k$ neighbours in level $k+1$.

Proof of Theorem 5(a). Trivially, for any vertex $v \in V(Q_n)$, we deterministically have that $pt_{pzf}(Q_n, v) \geqslant n$. Hence, in order to show that $pt_{pzf}(Q_n) \sim n$, it is enough to show that a.a.s. $pt_{pzf}(Q_n, v) \leqslant n + o(n)$ for some vertex $v \in V(Q_n)$. Let v be the vertex $(0, 0, \ldots, 0)$ on level 0; in fact, since Q_n is a vertex-transitive graph, v can be any vertex.
    We will use the following Lemma.

Lemma 7. *Suppose that for some integer k, $0 \leqslant k < n/\ln^2 n$, the following property $\mathcal{P}(k)$ holds at some point of the process: all vertices at levels up to k are blue (including level k). Then with probability $1 - o(1/n)$ property $\mathcal{P}(k+1)$ holds after an additional $O(\ln n)$ rounds.*

    Since trivially $\mathcal{P}(0)$ holds, by the union bound (over $n/\ln^2 n$ possible values of k), using this Lemma we will conclude that a.a.s. $\mathcal{P}(n/\ln^2 n)$ holds after $O(n/\ln n) = o(n)$ rounds.

Proof of Lemma 7. Suppose that property $\mathcal{P}(k)$ holds from some k, $0 \leqslant k < n/\ln^2 n$. By Lemma 6, we may assume that all vertices at level $k+1$ are white. It will be convenient to independently consider 3 phases after which property $\mathcal{P}(k+1)$ will hold with probability $1 - o(1/n)$. The first phase lasts $320 \ln n$ rounds. The probability that a given vertex at level $k+1$ stays white during this phase is at most

$$\left(1 - \frac{k+1}{n} \right)^{320 \ln n} \leqslant \exp\left(-\frac{320(k+1)\ln n}{n} \right) \leqslant 1 - \frac{160(k+1)\ln n}{n} =: p,$$

since $k \leqslant n/\ln^2 n$. Hence, for a given vertex w at level k, the number of neighbours at level $k+1$ that turned blue during this phase is stochastically lower bounded by $\mathrm{Bin}(n-k, 1-p) \geqslant \mathrm{Bin}(n/2, 1-p) \ni X$ with $\mathbb{E}[X] = (1-p)n/2 = 80(k+1)\ln n$. It follows from Chernoff's bound (2) (applied with $\varepsilon = 1/2$) that w has at most $40(k+1)\ln n$ blue neighbours at level $k+1$ with probability at most

$$2 \exp\left(-\frac{\mathbb{E}[X]}{12} \right) = \frac{2}{n^{(80/12)(k+1)}} = o\left(\frac{1}{n^{k+1}} \right).$$

By the union bound (over at most n^k vertices at level k), with probability $1 - o(1/n)$, each vertex at level k has at least $40(k+1)\ln n$ blue neighbours at level $k+1$ at the end of the first phase. As we aim for a statement that holds with probability $1 - o(1/n)$, we may assume that this property holds once we enter the second phase.

The second phase lasts $\log_{5/4} n$ rounds. As before, let us concentrate on a given vertex w at level k. Suppose that w has ℓ blue neighbours for some integer ℓ such that $40(k+1)\ln n \leqslant \ell \leqslant n/2$ (w has only k neighbours at level $k-1$ so, of course, it includes neighbours at level $k+1$). The number of white neighbours of w (at level $k+1$) that turned blue in one round of the process is stochastically lower bounded by $\mathrm{Bin}(n-\ell, (\ell+1)/n) \geqslant \mathrm{Bin}(n/2, (\ell+1)/n) \ni Y$ with $\mathbb{E}[Y] = (\ell+1)/2 \geqslant 20(k+1)\ln n$. We get from Chernoff's bound (2) (applied with $\varepsilon = 1/2$) that $Y \leqslant \frac{\ell+1}{4}$ with probability at most

$$2 \exp\left(-\frac{\mathbb{E}[Y]}{12} \right) = \frac{2}{n^{(20/12)(k+1)}} = o\left(\frac{1}{n^{k+1}\ln n} \right).$$

By the union bound (over at most n^k vertices at level k and at most $\log_{5/4} n$ rounds), with probability $1 - o(1/n)$, each vertex at level k increases the number of blue neighbours by a multiplicative factor of $5/4$ each round, reaching at least $n/2$ blue neighbours by the end of the second phase. We may assume that this property holds once we enter the third phase.

