Rumor Spreading and Conductance

FLAVIO CHIERICHETTI, Sapienza University of Rome
GEORGE GIAKKOUPIS, INRIA, Rennes
SILVIO LATTANZI, Google Research, New York
ALESSANDRO PANCONESI, Sapienza University of Rome

In this article, we study the completion time of the PUSH−PULL variant of rumor spreading, also known as randomized broadcast. We show that if a network has n nodes and conductance ϕ then, with high probability, PUSH−PULL will deliver the message to all nodes in the graph within $O(\log n/\phi)$ many communication rounds. This bound is best possible. We also give an alternative proof that the completion time of PUSH−PULL is bounded by a polynomial in $\log n/\phi$, based on graph sparsification. Although the resulting asymptotic bound is not optimal, this proof shows an interesting and, at the outset, unexpected connection between rumor spreading and graph sparsification. Finally, we show that if the degrees of the two endpoints of each edge in the network differ by at most a constant factor, then both PUSH and PULL alone attain the optimal completion time of $O(\log n/\phi)$, with high probability.

CCS Concepts: • **Mathematics of computing** → **Probabilistic algorithms**; • **Theory of computation** → **Graph algorithms analysis**; **Distributed algorithms**;

Additional Key Words and Phrases: Randomized broadcast, gossip algorithms, distributed algorithms, randomized algorithms, graph conductance, graph sparsification

1 INTRODUCTION

Rumor spreading, also known as randomized broadcast or randomized gossip, is a well-known distributed algorithm. Starting with one source node with a message, the algorithm proceeds in a sequence of synchronous rounds with the goal of broadcasting the message, i.e., delivering it to all nodes in the network. In each round, every node that knows the message selects a neighbor uniformly at random to which the message is forwarded. This is the so-called PUSH strategy. The PULL variant is symmetric. In each round, every node that does not yet have the message selects a

neighbor uniformly at random and asks for the information, which is then transferred to the node provided that the queried neighbor possesses it. Finally, the PUSH–PULL strategy is a combination of both: In each round, every node selects a random neighbor to perform a PUSH if it has the message, or a PULL if it does not have it.

Simplicity and effectiveness are among the many virtues of rumor spreading, making it an ideal candidate for quick dissemination of information inside a network. The algorithm has been extensively studied, since these three strategies were introduced by Demers et al. [12]. Naturally, one of the most studied questions concerns the completion time, i.e., how many rounds it takes to disseminate the information to all nodes in the graph, assuming a worst-case source.

Typically, this question has been approached on a case-by-case basis by giving careful estimates for particular classes of graphs. For instance, when the network is the complete graph we know that PUSH informs every vertex within $\log_2 n + \ln n + O(1)$ rounds almost surely [20, 33]. Other classes of graphs for which rumor spreading has been analyzed include hypercubes, constant degree graphs, and a large variety of random graphs (see Section 2 for more details).

Our approach here is quite different. Rather than focusing on special classes of graphs, we want to identify a set of necessary and/or sufficient conditions for rumor spreading to be fast in any given network satisfying such conditions, where by "fast" we mean that the completion time is poly-logarithmic in the size of the network with high probability. In this work, we make a step in this direction by providing a general sufficient condition for rumor spreading to be fast: *high conductance.*

The first result of this article is to show that if a network has n nodes and conductance ϕ then, with high probability, the completion time of PUSH–PULL is $O(\log n/\phi)$. This bound is optimal, for there are graphs of conductance ϕ and diameter $\Theta(\log n/\phi)$. The proof of this bound is based on a decoupling of the operations of PUSH and PULL. We show that, once the rumor reaches a vertex of maximum degree, every remaining vertex is informed within $O(\log n/\phi)$ many rounds with high probability by means of PULL alone. There is a clear symmetry between PUSH and PULL that can be exploited to show that the probability of spreading the message from a vertex u to another vertex v within t rounds via PULL is the same as that of spreading the message in the opposite direction, from v to u, within the same t rounds, via PUSH. Therefore, the argument above also shows that, starting from any source, a maximum degree vertex will be informed within $O(\log n/\phi)$ many rounds with high probability by using PUSH alone. Taken together, these two bounds for PUSH and PULL give the optimal $O(\log n/\phi)$ bound for PUSH–PULL.

The second result of this article is an alternative proof showing that the completion time of PUSH–PULL is bounded by a polynomial function in $\log n/\phi$. Unlike the previous approach, which is based on a decoupling of PUSH and PULL, this approach analyzes the PUSH–PULL process as a whole. It is based on what we consider to be an interesting and, at the outset, surprising connection between rumor spreading and the spectral sparsification theory. This connection is used to show that a small number of PUSH–PULL rounds (polynomial in $\log n/\phi$) expands the volume of the set of informed nodes by a factor of $(1 + \Theta(\phi^2))$, with high probability. In turn, this implies that with high probability, starting from any source, after a number of rounds of PUSH–PULL that is poly-logarithmic in n and polynomial in ϕ^{-1}, every node is informed. The resulting bound on the completion time is sub-optimal, but we believe that the proof remains interesting. In a broad sense, it draws a connection between the theory of spectral sparsification and the speed with which diffusion processes make progress in a network. This could potentially lead to further insights beyond the present work and is worth exploring.

The above bounds hold for neither PUSH, nor PULL, when taken in isolation. Perhaps the simplest counterexample is a star of n nodes. The star has maximum conductance, but PUSH takes at least $n - 1$ rounds to complete, regardless of the source, while a standard Coupon Collector's argument

shows that $\Theta(n \log n)$ many rounds are needed with high probability. Moreover, a basic expectation argument shows that when a leaf is the source, PULL requires n many rounds, on average. See also Reference [8] for examples of preferential attachment graphs with high conductance in which the two strategies are proven to be slow.

The final result of the article is to show a simple sufficient condition for the optimal $O(\log n/\phi)$ bound to be attained with high probability by PUSH and by PULL alone: For each edge in the graph, the degrees of its two endpoints differ by at most a constant factor.

One of the motivations for our work comes from social networks. Loosely stated, we are looking for a theorem of the form *"Rumor spreading is fast in social networks."* There is, however, no universally accepted definition of social networks. One way to circumvent this difficulty is to look for properties that are enjoyed by social networks and see whether they imply that rumor spreading is fast. The empirical evidence available suggests that social networks have high conductance [29]. Furthermore, this is true for preferential attachment models of social networks such as the one proposed by Mihail et al. [30]. Thus, our results can be considered as a possible explanation of why epidemic processes resembling rumor spreading spread virally in social networks.

We should note that the PUSH–PULL process has some sociological justification. PUSH corresponds to communicating in an active fashion, e.g., texting a message with a mobile phone or sending an email to a specific recipient. PULL on the other hand corresponds to, e.g., reading a tweet or a post on the wall of a Facebook account. Thus, PUSH–PULL seems to be a reasonable first approximation to model the spread of viral information in the age of the Internet.

The work presented here has been the first to show a connection between rumor spreading and graph expansion properties that, intuitively, should ensure fast dissemination.[1] This line of research has generated considerable follow-up work on the relationship between expansion properties and rumor speeding which we now know runs very deep. Following our work, optimal bounds for rumor spreading in terms of vertex expansion have been obtained [22, 23, 35], and extensions of the bounds with conductance and vertex expansions have been proposed to dynamic graph settings [24]. Other works have looked for general bounds for the completion time of rumor spreading without any dependence on conductance, showing that $O(D + \mathrm{polylog}(n))$ rounds are sufficient, where D is the diameter of the graph (and hence a trivial lower bound on the completion time) [5, 26]. Still, these results rely on conductance, in the sense that they show that every graph can be partitioned into regions each of which is of high conductance, and within which rumors spread fast for the reasons elucidated in this work.

The rest of the article is organized as follows. Section 2 discusses the relevant literature. Section 3 introduces some common concepts and definitions to be used in the rest of the article. Section 4 presents the optimal $O(\log n/\phi)$ bound for PUSH–PULL. Section 5 presents a sufficient condition for this bound to hold when PUSH and PULL operate by themselves. Finally, Section 6 shows the connection between rumor spreading and graph sparsification.

2   RELATED WORK

The literature on rumor spreading is huge and we confine ourselves to what appears to be more relevant to the present work.

Clearly, at least diameter-many rounds of rumor spreading are needed for the message to reach all nodes. On the other hand, Feige et al. [18] have shown that $O(n \log n)$ rounds are sufficient with

---

[1]This article is the result of a series of intermediate steps [9, 10, 21]. Reference [10] was the first to show the connection between rumor spreading and conductance. By using a different approach, Reference [9] followed up with an almost tight solution to the problem. Finally, Reference [21] improved the latter approach introducing new ideas, obtaining a tight solution.

high probability for any connected graph with n nodes. The problem has been studied on a number of graph classes, such as cliques, hypercubes, bounded-degree graphs, and various types of random graphs (see, for instance, References [18, 20, 33]). More recently, there has been a lot of work on "quasi-regular" expanders, i.e., expander graphs for which the ratio between the maximum and minimum degree is constant. It has been shown in different settings [3, 14, 15, 34] that $O(\log n)$ rounds are sufficient for the message to spread throughout the graph. Our work can be viewed as an extension of these studies to graphs of arbitrary degree distribution. Observe that many real-world graphs (e.g., Facebook, the Internet, etc.) have a very skewed degree distribution, i.e., the ratio between the maximum and minimum degrees is very high. In most graph models for social networks this ratio is polynomial in the order n of the graph.

