Randomized Rumor Spreading

R. Karp* C. Schindelhauer† S. Shenker‡ B. Vöcking§

Abstract

This paper deals with the problem of spreading rumors in a distributed environment using randomized communication. In particular, we envisage the class of so-called epidemic algorithms which are commonly used for the lazy transmission of updates to distributed copies of a database. We introduce the *random phone call model* in order to investigate the possibilities and limits of this class of broadcasting algorithms. In this model, n players communicate in parallel communication rounds. In each round, each player calls a randomly selected communication partner. Whenever communication is established between two players, each one must decide which rumors to transmit. The major problem (arising due to the randomization) is that players do not know which rumors their communication partners have already received. In order to illustrate this problem, we will give a simple example of a commonly used algorithm in which each individual rumor is transmitted $\Theta(n \ln n)$ times.

In this paper, we investigate whether a large communication overhead is inherent to epidemic algorithms using randomized rumor spreading or can be reduced significantly. We show that there is an algorithm using only $O(\ln n)$ rounds and $O(n \ln \ln n)$ transmissions. We prove the robustness of this algorithm against adversarial node failures and inaccuracies in the randomized selection of the communication partners. Furthermore, we show that our algorithm is optimal among those algorithms in which the actions of the players do not depend on the addresses of their communication partners. Finally, we give a lower bound for general algorithms showing that time- and communication-optimality cannot be achieved simultaneously. In particular, we prove that any algorithm (based on the random phone call model) that distributes a rumor in $O(\ln n)$ rounds needs to send $\omega(n)$ messages on expectation.

*Email: karp@icsi.berkeley.edu. International Computer Science Institute, Berkeley.

†Email: schindel@icsi.berkeley.edu. International Computer Science Institute, Berkeley. Parts of this work are supported by a stipend of the "Gemeinsames Hochschulsonderprogramm III von Bund und Länder" through the DAAD.

‡Email: shenker@icsi.berkeley.edu. International Computer Science Institute, Berkley.

§Email: voecking@cs.umass.edu. University of Massachusetts, Amherst.

1 Introduction

We investigate the problem of spreading rumors in a distributed environment using randomized communication. Suppose n players exchange information in parallel communication rounds over an indefinite time. In each round t, the players are connected by a communication graph G_t. This graph is generated at random in distributed fashion, that is, in each round, each player u selects a communication partner v at random and u calls v. Rumors can be started in any round by any player and can be transmitted along the edges in the graph G_t in round t. The goal is to spread the rumor among all participating players using a small number of rounds and a small number of transmissions.

The motivation for using randomized communication is that it naturally provides robustness, simplicity, and scalability. For example, consider the following so-called *push algorithm*. Starting with the round in which a rumor is generated, each player that holds the rumor forwards it to a communication partner selected independently and uniformly at random (i.u.r.). The algorithm is terminated after some fixed number of $O(\ln n)$ rounds. At this time all players are informed, with high probability (w.h.p.)[1].

Clearly, one can also inform all players in $O(\ln n)$ using a deterministic interconnection of constant degree, e.g., a shuffle network. (For an overview of deterministic information dissemination we refer to [5],[6].) The advantage of the push algorithm, however, is its implicit robustness against several kinds of failures compared to the deterministic case where either additional time is needed [4] or the error fraction is polynomial [11]. For example, consider node failures, i.e., a player (different from the player starting the rumor) fails to communicate or simply crashes and forgets its rumors. Obviously, when using a sparse deterministic network, even a single node failure can result in a large fraction of players not receiving the rumor. When using the randomized push algorithm, however, the effects of node failures are very limited. In fact, it is not difficult to prove that F node failures (specified by an oblivious adversary) result in only $O(F)$ uninformed players, w.h.p.

Unfortunately, the push algorithm produces a large communication overhead. In fact, it forwards each individual rumor for $\Theta(n \ln n)$ times until all players are informed, in comparison to a deterministic scheme which requires only $n-1$ transmissions. It seems that the large number of transmissions is the price for the robustness. This gives rise to the question whether this additional communication effort is a special property of the above push algorithm or is inherent to rumor spreading using randomly generated communication graphs in general.

1.1 Background

Demers et al. [2] introduced the idea of using so-called epidemic algorithms for the lazy update of data objects in a data base replicated at many sites, e.g., yellow pages, name servers, or server directories. In particular, they propose the following two concepts:

- *Anti-entropy:* Every site regularly chooses another site at random and by exchanging database contents with it resolves any differences between the two.
- *Rumor mongering:* When a site receives a new update it becomes a "hot rumor". While a site holds a "hot rumor", it periodically chooses another site at random and ensures that the other site has seen the rumor.

[1] The term with high probability (w.h.p.) means with probability at least $1 - O(n^{-\alpha})$ for some positive constant α.

It turns out that anti-entropy is extremely reliable but produces an enormous amount of communication such that it cannot be used too frequently. The idea of rumor mongering is to exchange only recent updates and thereby reducing the communication overhead significantly. In this paper, we investigate algorithms implementing the rumor mongering concept.

The original idea for rumor spreading was to send rumors only from the caller to the called player (*push transmission*) [2]. Several termination mechanisms deciding when a rumor becomes "cold" so that it transmission is stopped were investigated. All these algorithms share the same phenomenon: the fraction u of players that do not know a particular rumor decreases exponentially with the number of transmissions t (i.e., messages that contain this rumor). Mean field equations lead to the conjecture that $u \approx \exp(-t/n)$ for all investigated variants of the push algorithm. Thus, a push algorithm needs about $n \ln n$ transmissions for sending a rumor to all players.

A further idea introduced in [2] is to send rumors from the called to the calling player (*pull transmission)*. It was observed that the number of uninformed players decreases much faster using a pull scheme instead of a push scheme if updates occur frequently so that (almost) every player places a random call in each round. Experiments and mean field equation lead to the conjecture $u \approx \exp(-(t/n)^3)$ (for some specific pull algorithms) so that $n\sqrt[3]{\ln n}$ transmissions are sufficient to inform all players.

The work of Demers et al. initiated an enormous amount of experimental and conceptual studies of epidemic algorithms. For example, there is a variety of research issues for distributed epidemic algorithms like consistency, correctness, data structures, and efficiency [1, 7, 8, 9, 10]. In this extended abstract, we concentrate only on the efficiency of these randomized algorithms. In particular, we study their time and communication complexity using a simple model for the underlying randomized communication.

1.2 The random phone call model

Let V denote the set of players. The communication graph $G_t = (V, E_t \subseteq V \times V)$ of round t is obtained by a distributed, randomized process. In each round, each player u chooses a communication partner v from V at random and u *calls* v. Unless otherwise stated, we assume that all players choose their communication partners i.u.r. from V.

Even though we envisage an application (such as the lazy transmission of updates to distributed copies of a database) in which rumors are constantly generated by different players, our analysis is concerned with the distribution of a single rumor only. We focus on the lifetime of the rumor and the number of transmissions rather than the number of connections established because the latter cost is amortized over all the rumors using that connection.

In round t, the rumor and other information can be exchanged only along the edges of G_t. Whenever a connection is established between two players, each one of them (if holding the rumor) has to decide whether to transmit the rumor to the other player, typically without knowing whether this player has received the rumor already. Communication in each round is assumed to proceed in parallel, that is, any information received in a round cannot be forwarded to another player in the same round. We do not limit the size of the information exchanged. Each information exchange between neighboring players in a round is counted as a single transmission. (We point out that our algorithms only add small counter values to rumors, whereas our lower bounds hold even for algorithms in which players exchange their complete history whenever the rumor is sent in either direction.)

An algorithm is called *distributed* if all decisions (whether to sent a rumor) are based on local knowledge only. In particular, the decision whether player sends a message to a communication partner in round t depends only on the player's *state* in that round. The initial state of a player

is defined by the player's address, the number of players, and possibly a random bit string. In general the state of a player in round t is a function of its initial state, the addresses of the neighbors in the communication graphs $G_1, \ldots, G_t$, and the information received in the rounds 1 to $t-1$. (For our lower bounds one may also assume that the state depends also on a globally known round number as well as the birth date of the considered rumor.)