The third (and last) phase lasts $3 \log_2 n$ rounds. This time, let us concentrate on a given white vertex w at level $k+1$. This vertex has $k+1$ neighbours at level k, each of which has at least $n/2$ blue neighbours. Hence, vertex w stays white by the end of this phase with probability at most

$$\left(\left(\frac{1}{2} \right)^{k+1} \right)^{3\log_2 n} = 2^{-3(k+1)\log_2 n} = \frac{1}{n^{3(k+1)}} = o\left(\frac{1}{n^{k+2}} \right).$$

By the union bound (over at most n^{k+1} vertices at level $k+1$), with probability $1-o\left(1/n\right)$, all vertices at level $k+1$ turn blue by the end of this phase and so property $\mathcal{P}(k+1)$ holds. $\qquad\square$

Since we aim for a statement that holds a.a.s., we may assume that $\mathcal{P}(n/\ln^2 n)$ holds after $o(n)$ rounds, and continue the process from there. We say that a vertex at layer k is *happy* if all but at most $\ell = \ell(n) = \ln^5 n$ neighbours at layer $k-1$ are blue. Similarly, a vertex at layer k is *very happy* if not only it is happy but also all but at most ℓ neighbours at layer $k+1$ are blue. (Note that a happy or even a very happy vertex might still be white.) Trivially, all vertices at levels up to $n/\ln^2 n + 1$ are happy (including level $n/\ln^2 n + 1$), and all vertices at levels before level $n/\ln^2 n$ are very happy.

Suppose that for some integer k, $n/\ln^2 n \leqslant k \leqslant n-1$, all vertices at levels up to k are happy (including level k). We will show that after one single round all vertices at layer $k+1$ are going to be happy and all vertices at layer $k-1$ are going to be very happy with probability $1 - o\left(1/n\right)$. By the union bound (over all possible values of k), we will get that a.a.s. after less than n rounds all vertices of the hypercube are going to be very happy.

For simplicity, by Lemma 6 we may assume that all vertices at level k are white (despite the fact that they are happy). Let us concentrate on a given vertex w at level k. Since w is happy and all of its blue neighbours at level $k-1$ are happy too, the probability that w stays white is at most

$$\left(1 - \frac{k-\ell}{n}\right)^{k-\ell} \leqslant \exp\left(-(1+o(1))\frac{k^2}{n}\right) =: p.$$

Now, the probability that a given vertex at level $k+1$ is *not* happy is at most

$$\binom{k+1}{\ell}p^\ell \leqslant n^\ell \exp\left(-(1+o(1))\frac{k^2}{n}\ell\right)$$
$$= \exp\left(\ell\ln n - (1+o(1))\frac{k^2}{n}\ell\right)$$
$$\leqslant \exp\left(\ln^6 n - (1+o(1))n\ln n\right) \qquad \text{using } k \geqslant \frac{n}{\ln^2 n}$$
$$= o\left(\frac{1}{2^n\, n}\right).$$

Then the property that all vertices at level $k+1$ are happy holds with probability $1-o(1/n)$ by the union bound (over at most 2^n vertices at level $k+1$).

We may also show that all vertices at level $k-1$ are very happy with probability $1 - o(1/n)$. If $k \geqslant n-\ell$, then all vertices at level $k-1$ are trivially very happy as they have less than ℓ neighbours at level k, so there is nothing to show. If $k \leqslant n-\ell$ (but still $k \geqslant n/\ln^2 n$), then a given vertex at level $k-1$ is *not* very happy with probability at most

$$\binom{n-(k-1)}{\ell}p^\ell \leqslant n^\ell \exp\left(-(1+o(1))\frac{k^2}{n}\ell\right) = o\left(\frac{1}{2^n\, n}\right),$$

and the desired bound holds by the union bound (over at most 2^n vertices at level $k-1$).

At this point we may assume that all vertices of the hypercube are very happy. It is easy to see that a.a.s. all remaining white vertices will turn blue in one single round. Indeed, since all vertices are very happy, the probability that some vertex stays white is at most

$$
\begin{aligned}
2^n \left(1 - \frac{n - 2\ell}{n}\right)^{n-2\ell} &= 2^n \left(\frac{2\ell}{n}\right)^{(1+o(1))n} \\
&= \exp\left(O(n) - (1 + o(1))n \ln n\right) \\
&= o(1).
\end{aligned}
$$

The proof is finished. □

4 Grids

This section is devoted to proving part (b) of Theorem 5. The theorem is stated for grids but we also consider zero-forcing on tori, which we introduce both for convenience and because we get for free the same result on tori as for grids. Let us start with a formal definition of grids and tori. The *m by n grid graph* $G_{m \times n}$ is the mn vertex graph that is the Cartesian product of a path on m vertices and a path on n vertices. The *m by n torus graph* $T_{m \times n}$ is the Cartesian product of a cycle on m vertices and a cycle on n vertices.

We can restrict our focus to the square grid $G_{n \times n}$ or the square torus graph $T_{n \times n}$. For a non-square grid $G_{n \times m}$ with $n < m$, once a central $n \times n$ square is all blue, the two adjacent columns turn blue with probability one on the next time-step and so on. Thus $pt(G_{n \times m}) \leqslant pt(G_{n \times n}) + \lceil \frac{m-n}{2} \rceil$. A similar argument holds for the non-square torus. Hence, it will be straightforward to generalize the results for asymmetric cases which we will do once we are done with symmetric cases.

We define the *origin* of $G_{n \times n}$ to be the central vertex if n is odd and one of the four central vertices if n is even (the choice can be made arbitrarily as all of them are the same up to symmetry). Since $T_{n \times n}$ is vertex transitive, we may fix any vertex of $T_{n \times n}$ to be the origin of $T_{n \times n}$. For a given positive integer $k < \frac{n}{2}$ we define the *k-principal-square* to be the $2k + 1$ by $2k + 1$ sub-grid centred on the origin.