Mihail et al. [30] have studied the edge expansion and conductance of graphs that are very similar to preferential attachment graphs and showed that their edge expansion and conductance are constant. For "genuine" preferential-attachment graphs, Chierichetti et al. [8] showed that rumor spreading finishes in $O(\log^2 n)$ rounds. A subsequent article by Doerr et al. [13] improved that by showing a tight bound of $O(\log n)$ rounds, and, moreover, if one is interested in informing only $n - o(n)$ nodes, then one has to wait for just $O(\log n / \log \log n)$ many rounds.

Fountoulakis et al. [19] studied another class of graphs with power-law degree distribution, for which they showed that $O(\log \log n)$ rounds suffice with high probability to inform $n - o(n)$ nodes.

Boyd et al. [4], and Mosk-Aoyama and Shah [32] have shown that high conductance implies that *non-uniform* (over neighbors) rumor spreading is fast. By non-uniform we mean that, for every ordered pair of neighbors i and j, node i will select j with probability p_{ij} for the rumor spreading step (in general, $p_{ij} \neq p_{ji}$). This result does not extend to the case of uniform probabilities studied in this article. In our setting (but not in theirs), the existence of a non-uniform distribution that makes rumor spreading fast is a rather trivial matter. A graph of conductance ϕ has diameter bounded by $O(\log n / \phi)$. Observe that in a synchronous network, it is possible to elect a leader in $O(\log n / \phi)$ many rounds and set up a BFS tree originating from it. Assigning probability 1 to the edge between a node and its parent gives the desired non-uniform probability distribution.

Giakkoupis et al. [25] have shown an $O(\log n / \phi)$ bound for a "low-randomness" variant of PUSH on regular graphs. In their algorithm, the random choices of nodes are not independent but rely instead on a pseudorandom generator whose random seed is attached to the message.

Another standard measure of graph expansion that has been studied in connection with rumor spreading is vertex expansion [22, 23, 35]. The main result of this line of work is a tight bound of $O(\log n \cdot \log \Delta / \alpha)$ on the number of rounds for PUSH–PULL to spread a message with high probability, for any graph with vertex expansion α and maximum degree Δ. This bound is incomparable to the $O(\log n / \phi)$ bound, as the vertex expansion is incomparable to conductance for general graphs.

An interesting application of the $O(\log n / \phi)$ bound has been in the design of a new class of information dissemination protocols [5, 6, 26]. These protocols rely on the fact that PUSH–PULL spreads information fast in subgraphs of high conductance, and they combine PUSH–PULL with more sophisticated rules on how each node chooses the neighbor to contact in a round. These new protocols achieve fast information spreading for broad families of networks, and some of these protocols achieve for all graphs time bounds that are close (within polylogarithmic factors or polylogarithmic terms) to the network diameter.

For the message complexity of PUSH–PULL, we observe first that it has been determined precisely only for very special classes of graphs (namely, cliques [28] and Erdös-Rényi random graphs [17]). Apart from that, given the generality of the graph class we consider, it is not possible to obtain a non-trivial upper bound on the number of messages—where the trivial upper bound is the number of rounds times n. Consider, for instance, the "lollipop graph." Fix $\omega(n^{-1}) < \phi < O(\log^{-1} n)$, and suppose that we have a path of length ϕ^{-1} connected to a clique of size $n - \phi^{-1} = \Theta(n)$. This graph

has conductance $O(\phi)$. Let the source be an arbitrary node in the clique. After $\Theta(\log n)$ rounds each node in the clique will have the information with high probability. Further it will take at least ϕ^{-1} steps for the information to be sent to each node in the path. So, at least $n - \phi^{-1} = \Theta(n)$ messages are pushed (by the nodes in the clique) in each round, for at least $\phi^{-1} - \Theta(\log n) = \Theta(\phi^{-1})$ rounds. Thus, the total number of messages sent will be $\Omega(n\phi^{-1})$. Observing that the number of rounds is $\Theta(\phi^{-1} + \log n) = \Theta(\phi^{-1})$, we obtain that the number of rounds times n is (asymptotically) less than or equal to the number of transmitted messages.

We also note that one cannot give fault-tolerance guarantees (that is, the ability of the protocol to resist to node and/or edge deletions) based only on conductance. A star has high conductance, but failure of the central node destroys connectivity.

3 PRELIMINARIES

In this section, we introduce some, mostly standard, notation and provide some basic technical lemmas to be used in the analysis.

Let $G = (V, E)$ be an undirected graph with $n = |V|$ vertices. For a vertex $v \in V$, $\Gamma_G(v)$ denotes the set of v's neighbors in G, and $\deg_G(v) := |\Gamma_G(v)|$ is the degree of v. By Δ_G we denote the maximum degree of G. For a set of vertices $S \subseteq V$, we let

$$\Gamma_G(S) := \bigcup_{v \in S} \Gamma_G(v) - S.$$

We define the *volume* of set $S \subseteq V$ to be the sum of the degrees of the vertices contained in S,

$$\mathrm{vol}_G(S) := \sum_{v \in S} \deg_G(v).$$

Observe that $\mathrm{vol}_G(V) = 2 \cdot |E|$. Given two disjoint sets of vertices $A, B \subseteq V$, we let $\mathrm{cut}_G(A, B)$ be the set of edges with one endpoint in A and the other in B.

Definition 1 (Conductance [27]). The conductance of a non-empty vertex set S in graph G is

$$\phi_G(S) := \frac{|\mathrm{cut}_G(S, V - S)|}{\mathrm{vol}_G(S)}.$$

The conductance of G is the minimum $\phi_G(S)$, over all sets S for which $\mathrm{vol}(S) \leq \mathrm{vol}_G(V)/2$,

$$\phi_G := \min_{S \subseteq V, \ 0 < \mathrm{vol}(S) \leq \mathrm{vol}(V)/2} \frac{|\mathrm{cut}_G(S, V - S)|}{\mathrm{vol}_G(S)}.$$

In all the notations above we will omit G whenever the graph is clear from the context.

We say that an event occurs *with high probability*, abbreviated as *w.h.p.*, if it happens with probability $1 - n^{-\Omega(1)}$.

Let $B(k, p)$ denote a binomial random variable with parameters k and p, i.e., $B(k, p)$ is the number of successes among k independent trials each of which succeeds with probability p. We will use the next standard lemma, which can be found in Reference [1, Lemma 3.1].

LEMMA 2. *Let $Z_1, \ldots, Z_k$ be random variables with values in some arbitrary domain, and let $X_1, \ldots, X_k$ be binary random variables such that $X_i = X_i(Z_1, \ldots, Z_i)$, for $1 \leq i \leq k$. Let $X := \sum_{i=1}^{k} X_i$. If $\Pr(X_i = 1 \mid Z_1, \ldots, Z_{i-1}) \geq p$, for all $1 \leq i \leq k$, then for any t,*

$$\Pr(X < t) \leq \Pr(B(k, p) < t).$$

The following lemma gives simple bounds on the number of coupons collected in the Coupon Collector's process, within a given number of trials. We denote that number by $C(k, t)$, where k is the total number of coupons and t is the number of trials.

LEMMA 3. *Consider the standard Coupon Collector's process, where we are given an urn of k different coupons, and we draw coupons from the urn, equally likely with replacement. Let $C(k, t)$ denote the number of distinct coupons drawn in t trials. Then for any $t \geq 0$,*

$$\Pr(C(k, t) < k) \leq ke^{-t/k}, \tag{1}$$

and for any $\lambda < k$,

$$\Pr(C(k, t) \leq \lambda) \leq \Pr(B(t, 1 - \lambda/k) \leq \lambda). \tag{2}$$

PROOF. The probability that a given coupon is not selected in t trials is

$$(1 - 1/k)^t \leq e^{-t/k},$$

and by a union bound, the probability at least one coupon is not selected, i.e., $C(k, t) < k$, is at most $ke^{-t/k}$. This proves (1).

Let X_i, $1 \leq i \leq t$, be a binary random variable which is 1 if a new coupon is selected in the ith trial or if the total number of distinct coupons selected in the previous $i - 1$ trials is larger than λ; otherwise, $X_i = 0$. Then

$$\Pr(C(k, t) \leq \lambda) = \Pr\left(\sum_{i=1}^{t} X_i \leq \lambda\right).$$

The probability that $X_i = 1$ is at least $1 - \frac{\lambda}{k}$, independently of the outcome of the previous trials, because either more than λ distinct coupons have been selected in the first $i - 1$ trials and then $X_i = 1$, or at most λ coupons have been selected, and thus the probability a new coupon is selected in the ith trial is at least $\frac{k-\lambda}{k}$. Therefore, we can apply Lemma 2 to obtain

$$\Pr\left(\sum_{i=1}^{t} X_i \leq \lambda\right) \leq \Pr(B(t, 1 - \lambda/k) \leq \lambda).$$

Combining the last two equations above gives Equation (2).   □

Finally, we will use the following Chernoff bounds for binomial random variables, which can be found, e.g., in References [16, 31].

