Quasirandom Rumor Spreading

BENJAMIN DOERR, École Polytechnique, Palaiseau, France
TOBIAS FRIEDRICH, Friedrich-Schiller-Universität Jena, Germany
THOMAS SAUERWALD, University of Cambridge, United Kingdom

9

We propose and analyze a quasirandom analogue of the classical push model for disseminating information in networks ("randomized rumor spreading"). In the classical model, in each round, each informed vertex chooses a neighbor at random and informs it, if it was not informed before. It is known that this simple protocol succeeds in spreading a rumor from one vertex to all others within $\mathcal{O}(\log n)$ rounds on complete graphs, hypercubes, random regular graphs, Erdős-Rényi random graphs, and Ramanujan graphs with probability $1-o(1)$. In the quasirandom model, we assume that each vertex has a (cyclic) list of its neighbors. Once informed, it starts at a random position on the list, but from then on informs its neighbors in the order of the list. Surprisingly, irrespective of the orders of the lists, the above-mentioned bounds still hold. In some cases, even better bounds than for the classical model can be shown.

Categories and Subject Descriptors: F.2.2 [**Analysis of Algorithms and Problem Complexity**]: Nonnumerical Algorithms and Problems—*Computations on discrete structures*; G.2.2 [**Discrete Mathematics**]: Graph Theory—*Graph algorithms*

Additional Key Words and Phrases: Distributed computing, rumor spreading, broadcasting, quasirandomness, random graphs, expander

ACM Reference Format:
Benjamin Doerr, Tobias Friedrich, and Thomas Sauerwald. 2014. Quasirandom rumor spreading. ACM Trans. Algor. 11, 2, Article 9 (October 2014), 35 pages.
DOI: http://dx.doi.org/10.1145/2650185

1. INTRODUCTION

Randomized rumor-spreading or *random phone call protocols* are simple randomized epidemic algorithms designed to distribute a piece of information in a network. They build on the basic paradigm that informed vertices call random neighbors to inform them (*push model*) or that uninformed vertices call random neighbors to become informed if the neighbor is informed (*pull model*). Despite the simple concept, these algorithms succeed in distributing information extremely quickly. In contrast to many

---

Part of this work was done while Tobias Friedrich and Thomas Sauerwald were postdoctoral fellows at ICSI Berkeley supported by the German Academic Exchange Service (DAAD) or research associates at Max-Planck-Institut für Informatik. Benjamin Doerr was partially supported by project DO 749/4 in the German Research Foundation's (DFG) Priority Program (SPP) 1307.
Parts of the results appeared in the 19th ACM-SIAM Symposium on Discrete Algorithms (SODA '08) [Doerr et al. 2008] and the 36th International Colloquium on Automata, Languages and Programming (ICALP '09) [Doerr et al. 2009].
Authors' addresses: B. Doerr, LIX, Campus de l'École Polytechnique, 1, rue Honoré d'Estienne d'Orves, Bâtiment Alan Turing, CS35003, 91120 Palaiseau, France; email: doerr@lix.polytechnique.fr; T. Friedrich, Friedrich-Schiller-Universität Jena, Ernst-Abbe-Platz 2, 07743 Jena, Germany; email: friedrich@uni-jena.de; T. Sauerwald, Computer Laboratory, William Gates Building, 15 JJ Thomson Avenue, Cambridge CB3 0FD, United Kingdom; email: thomas.sauerwald@cl.cam.ac.uk.

ACM 1549-6325/2014/10-ART9 $15.00
DOI: http://dx.doi.org/10.1145/2650185

natural deterministic approaches, they are also highly robust against transmission failures [Feige et al. 1990; Karp et al. 2000; Elsässer and Sauerwald 2009].

Such algorithms have been applied successfully both in the context where a single item of news has to be distributed from one processor to all others [Hedetniemi et al. 1988] and in the case where news may be injected at various vertices at different times. The latter problem occurs when maintaining data integrity in distributed databases (e.g., name servers in large corporate networks [Demers et al. 1988; Kempe et al. 2003]). For a more extensive, but still concise discussion of various central aspects of this area, we refer the reader to the paper by Karp et al. [2000].

1.1. Randomized Rumor Spreading

Rumor-spreading protocols often assume that all vertices have access to a central clock. The protocols then proceed in rounds, in each of which each vertex, independent of the others, can perform certain actions. In the classical randomized rumor-spreading protocols, in each round, each vertex contacts a neighbor chosen independently and uniformly at random. In the push model, which we focus on here, this results in the contacted vertex becoming informed, provided it was not already. Because all communications are done independently at random, in the following description, we call this the *fully random model* to distinguish it from the quasirandom one we propose in this article. The first graphs for which the fully random model was analyzed are complete graphs [Frieze and Grimmett 1985; Pittel 1987]. Pittel [1987] proved that with probability $1-o(1)$, $\log_2 n+\ln n+f(n)$ rounds suffice, where $f(n)$ can be any function tending to infinity.

Feige et al. [1990] showed that on almost all random graphs $\mathcal{G}(n,p)$, $p \geqslant (1+\varepsilon)\log n/n$, the fully random model runs in $\mathcal{O}(\log n)$ time with probability $1-n^{-1}$. They also showed that this failure probability can be achieved for $p = (\log n + \mathcal{O}(\log\log n))/n$ only in $\Omega(\log^2 n)$ rounds. In addition, Feige et al. [1990] also considered hypercubes and proved a runtime bound of $\mathcal{O}(\log n)$ with probability $1-n^{-1}$.

For expanders, in which the maximum and minimum degree satisfy $\Delta/\delta = \mathcal{O}(1)$, it was shown in Sauerwald [2010] that the fully random model completes its broadcast campaign in $\mathcal{O}(\log n)$ rounds with probability $1-n^{-1}$ (similar results were shown earlier [Boyd et al. 2006; Mosk-Aoyama and Shah 2006], but these hold only for the push-pull model). Recently, Fountoulakis et al. [2010] and Fountoulakis and Panagiotou [2010] derived precise bounds on the runtime for random and pseudorandom regular graphs, extending the result of Frieze and Grimmett [1985] for complete graphs.

Demers et al. [1988] and Karp et al. [2000] introduced the push-pull model, which combines push and pull transmissions. For this model, Chierichetti et al. [2010a, 2010b] and Giakkoupis [2011] proved tight runtime bounds in terms of the conductance. In particular, for any graph with constant conductance and arbitrary degree distribution, a runtime bound of $\mathcal{O}(\log n)$ was shown in Giakkoupis [2011].

Rumor spreading has recently been studied intensively on social networks modeled by random graphs that have a power-law degree distribution. Chierichetti et al. [2011] showed that the push model with nonvanishing probability needs $\Omega(n^{\alpha})$ rounds on preferential attachment graphs [Barabási and Albert 1999] for some $\alpha > 0$. For such power-law networks, however, the push-pull strategy is much better than push or pull alone. With this strategy, $\mathcal{O}(\log n)$ rounds suffice with high probability [Doerr et al. 2011a]. Doerr et al. [2011a] further proved that, for a slightly adjusted process, where contacts are chosen uniformly at random among all neighbors except the one that was chosen just in the round before, $\mathcal{O}(\log n/\log\log n)$ rounds suffice. This is asymptotically optimal because the diameter of such preferential attachment graphs, with power-law exponent 3, is $\Theta(\log n/\log\log n)$ [Bollobás et al. 2001]. Fountoulakis et al. [2012] showed that push-pull requires $\Omega(\log n)$ on Chung-Lu random graphs [Chung and Lu 2002] with power-law exponent > 3, whereas for power-law exponent $\in (2, 3)$, the rumor spreads to almost all nodes in time $\Theta(\log \log n)$ rounds with high probability.

1.2. Our Results

In this work, we propose a quasirandom analogue of the randomized rumor-spreading algorithm. In this *quasirandom model*, every vertex is equipped with a cyclic list of its neighbors. If a vertex becomes informed, then, in the next round, it chooses a position on the list uniformly at random and informs the neighbor corresponding to this position. In the subsequent rounds, the vertex continues sending out messages in the order of its list. Clearly, by introducing these dependencies, we gain some natural advantages like the fact that an informed vertex does not call a neighbor a second time before having called all neighbors once. In consequence, we obtain an absolute guarantee that after $\Delta \operatorname{diam}(G)$ rounds, all vertices are informed (see Theorem 3.1), thus improving over the corresponding $O(\Delta(\operatorname{diam}(G) + \log n))$ bound of [Feige et al. 1990] for the fully random model.

Surprisingly, we do not observe that the newly introduced dependencies are harmful. More precisely, we show that the $\mathcal{O}(\log n)$ bound (valid with probability $1 - n^{-1}$) for complete graphs, hypercubes, random graphs, random regular graphs, and Ramanujan graphs in the classical protocol also holds in the quasirandom model *regardless of which lists are used*. In addition to its theoretical interest, this implies that in an implementation of the quasirandom protocol, one may reuse any lists that are already present (e.g., to encode the network structure).

Our $\mathcal{O}(\log n)$ runtime bound also applies to very sparse connected random graphs with $p = (\log n + \omega(1))/n$. This contrasts with a lower bound of $\Omega(\log^2 n)$ steps required by the fully random model to inform all vertices with probability $1 - n^{-1}$ Feige et al. [1990, Theorem 4.1] and with a lower bound on the expected time of $\Omega(\log n \log \log n)$ shown in this article. Similarly for hypercubes, we show that the quasirandom model completes in $\mathcal{O}(\log n)$ rounds with probability $1 - n^{-\Omega(\log n)}$, whereas the fully random model is easily seen to require $\Omega(\log^2 n)$ steps to achieve the same probability of success. The interesting aspect of these improvements is not so much their actual magnitude, but rather that they can be achieved for free by using a very natural protocol. Note also that speedups not visible by asymptotic analyses have been observed; see the experimental analysis in Doerr et al. [2011b]. For example, the quasirandom protocol was seen to be around 10% faster on the hypercube on 4,096 vertices and around 15% faster on random 12-regular graphs on 4,096 vertices.

To prove the results in this article, we need to cope with the more dependent random experiments. Recall that once a vertex has sent out a message, all its future transmissions are determined. The methods we develop to cope with these difficulties (e.g., suitably delaying independent random decisions so as to have enough independent randomness at certain moments to allow the use of Chernoff-type inequalities) might be useful in the analysis of other dependent settings as well.

Our analysis employs a certain graph class called *expanding graphs*, which are defined by three natural expansion properties. Roughly speaking, these properties require that small sets of vertices have many neighbors; that for large sets of vertices, the external vertices have many neighbors in the set; and, finally, that the vertex degrees are of similar order (see Definition 4.1 for details). This graph class has been used by other authors (e.g., Cooper et al. [2012]). We prove that complete graphs, random graphs, random regular graphs, and Ramanujan graphs are expanding. After that, we show that the quasirandom model succeeds in $\mathcal{O}(\log n)$ rounds on every expanding graph with probability $1 - n^{-\gamma}$, where $\gamma > 0$ is an arbitrary constant.

1.3. Related Work on Quasirandomness

We call an algorithm quasirandom if it imitates (or achieves in an even better way) a particular property of a randomized algorithm deterministically. The concept of quasirandomness occurs in several areas of mathematics and computer science. Prominent examples are low-discrepancy point sets and quasi-Monte Carlo methods [Niederreiter 1992], which imitate the property of a random point set to be evenly distributed in their domain.

Our quasirandom rumor-spreading protocol imitates two properties of the fully random counterpart; namely, that over a short period of time, a vertex does not contact neighbors twice and, over a long period of time, it calls all neighbors roughly equally often.

This is very much related to a quasirandom analogue of the classic random walk, which is also known as Eulerian walker [Priezzhev et al. 1996], edge ant walk [Wagner et al. 1999], whirling tour [Dumitriu et al. 2003], Propp machine [Kleber 2005; Cooper and Spencer 2006], and deterministic random walk [Cooper et al. 2007b; Doerr and Friedrich 2009]. Unlike in a random walk, in a quasirandom walk, each vertex serves its neighbors in a fixed order. The resulting (completely deterministic) walk nevertheless closely resembles a random walk in several respects [Cooper and Spencer 2006; Doerr and Friedrich 2009; Cooper et al. 2007b, 2010; Friedrich and Sauerwald 2010]. Other algorithmic applications of the idea of quasirandom walks are autonomous agents patroling a territory [Wagner et al. 1996], external mergesort [Barve et al. 1997], and iterative load balancing [Friedrich et al. 2010].

1.4. Results Obtained after This Work

Subsequent to the conference versions [Doerr et al. 2008, 2009] and during the preparation of this journal version, the following results appeared that answer some questions left open in this work. In Angelopoulos et al. [2009], it is proven that with probability $1 - o(1)$, the quasirandom model succeeds in informing all vertices of a complete graph on n vertices in $(1 + o(1))(\log_2 n + \ln n)$ rounds. Hence, for the complete graph, the quasirandom model achieves the same runtime as the fully random one [Frieze and Grimmett 1985] up to lower order terms. This was strengthened by Fountoulakis and Huber [2009], who nearly showed that Pittel's bounds [Pittel 1987] also hold for the quasirandom model—their upper and lower bounds deviate by only a $\Theta(\log\log n)$ term.

A second important aspect of broadcasting protocols is their robustness. The fully random model, due to its high use of independent randomness, is usually considered to be very robust. See Karp et al. [2000] and Elsässer and Sauerwald [2009] for some results in this direction. A very precise result, valid for both the fully random and the quasirandom model, was recently given in Doerr et al. [2013]. They consider the setting that each message reaches its destination only with an (independently sampled) probability of $0 < p < 1$. Again, for the complete graph on n vertices, they show that both protocols succeed in $(1 + o(1))\,(\log_{1+p} n + p^{-1} \ln n)$ rounds with probability $1 - o(1)$. Together with a corresponding lower bound for the fully random model, this shows that both models are equally robust against transmission failures, in spite of the greatly reduced use of independent randomness in the quasirandom model.

The question of how much randomness is needed in such protocols was first considered by Doerr and Fouz [2011] and Giakkoupis and Woelfel [2011]. Among other results, the latter work presents a variant of the quasirandom model that requires on average only $\mathcal{O}(\log\log n)$ instead of $\mathcal{O}(\log n)$ random bits per vertex in order to spread the rumor

Table I. Upper and Lower Bounds on the Broadcast Time
These times hold with probability at least $1-1/n$ for different graph classes in the fully random and the quasirandom model. More detailed analyses for sparse random graphs can be found in Table II on page 24.

	Broadcast Time	
Graph Class	Fully Random Model	Quasirandom Model
all graphs	$\mathcal{O}(\Delta\,(\mathrm{diam}(G)+\log n))$ [Feige et al. 1990]	$\leqslant \Delta\ \mathrm{diam}(G)$ (Thm. 3.1)
	$\leqslant 12n\log n$ [Feige et al. 1990]	$\leqslant 2n-3$ (Thm. 3.1)
Complete k-ary trees	$\Theta(k\log n)$ (Thm. 4.15)	$\Theta(k\log n/\log k)$ (Thm. 4.15)
Hypercubes	$\Theta(\log n)$ [Feige et al. 1990]	$\Theta(\log n)$ (Thm. 7.1)
Complete graphs	$\Theta(\log n)$ [Pittel 1987; Frieze and Grimmett 1985]	$\Theta(\log n)$ (Thm. 4.2 and 5.1)
Ramanujan	$\Theta(\log n)$ [Giakkoupis 2011]	$\Theta(\log n)$ (Thm. 4.9 and 5.1)
Almost all random graphs with fixed deg. seq.	$\Theta(\log n)$ [Giakkoupis 2011]	$\Theta(\log n)$ (Thm. 4.12 and 5.1)
Almost all random graphs $G(n,p)$ with $pn=\log n+\omega(1)$, $pn=\log n+\mathcal{O}(\log\log n)$	$\Theta(\log^2 n)$ Feige et al. [1990, Thm. 4.1]	$\Theta(\log n)$ (Thm. 4.2 and 5.1)
Almost all random graphs $G(n,p)$ with $pn=c\log n$, $c>1$	$\Theta(\log n)$ [Feige et al. 1990]	$\Theta(\log n)$ (Thm. 4.2 and 5.1)

in $\mathcal{O}(\log n)$ rounds on a complete graph with probability $1-n^{-\Omega(1)}$. Giakkoupis et al. [2012] present two protocols that are based on hashing and pseudorandom generators, respectively. Although these protocols only require a logarithmic number of random bits in total on many networks, they are more complicated; for instance, they require that random bits are appended to the rumor.

In order to bound the number of messages, Berenbrink et al. [2010] analyze another variant of the quasirandom model based on the combination of push and pull calls. This variant is shown to succeed in $\mathcal{O}(\log n)$ rounds on random graphs and hypercubes, while requiring only $\mathcal{O}(n\log\log n)$ messages on random graphs and $\mathcal{O}(n(\log\log n)^2)$ on hypercubes (all these results hold with probability $1-n^{-1}$).

The worst-case behavior of the quasirandom model was very recently addressed by Baumann et al. [2012]. Among other results, the authors present a polynomial-time algorithm to compute the configuration of lists and initial neighbors that maximizes the time to spread the rumor.

1.5. Organization

The rest of this article is organized as follows. In Section 2, we describe our model more formally and introduce some basic notation. In Section 3, we derive bounds on the broadcast time that hold for all graphs. After that, in Section 4, we describe the class of graphs we consider in this work. The runtime analysis of quasirandom rumor spreading on this graph class is deferred to Section 5. To highlight the efficiency of our new quasirandom model, we also derive some lower bounds for the fully random model in Section 6. In Section 7, we analyze the quasirandom model on hypercubes. We close in Section 8 with a brief summary of our results.