Let $s = \lfloor n/\ln^2 n \rfloor$. We will work in phases, where in phase i we start with all vertices in an is-principal-square being blue and end with all vertices in an $(i + 1)s$-principal-square being blue. For simplicity, by Lemma 6 we may assume that at the beginning of each phase the only vertices that are blue are the ones that belong to the corresponding principal-square. In fact, during each phase we will turn a few more vertices white (if needed) so that the process behaves more predictably. We aim to bound the number of time-steps it takes to go from the is-principal-square to the $(i + 1)s$-principal square. As a result, in total there will be at most $\frac{1}{2} \ln^2 n$ phases. Note that the only phase that potentially differs between the grid and the torus is the final phase.

Let us start with the following simple observation that will allow us to ignore a few initial and final phases.

Lemma 8. *Let $k \geqslant 1$. Suppose that at time t_0 all vertices in a k-principal-square are blue. Then with probability $1 - o(1/n)$ at time $t_0 + 6 \ln n$ all vertices in a $(k+1)$-principal-square are blue.*

Proof. Let $k \geqslant 1$, and suppose that at time t_0 all vertices in a k-principal-square are blue. Note that vertices adjacent to a blue vertex with three blue neighbours turn blue with probability 1, and so all vertices adjacent to a non-corner vertex of the k-principal-square will turn blue at time $t_0 + 1$. There are twelve (eight if $k = 1$) vertices in the $(k+1)$-principal square that do not turn blue deterministically: the four corner vertices and the eight (four if $k = 1$) vertices adjacent to the corners. We will call these *corner–adjacent*.

A corner–adjacent vertex has a blue neighbour at time t_0 and so the probability that it stays white in one time-step is at most $3/4$. The probability it remains white after $4 \ln n$ steps is at most $\left(\frac{3}{4}\right)^{4 \ln n}$ and so the probability that all corner–adjacent vertices are blue at time $t_0 + 4 \ln n$ is at least $1 - 8 \left(\frac{3}{4}\right)^{4 \ln n} = 1 - o(1/n)$.

Conditioning on a corner vertex having two blue neighbours at time $t_1 = t_0 + 4 \ln n$, the probability it stays white after the next time-step is at most $(3/4)^2 = 9/16$. The probability it is white after a further $2 \ln n$ time-steps is at most $\left(\frac{9}{16}\right)^{2 \ln n}$. In particular, the probability that all four corner vertices are blue at time $t_1 + 2 \ln n$ steps is at least $1 - 4 \left(\frac{9}{16}\right)^{2 \ln n} = 1 - o(1/n)$. $\qquad\square$

Lemma 8 tells us that with high probability any single phase takes $O(6s \ln n) = o(n)$ time-steps, so we may safely ignore what happens in the first five phases and the last phase. In particular, our argument is the same whether we are run the process on the grid or the torus. Hence, we may focus on the square grid graph $G_{n \times n}$.

Recall that $s = \lfloor n / \ln^2 n \rfloor$. Let $5 \leqslant i < n/2s$ and suppose that at time t_0 all vertices in the (is)-principal-square are blue. By Lemma 6, we may assume that these are the only blue vertices at that point of the process. We will show that with probability $1 - o(1/n)$, at time $t_0 + (1 + \varepsilon + o(1))s$ all vertices in the $(i+1)s$-principal-square are blue, where $\varepsilon > 0$ will be an explicit, very small constant that we are not ready to introduce yet.

For simplicity, we identify the vertices of $G_{n \times n}$ with pairs (a, b) from the set

$$\left\{ -\left\lfloor \frac{n-1}{2} \right\rfloor, \ldots, \left\lfloor \frac{n}{2} \right\rfloor \right\} \times \left\{ -\left\lfloor \frac{n-1}{2} \right\rfloor, \ldots, \left\lfloor \frac{n}{2} \right\rfloor \right\}$$

in the natural way with the origin at $(0, 0)$. In particular, note that a k-principal-square contains all vertices (a, b) with $|a| \leqslant k$ and $|b| \leqslant k$. We will focus on the top-right quadrant where both coordinates are positive; by symmetry, the argument for the other quadrants will be exactly the same.

In order to control how the process progresses with time, we will pay attention to a few blue vertices at the top-right corner that are at the same distance from the origin. Hence, the following definition will be useful. For fixed positive integer d, we define a *d-window (rooted at vertex $(a-d+1, b)$)* to be a d-tuple of vertices

$$\Big((a-d+1, b), (a-d+2, b-1), (a-d+3, b-2), \ldots, (a, b-d+1) \Big),$$

where $d < a, b < n/2$. Note that the distance of all vertices from such a d-window from the origin is $a + b - d + 1$.

Blue vertices from a d-window at distance ℓ from the origin will most likely turn some other vertices at distance $\ell + 1$ from the origin to be blue. If that happens, we will simply move the window appropriately and continue the process. In order to compute the transition probabilities between given configurations that might occur in a d-window, it will be convenient to assume that each blue vertex in the window has exactly two blue neighbours, namely the bottom and the left neighbours. We may do so based on the following observation.