Finally, an algorithm is called *address-independent* if a player's state in round t does not depend on the addresses of the neighbors in G_t but only on the number of neighbors in G_t. (For example, all rumor spreading algorithms proposed by Demers et al. [2] are address-independent.)

1.3 New results

We prove that the number of transmissions can be reduced significantly when the rumor is sent in either direction, that is, when using push and pull rather than only push operations. We introduce a simple *push&pull algorithm* spreading the rumor to all players in $O(\ln n)$ rounds using only $O(n \ln\ln n)$ transmissions rather than $O(n \ln n)$ as the push algorithm

The drawback of the push&pull-algorithm is that its success heavily relies on a very exact, global estimation of the right termination time. This mechanism is very sensitive to any kind of errors that influence the expansion of the set of informed players. We devise a distributed termination scheme, called the *median-counter algorithm*, that is provably robust against adversarial node failures and stochastic inaccuracies in establishing the random connections.

In particular, we show that the efficiency of the algorithm does not rely on the fact that players choose their communication partners uniformly from the set of all players. Suppose all players use the same arbitrary probability distribution $\mathcal{D} : V \to [0,1]$ rather than the uniform distribution. We show that the median-counter algorithm takes $O(\ln n)$ rounds and needs only $O(n \ln\ln n)$ transmissions regardless of the probability distribution used for establishing the random connections. For example, this feature allows sampling from an arbitrary address directory (possibly with redundant addresses and some non-listed players as in a telephone book) rather than sampling uniformly from the set of players itself. Thus, the algorithm can be executed even without global knowledge about the set of players.

In addition, we provide lower bounds on the number of required transmissions assuming that the communication graphs are obtained using the uniform probability distribution. The algorithms above are address-independent and perform $O(n \ln\ln n)$ transmissions. We prove a corresponding lower bound showing that any address-independent algorithm needs $\Omega(n \log\log n)$ transmissions in order to inform all players. We point out that this bound holds independently of the number of rounds executed.

The situation changes substantially when considering address-dependent algorithms. Allowing $\Theta(n \log n)$ rounds, an address-dependent algorithm can spread the rumor using only $n-1$ transmissions. For example, the player initiating the rumor can simply wait until each of the other players appears as communication partner for the first time and then forward the rumor to this player. Clearly, this is not a practical algorithm as it takes too many rounds. Nevertheless, it illustrates the additional possibilities of address-dependent algorithms.

The above example leads to the question of whether address-independent algorithms can spread a rumor in a small number of rounds while using only a linear number of transmissions. We give a lower bound answering this question negatively. In particular, we show that any randomized rumor spreading algorithm running $O(\log n)$ rounds requires $\omega(n)$ transmissions, regardless of the amount of information that can be attached to the rumors. Thus, there is a fundamental gap between rumor spreading algorithms based on random interconnections and deterministic broadcasting schemes.

2 Upper Bounds

2.1 The advantage of push&pull

First, let us explain the differences in the propagation of the rumor obtained by push transmissions on the one hand and pull transmissions on the other hand.

- Consider a *push scheme* in which every informed player, in every round, forwards the rumor to the player it calls until all players are informed. In this case the set of informed player grows exponentially until about $n/2$ players are informed. At about this time the exponential growth of the set of informed players stops. Starting from this point of time, let us consider the set of uninformed players. Once half of the players are informed, this set shrinks by a constant factor in each round. At the end of the rumor spreading process this factor is about $1 - 1/\mathrm{e}$ since the fraction of players that do not receive a call in a round is about $1/\mathrm{e}$. Thus, the shrinking phase takes $\Theta(\ln n)$ rounds until every player has received the rumor, and the push algorithm sends $\Theta(n)$ messages in each of these rounds.

- Now consider a *pull scheme* in which only called players send the rumor towards the calling players. In this case, the player starting the rumor may have to wait some rounds until it is called for the first time so that the propagation in the first rounds becomes unpredictable. But eventually (after $O(\ln n)$ rounds, w.h.p.) about $n/2$ of the players will be informed. From this time on, the pull algorithm has an advantage against the push algorithm as the fraction of uninformed players roughly squares from round to round. This is because in a round starting with ϵn uninformed players, each individual player has probability $1 - \epsilon$ to receive the rumor, so that the probability of staying uninformed is ϵ, resulting in an expected number of $\epsilon^2 n$ uninformed players at the end of the round. Thus, we can expect that the shrinking phase only takes $\Theta(\ln \ln n)$ rounds so that only $\Theta(n \ln \ln n)$ messages are sent during this phase.

In order to combine the predictability of the push scheme with the quadratic-shrinking property of the pull scheme, we simply sent the rumor in both directions whenever possible. In detail, our *push&pull scheme* works as follows. The creator of the rumor initiates a time-counter with 0 representing the *age* of the rumor. The age is incremented in every round and distributed with the rumor. In every round every informed player pushes and pulls unless the age of the rumor is higher than $t_{\max} = \log_3 n + O(\log \log n)$. In the following theorem, we assume the uniform distribution and a perfect interconnection without failures.

Theorem 2.1 *The push&pull-scheme informs all players in time* $\log_3 n + O(\log \log n)$ *using* $O(n \log \log n)$ *messages w.h.p.*

Proof. Let S_t be the set of informed players and U_t the set of uninformed players at the end of round t. Define $s_t = |S_t|$ and $u_t = |U_t|$. We distinguish four consecutive phases.

1. The *startup phase* starts in the round in which the rumor is created and ends with the first round after which execution there are at least $(\ln n)^4$ informed players for the first time. At the beginning of the first round only one player holds the rumor. If we execute c rounds then the probability that this player has called at least once an uninformed player (i.e., did not call itself) is $1 - n^{-c}$. Thus, we double the number of players in c rounds, w.h.p. In general, starting with at most $(\ln n)^4$ informed players, we need at most c rounds to double the number of informed players, w.h.p. Thus $O(\ln \ln n)$ rounds are sufficient to achieve $(\ln n)^4$ informed players.

2. The *exponential-growth phase* ends with the round after which execution there are at least $n/\ln n$ informed players for the first time. The expected number of messages (containing the rumor) sent during round t in this phase is $2s_{t-1}$ because each player holding the rumor calls one player and is called by one player on expectation. Applying a Chernoff bound yields that the number of actually sent messages is $m = (2 \pm o(1/\ln n))s_{t-1}$, w.h.p, applying $s_{t-1} \geq (\ln n)^4$. (Due to space limitations, we dot not explain the mathematical details behind the application of Chernoff bounds in this extended abstract.) Unfortunately, some of these messages are *wasted* as they are directed to the same player or an informed player. As interconnections are chosen at random, the probability that a particular message is wasted is at most $s_{t-1}/n + m/n$. This expression is bounded above by $(3 + o(1/\ln n))/\ln n$ because $s_{t-1} \leq n/\ln n$. As a consequence,

$$\mathbf{E}[s_t] = s_{t-1} + m\left(1 - \frac{3 + o(1/\ln n)}{\ln n}\right) = s_{t-1}\left(3 - O(1/\ln n)\right) \quad .$$

Applying a Chernoff bound yields

$$s_t = (1 \pm o(1/\ln n))\mathbf{E}[s_t] = s_{t-1}\left(3 \pm O(1/\ln n)\right) \quad ,$$

since $\mathbf{E}[s_t] \geq (\ln n)^4$. Assuming this expansion factor in each round, we can observe that this phase takes $\log_3 n \pm O(\ln\ln n)$ rounds.

3. The *quadratic-shrinking phase* ends with the round after which execution there are at most $\sqrt{n}(\ln n)^4$ uninformed players for the last time. Even if we only take into account pull transmissions we obtain (by following the arguments explaining the general properties of pull algorithms) that

$$E\left(\frac{u_t}{n}\right) \leq \left(\frac{u_{t-1}}{n}\right)^2 .$$

Applying a Chernoff bound yields

$$u_t \leq \left(1 + \frac{1}{\ln n}\right)\frac{(u_{t-1})^2}{n} \quad ,$$

w.h.p., provided $u_t \geq \sqrt{n}(\ln n)^4$. Now some easy calculations show that we need $O(\ln\ln n)$ rounds until the number of uninformed players drops from $n/\ln n$ to $\sqrt{n}(\ln n)^4$.

