Communication Complexity of Quasirandom Rumor Spreading *

Petra Berenbrink Robert Elsässer Thomas Sauerwald

May 23, 2013

Abstract

We consider rumor spreading on random graphs and hypercubes in the quasirandom phone call model. In this model, every node has a list of neighbors whose order is specified by an adversary. In step i every node opens a channel to its ith neighbor (modulo degree) on that list, beginning from a randomly chosen starting position. Then, the channels can be used for bi-directional communication in that step. The goal is to spread a message efficiently to all nodes of the graph.

For random graphs (with sufficiently many edges) we present an address-oblivious algorithm with runtime $O(\log n)$ that uses at most $O(n \log \log n)$ message transmissions. For hypercubes of dimension $\log n$ we present an address-oblivious algorithm with runtime $O(\log n)$ that uses at most $O(n(\log \log n)^2)$ message transmissions.

Together with a result of [9], our results imply that for random graphs the communication complexity of the quasirandom phone call model is significantly smaller than that of the standard phone call model.

1 Introduction

In this paper we consider rumor spreading (a.k.a. randomized broadcasting) in random graphs and hypercubes. This problem is motivated by overlay topologies in peer-to-peer (P2P) systems, in which each node possesses a list of neighboring peers. Our goal is to develop time-efficient rumor spreading algorithms which produce a minimal number of message transmissions and use a small amount of randomness. Since P2P networks are decentralized platforms for sharing data and computing resources, it is very important to provide efficient, simple, and robust rumor spreading algorithms for P2P overlays. Minimization of the number of transmission (communication complexity) is important for applications such as the maintenance of replicated databases in which often huge amounts of broadcasts are necessary to deal with frequent updates in the system.

We consider the *quasirandom phone call model*, a variant of the *standard phone call model*. Let us first introduce the standard phone call model (also known as random phone call model, see [6]). In this model, each node v may perform the following actions in every step: 1) create a new rumor to be spread, 2) establish a communication channel between itself and one randomly chosen neighbor, 3) transmit a message over incident channels (opened by v or by some neighbor of v)

*An extended abstract of this paper appeared in the 18th Annual European Symposium (ESA'10) [3].

and 4) close the channel opened in the current step. Note that open channels can be used for bi-directional communications. *Calling nodes* (i.e., the nodes that opened the channels) can send their messages to their neighbors. These are called `push` transmissions. *Called nodes* can also perform so called pull transmissions, i.e., they send the message to the calling nodes. These transmissions are simply called `pull` transmissions. If there are many rumors to be spread among the nodes, then it is assumed that nodes can combine several rumors to one larger message. Following [17], we therefore focus on the running time and message complexity produced by our algorithms w.r.t. one single message. Nodes can send messages over all their open channels in one time step.

The major challenge for rumor spreading algorithms in the phone call model is to decide whether or not a node should forward the rumor over an open communication channel. An algorithm is called *address-oblivious* (see [17]) if the decision of node v to send a rumor over an open channel (v, w) or (w, v) does not depend on w. However, this decision can depend on the communication partners chosen in earlier rounds or on decisions made so far. Hence, according to such an algorithm a node has to decide whether to use a channel without knowing if the rumor is already known by the neighbor in question. If there are only very few rumors in the network, then many communication channels may be established without ever being used for transmissions. Thus, the phone call model is especially of interest in situations where rumors are frequently generated. Then, the cost of establishing communication channels amortizes over all message transmissions.

In the case of the *quasirandom* phone call model it is assumed that every node has a cyclic list of all its neighbors, numbered from 1 to d, whose order is specified by an adversary. At the beginning, each node v chooses a random position in the list, independently of the other nodes. Assume that $1 \leqslant \ell \leqslant d$ is the random choice of node v, where d is the degree of v. Then v communicates in step i with the neighbor $((i + \ell - 2) \bmod d) + 1$ from the list. To create the list we assume that the adversary has total knowledge about the topology of the network, but cannot foresee any node's random choice such as the position selected at the beginning (cf. [7]).

1.1 Related Work

Due to space constraints, we mention only results which focus on the theoretical study of `push` and `push & pull` algorithms.

Runtime. Most rumor spreading studies analyze the runtime of the `push` algorithm in the standard phone call model for different graph classes. For complete graphs of size n, Pittel [20] shows that (with probability $1 - o(1)$) it is possible to spread a rumor in time $\log_2(n) + \ln(n) + f(n)$, where $f(n)$ is a slowly growing function, improving a result of Frieze and Grimmett [13]. In [12], Feige et al. determine asymptotically optimal upper bounds for the runtime on $G(n, p)$ graphs (i.e., traditional Erdös-Rényi random graphs [11]), bounded degree graphs, and hypercubes, which all hold w.h.p.[1]. Recently, Fountoulakis et al. [14] prove a tighter bound for the runtime on sufficiently dense $G(n, p)$ graphs, similar to the result of [20] for complete graphs. Also recently, Chierichetti et al. [5] show that the runtime of the combined `push & pull` model is $O(\Phi^{-1} \cdot \log n \cdot \log^2(\Phi^{-1}))$ w.h.p. for any graph G, where Φ denotes the conductance of G. This runtime bound has been recently tightened in [15] who proved a runtime bound of $O(\Phi^{-1} \cdot \log n)$ for any graph G with conductance Φ.

[1] W.h.p. or "with high probability" means with probability at least $1 - n^{-c}$ for some constant $c > 0$.

In [7], Doerr et al. analyze the so called quasirandom rumor spreading. They show that for hypercubes and $G(n,p)$ graphs $O(\log n)$ steps suffice to inform every node, w.h.p. These bounds are similar to the ones in the standard phone call model (`push` model). The results of [7] are extended to further graph classes with good expansion properties in [8]. Observe that in [7, 8] the authors mainly focus on the runtime efficiency, and the algorithms therein require $\Theta(n \log n)$ message transmissions for hypercubes and $G(n,p)$ graphs.

Number of Message Transmissions. Karp et al. [17] observe that in complete networks the `pull` approach is inferior to the `push` approach until roughly $n/2$ nodes receive the rumor. Then the `pull` approach becomes superior. They present a `push&pull` algorithm, together with a termination mechanism, which bounds the number of total transmissions to $O(n \log \log n)$ (w.h.p.), and show that this result is asymptotically optimal.

For sparser graphs and the standard phone call model it is not possible to get an oblivious algorithm that uses $O(n \log \log n)$ message transmissions, together with a runtime of $O(\log n)$. In [9], the second author considers random $G(n,p)$ graphs and shows a lower bound of $\Omega(n \log n / \log(pn))$ message transmissions for oblivious rumor spreading algorithms with a runtime of $O(\log n)$. For $p > \log^2 n/n$ he develops an oblivious algorithm that spreads a rumor in time $O(\log n)$ using $O(n \cdot (\log \log n + \log n/\log(pn)))$ transmissions, w.h.p.

In [10] the authors consider a simple modification of the standard phone call model, called RANDOM[4], where every node is allowed to open a channel to *four different* randomly chosen neighbors in every time step. For $G(n,p)$ graphs with $p > \log^2 n/n$, they show that this modification results in a reduction of the number of message transmissions down to $O(n \log \log n)$. Similar results are shown for random d-regular graphs in [1].

The authors of [2] present an extension of RANDOM[4] which they call RR model. In their model each node has a randomly ordered cyclic list with all its neighbors. In step i, the node opens a communication channel to the ith neighbor in its list. The RR model is the same as the quasirandom model except that the adversarial order is replaced by the random order. The authors present an oblivious algorithm for graphs with very good edge and node expansion properties which has a runtime $O(\log n)$ and which uses $O(n\sqrt{\log n})$ message transmissions, w.h.p. The authors establish a lower bound of $\Omega(n\sqrt{\log n/\log d})$ on the number of message transmissions for oblivious rumor spreading algorithms (assuming a runtime of $O(\log n)$), showing that their upper bound is tight up to a $\sqrt{\log \log n}$ factor if d is polylogarithmic in n.

The algorithms of [2, 9, 10, 17] spread the rumor using `push` transmissions until a constant fraction of the nodes receives the rumor (we call these nodes *informed* in the following). Then the algorithms spread the rumor via `pull` transmissions until every node is informed. To save on communications, the algorithms of [1, 2, 9, 10] only allow each node v a certain number of transmissions which depends on the age the rumor had at the time v received it for the first time.

1.2 Model

In this paper we consider random graphs $G(n,p) = (V,E)$ and hypercubes H_d of dimension d. A random graph $G(n,p)$ consists of n nodes. The probability that any pair of nodes is connected is p, we assume that $(\log^2 n)/n \leqslant p \leqslant 2^{o(\sqrt{\log n})}/n$. The expected number of edges for $G(n,p)$ is $pn \cdot (n-1)/2$. Let $d(v)$ be the degree of node v and $N(v)$ be the set of neighbors in V. For $S \subset V$, let $N(S)$ be the set of neighbors of nodes in S. Let α be the node expansion value of $G(n,p)$. Then

$\alpha = \min_{S \in V, |S| \leqslant n/4} N(S)/|S|$. It is known that for our choice of p, α is a constant close to 1 w.h.p. ([4]).

The d-dimensional hypercube H_d consists of $n = 2^d$ many nodes. A binary string of length 2 is assigned to every node and two nodes are connected if their binary strings differ in *exactly* one bit. Hence, the degree of any node of H_d is d. Note that hypercubes have much smaller expansion than random graphs.

We assume that every node has an estimation of n which is accurate to within a constant factor. We also assume that all nodes have access to a global clock, and that they work synchronously. As communication model we assume a variant of the phone call model. In the standard phone call model (see [6]) in each step t every node can create an arbitrary amount of rumors to be spread. To measure the communication cost we only count the number of message transmissions, i.e., opening a channel is not counted. Following [1, 2, 9, 17], we assume here that new pieces of information are generated frequently in the network, and then the cost of establishing communication channels amortizes over all message transmissions. However, we only concentrate on the distribution and lifetime of a single rumor.

The quasirandom variant of the phone call model considered in this paper was introduced in [7]. In the quasirandom phone call model every node v has a list $\tilde{L}_v = \tilde{L}_v[0], \tilde{L}_v[1], \ldots \tilde{L}_v[d(v)-1]$ of length $d(v)$ with all its neighbors. The order of that list is arbitrary, i.e., it may be determined by an adversary. We assume that the rumor is initiated on a node at step 0. For spreading the rumor, every node v chooses a random position i_v in the list, independently of the other nodes. For its j-th communication v will open a channel to node $L_v[(i_v + j - 1) \bmod d(v)]$. We define $L_v = L_v[0], L_v[1], \ldots, L_v[d(v)]$ as the list beginning at neighbor i_v.

Nodes that received the rumor will be called *informed*. By I_t (H_t) we denote the set of informed (uninformed) nodes in step t. Furthermore, let I_t^+ be the set of nodes that reveive the rumor *for the first time* in step t. These nodes will also be called *newly informed* nodes.