Lemma 9. *Suppose that at time t, a $k \times k$ sub-grid of vertices are all blue for some integer $k \geqslant 2$. Let v be the top-right corner (blue) vertex of this sub-grid. If a neighbour of v turns blue at time $t + 1$, then it is the top-right corner of a $(k - 1) \times (k - 1)$ sub-grid of vertices that are all blue.*

Proof. Consider the four vertices adjacent to v. Clearly, the vertex below v and the vertex to the left of v are each the top right corner of a $(k - 1) \times (k - 1)$ all blue grid. Let u be the vertex to the right of v. The column of $k - 2$ vertices directly below u are all blue at time $t + 1$ with probability 1 (as each is adjacent to a blue vertex with three blue neighbours at time t). Thus, if u turns blue at time $t + 1$, then it is the top-right corner of a $(k - 1) \times (k - 1)$ blue sub-grid—see Figure 1. An analogous argument holds for the vertex above v. $\square$
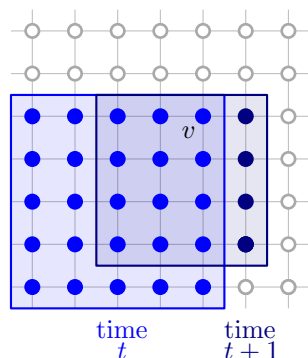

Figure 1: A 5×5 blue sub-grid at time t and a 4×4 blue sub-grid at time $t + 1$ as described in Lemma 9.

Now, we are ready to define an *auxiliary* process, which monitors the behaviour of the original process, giving a sequence of triples (D_j, C_j, t_j) of d-windows D_j at times t_j, and associated binary d-tuples C_j indicating which vertices in the d-window are blue at time t_j and which ones are white (1 represents blue and 0 represents white). This process will control the expansion of the blue principal squares and so can be used to upper bound the number of rounds needed to reach the end of a given phase. We will run this process for at most $3s$ steps and stop it *prematurely* if the end of the phase is not reached after $3s$ rounds. However, we will show that we do *not* stop prematurely with probability

$1 - o\left(1/n\right)$ and, in fact, when d is large enough, with this probability we will finish in almost s rounds which is a trivial lower bound for the number of rounds needed to finish a given phase.

(1) We start with $D_0 = ((is - d + 1, is), \ldots, (is, is - d + 1))$ and time-step t_0. Since all vertices in D_0 are blue, the corresponding configuration is $C_0 = (1, \ldots, 1)$. (In fact, the initial triple is not important. Another natural starting point would be to start with a d-window including vertex (is, is) and so would have only one blue vertex in the corresponding configuration.)

(2) Given the triple (D_j, C_j, t_j), we define triple $(D_{j+1}, C_{j+1}, t_{j+1})$ differently according to whether C_j contains a blue vertex or not. Suppose that D_j is a d-window rooted at vertex $(a - d + 1, b)$.

(3) If C_j contains a blue vertex, let $t_{j+1} = t_j + 1$ and consider the d-windows

$$
\begin{aligned}
X &= ((a - d + 1, b + 1), \ldots, (a, b - d + 2)) \quad \text{and} \\
Y &= ((a - d + 2, b), \ldots, (a + 1, b - d + 1)).
\end{aligned}
$$

Let D_{j+1} be whichever of X or Y has more blue vertices at time t_{j+1} (see Figure 2 for an illustration). If they have an equal number of blue vertices, then pick one of the two uniformly at random. Let C_{j+1} be the corresponding configuration capturing the information about which vertices in the window are blue at time t_{j+1}. Go to step (2).
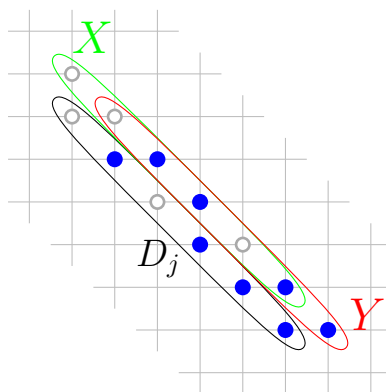

Figure 2: A 6-window D_j and 6-windows X, Y as defined in Step (3) of the auxiliary process. Since Y has more blue vertices, $D_{j+1} = Y$.

(4) If C_j is all white (that is, all vertices in D_j are white at time t_j), we set $t_{j+1} = t_j$. In this situation we need to introduce an *auxiliary* triple $(D_{j+1}, C_{j+1}, t_{j+1})$ that corresponds to an earlier d-window that is not all white at time t_j. To that end we chose D_{j+1} to have the same distance from the origin as D_{j-1}. Note that C_{j-1} is not all zeros and represents a configuration in the d-window D_{j-1} at time t_{j-1}; indeed,

the process is designed in such a way that no C_{j-1} and C_j can be all zeros at any step in this process. We may simply fix $D_{j+1} = D_{j-1}$ and $C_{j+1} = C_{j-1}$ but our goal is to create a memory-less Markov chain so we will follow a different strategy.