4. In the *final phase* we inform the few remaining uninformed players. Since the number of uninformed players in this phase is guaranteed to be smaller than $\sqrt{n}(\ln n)^4$, each player has probability at least $(\ln n)^4/\sqrt{n}$ to receive a rumor due to a pull transmission in each round of this phase. Consequently, we need only a constant number of rounds until all players are informed, w.h.p.

The exponential-growth phase takes $\log_3 n \pm O(\ln n)$ rounds. During this phase the number of transmissions grows exponentially from round to round. Therefore, we send only $O(n)$ messages during this phase. All other phases have length only $O(\ln\ln n)$. Thus, even if we assume $2n$ transmissions in each of these rounds, the total number of transmissions is only $O(n\ln\ln n)$. This completes the proof of Theorem 2.1.

□

2.2 The median-counter algorithm

The push&pull-algorithm heavily relies on a very exact estimation of the expansion of the set of informed players. The algorithm has to be executed exactly $\log_3 n + \Theta(\ln \ln n)$ rounds. For example, a constant fraction of players remains uninformed if the algorithm terminates after $(1-\epsilon)\log_3 n$ rounds, and the algorithm uses $\Theta(n \ln n)$ transmissions when terminating after $(1+\epsilon)\log_3 n$ rounds, for any constant $\epsilon > 0$. A robust algorithm requires a more flexible, distributed termination mechanism that recognizes when all players are informed. This termination mechanism is described in the following.

Median-Counter Algorithm

Let r denote the considered rumor. During the course of the algorithm each player v can be in one out of four states A, B, C, or D (with respect to the considered rumor r). State A means the player has not yet received the rumor. In all other states, the player knows the rumor. When a player is in one of the states B or C it pushes and pulls the rumor r along every established connection. In state D the player does not propagate the rumor anymore. Each player in state B holds a counter $\texttt{ctr}(v,r)$. We say a player v is in state B-m if $\texttt{ctr}(v,r) = m$. These counters are irrelevant in other states. The transitions between different states are defined as follows.

- State A: The player v does not know r. (For the purpose of analysis, we assume that $\texttt{ctr}(v,r) = 0$ in this state.) If a player v in state A receives r from a player in state B then it switches to state B-1. If a player in state A receives r from a player in state C then it switches to state C.
- State B-m: The player v knows r and $\texttt{ctr}(v,r) = m$. (The player injecting the rumor starts in state B-1.)

  Median rule: If during a round a player v in state B-m receives r from more players in state B-m' with $m' \geq m$ than from players in state A and B-m'' with $m'' < m$ then it switches to state B-$(m+1)$, i.e., increases its counter.

  There is one exception to this rule. If $\texttt{ctr}(v,r)$ is increased to $\texttt{ctr}_{\max}$ (where $\texttt{ctr}_{\max} = O(\ln \ln n)$ is a suitable integer) then v switches to state C. Furthermore, if a player in state B receives the rumor from a player in state C then it switches to state C, too.
- State C: Every player stays in this phase for at most $O(\ln \ln n)$ rounds, and then switches to state D, i.e., it terminates the rumor spreading.

Roughly speaking, the counters in state B are used in order to determine the point of time when the algorithm switches from the exponential-growth phase into the quadratic-shrinking phase. A counter value of $\texttt{ctr}_{\max}$ indicates that $n/\text{polylog}(n)$ players are informed so that it is sufficient to continue the propagation for only $O(\ln \ln n)$ rounds (which is done in state C). In order to make sure that the median-counter algorithm terminates even in case of the very unlikely event that the counter mechanism fails, we determine that every player stops propagating the rumor after some fixed number of $O(\ln n)$ rounds, regardless of its current state.

We investigate the robustness of the median-counter algorithm against different sources of errors and inaccuracies.

- First, we assume the random connections in each round are established using an arbitrary (possibly non-uniform) probability distribution $\mathcal{D} : V \to [0,1]$.

- Second, we assume that an oblivious adversary can specify up to F node failures occurring during the execution of the algorithm. The adversary specifies a set $\mathcal{F}$ of players (not containing the player starting the rumor) that fail to exchange information in some of the rounds (as specified by the adversary). We assume $|\mathcal{F}| \leq F$ and $n \sum_{v \in \mathcal{F}} \mathcal{D}(v) \leq F$.

Clearly, we cannot hope to inform all players when allowing adversarial node failures. Therefore, we are satisfied if the algorithm informs all but $O(F)$ players. (Alternatively, one may assume stochastic rather than adversarial failures, e.g., each random phone call fails with probability F/n. In this case, staying for $\tau = \Theta(\ln \ln n + \ln_{n/F} F)$ rounds in stage C ensures that all players are informed within $O(\ln n + \tau)$ rounds using $O(\tau n)$ transmissions, w.h.p.)

Theorem 2.2 *Assuming an arbitrary distribution $\mathcal{D}$ and up to F node failures as described above, the median-counter algorithms informs all but $O(F)$ players in $O(\ln n)$ rounds using $O(n \ln \ln n)$ transmissions, w.h.p.*

Due to space limitations we defer the proof of this theorem to the appendix.

3 Lower Bounds

3.1 Lower bound for address-independent algorithms

Our first lower bound shows that the push&pull scheme achieves optimal results for the class of address-independent algorithms. In particular, we show that any address-independent algorithm requires $\Omega(n \ln \ln n)$ transmissions in order to inform all players. Observe that this lower bound holds regardless of the number of rounds taken to inform all players. We assume the random phone call model using the uniform distribution.

Theorem 3.1 *Any address-independent rumor spreading algorithm guaranteeing that "all but a fraction f of the players receive the rumor with constant probability" needs to perform $\Omega(n \ln \ln f)$ transmissions on expectation.*

Proof. Fix an address-independent algorithm $\mathcal{A}$. Depending on the execution of $\mathcal{A}$, we partition the rounds into contiguous phases such that the total number of transmissions in the phases $1, \ldots, i$ is $(i-1)n/4 = \Omega(in)$. Let U_i denote the number of uninformed players at the end of phase i, and define $u(i) = n \exp(-2^i + \frac{3}{2})$. We will show by induction that $U_i \geq u(i)$, w.h.p. Consequently, $\mathcal{A}$ needs $\Omega(\ln \ln f)$ phases and, hence, $\Omega(n \ln \ln f)$ transmissions in order to inform all but a fraction f of the players, which yields the Theorem.

Phases are defined as follows. Phase 1 starts with the round in which the rumor is generated. If phase i ends in round t then phase $i+1$ starts in round $t+1$. We distinguish sparse and dense phases. A *sparse phase* contains at most $n/2$ transmissions. The length of these phases is maximized, that is, a sparse phase ends in round t if adding round $t+1$ to the phase would result in more than $n/2$ transmissions. A *dense phase* consists of only one round containing more than $n/2$ transmissions. Observe that the number of transmissions during the phases 0 to i is at least $(i-1)n/4$ because any pair of consecutive phases contains at least $n/2$ transmissions by construction.

Now assume by induction that the number of uninformed players at the beginning of phase i is at least $u(i-1)$. We have to show that the number of uninformed players at the end of phase i is at most $u(i)$, w.h.p.

For $1 \leq k \leq u(i-1)$, let x_k denote a 0-1 random variable indicating whether the kth of those players that are uninformed at the beginning of round i receives a message containing the rumor during the round. We claim

$$\mathbf{Pr}\left[x_k = 0\right] \geq \frac{u(i-1)}{\mathrm{e}n} \ .$$

The arguments leading to this inequality are different for sparse and dense rounds.