Note that we omit rounding of non-integers in our proofs but not in the statement of our algorithms. Furthermore, we assume that $\log n$ is the logarithm with base 2.

1.3 Our Contribution

In this paper we show the following results. For random graphs with $(\log^2 n)/n \leqslant p \leqslant 2^{o(\sqrt{\log n})}/n$ we present an oblivious algorithm (in the quasirandom model) that spreads a rumor in time $O(\log n)$ using $O(n \log\log n)$ message transmissions, w.h.p. Compared to [7], we reduce the number of message transmissions by a factor of $\log n/\log\log n$. Moreover, our upper bound in the quasirandom model is significantly smaller than the lower bound for the standard phone call model (cf. [9]).

For the hypercube we show a result that is slightly weaker than our result for random graphs. We present an oblivious algorithm (which is similar to the algorithm for random graphs) that spreads a rumor in time $O(\log n)$ using $O(n \cdot (\log\log n)^2)$ message transmissions, w.h.p. The communication complexity on the hypercube has not been analyzed before, neither in the standard nor in the quasirandom phone call model. Therefore the best known algorithms require $O(\log n)$ time, but produce $\Omega(n \log n)$ message transmissions. In comparison to that, we reduce the number of message transmissions by a factor of $\log n/(\log\log n)^2$.

Our results demonstrate that on two important networks rumor spreading can be done much more efficiently in the quasirandom phone model than in the standard phone call model. Moreover, the results provide evidence that avoiding previously chosen communication partners is more

important than choosing all communication partners independently and uniformly at random.

Note that the conference version of this paper ([3]) contained a lower bound on the message complexity for hypercubes. Unfortunately there is a mistake in the proof on the slower bound which we were not able to repair.

2 Random Graphs

In this section we present an algorithm with runtime $\mathcal{O}(\log n)$ and communication complexity $\mathcal{O}(n \log\log n)$ for random graphs.

2.1 Our Algorithm

We assume that the rumor we want to spread is generated at time 0, i.e., at time t the age of the rumor equals t. The algorithm describes the behavior of the nodes w.r.t. one specific rumor. Depending on the age of the rumor, each node is in one of the following phases (in our algorithm, ρ is a sufficiently large constant):

Phase 0: $[\mathbf{age} \leqslant \lceil \rho \log n \rceil]$ The node which generates the rumor performs `push` in each step of this phase. No other node transmits the rumor in this phase.

Phase 1: $[\lceil \rho \log n \rceil + 1 \leqslant \mathbf{age} \leqslant 2 \cdot \lceil \rho \log n \rceil + 320]$ Nodes that received the rumor in Phase 0 use the first 320 steps of this phase to perform `push` in each of these steps. If a node receives the rumor for the *first* time in some step $t \in \{\lceil \rho \log n \rceil + 1, \ldots, 2 \cdot \lceil \rho \log n \rceil\}$, then the node performs `push` in the steps $t+1, \ldots, t+320$.

Phase 2: $[2\lceil \rho \log n \rceil + 321 \leqslant \mathbf{age} \leqslant 2 \cdot \lceil \rho \log n + \rho \log\log n \rceil]$ Every informed node performs `push` in every step of this phase.

Phase 3: $[2\lceil \rho \log n + \rho \log\log n \rceil + 1 \leqslant \mathbf{age} \leqslant 3 \cdot \lceil \rho \log n \rceil]$ Every node that becomes informed in some step of this phase performs `pull` transmissions for the rest of the phase, i.e., after receiving the message it transmits over all incoming channels in every step of the phase. All other informed nodes perform `pull` over all incoming channels with probability $1/\log n$.

Phase 4: $[3\lceil \rho \log n \rceil + 1 \leqslant \mathbf{age} \leqslant 3 \cdot \lceil \rho \log n + \rho \log\log n \rceil]$ All informed nodes perform `pull` transmissions.

It is easy to see that at the end of Phase 0, exactly $\rho \log n + 1$ nodes are informed (Observation 2.2). In Phase 1 we inform half of the nodes (see Lemma 2.3). At the end of Phase 2 we have $n \cdot (1 - 2\log\log n / \log n)$ informed nodes, w.h.p. (Lemma 2.5). Phase 3 and Phase 4 are analyzed in Lemma 2.6. There we show that w.h.p. at the end of Phase 4 all nodes are informed.

2.2 Analysis of the Algorithm

For a graph $G(n,p)$ and our choice of p the degree of each node is in the range $[np\cdot(1-1/\log n), np\cdot(1+1/\log n)]$, with probability at least $1-n^{-3}$. In the following, we condition on this event, and for simplicity we assume in our analysis that $d = pn$.

Theorem 2.1. *Consider $G = G(n,p)$ with $(\log^2 n)/n \leqslant p \leqslant 2^{o(\sqrt{\log n})}/n$. Then, the algorithm above spreads a rumor in G in time $O(\log n)$ using $O(n \log\log n)$ message transmissions, w.h.p.*

In the rest of this section we will prove the above theorem. The proof is split into several lemmas. It is easy to see that in Phase 0 the node that generated the rumor informs $\rho \log n$ different neighbors, which results in the following observation.

Observation 2.2. *At the end of Phase 0 there are $\rho \log n$ informed nodes.*

Now we concentrate on Phase 1 and show the following lemma.

Lemma 2.3. *With probability $1-n^{-2}$, at least $n/2$ nodes are informed at the end of Phase 1.*

Proof. Assuming that the nodes all have degree d we show that

1. After the first $\rho\cdot(\log n)/2$ steps at least $6n/d$ nodes are informed, where $\rho > 8$.
2. After $\rho\cdot((\log n)/2 - 320)$ additional steps we have at least $n/40$ informed nodes.
3. After the last $320\cdot\rho$ steps we have $n/2$ informed nodes for ρ large enough.

Part 1). The statement follows from the claim below we adapted from Claim A.1 of [2] with expansion factor $\alpha > 0.6$. This (node-)expansion holds for the random graphs considered here, if the size of the set is bounded by n/d [4].

Claim 2.4. *Let $\tau_1, \tau_2, \ldots$ be consecutive time intervals in this phase, each of them consisting of 160 time steps. Furthermore, let t_i be the beginning of interval τ_i, and let $I^+_{\tau_i}$ be the set of nodes, which are informed in time interval τ_i for the first time. Assume that*

$$|I_{t_i}| \leqslant \frac{n}{d} \text{ and } |I_{t_i}| \geqslant \frac{8}{\alpha^2}\cdot|I_{t_{i-1}}|$$

Then, with probability at least $1-n^{-3}$ we have

$$|I^+_{\tau_i}| \geqslant \frac{8}{\alpha^2}\cdot|I_{t_i}|.$$

In the algorithm considered in [2], we assume that each node transmits the message for $80/\alpha^2$ steps; in our algorithm we assume that the nodes transmit for 320 time steps, and therefore the claim applies here too. In order to have at $6n/d$ informed nodes at the end of Phase 1, we apply the claim to a time interval which starts right after $n/d-1$ nodes are informed.

Part 2). In this case the number of informed nodes lies in the range $[6n/d, n/40]$. We show inductively that with a very high probability the number of informed nodes grows by a factor of 2.1 every 160 steps. To do so we divide the time into $\ell = (\rho \cdot ((\log n)/2 - 320)/160$ subphases. For $0 \leqslant i \leqslant \ell$, subphase τ_i starts in step $\rho \cdot (\log n)/2 + 160i + 1$ and ends in step $\rho \cdot (\log n)/2 + 160(i+1)$. Let $I^+_{\tau_i}$ be the newly informed nodes in Subphase τ_i, and I_{τ_i} are the informed nodes at the *beginning* of Subphase τ_i. Note that all nodes in $I^+_{\tau_i}$ perform a `push` transmissions in Subphase τ_{i+1}.

We show by induction that for $0 \leqslant i \leqslant \ell$ we have $|I^+_{\tau_i}| \geqslant 2.1 \cdot |I_{\tau_i}|$, which then implies that $|I^+_{\tau_i}| \geqslant |I_{\tau_{i+1}}|/2$.

Fix a subphase τ_{i+1}. From Lemma 4.5 (Part 1) we know that there are $n/6$ uninformed nodes at the beginning of the subphase so that, with probability at least $1 - \varepsilon^n$, all of these nodes have at least $|I^+_{\tau_i}| \cdot d/(2n)$ neighbors in the set of nodes $I^+_{\tau_i}$. Hence, such an uninformed node remains uninformed in the time interval τ_{i+1} with probability at most $(1 - 160/d)^{|I^+_{\tau_i}| \cdot d/(2n)}$. This holds since the first positions are chosen independently and uniformly at random, and a neighbor misses a specific node in 160 steps with probability $1 - 160/d$. Thus,

$$\begin{aligned}
\mathbf{E}\left[\,|I^+_{\tau_{i+1}}|\,\right] &\geqslant \left(1 - \left(1 - \frac{160}{d}\right)^{|I^+_{\tau_i}| \cdot d/(2n)}\right) \cdot \frac{n}{6} \\
&\geqslant \left(1 - \left(\frac{1}{e}\right)^{80|I^+_{\tau_i}|/n}\right) \cdot \frac{n}{6} \geqslant \left(1 - \left(\frac{1}{e}\right)^{40 \cdot |I_{\tau_{i+1}}|/n}\right) \cdot \frac{n}{6} \\
&\geqslant \left(1 - \left(1 - \frac{1}{n/(40 \cdot |I_{\tau_{i+1}}|) + 1}\right)\right) \cdot \frac{n}{6} > 2.2 \cdot |I_{\tau_{i+1}}|
\end{aligned}$$

Here, the third equation uses the induction hyphothesis. Now, we can construct a Martingale sequence Y_0, Y_1, …, $Y_{n/6}$ on these $n/6$ uninformed nodes, where Y_j is the expected value on the number of newly informed nodes in time interval τ_{i+1} after the first j uniformed nodes are exposed. Since this Martingale sequence satisfies the 160-Lipschitz condition, applying the Azuma-Hoeffding bound (see e.g. [18])

$$\mathbf{Pr}\left[\,|Y_{n/6} - Y_0| \geqslant 0.1 \cdot \mathbf{E}\left[\,|I^+_{\tau_{i+1}}|\,\right]\,\right] \leqslant 2\exp\left(\frac{0.01 \cdot \left(\mathbf{E}\left[\,|I^+_{\tau_{i+1}}|\,\right]\right)^2}{(160)^2 \cdot 2n/6}\right) \tag{1}$$

we obtain with probability $1 - o(n^{-3})$ that $|I^+_{\tau_{i+1}}| \geqslant 2.1 \cdot |I_{\tau_{i+1}}|$.