LEMMA 4 (CHERNOFF BOUNDS).

(a) *For $t \leq k/2$,*

$$\Pr(B(k, 1/2) \leq t) \leq e^{-2(k/2-t)^2/k}.$$

(b) *For $0 < \epsilon \leq 1$,*

$$\Pr(B(k, p) \geq (1 + \epsilon)np) \leq e^{-\epsilon^2 np/3}.$$

4 OPTIMAL BOUND FOR PUSH-PULL

In this section, we prove a general upper bound of $O(\log n/\phi)$ rounds for PUSH−PULL on any n-node graph with conductance ϕ. Further, we argue that this bound is best possible if we have no additional information for the graph other than ϕ and n.

The formal statement of the upper bound is as follows.

THEOREM 5. *Let $G = (V, E)$ be a connected n-node graph with conductance ϕ. Then $O(\log n/\phi)$ rounds of PUSH−PULL suffice w.h.p. to spread to all nodes of G a message originated at an arbitrary source node.*

The proof of Theorem 5 relies on the next key lemma, which states that a message spreads from a node of maximum degree to all other nodes in $O(\log n/\phi)$ rounds w.h.p. This statement is true even if only PULL is used. Recall that Δ denotes the maximum degree of G.

LEMMA 6. *Using* PULL*, a message originated at a node of degree Δ spreads to all other nodes in $O(\log n/\phi)$ rounds with probability at least $1 - n^{-c}$, for any constant $c > 0$.*

The proof of Lemma 6 is based on an analysis of the growth of the volume of the set of informed nodes and is given in Section 4.1.

We will also use the following basic lemma, which underlines the symmetries in the PUSH and PULL strategies.

LEMMA 7. *Let $u \xrightarrow{t}_{push} v$ denote the event that a message originated at vertex u reaches vertex v in at most t rounds when* PUSH *is used, and define $u \xrightarrow{t}_{pull} v$ similarly. For any $u, v \in V$, and $t \geq 0$, we have $\Pr(u \xrightarrow{t}_{push} v) = \Pr(v \xrightarrow{t}_{pull} u)$.*

PROOF. W.l.o.g. we can assume that in each round, *every* node (informed or not) chooses a random neighbor. Of course only informed nodes actually establish a connection to the neighbor they choose in the PUSH strategy, and only uninformed nodes establish a connection in the PULL strategy. Let Ω_t be the space of possible choices for all nodes over t rounds, i.e., each $\omega \in \Omega_t$ specifies for every node $u \in V$ and round $i \in \{1 \ldots t\}$, the neighbor that u chooses in round i. Observe that all $\omega \in \Omega_t$ have the same probability: $\Pr(\omega) = \prod_{u \in V} (\deg(u))^{-t}$. Further, each $\omega \in \Omega_t$ completely determines the spread of a message in t rounds, given the source node and the rumor spreading strategy. Let $\Omega_t(u, v, \text{push})$ be the set of all $\omega \in \Omega_t$ for which a message originated at u reaches v within t rounds when PUSH is used, and define $\Omega_t(u, v, \text{pull})$ similarly. Since all $\omega \in \Omega_t$ have the same probability, it suffices to show that $|\Omega_t(u, v, \text{push})| = |\Omega_t(v, u, \text{pull})|$ to prove $\Pr(u \xrightarrow{t}_{push} v) = \Pr(v \xrightarrow{t}_{pull} u)$.

For $\omega \in \Omega_t$, let $rev(\omega) \in \Omega_t$ contain the same collection of choices as ω but in reverse round order, i.e., if node u chooses node v in round i in ω, then u chooses v in round $t - i$ in $rev(\omega)$. We now argue that $\omega \in \Omega_t(u, v, \text{push})$ iff $rev(\omega) \in \Omega_t(v, u, \text{pull})$. A *push path* for $\omega \in \Omega_t$ is a list of nodes $u_0, \ldots, u_t$ such that for any two consecutive nodes u_{i-1} and u_i, either (i) $u_{i-1} = u_i$, or (ii) u_{i-1} chooses u_i in round i. A *pull path* is defined similarly except that (ii) is replaced by (ii') u_i chooses u_{i-1} in round i. For a given $\omega \in \Omega_t$, PUSH (or PULL) informs node v within t rounds of a message originated at u iff there is a push path (respectively, pull path) from u to v. Further, $u_0, u_1, \ldots, u_t$, with $u_0 = u$ and $u_t = v$, is a push path for ω iff $u_t, u_{t-1}, \ldots, u_0$ is a pull path for $rev(\omega)$. From the last two observations, it follows that $\omega \in \Omega_t(u, v, \text{push})$ iff $rev(\omega) \in \Omega_t(v, u, \text{pull})$. From this and the fact that for distinct $\omega, \omega' \in \Omega_t$ we have $rev(\omega) \neq rev(\omega')$, we obtain $|\Omega_t(u, v, \text{push})| = |\Omega_t(v, u, \text{pull})|$. □

Theorem 5 follows easily from the two lemmas above: From Lemma 6 we have that a message originating from an arbitrary source node u needs $O(\log n/\phi)$ rounds of PULL to spread to all nodes w.h.p., *after it has reached a node v of degree Δ*. Further, by combining Lemma 6 with Lemma 7, we obtain that $O(\log n/\phi)$ rounds suffice w.h.p. for u's message to reach the max-degree node v using PUSH.

PROOF OF THEOREM 5 (ASSUMING LEMMA 6 HOLDS). Let u be an arbitrary node, and v be a node with $\deg(v) = \Delta$. By Lemma 6, there is some $t = O(\log n/\phi)$ such that t rounds of PULL suffice with probability at least $1 - n^{-c}$ to spread a message from v to all other nodes and thus to node u. From this and Lemma 7, it follows that t rounds of PUSH suffice to spread a message from u to v with probability at least $1 - n^{-c}$; thus the same is true when PUSH–PULL is used instead of just PUSH. Once v has learned u's message, Lemma 6 implies, as mentioned above, that all nodes will learn the message with probability at least $1 - n^{-c}$ after at most t additional rounds of PULL (and thus of PUSH–PULL). A union bound then gives that with probability at least $1 - 2n^{-c}$, v learns u's message

within t rounds and then all remaining nodes learn the message within t additional rounds. Thus all nodes get informed after a total number of at most $2t = O(\log n/\phi)$ rounds w.h.p.   □

Below, we give the proof of Lemma 6, in Section 4.1, and we argue about the optimality of Theorem 5, in Section 4.2.

4.1 Analysis of Pull (Proof of Lemma 6)

In this section, we show that $O(\log n/\phi)$ rounds of PULL suffice w.h.p. to spread a message to all nodes if the source node has degree Δ.

The proof analyses the growth rate of the volume of informed vertices. Let I_t be the set of informed nodes after the first t rounds, and let $U_t = V - I_t$ be the set of uninformed nodes at that time. Roughly speaking, we argue that while $\mathrm{vol}(I_t)$ is at most some constant fraction of the total volume $\mathrm{vol}(V)$, $\mathrm{vol}(I_t)$ increases by a factor of at least $1 + \Omega(\phi)$ per round on average; and after that, $\mathrm{vol}(U_t)$ decreases by a factor of at most $1 - \Omega(\phi)$. It follows that $O(\log(\mathrm{vol}(V))/\phi) = O(\log n/\phi)$ rounds suffice for $\mathrm{vol}(I_t)$ to increase to a constant fraction of $\mathrm{vol}(V)$ and then $O(\log n/\phi)$ additional rounds suffice for $\mathrm{vol}(U_t)$ to drop to zero.

To convey the main intuition let us compute the expected increase of $\mathrm{vol}(I_t)$ in round $t + 1$ (see Lemma 8(a) for the formal argument): Suppose that $\mathrm{vol}(I_t) \leq \mathrm{vol}(V)/2$, and for each uninformed node $v \in U_t$ let $\deg'(v)$ be the number of its informed neighbors. The contribution of v to the expected increase of $\mathrm{vol}(I_t)$ is then $(\deg'(v)/\deg(v)) \cdot \deg(v) = \deg'(v)$, as v chooses some of its informed neighbors in round $t + 1$ (and pulls the message from it) with probability $\deg'(v)/\deg(v)$. It follows

$$\mathbf{E}[\mathrm{vol}(I_{t+1}) \mid I_t] - \mathrm{vol}(I_t) = \sum_{v \in U_t} \deg'(v) = |\mathrm{cut}(I_t, U_t)| \geq \phi \cdot \mathrm{vol}(I_t),$$

by the definition of conductance. Thus, $\mathbf{E}[\mathrm{vol}(I_{t+1}) \mid I_t] \geq (1 + \phi)\,\mathrm{vol}(I_t)$. Similarly, we obtain $\mathbf{E}[\mathrm{vol}(U_{t+1}) \mid U_t] \leq (1 - \phi)\,\mathrm{vol}(U_t)$, when $\mathrm{vol}(I_t) \geq \mathrm{vol}(V)/2$. These simple expectation bounds, together with Markov's Inequality, suffice to obtain an $O(\log n/\phi)$ bound w.h.p. on the number of rounds until all nodes get informed, from the time when a constant fraction of the total volume has been informed (Lemma 11). To bound the time to inform that constant fraction, however, more work is required. For that we bound also the variance of the increase of $\mathrm{vol}(I_t)$ (Lemma 8(b)) and use the second moment method to show that the volume doubles in $O(1/\phi)$ rounds with constant probability (Lemma 9). From this last result and Chernoff Bounds, we obtain the desired high-probability bound on the time to inform a constant fraction of the total volume (Lemma 10).