2. PRECISE MODEL AND PRELIMINARIES

Our aim is to spread a rumor in an undirected graph $G=(V,E)$. Let always $V=\{1,\dots,n\}$ and n be the number of vertices. In the quasirandom model, each vertex $v\in V$ is equipped with a cyclic permutation $\pi_v:\Gamma(v)\rightarrow\Gamma(v)$ of its neighbors $\Gamma(v)$. We call this its list of neighbors.

The quasirandom rumor-spreading process then works as follows. In time step 0, an arbitrary vertex s is informed initially. If a vertex v becomes informed in time step t, then, in time step $t+1$, it contacts one of its neighbors w, chosen uniformly at random. From then on, it respects the order of the list; that is, in time step $t+1+\tau$, $\tau \in \mathbb{N}$, it contacts vertex $\pi_v^{\tau}(w)$. To simplify the analysis, we assume that every vertex never stops contacting its neighbors. However, it is easily seen that the propagation of the rumor is exactly the same as if every vertex v stops contacting its neighbors $\deg(v)$ rounds after it got informed. We denote by I_t the set of vertices that are informed at the end of time step t.

Note that the assumption that the initial vertex contacted first by an informed vertex is chosen uniformly at random is crucial for the quasirandom protocol. If the adversary was allowed to specify the initial vertices also, then the time to inform all vertices could take up to $n-1$ steps, for example, on a complete graph.

In the remainder of this article, it will be convenient to consider a model equivalent to the quasirandom model. This model uses the so-called *ever-rolling lists assumption*, in which we assume that vertices contact neighbors at all times, informing the neighbors (if the vertex is informed itself). Hence, here each vertex v, already at the start of the protocol, chooses a neighbor i_v uniformly at random from $\Gamma(v)$. This is the neighbor it contacts at time $t=1$. In each following time step $t=2,3,\ldots$, the vertex v contacts the vertex $\pi_v^{t-1}(i_v)$ and informs it, if it was not yet informed and if v is informed at that time (here, π_v^{t-1} is the $(t-1)$-th composition of π with itself).

From the viewpoint of how the information spreads, the model with the ever-rolling lists assumption yields a process equivalent to the standard quasirandom rumor-spreading model. Hence, in the remainder of this article, we are always discussing the model with ever-rolling lists unless we say otherwise.

We next analyze how long it takes until a rumor known to a single vertex is spread to all other vertices. We adopt a worst-case view in that we aim at bounds that are independent of the starting vertex and of all lists present in the model. This suggests the following definitions.

Definition 2.1. Let $G=(V,E)$ be a graph and $s \in V$. Then, by R_s, we denote the random variable describing the first time t at which the random rumor-spreading process started in the vertex s leads to all vertices being informed. Let $\mathcal{R}(G)$ be the (unique) minimal integer-valued random variable that dominates all R_s ; that is, for every $s \in V$ and $t \in \mathbb{N}$ it holds that

$$\mathbf{Pr}\left[\mathcal{R}(G) \geqslant t\right] \geqslant \mathbf{Pr}\left[R_s \geqslant t\right].$$

We call $\mathcal{R}(G)$ the broadcast time of the randomized rumor-spreading protocol on the graph G[1].

Let $\mathcal{L}=(\pi_v)_{v\in V}$ be a family of lists. By $Q_{\mathcal{L},s}$ we denote the (random) first time that the quasirandom rumor-spreading protocol with lists $\mathcal{L}$ started in s succeeds in informing all vertices. Let $\mathcal{Q}(G)$ be the (unique) minimal integer-valued random variable that dominates all $Q_{\mathcal{L},s}$; that is, for every family of lists $\mathcal{L}$, every $s \in V$ and $t \in \mathbb{N}$, it holds that

$$\mathbf{Pr}\left[\mathcal{Q}(G) \geqslant t\right] \geqslant \mathbf{Pr}\left[Q_{\mathcal{L},s} \geqslant t\right].$$

We call $\mathcal{Q}(G)$ the broadcast time of the quasirandom rumor-spreading protocol on the graph G.

[1]In order to see that $\mathcal{R}(G)$ is well-defined, note that for every t there exists one vertex $s=s(t)$ such that $\mathbf{Pr}\left[R_s(G) \geqslant t\right]$ is maximized. Then, we let $\mathcal{R}(G)$ satisfy $\mathbf{Pr}\left[\mathcal{R}(G) \geqslant t\right] = \mathbf{Pr}\left[R_s(G) \geqslant t\right]$. Doing this for all integers $t \in \mathbb{N}$ yields a sequence $\{\mathbf{Pr}\left[\mathcal{R}(G) \geqslant t\right] : t \in \mathbb{N}\}$ of nonincreasing values in $[0,1]$. Hence, $\mathbf{Pr}\left[\mathcal{R}(G) = t\right] := \mathbf{Pr}\left[\mathcal{R}(G) \geqslant t\right] - \mathbf{Pr}\left[\mathcal{R}(G) \geqslant t+1\right]$ completes the definition of $\mathcal{R}(G)$.

In the analysis, it will often be convenient to assume that after receiving the rumor, a vertex does not pass it on for a certain number of time steps (*delaying*). Also, it will be helpful to ignore all messages that certain vertices send out from a certain time onward (*ignoring*). Because we assumed all random decisions done by the vertices before the start of the protocol (ever-rolling list assumption), an easy induction shows that any delaying and ignoring assumptions (possibly even relying on the random choices done by the vertices that have not been active yet) for each vertex can only increase the round in which it becomes informed. In consequence, these assumptions can only increase the time needed to inform all vertices. More precisely, the random variable describing the broadcast time of any model with delaying and ignoring assumptions dominates the original one (see Definition A.1 for the precise definition of stochastic domination).

LEMMA 2.1. *For all possible delaying and ignoring assumptions, the random variable describing the broadcast time of the quasirandom model with these assumptions is stochastically larger than the broadcast time of the true quasirandom model.*

We use both delaying and ignoring to reduce the number of dependencies in the analysis. We do this by splitting the analysis into *phases*. All vertices that receive the rumor within this phase (*newly informed vertices*) are assumed to delay their actions until the beginning of the next phase. From this next phase on, all messages from vertices that previously sent out messages are ignored. Thus, we start each phase with only newly informed vertices acting. Because they have not actively participated in the rumor-spreading process, the first neighbors to which they send the rumor are chosen independently.

We also need chains of contacting vertices. That is, we say a vertex $u_1 \in V$ *reaches* another vertex $u_m \in V$ within the time interval $[a, b]$ if there is a path $(u_1, u_2, \ldots, u_m)$ in G and $t_1 < t_2 < \cdots < t_{m-1} \in [a, b]$, such that for all $j \in [1, m-1]$, $\pi_{u_j}^{t_j - 1}(i_{u_j}) = u_{j+1}$. For a vertex $w \in V$, we denote by $U_{[a,b]}(w)$ the set of vertices that reach w within the time interval $[a, b]$.

Other Notation

Throughout the article, we use the following graph-theoretical notation. For a vertex v of a graph $G = (V, E)$, let $\Gamma(v) := \{u \in V : \{u, v\} \in E\}$ be the set of its *neighbors* and $\deg(v) := |\Gamma(v)|$ its *degree*. For any $S \subseteq V$, let $\deg_S(v) := |\Gamma(v) \cap S|$. For any $S_1, S_2 \subseteq V$, let $E(S_1, S_2) := \{(u, v) \in E? : u \in S_1 \wedge v \in S_2\}$. Let $\delta := \min_{v \in V} \deg(v)$ be the *minimum degree*, $d := 2|E|/n$ the *average degree*, and $\Delta := \max_{v \in V} \deg(v)$ the *maximum degree*. The *distance* $\mathrm{dist}(x, y)$ between vertices x and y is the length of a shortest path from x to y. The *diameter* $\mathrm{diam}(G)$ of a connected graph G is the largest distance between two vertices in G. We also use $\Gamma^k(u) := \{v \in V\colon \mathrm{dist}(u, v) = k\}$ and $\Gamma^{\leqslant k}(u) := \{v \in V\colon \mathrm{dist}(u, v) \leqslant k\}$. For sets S, we define $\Gamma(S) := \{v \in V : \exists u \in S, \{u, v\} \in E\}$ as the set of neighbors of S. The complement of a set S is denoted $S^c := V \setminus S$.

All logarithms $\log n$ are natural logarithms to the base e. Because we are only interested in the asymptotic behavior, we sometimes assume that n is sufficiently large.

3. QUASIRANDOM RUMOR SPREADING ON GENERAL GRAPHS

In this section, we prove two bounds for the broadcast time valid for all graphs. The corresponding upper bounds for the fully random model are $\mathcal{O}(\Delta\,(\mathrm{diam}(G) + \log n))$ and $12n \log n$, both satisfied with probability $1 - 1/n$ [Feige et al. 1990].

THEOREM 3.1. *For any graph $G = (V, E)$, the broadcast time of the quasirandom model is at most*

(1) $\Delta \cdot \mathrm{diam}(G)$ *with probability 1, and*
(2) $2n - 3$ *with probability 1.*

Proof. Let u be the vertex initially informed.

Let $v \in V$ and $P = (u = u_0, u_1, \ldots, u_\ell = v)$ be a shortest path from u to v. Clearly, for all $i \leqslant \ell$, u_i becomes informed at most $\deg(u_{i-1}) \leqslant \Delta$ time-steps after u_{i-1} became informed. Claim *(i)* follows.

To prove claim *(ii)*, again let $v \in V$ and let $P = (u = u_0, u_1, \ldots, u_\ell = v)$ be a shortest path from u to v. Let w be a vertex not lying on P. Then, as observed already in Feige et al. [1990], w has at most three neighbors on P, and these are contained in $\{u_{i-1}, u_i, u_{i+1}\}$ for some $i < \ell$. If w has exactly three neighbors u_{i-1}, u_i, u_{i+1} on P, we call it a counterfeit of u_i (as u_i and w have, apart from each other, the same neighbors on P). Denote by $C(u_i)$ the set of counterfeits of u_i. Without loss of generality, we may choose P in such a way that for all $i < \ell$, u_i is informed no later than any of its counterfeits.

Note also that any vertex u_i on the path has only u_{i-1} and u_{i+1} (if existent) as neighbors on the path.

Let t_i denote the time that vertex u_i becomes informed. Then, $t_0 = 0$. By definition of our algorithm and choice of P, we have $t_1 \leqslant t_0 + |\Gamma(u_0) \setminus C(u_1)| = t_0 + |\Gamma(u_0) \setminus P| + 1 - |C(u_1)|$. For $2 \leqslant i \leqslant \ell - 1$, similarly, we have $t_i \leqslant t_{i-1} + |\Gamma(u_{i-1}) \setminus C(u_i)| = t_{i-1} + |\Gamma(u_{i-1}) \setminus P| + 2 - |C(u_i)|$. Finally, $t_\ell \leqslant t_{\ell-1} + |\Gamma(u_{\ell-1}) \setminus P| + 2$. We conclude

$$t_\ell \leqslant \sum_{i=0}^{\ell-1} |\Gamma_V(u_i) \setminus P| - \sum_{i=1}^{\ell-1} |C(u_i)| + 2\ell - 1.$$

Now each vertex w not lying on P can contribute at most 2 to the above expression (if it has three neighbors on P, then it is also a counterfeit). Hence, $t_\ell \leqslant 2(n - \ell - 1) + 2\ell - 1 = 2n - 3$. □

It is easy to verify that for a path of length $n - 1$ there are lists and initial vertices such that $2n - 3$ rounds are needed. Hence, the second bound is tight. The first bound is matched by k-ary trees (up to constant factors), as shown in Section 4.3, where we also demonstrate that the quasirandom model is faster than the fully random one on these graphs.

4. GRAPH CLASSES

Our results cover hypercubes, many expander graphs, random regular graphs, and Erdős-Rényi random graphs. The three latter graph classes have three properties in common, which we refer to as "expanding." This allows us to examine quasirandom rumor spreading on them from a higher level just using these three properties defined in Section 4.1.

4.1. Expanding Graphs

In order to analyze our quasirandom rumor-spreading model for a larger class of graphs at once, we distill three simple properties of graphs, which are satisfied by several common graph classes. Given these three properties, we can later prove in Theorem 5.1 that quasirandom rumor spreading successfully informs all vertices in a logarithmic runtime. Roughly speaking, these properties concern the vertex expansion of not too large subsets **(P1)**, the edge expansion **(P2)** and the regularity of the graph **(P3)**.

Definition 4.1 (*expanding graphs*). We call a connected graph *expanding* if the following properties hold:

(P1). For any constant C_α with $0 < C_\alpha \leqslant d/2$ there is a constant $C_\beta \in (0, 1)$ such that, for any connected subset $S \subseteq V$ with $3 \leqslant |S| \leqslant C_\alpha (n/d)$, it holds that $|\Gamma(S) \setminus S| \geqslant C_\beta \, d \, |S|$.

(P2). There are constants $C_\delta \in (0, 1)$ and $C_\omega > 0$ such that for any subset $S \subseteq V$, the number of vertices in S^c that have at least $C_\delta d(|S|/n)$ neighbors in S is at least $|S^c| - \frac{C_\omega n^2}{d|S|}$.

(P3). $d = \Omega(\Delta)$ and if $d = \omega(\log n)$, then also $d = \mathcal{O}(\delta)$.

We now describe the properties in detail and argue why each of them is intrinsic for the analysis. **(P1)** describes a vertex expansion, which means that connected sets have a neighborhood that is roughly in the order of the average degree larger than the set itself. Without this property, the broadcasting process could end up in a set with a tiny neighborhood and thereby slow down too much. Note that in **(P1)**, C_β depends on C_α. Becausse C_α has to be a constant, the upper limit on C_α only applies for constant d.

(P2) is a certain edge expansion property, implying that a large portion of uninformed vertices have a sufficiently large number of informed neighbors. This avoids the situation where the broadcasting process stumbles upon a point when it has informed many vertices but most of the remaining uninformed vertices have very few informed neighbors and therefore only a small chance to get informed. Note that **(P2)** is only useful for $|S| = \omega(n/d)$.

The last property **(P3)** demands a certain regularity of the graph. It is trivially fulfilled for regular graphs, which many definitions of expanders require. The condition $d = \Omega(\Delta)$ for the case $d = \mathcal{O}(\log n)$ does not limit any of our graph classes below. If the average degree is at most logarithmic, **(P3)** implies no further restrictions. Otherwise, we require δ, d, and Δ to be of the same order of magnitude. Without this condition, there could be an uninformed vertex with δ informed neighbors of degree $\omega(\delta)$ that does not get informed in logarithmic time with a good probability. With an additional factor of Δ/δ, this could be resolved, but because we aim at a logarithmic bound, we require $\delta = \Theta(\Delta)$ for $d = \omega(\log n)$. Note that we do not require $d = \omega(1)$, but the proof techniques for constant and nonconstant average degrees will differ in Section 5.

We now describe several important graph classes that are expanding (i.e., satisfy all three properties of Definition 4.1 with high probability).

4.1.1. Complete Graph. It is not difficult to show that complete graphs are expanding.

THEOREM 4.1. *Complete graphs are expanding.*

PROOF. We first prove that **(P1)** holds. Let C_α be an arbitrary constant. Take any subset $S \subseteq V$ with $3 \leqslant |S| \leqslant C_\alpha n/(n-1)$. Then

$$|\Gamma(S) \setminus S| = n - |S| \geqslant |S|\,(n-1)\,\frac{n - |S|}{|S|\,n} = |S|\,(n-1)\left(\frac{1}{|S|} - \frac{1}{n}\right),$$

so **(P1)** holds with $C_\beta = \frac{1}{|S|} - \frac{1}{n} \geqslant \frac{n-1}{C_\alpha n} - \frac{1}{n} > 0$. We now show that **(P2)** holds. Let $C_\delta \in (0, 1)$ be an arbitrary constant. Take any subset $S \subseteq V$. Then every vertex $v \in S^c$ has exactly $|S| \geqslant C_\delta d(|S|/n)$ neighbors in S which implies that **(P2)** is satisfied.

Property **(P3)** is trivially fulfilled because a complete graph is regular. □

4.1.2. Random Graphs $\mathcal{G}(n, p)$, $p \geqslant (\log n + \omega(1))/n$. In this section, we show that a large class of random graphs is expanding with probability $1 - o(1)$. We use the popular random graph model $\mathcal{G}(n, p)$, where between each two vertices out of a set of n vertices an edge is present independently with probability p. This model is usually called the Erdős-Rényi random graph model.

We distinguish two kinds of random graphs with slightly different properties:

Definition 4.2 (*sparse and dense random graph*). We call a random graph $\mathcal{G}(n, p)$ *sparse* if $p = (\log n + f_n)/n$ with $f_n = \omega(1)$ and $f_n = \mathcal{O}(\log n)$ and *dense* if $p = \omega(\log(n)/n)$.

Note that our definition of a sparse random graph coincides with the one of Cooper and Frieze [2007] who set $p = c_n \log(n)/n$ with $(c_n - 1)\log n = \omega(1)$ and $c_n = \mathcal{O}(1)$. In the remainder of this section, we prove the following theorem.

THEOREM 4.2. *Sparse and dense random graphs are expanding with probability* $1 - o(1)$.

The proof can be skipped at a first reading of the article because the following sections do not depend on the proven results of this section.

PROOF. Note that for random graphs, $d = p(n-1)(1 \pm o(1))$ holds with probability $1 - n^{-1}$. To simplify the presentation of the proof, we ignore the factor $(1 \pm o(1))$ because we do not try to optimize the used constants.