Part 3). Now the number of informed nodes lies in the range $[n/40, n/2]$. We divide the time into $\ell = 2\rho$ subphases. For $0 \leqslant i \leqslant \ell$, subphase τ_i starts in step $\rho \cdot (\log n - 320) + 160i + 1$ and ends in step $\rho \cdot (\log n - 320) + 160(i+1)$. Our goal is to show inductively that for all but the last phase $|I^+_{\tau_i}| \geqslant 2.1 \cdot |I_{\tau_i}|$. In the last phase we inform enough nodes so that half of the nodes are informed at the end of this phase.

Similar to Part 2) we fix a subphase τ_{i+1} and define $H_{\tau_{i+1}}$ as the number of uninformed nodes at the *beginning* of Subphase τ_{i+1}. From Lemma 4.5 (Part 2) it follows that $|H_{\tau_{i+1}}|/2$ of the uninformed nodes have at least $|I^+_{\tau_i}|d/(2n)$ neighbors in the set of nodes $I^+_{\tau_i}$, with probability $1 - \varepsilon^n$. Such an

uninformed node remains uninformed in τ_{i+1} with probability at most $(1-160/d)^{|I^+_{\tau_i}|d/(2n)}$. Thus,

$$\begin{aligned}\mathbf{E}\left[\,|I^+_{\tau_{i+1}}|\,\right] &\geqslant \left(1-\left(1-\frac{160}{d}\right)^{|I^+_{\tau_i}|d/(2n)}\right)\cdot\frac{|H_{\tau_{i+1}}|}{2}\\ &\geqslant \left(1-\left(\frac{1}{e}\right)^{80|I^+_{\tau_i}|/n}\right)\cdot\frac{|H_{\tau_{i+1}}|}{2}\geqslant\left(1-\left(\frac{1}{e}\right)^{40|I_{\tau_{i+1}}|/n}\right)\cdot\frac{|H_{\tau_{i+1}}|}{2}.\end{aligned}$$

The remainder of the proof is a case analysis depending on $|I_{\tau_{i+1}}|$. If $n/40\leqslant|I_{\tau_{i+1}}|\leqslant n/10$, then

$$\left(1-\left(\frac{1}{e}\right)^{40|I_{\tau_{i+1}}|/n}\right)\cdot\frac{|H_{\tau_{i+1}}|}{2}\geqslant\left(1-\left(\frac{1}{e}\right)\right)\cdot\frac{9n}{20}\geqslant\frac{2.2\cdot n}{10}.$$

Using the method of bounded independent differences [18] as in inequality (1) (with the adapted length of the Martingale to $|H_{\tau_{i+1}}|/2$) one can show that with probability $1-o(n^{-3})$ we obtain $|I^+_{\tau_{i+1}}|\geqslant 2.1\cdot|I_{\tau_{i+1}}|$. For $n/10<|I_{\tau_{i+1}}|\leqslant n/6$

$$\left(1-\left(\frac{1}{e}\right)^{40|I_{\tau_{i+1}}|/n}\right)\cdot\frac{|H_{\tau_{i+1}}|}{2}\geqslant\left(1-\left(\frac{1}{e}\right)^{4}\right)\cdot\frac{5n}{12}\geqslant\frac{2.2\cdot n}{6}.$$

Then, with probability $1-o(n^{-3})$ we have $|I^+_{\tau_{i+1}}|\geqslant 2.1\cdot|I_{\tau_{i+1}}|$ [18].

For $|I_{\tau_{i+1}}|\geqslant n/6$ we get

$$\begin{aligned}&|I_{\tau_{i+1}}|+\left(1-\left(\frac{1}{e}\right)^{40|I_{\tau_{i+1}}|/n}\right)\cdot\frac{|H_{\tau_{i+1}}|}{2}\\ &\geqslant|I_{\tau_{i+1}}|+\left(1-\left(\frac{1}{e}\right)^{40/6}\right)\cdot\left(\frac{n-|I_{\tau_{i+1}}|}{2}\right)\geqslant\frac{41n}{80}.\end{aligned}\tag{2}$$

Again, we obtain with probability $1-o(n^{-3})$ that $|I_{\tau_{i+2}}|>n/2$. □

Lemma 2.5. *Assume $\rho\geqslant 30$. With probability at least $1-n^{-2}$, there are at most $(n\cdot 2\log\log n/\log n)$ uninformed nodes at the end of Phase 2.*

Proof. Note that in this phase every informed node performs a `push` transmission in every step. Let T be a random variable defined as the first time step directly after the subphase τ of the previous phase, for which $|I_\tau\cup I^+_\tau|\geqslant n/2$ for the first time (this happens w.h.p. in Phase 1). For the sake of this proof we assume that only the nodes of I^+_τ perform `push` transmissions in this phase. Due to the second term in the left hand side of inequality (2), we have $|I^+_\tau|>n/5$ with probability at least $1-n^{-3}$.

According to Lemma 4.6, with probability $1-\varepsilon^n$ ($\varepsilon>0$ is a constant) there are at most $n\cdot\log\log n/\log n$ nodes in H_T which have fewer than $d/10$ neighbors in I^+_τ. After $\rho\log\log n$ additional steps each of the other (uninformed) nodes remains uninformed with probability at least

$$\left(1-\frac{\rho\log\log n}{d}\right)^{d/10}\leqslant e^{-\rho\log\log n/10}\leqslant\log^{-3}n,$$

for $\rho \geqslant 30$. Thus, if there are at most $n \cdot \log\log n/\log n$ nodes in H_T which have fewer than $d/10$ neighbors in I_τ^+, the expected number of newly informed nodes in Phase 2 is at least

$$\left(|H_T| - \frac{n\log\log n}{\log n}\right) \cdot (1 - \log^{-3} n).$$

Then, using [18] one can show that with probability at least $1-n^{-2}$, the number of newly informed nodes in this phase is at least

$$\left(|H_T| - \frac{2n\log\log n}{\log n}\right).$$

Hence with probability at least $1-n^{-2}$, the number of uninformed nodes after this phase is at most $n \cdot 2\log\log n/\log n$. □

Finally, we concentrate on Phases 3 and 4.

Lemma 2.6. *Assume $\rho \geqslant 30$. With probability at least $1-n^{-2}$ all nodes are informed at the end of Phase 4.*

Proof. For a node u and time interval $\tau = [t, t']$, let $L_u(\tau)$ be the set of nodes chosen by u in steps $\tau = t, t+1, \ldots, t'$. Define $t_2 = 3\rho \cdot (\log n + \log\log n)$ as the end of Phase 4, $t_1 = 3 \cdot \rho \log n$ as the beginning of Phase 4, and $t_0 = 2\rho(\log n + \log\log n)$ as the beginning of Phase 3.

First we consider Phase 4 and divide the time interval $[t_1+1, t_2]$ into $k' = (t_2 - t_1)/320$ subintervals of length 320. For any $0 \leqslant i \leqslant k'-1$ we define

$$\tilde{\tau}_i = [t_2 - 320i, t_2 - 320 \cdot (i+1) + 1].$$

For a node v, let

$$U_0(v) = L_v(\tilde{\tau}_0) \text{ and } U_i(v) = \cup_{w \in U_{i-1}} L_w(\tilde{\tau}_i).$$

We can visualize $\cup_{i \leqslant k'-1} U_i(v)$ as tree of depth $k'-1$ rooted in v (cf. Lemma 4.7). The level i nodes are the nodes in $U_i(v)$. Then, according to Lemma 4.7 $|U_{k'-1}(v)| = \Omega(\log^3 n)$ with probability $1 - o(n^{-3})$.

In the following we consider two cases. In the first case, we assume that $\cup_{i \leqslant k'-1} U_i(v) \cap I_{t_1} \neq \emptyset$ for some node v. Then v is informed in Phase 4 since all informed nodes perform `pull` transmissions in that phase. In the second case, let $U_{k'-1}(v) \cap I_{t_1} = \emptyset$. For this case we know from Lemma 4.7 that in Phase 4 v will be the root of a communication tree consisting of nodes which are still all uninformed in step t_1. Then we will show that w.h.p. at least one of the leaves of the tree will be informed in Phase 3. The rumor will be propagated to v via the path between v and the informed leaf.

Now we need some additional definitions. We divide the time interval $[t_0+1, t_1]$ into $k'' = (t_1 - t_0)/160$ rounds of length 160. For any $0 \leqslant i \leqslant k''-1$

$$\tilde{\tau}'_i = [t_1 - 160i, t_1 - 160 \cdot (i+1) + 1].$$

For $0 \leqslant i \leqslant \rho \log n$, let

$$\begin{aligned}
\tilde{U}^H_{-1}(v) &= U_{k'-1}(v) \\
\tilde{U}^H_i(v) &= \cup_{w \in \tilde{U}^H_{i-1}(v)} L_w(\tilde{\tau}'_i) \cap H_{t_0} \\
\tilde{U}^I_i(v) &= \cup_{w \in \tilde{U}^H_{i-1}(v)} L_w(\tilde{\tau}'_i) \cap I_{t_0}.
\end{aligned}$$

A node $\tilde{w}_i \in \tilde{U}_i^I(v)$ is connected to a node $\tilde{w}_{-1} \in \tilde{U}_{-1}^H(v)$ by a path $P = (\tilde{w}_i, \ldots, \tilde{w}_0, \tilde{w}_{-1})$, where $\tilde{w}_{i-1}, \ldots, \tilde{w}_0, \tilde{w}_{-1} \in H_{t_0}$, and for $j = -1, \ldots, i-1$ we have $\tilde{w}_{j+1} \in L_{\tilde{w}_j}(\tilde{\tau}'_{j+1})$. Now define

$$\tilde{U}_{0\to i}^H(v) = \cup_{j=0}^i \tilde{U}_j^H(v) \qquad \text{and} \qquad \tilde{U}_{0\to i}^I(v) = \cup_{j=0}^i \tilde{U}_j^I(v).$$

Since $|\tilde{U}_{-1}^H(v)| = \Omega(\log^3 n)$, we can apply the same techniques as in Lemma 2.3 and obtain that

$$|\tilde{U}_i^H(v) \cup \tilde{U}_i^I(v)| \leqslant 2.1 \cdot |\tilde{U}_{i-1}^H(v)|$$

for any $i \geqslant 1$ as long as $|\tilde{U}_{i-1}^I(v)| = O(\log^2 n)$ and $|\tilde{U}_i^H(v)| < n/40$. However, since $|H_{t_0}| \leqslant 2n \log\log n / \log n$, there exists some $i < k''$ such that $|\tilde{U}_{0\to i}^I(v)| > \rho \log^2 n$. Then, we can argue that every node $u \in \tilde{U}_{0\to i}^I(v)$ performs pull transmissions with probability $1/\log n$. Since for every u there is a $s < k''$ such that a path $(u, \tilde{w}_s, \ldots, \tilde{w}_0, \ldots, v)$ exists, that consists of nodes of H_{t_0} that perform pull transmissions in the corresponding rounds. Hence, there is a node $u \in L_{\tilde{w}_s}(\tilde{\tau'}_{s+1})$ which transmits the rumor at the right time, with probability at least

$$1 - \left(1 - \frac{1}{\log n}\right)^{\rho \log^2 n} = 1 - o(n^{-3}),$$

if ρ is large enough. □

Let us now prove Theorem 2.1. The correctness (every node gets informed w.h.p.) follows from the lemmas above. It remains to analyze the total number of message transmissions. In Phase 0, the algorithm uses $\mathcal{O}(\log n)$ message transmissions. In Phases 1, 2 and 4, the algorithm uses $\mathcal{O}(n \log\log n)$ message transmissions. By Lemma 2.5, we know that after Phase 2 at most $\mathcal{O}(n \log\log n/(\log n))$ uninformed nodes remain. These nodes generate at most $\mathcal{O}(n \log\log n)$ message transmissions in Phase 3. Using a Chernoff bound, we can show that the nodes that are informed at the end of Phase 2 use at most $\mathcal{O}(n)$ message transmissions. Hence the total number of message transmissions is $\mathcal{O}(n \log\log n)$. □

3 Hypercubes

In this section we present an algorithm with runtime $\mathcal{O}(\log n)$ and communication complexity $\mathcal{O}(n(\log\log n)^2$ for hypercubes.