Next we give the detailed analysis.

Consider the volume L of nodes that get informed in a round by pulling the message through a given set D of edges between informed and uninformed nodes (D does not necessarily contain all edges between informed and uninformed nodes and is fixed before the round starts). The next lemma states that the expectation of L is equal to the number of edges contained in D, and L's variance is bounded by the number of edges times the maximum degree of their endpoints in U_t.

LEMMA 8. *Fix an arbitrary subset D of* $\mathrm{cut}(I_t, U_t)$ *before round $t + 1$, and let L be the volume of the set of nodes that get informed in round $t + 1$ by pulling the message through edges from D. Then*

(a) $\mathbf{E}[L] = |D|$.
(b) $\mathbf{Var}[L] \leq |D| \cdot \max_{v \in N} \deg(v)$, *where N is the set of nodes $v \in U_t$ incident to some edge in D.*

PROOF. For each node $v \in N$, let $\deg_D(v)$ be the number of edges from set D that are incident to v. Clearly, $\sum_{v \in N} \deg_D(v) = |D|$. For each $v \in N$, let X_v be a binary random variable that is 1 if v gets informed in round $t + 1$ by pulling the message through some edge from D and is 0 otherwise.

Then, $\Pr(X_v = 1) = \deg_D(v)/\deg(v)$. We have $L = \sum_{v \in N}(X_v \cdot \deg(v))$, and thus

$$\mathbf{E}[L] = \sum_{v \in N} \mathbf{E}[X_v \cdot \deg(v)] = \sum_{v \in N} \Pr(X_v = 1) \cdot \deg(v)$$

$$= \sum_{v \in N} (\deg_D(v)/\deg(v)) \cdot \deg(v)$$

$$= \sum_{v \in N} \deg_D(v)$$

$$= |D|.$$

Next we bound the variance of L. Since $L = \sum_{v \in N}(X_v \cdot \deg(v))$ and the random variables X_v are independent,

$$\mathbf{Var}[L] = \sum_{v \in N} \mathbf{Var}[X_v \cdot \deg(v)] = \sum_{v \in N} \mathbf{Var}[X_v] \cdot (\deg(v))^2.$$

Using also that $\mathbf{Var}[X_v] = \Pr(X_v = 1) \cdot \Pr(X_v = 0) \le \Pr(X_v = 1) = \deg_D(v)/\deg(v)$, we obtain

$$\mathbf{Var}[L] \le \sum_{v \in N} (\deg_D(v) \cdot \deg(v)) \le \left(\sum_{v \in N} \deg_D(v) \right) \cdot \max_{v \in N} \deg(v) = |D| \cdot \max_{v \in N} \deg(v). \qquad \square$$

In the following, we first bound the number of rounds to inform a constant fraction of the total volume, namely a 1/3 fraction, and then bound the number of rounds to inform the remaining nodes.

Informing a Constant Fraction of the Volume. We prove that $O(\ln n/\phi)$ rounds of PULL suffice w.h.p. to inform a subset of nodes with volume at least $\mathrm{vol}(V)/3$, when the message starts from a node of degree Δ.

Our key claim is that if $\Delta \le \mathrm{vol}(I_t) \le \mathrm{vol}(V)/3$, then $O(1/\phi)$ rounds suffice with constant probability to double the volume of informed nodes. The proof proceeds as follows. Before each round $t + k > t$, we fix an edge set $D_k \subseteq \mathrm{cut}(I_{t+k-1}, U_{t+k-1})$ of size $|D_k| = \phi \cdot \mathrm{vol}(I_t)$; this set can be arbitrary but must be fixed before round $t + k$ starts. Then we compute the volume L_k of the set of nodes that get informed in round $t + k$ by pulling the message through edges from D_k. Clearly, $\sum_{1 \le i \le k} L_i$ is a lower bound on $\mathrm{vol}(I_{t+k}) - \mathrm{vol}(I_t)$ and thus to prove the claim it suffices to show that $\sum_{1 \le i \le k} L_i \ge \mathrm{vol}(I_t)$ holds with constant probability. To show that, we use Lemma 8 to compute the expectation of $\sum_{1 \le i \le k} L_i$ and to bound its variance and then apply Chebyshev's Inequality.

Lemma 9 below proves a slightly more general statement, which does not assume that volume $\mathrm{vol}(I_t)$ is at least Δ, and gives a lower bound on the probability that the volume doubles within a given number k of steps, as a function of k and the ratio $\Delta/\mathrm{vol}(I_t)$.

LEMMA 9. *If* $\mathrm{vol}(I_t) \le \mathrm{vol}(V)/3$, *then for any* $k > 1/\phi$,

$$\Pr\left(\mathrm{vol}(I_{t+k}) \ge 2\,\mathrm{vol}(I_t) \mid I_t \right) > 1 - \frac{k\phi}{(k\phi - 1)^2} \cdot \frac{\Delta}{\mathrm{vol}(I_t)}.$$

PROOF. Throughout the proof, we assume that I_t is fixed and $\mathrm{vol}(I_t) \le \mathrm{vol}(V)/3$. Let

$$\tau = \min \left\{ k : \mathrm{vol}(I_{t+k}) \ge 2\,\mathrm{vol}(I_t) \right\}.$$

For each $1 \le k \le \tau$, we define random variable L_k as follows. Let D_k denote the set of the first $M := \lceil \phi \cdot \mathrm{vol}(I_t) \rceil$ edges in $\mathrm{cut}(I_{t+k-1}, U_{t+k-1})$, according to some fixed total order over E.[2] (We

---

[2]In fact, it is not important how set $D_k \subseteq \mathrm{cut}(I_{t+k-1}, U_{t+k-1})$ is chosen, as long as it is fixed before round $t + k$ and has size M.

will see in Equation (3) that $|\operatorname{cut}(I_{t+k-1}, U_{t+k-1})| \geq M$.) Then L_k is the volume of the set of nodes that get informed in round $t + k$ by pulling the message through edges from D_k. Further, for each $k > \tau$, we let $L_k = M$.

We argue now that for any t' with $t \leq t' < t + \tau$,

$$|\operatorname{cut}(I_{t'}, U_{t'})| \geq M. \tag{3}$$

If $\operatorname{vol}(I_{t'}) \leq \operatorname{vol}(V)/2$, then $\phi(I_{t'}) \geq \phi$. Also $\operatorname{vol}(I_{t'}) \geq \operatorname{vol}(I_t)$. Then

$$|\operatorname{cut}(I_{t'}, U_{t'})| = \phi(I_{t'}) \cdot \operatorname{vol}(I_{t'}) \geq \phi \cdot \operatorname{vol}(I_t).$$

If $\operatorname{vol}(I_{t'}) > \operatorname{vol}(V)/2$, then $\operatorname{vol}(U_{t'}) < \operatorname{vol}(V)/2$, and thus $\phi(U_{t'}) \geq \phi$. Also $\operatorname{vol}(I_{t'}) < 2\operatorname{vol}(I_t)$ and thus $\operatorname{vol}(U_{t'}) > \operatorname{vol}(V) - 2\operatorname{vol}(I_t) \geq \operatorname{vol}(I_t)$, as $\operatorname{vol}(I_t) \leq \operatorname{vol}(V)/3$. Then

$$|\operatorname{cut}(I_{t'}, U_{t'})| = \phi(U_{t'}) \cdot \operatorname{vol}(U_{t'}) \geq \phi \cdot \operatorname{vol}(I_t).$$

For each $k \geq 0$, let $Y_k = \sum_{1 \leq i \leq k} L_i$. Then for $k \leq \tau$, Y_k is a lower bound on the volume of nodes informed in rounds $t + 1$ up to $t + k$, i.e., $Y_k \leq \operatorname{vol}(I_{t+k}) - \operatorname{vol}(I_t)$. It follows that if $\operatorname{vol}(I_{t+k}) < 2\operatorname{vol}(I_t)$, then $k < \tau$ and $Y_k \leq \operatorname{vol}(I_{t+k}) - \operatorname{vol}(I_t) < \operatorname{vol}(I_t)$. Therefore,

$$\Pr(\operatorname{vol}(I_{t+k}) < 2\operatorname{vol}(I_t)) \leq \Pr(Y_k < \operatorname{vol}(I_t)). \tag{4}$$

Next we bound the probability on the right side that $Y_k < \operatorname{vol}(I_t)$. For that, we compute the expectation and variance of Y_k and then apply Chebyshev's Inequality.