- Suppose phase i is sparse. Then $\mathcal{A}$ sends at most $\frac{n}{2}$ messages. Each of these messages is initiated without knowing the receiver because decisions are placed address-independently. As connections are chosen uniformly at random, the probability that a particular message reaches a particular player is $\frac{1}{n}$. Consequently, $\mathbf{Pr}\left[x_k = 1\right] \leq \frac{n}{2} \cdot \frac{1}{n} \leq \frac{1}{2}$ so that $\mathbf{Pr}\left[x_k = 0\right] \geq \frac{1}{2} \geq \frac{u(i-1)}{\mathrm{e}n}$.
- Now suppose phase i is dense. Then the phase consists of only one round. In this case, the probability that a player does not call an informed player is lower bounded by $\frac{u(i-1)}{n}$. Furthermore, the probability that a player is not called by any other player is at least $\frac{1}{\mathrm{e}}$. Thus, the probability that a player is not connected to an informed player is at least $\frac{u(i-1)}{\mathrm{e}n}$. Clearly, this implies $\mathbf{Pr}\left[x_k = 0\right] \geq \frac{u(i-1)}{\mathrm{e}n}$.

Since $u(i) = \sum_{k=1}^{u(i-1)} (1 - x_k)$, we obtain

$$\mathbf{E}\left[u(i)\right] = \sum_{k=1}^{u(i-1)} \mathbf{Pr}\left[x_k = 0\right] \geq \frac{u(i-1)^2}{\mathrm{e}n} \geq \frac{(n \exp(-2^{i-1} + \frac{3}{2}))^2}{\mathrm{e}n} = n \exp(-2^i + 2) \ = \ \sqrt{\mathrm{e}} u(i) \ .$$

Observe, that the random variables x_k are slightly dependent since the random interconnections used for transmissions in phase i form partial permutations on the caller site. This dependence, however, is negative [3] so that we can apply a Chernoff bound. Assuming $u(i) \geq (\ln n)^2$, we obtain

$$\mathbf{Pr}\left[U_i \geq u(i)\right] \leq \exp\left(\frac{(\sqrt{\mathrm{e}} - 1)^2}{2} u(i)\right) = O(n^{-\alpha}) \ ,$$

for any positive constant α. This completes the proof of Theorem 3.1. □

3.2 Lower bound for general algorithms

The above lower bound for address-independent algorithms does not hold for those distributed algorithms that place their decisions based on the addresses of their communication partners. In the introduction, we give an example showing how all players can be informed in $\Theta(n \ln n)$ rounds using only $O(n)$ of transmissions. Now we investigate whether there is an algorithm that is both time-optimal (i.e., using only $O(\log n)$ rounds) and communication-optimal (i.e., using only $O(n)$ transmissions) The following lower bound answers this question negatively. Again, we assume the random phone call model using the uniform distribution.

Theorem 3.2 *Any distributed rumor spreading algorithm guaranteeing that "all but a fraction $o(1)$ of the players receive the rumor within $O(\ln n)$ rounds with constant probability" needs to perform $\omega(n)$ transmissions on expectation.*

Proof. The difficulty in analyzing arbitrary distributed rumor spreading algorithms is that the distribution of the rumor can be a highly dependent process although the underlying random

calling mechanism is generated by n independent experiments in each round. For example, if player 1 is the only player with an odd address sending the rumor to players with even addresses then the success of the algorithm is highly dependent on the event that player 1 receives the rumor. This small example (not even involving any additional communication) shows that the analysis needs more than simply applying martingales or Chernoff bounds.

Our basic trick in the following analysis is that we choose a random sample of the players that can be guaranteed to act independently during the execution. This independence, can be guaranteed only for $T = \lfloor \frac{1}{8} \log n \rfloor$ rounds. Of course, this number of rounds is not enough to inform all players about a rumor initiated by a single player. Therefore, we assume that the rumor is spread already to at least half of the player and we consider the next T rounds.

Let $U_V \leq n/2$ denote the number of initially uninformed players. (In order to be able to extend our result to more than $\lfloor \frac{1}{8} \log n \rfloor$ rounds, we assume that the initially uninformed players are known by all players in the system. For example, assume the players $1, \ldots, U_V$ are these players.) Let X_V denote a random variable describing the number of messages sent during the T rounds. Furthermore, let U'_V denote a random variable describing the number of uninformed players after round T. (These random variables are with respect to the random phone calls.)

Let A denote a set of $m = \lfloor n^{1/8} \rfloor$ players chosen randomly from V. The set A will be our random sample. Let U_A denote the random variable describing the number of initially uninformed players in A (with respect to the random choice of A.) Let X_A denote a random variable describing the number of messages received by the players in A, and let U'_A denote the random variable describing the number of uninformed players in the set A after the last round. (These random variables are with respect to the random choice of A and the random phone calls.)

The communication graph G_t in round t is obtained by a distributed random process, i.e., each player v chooses a player u from V at random and v calls u. This random process generates a probability distribution $\mathcal{D}$ on the set $\mathcal{G}$ of possible communication graphs. Repeating this random process for T rounds extends the probability distribution $\mathcal{D}$ to $\mathcal{G}^T$.

For the analysis, we assume a slightly different probability distribution $\mathcal{D}'$ on $\mathcal{G}$. Instead of letting each player call a random other player, we establish the connections as follows. In each round t,

- we choose uniformly at random a collection of m disjoint subsets $B_t(v)$ $(v \in A)$, each containing m players from $V \setminus A$; (once these sets are chosen, the players in A can act fully independently)
- each player $v \in A$, chooses at random an integer $\delta(v) \geq 0$ with $\mathbf{Pr}[\delta(v) = i] = \dfrac{1}{\mathrm{e} i!}$; if $\delta(v) \geq m$, we set $\delta(v) = m - 1$;
- each player $v \in A$, chooses i.u.r. a set of $\delta(v) + 1$ different players $u_0(v), \ldots, u_{\delta(v)}(v)$ from $B_t(v)$.

We determine that every player $v \in A$ calls player $u_0(v)$, and the players $u_1(v), \ldots, u_{\delta(v)}(v)$ call v. Every player for which we have not yet specified whom to call simply chooses a communication partner from $V \setminus A$ i.u.r. Clearly, $\mathcal{D}$ and $\mathcal{D}'$ are different distributions. The following lemma, however, shows that these distributions are closely related.

Lemma 3.3 *The total variation distance between $\mathcal{D}$ and $\mathcal{D}'$ on $\mathcal{G}^T$ is $O(n^{-1/4})$.*

Based on this bound, we are able to give the following lemma comparing the behavior of the complete system $V|\mathcal{D}$ with that of the small system $A|\mathcal{D}'$.

Lemma 3.4 *For $\beta \geq 0$, $u \geq n^{-1/16}$, $0 \leq p \leq 1$,*

a) $\mathbf{E}[X_V|\mathcal{D}] \leq \beta n \Rightarrow \mathbf{Pr}\left[X_A > \frac{\beta m}{p}|\mathcal{D}'\right] = p + O(n^{-1/4})$,

b) $U_V \geq un \Rightarrow \mathbf{Pr}\left[U_A < \frac{um}{2}\right] = O(n^{-1})$, *and*

c) $\mathbf{Pr}[U'_A \geq um|\mathcal{D}'] < p \Rightarrow \mathbf{Pr}\left[U'_V < \frac{un}{2}|\mathcal{D}\right] = p + O(n^{-1/4})$.

Informally, this lemma states that it is sufficient to analyze $A|\mathcal{D}'$ in order to estimate $V|\mathcal{D}$. In fact, restricting to the smaller system $A|\mathcal{D}'$ enables us to deal with the dependencies. The following lemma summarizes our analysis for $A|\mathcal{D}'$.

Lemma 3.5 *Suppose $U_A \leq m/\alpha$ and $X_A \leq \beta m$ with $\alpha \geq 4$ and $\beta \geq 1$. Let c denote a suitable constant. Then*

$$\max\{U'_A, mn^{-1/16}\} \geq m\alpha^{-\exp(c\alpha\beta)} \ ,$$

with probability $1 - O(n^{-1/4})$.

(Lemma 3.3 and 3.4 require only applying standard methods from probability theory like the comparison of distributions and applying Chernoff bounds, the Markov inequality, etc. We omit these proofs due to space limitations. The proof of Lemma 3.5 is moved to the appendix. It shows how the highly dependent probabilistic rumor spreading process can finally be reduced to a simple, deterministic token game.)