The easiest property to check is **(P3)**. That $d = \Omega(\Delta)$ holds with probability $1 - o(1)$ is a well-known property of random graphs and can be shown by union and Chernoff bounds (cf. Lemma A.1) as follows:

$$\mathbf{Pr}\left[\Delta \geqslant 5d\right] = \mathbf{Pr}\left[\exists v \in V : \deg(v) \geqslant 5d\right] \leqslant n \exp(-4d/3) = o(1).$$

Analogously for $d = \omega(\log n)$,

$$\mathbf{Pr}\left[\delta \leqslant d/2\right] = \mathbf{Pr}\left[\exists v \in V : \deg(v) \leqslant d/2\right] \leqslant n \exp(-d/8) = o(1).$$

For the proof of **(P2)** it suffices to bound the number of neighbors of a set by Chernoff bounds. The following lemma does this for sparse and dense random graphs at once.

LEMMA 4.3. *Sparse and dense random graphs satisfy* **(P2)** *with probability* $1 - o(1)$.

PROOF. We choose $C_\delta = 1/2$ and $C_\omega = 32$. Consider a set $S \subseteq V$ of arbitrary size $|S| = s$. We want to show that the number of vertices in S^c that have at least $C_\delta ds/n$ neighbors in S is at least $|S^c| - C_\omega \frac{n^2}{ds}$.

Fix a vertex $v \in S^c$. Linearity of expectations implies $\mathbf{E}[\deg_S(v)] = \sum_{u \in S} p = ps$. Hence, a Chernoff bound (Lemma A.1) gives

$$\mathbf{Pr}\left[\deg_S(v) \leqslant (1/2)\,\mathbf{E}\left[\deg_S(v)\right]\right] \leqslant \exp\left(-\frac{ds}{8n}\right).$$

Hence, the probability for the existence of a subset of vertices in S^c of size $C_\omega n^2/(ds)$ being *bad* (i.e., the set has more than $\frac{C_\omega n^2}{ds}$ vertices with less than $C_\delta ds/n$ neighbors in S) can be bounded by

$$\binom{n-s}{\frac{C_\omega n^2}{ds}} \exp\left(-\frac{ds}{8n}\right)^{C_\omega n^2/(ds)} \leqslant 2^n \exp(-4n).$$

Taking the union bound over all possible sets S, we obtain

$$\mathbf{Pr}\left[\exists \,\text{bad}\; S\right] \leqslant 2^n \cdot 2^n \exp(-4n) \leqslant \left(\frac{4}{e^4}\right)^n. \quad \square$$

We now turn to **(P1)**. We first prove that **(P1)** holds for dense random graphs. After that, we extend it to sparse random graphs, which requires slightly more involved arguments.

LEMMA 4.4. *Dense random graphs satisfy* **(P1)** *with probability* $1 - o(1)$.

PROOF. Let $C_\alpha > 0$ be an arbitrary constant. Fix a set $S \subseteq V$ of size $s = |S|$ with $1 \leqslant s \leqslant C_\alpha(n/d)$. We show that $|\Gamma(S) \setminus S| \geqslant C_\beta ds$ with $C_\beta := 1/(4(C_\alpha + 1))$.

The probability that a vertex $v \in S^c$ is connected to a vertex in S is

$$1 - (1-p)^s \geqslant 1 - \exp(-ps).$$

Linearity of expectation and using the fact that $e^{-x} \leqslant \frac{1}{x+1}$ for any number $x \geqslant 0$ gives

$$\begin{aligned}\mathbf{E}\left[|\Gamma(S) \setminus S|\right] &\geqslant (n-s)\left(1 - \frac{1}{ps+1}\right)\\ &= \left(n - o\left(\frac{n}{\log n}\right)\right)\frac{ps}{ps+1} \geqslant \frac{n}{2}\,\frac{ps}{C_\alpha + 1} = 2\,C_\beta ds.\end{aligned}$$

Applying Chernoff bounds (Lemma A.1), we obtain

$$\mathbf{Pr}\left[|\Gamma(S) \setminus S| \leqslant C_\beta\, ds\right] \leqslant \exp\left(-C_\beta ds/4\right).$$

It remains to show that this holds for all sets S. First, taking a union bound over all sets of size s, we obtain

$$\mathbf{Pr}\left[\exists\, S \subseteq V : |S| = s,\ |\Gamma(S) \setminus S| \leqslant C_\beta\, ds\right] \leqslant n^s \exp\left(-C_\beta ds/4\right) \leqslant n^{-\omega(1)},$$

where the last inequality uses the assumption $d = \omega(\log n)$. Finally, a union bound over all possible values of s yields

$$\mathbf{Pr}\left[\exists\, S \subseteq V : |\Gamma(S) \setminus S| \leqslant C_\beta\, ds\right] \leqslant \sum_{s=1}^{n} n^{-\omega(1)} = n^{-\omega(1)}. \quad \square$$

We now consider sparse random graphs. For this, we need the following three technical lemmas. The first one proves a slightly stronger bound compared to the original lemma in Cooper and Frieze [2007, Property P2].

LEMMA 4.5. *Sparse random graphs satisfy with probability $1 - o(1)$ that for every subset $S \subseteq V$ of size $s = \mathcal{O}(n/d)$ it holds that $|E(S,S)| = o(s \log n)$.*

PROOF. We assume without loss of generality $S \neq \emptyset$. We bound the probability for the existence of a set S of size s with $|E(S,S)| \geqslant s\frac{\log n}{\sqrt{\log\log n}}$ as follows:

$$\begin{aligned}&\mathbf{Pr}\left[\exists S : |E(S,S)| \geqslant s\frac{\log n}{\sqrt{\log\log n}}\right]\\ &\leqslant \binom{n}{s}\binom{\binom{s}{2}}{s\frac{\log n}{\sqrt{\log\log n}}} p^{s\frac{\log n}{\sqrt{\log\log n}}}\\ &\leqslant n^s\left(\frac{s^2\, e}{s\frac{\log n}{\sqrt{\log\log n}}}\right)^{s\frac{\log n}{\sqrt{\log\log n}}} p^{s\frac{\log n}{\sqrt{\log\log n}}} = n^s\left(\frac{s\, e\, p\sqrt{\log\log n}}{\log n}\right)^{s\frac{\log n}{\sqrt{\log\log n}}}\\ &= \exp\left(-s\left(\frac{\log n}{\sqrt{\log\log n}}\log\left(\frac{\log n}{s\, e\, p\sqrt{\log\log n}}\right) - \log n\right)\right)\\ &\leqslant \exp\left(-\Omega(\log n\sqrt{\log\log n}) - \log n\right)\\ &= n^{-\omega(1)},\end{aligned}$$

where, in the third inequality, we used that $s = \mathcal{O}(n/d)$ and $p = \Theta(d/n)$ together imply that $s\, e\, p = \mathcal{O}(1)$. Taking the union bound over all values of s completes the proof. □

It is known that in very sparse random graphs vertices with small degree are rare and far away. To prove **(P1)** we need the following statement.

LEMMA 4.6. *Sparse and dense random graphs satisfy with probability* $1 - o(1)$ *that no two vertices of degree at most* $d/50$ *are within distance at most 3.*

PROOF. We prove a slightly stronger statement: that there are no two vertices of degree at most $d/50$ within distance at most $\log(n)/(\log\log n)^2$ with probability $1 - o(1)$.

For $d \leqslant 2.5 \log n$, we use property P2 of Lemma 1 of Cooper and Frieze [2007], which states that no two vertices of degree at most $\log n/20$ are within distance at most $\log(n)/(\log\log n)^2$ with probability $1 - o(1)$.

For $d \geqslant 2.5 \log n$, we calculate by Chernoff bounds that the probability that an arbitrary vertex has at most $d/50$ neighbors is $\exp\left(-(49^2\, d)/(2 \cdot 50^2)\right) \leqslant n^{-1.2}$. Therefore, the probability that there exists a vertex with at most $d/50$ neighbors is $n \cdot n^{-1.2} = o(1)$, and the claim is satisfied. □

We also need the following simple graph-theoretical lemma. We use it later, with d being the average degree, but it holds for d being an arbitrary number.

LEMMA 4.7. *Let* $d \in \mathbb{N}$ *and* G *be a graph where no two vertices of degree at most* $d/50$ *are within distance at most* 2. *Then, for any connected* $S \subseteq V$ *having at least two vertices,* $\sum_{v \in S} \deg(v) \geqslant (d/100)|S|$.

PROOF. Call a vertex *small* if it has degree less than $d/50$; otherwise, we call it *big*. Let T be a spanning tree of S. Let x be any vertex in S that is not small (i.e., big). For any small vertex $u \in S$, let $\pi(u)$ be the unique neighbor of u that is on the unique path from u to x in T. Because two small vertices have distance at least three, $\pi(u)$ is big, and, for different small vertices u_1, u_2, we have $\pi(u_1) \neq \pi(u_2)$. Hence, π is an injective mapping of small vertices into big vertices. In consequence, S contains at least $|S|/2$ big vertices. Hence, $\sum_{v \in S} \deg(v) \geqslant (|S|/2)(d/50) = (d/100)|S|$. □

Using all three above lemmas, we prove **(P1)** for sparse graphs.

LEMMA 4.8. *Sparse random graphs satisfy* **(P1)** *with probability* $1 - o(1)$.

PROOF. To prove **(P1)**, let $C_\alpha > 0$ be an arbitrary constant and let $S \subseteq V$ with $s = |S|$ be a subset with

—$3 \leqslant s \leqslant C_\alpha \frac{n}{d}$,
—$|E(S, S)| = o(s \log n)$, and
—$\sum_{v \in S} \deg(v) \geqslant s \frac{d}{100}$.

The last two conditions follow from Lemmas 4.5, 4.6, and 4.7. We show that $|\Gamma(S) \setminus S| > C_\beta d s$ with $C_\beta = \min\{1/200, e^{-500}/C_\alpha\}$.

We may assume that all $\sum_{v \in S} \deg(v) - o(s \log n)$ outgoing edges from S hit a uniformly chosen vertex among $V \setminus S$. This is a valid assumption because it may only lead to an underestimation of the number of outgoing edges since a vertex in S may actually only hit the same vertex once. We call a set S of size s *bad* if $|\Gamma(S) \setminus S| \leqslant C_\beta ds$. We compute

$$\begin{aligned}
\mathbf{Pr}\left[\exists \text{ bad set } S \text{ with } |S| = s\right] &\leqslant \binom{n}{s}\binom{n-s}{C_\beta ds}\left(\frac{C_\beta ds}{n}\right)^{\sum_{v\in S}\deg(v) - o(s\log n)} \\
&\leqslant \left(\frac{en}{s}\right)^s \left(\frac{en}{C_\beta ds}\right)^{C_\beta ds} \left(\frac{C_\beta ds}{n}\right)^{ds/110} \\
&= \left(\frac{en}{s}\right)^s e^{C_\beta\, ds} \left(\frac{C_\beta ds}{n}\right)^{(\frac{1}{110} - C_\beta)\, ds} \\
&\leqslant \left(\frac{en}{s}\right)^s e^{C_\beta\, ds} \left(\frac{C_\beta ds}{n}\right)^{ds/11000} \left(\frac{C_\beta ds}{n}\right)^{ds/250}.
\end{aligned}$$

Plugging in the definition of s and C_β, we observe that the two middle terms of the last expression can together be upper-bounded by 1 since

$$e^{11000\, C_\beta}\left(\frac{C_\beta ds}{n}\right) \leqslant e^{11000\, C_\beta}\, C_\beta C_\alpha \leqslant e^{11000/200}\, e^{-500} = e^{-445} < 1.$$

Hence,

$$\begin{aligned}
\mathbf{Pr}\left[\exists \text{ bad set } S \text{ with } |S| = s\right] &\leqslant \left(\frac{en}{s}\right)^s \left(\frac{C_\beta ds}{n}\right)^{ds/250} \\
&= \exp\left(-s\left(\frac{d}{250}\log\left(\frac{n}{C_\beta ds}\right) - \log\left(\frac{en}{s}\right)\right)\right) \\
&\leqslant \exp\left(-3\left(\frac{\log n}{250}\log\left(\frac{1}{C_\alpha\, C_\beta}\right) - \log\left(\frac{en}{3}\right)\right)\right) \\
&\leqslant n^{-3},
\end{aligned}$$

where the second to last inequality holds due to our assumptions on s, $d \geqslant \log n$, and $C_\beta \leqslant e^{-500}/C_\alpha$. A union bound over all values for s proves the claim of Lemma 4.8. □

This proves that sparse and dense random graphs satisfy all three properties of expanding graphs with probability $1 - o(1)$ and therefore also completes the proof of Theorem 4.2.

4.1.3. Strong Expander Graphs. Expander graphs (see Hoory et al. [2006] for a survey) are "perfect" networks in the sense that they unite several desirable properties, such as low diameter, small degree, and high connectivity. They are therefore attractive for routing [Broder et al. 1994], load balancing [Rabani et al. 1998], and communication problems such as the rumor-spreading task considered here.

To define a strong expander graph more formally, we have to introduce a bit of notation. For a d-regular graph G, its adjacency matrix A is symmetric and has n real eigenvalues $d = \lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$. Define $\lambda := \max\{|\lambda_2|, |\lambda_n|\}$. It is well-known that λ captures the expansion of G in the sense that a small λ implies good expansion (cf. Lemmas 4.10 and 4.11) and vice versa Hoory et al. [2006, Theorem 2.4].

Definition 4.3 (*expander*). We call a d-regular graph $G = (V, E)$ a *strong expander* if there is a constant $C > 0$ (independent of d) such that $C < \sqrt{d}$ and $\lambda(G) \leqslant C\sqrt{d}$.

We remark that graphs that satisfy the even stronger condition $\lambda \leqslant 2\sqrt{d-1}$ are called *Ramanujan graphs,* and the construction of such graphs has received a lot of attention (cf. Hoory et al. [2006] for more details). It is known that for any d-regular graph, $\lambda \geqslant 2\sqrt{d-1} - \frac{2\sqrt{d-1}-1}{nd/2}$. Hence, as $n \to \infty$, the smallest possible value for the constant C in Definition 4.3 is $2\sqrt{(d-1)/d}$; in particular, we may assume in the following that $C > 1$.

We prove the following theorem, which has been used in Cooper et al. [2012].

THEOREM 4.9. *Strong expanders are expanding.*

We first state two auxiliary lemmas that relate the second largest eigenvalue in absolute value λ to the expansion of G.

LEMMA 4.10 (FROM [KAHALE 1995; TANNER 1984]). *For any subset $S \subseteq V$ of a d-regular graph G,*

$$|\Gamma(S)| \geqslant \frac{d^2\,|S|}{\lambda^2 + (d^2-\lambda^2)\,|S|/n}.$$

We also need the expander mixing lemma.

LEMMA 4.11 (EXPANDER MIXING LEMMA HOORY ET AL. [2006, LEMMA 2.5]). *For any two subsets $A, B \subseteq V$ of a d-regular graph G, we have*

$$\left| |E(A,B)| - \frac{d|A|\cdot|B|}{n} \right| \leqslant \lambda \cdot \sqrt{|A|\cdot|B|}.$$

We are now ready to prove Theorem 4.9, that strong expanders are expanding.

PROOF OF THEOREM 4.9. **(P3)** is trivially satisfied because the graph is regular. We first prove **(P1)** and afterward **(P2)**.

(P1): Let $S \subseteq V$ be any set of size $s = |S| \leqslant C_\alpha \frac{n}{d}$, where $C_\alpha \leqslant d/2$ is an arbitrary constant. Consider first the case $d = \omega(1)$. Then, using Lemma 4.10 and $\lambda \leqslant C\sqrt{d}$ gives

$$|\Gamma(S)| \geqslant \frac{d^2\,s}{\lambda^2 + (d^2-\lambda^2)\,s/n} \geqslant \frac{d^2\,s}{C^2 d + d^2\frac{C_\alpha}{d}} = \frac{ds}{C^2 + C_\alpha}$$

and therefore

$$|\Gamma(S) \setminus S| \geqslant \left(\frac{1}{C^2 + C_\alpha} - \frac{1}{d}\right) ds.$$

This proves **(P1)** because the factor in front of ds is at least a constant (since $d = \omega(1)$).

For $d = \mathcal{O}(1)$, we use Lemma 4.10 slightly differently to get

$$\begin{aligned} |\Gamma(S)| &\geqslant \frac{d^2\,s}{\lambda^2 + (d^2-\lambda^2)\,s/n} \\ &= \frac{d^2 s}{\lambda^2\,(1-(s/n)) + d^2\,(s/n)} \\ &\geqslant \frac{d^2 s}{C^2 d\,(1-(s/n)) + d^2\,(s/n)}. \end{aligned}$$

Hence,

$$\begin{aligned} |\Gamma(S) \setminus S| &\geqslant \frac{d^2 s}{C^2 d\,(1-(s/n)) + d^2\,(s/n)} - s \\ &= \frac{d - C^2\,(1-(s/n)) - d\,(s/n)}{C^2 d\,(1-(s/n)) + d^2\,(s/n)} \cdot ds. \end{aligned}$$

The denominator is bounded here by a constant, since $d = \mathcal{O}(1)$ and $s \leqslant n/2$. The numerator is at least a constant, since, by assumption, C is a constant that is strictly smaller than $\sqrt{d}$. This proves **(P1)**.