If $1 \leq k \leq \tau$, then from Lemma 8(a) it follows that

$$\mathbf{E}[L_k \mid I_{t+k-1}] = |D_k| = M.$$

This equation is also trivially true if $k > \tau$, as $L_k = M$ in this case. It follows that for all $k \geq 1$,

$$\mathbf{E}[L_k] = \mathbf{E}[\mathbf{E}[L_k \mid I_{t+k-1}]] = M,$$

and by linearity of expectation,

$$\mathbf{E}[Y_k] = kM. \tag{5}$$

For the variance of Y_k we have

$$\mathbf{Var}[Y_k] = \mathbf{Var}[Y_{k-1} + L_k] = \mathbf{Var}[Y_{k-1}] + \mathbf{Var}[L_k] + 2\,\mathbf{Cov}[Y_{k-1}, L_k].$$

The covariance term is zero, as

$$\mathbf{E}\left[(Y_{k-1} - \mathbf{E}[Y_{k-1}]) \cdot (L_k - \mathbf{E}[L_k]) \,\Big|\, I_t \ldots I_{t+k-1}\right]$$
$$= (Y_{k-1} - \mathbf{E}[Y_{k-1}]) \cdot (\mathbf{E}[L_k \mid I_{t+k-1}] - \mathbf{E}[L_k])$$
$$= (Y_{k-1} - \mathbf{E}[Y_{k-1}]) \cdot (M - M) = 0,$$

and thus $\mathbf{Cov}[Y_{k-1}, L_k] = \mathbf{E}[\mathbf{E}[(Y_{k-1} - \mathbf{E}[Y_{k-1}]) \cdot (L_k - \mathbf{E}[L_k]) \mid I_t \ldots I_{t+k-1}]] = 0$. Substituting this to the formula for $\mathbf{Var}[Y_k]$ above yields

$$\mathbf{Var}[Y_k] = \mathbf{Var}[Y_{k-1}] + \mathbf{Var}[L_k].$$

We now bound $\mathbf{Var}[L_k]$. If $1 \leq k \leq \tau$, then from Lemma 8(b) it follows

$$\mathbf{Var}[L_k \mid I_{t+k-1}] \leq |D_k| \cdot \max_{v \in N_k} \deg(v) \leq M\Delta, \tag{6}$$

where N_k is the set of nodes $v \in U_{t+k-1}$ incident to some edge in D_k. The above inequality is also trivially true if $k > \tau$, as $\mathbf{Var}[L_k \mid I_{t+k-1}] = 0$ in this case. It follows that $\mathbf{Var}[L_k] \leq M\Delta$, for all $k \geq 1$. Substituting this to the equation for $\mathbf{Var}[Y_k]$ above gives $\mathbf{Var}[Y_k] \leq \mathbf{Var}[Y_{k-1}] + M\Delta$, and applying this inequality repeatedly yields

$$\mathbf{Var}[Y_k] \leq kM\Delta. \tag{7}$$

We use now Chebyshev's Inequality to bound the probability that $Y_k < \mathrm{vol}(I_t)$. We assume that $k > 1/\phi$. Then from Equation (5), $\mathbf{E}[Y_k] = kM > M/\phi \geq \mathrm{vol}(I_t)$. We have

$$\Pr(Y_k < \mathrm{vol}(I_t)) = \Pr(\mathbf{E}[Y_k] - Y_k > \mathbf{E}[Y_k] - \mathrm{vol}(I_t)) < \frac{\mathbf{Var}[Y_k]}{(\mathbf{E}[Y_k] - \mathrm{vol}(I_t))^2}.$$

Substituting the value of $\mathbf{E}[Y_k]$ from Equation (5) and the bound for $\mathbf{Var}[Y_k]$ from Equation (7), we obtain

$$\Pr(Y_k < \mathrm{vol}(I_t)) < \frac{kM\Delta}{(kM - \mathrm{vol}(I_t))^2}$$
$$= \frac{\Delta}{\mathrm{vol}(I_t)} \cdot \frac{kM/\mathrm{vol}(I_k)}{(kM/\mathrm{vol}(I_k) - 1)^2}$$
$$\leq \frac{\Delta}{\mathrm{vol}(I_t)} \cdot \frac{k\phi}{(k\phi - 1)^2},$$

where the last inequality holds because $kM/\mathrm{vol}(I_t) \geq k\phi$, and the function $f(x) = x/(x-1)^2$ is decreasing for $x > 1$. Combining the above bound on $\Pr(Y_k < \mathrm{vol}(I_t))$ with Equation (4) completes the proof of Lemma 9. □

Using Lemma 9 and Chernoff Bounds, we show next our main claim, that $O(\ln n/\phi)$ rounds suffice w.h.p. to inform a subset of nodes with volume at least $\mathrm{vol}(V)/3$.

LEMMA 10. *If the source of the message has degree Δ, then $\Pr(\mathrm{vol}(I_t) \geq \mathrm{vol}(V)/3) \geq 1 - n^{-c}$, for $t = \lceil 2c \ln n + 4 \log n \rceil \cdot \lceil 4/\phi \rceil$.*

PROOF. We divide rumor spreading into phases of $k := \lceil 4/\phi \rceil$ rounds each and consider whether the volume of informed vertices doubles in these phases. For each $i \geq 1$, let X_i be a binary random variable that is 1 if $\mathrm{vol}(I_{ik}) \geq 2\,\mathrm{vol}(I_{(i-1)k})$ or $\mathrm{vol}(I_{(i-1)k}) \geq \mathrm{vol}(V)/3$ and is 0 otherwise. From Lemma 9 it follows

$$\Pr\left(X_i = 1 \mid I_{(i-1)k}\right) > 1 - \frac{k\phi}{(k\phi - 1)^2} \cdot \frac{\Delta}{\mathrm{vol}(I_{(i-1)k})} \geq 1 - \frac{4}{(4-1)^2} \cdot \frac{\Delta}{\Delta} = 5/9 > 1/2. \qquad (8)$$

We bound the probability that $\mathrm{vol}(I_t) < \mathrm{vol}(V)/3$ in terms of the sum of X_i: If $\mathrm{vol}(I_{ik}) < \mathrm{vol}(V)/3$, then fewer than $\log(\mathrm{vol}(V)/(3\Delta))$ among the random variables $X_1 \ldots X_i$ are 1, because, otherwise, the volume of informed nodes would double at least $\log(\mathrm{vol}(V)/(3\Delta))$ times in the first ik rounds and thus increase from Δ to at least $2^{\log(\mathrm{vol}(V)/(3\Delta))}\Delta = \mathrm{vol}(V)/3$. Therefore,

$$\Pr\left(\mathrm{vol}(I_{ik}) < \mathrm{vol}(V)/3\right) \leq \Pr\left(\sum_{j \leq i} X_j < \log(\mathrm{vol}(V)/(3\Delta))\right) \leq \Pr\left(\sum_{j \leq i} X_j < \log n\right),$$

as $\mathrm{vol}(V) \leq n\Delta$. To bound the sum of the (non-independent) random variables X_i, we use the fact we showed in Equation (8) that for all $i \geq 1$, $\Pr(X_i = 1 \mid I_{(i-1)k}) \geq 1/2$, and apply Lemma 2 to obtain $\Pr(\sum_{j \leq i} X_j < \log n) \leq \Pr(B(i, 1/2) < \log n)$. Then Lemma 4(a) gives for $i \geq 2 \log n$,

$$\Pr\left(\sum_{j \leq i} X_j < \log n\right) \leq e^{-2(i/2 - \log n)^2/i} \leq e^{-(i/2 - 2\log n)}.$$

Letting $i \geq 2c \ln n + 4 \log n$, the above bound becomes at most $e^{-c \ln n} = n^{-c}$. Therefore, we have that $\Pr(\mathrm{vol}(I_{ik}) < \mathrm{vol}(V)/3) \leq n^{-c}$ for $i = \lceil 2c \ln n + 4 \log n \rceil$. □

Informing the Remaining Nodes. We show that $O(\ln n/\phi)$ rounds of PULL suffice w.h.p. to inform all the remaining nodes, after a set of nodes with volume at least $\mathrm{vol}(V)/3$ has been informed.

LEMMA 11. *If* $\mathrm{vol}(I_t) \geq \mathrm{vol}(V)/3$, *then* $\Pr(U_{t+k} = \emptyset \mid I_t) \geq 1 - n^{-c}$, *for* $k = \lceil 2(c+2) \ln n/\phi \rceil$.

PROOF. In the following, we assume I_t is fixed and $\mathrm{vol}(I_t) \geq \mathrm{vol}(V)/3$. We first argue that for $t' \geq t$,

$$| \mathrm{cut}(I_{t'}, U_{t'})| \geq \phi \cdot \mathrm{vol}(U_{t'})/2.$$

If $\mathrm{vol}(U_{t'}) \leq \mathrm{vol}(V)/2$, then $\phi(U_{t'}) \geq \phi$ and $| \mathrm{cut}(I_{t'}, U_{t'})| = \phi(U_{t'}) \cdot \mathrm{vol}(U_{t'}) \geq \phi \cdot \mathrm{vol}(U_{t'})$. If $\mathrm{vol}(U_{t'}) > \mathrm{vol}(V)/2$, then $\phi(I_{t'}) \geq \phi$ and $| \mathrm{cut}(I_{t'}, U_{t'})| = \phi(I_{t'}) \cdot \mathrm{vol}(I_{t'}) \geq \phi \cdot \mathrm{vol}(I_{t'}) \geq \phi \cdot \mathrm{vol}(U_{t'})/2$, as $\mathrm{vol}(I_{t'}) \geq \mathrm{vol}(V)/3$.