Let (x, y) be a random vertex in D_{j-1} that was blue at time t_{j-1}. Let D_{j+1} be the d-window centred on (x, y). To be specific, if d is odd then let

$$D_{j+1} = \left(\left(x - \frac{d-1}{2}, y + \frac{d-1}{2} \right), \ldots, \left(x + \frac{d-1}{2}, y - \frac{d-1}{2} \right) \right),$$

and if d is even then let D_{j+1} be one of

$$\left(\left(x - \frac{d-2}{2}, y + \frac{d-2}{2} \right), \ldots, \left(x + \frac{d}{2}, y - \frac{d}{2} \right) \right)$$

and

$$\left(\left(x - \frac{d}{2}, y + \frac{d}{2} \right), \ldots, \left(x + \frac{d-2}{2}, y - \frac{d-2}{2} \right) \right)$$

chosen uniformly at random. At time t_{j-1}, (x, y) is blue and it follows from Lemma 9 that $(x - 1, y)$ and $(x, y - 1)$ are blue and have three blue neighbours. (See the discussion below for details about implications of Lemma 9.) This means at time $t_{j+1} = t_j = t_{j-1} + 1$, the vertices $(x - 1, y + 1), (x, y), (x + 1, y - 1)$ are each blue and in D_{j+1}. Applying Lemma 6, we may assume that any vertices in D_{j+1} that are not one of $(x - 1, y + 1), (x, y), (x + 1, y - 1)$ are white at time $t_{j+1} = t_j$. Let us stress again that by our choice of (x, y), D_{j+1} is not all white at time t_{j+1}. Go to step (2).

Recall that we start the current phase at time t_0 with all vertices in the (is)-principal-square being blue. Moreover, we already dealt with the first few phases so $i \geqslant 5$. By this assumption, every vertex in the starting d-window D_0 is the top-right corner of a $(2is + 1 - d) \times (2is + 1 - d)$ blue sub-grid at time t_0. Applying Lemma 9 recursively for each j, we see that if a vertex in D_j turns blue at time t_j because of its blue neighbour in D_{j-1}, then it is the top-right corner of a $(2is + 1 - d - j) \times (2is + 1 - d - j)$ blue sub-grid. In particular, since we run the auxiliary process for at most $3s$ steps, it will be the top-right corner of a 3×3 blue sub-grid.

We are now ready to investigate the transition probabilities between configurations C_j and guide the auxiliary process so that it yields a Markov chain. Suppose that $D_j = ((a - d + 1, b), \ldots, (a, b - d + 1))$. Let us first investigate step (4) of the auxiliary process. By the assumption of this step, C_j is not all zeros; that is, D_j contains a blue vertex at time t_j. By Lemma 6 we may assume that at time t_j all vertices in the next d-windows $X = ((a - d + 1, b + 1), \ldots, (a, b - d + 2))$ and $Y = ((a - d + 2, b), \ldots, (a + 1, b - d + 1))$ are white, and that the vertices $(a - d, b + 1)$ and $(a + 1, b - d)$ are white. Moreover, as discussed above, Lemma 9 implies that any vertex in D_j that is blue at time t_j has exactly two blue neighbours. Thus the probability that a vertex in $X \cup Y$ turns blue at time $t_j + 1$ is exactly $1 - (1/4)^2 = 15/16$ if it has two blue neighbours in D_j at time t_j, $1 - 1/4 = 3/4$ if it

has one blue neighbour in D_j at time t_j, and 0 otherwise. In particular, the configuration C_{j+1} representing the state of D_{j+1} at time t_{j+1} depends only on the configuration C_j representing the state of D_j at time t_j. The transition probability $P_d(C, C')$ between any two possible configurations C and C' can be easily computed.

Now, let us investigate step (4). This is set up so that if C_j is all white (that is, D_j contains no blue vertices at time t_j) then deterministically C_{j+1} consists of three centred consecutive blue vertices and all other vertices white. In particular, for $d \geqslant 3$ odd and C all white, $P_d(C, C') = 1$ if $C' = (\underbrace{0, \ldots, 0}_{(d-3)/2}, 1, 1, 1, \underbrace{0, \ldots, 0}_{(d-3)/2})$ and 0 otherwise. For $d \geqslant 4$ even and C all white,

$$
P_d(C, C') = \begin{cases} 1/2 & \text{if } C' = (\underbrace{0, \ldots, 0}_{(d-4)/2}, 1, 1, 1, \underbrace{0, \ldots, 0}_{(d-2)/2}) \text{ or } (\underbrace{0, \ldots, 0}_{(d-2)/2}, 1, 1, 1, \underbrace{0, \ldots, 0}_{(d-4)/2}) \\ 0 & \text{otherwise.} \end{cases}
$$

Finally, for the degenerate case $d = 2$ and C all white, $P_2(C, C') = 1$ if C' is all blue, and 0 otherwise.

Since the state of D_{j+1} at time t_{j+1} (namely, configuration C_{j+1}) depends only on the state of D_j at time t_j (namely, configuration C_j) independent of the time and the choice of vertices, we can capture the behaviour of the process using a Markov chain $(C_j)_{j=0}^{\infty}$ on state space $\Omega = \{0, 1\}^d$. For example, when $d = 2$ the Markov chain has transition matrix

$$
P_2 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1/16 & 9/32 & 3/32 & 9/16 \\ 1/16 & 3/32 & 9/32 & 9/16 \\ 1/256 & 15/256 & 15/256 & 225/256 \end{pmatrix}
$$

where the states are in the order $(0, 0), (0, 1), (1, 0), (1, 1)$. It is easy to see that for any d this Markov chain is irreducible and aperiodic, and so it has a limiting distribution which is equal to the stationary distribution. We define μ_d to be the probability of the all white state $C = (0, 0, \ldots, 0)$ in the limiting distribution.