(P2): We may assume that $|S^c| \geqslant \lceil \frac{4n^2C^2}{ds} \rceil$, because otherwise $|S^c| = \mathcal{O}(\frac{n^2}{ds})$, and **(P2)** holds trivially by choosing the constant C_ω sufficiently large; for instance, $C_\omega := 10 \cdot \max\{C^2, 1\}$. Let us now order the vertices in S^c according to the number of neighbors in S in decreasing order. Let N^- be the last $\lceil \frac{4n^2C^2}{ds} \rceil$ vertices in that list (i.e., the $\lceil \frac{4n^2C^2}{ds} \rceil$ vertices with the least number of neighbors in S) and let $N^+ := S^c \setminus N^-$ be the remaining set of vertices in S^c. Observing that $\lceil \frac{4n^2C^2}{ds} \rceil \leqslant \frac{3}{2} \cdot \frac{4n^2C^2}{ds}$ (since $ds \leqslant n^2$ and $C \geqslant 1$) and applying Lemma 4.11, we obtain

$$
\begin{aligned}
|E(S, N^-)| &\geqslant d\,\frac{|S|\,|N^-|}{n} - \lambda\sqrt{|S|\,|N^-|} \\
&\geqslant d\,\frac{s\,\frac{4n^2C^2}{ds}}{n} - C\sqrt{d}\sqrt{s\,\frac{3}{2}\cdot\frac{4n^2C^2}{ds}} \\
&= 4C^2n - \sqrt{6}\cdot C^2 n \geqslant C^2 n.
\end{aligned}
$$

This implies that the average number of neighbors in S of vertices in N^- is at least

$$
\frac{C^2 n}{\gamma\,\frac{\frac{3}{2}\cdot 4n^2C^2}{ds}} \geqslant \frac{ds}{6n},
$$

and all vertices N^+ must have at least this degree. Hence, we have shown that for every subset S, at least $|S^c| - |N^-| \geqslant n - s - \frac{3}{2}\cdot\gamma\,\frac{4n^2C^2}{ds} \geqslant n - s - 6\frac{n^2C^2}{ds}$ vertices in S^c have at least $ds/(6n)$ neighbors in S, and property **(P2)** follows with $C_\delta = 1/6$ and $C_\omega = 6C^2 > 0$. □

4.1.4. Random Graphs with Fixed Degree Sequence.

Definition 4.4 (*random graph with fixed degree sequence*). Let $d_1, d_2, \ldots, d_n$ be a degree sequence with maximum degree $\Delta = o(\sqrt{n})$ and $\Delta/\delta = \mathcal{O}(1)$. Then, a random graph with this degree sequence is chosen uniformly at random from the set of all simple graphs with this degree sequence.

Note that a random d'-regular graph is a random graph with a fixed-degree sequence $d_1 = d_2 = \cdots = d_n = d'$. Random regular graphs have gained increasing interest in the context of peer-to-peer networks (e.g., they appear quite naturally as a limiting distribution of certain graph transformations [Mahlmann and Schindelhauer 2006; Cooper et al. 2007a]).

For a random graph with a fixed-degree sequence as defined earlier, Broder et al. [1998, Lemma 18] showed that $\lambda = \mathcal{O}(\sqrt{d})$ with probability $1 - \mathcal{O}(n^{-\operatorname{poly}(n)})$ and hence gave the following theorem:

THEOREM 4.12. *A random graph with fixed-degree sequence is expanding with probability* $1 - o(1)$.

4.2. Hypercubes

We now recall the definition of hypercubes.

Definition 4.5 (*Hypercube*). For any d, a d-dimensional hypercube $H = (V, E)$ has $n = 2^d$ vertices $V = \{0, 1\}^d$ and edges $E = \{\{u, v\} : \|u - v\|_1 = 1\}$.

The i-th bit of a bitstring $x \in \{0,1\}^d$ will be denoted as $x[i]$. We observe that the hypercube is not expanding.

THEOREM 4.13. *Hypercubes are not expanding.*

PROOF. Define $S := \bigcup_{i=1}^{\log d} L_i$, where L_i is the set of vertices x with $\|x\|_1 = i$. Then, $3 \leqslant |S| = o(n/\log n)$ and

$$|\Gamma(S) \setminus S| = |L_{\log(d)+1}| = \binom{d}{\log(d)+1} = \frac{d-\log d}{\log(d)+1}\binom{d}{\log d} \leqslant \frac{d}{\log(d)+1}\,|S| = o(d\,|S|),$$

which violates **(P1)**. □

Hence, a separate analysis is needed, and this is given in Section 7.

4.3. *k*-ary Trees

For complete k-ary trees ($k \geqslant 2$), it is easy to verify that they are not expanding.

LEMMA 4.14. *k-ary trees are not expanding.*

PROOF. Consider a k-ary tree and let $C_\alpha = 1/2$ and S be the set of vertices which are in the subtree of fixed children of the root. Then, $|S| \leqslant (n-1)/k \leqslant n/2 \leqslant C_\alpha(n/d)$, but $|\Gamma(S) \setminus S| = 1$ violating **(P1)**.

However, it is also not difficult to show the following theorem:

THEOREM 4.15. *For complete k-ary trees, the broadcast time of the quasirandom model is $\mathcal{O}(k\log(n)/\log k)$ with probability 1, whereas the expected broadcast time of the fully random model is $\Omega(k\log n)$.*

PROOF. Because a k-ary tree has a diameter of $\Theta(\log(n)/\log k)$ and maximum degree of $k+1$, plugging these values into the bound of Theorem 3.1, we obtain the first claim.
To see the lower bound for the fully random model, define a path P of length $\mathrm{diam}(G)/2$ inductively as follows: Assume that the root u_0 is initially informed. Then let $P = (u_0, u_1, \ldots, u_i)$ for $1 \leqslant i \leqslant \mathrm{diam}(G)/2$, where u_i is the last vertex informed by u_{i-1}. By the coupon collector's problem, the expected time it takes for u_{i-1} to inform u_i is at least $k\log k$ and, therefore, the expected time to inform $v_{\mathrm{diam}(G)/2-1}$ is at least $\Omega(\mathrm{diam}(G)\,k\log k) = \Omega(k\log n)$. □

5. QUASIRANDOM RUMOR SPREADING ON EXPANDING GRAPHS

In this section, we prove our main result, that quasirandom rumor spreading informs all vertices in an expanding graph in a logarithmic number of rounds.

THEOREM 5.1. *Let $\gamma \geqslant 1$ be a constant. The broadcast time of the quasirandom model on expanding graphs is $\mathcal{O}(\log n)$ with probability $1 - \mathcal{O}(n^{-\gamma})$.*

To analyze the propagation process, we decompose it into a forward part (Sections 5.1 and 5.2) and a backward part (Sections 5.3 and 5.4). In the analysis of the forward part, we show that if a vertex is informed at some time, then $\mathcal{O}(\log n)$ steps later, only $\mathcal{O}(n/d)$ vertices remain uninformed (cf. Theorem 5.2). In the analysis of the backward part, we show that if a vertex is uninformed at some time, then $\mathcal{O}(\log n)$ steps earlier, at least $\omega(n/d)$ vertices must be uninformed as well (cf. Theorem 5.7). Combining both yields Theorem 5.1.
We show that all this holds with probability $1 - n^{-\gamma}$ for an arbitrary $\gamma \geqslant 1$. Because Theorem 5.1 is considerably easier to show for $d = \mathcal{O}(1)$, we handle this case separately in Section 5.5 and now concentrate on the case $d = \omega(1)$. This makes the proofs of the

lemmas of this section slightly shorter. Therefore, in this section, apart from the last subsection, we use the following adjusted property:

(P3'). $d = \omega(1)$ and $d = \Omega(\Delta)$. If $d = \omega(\log n)$ then $d = \mathcal{O}(\delta)$.

Because the precise constants will be crucial in parts of the following proofs, we use the following notation. Constants with a lowercase Greek letter index (e.g., C_α and C_β) stem from Definition 4.1. Constants without an index or with a numbered index (e.g., C and C_1) are local constants in lemmas. K is used to denote a number of time steps.

5.1. Forward Analysis

In this section, we prove the following theorem:

THEOREM 5.2. *Let $\gamma \geqslant 1$ be a constant. The probability that the quasirandom model started in a fixed vertex u informs $n - \mathcal{O}(n/d)$ vertices within $\mathcal{O}(\log n)$ rounds is at least $1 - n^{-\gamma}$.*

In our analysis, we use the following two notations for sets of informed vertices: Let I_t be the set of vertices that know the rumor after the t-th step. Let $N_t \subseteq I_t$ be the set of "newly informed" vertices—that is, those that know the rumor after the t-th step, but have not spread this information yet. The latter set will be especially important because these are the vertices that have preserved their independent random choice.

Each of the following Lemmas 5.3–5.6 examines one phase consisting of several steps. Within each phase, we only consider information spread from vertices that became informed in the previous phase. This is justified by Lemma 2.1.

Let u be (newly) informed at time step 0. To get a sufficiently large set of newly informed vertices to start with, we first show how to obtain a set N_t of size $\Theta(\log n)$ within $t = \mathcal{O}(\log n)$ steps. This is simple if $d = \omega(\log n)$—after $c \log n$ rounds, the first vertex has informed exactly $c \log n$ new vertices. Otherwise, we use the fact that **(P1)** implies that the neighborhoods $\Gamma^k(u)$ grow exponentially with k. Since within Δ steps $\Gamma^k(u)$ becomes informed if $\Gamma^{k-1}(u)$ was informed beforehand, this yields the claim in this case. The precise statement is as follows:

LEMMA 5.3. *Let $C > 0$ be an arbitrary constant. Then, with probability 1, there is a time step $t = \mathcal{O}(\log n)$ such that*

—$|N_t| \geqslant C \log n$ *and*
—$|I_t \setminus N_t| = o(|N_t|)$.

The proof of Lemma 5.3 and all following lemmas can be found in Section 5.2. We now assume that we have a set N_t of size $\Omega(\log n)$. We aim at informing $\Omega(n/d)$ vertices. For the very dense case of $d = \Omega(n/\log n)$, this is a trivial statement. Note that in the following argument we can always assume that we have not informed *too many* vertices as the number of informed vertices can at most double in each time step. The following lemma shows that, given a set of informed vertices matching the conditions of **(P1)**, within a constant number of steps the set of informed vertices increases by a factor strictly larger than 1.

LEMMA 5.4. *For any constants $\gamma \geqslant 1$ and $C_\alpha > 0$ there are constants $K \geqslant 1$, $C_1 > 1$, $C_2 > 1$, and $C_3 \in (3/4, 1)$ such that for all time steps t, if*

—$C_1 \log n \leqslant |I_t| \leqslant C_\alpha\, (n/d)$ *and*
—$|N_t| \geqslant C_3\, |I_t|$,

then with probability $1 - n^{-\gamma}$,

—$|I_{t+K}| \geqslant C_2\, |I_t|$ *and*
—$|N_{t+K}| \geqslant C_3\, |I_{t+K}|$.

Because the precondition of the next Lemma 5.5 is $|I_t| \geqslant 16\,C_\omega(n/d)$, let $C_\alpha = 16\,C_\omega$. Then, Lemma 5.4 yields a constant $C_2 > 1$ such that applying this lemma at most $\log_{C_2}\big(16\,C_\omega\,(n/d)\big) = \mathcal{O}(\log n)$ times leads to at least $16\,C_\omega(n/d)$ informed vertices, a constant fraction of which is newly informed.

The next aim is informing a linear number of vertices. Note that as long as that is not achieved, **(P2)** implies that there is a large set of uninformed vertices that have many neighbors in N_t. This is the main ingredient of the following Lemma 5.5. It shows that, under these conditions, a phase of a constant number of steps suffices to triple the number of informed vertices.

LEMMA 5.5. *For any constant $\gamma \geqslant 1$ there are constants $K \geqslant 1$, $C > 1$, and $C_\omega > 0$ such that for all time steps t, if*

—$\max\{C \log n, 16\,C_\omega(n/d)\} \leqslant |I_t| \leqslant n/16$ *and*
—$|N_t| \geqslant (3/4)\,|I_t|$,

then with probability $1 - n^{-\gamma}$,

—$|I_{t+K}| \geqslant 3\,|I_t|$ *and*
—$|N_{t+K}| \geqslant (3/4)\,|I_{t+K}|$.

Applying Lemma 5.5 at most $\mathcal{O}(\log n)$ times, a linear fraction of the vertices gets informed. In a final phase of $\mathcal{O}(\log n)$ steps, one can then inform all but $\mathcal{O}(n/d)$ vertices, as shown in Lemma 5.6.

LEMMA 5.6. *For any constants $\gamma \geqslant 1$ and $C > 0$, there is a $K = \mathcal{O}(\log n)$ such that for all time steps t, if*

—$|N_t| \geqslant C\,n$,

then with probability $1 - n^{-\gamma}$,

—$|I_{t+K}| = n - \mathcal{O}(n/d)$.

Combining all these phases, a union bound gives $|I_{\mathcal{O}(\log n)}| = n - \mathcal{O}(n/d)$ with probability $1 - \mathcal{O}(\log(n)\,n^{-\gamma})$. Because γ was arbitrary in all lemmas, Theorem 5.2 follows.

5.2. Proofs of the Lemmas Used in the Forward Analysis

PROOF OF LEMMA 5.3. Let u be informed at time step 0. If $d = \omega(\log n)$, then by **(P3)** $\delta = \Theta(d)$, and a single phase of $C \log n$ rounds suffices; that is, we have $N_{C \log n} = C \log n$, and the lemma follows.

We now describe how to obtain $C \log n$ newly informed vertices for $d = \mathcal{O}(\log n)$. For this, we choose a C_α such that $C_\alpha n/d \geqslant C \log n$ and get, by **(P1)** for $k \geqslant 3$, as long as $|\Gamma^{\leqslant k}(v)| = \mathcal{O}(n/d)$,

$$\begin{aligned} |\Gamma^{\leqslant k+1}(v)| &= |\Gamma^{\leqslant k}(v)| + |\Gamma^{k+1}(v)| = |\Gamma^{\leqslant k}(v)| + |\Gamma(\Gamma^{\leqslant k}(v)) \setminus \Gamma^{\leqslant k}(v)| \\ &\geqslant (1 + C_\beta d)\,|\Gamma^{\leqslant k}(v)|. \end{aligned} \tag{1}$$

Subtracting $|\Gamma^{\leqslant k}(v)|$ on both sides yields

$$|\Gamma^{k+1}(v)| \geqslant C_\beta\, d\,|\Gamma^{\leqslant k}(v)|.$$

Because $|\Gamma^{\leqslant 3}(v)| \geqslant 3$, by induction,

$$|\Gamma^{k}(v)| \geqslant 3\,(C_\beta\, d)^{k-3}$$

for all k with $k \geqslant 3$ and $|\Gamma^{\leqslant k-1}(v)| \leqslant C \log n$. Therefore, we can choose a $k = \mathcal{O}(\log\log(n)/\log d)$ such that $|\Gamma^{k}(v)| \geqslant C \log n$.

We use the delaying and ignoring assumption (cf. Lemma 2.1) to perform k phases of Δ rounds each. Then, after these $t = \Delta k = \mathcal{O}(\Delta\,(\log\log n)/\log d) = \mathcal{O}(\log n)$ steps (as $\Delta = \mathcal{O}(d)$ by **(P3)** and $d/\log d = \mathcal{O}(\log(n)/\log\log n)$ by $d = \mathcal{O}(\log n)$), all vertices in $\Gamma^{\leqslant k}(v)$ get informed, but no vertex of $\Gamma^k(v)$ has been active. In consequence, we have

$$\begin{aligned} |N_t| &= |\Gamma^k(v)| \geqslant C\log n, \\ |I_t \setminus N_t| &= |\Gamma^{\leqslant k-1}(v)| \leqslant |\Gamma^k(v)|/(C_\beta\, d) = o(|N_t|), \end{aligned} \tag{2}$$

where the last equation stems from **(P3')**. □

PROOF OF LEMMA 5.4. We choose the following constants:

$$C_1 := \frac{8\,\gamma\,\Delta^2}{C_\beta^2\, d^2} > 1, \qquad C_2 := \frac{4\,\Delta}{C_\beta\, d} > 1,$$

$$C_3 := \left(1 - \frac{C_\beta\, d}{4\,\Delta}\right) \in (3/4, 1), \qquad K := \left\lceil \left(\frac{3\,\Delta}{C_\beta\, d}\right)^2 \right\rceil \geqslant 1,$$

where the C_β is from **(P1)** and depends on the given C_α. K and C_1 to C_3 are all $\Theta(1)$ by **(P3)**. Because I_t is a connected set of appropriate size, **(P1)** gives

$$|\Gamma(I_t) \setminus I_t| \geqslant C_\beta\, d\, |I_t|. \tag{3}$$

Because we are interested in the expansion of N_t and not of I_t, we calculate

$$\begin{aligned} |\Gamma(I_t) \setminus I_t| &= \big|\big(\Gamma(I_t \setminus N_t) \setminus I_t\big) \cup \big(\Gamma(N_t) \setminus I_t\big)\big| \\ &\leqslant \big|\Gamma(I_t \setminus N_t) \setminus I_t\big| + \big|\Gamma(N_t) \setminus I_t\big| \\ &\leqslant \Delta |I_t \setminus N_t| + \big|\Gamma(N_t) \setminus I_t\big|. \end{aligned} \tag{4}$$

Combining Equations (3) and (4) with the assumption $|I_t \setminus N_t| \leqslant \frac{C_\beta\, d}{4\,\Delta} |I_t|$,

$$\big|\Gamma(N_t) \setminus I_t\big| \geqslant C_\beta\, d\, |I_t| - \Delta |I_t \setminus N_t| \geqslant 3\, C_\beta\, d\, |I_t|/4.$$