3.1 Our Algorithm

In the algorithm below, the total number of message transmissions is $\mathcal{O}(n(\log\log n)^2)$, which can be shown as in the proof of Theorem 2.1 above ($\rho > 0$ is a sufficiently large constant).

Phase 1: $[1 \leqslant$ **age** $\leqslant \lceil \rho \log n \rceil]$ If a node receives a rumor for the *first* time in step $t \in \{1, \ldots, \lceil \rho \log n \rceil\}$, then the node performs `push` for the next $C \log\log n$ consecutive steps, where $C > 0$ is a sufficiently large constant.

Phase 2: $[\lceil \rho \log n \rceil + 1 \leqslant$ **age** $\leqslant 2 \cdot \lceil \rho \log n \rceil]$ Every node which becomes informed in this phase performs `pull` over each incoming channel. All other informed nodes perform `pull` with probability $1/\log n$ over each incoming channel.

Phase 3: $[2\lceil \rho \log n \rceil + 1 \leqslant \textbf{age} \leqslant 2 \cdot \lceil \rho \log n + \rho (\log \log n)^2 \rceil]$ All informed nodes perform `pull` transmissions in every step of this phase.

3.2 Analysis of the Algorithm

Theorem 3.1. *Assume that H_d is a hypercube of dimension $d = \log n$. The algorithm above spreads a rumor in H_d in time $O(\log n)$ using $O(n(\log \log n)^2)$ message transmissions, w.h.p.*

The analysis of the above theorem is similar to the one for random graphs. However, the lack of strong expansion properties makes it more difficult and one has to resort to the special structure and the symmetries of hypercubes.

For any integer $0 \leqslant k \leqslant d$, let $N_k(0)$ be the set of nodes with distance k to the node 0^d, i.e. the set of nodes with k ones. Sometimes we also simply write N_k for $N_k(0)$ if the reference to 0^d is clear from the context. In addition, we define $N_{\geqslant k} := \cup_{j=k}^{d} N_j$. For any node v and any subset $S \subseteq V$, we denote by $d_S(v)$ the number of neighbors of v within the set S.

3.2.1 Analysis of Phase 1

For the analysis we first subdivide this phase into two intervals $[1, t_1), [t_1, t_2)$, where $t_1 := (C^2/2)(\log \log n)^2$, $t_2 := t_1 + K \cdot (\log n/2 - (C/2) \log \log n)$, where $K > 0$ is a sufficiently large constant. We further define $\ell_1 := (C/2) \log \log n$.

Lemma 3.2. *For all possible lists $\cup_{v \in V} L_v$ we have*

$$|I_{t_1}^+ \cap N_{\ell_1}| \geqslant (\log n)^{C/2}.$$

Proof. To prove this lemma we consider a delayed version of our algorithm. We assume that a node in level N_i does not transmit before timestep $iC \log \log n$. Observe that any node in N_i has exactly $d - i$ neighbors in N_{i+1} and i neighbors in N_{i-1}. Let $\ell := (C/2) \log \log n$. Any informed node in N_i informs at least $C \log \log n - i$ nodes in N_{i+1} and any node in N_{i+1} is informed by at most $i + 1$ nodes in N_i. Hence we obtain that the number of informed nodes in N_ℓ within $\ell \cdot C \log \log n$ rounds is at least

$$\prod_{i=0}^{\ell-1} \left(\frac{C \log \log n - i}{i+1} \right) \geqslant \frac{((C/2) \log \log n)^{(C/2) \log \log n}}{((C/2) \log \log n)!} \geqslant 2^{(C/2) \log \log n} = (\log n)^{C/2},$$

where we have used the fact that $n! \leqslant (n/2)^n$ for every integer n.

□

Lemma 3.3. *Let $\varepsilon > 0$ be a constant. Assume that an adversary (who knows all random choices of the protocol) is allowed to choose a set of nodes of size $\log^3 n$ so that these nodes never transmit a rumor within the time-interval $[t_1, t_2]$. Then there exists a constant $\delta = \delta(\varepsilon) > 0$ and $\ell_2 := (1/2 - \delta) \cdot d$, so that with probability $1 - n^{-\omega(1)}$,*

$$|I_{t_2}^+ \cap N_{\ell_2}| \geqslant 2^{(1-\varepsilon) \cdot d},$$

and $I_{t_2} \cap N_{>\ell_2} = \emptyset$.

Proof. The proof is similar to the one of [7, Theorem 2]. From Lemma 3.2 we get

$$|I_{t_1}^+ \cap N_{\ell_1}| \geqslant (\log n)^{C/2}.$$

In the following we consider a slowed-down version of our algorithm. The algorithm works in phases of K steps each, where K is a sufficiently large constant to be determined later. In phase i, only nodes communicate that are in level $N_{i+\ell_1}$. Fix an arbitrary phase i with $0 \leqslant i \leqslant \log n - t_1$ and a timestep $t = t_1 + K \cdot i$ (first round of phase i). Let us denote by $\tilde{I}_{t+K}$ the nodes that would get informed if there was no adversary. Consider N_j with $j = i + \ell_1$ and $1 \leqslant j \leqslant \ell_2 \leqslant d/2$. Our goal is to show that a large subset of the nodes in N_{j+1} will be informed in phase i. The probability that a node $v \in N_{j+1}$ is still uninformed at the end of phase i is

$$\mathbf{Pr}\left[\, v \notin \tilde{I}_{t+K} \,\right] \leqslant \prod_{u \in \Gamma(v) \cap I_t \cap N_j} \left(1 - \frac{K}{d}\right) = \left(1 - \frac{K}{d}\right)^{d_{\tilde{I}_t \cap N_j}(v)}.$$

By linearity of expectation we get

$$\begin{aligned}
\mathbf{E}\left[\, |\tilde{I}_{t+K} \cap N_{j+1}| \,\right] &= \sum_{v \in N_{j+1}} \mathbf{Pr}\left[\, v \in \tilde{I}_{t+K} \,\right] \\
&\geqslant \sum_{v \in N_{j+1}} 1 - \left(1 - \frac{K}{d}\right)^{d_{I_t \cap N_j}(v)} \\
&\geqslant \sum_{v \in N_{j+1} \cap N(I_t)} 1 - \exp\left(-\frac{K}{d} \cdot d_{I_t \cap N_j}(v)\right) \\
&= |N_{j+1} \cap N(I_t)| - \sum_{v \in N_{j+1} \cap N(I_t)} \exp\left(-\frac{K}{d} \cdot d_{I_t \cap N_j}(v)\right).
\end{aligned}$$

Applying Lemma 4.3 with $|E(I_t \cap N_j, N_{j+1})| = \sum_{v \in N_{j+1}} d_{I_t \cap N_j}(v) = |I_t \cap N_j| \cdot (d - j)$ and for any $v \in N_{j+1}$, $d_{I_t \cap N_j}(v) \in [0, j+1]$, we get

$$\begin{aligned}
\mathbf{E}\left[\, |\tilde{I}_{t+K} \cap N_{j+1}| \,\right] &\geqslant |N_{j+1} \cap N(I_t)| - \frac{|E(I_t \cap N_j, N_{j+1})|}{j+1} \cdot \exp\left(-\frac{K}{d} \cdot (j+1)\right) - \\
&\qquad \left(|N_{j+1} \cap N(I_t)| - \frac{|E(I_t \cap N_j, N_{j+1})|}{j+1}\right) \cdot 1 \\
&= \frac{|E(I_t \cap N_j, N_{j+1})|}{j+1} \cdot \left(1 - \exp\left(-\frac{K}{d} \cdot (j+1)\right)\right) \\
&\geqslant |I_t \cap N_j| \cdot \frac{d-j}{j+1} \cdot \left(1 - \frac{1}{\frac{K(j+1)}{d} + 1}\right),
\end{aligned}$$

where we have used in the last inequality the fact that $\exp(-x) \leqslant \frac{1}{x+1}$ for any $x \in \mathbb{R}$.