Lemma 8(a), for $D = \mathrm{cut}(I_{t'}, U_{t'})$, gives $\mathbf{E}[\mathrm{vol}(U_{t'} - U_{t'+1}) \mid I_{t'}] = | \mathrm{cut}(I_{t'}, U_{t'})|$. Substituting the bound for $| \mathrm{cut}(I_{t'}, U_{t'})|$ from above and rearranging yields

$$\mathbf{E}[\mathrm{vol}(U_{t'+1}) \mid I_{t'}] \leq (1 - \phi/2) \cdot \mathrm{vol}(U_{t'}).$$

Taking the expectation on both sides gives $\mathbf{E}[\mathrm{vol}(U_{t'+1})] \leq (1 - \phi/2) \cdot \mathbf{E}[\mathrm{vol}(U_{t'})]$, and applying this inequality repeatedly yields, for $t' = t + k$,

$$\mathbf{E}[\mathrm{vol}(U_{t'})] \leq (1 - \phi/2)^k \cdot \mathrm{vol}(U_t) \leq e^{-k\phi/2} \cdot n^2.$$

For $k \geq 2(c+2) \ln n/\phi$, we get $\mathbf{E}[\mathrm{vol}(U_{t'})] \leq n^{-c}$, and then Markov's Inequality gives

$$\Pr(U_{t'} \neq \emptyset) = \Pr(\mathrm{vol}(U_{t'}) \geq 1) \leq \mathbf{E}[\mathrm{vol}(U_{t'})]/1 \leq n^{-c}. \qquad \square$$

Lemma 6 follows now immediately from Lemmas 10 and 11 and a union bound.

4.2 Tightness of the Upper Bound

We observe that the upper bound of Theorem 5 is tight. Indeed, there exist graphs on n nodes, with conductance ϕ, that have diameter $\Omega(\log n/\phi)$. We now present one example of such a graph. Let $\phi > 1/n^{1-\epsilon}$, and take any graph on $\Theta(\phi n)$ nodes with constant degree and constant conductance (e.g., a random cubic graph on $\Theta(\phi n)$ nodes.) Then, replace each edge of the graph with a path of length $\Theta(\phi^{-1})$. The resulting graph will have conductance $\Theta(\phi)$ and will contain $\Theta(n)$ nodes. Moreover, the diameter of the new graph will be equal to the diameter of the original graph times $\Theta(\phi^{-1})$—since the original graph had constant degree, its diameter is at least logarithmic in its order, that is, at least $\Omega(\log(\phi n))$. Therefore, the new graph will have diameter at least $\Omega(\log(\phi n)/\phi)$. Since $\phi > 1/n^{1-\epsilon}$, the lower bound reduces to $\Omega(\log n/\phi)$.

5 A SUFFICIENT CONDITION FOR PUSH AND PULL

The $O(\log n/\phi)$ bound we have shown for PUSH–PULL in Theorem 5 does not hold if only one of the PUSH or PULL strategies is used. As already mentioned, for the n-vertex star, which has conductance $\phi = 1$, at least $n - 1$ rounds of PUSH are needed to inform all nodes, regardless of the source. If PULL is used instead, and the message originates at a leaf of the star, then n rounds are needed in expectation ($n - 1$ expected rounds until the center of the star pulls the message from the source, and 1 additional round for all leaves to pull the message from the center). Observe, however, that if the message originates at the center, then all nodes get informed in just a single round of PULL. As shown in Lemma 6, for any graph $G = (V, E)$ there is some node $u \in V$ (namely, any max-degree node u) such that a message originated at u spreads to all nodes within $O(\log n/\phi)$ rounds of PULL w.h.p. The same is not true for PUSH, as the example of the star shows.

The next theorem identifies a simple graph property which guarantees that the $O(\log n/\phi)$ bound holds for PUSH and PULL, regardless of the source of the message. The property states that for any edge of the graph, the ratio of the degrees of its two endpoints is bounded by some constant.

THEOREM 12. *Let $G = (V, E)$ be a connected n-node graph with conductance ϕ and the property that for any edge $uv \in E$, $\deg(u)/\deg(v) = \Theta(1)$. Then $O(\log n/\phi)$ rounds of PUSH or PULL suffice w.h.p. to spread to all nodes of G a message originated at an arbitrary source node.*

PROOF. The analysis for PULL is the same as that for Lemma 6, except for a few small modifications we point out now. In place of Lemma 9, we use the following statement.

LEMMA 13. *Suppose that for every edge $uv \in E$, we have $\deg(u)/\deg(v) \in [1/\alpha, \alpha]$ for some $\alpha \geq 1$. If $\mathrm{vol}(I_t) \leq \mathrm{vol}(V)/3$, then for any $k > 1/\phi$,*

$$\Pr\left(\mathrm{vol}(I_{t+k}) \geq 2\,\mathrm{vol}(I_t) \mid I_t\right) > 1 - \frac{k\phi}{(k\phi - 1)^2} \cdot \alpha.$$

The proof of Lemma 13 is identical to that of Lemma 9, except that, in Equation (6), we do not bound $\max_{v \in N_k} \deg(v)$ by Δ but use instead that

$$\max_{v \in N_k} \deg(v) \leq \alpha \cdot \mathrm{vol}(I_t).$$

This inequality holds because each node $v \in N_k$ has degree bounded by $\alpha \cdot \max_{u \in I_{t+k-1}} \deg(u)$, and each node $u \in I_{t+k-1}$ has degree at most $\mathrm{vol}(I_t)$: If $u \in I_t$, then $\deg(u) \leq \mathrm{vol}(I_t)$, and if $u \in I_{t+k-1} - I_t$ then $\deg(u) \leq \mathrm{vol}(I_{t+k-1}) - \mathrm{vol}(I_t) < \mathrm{vol}(I_t)$, since $\mathrm{vol}(I_{t+k-1}) < 2\,\mathrm{vol}(I_t)$.

Using Lemma 13, we show, as in Lemma 10, that $O(\log n/\phi)$ rounds suffice w.h.p. to inform a $1/3$ fraction of the total volume, but now we do not need the assumption that the source node has degree Δ. Precisely, we show that $\Pr(\mathrm{vol}(I_t) \geq \mathrm{vol}(V)/3) \geq 1 - n^{-c}$ for $t = \lceil 2c \ln n + 4 \log n \rceil \cdot \lceil 4\alpha/\phi \rceil$. From this and Lemma 11, the bound of Theorem 12 follows for PULL.

For PUSH, the $O(\log n/\phi)$ bound is obtained by combining the above $O(\log n/\phi)$ bound for PULL, with Lemma 7 on the symmetry between the two strategies. □

6 RUMOR SPREADING AND GRAPH SPARSIFICATION

In this section, we give an alternative proof that PUSH–PULL finishes in a number of rounds that is at most polynomial in $\log n/\phi$. The proof relies on an interesting connection between rumor spreading and the Spielman-Teng sparsification process [36, 37].

THEOREM 14. *Let $G = (V, E)$ be a connected n-node graph with conductance ϕ. Then $O(\log^4 n/\phi^6)$ rounds of PUSH–PULL suffice w.h.p. to spread to all nodes in G a message originated at an arbitrary source node.*

The proof of Theorem 14 relates the *volume expansion* of the Spielman-Teng sparsification process, denoted ST, to that of an edge sampling process, PP, which selects edges similarly to PUSH–PULL. Roughly, the volume expansion of an edge sampling process on G is the minimum, over all $S \subseteq V$, of the ratio $\mathrm{vol}_G(S')/\mathrm{vol}_G(S)$, where S' is the set of vertices in $V - S$ that are reachable from S via a sampled edge (Definition 15). Process ST selects each edge uv independently with probability $\Upsilon/\min\{\deg(u), \deg(v)\}$, where $\Upsilon = \Theta(\log^2 n/\phi^4)$. Process PP on the other hand, chooses for each node u a number of $\Psi = \Theta(\Upsilon \cdot \log n)$ edges incident to u at random with replacement, imitating the random choices of nodes in Ψ rounds of PUSH–PULL.

It has been shown that the subgraph sampled by ST almost preserves certain spectral parameters of the original graph G. We use this result to show that the volume expansion of ST is $\Omega(\phi^2)$ (Lemma 21). We relate ST and PP by showing that process PP stochastically dominates ST, in the sense that there is a coupling of the two processes for which the edges selected by ST are w.h.p. a subset of those selected by PP (Lemma 23); this is our key technical claim. Once we have established this dominance, we infer that the volume expansion of PP is at least that of ST, i.e., $\Omega(\phi^2)$. It follows that Ψ rounds of PUSH–PULL increase the volume of informed nodes by a factor of $1 + \Omega(\phi^2)$, and

thus $O(\Psi \cdot \log n/\phi^2) = O(\log^4 n/\phi^6)$ rounds suffice to inform half of the total volume. A symmetric argument gives that $O(\log^4 n/\phi^6)$ additional rounds suffice to inform the remaining graph.

Below, we first set some definitions and notation, in Section 6.1. Then we review Spielman-Teng sparsification and analyze its volume expansion in Section 6.2. Last, we relate that process to rumor spreading and prove Theorem 14 in Section 6.3.