We now perform one phase consisting of K rounds. We compute the size of the resulting sets I_{t+K} and N_{t+K} as follows:

Let $v \in \Gamma(N_t) \setminus I_t$. Then there is a $u \in N_t$ such that $(u, v) \in E$. The probability that u contacts v within this time interval is $\min\{K/\deg(u), 1\} \geqslant K/\Delta$ (as $\Delta = \omega(1)$ by **(P3')**), which naturally is a lower bound for v becoming contacted by an arbitrary vertex of N_t. By linearity of expectation, the expected number of vertices becoming contacted is at least

$$\mathbf{E}\left[|N_{t+K}|\right] \geqslant K\, |\Gamma(N_t) \setminus I_t|/\Delta \geqslant 3\, C_\beta K\, d\, |I_t|/(4\Delta).$$

Because every vertex can only contact at most K vertices in this time interval, Azuma's inequality (cf. Lemma A.2) gives a probabilistic lower bound on the number of newly informed vertices. More precisely,

$$\mathbf{Pr}\left[|N_{t+K}| \leqslant \frac{C_\beta\, K\, d\, |I_t|}{2\Delta}\right] \leqslant \exp\left(-\frac{C_\beta^2\, d^2\, |I_t|^2}{8\,\Delta^2\, |N_t|}\right) \leqslant n^{-C_1\, C_\beta^2\, d^2/(8\,\Delta^2)} = n^{-\gamma}.$$

It remains to check that $|N_{t+K}| \geqslant \frac{C_\beta\, K\, d\, |I_t|}{2\,\Delta}$ implies the two parts of the claim. First,

$$|I_{t+K}| \geqslant |N_{t+K}| \geqslant \frac{C_\beta\, K\, d}{2\,\Delta}\, |I_t| \geqslant \frac{4\,\Delta}{C_\beta\, d}\, |I_t| = C_2\, |I_t|.$$

For the second part, observe that

$$|N_{t+K}| \geqslant \frac{C_\beta\, K d\, |I_t|}{2\,\Delta} \geqslant \frac{C_\beta\, K d\,(|I_{t+K}| - |N_{t+K}|)}{2\,\Delta} = \frac{C_\beta\, K d}{2\,\Delta}\,|I_{t+K}| - \frac{C_\beta\, K d}{2\,\Delta}\,|N_{t+K}|.$$

Rearranging yields

$$|N_{t+K}| \geqslant \frac{C_\beta\, K d}{2\,\Delta + C_\beta\, K d}\,|I_{t+K}| \geqslant \frac{C_\beta\big(\frac{3\,\Delta}{C_\beta\, d}\big)^2 d}{2\Delta + C_\beta\big(\frac{3\,\Delta}{C_\beta\, d}\big)^2 d}\,|I_{t+K}|$$
$$= \frac{9\Delta}{2\,C_\beta\, d + 9\,\Delta}\,|I_{t+K}| \geqslant \left(1 - \frac{C_\beta\, d}{4\,\Delta}\right)|I_{t+K}|. \quad \square$$

Proof of Lemma 5.5. We choose $C := \frac{512\,\gamma^3\,\Delta^2}{3\,C_\delta^2\, d^2} > 1$, $K := \lceil \frac{16\,\gamma\,\Delta}{C_\delta\, d} \rceil \geqslant 1$, and $C_\omega > 0$ according to **(P2)**.

By property **(P2)**, the number of vertices in N_t^c that have at least $C_\delta d\,|N_t|/n$ neighbors in N_t is at least $|N_t^c| - \frac{C_\omega n^2}{d\,|N_t|}$. Therefore, the number of vertices in I_t^c that have at least $C_\delta d(|N_t|/n)$ neighbors in N_t is at least

$$|N_t^c| - |I_t| - \tfrac{C_\omega n^2}{d\,|N_t|} \geqslant n - 2\,|I_t| - n/12 \geqslant 19n/24 \geqslant 3n/4,$$

where the first inequality is due to $16\,C_\omega(n/d) \leqslant |I_t| \leqslant 4/3\,|N_t|$.

We call a vertex $v \in I_t^c$ *good* if it has at least $C_\delta d\,|N_t|/n$ neighbors in N_t. The probability that a good vertex gets informed in a phase of K rounds (again using $K \leqslant \Delta = \omega(1)$ by **(P3')**) is at least

$$1 - \left(1 - \frac{K}{\Delta}\right)^{C_\delta d\,|N_t|/n} \geqslant 1 - \exp\left(-\frac{K C_\delta d\,|N_t|}{\Delta n}\right) \geqslant 1 - \exp(-16\,\gamma\,|N_t|/n)$$
$$\geqslant 1 - \frac{1}{(16\,\gamma\,|N_t|/n) + 1} = \frac{16\,\gamma\,|N_t|}{16\,\gamma\,|N_t| + n}.$$

By linearity of expectation,

$$\mathbf{E}\left[|N_{t+K}|\right] \geqslant \frac{16\,\gamma\,|N_t|}{16\,\gamma\,|N_t| + n}\frac{3n}{4} \geqslant \frac{16\,\gamma\,|N_t|}{16\,\gamma\, n/16 + n}\frac{3n}{4} = \frac{\gamma\,|N_t|}{(\gamma/16) + 1/16}\frac{3}{4} \geqslant 6\,|N_t|.$$

Azuma's inequality (cf. Lemma A.2) gives

$$\mathbf{Pr}\left[|N_{t+K}| \leqslant 4\,|N_t|\right] \leqslant \exp\left(-\frac{2\,(2|N_t|)^2}{|N_t|\,K^2}\right) = \exp\left(-\frac{8|N_t|}{K^2}\right)$$
$$\leqslant \exp\left(-\frac{|N_t| C_\delta^2 d^2}{128\,\gamma^2\Delta^2}\right) \leqslant \exp\left(-\frac{3C\log(n) C_\delta^2 d^2}{512\,\gamma^2\Delta^2}\right) = n^{-\gamma}.$$

Therefore, with probability $1 - n^{-\gamma}$,

$$|N_{t+K}| \geqslant 4\,|N_t| \geqslant 3\,|I_t| = 3\,|I_{t+K}| - 3\,|N_{t+K}|$$

and after rearranging,

$$|N_{t+K}| \geqslant \frac{3}{4}|I_{t+K}|.$$

This proves the first claim. The second claim follows from

$$|I_{t+K}| \geqslant |N_{t+K}| \geqslant 4|N_t| \geqslant 3\,|I_t|. \quad \square$$

PROOF OF LEMMA 5.6. Let $X \subseteq N_t^c$ be the set of vertices in N_t^c that have at least $C_\delta\, d\,|N_t|/n$ neighbors in N_t. By **(P2)**,

$$|X| \geqslant (n - |N_t|) - \frac{C_\omega\, n^2}{d|N_t|} \geqslant n - |N_t| - \Theta\big(\tfrac{n}{d}\big).$$

Let $v \in X$ and consider a phase of $K := \lceil \frac{2\gamma\,\Delta\, n}{C_\delta\,|N_t|\, d} \log n \rceil$ rounds. Note that $K = \mathcal{O}(\log n)$ by **(P3)**.

If $K \geqslant \Delta$, v becomes informed in this phase with probability 1. Otherwise, the probability that v will not be informed in this phase is at most

$$\mathbf{Pr}\left[v \notin N_{t+K}\right] \leqslant \left(1 - \frac{K}{\Delta}\right)^{C_\delta |N_t| d/n} \leqslant \exp(-2\,\gamma\;\log n) = n^{-2\,\gamma}.$$

Taking the union bound over all vertices in X, we obtain that all vertices in X get informed with probability $1 - n^{-\gamma}$. The claim follows. □

5.3. Backward Analysis

The forward analysis has shown that within $\mathcal{O}(\log n)$ steps, at most $\mathcal{O}(n/d)$ vertices stay uninformed. We now analyze the reverse. The question here is how many vertices have to be uninformed at time $t - \mathcal{O}(\log n)$ if there is an uninformed vertex at time t. We show that this is at least $\omega(n/d)$. To formalize this, recall that $U_{[t_1,t_2]}(w)$ is the set of vertices that reach the vertex w within the time interval $[t_1, t_2]$ (using the usual meaning of "reach" as defined on page 7). We prove the following theorem.

THEOREM 5.7. *Let $\gamma \geqslant 1$ be a constant. If the quasirandom rumor-spreading process does not inform a fixed vertex w until some time t, then there are $\omega(n/d)$ uninformed vertices at time $t - \mathcal{O}(\log n)$ with probability at least $1 - n^{-\gamma}$.*

To prove Theorem 5.7, we fix an arbitrary vertex w and a time t. Ignoring some technicalities, our aim is to prove a lower bound on the number of vertices that have to be uninformed at times before t to keep w uninformed at time t. We first show that the set of uninformed vertices at time $t - \mathcal{O}(\log n)$ is at least of logarithmic size.

For $d = \mathcal{O}(\log n)$, this follows from **(P1)** because all vertices of $\Gamma^{\mathcal{O}(\log\log n/\log d)}(w)$ (and there are at least $\Omega(\log n)$ of these) reach w within $\mathcal{O}(\log n)$ steps. For $d = \omega(\log n)$, a simple Chernoff bound shows that enough vertices of $\Gamma(w)$ contact w within $\mathcal{O}(\log n)$ steps. This is summarized in the following lemma. The proofs of all three lemmas of this section can be found in Section 5.4.

LEMMA 5.8. *Let $\gamma \geqslant 1$ and $C \geqslant 1$ be constants, w a vertex, and $t_2 = \Omega(\log n)$ a time step. Then, with probability $1 - 2\,n^{-\gamma}$ there is a time step $t_1 = t_2 - \mathcal{O}(\log n)$ such that*

$$|U_{[t_1,t_2]}(w)| \geqslant C\;\log n.$$

We now know that within a logarithmic number of time steps there are at least $c \log n$ vertices that have reached w. Very similarly to Lemmas 5.4 and 5.5 in the forward analysis, we can increase the set of vertices that reach w by a multiplicative factor by going back a constant number of time steps. The following lemma again mainly uses **(P1)**. For the very dense case of $d = \Omega(n/\log n)$, there is nothing to show.

LEMMA 5.9. *For any constant $\gamma \geqslant 1$ there is a constant K such that for all vertices w and time steps t_1, t_2, if*

$$\log n \leqslant |U_{[t_1,t_2]}(w)| = \mathcal{O}(n/d),$$

then with probability $1 - n^{-\gamma}$,

$$|U_{[t_1-K,t_2]}(w)| \geqslant 4\,|U_{[t_1,t_2]}(w)|.$$

Using Lemma 5.9 at most $\mathcal{O}(\log n)$ times, we obtain a set of vertices that reach w of size $\Omega(n/d)$. If these are $\omega(n/d)$ vertices, we are done. Otherwise, Lemma 5.10 shows that a phase consisting of $\mathcal{O}(\log n)$ steps suffices to get to this point. This is the only lemma that substantially uses **(P3')**.

LEMMA 5.10. *Let $\gamma \geqslant 1$ be a constant, w a vertex, and t_1, t_2 time steps such that*

$$|U_{[t_1,t_2]}(w)| = \Theta(n/d).$$

Then, with probability $1 - n^{-\gamma}$,

$$|U_{[t_1-\mathcal{O}(\log n),t_2]}(w)| = \omega(n/d).$$

This finishes the backward analysis and shows that $\omega(n/d)$ vertices have to be uninformed to keep a single vertex uninformed for $\mathcal{O}(\log n)$ steps. Together with the forward analysis, which proved that only $\mathcal{O}(n/d)$ vertices remain uninformed after $\mathcal{O}(\log n)$ steps, this finishes the proof of Theorem 5.1 for $d = \omega(1)$.

5.4. Proofs of the Lemmas Used in the Backward Analysis

PROOF OF LEMMA 5.8. Consider first the case that $d = \mathcal{O}(\log n)$. In this case, we choose, as in the proof of Lemma 5.3, a constant C_α such that $C_\alpha n/d \geqslant C \log n$ and apply **(P1)**. By Equation (2) from page 19, there exists a $k = \mathcal{O}(\log\log(n)/\log d)$ such that

$$|\Gamma^{\leqslant k}(w)| \geqslant |\Gamma^k(w)| \geqslant C \log n.$$

Since within Δ rounds each vertex has contacted all neighbors, we have $\Gamma^{\leqslant i}(w) \subseteq U_{[t_2-i\Delta,t_2]}(w)$ for $i \geqslant 1$ and therefore $\Gamma^{\leqslant k}(w) \subseteq U_{[t_2-k\Delta,t_2]}(w)$. As $k\Delta = \mathcal{O}(\log n)$, we see that $|U_{[t_2-\mathcal{O}(\log n),t_2]}| \geqslant C \log n$ with probability 1.

In the remaining case $d = \omega(\log n)$, we estimate the number of neighbors of w that reach w in the previous $K := \lceil 4C^2\gamma\Delta\log(n)/\delta\rceil$ steps. Note that $K = \mathcal{O}(\log n)$ by **(P3)**. For each neighbor $u \in \Gamma(w)$, define a random variable $X(u)$, which is 1 if u contacts v within the time interval $[t_2 - K, t_2]$ and zero otherwise. Then, for each $u \in \Gamma(w)$, $\mathbf{Pr}\left[X(u) = 1\right] \geqslant K/\Delta$. We define $X := \sum_{u\in\Gamma(w)} X_u$. Linearity of expectation gives $\mathbf{E}\left[X\right] \geqslant K\delta/\Delta \geqslant 4C^2\gamma\log n$. Since $\{X(u) : u \in \Gamma(w)\}$ is a set of independent random variables, we obtain by a Chernoff bound that

$$\begin{aligned}\mathbf{Pr}\left[X \leqslant C\ \log n\right] &\leqslant \mathbf{Pr}\left[X \leqslant \frac{1}{4}\mathbf{E}\left[X\right]\right]\\ &\leqslant \exp(-(3/4)^2\,\mathbf{E}\left[X\right]/2)\\ &= \exp(-(9/32)\,4C^2\gamma\log n) \leqslant n^{-\gamma},\end{aligned}$$

where we used the assumption $C \geqslant 1$. This implies that with probability $1 - n^{-\gamma}$, we have

$$\left|U_{[t_2-\mathcal{O}(\log n),t_2]}(w)\right| \geqslant C \log n. \quad \square$$

PROOF OF LEMMA 5.9. Let $S := U_{[t_1,t_2]}(w)$ and let $|S| \leqslant C_\alpha\,(n/d)$ for a constant C_α. Because S is a connected set, **(P1)** gives

$$|\Gamma(S) \setminus S| \geqslant C_\beta\, d\, |S|$$

for a suitable constant C_β. Let $K = \lceil \frac{8\gamma}{C_\beta}\frac{\Delta}{d}\rceil = \mathcal{O}(1)$ (by **(P3)**). Because every vertex $u \in \Gamma(S)\setminus S$ has at least one edge to a vertex $v \in S$, the probability that a vertex $u \in \Gamma(S)\setminus S$ contacts a $v \in S$ in the interval $[t_1 - K, t_1 - 1]$ is at least K/Δ and

$S' := U_{[t_1-K,t_2]}(w)$. By linearity of expectation, the expected number of vertices in $S' \setminus S$ is at least

$$\mathbf{E}\left[|S' \setminus S|\right] \geqslant K|\Gamma(S) \setminus S|/\Delta \geqslant C_\beta Kd\,|S|/\Delta.$$

A simple application of the Chernoff bound gives

$$\mathbf{Pr}\left[|S' \setminus S| \leqslant \frac{C_\beta Kd\,|S|}{2\Delta}\right] \leqslant \exp\left(-\frac{C_\beta Kd\,|S|}{8\Delta}\right) \leqslant n^{-\frac{C_\beta Kd}{8\Delta}}.$$

Hence, with probability $1 - n^{-\gamma}$,

$$|S'| \geqslant \frac{C_\beta Kd\,|S|}{2\Delta} \geqslant 4\gamma|S| \geqslant 4\,|S|. \quad \square$$

Proof of Lemma 5.10. Let $S := U_{[t_1,t_2]}(w)$ with $|S| \leqslant C_\alpha\,(n/d)$ for a constant C_α. Also let $K := \lceil \frac{8\gamma}{C_\beta}\frac{\Delta}{d}\frac{n}{|S|d}\log n \rceil$ and $S' := U_{[t_1-K,t_2]}(w)$. Note that $K = \mathcal{O}(\log n)$ by **(P3)**. We examine a phase of K steps.

Because S is a connected set, **(P1)** gives, as in the proof of Lemma 5.9, $|\Gamma(S) \setminus S| \geqslant C_\beta\,d\,|S|$. If $K \geqslant \Delta$, the lemma immediately follows from the observation

$$|S'| = |\Gamma^{\leqslant 1}(S)| = \Theta(d\,|S|) = \Theta(n) = \omega(n/d).$$

The last equality is based on $d = \omega(1)$, as given by **(P3')**.