Combining Lemma 3.4 and 3.5, we obtain the following result for $V|\mathcal{D}$. Suppose $U_V \leq n/\alpha$ and $\mathbf{E}[X_V \leq \beta n]$ with $2 \leq \alpha \leq n^{1/16}$ and $\beta \geq 0$. Applying Lemma 3.4 a) and b) yields

$$X_A \leq \kappa\beta \qquad \text{and} \qquad U_A \geq \frac{m}{2\alpha} \ .$$

with probability at least $1 - \frac{1}{\kappa} - O(n^{-1/4})$. Now applying Lemma 3.5 yields

$$U'_A \ \geq \ m\alpha^{-\exp(c\alpha(\kappa\beta+1))} \ ,$$

with probability $1 - \frac{1}{\kappa} - O(n^{-1/4})$. Finally, we can conclude from Lemma 3.4 c) that

$$U'_V \ \geq \ \frac{n}{2}\alpha^{-\exp(c\alpha(\kappa\beta+1))} \ , \tag{1}$$

with probability $1 - \frac{1}{\kappa} - O(n^{-1/4})$. Observe that this probability is lowerbounded by $1 - \frac{2}{\kappa}$, provided that n is sufficiently large.

In other words, for any constants $\alpha \geq 2$ and $\beta \geq 0$, starting with n/α uninformed players (possibly known by all players), performing $X_v \leq \beta n$ transmissions reduces the number of uninformed players only by some constant factor over $\frac{1}{8}\log\log n$ rounds, with probability at least $1 - \frac{2}{\kappa}$. Now suppose we execute a constant number of c phases of length $\frac{1}{8}\log\log n$. Then spending $O(n)$ transmissions in c phases reduces the number of uninformed players by only a constant factor, too. This result holds with probability $1 - \frac{2c}{\kappa}$. (Recall that κ can be chosen to be an arbitrary constant.) Consequently, performing $O(n)$ transmissions over $O(\ln n)$ rounds leaves a constant fraction of the players uninformed. (A rigorous analysis based on inequality 1 shows that informing all but a fraction $f(n)$ of the players with constant probability requires $\mathbf{E}[X_V] = \Omega(\ln^{[2k]} f(n))$, where $\ln^{[x]}$ denotes the natural logarithm iterated for x times.) Hence, Theorem 3.2 is shown. □

References

[1] D. Agrawal, A. El. Abbadi, R. C. Steinke. Epidemic Algorithms in Replicated Databases. In *Proceedings of the sixteenth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems*, pages 161-172, 1997.

[2] 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 *Proceedings of the 6th ACM Symposium on Principles of Distributed Computing*, pages 1-12, 1987.

[3] D. Dubhashi, D. Ranjan. Balls and Bins: A Study in Negative Dependence. In *Random Structures & Algorithms*, 13(2):99-124, 1998.

[4] L. Gasieniec, A. Pelc, Adaptive broadcasting with faulty nodes, In *Parallel Comput. 22 (1996) 903-912., 1996.*

[5] S. Hedetniemi, S. Hedetniemi, and A. Liestman. A Survey of Gossiping and Broadcasting in Communication Networks. Networks 18, 1988, 319-349.

[6] J. Hromkovic, R. Klasing, B. Monien, and R. Peine. Dissemination of Information in Interconnection Networks (Broadcasting & Gossiping). In *Combinatorial Network Theory*, pp. 125–212, D.-Z. Du and D.F. Hsu (Eds.), Kluwer Academic Publishers, Netherlands, 1996.

[7] R. Golding, D. Long. Accessing Replicated Data in a Large-Scale Distributed System. UCSC-CRL-91-01, Santa Cruz CA, 1991.

[8] R. Guy, G. Popek, T. Page, Jr. Consistency Algorithms for Optimistic Replication. In *Proceedings of the First International Conference on Network Protocols*. IEEE, October 1993.

[9] R. Ladin, B. Liskov, L. Shrira, S. Ghemawat. Providing high availability using lazy replication. In *ACM Transaction on Computer Systems*, 10(4):360, November 1992.

[10] M. Rabinovich, N. Gehani, A. Kononov. Scalable update propagation in epidemic replicated databases. AT&T Bell Labs Technical Memorandum 112580951213-11TM, 1995.

[11] T. Leighton, B. Maggs, R. Sitamaran. On the fault tolerance of some popular bounded-degree networks. In *Proceedings of the 33rd Annual Symposium on Foundations of Computer Science*, pages 542-552, October 1992.

A Proof of Theorem 2.2

First we investigate the errorless case. The median-counter-algorithm spreads the rumor in similar phases as the push&pull-algorithm. Let w_i be the probability that a player calls player i, let $S_t, s_t, U_t,$ and u_t be defined as above and let g_t be the weight of all informed players: $g_t := \sum_{i \in S_t} w_i$.

1. *Startup*-phase. We want to ensure at least $s_t \in \Omega(\log n)$ informed players with weight $g_t \geq \frac{\log n}{n}$ are established. First, we concentrate on some $\Theta(\log \log n)$ rounds of push communication. If only players with weight smaller than $\frac{\log n}{n}$ are informed, then after $O(\log \log n)$ rounds $c \log n$ nodes are informed which will push their rumors to players with a total weight of at least $\frac{\log n}{n}$ after some constant rounds.

   Consider now the case that the weight after these rounds exceeds $g_t \geq \frac{\log n}{n}$, but the number of informed players is small $s_t < c \log n$. The probability that an uninformed player calls an informed one is g_t. The expected number of informed players is therefore: $E(s_{t+1}) \geq s_t + (n - s_t)g_t \geq s_t(2 - \frac{1}{\log n})$. Applying Chernoff bounds it follows w.h.p. $s_{t+1} \geq s_t(2 - \epsilon)$ for some arbitrary small $\epsilon > 0$.

   So, the startup phase lasts at most $O(\log \log n)$ rounds.

2. *Exponential growth*: This phase ends, when $g_t \geq \frac{1}{\log n}$.

   For this phase the weight h_t of all uninformed players H_t with larger weight than $\frac{1}{s_t}$ is of special interest: $h_t := \sum_{i \in U_t : w_i \geq 1/s_t} w_i$. Note that $|H_t| \leq s_t$ and that the probability of a member of H_t being called by an informed player in S_t is larger than the constant $1 - 1/e$. Therefore, push-operations cause an increase of the weight of informed players $g_{t+1} \geq g_t + (1 - \epsilon)(1 - 1/e)h_t$ for some constant $\epsilon > 0$ w.h.p.

   In $U_t \setminus H_t$ there is a constant fraction of at least $1/e - \epsilon$ of players which only get one call in a round for an arbitrary small constant $\epsilon > 0$ w.h.p. The probability that one of these players gets the rumor pushed from S_t is $\frac{s_t}{n}$. The expected number of informed players in the next round is therefore $E(s_{t+1}) \geq s_t + \frac{s_t}{n}(1/e - \epsilon)(n - s_t - |H_t|) \geq s_t(1 + (1/e - \epsilon)(1 - \frac{2s_t}{n}))$.

   If $s_t \leq \frac{n}{\log n}$ for $h_t \leq \frac{1}{2}$ this implies $s_{t+1} \geq s_t(1 + \frac{1}{e} - \epsilon')$ and in the other case $g_{t+1} \geq g_t(\frac{3}{2} - \frac{1}{2e} - \epsilon')$ for some arbitrary small $\epsilon' > 0$.

   So after some $O(\log n)$ rounds it holds either $g_t \geq \frac{n}{\log n}$ or $s_t \geq \frac{2n}{\log n}$. In the second case every player with weight larger than $\frac{c \log^2 n}{n}$ is informed in the next round w.h.p. Furthermore, the expected weight of all informed players is $E(g_{t+1}) \geq \sum_{i=1}^{n} w_i^2 s_t$. It turns out that this sum is minimal for the uniform probability distribution. Hence, $E(g_{t+1}) \geq\geq \frac{s_t}{n}$. Because the weights are upperbounded we can apply Chernoff bounds and get $g_{t+1} \geq \frac{s_t}{2n} \geq \frac{1}{\log n}$.

   Note for the number of messages that in all but one rounds $s_t \leq \frac{2n}{\log n}$. Therefore, the number of messages is bounded by $O(n)$.