Let the nodes in $I_t \cap N_j$ be $u_1, u_2, \ldots, u_{|I_t \cap N_j|}$. Consider the random variable $|\tilde{I}_{t+K} \cap N_{j+1}|$ as a function of $i_{u_1}, i_{u_2}, \ldots, i_{u_{|I_t \cap N_j|}}$, where i_{u_j} is the randomly chosen starting point of the list of node

u_j. Since the i_{u_j} are independent random variables and changing one i_{u_j} can change $\tilde{I}_{t+K} \cap N_{j+1}$ by at most K, Lemma 4.8 gives

$$\begin{aligned}
&\mathbf{Pr}\left[\,|\tilde{I}_{t+K} \cap N_{j+1}| \leqslant \left(1 - \frac{1}{\log n}\right) \cdot \mathbf{E}\left[\,|\tilde{I}_{t+K} \cap N_{j+1}|\,\right]\right] \\
&\leqslant \exp\left(-\frac{\left(\frac{1}{\log n} \cdot \mathbf{E}\left[\,|I_{t+K} \cap N_{j+1}|\,\right]\right)^2}{K \cdot |I_t \cap N_j|}\right) \\
&\leqslant \exp\left(-\frac{\frac{1}{\log^2 n}\left(|I_t \cap N_j| \cdot \left(1 - \frac{1}{\frac{K(j+1)}{d}+1}\right)\right)^2}{K \cdot |I_t \cap N_j|}\right) \\
&\leqslant \exp\left(-\Omega\left(\frac{1}{\log^4 n}\,|I_t \cap N_j|\right)\right) = \exp\left(-\Omega(\log^2 n)\right),
\end{aligned}$$

provided that $|I_t \cap N_j| = \Omega(\log^6 n)$ (which holds initially for timestep t_1 by assumption). Recall now that the adversary is allowed to choose $\log^3 n$ nodes that never transmit the rumor within the time-interval $[t_1, t_2]$. Since $|I_{t+K} \cap N_{j+1}| \geqslant |\tilde{I}_{t+K} \cap N_{j+1}| - \log^3 n$,

$$\mathbf{Pr}\left[\,|I_{t+K} \cap N_{j+1}| \leqslant \left(1 - \frac{2}{\log n}\right) \cdot \mathbf{E}\left[\,|I_{t+K} \cap N_{j+1}|\,\right]\right] = \exp(-\Omega(\log^2 n)).$$

Let us now compute how many nodes are finally informed in N_{ℓ_2}, provided that $|I_{t+K} \cap N_{j+1}|$ is always close to its expectation. Taking the union bound, it holds with probability $1 - \exp\left(-\Omega(\log^2 n)\right)$ that, for sufficiently large constant $\delta = \delta(\varepsilon) > 0$,

$$\begin{aligned}
&|I_{t_0+K\cdot(\ell_2-(C/2)\log\log n)}| \\
&\geqslant \prod_{j=(C/2)\log\log n}^{\ell_2} \left(1 - \frac{2}{\log n}\right) \cdot \frac{d-j}{j+1} \cdot \left(1 - \frac{1}{\frac{K(j+1)}{d}+1}\right) \cdot |I_{t_1} \cap N_{\ell_1}| \\
&\geqslant \left(1 - \frac{1}{\log n}\right)^{\log n} \cdot \frac{\prod_{j=1}^{\ell_2} \frac{d-j}{j+1}}{\prod_{j=1}^{(C/2)\log\log n} \frac{d-j}{j+1}} \cdot \prod_{j=1}^{d/2} \left(1 - \frac{1}{\frac{K(j+1)}{d}+1}\right) \cdot 1 \\
&\geqslant \frac{1}{16} \cdot \frac{\binom{d}{d/2} \prod_{r=0}^{d/2-\ell_2-1} \frac{d/2-r}{d/2+(r+1)}}{\binom{d}{(C/2)\log\log n}} \cdot \prod_{j=1}^{d/2} \left(1 - \frac{1}{\frac{K(j+1)}{d}+1}\right) \\
&\geqslant \frac{1}{16} \cdot \frac{\frac{1}{2}\frac{n}{\sqrt{\pi d/2}} \cdot 2^{-(\varepsilon/2)n}}{d^{\mathcal{O}(\log\log n)}} \cdot \prod_{j=1}^{d/2} \left(\frac{\frac{K(j+1)}{d}}{1+\frac{K(j+1)}{d}}\right) \\
&= \frac{1}{32} \cdot \frac{\frac{n}{\sqrt{\pi d/2}} \cdot 2^{-(\varepsilon/2)n}}{d^{\mathcal{O}(\log\log n)}} \cdot \frac{1}{\prod_{j=1}^{d/2}\left(1+\frac{d}{K(j+1)}\right)} \\
&\geqslant \frac{1}{32} \cdot \frac{\frac{n}{\sqrt{\pi d/2}} \cdot 2^{-(\varepsilon/2)n}}{d^{\mathcal{O}(\log\log n)}} \cdot \frac{1}{2^{(\varepsilon/3)d}} \\
&\geqslant 2^{(1-\varepsilon)d},
\end{aligned}$$

where the penultimate inequality holds by Lemma 4.2. □

Lemma 3.4. *Let $\varepsilon > 0$ and $\delta(\varepsilon) > 0$ be the constant from Lemma 3.3. Let w be an arbitrary node, $\ell_2 := (1/2 - \delta) \cdot d$ and $\ell_3 := (1/2 - \delta/2) \cdot d > \ell_2$. Suppose that there is a time step t_2 and a subset $U \subseteq V$ satisfying the following three conditions:*

- $|I^+_{t_2} \cap N_{\ell_2}| \geqslant n^{1-\varepsilon}$,
- $I_{t_2} \cap N_{>\ell_2} = \emptyset$,
- $|U \cap N_{\geqslant \ell_3}| \geqslant n^{1-\varepsilon}$.

Assume further that ε is chosen small enough so that $3/2 - 2\varepsilon - \log(e) > 0$. Then in step $t_2 + d$, there exists an informed node in U with probability at least $1 - 2n^{-3}$.

Proof. We define a relation between random walks starting randomly chosen vertices and the spread of the information. The idea is that, if a random walk does not collide with another random walk (or with itself) *and* it starts from an informed node, then the random walk follows a path by which the message is spread. Note that the assumptions that the random walks start at random nodes is for technical reasons only; it allows us to conveniently bound the number of collisions while there will be still enough random walks that start from an uniformed node. We continue to define this relation precisely.

Set $x := n^{1/2-\gamma}$ ($\gamma > 0$ is a sufficiently small constant) and choose x vertices $X^{t_2}_1, X^{t_2}_2, \ldots, X^{t_2}_x$ independently and uniformly at random from V. From each X_j with $X^{t_2}_j \in I_{t_2}$ we consider a walk of maximum length $\log n$ that is defined recursively for $t \geqslant t_2$ as follows. If X^t_j transmits to a node $v \in H_t \cap N_{>\ell_2}$ then we set $X^{t+1}_j := v$. Otherwise the walk X_j terminates at step t. Our goal is to prove that with high probability there is at least one walk X_j that reaches U before step $t_2 + d$, implying that U contains at least one informed node at step $t_2 + d$.

Note that it follows that with probability $1 - n^{-\Omega(1)}$, at least

$$\frac{x}{2} \cdot \frac{n^{1-\varepsilon}}{n} = \frac{1}{2} n^{1/2-\gamma-\varepsilon}$$

random walks start from $I^+_{t_2} \cap N_{\ell_2}$.

Let us first consider the number of collisions between the random walks. Here, a collision occurs if two different random walks visit the same node. Let us expose the x random walks one after the other to estimate their collision probability. We say that a *failure* occurs if the random walk collides with a previously exposed one. Note that in order to have more than C^2 failures ($C > 0$ is a constant to be specified later), there must be at least C different random walks which are involved in a (not necessarily the same) collision. Therefore, we can upper bound the probability for having more than C^2 failures by

$$\binom{n^{1/2-\gamma}}{C} \cdot \left(\frac{(C-1) \cdot \log n \cdot n^{1/2-\gamma}}{n} \right)^C .$$

This bound holds since (i) we have $\binom{n^{1/2-\gamma}}{C}$ possibilities to choose the C random walks which should be involved in a collision, (ii) a random walk involved in a collision has to collide with one of the at

most $(C-1)\log n \cdot n^{1/2-\gamma}$ nodes of the previous random walks and (iii) each random walks starts from a randomly chosen vertex. Let us define the constant $C := \lceil 2/\gamma \rceil$ and let $\mathcal{A}$ be the event that there are at most C^2 collisions among the $n^{1/2-\gamma}$ random walks. It follows that

$$\begin{aligned}\mathbf{Pr}\left[\neg A\right] &\leqslant \binom{n^{1/2-\gamma}}{\lceil \frac{2}{\gamma} \rceil} \cdot \left(\frac{(\lceil \frac{2}{\gamma} \rceil - 1) \cdot \log n}{n^{1/2+\gamma}} \right)^{\lceil \frac{2}{\gamma} \rceil} \\ &\leqslant \left(\frac{e \cdot n^{1/2-\gamma}}{\lceil \frac{2}{\gamma} \rceil} \right)^{\lceil \frac{2}{\gamma} \rceil} \cdot \left(\frac{(\lceil \frac{2}{\gamma} \rceil - 1)^{\lceil \frac{2}{\gamma} \rceil} \log^{\lceil \frac{2}{\gamma} \rceil} n}{n^{(1/2+\gamma)\cdot \lceil \frac{2}{\gamma} \rceil}} \right) \\ &= \mathcal{O}\left(n^{\lceil \frac{2}{\gamma} \rceil \cdot (-2\gamma)} \cdot \log^{\lceil \frac{2}{\gamma} \rceil} n \right) = \mathcal{O}\left(n^{-3.5} \right).\end{aligned}$$

Next we compute the probability that a particular walk X_j with starting node $v := X_j^{t_2} \in I_{t_2}$ reaches a node in U. Note that there is a subset $U' \subseteq U$ with $|U'| \geqslant |U|/d$ such that for all $u' \in U'$, $\mathrm{dist}(v, u')$ is fixed (independent of u'). Let us set $D := \mathrm{dist}(v, u')$. Let $p_{v,u'}$ be the probability that the walk $X_j^{t_2}$ reaches the node u' at step $t_2 + D$ and let $p_{v,U'}$ be the probability that the walk $X_j^{t_2}$ reaches any node in U' at step $t_2 + D$. By disjointness of events, $p_{v,U'} = \sum_{u' \in U'} p_{v,u'}$, and by symmetry of the hypercube, $p_{v,u'}$ is the same for each $u' \in U'$. Let us now lower bound $p_{v,u'}$ for a fixed $u' \in U'$. Certainly, the path X_j reaches u' from v if it follows a shortest path from v to u' that does not return to $N_{\leqslant \ell_2}$. Note that since by assumption $|I_{t_2} \cap N_{>\ell_2}| = 0$, a shortest path from v to u' will not contain any node in I_{t_2} (except the starting node v) provided that the path always uses nodes in $N_{>\ell_2}$.