We now assume $K \leqslant \Delta$. Because every vertex $u \in \Gamma(S) \setminus S$ has at least one edge to a vertex $v \in S$, the probability that a vertex $u \in \Gamma(S) \setminus S$ contacts a $v \in S$ in the interval $[t_1 - K, t_1 - 1]$ is at least K/Δ. By linearity of expectation, the expected number of vertices in $S' \setminus S$ is at least

$$\frac{K}{\Delta}|\Gamma(S) \setminus S| \geqslant \frac{C_\beta Kd\,|S|}{\Delta} \geqslant \frac{8\gamma\,n\log n}{d}$$

Again, a Chernoff bound gives

$$\mathbf{Pr}\left[|S' \setminus S| \leqslant \frac{4\gamma\,n\log n}{d}\right] \leqslant \exp\left(-\frac{\gamma\,n\log n}{d}\right) \leqslant n^{-\gamma}.$$

Hence, $|S'| \geqslant |S' \setminus S| = \Omega(n\log(n)/d) = \omega(n/d)$ with probability $1 - n^{-\gamma}$ for $K \leqslant \Delta$. $\square$

5.5. Analysis for Graphs with Constant Degree

It remains to show that the quasirandom model also works well on expanding graphs with constant degree $d = \mathcal{O}(1)$. To do this, we apply Theorem 3.1 to see that, for any graph, the quasirandom model succeeds in $\Delta \cdot \mathrm{diam}(G)$ steps. The corresponding bound for the fully random model is $\mathcal{O}(\Delta\,(\mathrm{diam}(G) + \log n))$ with probability $1 - n^{-1}$ Feige et al. [1990, Theorem 2.2].

Naturally, the diameter of expanding graphs can be bounded easily as follows (cf. Hoory et al. [2006, p. 455] for a related result). Plugging Lemma 5.11 into the upper bound of $\Delta \cdot \mathrm{diam}(G)$ yields Theorem 5.1 for $d = \mathcal{O}(1)$.

Lemma 5.11. *For any expanding graph G with $d = \mathcal{O}(1)$, $\mathrm{diam}(G) = \mathcal{O}(\log n)$.*

Proof. Fix two vertices v and w. We show that $\Gamma^{\leqslant \mathcal{O}(\log n)}(v) \cup \Gamma^{\leqslant \mathcal{O}(\log n)}(w) \neq \emptyset$. Because G is connected, $|\Gamma^{\leqslant 3}(v)| \geqslant 3$. Now we choose $C_\alpha = d/2$ (which is valid since d is a constant) and proceed as in the proof of Lemma 5.3. By **(P1)** we again get Equation (1) for $k > 3$, and therefore by induction

$$|\Gamma^{\leqslant k}(v)| \geqslant 3\,(1 + C_\beta d)^{k-3}$$

Table II. Summary of the Broadcast Times for Almost All Random Graphs $G(n, p)$ with $p\,n = \log n + \omega(1)$ and $p\,n = \log n + \mathcal{O}(\log\log n)$.

Broadcast Time	
Random Model	Quasirandom Model
$\mathcal{O}(\log^2 n)$ with probability $\geqslant 1 - n^{-1}$ [Feige et al. 1990] $\Omega(\log^2 n)$ with probability $\geqslant n^{-1}$ [Feige et al. 1990] $\Omega(\log(n)\,\log\log n)$ with probability $\geqslant 1 - o(1)$ (Thm. 6.2)	$\mathcal{O}(\log n)$ with probability $\geqslant 1 - n^{-\gamma}$ $\forall\gamma = \mathcal{O}(1)$ (Thm. 4.2 and 5.1)

for all k with $k > 3$ and $|\Gamma^{\leqslant k-1}(v)| \leqslant n/2$. Therefore we can choose a k such that $|\Gamma^{\leqslant k}(v)| \geqslant n/2$ and $k = \mathcal{O}(\log n)$. Because analogously $|\Gamma^{\mathcal{O}(\log n)}(w)| \geqslant n/2$, we can conclude that there is a path of length $\mathcal{O}(\log n)$ from v to w. $\square$

6. LOWER BOUNDS FOR THE FULLY RANDOM MODEL ON SPARSE RANDOM GRAPHS

In this section, we discuss lower bounds for the fully random model on sparse random graphs. They show that the quasirandom model is superior on such graphs. Feige et al. [1990] proved the following bound.

THEOREM 6.1 ([FEIGE ET AL. 1990, THEOREM 4.1]). *Let $p = (\log n + f(n))/n$, where $f(n) = \omega(1)$ and $f(n) = \mathcal{O}(\log\log n)$. Then, for almost all random graphs $G(n, p)$, the broadcast time of the fully random model is $\Omega(\log^2 n)$ with probability at least n^{-1}.*

Theorem 6.1 stems simply from the fact that with high probability such graphs contain a vertex having constant degree with all neighbors having logarithmic degree. Although the expected time to inform such a vertex, given that all its neighbors are informed, is logarithmic, we need $\Omega(\log^2 n)$ rounds to do so with probability at least n^{-1}. The following result shows that we need $\omega(\log n)$ rounds with probability $1 - o(1)$ (see also Table II for a survey).

THEOREM 6.2. *Let $p = (\log n + f(n))/n$, where $f(n) = \omega(1)$ and $f(n) \leqslant C \log\log n$ for some constant $C \geqslant 1$. Then, for almost all random graphs $G(n, p)$, the broadcast time of the fully random model is $\Omega(\log(n)\,\log\log n)$ with probability $1 - o(1)$.*

PROOF. Fix an arbitrary vertex v. Then, for any $x \geqslant 1$ we have,

$$\begin{aligned}
\mathbf{Pr}\left[\deg(v) \leqslant x\right] \geqslant \mathbf{Pr}\left[\deg(v) = x\right] &= \binom{n-1}{x} p^x (1-p)^{n-1-x} \\
&\geqslant \left(\frac{n-1}{x}\right)^x \left(\frac{\log n}{n}\right)^x \left(1 - \frac{\log n + C\log\log n}{n}\right)^{n-1}.
\end{aligned}$$

Now, using the fact that $(1 - \frac{1}{n})^{n-1} \geqslant e^{-1}$ twice gives

$$\begin{aligned}
\mathbf{Pr}\left[\deg(v) \leqslant x\right] &\geqslant \left(\frac{n-1}{n}\right)^x \left(\frac{\log n}{x}\right)^x e^{-\log n - C\log\log n} \\
&\geqslant e^{-1} \left(\frac{\log n}{x}\right)^x e^{-\log n - C\log\log n}.
\end{aligned}$$

We now argue that, with high probability, we have sufficiently many vertices of this small degree. The basic idea is to inspect the degree of the vertices in a careful manner. First, in order to verify whether a vertex v_1 has degree larger than x or not, we only have to expose at most $x + 1$ edges incident to v_1. Then, the next vertex we pick will be a vertex for which we have not exposed any edge so far. Using this way of exposing the vertices allows us to use a Chernoff bound and to conclude that there are enough vertices of small degree.

More precisely, start with an arbitrary vertex $v_1 \in V$. In the first iteration, we check sequentially for all vertices $u \in V$ whether $\{v_1, u\} \in E$ until we know whether $\deg(v_1) \leqslant x$ holds or not. Although we may have to check for up to $n-1$ vertices u whether $\{v_1, u\}$ exists, we will never expose more than $x+1$ edges. This holds because after we have found $x+1$ edges incident to v_1, the event $\deg(v_1) \leqslant x$ does not hold. Then, in the second iteration, we pick a new vertex $v_2 \neq v_1$ for which we have not exposed the existence of any edge (but we may already know that $\{v_2, v_1\} \notin E$). Again, we sequentially check for all vertices $u \in V$ whether $\{v_2, u\} \in E$ holds until we know whether $\deg(v_2) \leqslant x$ holds or not. Observe that we can continue in this manner as long as there is a new vertex v_i for which we have not exposed the existence of any edge. Since in each iteration at most $x+1$ edges are exposed, the number of vertices with no exposed edge is reduced by at most $x+2$ per iteration. As a consequence, the whole procedure can be run for at least $n/(x+2)$ iterations. In each iteration $1 \leqslant i \leqslant n/(x+2)$, we have

$$\mathbf{Pr}\left[\deg(v_i) \leqslant x\right] \geqslant e^{-1} \left(\frac{\log n}{x}\right)^x e^{-\log n - C\log\log n},$$

by the same reasoning as given previously.

Let X be the number of vertices with degree at most x. By the preceding arguments, it follows that X is stochastically larger (cf. Definition A.1 for a definition of stochastically larger) than the sum of $n/(x+2)$ independent Bernoulli-random variables, each of which has success probability $e^{-1}(\frac{\log n}{x})^x e^{-\log n - C\log\log n}$. Therefore, it follows by a Chernoff bound (Lemma A.1) that

$$\mathbf{Pr}\left[X \leqslant \frac{1}{2}\,\mathbf{E}\left[X\right]\right] \leqslant e^{-(1/2)^2\,\mathbf{E}[X]/2}. \tag{5}$$

Now choose $x := (\log n)^\varepsilon$ for an arbitrary constant $0 < \varepsilon < 1$. By the previous paragraph, we obtain

$$\begin{aligned}\mathbf{E}\left[X\right] &\geqslant \frac{n}{(\log n)^\varepsilon + 2}\, e^{-1} \left(\frac{\log n}{(\log n)^\varepsilon}\right)^{(\log n)^\varepsilon} e^{-\log n - C\log\log n}\\ &\geqslant \frac{1}{3}(\log n)^{-\varepsilon - C + (1-\varepsilon)(\log n)^\varepsilon} = (\log n)^{\Omega((\log n)^\varepsilon)}.\end{aligned}$$

Plugging this into Equation (5), we obtain

$$\mathbf{Pr}\left[X \leqslant (\log n)^{\Omega((\log n)^\varepsilon)}\right] = o(1).$$

By Cooper and Frieze [2007, Lemma 1, Property 2], we know that for almost all random graphs, any two vertices with a degree of less than $\log n/20$ have a distance of at least $\log n/(\log\log n)^2$ from each other. Hence, all neighbors of vertices in X have a degree of more than $\log n/20$. In particular, the time until a vertex $u \in X$ gets contacted by a fixed neighbor $v \in N(u)$ is stochastically larger than a geometric random variable with parameter $\log n/20$. Hence, the time until u gets contacted by any of its neighbors is stochastically larger than the minimum of $\deg(u) \leqslant x = (\log n)^\varepsilon$ independent such geometric variables. Since any two vertices in X have a distance of at least 3, these times are independent for all $u \in X$.

Now recall that $\mathcal{R}(G)$ is the random variable describing the runtime of the fully random model. Furthermore, let $\mathsf{Geo}(p)$ be the geometric distribution defined by

$\mathbf{Pr}\left[\mathsf{Geo}(p) = i\right] = p \cdot (1-p)^i$ for any integer $i \geqslant 0$. Denoting with $\succeq$ "stochastically larger" and using Lemma A.4, we obtain

$$\begin{aligned}\mathcal{R}(G) &\succeq \max_{u \in X} \min_{v \in N(x)} \left\{\mathsf{Geo}(20/\log n)\right\} \\ &\succeq \max_{u \in X} \left\{\mathsf{Geo}\left(1 - \prod_{v \in N(x)} (1 - 20/\log n)\right)\right\} \\ &\succeq \max_{u \in X} \left\{\mathsf{Geo}\left(1 - (1 - 20/\log n)^{(\log n)^{\varepsilon}}\right)\right\} \\ &\succeq \max_{i=1}^{(\log n)^{\Omega((\log n)^{\varepsilon})}} \left\{\mathsf{Geo}\left(1 - e^{-20(\log n)^{\varepsilon-1}}\right)\right\}.\end{aligned}$$

Hence

$$\begin{aligned}\mathbf{Pr}\left[\mathcal{R}(G) \leqslant t\right] &\leqslant \mathbf{Pr}\left[\mathsf{Geo}\left(1 - e^{-20(\log n)^{\varepsilon-1}}\right) \leqslant t\right]^{(\log n)^{\Omega((\log n)^{\varepsilon})}} \\ &= \left(1 - \left(e^{-20(\log n)^{\varepsilon-1}}\right)^t\right)^{(\log n)^{\Omega((\log n)^{\varepsilon})}} \\ &\leqslant \exp\left(-e^{-20(\log n)^{\varepsilon-1}t} \, (\log n)^{\Omega((\log n)^{\varepsilon})}\right).\end{aligned}$$

Setting $t = c \log n \log\log n$ with a sufficiently small constant c finally gives

$$\begin{aligned}\mathbf{Pr}\left[\mathcal{R}(G) \leqslant t\right] &\leqslant \exp\left(-(\log n)^{-20c(\log n)^{\varepsilon}} \, (\log n)^{\Omega((\log n)^{\varepsilon})}\right) \\ &= \exp\left(-(\log n)^{\Omega((\log n)^{\varepsilon})}\right). \quad \square\end{aligned}$$

7. QUASIRANDOM RUMOR SPREADING ON HYPERCUBES

In this section, we analyze the quasirandom model on hypercubes. We prove that the quasirandom model informs all vertices in $\mathcal{O}(\log n)$ rounds with high probability. This extends a corresponding runtime bound of $\mathcal{O}(\log n)$ for the fully random model in Feige et al. [1990]. The difficulty in our analysis is that the hypercube is not an expanding graph (cf. Theorem 4.13), and, also, an application of the bound of Theorem 3.1 yields only a much weaker upper bound of $\mathcal{O}(\log^2 n)$.

We now state and prove our runtime bound for the quasirandom model on hypercubes. Finally, we also examine the failure probability more closely to reveal that there is again a slight superiority of the quasirandom model over the fully random model (Section 7.4).

THEOREM 7.1. *The broadcast time of the quasirandom model on the hypercube is* $\mathcal{O}(\log n)$ *with probability* $1 - n^{-\Omega(\log n)}$.

Similarly to the proof for expanding graphs in Section 5, the analysis consists of a forward part and a backward part. Although the analysis of the forward part borrows several concepts from the analysis of the fully random model [Feige et al. 1990], the idea of analyzing the process in reversed order was not used in that work.

The forward part informs sufficiently many vertices in $\mathcal{O}(\log n)$ time. The backward part shows that if there is an uninformed vertex, then $\mathcal{O}(\log n)$ steps earlier, every ball of small radius in the hypercube contains at least one uninformed vertex. To prove that one of these uninformed vertices gets informed eventually, we need a third part in between, which we call coupling. A graphical illustration of our proof can be found in Figure 1 on page 27.

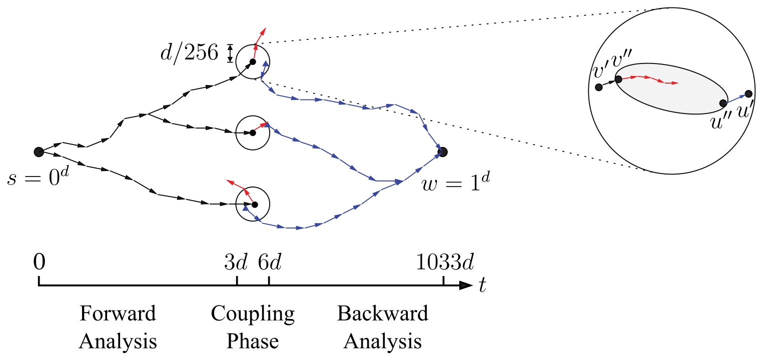

Fig. 1. The left side contains a sketch of the proof of Theorem 7.1. The black circles represent $I'_{3d,}$, and the triangles represent $\bigcup_{v \in I'_{3d,}} \Phi(v)$. The right side illustrates the analysis of the coupling part. We find two vertices v'' and u'' such that every shortest path between them is included in a subcube of vertices whose initially contacted neighbors are not exposed.

To formally prove Theorem 7.1, we assume that the following three lemmas hold. We state them here and prove them in the remainder of this section. Recall that $n = 2^d$.

LEMMA 7.2. *The probability that the quasirandom rumor-spreading process started in a fixed vertex s informs $2^{d/6}$ vertices in $3d$, steps is at least $1 - n^{-\Omega(\log n)}$.*

Let $s = 0^d$ be initially informed. By Lemma 7.2, at least $2^{d/6}$ vertices get informed in $3d$, with probability at least $1 - n^{-\Omega(\log n)}$. Now fix an arbitrary vertex $w \in V$. Recall that $U_{[t_1,t_2]}(w)$ is the set of vertices that reach the vertex w within the time interval $[t_1, t_2]$ (cf. definition on page 7).

LEMMA 7.3. *For any vertex w and $t_2 = 1033d$, with probability at least $1 - n^{-\Omega(\log n)}$, there is for every vertex v a vertex $u(v) \in U_{[6d,t_2]}(w)$ with $\mathrm{dist}(u, v) \leqslant d/256$.*

By applying Lemma 7.3, there is, with probability at least $1-n^{-\Omega(\log n)}$ for each $v \in I_{3d,}$ a vertex $u(v) \in U_{[6d,t_2]}(w)$ with $\mathrm{dist}(u, v) \leqslant d/256$.

LEMMA 7.4. *Let s be the initially informed vertex and w be an arbitrary vertex. Assume that the following two conditions hold:*

—*there are at least $2^{d/6}$ informed vertices at step $3d$, and*
—*there is for every vertex v a vertex $u(v) \in U_{[6d,t_2]}(w)$ with $\mathrm{dist}(u, v) \leqslant d/256$ and $t_2 = 1033d$.*

Then, with probability $1-e^{-\mathrm{poly}(n)}$, at least one vertex in $U_{[6d,t_2]}(w)$ is informed at step $6d$.

Now, if the two former conditions hold, Lemma 7.4 implies that a vertex in $U_{[6d,t_2]}(w)$ gets informed with (conditional) probability at least $1 - n^{-\Omega(\log n)}$. By definition, this implies that the vertex w gets informed at step t_2. Taking the union bound over the success of the forward and backward part (Lemma 7.2 and Lemma 7.3), it follows that at step t_2 the vertex w gets informed with probability at least $1 - n^{-\Omega(\log n)}$. Taking the union bound over all possible vertices $w \in V$ yields Theorem 7.1.

7.1. Proof of the Forward Analysis

In this section, we prove Lemma 7.2.