The next lemma is the crux of our argument.

Lemma 10. *Take $s = \lfloor n/\ln^2 n \rfloor$. Let $5 \leqslant i < n/(2s)$ and suppose that at time t_0 all vertices in the (is)-principal-square are blue. With probability $1 - o(1/n)$, before the end of round $t_0 + (1 + o(1))\frac{1 - \mu_d}{1 - 2\mu_d} 2s$ all vertices in the $(i+1)s$-principal-square are blue, where μ_d is the probability of the all white state $C = (0, 0, \ldots, 0)$ in the limiting stationary distribution of the Markov's chain P_d defined above.*

Applying the Lemma together with earlier observations we immediately obtain the following corollary.

Corollary 11. *The following bound holds a.a.s.:*

$$
pt(G_{n \times n}) \leqslant (1 + o(1))\frac{1 - \mu_d}{1 - 2\mu_d} n,
$$

where μ_d is the probability of the all white state $C = (0, 0, \dots, 0)$ in the limiting stationary distribution of the Markov's chain P_d defined above.

Proof of Lemma 10. Let us fix an integer $d \geqslant 2$. Suppose we run the auxiliary process (D_j, C_j, t_j) monitoring the trajectory of the d-window for ℓ steps. Let w be the number of steps we spend in the all white state $C = (0, 0, \dots, 0)$. We have that $t_\ell = t_0 + \ell - w$, since when we are in the all white state $t_{j+1} = t_j$, and $t_{j+1} = t_j + 1$ otherwise. The distance of D_0 from the origin is $2(is) - d + 1$. The distance of D_ℓ from the origin is $2(is) - d + 1 + \ell - 2w$, since when we are in the all white state D_{j+1} is one closer to the origin than D_j, and otherwise D_{j+1} is one further from the origin than D_j.

Applying Chernoff–Heoffding bounds for Markov chains (equation (4)), we have that

$$\mathbb{P}\Big((1-\delta)\mu_d\ell \leqslant w \leqslant (1+\delta)\mu_d\ell\Big) \geqslant 1 - c\exp\left(-c\delta^2\ell\right)$$

for $0 \leqslant \delta \leqslant 1$, where c is some positive constant depending only on the Markov chain (in particular, independent of ℓ and δ). Set

$$\ell = \left(\frac{2}{1-2\mu_d} + \frac{\ln^4 n}{\sqrt{n}}\right)s \geqslant s \quad \text{and} \quad \delta = \frac{\ln^2 n}{\sqrt{n}}.$$

(Note that $\mu_d \leqslant 1/77$ implies that $\ell < 3s$, which is the number of steps taken. We will show that μ_d satisfies this property below and so we may assume that the process does *not* stop prematurely.) Then, using that $s = \lfloor n/\ln^2 n \rfloor$, we have that with probability $1 - o(1/n)$ the distance of D_ℓ from the origin is at least

$$2is - d + 1 + \ell - 2(1+\delta)\mu_d\ell$$

$$= 2is + O(1) + (1 - 2\mu_d)\left(1 - \Theta\left(\frac{\ln^2 n}{\sqrt{n}}\right)\right)\ell$$

$$= 2is + O(1) + (1 - 2\mu_d)\left(1 - \Theta\left(\frac{\ln^2 n}{\sqrt{n}}\right)\right)\frac{2}{1-2\mu_d}\left(1 + \Theta\left(\frac{\ln^4 n}{\sqrt{n}}\right)\right)s$$

$$= 2(i+1)s + \Theta(\sqrt{n}\ln^2 n),$$

and the number of rounds in this phase that passed in the original zero-forcing process is equal to

$$t_\ell - t_0 = \ell - w \leqslant (1 - (1-\delta)\mu_d)\ell = (1-\mu_d)\left(1 + \Theta\left(\frac{\ln^2 n}{\sqrt{n}}\right)\right)\ell$$

$$= (1 + o(1))\frac{1-\mu_d}{1-2\mu_d}2s.$$

It remains to show that the d-windows D_j are travelling broadly North-East rather than North or East, in order to be sure of obtaining a principal square as opposed to a rectangle. To that end, let us define the discrepancy $\mathrm{disc}(D)$ of a d-window $D = (a-d+1, b), \dots, (a, b-d+1)$ to be $a - b$. The discrepancy captures how far from the

North-East diagonal the d-window is shifted. The discrepancy of D_0 is 0. By definition, $\mathrm{disc}\,(D_{j+1})$ differs from $\mathrm{disc}\,(D_j)$ by at most one if D_j is not all white at time t_j, and differs from $\mathrm{disc}\,(D_j)$ by at most d if D_j is all white at time t_j.