   Now we discuss how often a counter of a player will be increased during this phase. We consider a player i with weight w_i who is informed during this phase.

   (a) $w_i \geq \frac{3 \log n}{n}$

   In every round at least $2 \log n$ uninformed call i, while i receives a call only from at most $\log n$ informed players ($s_t \leq \frac{2n}{\log n}$). i's push call can be neglected. So, this player will communicate with more uninformed than informed players in each round and the median rule prevents an incrementation of i's counter.

(b) $w_i \leq \frac{3\log n}{n}$

We allow that during the time interval $t \in \{a, \ldots, b\}$ for which it holds $\frac{1}{\log^2 n} \leq w_i s_t \leq c \log n$ the counter of P_i is increased in every round.

In every round g_t or s_t grows by a factor $\alpha > 1$, but possibly not both of them. Nevertheless they interact pairwise, since the expected number of uninformed nodes informed by a pull is $u_t g_t$. Therefore it holds $s_{t+1} \geq (1-\epsilon) u_t g_t \geq n g_t (1-\epsilon')$ for $\epsilon, \epsilon' > 0$ with high probability. On the other hand, every informed nodes pushes in every round it holds $g_{t+1} \geq \frac{s_t}{2n \log n}$ w.h.p. So, this time interval is bounded by $O(\log \log n)$.

For any time step after b the number of uninformed players calling P_i is higher than those of the informed players for the same reasons as in (a).

For every round t before a we concentrate on weights w_i with $w_i \leq \frac{1}{s_t \log^2 n}$. The probability that a player with such a weight is called by an informed player is smaller than $1 - (1 - \frac{1}{s_t \log^2 n})^{s_t} \leq \frac{1}{\log^2 n}$. Let q_i be the number of players which at least i-times increase their counter before point a and let $q_0 = s_t$. In the worst case all players stay in this situation for the whole phase. Only q_i players can cause an increase for a counter larger than i. The probability that such a player calls another is $\frac{q_i}{s_t \log^2 n}$. Therefore, it holds $E(q_{i+1}) \leq \frac{q_i^2}{s_t \log^2 n}$. It follows $\frac{q_{i+1}}{s_t} \leq c \frac{q_i^2}{s_t^2 \log^2 n}$ if $q_i \in \Omega(\log n)$; and if $q_i \leq O(\log n)$, then $q_{i+c'} = 0$ for some constants c, c' w.h.p. This proves $q_{O(\log \log n)} = 0$. So, there are no players whose counters will be increased more than some $c \log \log n$ time during this phase.

3. *Quadratic-shrinking*: This phase ends, when all players have left states A or B.

The probability for each uninformed player to remain uninformed is at most $1 - g_t$, if we consider only pull-communication. Therefore it holds $E(u_{t+1}) \leq u_t(1 - g_t)$, which implies $u_{t+1} \leq u_t(1 - g_t)(1 + \frac{\log n}{\sqrt{n}})$ w.h.p. The expected weight of the uninformed player of the next round is $E(1 - g_{t+1}) = (1 - g_t)^2$. Note that $\max_{i \in U_t} w_i \leq \frac{c \log^2 n}{n}$. Therefore, applying Chernoff bounds it follows that $1 - g_{t+1} \leq (1 - g_t)^2 (1 + \frac{\log^2 n}{\sqrt{n}})$ w.h.p. It is clear that after some $O(\log \log n)$ rounds it holds $1 - g_{t+1} \leq \frac{2 \log^2 n}{\sqrt{n}}$. Then, some constant rounds of pull will sufficiently decrease the probability of an uninformed player remaining in state A.

Since in every round each counter may be incremented only once, it suffices to choose $\mathtt{ctr}_{\max} \geq c \log \log n$ for some constant c independent from $\mathcal{D}$.

It remains to show that after some additional $O(\log \log n)$ rounds all counters reach $\mathtt{ctr}_{\max}$. Consider the time point at which all players are informed. Clearly, all counters are at least 1. Then, in every step i each counter is at least $i + 1$. Therefore the distributional algorithm ends after $O(\log \log n)$ rounds.

Since every player produces only one random call in each round the overall number of messages in this phase is bound by $O(n \log \log n)$.

Now we focus on the case of $F \leq \frac{1}{4} n$ node failures with weight F/n. We assume, that if a node failure occurs on v that v terminates, i.e. switches to D without learning the rumor. The analysis of the startup phase and exponential can be easily adapted to this case, since the growth of informed nodes proceeds slower but even though exponential. We now investigate the situation in the double exponential shrinkage phase.

Let $\mathcal{F}$ be the set of nodes, which may be disconnected in some rounds. Then S_t and U_t are defined as the set of informed and uninformed nodes, being disjunct with $\mathcal{F}$; u_t, s_t, and g_t are

defined as before. The probability that a node remains uninformed is at most $1 - g_t$ per round. Therefore we can conclude w.h.p. $u_{t+1} \leq (1 - g_t)u_t$. Similary as in the error-free case we can conclude that $1 - g_{t+1} \leq \frac{F}{n} + (1 + \frac{\log^2 n}{n})(1 - g_t)^2$ w.h.p. This recursion converges in $O(\log\log n)$ rounds to $1 - g_{t'} \in O(\frac{F}{n})$. This implies a maximum number of $O(F)$ uninformed nodes within the next round.

The main problem for the error case is to verify that the number of messages does not exceed $O(n \log\log n)$. We prove this by showing that at least $O(n/\log n)$ players reach state C or D, when the first error-free players reach state D. The remaining error-free players can only cause $O(\log n)$ messages each, where faulty F players do not add further messages. We start our analysis at the moment when only $F' \in O(F)$ nodes with weight F'/n remained uninformed. Let us assume that all informed players are in the state B-1.

Let $Z_{z,m}$ be the set and $y_{t,m}$ the weight of error-free nodes in round t with $\mathtt{ctr}(v) = m$. The probability that a node in $Z_{t,m}$ is increased is at least $\sum_{i=m}^{\mathtt{ctr}_{\max}} y_{t,i}$. We want to prove that in the triangular section where $t \leq km$ for some constant k, $y_{t,m}$ decreases exponentially in t. For the analysis we allow that some of the counter may be decreased. The aim of this modification is that the series $y_{t,1}, \ldots, y_{t,m_t}$ is exponentially increasing, the series $y_{m_t,t}, y_{m_t+1,t}, \ldots$ is exponentially decreasing, and the weight $y_{t,m_t+1} \geq \frac{1}{2}$ contains the rest of the weight. More formally, $\forall i \leq m_t \ : \ y_{t,i} \leq \alpha y_{t,i+1}$ and $y_{t,m_t+1} = 1 - F'/n - \sum_{i=0}^{m_t} y_{t,i}$ for some $\alpha > 1$.

By decreasing some of the counters it can be ensured that in the next round it holds $\forall i \leq m_t \ : \ y_{t,i} \leq \alpha \ y_{t,i+1}$ and $y_{t+1,i} \leq \frac{1+\alpha}{2} \ y_{t,i}$. This follows by the fact that $\sum_{i=j}^{m_t} y_{t,j} \geq \frac{1}{2}$ and by reducing the number of players increasing their counter to a fraction of $\frac{1}{2}$ each. After some constant rounds c it holds $y_{t+c,m_t+1} \geq \alpha y_{t+c,m_t}$. Then, we increase $m_{t+c} := m_t + 1$ and get the claimed triangular section.

Therefore, after some $O(\log\log n)$ rounds only a fraction of $O(n/\log n)$ has a smaller counter than $c \log\log n$.

B Proof of Lemma 3.5

We consider the execution of $T = \lfloor \frac{1}{8} \log n \rfloor$ rounds of a distributed algorithm assuming that the communication partners in each round are selected according to distribution $\mathcal{D}'$. Let $u_0 \leq \frac{1}{4}$ denote the fraction of initially uninformed players in A. Let u_t, for $1 \leq t \leq T$, denote a random variable describing the fraction of uninformed players after the execution of round t. We assume that X_A, the number of messages received by the players in A, is upperbounded by βm with $\beta \geq 1$. We have to show that

$$\max\{u_T, n^{-1/16}\} \geq u_0^{\exp(cu_0\beta)} \ , \tag{2}$$

with probability $1 - O(n^{-1/4})$, for some suitable constant $c > 0$.