PROOF OF LEMMA 7.2. By symmetry, we may assume that $s = 0^d$ is initially informed. Let L_i be the set of vertices with $\|x\|_1 = i$. Note that after two phases of d steps each, we have $I_{2d} = \{s\} \cup L_1 \cup L_2$.

Consider some time-step $t \geqslant 2d$. Assume that all initially contacted neighbors of $I_t \cap L_i$ are still to be chosen uniformly at random for $i \geqslant 2$. Notice that the number of edges between $I_t \cap L_i$ and L_{i+1} is $|E(I_t \cap L_i, L_{i+1})| = \sum_{v \in L_{i+1}} \deg_{I_t \cap L_i}(v) = |I_t \cap L_i|\,(d-i)$. Our goal is to show that a large set of vertices in L_{i+1} will be informed after a phase of four additional steps. The probability that a vertex $v \in L_{i+1}$ is still uninformed after this phase is

$$\mathbf{Pr}\left[v \notin I_{t+4}\right] \leqslant \prod_{u \in \Gamma(v) \cap I_t \cap L_i} \left(1 - \frac{4}{d}\right) = \left(1 - \frac{4}{d}\right)^{\deg_{I_t \cap L_i}(v)}.$$

By linearity of expectations we get

$$\mathbf{E}\left[|I_{t+4} \cap L_{i+1}|\right] = \sum_{v \in L_{i+1}} \mathbf{Pr}\left[v \in I_{t+4}\right] \geqslant \sum_{v \in L_{i+1}} 1 - \left(1 - \frac{4}{d}\right)^{\deg_{I_t \cap L_i}(v)}$$
$$\geqslant \sum_{v \in L_{i+1}} 1 - \exp\left(-\frac{4 \deg_{I_t \cap L_i}(v)}{d}\right).$$

Let us now assume that $1 \leqslant i \leqslant d/4 - 1$. Then, since $\deg_{I_t \cap L_i}(v) \leqslant i + 1$ for $v \in L_{i+1}$ and $1 + \frac{x}{2} \geqslant e^x$ for any $-1 \leqslant x \leqslant 0$, we get

$$\mathbf{E}\left[|I_{t+4} \cap L_{i+1}|\right] \geqslant \sum_{v \in L_{i+1}} \frac{2 \deg_{I_t \cap L_i}(v)}{d} = \frac{2}{d}\,|I_t \cap L_i|\,(d-i) = 2\,\frac{d-i}{d}\,|I_t \cap L_i|.$$

Since any vertex of $|I_t \cap L_i|$ can only inform at most four vertices within four steps, an application of Azuma's inequality (cf. Lemma A.2) gives, for any constant $0 < \varepsilon \leqslant 2/3$,

$$\mathbf{Pr}\left[|I_{t+4} \cap L_{i+1}| \leqslant (2-\varepsilon)\,\frac{d-i}{d}\,|I_t \cap L_i|\right]$$
$$\leqslant \exp\left(-\frac{(\varepsilon\,\frac{d-i}{d}\,|I_t \cap L_i|)^2}{16\,|I_t \cap L_i|}\right) = \exp(-\Omega(d^2)) = n^{-\Omega(\log n)},$$

as long as $|I_t \cap L_i| \geqslant \frac{d(d-1)}{2}$ holds. Observe that if the condition $|I_t \cap L_i| \geqslant \frac{d(d-1)}{2}$ holds initially, then $|I_{t+4} \cap L_{i+1}| \geqslant (2-\varepsilon)\frac{d-i}{d}|I_t \cap L_i|$ implies that $|I_{t+4} \cap L_{i+1}| \geqslant \frac{d(d-1)}{2}$, since $(2-\varepsilon)\frac{d-i}{d} \geqslant (2-\varepsilon)\frac{3}{4} \geqslant 1$ by definition of i and ε.

Recall that we first spent $2d$ steps in the first two phases to inform L_2 completely. Then, in the preceding analysis, we spent, for each level i with $2 \leqslant i \leqslant d/4 - 1$, a phase of exactly four steps. Hence, the total time consumption is

$$2d + (d/4 - 2) \cdot 4 \leqslant 3d.$$

Now, taking the union bound over all levels $2 \leqslant i \leqslant d/4 - 1$, with probability $1 - (d/4 - 1)\,n^{-\Omega(\log n)} = 1 - n^{-\Omega(\log n)}$ it holds that

$$|I_{4d} \cap L_{d/4}| \geqslant \frac{d\,(d-1)}{2} \prod_{i=2}^{d/4-1} \left((2-\varepsilon)\,\frac{d-i}{d}\right)$$
$$= \frac{d\,(d-1)}{2}\,(2-\varepsilon)^{d/4-2} \prod_{i=2}^{d/4-1} \left(1 - \frac{i}{d}\right).$$

We now use the fact that $(1-x)^{1/x}$ is nonincreasing in $0 < x < 1$, implying $(1-x) \geqslant 4^{-x}$ for any $x \leqslant 1/4$. Plugging this into the previous inequality yields

$$|I_{4d} \cap L_{d/4}| \geqslant (2-\varepsilon)^{d/4}\, 4^{-\sum_{i=2}^{d/4-1} \frac{i}{d}} \geqslant (2-\varepsilon)^{d/4}\, 4^{-d/32} \geqslant 2^{d/6},$$

if $\varepsilon > 0$ is a sufficiently small constant. $\square$

7.2. Proof of the Backward Analysis

In this section, we prove Lemma 7.3. We use the notation that $x[j]$ denotes the j-th bit of a vertex $x \in V$.

PROOF OF LEMMA 7.3. We now analyze the propagation of the rumor in the reverse order. Due to the symmetry of H, we restrict our attention to the case $w = 1^d$.

Let us first consider the case where $v = 0^d$. So we have to show that $U_{[6d,t_2]}(w)$ contains a vertex u such that $\mathrm{dist}(0^d, u) \leqslant d/256$ with probability at least $1 - n^{-\Omega(\log n)}$. In order to achieve such a large success probability, we construct $d/512$ vertex-disjoint paths that start from a vertex in $\Gamma(w)$ and move toward the vertex v. For each neighbor of w that differs from w in one of the last $d/512$ bits, we associate a path starting from that vertex and moving toward the vertex v. The disjointness is ensured by not allowing the path to change any of the last $d/512$ bits.

First note that $U_{[t_2-d,t_2]}(w) \supseteq \Gamma(w)$, since within a time interval of d steps, every neighbor of w contacts w. Let $\mathcal{J} := [(511/512)d, d]$. For each $j \in \mathcal{J}$, we define a set of vertices

$$V(j) := \left\{x \in \{0,1\}^d \text{ with } x[j] = 0 \text{ and } x[i] = 1 \text{ for } i \in [(511/512)d, d]\backslash\{j\}\right\}.$$

For each $j \in \mathcal{J}$ we consider a path $P(j) = (v_1, v_2, \ldots, v_\ell) \subseteq V(j)$ of length $\ell := (255/256)d$, which is defined inductively as follows:

—The first vertex of $P(j)$ is defined by $v_1 \in \Gamma(w) \cap V(j)$.
—If s_i denotes the time-step when $P(j)$ has reached the vertex v_i, then $P(j)$ is extended to a vertex $v_{i+1} \in \Gamma(v_i) \cap V(j)$ with $\|v_{i+1}\|_1 = d - i - 1$ such that v_{i+1} is the last vertex before time-step s_i that contacts v_i.

Fix an arbitrary $j \in \mathcal{J}$ and consider the path $P(j)$. Recall that $v_1 \in U_{[t_2-d,t_2]}(w)$. Fix any i with $1 \leqslant i \leqslant \ell$ and consider the vertex v_i. Note that there are $d - i - (1/512)d$ vertices $u \in \Gamma(v_i) \cap V(j)$ with $\|u\|_1 = \|v_i\|_1 - 1$. Let us denote by $\Delta_i(v_i)$ the waiting time (going back in time) until such a fixed vertex u contacts v_i, in symbols,

$$\Delta_i(u, v_i) := s_i - \max\{s \leqslant s_i - 1 : u \in U_{[s,s_i]}(v_i)\}.$$

Note that $\Delta_i(u, v_i)$ is a uniform random variable in $\{1, \ldots, d\}$. In particular, the distribution is the same for every u, and since the initially contacted neighbors are chosen independently and uniformly at random, $\{\Delta_i(u, v_i) : u \in \Gamma(v_i) \cap V(j), \|u\|_1 = \|v_i\|_1 - 1\}$ is a set of mutually independent random variables. The waiting time Δ_i until the first vertex $u \in \Gamma(v_i) \cap V(j)$ with $\|u\|_1 = \|v_i\|_1 - 1$ contacts v_i satisfies

$$\Delta_i := \min_{\substack{u \in \Gamma(v_i) \cap V(j): \\ \|u\|_1 = \|v_i\|_1 - 1}} \Delta_i(u, v_i).$$

To bound this random variable, let $X_{i,u} \sim \mathsf{Geo}(1/d)$; that is, a geometric random variable with parameter $1/d$. By Lemma A.4, the minimum of $d - i - (1/512)d$ independent geometric random variables with parameter $1/d$ is itself a geometric random variable X_i with parameter

$$1 - \left(1 - \frac{1}{d}\right)^{d-i-(1/512)d} \geqslant 1 - \exp(-1/512) \geqslant 1 - \frac{1}{1/512 + 1} = \frac{1}{513}.$$

Hence, with "$\preceq$" denoting "stochastically smaller than" we obtain by Lemma A.3 that

$$\Delta_i = \min_{\substack{u \in \Gamma(v_i) \cap V(j): \\ \|u\|_1 = \|v_i\|_1 - 1}} \Delta_i(u, v_i) \preceq \min_{\substack{u \in \Gamma(v_i) \cap V(j): \\ \|u\|_1 = \|v_i\|_1 - 1}} X_{i,u} = X_i.$$

Hence, the time $\Delta(j) := \sum_{i=1}^{\ell} \Delta_i$ until we reach the end of $P(j)$ is stochastically smaller than $\sum_{i=1}^{\ell} X_i$, where the X_i's are independent geometric random variables with parameter $1/513$.

Let us first note that $\mathbf{E}[X_i] \leqslant 513$ and therefore with $X := \sum_{i=1}^{\ell} X_i$,

$$\mathbf{E}[X] = \sum_{i=1}^{\ell} \mathbf{E}[X_i] \leqslant 513\, d.$$

Now we apply a Chernoff bound for a sum of independent geometric random variables (Lemma A.5 with $\varepsilon := 1$) to obtain

$$\mathbf{Pr}[X \geqslant 1026\, d] \leqslant \exp\left(-\frac{1}{4}\ell\right),$$

and since $\Delta(j) \preceq X$,

$$\mathbf{Pr}[\Delta(j) \geqslant 1026\, d] \leqslant \exp\left(-\frac{1}{4}\ell\right).$$

Hence, with probability $1 - \exp(-\frac{1}{4}\ell)$, the endpoint of a path $P(j)$ for a fixed j contacts w within the time interval $[6d, t_2]$.

Note that $\{\Delta(j) : j \in \mathcal{J}\}$ is a set of independent random variables, since for any $j_1, j_2 \in \mathcal{J}$ with $j_1 \neq j_2$, the vertex sets $V(j_1)$ and $V(j_2)$ are disjoint. Using this independence, we can lower bound the probability that there is a vertex u with $\|u\|_1 \leqslant d/256$ and $u \in U_{[6d,t_2]}(w)$ by

$$1 - \left(\exp\left(-\frac{1}{4}\ell\right)\right)^{|\mathcal{J}|} \geqslant 1 - e^{-\Omega(d^2)} = 1 - n^{-\Omega(\log n)}.$$

So far, we have considered the case where $v = 0^d$. With the same arguments, we can prove that for an arbitrary vertex v there is a vertex $u(v)$ satisfying $\operatorname{dist}(u(v), v) \leqslant d/256$ and $u(v) \in U_{[6d,t_2]}(w)$ with probability $1 - n^{-\Omega(\log n)}$. It follows by a union bound that with probability $1 - n^{-\Omega(\log n)}$, there is for every vertex $v \in V(G)$ a vertex $u(v) \in U_{[6d,t_2]}(w)$ with $\operatorname{dist}(v, u(v)) \leqslant d/256$.

7.3. Proof of the Coupling Part

In this section, we prove Lemma 7.4.

Proof of Lemma 7.4. Let w be an arbitrary, fixed vertex. By the first condition in Lemma 7.4, we have $|I_{3d,}| \geqslant 2^{d/6}$. By definition of the hypercube, there are for every vertex u exactly $\sum_{k=0}^{d/64} \binom{d}{k}$ vertices with distance at most $d/64$ to u. Hence, there is subset $I'_{3d,} \subseteq I_{3d,}$ such that two vertices in $I'_{3d,}$ have a distance at least $d/64$ from each other which is of size

$$\frac{2^{d/6}}{\sum_{k=0}^{d/64} \binom{d}{k}} \geqslant \frac{2^{d/6}}{(64e)^{d/64}} \geqslant \frac{2^{d/6}}{(2^8)^{d/64}} = 2^{d/24},$$

where we have used the inequality $\sum_{i=0}^{m} \binom{n}{i} \leqslant (\frac{e\,n}{m})^m$.

By our second condition in Lemma 7.4, there is for each vertex $v \in I'_{3d,}$ at least one vertex $u = u(v) \in U_{[6d,t_2]}(w)$ such that $\mathrm{dist}(u, v) \leqslant d/256$.

Let $\Phi : I'_{3d,} \to U_{[6d,t_2]}(w)$ be a function that assigns each vertex $v \in I'_{3d,}$ a vertex $u = u(v) \in U_{[6d,t_2]}(w)$ such that $\mathrm{dist}(u, v) \leqslant d/256$. Using the fact that two vertices in $I'_{3d,}$ have distance at least $d/64$ from each other, we observe that Φ is an injective function.

Let us now fix a pair of vertices $v \in I'_{3d,}$ and $\Phi(v) \in U_{[6d,t_2]}(w)$. Note that the set of all shortest paths between v and $\Phi(v)$ form a subcube $H' = H'(v, \Phi(v))$ whose dimension is equal to the distance between v and $\Phi(v)$. Now choose a pair of vertices $v' \in H' \cap I_{3d,}$ and $u' \in H' \cap U_{[6d,t_2]}(w)$ such that $\mathrm{dist}(v', u')$ is minimized. Our aim is to lower bound the probability that v' reaches u' within the time interval $[3d, 6d]$.

First, let us assume that $\mathrm{dist}(v', u') \leqslant 3$. In this case, u' is informed within $3d$ steps with probability 1. Otherwise, we have $\mathrm{dist}(v', u') \geqslant 4$. In this case, let $v'' \in \Gamma(v')$ and $u'' \in \Gamma(u')$ be two vertices such that $\mathrm{dist}(v'', u'') = \mathrm{dist}(v', u') - 2$. Note that $v'' \in I_{4d}$ and $u'' \in U_{[5d,t_2]}(w)$. By our construction, every vertex on a shortest path between v'' and u'' (except u'') has distance at least 1 to $I_{3d,}$ and distance at least 2 to $U_{[6d,t_2]}(w)$. Hence, for each vertex on such a shortest path, the initially contacted neighbor is still chosen uniformly at random and independently of all other vertices.

Similarly to the proof of Lemma 7.3, we lower bound the probability that there exists a path $P = P(v) = (v_1 = v'', v_2, \ldots, v_{\mathrm{dist}(v'',u'')} = u'')$, which satisfies the following two conditions for any $1 \leqslant i < \mathrm{dist}(v'', u'')$:

—v_{i+1} is closer to u'' than v_i and
—v_i informs v_{i+1} at step $4d + i$.

Note that once the rumor has reached a vertex v_i for the first time, the vertex v_i forwards it to a vertex v_{i+1} closer to v'' with probability at least $(\mathrm{dist}(v'', u'') - i + 1)/d$. Repeating this argument gives the following lower bound for the existence of P:

$$\prod_{i=1}^{d/256} \frac{i}{d} = \frac{(d/256)!}{d^{d/256}} \geqslant \frac{d^{d/256}}{(768\,d)^{d/256}} \geqslant (2^{-10})^{d/256} \geqslant 2^{-d/25},$$

where we have used the fact that $n! \geqslant (n/3)^n$ for any integer n in the left inequality.

Our next claim is that $\{P(v) \text{ exists} : v \in I'_{3d,}\}$ is a set of mutually independent events. In order to prove this, let us consider two arbitrary vertices $v_1, v_2 \in I'_{3d,}$, $v_1 \neq v_2$. Recall that by definition of $I'_{3d,}$, $\mathrm{dist}(v_1, v_2) \geqslant d/64$. Because every vertex on a shortest path between v_i and $\Phi(v_i)$ has a distance of at most $d/256$ to v_i, it holds by the triangle inequality that the two paths $P(v_1)$ and $P(v_2)$ always have a distance of at least $d/128$ from each other, which proves the claimed independence.

Using this, we can lower bound the probability that at least one $P(v)$ exists by

$$1 - \left(1 - 2^{-d/25}\right)^{2^{d/24}} \geqslant 1 - \exp\left(-2^{d/(24\cdot 25)}\right) = 1 - \exp\left(-\operatorname{poly}(n)\right).$$

If there is a $v \in I'_{3d,}$ for which $P(v)$ exists, then we know that there is a vertex $v''(v) \in I'_{4d}$ which reaches a vertex $u''(v) \in U_{[5d,t_2]}(w)$ within the time interval $[3d, 6d]$. This implies that $u(v) \in U_{[6d,t_2]}(w)$ is informed at step $6d$, and, as a consequence, w will become informed at step t_2.