Let us now ignore all other paths (which could possibly result in a premature termination of X_j) and focus on the event that the path never visits a node in $N_{\leqslant \ell_2}$ (except for the starting node v). Let ℓ be the number of ones in u' and recall that $D = \mathrm{dist}(v, u')$. Hence, v has exactly $(1/2-\delta)d$ ones and u' has at least $\ell_3 = (1/2 - \delta/2)d$ ones. It follows that any shortest path from v to u' has to change α ones into a zero, and has to change $\beta = \alpha + \ell - (1/2 - \delta)d$ zeros into a one. Since $\ell \geqslant (1/2 - \delta/2)d$, it follows that $\beta \geqslant \alpha + (1/2)\delta d$. Clearly, the number of shortest paths between v and u' equals $D!$. Using Bertrand's ballot theorem ([19, pages 299–300]), it follows that the number of shortest paths between v and u' for which only the starting vertex v lies in $N_{\leqslant \ell_2}$ equals

$$D! \cdot \frac{\beta - \alpha}{\beta + \alpha} \geqslant D! \cdot \frac{(1/2)\delta d}{d} = D! \cdot (1/2)\delta.$$

Hence if we were to ignore all other paths (which could result in the termination of X_j), then we would get a lower bound on the probability that the path X_j reaches u' of

$$p_{v,u'} \geqslant D! \cdot (1/2)\delta \cdot d^{-D} \geqslant d! \cdot (1/2)\delta \cdot d^{-d},$$

since $D \leqslant d$. For sufficiently large integer $n \in \mathbb{N}$, Stirling's formula gives $n! \geqslant \frac{1}{2}\sqrt{2\pi n} \cdot (n/e)^n \geqslant (n/e)^n$ and thus

$$p_{v,u'} \geqslant (1/2)\delta \cdot e^{-d}.$$

Consequently,

$$\sum_{u' \in U'} p_{v,u'} \geqslant |U'| \cdot (1/2)\delta e^{-d} \geqslant \frac{|U|}{d} \cdot (1/2)\delta e^{-d} \geqslant (1/2)\delta \frac{n^{1-\varepsilon}}{d} \cdot e^{-d}.$$

Define now a random variable Z that counts the number of paths from $I^+_{t_2} \cap N_{\ell_2}$ to U, if we were to ignore that paths could terminate due to a collision. Linearity of expectation yields

$$\mathbf{E}[Z] \geqslant \frac{1}{2} n^{1/2-\varepsilon-\gamma} \cdot (1/2)\delta \frac{n^{1-\varepsilon}}{d} \cdot e^{-d} = \frac{\delta}{4d} \cdot n^{3/2-2\varepsilon-\log(e)-\gamma}.$$

Let $\mathcal{B} := \{Z \geqslant (1/2) \cdot \mathbf{E}[Z]\}$. Using a Chernoff bound, we obtain that

$$\mathbf{Pr}[\neg\mathcal{B}] \leqslant \exp\left(-\frac{1}{8} \cdot \mathbf{E}[Z]\right) = \exp\left(-\Omega(\mathrm{poly}(n))\right),$$

as $3/2 - 2\varepsilon - \log(e) > 0$ by assumption and $\gamma > 0$ can be made arbitrarily small. Note that the event $\mathcal{A} \wedge \mathcal{B}$ implies that at least one random walk starting from a node in $I^+_{t_2} \cap N_{\ell_2}$ will reach a node in $U' \subseteq U$ before terminating. Taking a union bound, we obtain

$$\mathbf{Pr}[\mathcal{A} \wedge \mathcal{B}] \geqslant 1 - \mathbf{Pr}[\neg A] - \mathbf{Pr}[\neg B] \geqslant 1 - n^{-3} - \exp\left(-\Omega(\mathrm{poly}(n))\right) \geqslant 1 - 2n^{-3}.$$

This completes the proof. □

3.2.2 Analysis of Phase 2 and Phase 3

Lemma 3.5. *With probability at least $1 - n^{-2}$ all nodes are informed at the end of Phase 3.*

Proof. The proof is similar to the proof of Lemma 2.6. Again, define $t_5 := 2\rho \cdot (\log n + (\log\log n)^2)$ as the end of Phase 3, $t_4 := 2 \cdot \rho \log n$ as the beginning of Phase 3, and $t_3 := \rho \log n \geqslant t_2$ as the beginning of Phase 2 (we ignore the $+1$ at the beginning of Phases 2 and 3). We set $\sigma := \rho(\log\log n)^2$.

First we concentrate on Phase 3 and divide the time interval $[t_4 + 1, t_5]$ into $k' = (t_5 - t_4)/\sqrt{\sigma}$ subintervals of length $\sqrt{\sigma}$. For any $0 \leqslant i \leqslant k' - 1$ we define

$$\tilde{\tau}_i = [t_5 - \sqrt{\sigma} i, t_5 - \sqrt{\sigma} \cdot (i+1) + 1].$$

We assume w.l.o.g. that node $v = 0^d$ is uninformed at the end of Phase 3, and show that then there are $\log^3 n$ uninformed nodes in $N_{\sqrt{\sigma}/2}$ at the beginning of Phase 3. For this, for $1 \leqslant i \leqslant \sqrt{\sigma}/2$ we define

$$U_0(v) = L_v[\tilde{\tau}_0] \text{ and } U_i(v) = \cup_{w \in U_{i-1}(v)} L_w[\tilde{\tau}_i].$$

Then, according to Lemma 3.2, we have $|U_{\sqrt{\sigma}/2}(v)| \geqslant \log^{C/2} n$, if ρ is large enough. Now, all nodes of $U_{\sqrt{\sigma}/2}(v)$ must be uninformed at the beginning of Phase 3, since otherwise v becomes informed in the time-interval $[t_4 + 1, t_5]$.

Now we consider the channels opened by the nodes of $U_{\sqrt{\sigma}/2}(v)$ in Phase 2. Recall that in this phase, the nodes informed in Phase 1 perform `pull` with probability $1/\log n$. The nodes which become informed in this phase, perform `pull` with probability 1 in each step of this phase. For the analysis of Phase 2, we divide the time interval $[t_3 + 1, t_4]$ into $k'' = (t_4 - t_3)/\phi$ rounds of length ϕ, where ϕ is a large constant. For any $0 \leqslant i \leqslant k'' - 1$

$$\tilde{\tau}'_i = [t_4 - \phi i, t_4 - \phi \cdot (i+1) + 1].$$

For $0 \leqslant i \leqslant \rho \log n$, let

$$\begin{aligned}
\tilde{U}^H_{-1}(v) &= U_{\sqrt{\sigma}/2}(v) \\
\tilde{U}^H_i(v) &= \cup_{w \in \tilde{U}^H_{i-1}(v)} L_w[\tilde{\tau}'_i] \cap H_{t_3} \\
\tilde{U}^I_i(v) &= \cup_{w \in \tilde{U}^H_{i-1}(v)} L_w[\tilde{\tau}'_i] \cap I_{t_3}.
\end{aligned}$$

Note that a node $\tilde{w}_i \in \tilde{U}_i^I(v)$ is connected to a node $\tilde{w}_{-1} \in \tilde{U}_{-1}^H(v)$ by a path $P = (\tilde{w}_i, \dots, \tilde{w}_0, \tilde{w}_{-1})$, where $\tilde{w}_{i-1}, \dots, \tilde{w}_0, \tilde{w}_{-1} \in H_{t_3}$, and $\tilde{w}_{j+1} \in L_{\tilde{w}_j}[\tilde{\tau}'_{j+1}]$. Assume that there is some $i \leqslant \rho \log n/\phi$ for which $|\tilde{U}_i^I(v)| > \rho \log^2 n$. Then, for every $u \in \tilde{U}_i^I(v)$ there is a path $(u, \tilde{w}_i, \dots, \tilde{w}_0, \dots, v)$, where all nodes $\tilde{w}_i, \dots, \tilde{w}_0, \dots, v \in H_{t_3}$. If u performs `pull` when $\tilde{w}_i$ opens a channel to u, then v becomes informed at the end of Phase 3. However, there are more than $\rho \log^2 n$ such nodes u, and at least one of them performs `pull` at the right time with probability

$$1 - \left(1 - \frac{1}{\log n}\right)^{\rho \log^2 n} = 1 - o(n^{-3}),$$

if ρ is large enough.

If there is no $i \leqslant \rho \log n/\phi$ for which $|\tilde{U}_i^I(v)| > \rho \log^2 n$, then we can apply Lemma 3.3, and obtain that $|\tilde{U}^H_{\rho \log n}(v)| > n^{1-\varepsilon'}$, where ε' can be made an arbitrarily small constant by choosing ρ large enough. Since the hypercube is vertex-transitive, the same result holds for any v that is uninformed at the end of Phase 3.

Our goal is now to apply Lemma 3.4. In that notation, we let $U = \tilde{U}^H_{\rho \log n}(v)$. Using Lemma 3.2 and Lemma 3.3, it follows that $|I^+_{t_2} \cap N_{\ell_2}| \geqslant n^{1-\varepsilon}$ and $I_{t_2} \cap N_{>\ell_2} = \emptyset$, so that the first two preconditions of Lemma 3.4 are satisfied. We now choose $\varepsilon' := \min\{\varepsilon/2, \delta/2\}$, where $\varepsilon > 0$ and $\delta = \delta(\varepsilon) > 0$ are the constants from Lemma 3.3. By Lemma 4.4, $|U \cap N_{\geqslant (1/2-\delta/2)d}| \geqslant \frac{1}{2} n^{1-\varepsilon'} \geqslant n^{1-\varepsilon}$, so that the third precondition of Lemma 3.4 also holds. Applying now Lemma 3.4 we obtain that, with probability $1 - n^{-3}$, at least one node in $U = \tilde{U}^H_{\rho \log n}(v)$ becomes informed in Phase 1, and thus v cannot be uninformed by the end of Phase 3. □

4 Applied Results

In this section we first state some technical results (Section 4.1) that were used for the analysis of the algorithm for hypercubes. In Section 4.2 we show some properties of random graphs that were used for the analysis of the random graph algorithm. Finally, in Section 4.3 we state two tail estimates which we applied in this paper.

4.1 Technical Claims

Lemma 4.1. *There is a sufficiently large constant K, such that for any $d \in \mathbb{N}$ and for any $1 \leqslant j \leqslant d/\sqrt{K}$,*

$$\frac{d + Kj}{Kj} \leqslant \frac{d - j}{K^{2/3} j}.$$

Proof. The claim is equivalent to

$$dK^{2/3} + K^{5/3} j \leqslant Kd - jK,$$

and further rearranging gives

$$j \leqslant d \cdot \frac{K - K^{2/3}}{K^{5/3} + K} \leqslant \frac{d}{\sqrt{K}},$$

if K is sufficiently large. □

Lemma 4.2. *Let $\varepsilon > 0$ be any small constant. Then there is a sufficiently large constant $K > 0$, so that it holds for any $d \in \mathbb{N}$ that*

$$\prod_{j=1}^{d/2} \left(1 + \frac{d}{Kj}\right) \leqslant 2^{(\varepsilon/3)d}.$$

Proof. For sufficiently large K we have,

$$\begin{aligned}
\prod_{j=1}^{d/2} \left(1 + \frac{d}{Kj}\right) &\leqslant \prod_{j=1}^{d/\sqrt{K}} \left(1 + \frac{d}{Kj}\right) \cdot \prod_{j=d/\sqrt{K}+1}^{d/2} \left(1 + \frac{d}{\sqrt{K}d}\right) \\
&\leqslant \prod_{j=1}^{d/\sqrt{K}} \left(\frac{d+Kj}{Kj}\right) \cdot \left(1 + \frac{1}{\sqrt{K}}\right)^{d/2-d/\sqrt{K}} \\
&\leqslant \prod_{j=1}^{d/\sqrt{K}} \left(\frac{d-j}{K^{2/3}j}\right) \cdot \left(1 + \frac{1}{\sqrt{K}}\right)^{d/2-d/\sqrt{K}} && \text{(by Lemma 4.1)} \\
&\leqslant \binom{d}{d/\sqrt{K}} \cdot K^{-(2/3)\cdot(d/\sqrt{K})} \cdot \left(1 + \frac{1}{\sqrt{K}}\right)^{d/2} \\
&\leqslant \left(\sqrt{K}e\right)^{d/\sqrt{K}} \cdot K^{-(2/3)\cdot(d/\sqrt{K})} \cdot \left(1 + \frac{1}{\sqrt{K}}\right)^{d/2} \\
&\leqslant 2^{(1/100)\cdot d}.
\end{aligned}$$