We want to make use of the fact that the distributed algorithm has only local knowledge about the random communication graph. The following lemma shows that this knowledge is actually very limited. For a set $Y \subseteq V$, set $\Gamma_0(Y) = Y$, and define $\Gamma_t(Y)$ to be the set of all players connected to $\Gamma_{t-1}(Y)$ in round t plus the players in $\Gamma_{t-1}(Y)$ itself. Observe that, in round t, only the players in $\Gamma_t(Y)$ can receive information from the players in $\Gamma_{t-1}(Y)$. In other words, none of the players in $V \setminus \Gamma_t(Y)$ received any information sent by the players in Y during the rounds 1 to $t - 1$.

Lemma B.1

$$\mathbf{Pr}\left[\exists v \in A, 1 \le t \le T : B_t(v) \cap \Gamma_{t-1}\left(A \cup \bigcup_{1 \le \tau < t} \bigcup_{w \in A} B_\tau(w)\right) \neq \emptyset\right] = O(n^{-1/4}) .$$

Proof. $B_t(v)$ is a random set of size m chosen from $V \setminus A$. The probability that one of the elements in this set is contained also in $\Gamma_{t-1}(\cdot)$ is bounded above by

$$\frac{m \cdot |\Gamma_{t-1}(\cdot)|}{n-m} \le \frac{mTm^2 4^T}{n} \le \frac{m^3 n^{1/4} \log n}{n-m} = O(n^{-1/4})$$

applying $|\Gamma_{t-1}(\cdot)| \le Tm^2 4^T$, $T \le \frac{1}{8}\log n$, and $m \le n^{1/8}$. The bound on $|\Gamma_{t-1}(\cdot)|$ can be obtained as follows. The size of $\Gamma_0(\cdot)$ is $(t-1)m^2 + m \le Tm^2$ taking into account the set A and $(t-1)m$ sets $B_\tau(w)$ each of which having size m. In expectation, the Γ set grows by a factor of at most 3 in each round, that is, $\mathbf{E}\left[|\Gamma_\tau(\cdot)|\right] \le 3|\Gamma_{\tau-1}(\cdot)|$. Applying Chernoff bounds, we obtain $|\Gamma_\tau(\cdot)| \le 4|\Gamma_{\tau-1}(\cdot)|$, with probability $1 - O(1/n)$. Thus, $\Gamma_{t-1}(\cdot) \le Tm^2 4^{t-1} \le Tm^2 4^T$, with probability $1 - O(T/n)$. □

In the following we will assume that the unlikely event described in Lemma B.1 actually does not occur, that is, the players in $B_t(v)$ do not have any information about the outcome of the random edges connecting $w \in A$ with $B_\tau(w)$, for $1 \le \tau < t$. Let us describe this in more detail. When the distributed algorithm decides to send a message to a player $v \in A$ in round t, then this decision is actually placed by a player u in $B_t(v)$. This player u might have gathered some knowledge about which players are in the sets $B_\tau(w)$, for $w \in A$ and $1 \le \tau < t$. (For example, u knows that it is not in one of these sets itself.) Under the assumption above, however, the player u has no knowledge about the outcome of the random edges between w and $B_\tau(w)$ with $(w, \tau) \in A \times \{1, \ldots T\} \setminus \{v, t\}$.

For $v \in A$ and $1 \le t \le T$, assume the set $B_t(v)$ are fixed. Furthermore, assume that the informed players in $B_t(v)$ are known. In order to get an unambiguous model, we replace the distributed algorithm by a specification $\mathcal{S}$ describing for which outcome of the random edges between v and $B_t(v)$ in round t the rumor should be sent to v, for every $v \in A$ and $1 \le t \le T$. The specification $\mathcal{S}$ is assumed to be *oblivious* with respect to the random interconnections, i.e., it must be given without knowing the outcome of the random edges between $v \in A$ and $B_t(v)$, for any $v \in A$ and $1 \le t \le T$.

Suppose we fix the $B_t(v)$ sets and those players in these sets that hold the rumor at the beginning of round t. Then we can transform the oblivious specification $\mathcal{S}$ into a table of independent probabilities $b_t(v)$ ($v \in A$, $1 \le t \le T$) so that $b_t(v)$ is the probability that v receives a message in round t. Observe that $b_t(v)$ cannot be larger than $b_t^{\max}(v)$, i.e., the probability that v is connected to a player holding the rumor in round t. The following lemma upperbounds $b_t^{\max}(v)$ based on the fact that the fraction of informed players in $B_t(v)$ does not deviate too much from $1 - u_{t-1}$ (i.e., the fraction of uninformed players in A). The given probability is with respect to the random choice of the $B_t(v)$ sets.

Lemma B.2 $\mathbf{Pr}\left[\exists v \in A, 1 \le t \le T : b_t^{\max}(v) > 1 - \frac{u_{t-1}}{2\mathrm{e}}\right] = O(n^{-1/4})$.

Proof. We make regress to the distribution $\mathcal{D}$. Observe that all results we show for $\mathcal{D}$ hold for $\mathcal{D}'$ with probability $1 - O(n^{-1/4})$ because of Lemma 3.3.

We need the following technical lemma which is a straightforward consequence of Chernoff bounds.

Lemma B.3 *Suppose V contains $un \geq n^{-15/16}$ uninformed players Let $Y \subseteq V$ denote a randomly chosen subset of size $m = \lfloor n^{-1/8} \rfloor$. Then the expected number of uninformed players in Y is $um \geq n^{-1/16}/2$. Furthermore, for any constant $\delta > 1$, the probability that Y contains less than um/δ or more than δum uninformed players is $O(n^{-1})$.*

Applying this lemma we can upper- and lowerbound the number of uninformed persons in any randomly selected set of players.

Assuming $\mathcal{D}$, A is a set of size m selected at random from V. Let u denote the fraction of uninformed players in V at the beginning of round t. Recall u_{t-1} specifies the same fraction for the set A. Applying B.3 yields $u_{t-1} \leq \sqrt{2}u$, with probability $1 - O(n^{-1})$.

The set $B_t(v)$ is chosen at random from $V \setminus A$. Alternatively, we can view $B_t(v)$ as chosen at random from V since A only "eliminates" some random players in V. Let u^* denote the fraction of uninformed players in $B_t(v)$ at the beginning of round t. Then applying B.3 yields $u* \geq \frac{u}{\sqrt{2}}$, with probability $1 - O(n^{-1})$.

Combining, the bounds for A and $B_t(v)$, we obtain $\mathbf{Pr}\left[u^* > \frac{u_{t-1}}{2} | \mathcal{D}\right] = O(n^{-1})$, which implies $\mathbf{Pr}\left[u^* > \frac{u_{t-1}}{2} | \mathcal{D}'\right] = O(n^{-1/4})$.

Now let us assume $u^* \geq \frac{u_t}{2}$ and calculate $b_t^{\max}(v)$ under this assumption. The probability that v has only one neighbor in round t is $\mathbf{Pr}\left[\delta(v) = 0\right] = \frac{1}{\mathrm{e}}$. The probability that this neighbor is uninformed is at least $u^* \geq \frac{u_{t-1}}{2}$. If both of these independent events occur then we cannot send the rumor. Therefore, $b_t^{\max}(v) \leq 1 - \frac{u_{t-1}}{2\mathrm{e}}$. □

The relationship between the variables and parameters u_t, $b_t(v)$, and β can be described by a set of four equations E1 to E4. From Lemma B.2, we can conclude that

$$\text{E1:} \quad 0 \;\leq\; b_t(v) \;\leq\; 1 - \frac{u_{t-1}}{2\mathrm{e}} \;,$$

with probability $1 - O(n^{-1/4})$. We introduce some auxiliary variables. Let $p_t(v)$ denote the probability that player $v \in A$ does not know the rumor at the end of round t. Set $p_0(v) = 1$ if $v \in A$ is initially uninformed, 0 otherwise. For every $v \in A$ and $1 \leq t \leq T$,

$$\text{E2:} \quad p_t(v) \;=\; (1 - b_t(v))p_{t-1}(v)$$

because v is uninformed in round t if and only if it is uninformed in round $t-1$ and does not receive a message in round t. Observe that these two events are independent.