6.1 Definitions

We start with some terminology on edge sampling processes.

Definition 15 (Edge Sampling and Volume Expansion). Let f be a (randomized) process selecting edges in graph $G = (V, E)$. We denote by $f(G)$ the spanning subgraph of G whose edge set consists of the edges selected by f. Process f may also associate a weight with each edge it selects, in which case $f(G)$ is a weighted graph. The *volume expansion* of f on G, denoted $\beta_{f,G}$, is the minimum over all non-empty sets $S \subset V$, of the ratio[3]

$$\frac{\mathrm{vol}_G(\Gamma_{f(G)}(S))}{\min\{\mathrm{vol}_G(S), \mathrm{vol}_G(V-S)\}}.$$

Definition 15 above states that graph $f(G)$ obtained by process f may have weights associated to its edges. Next, we formalize various notions of weight we will use and also extend the standard definition of conductance to take weights into account.

Definition 16 (Weights). Let $G = (V, E, w)$ be a weighted graph where each edge $e \in E$ has weight w_e. The weight of a set of edges $E' \subseteq E$ is $w_G(E') := \sum_{e \in E'} w_e$; the weight of a vertex $u \in V$ is $w_G(v) := \sum_{v \in \Gamma(u)} w_{uv}$; and the weight of a set of vertices $S \subseteq V$ is $w_G(S) := \sum_{u \in S} w_G(u)$.

Definition 17 (Weighted Conductance). Consider a weighted graph $G = (V, E, w)$. The *weighted conductance* of a non-empty vertex set S in graph G is

$$\varphi_G(S) := \frac{w_G(\mathrm{cut}_G(S, V-S))}{w_G(S)}.$$

The *weighted conductance* φ_G of G is the minimum $\varphi_G(S)$ over all S for which $w_G(S) \leq w_G(V)/2$.

In the notations above, we will omit subscript G whenever the graph is clear from the context.

6.2 Spectral Sparsification

Next we recall the Spielman-Teng spectral sparsification procedure.

Definition 18 (Spielman-Teng Sparsification). Given a graph $G = (V, E)$, the Spielman-Teng sparsification process [36, 37], denoted ST, is an edge sampling process that selects each edge $uv \in E$ independently with probability

$$p_{uv} := \min\left\{1, \frac{\Upsilon}{\min\{\deg(u), \deg(v)\}}\right\},$$

where $\Upsilon > 0$ is a parameter of the process. The process also assigns to each selected edge e a weight $w_e := p_e^{-1}$.

Note that since the weight of each edge in ST(G) is inversely proportional to the edge selection probability, the expected weight of each vertex in ST(G) is the same as the degree of that vertex in G.

It was shown in Reference [37, Theorem 6.1] that if $\Upsilon = \Omega(\log^2 n/\phi_G^4)$, then with high probability, ST(G) is an $(1 + \epsilon)$-spectral approximation of graph G, i.e., it has the same Laplacian quadratic form

---

[3]Recall that $\Gamma_G(S) = \bigcup_{u \in S} \Gamma_G(u) - S$.

as G within a factor of $1 + \epsilon$. This implies the following result, stating that $\text{ST}(G)$ approximately preserves cut weights, as well as the conductance of G—the latter within a quadratic exponent.

LEMMA 19 (PROPERTIES OF ST). *Let $0 < \epsilon < 1/2$ and $c \geq 1$, and let $G = (V, E)$ be an n-node graph with conductance ϕ. Suppose that we apply the ST process on G with parameter*

$$\Upsilon \geq \left(\frac{24c \log(3n)}{\epsilon \phi^2} \right)^2,$$

and let $\tilde{G} := \text{ST}(G)$. Then with probability at least $1 - n^{-c}$, the following hold:

(a) *For any $S \subseteq V$,*

$$\frac{1}{1+\epsilon} \cdot |\text{cut}_G(S, V - S)| \leq w_{\tilde{G}}(\text{cut}_{\tilde{G}}(S, V - S)) \leq (1 + \epsilon) \cdot |\text{cut}_G(S, V - S)|. \tag{9}$$

(b) *The weighted conductance $\varphi_{\tilde{G}}$ of $\tilde{G}$ satisfies*

$$\varphi_{\tilde{G}} \geq \frac{\phi_G^2}{4(1+\epsilon)}. \tag{10}$$

Remark 20. As already mentioned, Reference [37, Theorem 6.1] is more general than our Lemma 19, proving that $\tilde{G}$ is an $(1 + \epsilon)$-spectral approximation of G. This immediately implies the weaker notion of $(1 + \epsilon)$-cut approximation stated in Equation (9) [2]. To obtain Equation (10), we use the fact that the $(1 + \epsilon)$-spectral approximation implies that the eigenvalues of the Laplacians of G and $\tilde{G}$ differ by at most a factor of $(1 + \epsilon)$. Then we apply Cheeger's Inequality [7, 11], stating that for any unweighted or weighted graph with (weighted) conductance ϕ the smallest non-zero eigenvalue λ of its normalized Laplacian satisfies $\phi^2/2 \leq \lambda \leq 2\phi$.

Using Lemma 19, we show now that the volume expansion of ST is $\Omega(\phi_G^2)$ w.h.p.

LEMMA 21 (VOLUME EXPANSION OF ST). *Under the same assumptions as in Lemma 19, with probability at least $1 - n^{-c}$, the volume expansion of ST on G is*

$$\beta_{\text{ST},G} \geq \frac{\phi_G^2}{4(1+\epsilon)^3}. \tag{11}$$

PROOF. Lemma 19(a), applied for a singleton set $S = \{u\}$, implies that for any $u \in V$,

$$\frac{1}{1+\epsilon} \cdot w_{\tilde{G}}(u) \leq \deg_G(u) \leq (1 + \epsilon) \cdot w_{\tilde{G}}(u).$$

From this, it follows that for any $T \subseteq V$, the volume of T in G: $\text{vol}_G(T) = \sum_{u \in T} \deg_G(u)$, satisfies

$$\frac{1}{1+\epsilon} \cdot w_{\tilde{G}}(T) \leq \text{vol}_G(T) \leq (1 + \epsilon) \cdot w_{\tilde{G}}(T). \tag{12}$$

We now proceed to derive Equation (11). Applying Equation (12) for $T = \Gamma_{\tilde{G}}(S)$, we obtain

$$\text{vol}_G(\Gamma_{\tilde{G}}(S)) \geq \frac{1}{1+\epsilon} \cdot w_{\tilde{G}}(\Gamma_{\tilde{G}}(S))$$

$$\geq \frac{1}{1+\epsilon} \cdot w_{\tilde{G}}(\text{cut}_{\tilde{G}}(S, V - S)), \tag{13}$$

where the last inequality holds because $\text{cut}_{\tilde{G}}(S, V - S)$ is just a subset of the edges incident to vertices in $\Gamma_{\tilde{G}}(S)$. From the definition of weighted conductance,

$$w_{\tilde{G}}(\text{cut}_{\tilde{G}}(S, V - S)) \geq \varphi_{\tilde{G}} \cdot \min\{w_{\tilde{G}}(S), \, w_{\tilde{G}}(V - S)\}$$

$$\geq \varphi_{\tilde{G}} \cdot \min\left\{\frac{\text{vol}_G(S)}{1 + \epsilon}, \, \frac{\text{vol}_G(V - S)}{1 + \epsilon}\right\}, \quad \text{by Equation (12)}$$

$$\geq \frac{\phi_G^2}{4\,(1 + \epsilon)^2} \cdot \min\{\text{vol}_G(S), \, \text{vol}_G(V - S)\}, \quad \text{by Equation (10).}$$

Substituting that to Equation (13) and using the definition of $\beta_{\text{ST}, G}$, we obtain Equation (11). □

6.3 Relating Sparsification to Rumor Spreading (Proof of Theorem 14)

We consider now the following edge sampling process, PP, which imitates the random choices of nodes in PUSH–PULL.

Definition 22 (Process pp). Let PP be an edge sampling processes which selects edges from a given graph $G = (V, E)$ as follows. Let E_u be the set of all edges $e \in E$ incident to vertex u, and let $\Psi > 0$ be an integer parameter of the process. For each vertex $u \in V$, process PP draws Ψ edges from E_u independently at random with replacement; let F_u be the set of edges drawn for u. The total set of edges selected by PP is the union of all sets F_u, $u \in V$.

Note that graph $\text{PP}(G)$ is unweighted, unlike graph $\text{ST}(G)$.