□

Lemma 4.3. *Let $x_1, x_2, \ldots, x_n \in [0, M]$ and $X := \sum_{i=1}^n x_i$. Then it holds for any $\lambda > 0$,*

$$\sum_{i=1}^n \lambda^{-x_i} \leqslant \frac{X}{M} \cdot \lambda^{-M} + \left(n - \frac{X}{M}\right) \cdot \lambda^0.$$

Proof. As the function $x \mapsto \lambda^{-x}$ is convex, we have that

$$\lambda^{-x_i} = \lambda^{-(x_i/M)\cdot M - (1-x_i/M)\cdot 0} \leqslant \frac{x_i}{M} \cdot \lambda^{-M} + \left(1 - \frac{x_i}{M}\right) \cdot \lambda^{-0}.$$

This implies

$$\sum_{i=1}^n \lambda^{-x_i} \leqslant \sum_{i=1}^n \frac{x_i}{M} \cdot \lambda^{-M} + \sum_{i=1}^n \left(1 - \frac{x_i}{M}\right) \cdot \lambda^{-0} = \frac{X}{M} \cdot \lambda^{-M} + \left(n - \frac{X}{M}\right) \cdot \lambda^0,$$

as needed. □

Lemma 4.4. *Let $U \subseteq \{0,1\}^d$ be any subset of the hypercube. For any $0 \leqslant i \leqslant d$, let $N_{\geqslant i} := \{v \in \{0,1\}^d \colon |v|_1 \geqslant i\}$. If $|U| \geqslant n^{(1-\varepsilon)}$ for a constant $\varepsilon > 0$, then it follows that $|U \cap N_{\geqslant (1/2-\varepsilon)d+1}| \geqslant (1/2)n^{1-\varepsilon}$ for sufficiently large d.*

Proof. We first upper bound the size $N_{\leqslant(1/2-\varepsilon)d}$. In order to do that, we use the probabilistic method. More specifically, let $X_1, \ldots, X_d$ be independent 0/1-random variables with $\mathbf{Pr}[X_i = 1] = \mathbf{Pr}[X_i = 0] = 1/2$. Let $X := \sum_{i=1}^{d} X_i$ and $\mu := \mathbf{E}[X] = d/2$. Then by the Hoeffding bound,

$$\left|N_{\leqslant(1/2-\varepsilon)d}\right| \leqslant 2^d \cdot \mathbf{Pr}[X \geqslant \mathbf{E}[X] - \varepsilon d] \leqslant 2^d \cdot e^{-\frac{(\varepsilon d)^2}{d}} \leqslant \frac{1}{2} \cdot 2^{(1-\varepsilon)d}.$$

Hence,

$$|U \cap N_{\geqslant(1/2-\varepsilon)d+1}| \geqslant |U| - \left|N_{\leqslant(1/2-\varepsilon)d}\right| \geqslant \frac{1}{2} n^{1-\varepsilon}.$$

□

4.2 Random Graph Properties

In this section we show some combinatorial properties of random graphs. Some of these properties (Lemmas 4.5 and 4.6) are also derived (in a modified form) in [8] for (almost) Ramanujan graphs. Recall that we consider $G = G(n,p)$ with $(\log^2 n)/n \leqslant p \leqslant 2^{o(\sqrt{\log n})}/n$.

Lemma 4.5. *Let $\varepsilon < 1$ be a suitably chosen constant, and fix $x \in [6n/d, n/2]$. Let $\mathcal{X} = \{(X,Y) |\ Y \subset X \subset V, |X| = x,$ and $|Y| = x/4\}$, and let $N(u,S) := \{v \in S \mid (u,v) \in E\}$ for some $u \in V$ and $S \subseteq V$.*

1. *For $x \leqslant n/40$ let $\mathcal{A}$ be the event that for all $(X,Y) \in \mathcal{X}$ there exists $Y' \subset V \setminus X$ such that 1) $|Y'| = n/6$, and 2) for all $y' \in Y' :\ |N(y',Y)| \geqslant |Y| \cdot d/(2n)$. Then*

$$\mathbf{Pr}[\mathcal{A}] \geqslant 1 - \varepsilon^n.$$

2. *For $n/40 < x \leqslant n/2$ let $\mathcal{B}$ be the event that for all $(X,Y) \in \mathcal{X}$ there exists $Y' \subset V \setminus X$ such that 1) $|Y'| = (n-x)/2$, and 2) for all $y' \in Y' :\ |N(y',Y)| \geqslant |Y| \cdot d/(2n)$. Then*

$$\mathbf{Pr}[\mathcal{B}] \geqslant 1 - \varepsilon^n.$$

Proof. We define $y = |Y|$. Now fix two subsets Y_1 of size y and X_1 of size x with $Y_1 \subset X_1$. For $1 \leqslant i \leqslant n - x$, let $Z_i = 1$ if the ith node of $V \setminus X_1$ has at most $yd/(2n)$ neighbors in Y_1 and 0 otherwise. Define $Z = Z_1 + Z_2 \cdots + Z_{n-x}$. Every node of $V \setminus X_1$ is connected to a fixed node of Y_1 with probability p, independently form all other nodes. With $d = pn$ we get for $1 \leqslant i \leqslant n - x$

$$\begin{aligned}
\mathbf{Pr}[Z_i = 1] &\leqslant \sum_{i=y-yd/(2n)}^{y} \binom{y}{i} \cdot (1-p)^i \cdot p^{y-i} \\
&\leqslant \left(\frac{1-p}{1-d/(2n)}\right)^{y(1-d/(2n))} \cdot \left(\frac{p}{d/(2n)}\right)^{yd/2n} \\
&\leqslant \left(1 - \frac{d}{2n}\right)^{y(1-d/(2n))} \cdot 2^{yd/2n} \leqslant \left(\frac{1}{e}\right)^{y \cdot \frac{d}{2n} \cdot (1-d/(2n))} \cdot 2^{yd/2n} \\
&\leqslant \left(\frac{2}{e \cdot (1-o(1))}\right)^{yd/(2n)} \leqslant \left(\frac{2 \cdot (1+o(1))}{e}\right)^{yd/(2n)}
\end{aligned}$$

Now we consider two cases, depending on x.

Case 1: $x \leqslant n/40$**.** In this case we have

$$\mathbf{Pr}\,[\,Z_i = 1\,] \leqslant \left(\frac{2\cdot(1+o(1))}{e}\right)^3 < 1/2.$$

This gives us $E[Z] \leqslant (n-x)/2$. Now we can use Chernoff bounds (Equation (12) in [16]) and show that with $x \leqslant n/40$ we get

$$\begin{aligned}
\mathbf{Pr}\,[\,Z \geqslant 5n/6 - x\,] &\leqslant \left(\frac{(n-x)/2}{5n/6-x}\right)^{5n/6-x} \cdot \left(\frac{(n-x)/2}{n-x-(5n/6-x)}\right)^{n/6} \\
&\leqslant \left(\frac{(n-n/40)/2}{5n/6-n/40}\right)^{5n/6-n/40} \cdot \left(\frac{(n)/2}{n-(5n/6)}\right)^{n/6} \\
&\leqslant \left(\frac{117}{194}\right)^{5n/6-n/40} \cdot 3^{n/6} \leqslant \left(\frac{4}{5}\right)^n
\end{aligned}$$

Now we can conclude that, with probability $1-(4/5)^{-n}$ we have at least $n/6$ nodes in $V \setminus X$ with $yd/(2n)$ (or more) neighbors in Y_1.

There are $\binom{n}{x}$ different ways to choose the nodes for the set X_1. Furthermore, there are $\binom{x}{x/4}$ possible ways to choose Y_1 as subsets of size $x/4$ from the nodes of X. Thus, for every subset Y of size $x/4$ of an arbitrary set X there exists a subset $Y' \subset V \setminus X$ of size $n/6$ such that each node of Y' has at least $yd/2n$ neighbors in Y with probability at least

$$\begin{aligned}
1 - \binom{n}{x}\cdot\binom{x}{x/4}\cdot\left(\frac{4}{5}\right)^n &\geqslant 1 - \left(\frac{n\cdot e}{x}\right)^x \cdot \left(\frac{x\cdot e}{x/4}\right)^{x/4} \cdot \left(\frac{4}{5}\right)^n \\
&\geqslant 1 - \left(\frac{n\cdot e}{n/40}\right)^{n/40} \cdot \left(\frac{n/40\cdot e}{n/160}\right)^{n/160} \cdot \left(\frac{4}{5}\right)^n \\
&= 1 - (40e)^{n/40}\cdot(40e)^{n/160}\cdot\left(\frac{4}{5}\right)^n \leqslant 1-\varepsilon^n.
\end{aligned}$$

Case 2: $n/40 \leqslant |X| \leqslant n/2$**.** In this case we have

$$\begin{aligned}
\mathbf{Pr}\,[\,Z_i = 1\,] &\leqslant \left(\frac{2}{e\cdot(1-o(1))}\right)^{yd/(2n)} \leqslant \left(\frac{2\cdot(1+o(1))}{e}\right)^{yd/(2n)} \\
&\leqslant \left(\frac{2\cdot(1+o(1))}{e}\right)^{d/(320)} \leqslant \frac{1}{n^2}
\end{aligned}$$

and

$$\mathbf{Pr}\,[\,Z \geqslant (n-x)/2\,] \leqslant \binom{n-x}{(n-x)/2}\cdot\left(\frac{1}{n^2}\right)^{(n-x)/2} \leqslant \left(\frac{2e}{n^2}\right)^{(n-x)/2}.$$

With the same probability there are at least $(n-|X|)/2$ nodes in $V \setminus X$ with $yd/(2n)$ or fewer neighbors in Y_1. Again, for every subset Y of size $x/4$ of an arbitrary set X there exists a subset

$Y' \subset V \setminus X$ of size $n/6$ such that each node of Y' has at least $yd/2n$ neighbors in Y with probability at least

$$1 - \binom{n}{x} \cdot \binom{x}{x/4} \cdot \left(\frac{2e}{n^2}\right)^{(n-x)/2} \geqslant 1 - 2^n \cdot (4e)^{x/4} \cdot \left(\frac{2e}{n^2}\right)^{(n-x)/2}$$
$$\geqslant 1 - 2^n \cdot (4e)^{n/8} \cdot \left(\frac{2e}{n^2}\right)^{n/4} \geqslant 1 - \varepsilon^n$$

□

Next, we consider the connectivity between two very large sets.