Define a sequence $(j_k)_{k\geqslant 0}$ where $j_0 = 0$ and for $k > 0$, j_k is the least $j > j_{k-1}$ such that D_j is all white at time t_j (that is, $C_j = (0, 0, \ldots, 0)$). For any j for which $C_j \neq (0, 0, \ldots, 0)$, the probability that $C_{j+1} = (0, 0, \ldots, 0)$ is at least $(1/16)^{d+1}$. Hence, the probability that we do not hit an all white state after $2 \cdot 16^{d+1} \ln n$ consecutive steps, is at most

$$\left(1 - (1/16)^{d+1}\right)^{2 \cdot 16^{d+1} \ln n} \leqslant \exp(-2 \ln n) = 1/n^2.$$

Thus, since there are at most $\ell = o(n)$ terms in the sequence, we have $j_k - j_{k-1} \leqslant 2 \cdot 16^{d+1} \ln n$ for all k with probability $1 - o\,(1/n)$. Since we aim for a statement that holds with probability $1 - o\,(1/n)$, we may assume that this property is deterministically satisfied.

Consider the sequence $(\mathrm{disc}\,(D_{j_k}))_{k\geqslant 0}$. Based on our assumption, for all $k \geqslant 1$

$$|\,\mathrm{disc}\,(D_{j_k}) - \mathrm{disc}\,(D_{j_{k-1}})\,| \leqslant d + t_k - t_{k-1} \leqslant \left(2 \cdot 16^{d+1} + o(1)\right) \ln n \leqslant 3 \cdot 16^{d+1} \ln n.$$

By symmetry of the auxiliary process, we know that given $\mathrm{disc}\,(D_{j_{k-1}})$ and integer a the probability that $\mathrm{disc}\,(D_{j_k}) = \mathrm{disc}\,(D_{j_{k-1}}) + a$ must be the same as the probability that $\mathrm{disc}\,(D_{j_k}) = \mathrm{disc}\,(D_{j_{k-1}}) - a$. Thus the sequence $(\mathrm{disc}\,(D_{j_k}))_{k\geqslant 0}$ is a martingale. We can apply the Azuma–Hoeffding inequality (3) with $b = \sqrt{\ell} \ln^2 n$ and $2 \cdot 16^{d+1} \ln^2 n \leqslant j_k \leqslant \ell$ to see that

$$\mathbb{P}\left(|\,\mathrm{disc}\,(D_{j_k})\,| \geqslant b\right) \leqslant 2\exp\left(\frac{-b^2}{2 \cdot k \cdot 3 \cdot 16^{d+1} \ln^2 n}\right) = 2\exp\left(\frac{-\ell \ln^4 n}{O(\ell \ln^2 n)}\right) = o\,(1/n),$$

and so

$$\mathbb{P}\left(|\,\mathrm{disc}\,(D_\ell)\,| \leqslant \sqrt{\ell} \ln^2 n + 3 \cdot 16^{d+1} \ln n\right) \geqslant 1 - o\,(1/n),$$

where the additional term $3 \cdot 16^{d+1} \ln n$ needs to be added because the last term in the sequence $(j_k)_{k\geqslant 0}$ can be smaller than ℓ but, with the desired probability, not smaller than $\ell - 3 \cdot 16^{d+1} \ln n$.

Putting all of this together, we conclude that with probability $1 - o\,(1/n)$, at time $t_\ell = t_0 + (1 + o(1))\frac{1-\mu_d}{1-2\mu_d}2s$ there is a blue vertex at distance $2(i+1)s + \Theta\left(\sqrt{n} \ln^2 n\right)$ from the origin with a discrepancy of less than $\sqrt{\ell} \ln^2 n = O\left(\sqrt{n} \ln n\right) = o\left(\sqrt{n} \ln^2 n\right)$. Using Lemma 9 for the last time, we get that this blue vertex is the top-right corner of a $(2is + 1 - d - \ell) \times (2is + 1 - d - \ell)$ all blue sub-grid at time t_ℓ. Since $i \geqslant 5$, we get that

$$2is + 1 - d - \ell \geqslant (i+1)s + 4s + 1 - d - 3s > (i+1)s,$$

and so this all blue sub-grid entirely contains the top-right quadrant of the $(i + 1)s$-principal-square.

By symmetry, the same conclusion holds for the other three quadrants and so with probability $1 - o\,(1/n)$ at time $t_0 + (1 + o(1))\frac{1-\mu_d}{1-2\mu_d}2s$ the $(i+1)s$-principal-square is entirely blue. The proof of the lemma is finished. $\qquad\square$

The only thing remaining to finish the main theorem is to calculate the stationary distribution of the Markov chain and thus μ_d. Based on Corollary 11, each value of μ_d implies that a.a.s.,

$$pt\left(G_{n\times n}\right) \leqslant \left(1 + \varepsilon_d + o(1)\right)n, \quad \text{where} \quad \varepsilon_d := \frac{1-\mu_d}{1-2\mu_d} - 1 = \frac{\mu_d}{1-2\mu_d}.$$