7.4. Failure Probability

We now examine the probabilities in the runtime bounds for the hypercube more closely. Recall that the runtime bound of $\mathcal{O}(\log n)$ for the quasirandom model holds with probability at least $1 - n^{-\Omega(\log n)}$. In the fully random model, however, a fixed vertex remains uninformed for x steps with probability at least $(1 - 1/d)^{dx} \geqslant 4^{-x}$. Hence, the runtime of the fully random model is at least $\rho \cdot \log_2 n$ with probability at least $n^{-2\rho}$ for any value of $\rho \geqslant 1$. Hence, if $\rho = (c/2)\log_2 n$ for some constant $c > 0$, this shows that the time for the fully random model to inform all n vertices with probability at least $1 - n^{-c\log_2 n}$ is at least $(c/2)(\log_2 n)^2 = \Omega(\log(n)^2)$. This should be compared with our upper bound of $\mathcal{O}(\log n)$ for the quasirandom model, which holds with probability at least $1 - n^{-\Omega(\log n)}$.

8. CONCLUSION AND OUTLOOK

In this article, we proposed and investigated a quasirandom analogue of the classical push model for spreading a rumor to all vertices of a network.

We showed that for many network topologies, after $\Theta(\log n)$ iterations, all vertices are informed with probability $1 - \mathcal{O}(\text{poly}(n))$. Hence, the quasirandom model achieves asymptotically the same bounds as the random one, or even better ones (e.g., for random graphs with p close to $\log(n)/n$).

This work is also interesting from the methodological point of view. Our proofs show, in particular, that the difficulties usually invoked by highly dependent random experiments can be overcome. From the general perspective of using randomized methods in computer science, our results, like a number of other recent results, can be viewed as suggesting that choosing the right dose of randomness might be a fruitful topic for further research.

An interesting open problem is to analyze the quasirandom push model on other graph classes. A natural candidate would be the class of regular graphs with constant conductance, for which it is known that the classical push model spreads a rumor in $\mathcal{O}(\log n)$ rounds [Giakkoupis 2011; Chierichetti et al. 2010a]. Another interesting target are preferential attachment graphs. Here, Doerr et al. [2011a] have shown that the fully random push-pull model has a broadcast time of $\Theta(\log n)$, whereas the variant with contactees chosen uniformly at random from all neighbors except the previous contactee has a broadcast time of only $\Theta(\log n/\log\log n)$. Since the quasirandom protocol automatically avoids the previous contactee, it seems likely that it also has this superior broadcast time.

Note, however, that it is not true that the quasirandom model always performs at least as good as the fully random model. For instance, consider the graph consisting of two cliques of size $n/2 - 1$ and an extra vertex that is connected to all other $n/2 - 2$ vertices. On this graph, the fully random model spreads a rumor in $\mathcal{O}(\log n)$ rounds with high probability, whereas the quasirandom model needs $\Omega(n)$ rounds with probability at least $1/4$ for appropriately chosen lists.

A. PROBABILISTIC TAIL BOUNDS USED FOR OUR ANALYSIS

As our analysis heavily relies on probabilistic tails bounds, we summarize them here for reference. The following bound can be found, e.g., in the textbook of Mitzenmacher and Upfal [2005].

LEMMA A.1 (CHERNOFF BOUNDS FOR SUMS OF BERNOULLI VARIABLES). *Let X_i, $1 \leqslant i \leqslant n$, be independent random variables. Let $X = \sum_{i=1}^{n} X_i$, $0 < p < 1$ and $\delta > 0$. If $\mathbf{Pr}\left[X_i = 1\right] = p$ and $\mathbf{Pr}[X_i = 0] = 1 - p$ for all $i \in \{1, \ldots, n\}$, then*

$$\mathbf{Pr}\left[X \leqslant (1-\delta)\,\mathbf{E}\left[X\right]\right] \leqslant \exp(-\delta^2\,\mathbf{E}\left[X\right]/2),$$

$$\mathbf{Pr}\left[X \geqslant (1+\delta)\,\mathbf{E}\left[X\right]\right] \leqslant \exp(-\min\{\delta, \delta^2\}\,\mathbf{E}\left[X\right]/3).$$

We also use the following concentration bound, which is also called the method of bounded differences McDiarmid [1989, Lemma 1.2].

LEMMA A.2 (AZUMA'S INEQUALITY). *Let $X_i : \Omega_i \to \mathbb{R}$, $1 \leqslant i \leqslant n$, be mutually independent random variables. Let $f : \prod_{i=1}^{n} \Omega_i \to \mathbb{R}$ satisfy the Lipschitz condition*

$$|f(\mathbf{x}) - f(\mathbf{x}')| \leqslant c_i$$

where $\mathbf{x}$ and $\mathbf{x}'$ differ only in the i-th coordinate, $1 \leqslant i \leqslant n$. Let Y be the random variable $f(X_1, \ldots, X_n)$. Then for any $t \geqslant 0$,

$$\mathbf{Pr}\left[Y > \mathbf{E}\left[Y\right] + t\right] \leqslant \exp\left(-2t^2 \Big/ \sum_{i=1}^{n} c_i^2\right).$$

We shall also use the concept of stochastic domination between random variables.

Definition A.1. A random variable X is *stochastically smaller* than Y, if for all $k \in \mathbb{R}$, $\mathbf{Pr}\left[X \geqslant k\right] \leqslant \mathbf{Pr}\left[Y \geqslant k\right]$. In this case, we also write $X \preceq Y$.

We list two obvious facts about stochastic domination.

LEMMA A.3. *Let X_1, X_2 be two independent random variables and let Y_1, Y_2 be two additional independent random variables with $X_1 \preceq Y_1$ and $X_2 \preceq Y_2$. Then,*

—$X_1 + X_2 \preceq Y_1 + Y_2$ *and*
—$\min\{X_1, X_2\} \preceq \min\{Y_1, Y_2\}$.

We continue with a simple fact about the geometric distribution.

LEMMA A.4. *Let $X_1, X_2, \ldots, X_n$ be n independent geometric random variables each with parameter $0 < p < 1$. Then $X := \min_{i=1}^{n} X_i$ is a geometric random variable with parameter $(1 - (1 - p)^n)$.*

We use the following standard Chernoff bound for sums of geometric random variables from Dubhashi and Panconesi [2009, Problem 3.6].

LEMMA A.5 (CHERNOFF BOUND FOR SUMS OF GEOMETRIC VARIABLES). *Let $Y_1, Y_2, \ldots, Y_n$ be independent geometric random variables, each with parameter $p > 0$. Let $Y := \sum_{i=1}^{n} Y_i$. Then for any $\varepsilon > 0$,*

$$\mathbf{Pr}\left[Y \geqslant (1 + \varepsilon)\frac{n}{p}\right] \leqslant \exp\left(-\frac{\varepsilon^2}{2(1 + \varepsilon)}n\right).$$

REFERENCES

S. Angelopoulos, B. Doerr, A. Huber, and K. Panagiotou. 2009. Tight bounds for quasirandom rumor spreading. *Electronic Journal of Combinatorics* 16, #R102.

A.-L. Barabási and R. Albert. 1999. Emergence of scaling in random networks. *Science* 286, 509–512.

R. D. Barve, E. F. Grove, and J. S. Vitter. 1997. Simple randomized mergesort on parallel disks. *Parallel Computing* 23, 4–5, 601–631.

H. Baumann, P. Fraigniaud, H. Harutyunyan, and R. de Verclos. 2012. The worst case behavior of randomized gossip. In *Proceedings of the 9th Annual Conference on Theory and Applications of Models of Computation (TAMC)*. 330–345.

P. Berenbrink, R. Elsässer, and T. Sauerwald. 2010. Communication complexity of quasirandom rumor spreading. In *Proceedings of the 18th European Symposium on Algorithms (ESA)*. 134–145.

B. Bollobás, O. Riordan, J. Spencer, and G. Tusnády. 2001. The degree sequence of a scale-free random graph process. *Random Structures & Algorithms* 18, 279–290.

S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. 2006. Randomized gossip algorithms. *IEEE Transactions on Information Theory and IEEE/ACM Transactions on Networking* 52, 2508–2530.

A. Z. Broder, A. M. Frieze, E. Shamir, and E. Upfal. 1994. Near-perfect token distribution. *Random Structures and Algorithms* 5, 4, 559–572.

A. Z. Broder, A. M. Frieze, S. Suen, and E. Upfal. 1998. Optimal construction of edge-disjoint paths in random graphs. *SIAM Journal on Computing* 28, 2, 541–573.

F. Chierichetti, S. Lattanzi, and A. Panconesi. 2010a. Almost tight bounds for rumour spreading with conductance. In *Proceedings of the 42nd ACM Symposium on Theory of Computing (STOC)*. 399–408.

F. Chierichetti, S. Lattanzi, and A. Panconesi. 2010b. Rumour spreading and graph conductance. In *Proceedings of the 21st ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 1657–1663.

F. Chierichetti, S. Lattanzi, and A. Panconesi. 2011. Rumor spreading in social networks. *Theoretical Computer Science* 412, 2602–2610.

F. Chung and L. Lu. 2002. Connected components in random graphs with given expected degree sequences. *Annals of Combinatorics* 6, 2, 125–145.

C. Cooper, M. Dyer, and C. Greenhill. 2007a. Sampling regular graphs and a peer-to-peer network. *Combinatorics, Probability & Computing* 16, 4, 557–593.

C. Cooper, R. Elsässer, H. Ono, and T. Radzik. 2012. Coalescing random walks and voting on graphs. In *Proceedings of the 31st ACM Symposium on Principles of Distributed Computing (PODC)*. 47–56.

C. Cooper and A. M. Frieze. 2007. The cover time of sparse random graphs. *Random Structures and Algorithms* 30, 1–2, 1–16.

J. Cooper, B. Doerr, T. Friedrich, and J. Spencer. 2010. Deterministic random walks on regular trees. *Random Structures and Algorithms* 37, 3, 353–366.

J. Cooper, B. Doerr, J. Spencer, and G. Tardos. 2007b. Deterministic random walks on the integers. *European Journal of Combinatorics* 28, 8, 2072–2090.

J. Cooper and J. Spencer. 2006. Simulating a random walk with constant error. *Combinatorics, Probability and Computing* 15, 815–822.

A. J. Demers, D. H. Greene, C. Hauser, W. Irish, J. Larson, S. Shenker, H. E. Sturgis, D. C. Swinehart, and D. B. Terry. 1988. Epidemic algorithms for replicated database maintenance. *Operating Systems Review* 22, 1, 8–32.

B. Doerr and M. Fouz. 2011. Quasi-random rumor spreading: Reducing randomness can be costly. *Information Processing Letters* 111, 5, 227–230.

B. Doerr, M. Fouz, and T. Friedrich. 2011a. Social networks spread rumors in sublogarithmic time. In *Proceedings of the 43rd ACM Symposium on Theory of Computing (STOC)*. 21–30.

B. Doerr and T. Friedrich. 2009. Deterministic random walks on the two-dimensional grid. *Combinatorics, Probability and Computing* 18, 123–144.

B. Doerr, T. Friedrich, M. Künnemann, and T. Sauerwald. 2011b. Quasirandom rumor spreading: An experimental analysis. *Journal of Experimental Algorithmics* 16, Article 3.3.

B. Doerr, T. Friedrich, and T. Sauerwald. 2008. Quasirandom rumor spreading. In *Proceedings of the 19th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 773–781.

B. Doerr, T. Friedrich, and T. Sauerwald. 2009. Quasirandom rumor spreading: Expanders, push vs. pull, and robustness. In *Proceedings of the 36th International Colloquium on Automata, Languages and Programming (ICALP)*. 366–377.

B. Doerr, A. Huber, and A. Levavi. 2013. Strong robustness of randomized rumor spreading protocols. *Discrete Applied Mathematics* 161, 6, 778–793.

D. Dubhashi and A. Panconesi. 2009. *Concentration of Measure for the Analysis of Randomized Algorithms*. Cambridge University Press.

I. Dumitriu, P. Tetali, and P. Winkler. 2003. On playing golf with two balls. *SIAM Journal on Discrete Mathematics* 16, 4, 604–615.

R. Elsässer and T. Sauerwald. 2009. On the runtime and robustness of randomized broadcasting. *Theoretical Computer Science* 410, 36, 3414–3427.

U. Feige, D. Peleg, P. Raghavan, and E. Upfal. 1990. Randomized broadcast in networks. *Random Structures and Algorithms* 1, 4, 447–460.

N. Fountoulakis and A. Huber. 2009. Quasirandom rumour spreading on the complete graph is as fast as randomized rumour spreading. *SIAM Journal on Discrete Mathematics* 23, 4, 1964–1991.

N. Fountoulakis, A. Huber, and K. Panagiotou. 2010. Reliable broadcasting in random networks and the effect of density. In *Proceedings of the 29th IEEE International Conference on Computer Communications (INFOCOM)*. 2552–2560.

N. Fountoulakis and K. Panagiotou. 2010. Rumor spreading on random regular graphs and expanders. In *Proceedings of the 14th International Workshop on Randomization and Computation (RANDOM)*. 560–573.

N. Fountoulakis, K. Panagiotou, and T. Sauerwald. 2012. Ultra-fast rumor spreading in social networks. In *Proceedings of the 23rd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 1642–1660.

T. Friedrich, M. Gairing, and T. Sauerwald. 2010. Quasirandom load balancing. In *Proceedings of the 21st Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 1620–1629.

T. Friedrich and T. Sauerwald. 2010. The cover time of deterministic random walks. *Electronic Journal of Combinatorics* 17, 1, 1–7. R167.

A. M. Frieze and G. R. Grimmett. 1985. The shortest-path problem for graphs with random arc-lengths. *Discrete Applied Mathematics* 10, 57–77.

G. Giakkoupis. 2011. Tight bounds for rumor spreading in graphs of a given conductance. In *Proceedings of the 28th International Symposium on Theoretical Aspects of Computer Science (STACS)*. 57–68.

G. Giakkoupis, T. Sauerwald, H. Sun, and P. Woelfel. 2012. Low randomness rumor spreading via hashing. In *Proceedings of the 29th International Symposium on Theoretical Aspects of Computer Science (STACS)*. 314–325.

G. Giakkoupis and P. Woelfel. 2011. On the randomness requirements of rumor spreading. In *Proceedings of the 22nd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 449–461.

S. M. Hedetniemi, S. T. Hedetniemi, and A. L. Liestman. 1988. A survey of gossiping and broadcasting in communication networks. *Networks* 18, 4, 319–349.

S. Hoory, N. Linial, and A. Wigderson. 2006. Expander graphs and their applications. *Bulletin of the American Mathematical Society* 43, 439–561.

N. Kahale. 1995. Eigenvalue and expansion of regular graphs. *Journal of the ACM* 42, 5, 1091–1106.

R. Karp, C. Schindelhauer, S. Shenker, and B. Vöcking. 2000. Randomized rumor spreading. In *Proceedings of the 41st IEEE Symposium on Foundations of Computer Science (FOCS)*. 565–574.

D. Kempe, A. Dobra, and J. Gehrke. 2003. Gossip-based computation of aggregate information. In *Proceedings of the 44th IEEE Symposium on Foundations of Computer Science (FOCS)*. 482–491.

M. Kleber. 2005. Goldbug variations. *The Mathematical Intelligencer* 27, 55–63.

P. Mahlmann and C. Schindelhauer. 2006. Distributed random digraph transformations for peer-to-peer networks. In *Proceedings of the 18th ACM Symposium on Parallel Algorithms and Architectures (SPAA)*. 308–317.

C. McDiarmid. 1989. On the method of bounded differences. In *Surveys in Combinatorics, 1989 (Norwich, 1989)*. London Mathematical Society Lecture Note Series, vol. 141. Cambridge University Press, Cambridge, 148–188.

M. Mitzenmacher and E. Upfal. 2005. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press.

D. Mosk-Aoyama and D. Shah. 2006. Computing separable functions via gossip. In *Proceedings of the 25th ACM-SIGOPT Principles of Distributed Computing (PODC)*. 113–122.

H. Niederreiter. 1992. *Random Number Generation and Quasi-Monte Carlo Methods*. SIAM, Philadelphia, PA.

B. Pittel. 1987. On spreading a rumor. *SIAM Journal on Applied Mathematics* 47, 1, 213–223.

V. B. Priezzhev, D. Dhar, A. Dhar, and S. Krishnamurthy. 1996. Eulerian walkers as a model of self-organized criticality. *Physical Review Letters* 77, 5079–5082.

Y. Rabani, A. Sinclair, and R. Wanka. 1998. Local divergence of Markov chains and the analysis of iterative load balancing. In *Proceedings of the 39th IEEE Symposium on Foundations of Computer Science (FOCS)*. 694–705.

T. Sauerwald. 2010. On mixing and edge expansion properties in randomized broadcasting. *Algorithmica* 56, 1, 51–88.

R. M. Tanner. 1984. Explicit concentrators from generalized N-gons. *SIAM Journal on Algebraic Discrete Methods* 5, 3, 287–293.

I. A. Wagner, M. Lindenbaum, and A. M. Bruckstein. 1996. Smell as a computational resource—a lesson we can learn from the ant. In *Proceedings of the 4th Israel Symposium on Theory of Computing and Systems (ISTCS)*. 219–230.

I. A. Wagner, M. Lindenbaum, and A. M. Bruckstein. 1999. Distributed covering by ant-robots using evaporating traces. *IEEE Transactions on Robotics and Automation* 15, 5, 918–933.

Received June 2012; revised July 2013; accepted July 2013