Lemma 4.6. *Let $\varepsilon < 1$ be a suitably chosen constant, and let*

$$\mathcal{X} = \{(X, Y) |\ Y \subset X \subset V, |X| = n/2, \text{ and } |Y| = n/5\}.$$

Let $\mathcal{C}$ be the event that for all $(X, Y) \in \mathcal{X}$ there exists $Y' \subset V \setminus X$ such that 1) $|Y'| = n/2 - n \log\log n / \log n$, and 2) for all $y' \in Y' :\ |N(y', Y)| \geqslant d/10$. Then

$$\mathbf{Pr}\,[\,\mathcal{C}\,] \geqslant 1 - \varepsilon^n.$$

Proof. The proof is similar to the proof of Lemma 4.5. Let X_1 and $Y_1 \subseteq X_1$ be two sets of size $n/2$ and $n/5$, respectively. For $1 \leqslant i \leqslant n - x$, let $Z_i = 1$ if the ith node of $V \setminus X_1$ has at most $d/10$ neighbors in Y_1 and 0 otherwise. Define $Z = Z_1 + Z_2 \cdots + Z_{n-x}$. Similar to Lemma 4.5 we get for $1 \leqslant i \leqslant n - x$

$$\mathbf{Pr}\,[\,Z_i = 1\,] \leqslant \left(\frac{2(1 + o(1))}{e}\right)^{d/10}.$$

Thus, there are more than $n \log\log n / \log n$ nodes in $V \setminus X_1$ with fewer than $d/10$ neighbors in Y_1 with probability

$$\sum_{i=\frac{n \log\log n}{\log n}}^{n/2} \binom{n/2}{i} p'^{d/10} (1 - p')^{n/2-i}$$
$$\leqslant \left(\frac{p'}{\log\log n / \log n}\right)^{\frac{n \log\log n}{\log n}} \left(\frac{1 - p'}{1 - \log\log n / \log n}\right)^{\frac{n}{2}\left(1 - \frac{\log\log n}{\log n}\right)},$$

which equals $e^{-\Omega(n \log\log n)}$. Now there are $\binom{n}{n/2}$ different subdivisions of V into two subsets of size $n/2$ and there are $\binom{n/2}{n/5}$ different subsets of size $n/5$ in a subset of size $n/2$. Thus, the statement of the lemma holds with probability

$$1 - \binom{n}{n/2} \cdot \binom{n/2}{n/5} \cdot e^{-\Omega(n \log\log n)}$$
$$\geqslant 1 - (2e)^{n/2} \cdot \left(\frac{5e}{2}\right)^{n/5} e^{-\Omega(n \log\log n)} \geqslant 1 - e^{-\Omega(n \log\log n)}.$$

□

The next lemma deals with the local neighborhood-structure around a node in a $G(n,p)$ graph.

Lemma 4.7. *Let $v \in V$ be an arbitrary node in $G(n,p)$ and let $T(v)$ be the graph induced by the nodes at distance at most $\rho \log\log n$ from v. Then, with probability $1 - o(n^{-3})$, the graph $T(v)$ is either a tree, or there are at most 4 edges which violate the tree property in $T(v)$.*

Proof. In the following we denote v as the root of $T(v)$. The nodes at distance ℓ from v are called nodes on level ℓ in the following. Then v is on level 0. For $w, w' \in T(v)$, w is called ancestor of node w' iff $dist(v,w) < dist(v,w')$. For $\leqslant i \leqslant \rho \log\log n$, N_i will be the set of nodes on level i.

We can assume that $G(n,p)$ is constructed by the following procedure. In the first step $v = v_1^0$ draws an edge to every other node with probability p. This gives our level 1 nodes $N_1 = \{v_1^1, v_2^1, \dots\}$. In the second step the edges between v_1^1 and $V \setminus \{N_0, N_1\}$ are chosen in the same way. Then we choose the edges between v_2^1 and $V \setminus \{N_0, N_1\}$, v_3^1 and $V \setminus \{N_0, N_1\}, \dots$, until all nodes in N_1 are considered. The nodes that are connected to nodes in N_1 in this way are the nodes of $N_2 = v_1^2, v_2^2, \dots$. We do the same for the nodes nodes in N_2 (considering only nodes in $V \setminus \{N_0, N_1, N_2\}$) , which gives us the set $N_3 = v_1^3, v_2^3, \dots$, and so on. For $1 \leqslant i \leqslant \rho \log\log n - 1$ we use the nodes in N_i to create $N_{i+1} = v_1^{i+1}, v_2^{i+1}, \dots$. For $v_j^i \in N_i$ we consider only the nodes in $V \setminus \{N_0, N_1, \dots N_i\}$.

So far we created a tree with some additional edges. Two different nodes on level j can be connected to the same node on level $j+1$. So far we did not evaluate all events for the nodes in the tree, since a node on level j can also be connected to other nodes on the same level. We evaluate these events at the end of the process. In the following we call these edges cycle edges and our goal is to upper bound the number of cycle edges in the tree.

First we calculate the maximum number of nodes of $T(v)$. Recall that $p \leqslant 2^{o(\sqrt{\log n})}/n$. We know that with probability $1 - o(n^{-5})$ the maximum degree of G is $2 \cdot pn$ [4]. Hence, w.h.p. $T(v)$ has at most

$$(2 \cdot pn)^{\rho \log\log n} \leqslant (2 \cdot 2^{o(\sqrt{\log n})})^{\rho \log\log n} \leqslant \log^{O(1)} n.$$

many nodes. Hence, the probability that there are more than 4 cycle edges is at most

$$\binom{(\log^{O(1)} n)^2}{5} \cdot p^5 \leqslant (\log n)^{O(1)} \cdot \left(\frac{2^{o(\sqrt{\log n})}}{n}\right)^5 \leqslant \frac{1}{n^4}$$

□

4.3 Probabilistic and Combinatorial Tools

Lemma 4.8 (Method of Bounded Independent Differences, [18])**.** *Let $X_i \colon \Omega_i \to \mathbb{R}$, $1 \leqslant i \leqslant n$, be mutually independent random variables. Let $f \colon \prod_{i=1}^n \Omega_i \to \mathbb{R}$ 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 $f(X_1, \dots, X_n)$. Then for any $t \geqslant 0$,

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

and

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

We need the following standard Chernoff bound.

Lemma 4.9. *Consider some fixed $0 < \delta < 1$. Suppose that $X_1, \ldots, X_n$ are independent geometric random variables on $\mathbb{N}$ with $\mathbf{Pr}\left[\, X_i = k \,\right] = (1-\delta)^{k-1}\delta$ for every $k \in \mathbb{N}$. Let $X = \sum_{i=1}^{n} X_i$, $\mu = \mathbf{E}\left[\, X \,\right]$. Then it holds for all $\varepsilon > 0$ that*

$$\mathbf{Pr}\left[\, X \geqslant (1+\varepsilon)n/\delta \,\right] \leqslant e^{-\varepsilon^2 n/2(1+\varepsilon)}$$

5 Conclusions

In this paper we consider rumor spreading on random graphs and hypercubes in the quasirandom phone call model. We show two results. For random graphs we present an address-oblivious algorithm with runtime $O(\log n)$ that uses at most $O(n \log\log n)$ message transmissions. For hypercubes of dimension $\log n$ we present an address-oblivious algorithm with runtime $O(\log n)$ that uses at most $O(n(\log\log n)^2)$ message transmissions. Together with a result of [9], our results imply that for random graphs the communication complexity of the quasi random phone call model is significantly smaller than that of the standard phone call model.

Open problems include a generalisation of our results for general random graphs where the nodes can have very different degrees. Also, it might be interesting to show results for additional deterministic graphs like star graphs or grids. And, of course, another open problem is to show (this time a correct proof) of the lower bound of that is stated in [3].

Acknowledgements. We would like to thank the reviewers of this journal version for their helpful comments.

References

[1] P. Berenbrink, R. Elsässer, T. Friedetzky. Efficient Randomised Broadcasting in Random Regular Networks with Applications in Peer-to-Peer Systems. In *Proc. of PODC'08*, pages 155–164, 2008.

[2] P. Berenbrink, R. Elsässer and T. Sauerwald. Randomised Broadcasting: Memory vs. Randomness. In *Proc. of LATIN'10*, pages 306–319, 2010.

[3] P. Berenbrink, R. Elsässer and T. Sauerwald. Communication Complexity of Quasirandom Rumor Spreading. In *Proc. of ESA'10*, pages 134–145, 2010.

[4] B. Bollobás. *Random Graphs*. Academic Press, 1985.

[5] F. Chierichetti, S. Lattanzi, and A. Panconesi. Almost Tight Bounds for Rumour Spreading with Conductance. In *Proc. of STOC'10*, pages 399–408, 2010.

[6] A. Demers, D. Greene, C. Hauser, W. Irish, J. Larson, S. Shenker, H. Sturgis, D. Swinehart, and D. Terry. Epidemic algorithms for replicated database maintenance. In *Proc. of PODC'87*, pages 1–12, 1987.

[7] B. Doerr, T. Friedrich, T. Sauerwald. Quasirandom Rumor Spreading In *Proc. of SODA'08*, pages 773–781, 2008.

[8] B. Doerr, T. Friedrich, T. Sauerwald. Quasirandom rumor spreading: expanders, push vs. pull, and robustness In *Proc. of ICALP'09*, pages 366-377, 2009.

[9] R. Elsässer. On the communication complexity of randomized broadcasting in random-like graphs. In *Proc. of SPAA'06*, pages 148–157, 2006.

[10] R. Elsässer and T. Sauerwald. The power of memory in randomized broadcasting. In *Proc. of SODA'08*, pages 218–227, 2008.

[11] P. Erdős and A. Rényi. On random graphs I. *Publ. Math. Debrecen*, 6:290–297, 1959.

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

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

[14] N. Fountoulakis, A. Huber and K. Panagiotou. Reliable broadcasting in random networks and the effect of density. In *Proc. of INFOCOM'10*, pages 2552–2560, 2010.

[15] G. Giakkoupis. Tight bounds for rumor spreading in graphs of a given conductance. In *Proc. of STACS'11*, pages 57-68, 2011.

[16] T. Hagerup and C. Rüb. A guided tour of Chernoff bounds. *Information Processing Letters*, 36(6):305–308, 1990.

[17] R. Karp, C. Schindelhauer, S. Shenker, and B. Vöcking. Randomized rumor spreading. In *Proc. of FOCS'00*, pages 565–574, 2000.

[18] C. McDiarmid. On the method of bounded differences. *Surveys in Combinatorics, 1989*, J. Siemons ed., London Mathematical Society Lectures Note Series 141, Cambridge University Press, 1989, 148-188.

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

[20] B. Pittel. On spreading a rumor. *SIAM Journal on Applied Mathematics*, 47(1):213–223, 1987.