When $d = 2$ or $d = 3$ the transition matrix is small enough to be calculated by hand and one can check that $\mu_2 = 1/77$ and $\mu_3 = 1861/491117$, respectively. It gives us $\varepsilon_2 = 1/75 \leqslant 0.01334$ and $\varepsilon_3 = 1861/487395 \leqslant 0.003819$. For $d \leqslant 7$, one can use a computer to calculate the exact fractions in the transition matrix and thus the exact values of μ_d. This gives

$$\mu_4 = \frac{11439524}{9092101243}$$
$$\mu_5 = \frac{1133763610798567}{2542177028478096119}$$
$$\mu_6 = \frac{1126668271831162358892325831063}{6861272362648644090193985407497 61}$$
$$\mu_7 = \frac{53677808692824989928312388350730928714869303434556 5}{86637916450461736904089899318924922661986521038146 70581}.$$

For larger values of d, we must move to numerical approximations. Rounding errors become a concern for $d > 14$ when the minimum entry in the transition matrix is close to the precision of the computer. Below we summarize these numerical values in the table. In particular, when $d = 14$ we obtain $\varepsilon_{14} < 10^{-7}$ and the main theorem holds.

d	μ_d	ε_d
2	0.012987012987012988	0.013333333333333334
3	0.0037893210782766634	0.0038182582915294574
4	0.0012581826460420552	0.001261356680213082
5	0.0004459813766302923	0.0004463795305453566
6	0.00016420690103551534	0.00016426084656466706
7	$6.1956486134 \cdot 10^{-5}$	$6.1964164298 \cdot 10^{-5}$
8	$2.3776197997 \cdot 10^{-5}$	$2.3777328666 \cdot 10^{-5}$
9	$9.2381456535 \cdot 10^{-6}$	$9.2383163433 \cdot 10^{-6}$
10	$3.6235531968 \cdot 10^{-6}$	$3.6235794573 \cdot 10^{-6}$
11	$1.4319129399 \cdot 10^{-6}$	$1.4319170406 \cdot 10^{-6}$
12	$5.6925354755 \cdot 10^{-7}$	$5.6925419565 \cdot 10^{-7}$
13	$2.2742611942 \cdot 10^{-7}$	$2.2742622287 \cdot 10^{-7}$
14	$9.1236746477 \cdot 10^{-8}$	$9.1236763126 \cdot 10^{-8}$

Acknowledgment

were obtained using the Julia language [3]. We would like to thank Bogumił Kamiński from SGH Warsaw School of Economics for helping us to implement it. The program is available on-line.[1]

References

[1] D. Bal, P. Bennett, S. English, C. MacRury, and P. Prałat. Zero-forcing in random regular graphs. *Journal of Combinatorics*, 12(1):85–116, 2021.

[2] F. Barioli, W. Barrett, S. Fallat, H. 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(2):401–411, 2010.

[3] J. Bezanson, A. Edelman, S. Karpinski, and V.B. Shah. Julia: A fresh approach to numerical computing. *SIAM review*, 59(1):65–98, 2017.

[4] B. Brimkov and I. Hicks. Complexity and computation of connected zero forcing. *Discrete Appl. Math.*, 229:31–45, 2017.

[5] D. Burgarth, D. D'Alessandro, L. Hogben, S. Severini, and M. Young. Zero forcing, linear and quantum controllability for systems evolving on networks. *IEEE Trans. Automat. Control*, 58(9):2349–2354, 2013.

[6] Y. Chan, E. Curl, J. Geneson, L. Hogben, K. Liu, I. Odegard, and M. Ross. Using markov chains to determine expected propagation time for probabilistic zero forcing. *Electronic Journal of Linear Algebra*, 36:318–333, 2020.

[7] K. Chung, H. Lam, Z. Liu, and M. Mitzenmacher. Chernoff-hoeffding bounds for markov chains: Generalized and simplified. In *29th International Symposium on Theoretical Aspects of Computer Science*, page 124, 2012.

[8] R. Davila and F. Kenter. Bounds for the zero forcing number of graphs with large girth. *Theory Appl. Graphs*, 2(2):Art. 1, 8, 2015.

[9] S. English, C. MacRury, and P. Prałat. Probabilistic zero forcing on random graphs. *European J. Combin.*, 91:103207, 22, 2021.

[10] A. Frieze and M. Karoński. *Introduction to random graphs*. Cambridge University Press, Cambridge, 2016.

[11] J. Geneson and L. Hogben. Propagation time for probabilistic zero forcing. arXiv:1812.10476, 2018.

[12] AIM Minimum Rank-Special Graphs Work Group. Zero forcing sets and the minimum rank of graphs. *Linear Algebra Appl.*, 428(7):1628–1648, 2008.

[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(13-14):1994–2005, 2012.

[14] D. Hu and A. Sun. Probabilistic zero forcing on grid, regular, and hypercube graphs. arXiv:2010.12343, 2020.

---

[1]https://math.ryerson.ca/~pralat/research.html#publications

[15] S. Janson, T. Łuczak, and A. Rucinski. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

[16] T. Kalinowski, N. Kamčev, and B. Sudakov. The zero forcing number of graphs. *SIAM J. Discrete Math.*, 33(1):95–115, 2019.

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

[18] S. Narayanan and A. Sun. Bounds on expected propagation time of probabilistic zero forcing. *European Journal of Combinatorics*, 98:103405, 2021.

[19] M. Zhao, L. Kang, and G. Chang. Power domination in graphs. *Discrete Math.*, 306(15):1812–1816, 2006.