Furthermore, the expected number of uninformed players in any round can be calculated easily by $\mathbf{E}\left[u_t m\right] = \sum_{v \in A} p_t(v)$. The probabilities $p_t(v)$ are independent as they are composed out of the independent probabilities $b_t(v)$. Thus, we can bound $u_T m$ using Chernoff bounds. Recall that we aim to give a lower bound on $\max\{u_T, n^{-1/16}\}$. Hence, we may assume w.l.o.g. that $u_t \geq n^{-1/16}$ so that $u_t m \geq m n^{-1/16} \geq \frac{1}{2} n^{1/16}$. Hence, applying a Chernoff bound yields

$$\text{E3:} \quad u_t m \;\geq\; \frac{1}{2} \sum_{v \in A} p_t(v) \;,$$

with probability $1 - O(n^{-1})$.

Finally, the expected number of messages received by the players in A during all rounds is $\mathbf{E}\left[X_A\right] = \sum_{t=1}^{T} \sum_{v \in A} b_t(v)$. By our initial assumptions $X_A \leq \beta m$ for $\beta > 1$. Thus, applying a Chernoff bound yields $\mathbf{E}\left[X_A\right] \leq \max\{2X_A, n^{-1/16}\} \leq 2\beta m$, with probability $1 - O(n^{-1})$. Consequently,

$$\text{E4:} \quad \sum_{t=1}^{T} \sum_{v \in A} b_t(v) \;\leq\; 2\beta m \;,$$

with probability $1 - O(n^{-1})$.

Lemma B.4 *Let u_T^* be an optimal solution to the optimization problem "minimize u_T under the constraints E1, E2, E3, E4" with parameters $u_0 \leq \frac{1}{4}$, $\beta \geq 1$ and variables $b_t(v)$, $p_t(v)$, and u_t ($1 \leq v \leq m$, $1 \leq t \leq T$). Then u_T^* satisfies inequality 2.*

Proof. First, we change some of the constraints. We redefine

$$\text{E1:} \quad b_t(v) \quad \in \quad \left\{0, 1 - \frac{u_{t-1}}{2\mathrm{e}}\right\} \quad .$$

Clearly, this restricts the set of allowed solutions, but some calculations show that this is compensated by relaxing E4 as follows.

$$\text{E4:} \quad \sum_{t=1}^{T} \sum_{v \in A} b_t(v) \quad \leq \quad 3\beta m \quad .$$

Next we rewrite the optimization problem by substituting

$$\beta_t(v) \quad = \quad b_t(v) \cdot \left(1 - \frac{u_{t-1}}{2\mathrm{e}}\right)^{-1} \quad .$$

Observe that $\beta_t(v) \leq \frac{4}{3} b_t(v)$. Therefore, we obtain the following relaxation of our optimization problem.

$$E1: \quad \beta_t(v) \in \{0, 1\}$$

$$E2: \quad p_t(v) = \left(1 - \beta_t(v) \left(1 - \frac{u_{t-1}}{2\mathrm{e}}\right)\right) p_{t-1}(v)$$

$$E3: \quad u_t m \geq \frac{1}{2} \sum_{v=1}^{m} p_t(v)$$

$$E4: \quad \sum_{t=1}^{T} \sum_{v \in A} \beta_t(v) \leq 4\beta m$$

Interpreting these constraints, we obtain the following token game. Let A_u denote the set of initially uninformed players in A.

- There are at most $4\beta m$ tokens which can be spent in T rounds.
  (Setting $\beta_t(v) = 1$ means spending a token for player v in round t.)
- In each round, each player $v \in A_u$ can receive at most one token.
- If player $v \in A_u$ receives a token in round t then $p_t(v) = \frac{u_{t-1}}{2\mathrm{e}} p_{t-1}(v)$,
  otherwise $p_{t-1}(v) = p_{t-1}(v)$.

The goal of this game is to minimize $u_T = \sum_{v=1}^{m} p_T(v)/(2m)$. Unfortunately, the optimal strategy for this game is not obvious. For example, the greedy strategy starting with giving a token to every player in the first round, then giving a token to every player in the second round, and so on ... does not lead to an optimal solution. In fact, the benefit of a token is maximized if the player receiving the token has been spared from previous tokens.

We can enforce an almost even distribution of the tokens, however, by spending some extra tokens. Suppose at least half of the nodes in A_u receive at least I tokens. Then

$$I \leq \frac{4\beta m}{u_0 m/2} = 8\beta \ . \tag{3}$$

Let t_i denote that round until whose completion at least half of the players in A_u received their ith token, for $1 \leq i \leq I$. Now we add the additional constraint that every player in A_u has to receive at least i tokens until the end of round t_i. Observe that we can easily satisfy this constraint by spending up to $|A_u|/2 = u_0 m/2$ free tokens in round t_i, for $1 \leq i \leq I$.

Next we analyze the number of uninformed nodes in A taking advantage of the additional constraint. Let $u(i)$ denote the fraction of uninformed players in A at the beginning of round t_i, i.e., $u(i) = u_{t_i - 1}$, for $1 \leq i \leq I$. Furthermore, set $u(I+1) = u_T$. Let $\rho(i, v)$ describe the effect of the ith token assigned to player $v \in A_u$, that is, if v receives its ith token in round t then $p_t(v) = \rho(v, t) \cdot p_{t-1}(v)$, for $1 \leq i \leq I$. As the ith token of every player $v \in A_u$ is spent before or in round t_i, we obtain by constraint E2 that

$$\rho(v, i) \geq \frac{u(i)}{2\mathrm{e}} \ , \tag{4}$$

for $1 \leq i \leq I$. Let $A_u(i)$ denote the set of those players receiving their ith token in round t_i, for $1 \leq i \leq I$. Let $A(I+1)$ denote the set of players receiving at most I tokens. Our construction ensures

$$|A_u(i)| > \frac{|A_u|}{2} \geq \frac{um}{2} \tag{5}$$

because less than half of the players in A_u receive their ith token before round t_i, for $1 \leq i \leq I$, and less than half of the players in A_u receive more than I tokens. Besides, as the players in $A_u(i)$ receive at most $i-1$ tokens during the rounds 1 to $t_i - 1$, we have

$$p_{t_i - 1}(v) \geq \prod_{j=1}^{i-1} \rho(v, j) \ , \tag{6}$$

for $v \in A_u(i), 1 \leq i \leq I+1$. Combining these inequalities yields

$$\begin{aligned}
u(i) &= u_{t_i - 1} \\
&\overset{(E3)}{\geq} \frac{1}{2m} \sum_{v \in A} p_{t_i - 1}(v) \\
&\overset{(6)}{\geq} \frac{1}{2m} \sum_{v \in A_u(i)} \prod_{j=1}^{i-1} \rho(v, j) \\
&\overset{(4)}{\geq} \frac{1}{2m} \sum_{v \in A_u(i)} \prod_{j=1}^{i-1} \frac{u(j)}{2\mathrm{e}} \\
&\overset{(5)}{>} \frac{u}{4} \prod_{j=1}^{i-1} \frac{u(j)}{2\mathrm{e}} \ ,
\end{aligned}$$

for $1 \leq i \leq I+1$. Furthermore, we have $u(1) > \frac{u}{2}$ because less than half of the um players in A_u received a token before round t_1. Consequently, solving the recurrence on $u(i)$ for $u_{(I+1)} = u_T$ yields

$$u_T > \frac{u^{2^I}}{2^{2^{I+1} + 2^{I-1} - 1} \mathrm{e}^{2^I - 1}} = u^{\exp(\kappa I)} \overset{(3)}{=} u^{\exp(6\kappa\beta)} \ ,$$

for some suitable constant $\kappa > 0$. $\square$

Summarizing, we have shown that the four equations E1 to E4 hold with probability $1 - O(n^{-1/4})$. Assuming these equations we have deduced inequality 2. Thus, this inequality holds with probability $1 - O(n^{-1/4})$, which completes the proof of Lemma 3.5.