The next lemma is our main technical claim, showing that w.h.p. the volume expansion of PP is at least equal to that of ST, if $\Psi = \Omega(\Upsilon \cdot \log n)$. The proof shows a coupling of PP and ST under which $\text{ST}(G)$ is a subgraph of $\text{PP}(G)$ w.h.p. For this coupling, we consider a variant of ST, denoted ST_2, which makes two (independent) attempts to select each edge uv: One attempt succeeds with probability $\Upsilon / \deg(u)$, and the other with probability $\Upsilon / \deg(v)$[4] (thus, the total success probability is at least as high as for a single attempt with the probability p_{uv} from Definition 18). We associate the first attempt with vertex u and the second with v. This allows us to couple ST_2 and PP by considering independently for each vertex u, the selections of the two processes: In ST_2, we select each edge incident to u with probability $\Upsilon / \deg(u)$, while in PP we make $\Psi = \Omega(\Upsilon \cdot \log n)$ independent choices selecting random edges incident to u with replacement; let L_u and F_u be the respective sets of selected edges. We are looking for a coupling that ensures $L_u \subseteq F_u$ w.h.p. We observe that given $|L_u| = k$, the set L_u is just a random k-subset of the edges incident to u, and the same is true for F_u, given $|F_u| = k$. It follows that the desired coupling exists, if $|L_u|$ is stochastically dominated by $|F_u|$. We have that $|L_u|$ is distributed like binomial random variable $B(\deg(u), \Upsilon / \deg(u))$, while $|F_u|$ is distributed like Coupon Collector's random variable $C(\deg(u), \Psi)$ (see Lemma 3). We then use Lemma 3 to show the desired stochastic dominance of $B(\deg(u), \Upsilon / \deg(u))$ by $C(\deg(u), \Psi)$.

LEMMA 23 (VOLUME EXPANSION OF PP). *Let $0 < \epsilon < 1/2$ and $c \geq 1$, and let $G = (V, E)$ be an n-node graph with conductance ϕ. Suppose that we apply the PP process on G with parameter*

$$\Psi \geq \left(\frac{24c \log(3n)}{\epsilon \phi^2}\right)^2 \cdot (4c + 8) \ln n. \tag{14}$$

Then with probability at least $1 - 2n^{-c}$, the volume expansion of PP on G is

$$\beta_{\text{PP}, G} \geq \frac{\phi_G^2}{4\,(1 + \epsilon)^3}.$$

---

[4]Formally, the probabilities are $\min\{1, \Upsilon / \deg(u)\}$ and $\min\{1, \Upsilon / \deg(v)\}$, respectively, but for simplicity we omit the "min" in this informal overview.

Proof. We show that there exists a coupling of processes PP and ST, where PP has parameter Ψ satisfying Equation (14) and ST has parameter

$$\Upsilon = \left(\frac{24c \log(3n)}{\epsilon \phi^2} \right)^2,$$

such that with probability at least $1 - n^{-c}$, graph $\tilde{G} := \text{ST}(G)$ is a subgraph of $\hat{G} := \text{PP}(G)$. This implies that with probability at least $1 - n^{-c}$, for any set $S \subseteq V$,

$$\text{vol}_G(\Gamma_{\hat{G}}(S)) \geq \text{vol}_G(\Gamma_{\tilde{G}}(S)).$$

Combining that with Lemma 21, and using a union bound, proves Lemma 23. The rest of the proof demonstrates the coupling we promised above.

Coupling of PP *and* ST: Let ST′ be a process identical to ST except that the probability of selecting edge uv is

$$p'_{uv} := \min \left\{ 1, \frac{\Upsilon}{\max\{\deg(u), \deg(v)\}} \right\}.$$

(Notice that we use $\max\{\deg(u), \deg(v)\}$ instead of min.) Consider now a process ST_2 that executes independently ST and ST′ on G and selects all edges picked by at least one of ST and ST′. We will show a coupling of PP and ST_2 for which $\text{ST}_2(G)$ is a subgraph of PP(G) with probability $1 - n^{-c}$. This immediately implies a coupling of PP and ST with the same property, since there is a trivial coupling of ST_2 and ST for which $\text{ST}(G)$ is a subgraph of $\text{ST}_2(G)$ (with probability 1).

There is a convenient equivalent way to describe ST_2, along the lines of Definition 22 of PP: For each vertex $u \in V$, process ST_2 chooses each edge $e \in E_u$ independently with probability

$$\min \left\{ 1, \frac{\Upsilon}{\deg(u)} \right\};$$

let L_u denote the set of edges chosen for u. Then the total set of edges selected by ST_2 is the union of all L_u, $u \in V$. This description is equivalent to the original definition of ST_2, because each edge $uv \in E$, where $\deg(u) \leq \deg(v)$, is added to L_u with probability p_{uv} and is independently added to L_v with probability p'_{uv}.

Observe that the random sets L_u, $u \in V$, are mutually independent, similar to the sets F_u in the definition of PP. For each $u \in V$, the cardinality of set L_u follows the binomial distribution $B(\deg(u), \min\{1, \frac{\Upsilon}{\deg(u)}\})$. Moreover, for any $0 \leq k \leq \deg(u)$, conditionally on $|L_u| = k$, we have that L_u is a uniformly random k-subset of E_u. For PP on the other hand, we have that $|F_u|$ has the same distribution as $C(\deg(u), \Psi)$, in the statement of Lemma 3, i.e., the number of coupons collected in Ψ trials in the Coupon Collector's process with $\deg(u)$ coupons in total. Similarly to L_u, we have that F_u is a uniformly random k-subset of E_u, given $|F_u| = k$.

We use the following coupling of the sets L_u and F_u. For each $u \in V$, we draw a sample ℓ_u from distribution $B(\deg(u), \min\{1, \frac{\Upsilon}{\deg(u)}\})$ and a sample f_u from the distribution of $C(\deg(u), \Psi)$, and, independently, we choose a random permutation π_u over the edges in E_u.[5] Then we let L_u be the set of the first ℓ_u edges in permutation π_u, and let F_u be the set of the first f_u edges in π_u. From the discussion above, it follows that the marginal distributions of L_u and F_u are the correct ones.

To show that $\text{ST}_2(G)$ is a subgraph of PP(G), i.e., $\bigcup_{u \in V} L_u \subseteq \bigcup_{u \in V} F_u$, it suffices to show that $L_u \subseteq F_u$ for all $u \in V$. And this is equivalent, under the coupling above, to showing that $\ell_u \leq f_u$ for all $u \in V$. We now show that $\ell_u \leq f_u$, for any $u \in V$, w.h.p.

For each vertex $u \in V$, we distinguish two cases depending on $\deg(u)$.

---

[5]Samples ℓ_u and f_u do not need to be independent.

- *Case $\Upsilon/\deg(u) \geq 1/4$:* We apply Equation (1) of Lemma 3, for $k = \deg(u)$ and $t = \Psi$, to get

$$\Pr\left(C(\deg(u), \Psi) < \deg(u)\right) \leq \deg(u) \cdot e^{-\Psi/\deg(u)}$$

$$\leq n \cdot e^{-\Upsilon \cdot (4c+8) \ln n/\deg(u)}$$

$$\leq n \cdot e^{-(4c+8) \ln n/4}, \quad \text{by the case hypothesis}$$

$$= n^{-c-1}.$$

It follows that $\Pr(f_u < \deg(u)) \leq n^{-c-1}$, and since $\ell_u \leq \deg(u)$, we have

$$\Pr(f_u \geq \ell_u) \geq 1 - n^{-c-1}.$$

- *Case $\Upsilon/\deg(u) < 1/4$:* Let

$$\lambda = \min\{\deg(u)/2, \; \Psi/4\}.$$

We apply Equation (2) of Lemma 3, for $k = \deg(u)$, $t = \Psi$ and the above value of λ, to get

$$\Pr\left(C(\deg(u), \Psi) \leq \lambda\right) \leq \Pr\left(B(\Psi, 1 - \lambda/\deg(u)) \leq \lambda\right)$$

$$\leq \Pr\left(B(\Psi, 1/2) \leq \lambda\right)$$

$$\leq \Pr\left(B(\Psi, 1/2) \leq \Psi/4\right).$$

Using Lemma 4(a) to bound the probability in the last line above, gives

$$\Pr\left(C(\deg(u), \Psi) \leq \lambda\right) \leq e^{-2(\Psi/2 - \Psi/4)^2/\Psi} = e^{-\Psi/8}.$$

Therefore,

$$\Pr(f_u \leq \lambda) \leq e^{-\Psi/8}. \tag{15}$$

Next we bound the probability that $\ell_u \geq \lambda$. We have $\lambda \geq 2\Upsilon$, because $2\Upsilon \leq \deg(u)/2$ by the case hypothesis, and $2\Upsilon \leq \Psi/4$ for the chosen values of Υ and Ψ. Then,

$$\Pr\left(B(\deg(u), \Upsilon/\deg(u)) \geq \lambda\right) \leq \Pr\left(B(\deg(u), \Upsilon/\deg(u)) \geq 2\Upsilon\right),$$

$$\leq e^{-\Upsilon/3},$$

by Lemma 4(b). Therefore,

$$\Pr(\ell_u \geq \lambda) \leq e^{-\Upsilon/3}.$$

Combining that with Equation (15), and applying a union bound, gives

$$\Pr(f_u > \lambda > \ell_u) \geq 1 - e^{-\Psi/8} - e^{-\Upsilon/3} > 1 - n^{-c-1}.$$

Therefore, in both cases above it holds

$$\Pr(f_u \geq \ell_u) \geq 1 - n^{-c-1}.$$

Then, by a union bound, the probability that $f_u \geq \ell_u$ for all $u \in V$ is at least $1 - n^{-c}$. This completes the proof of Lemma 23. $\square$

The next claim follows from Lemma 23, and the analogy between the random choices made by PP and PUSH–PULL.

LEMMA 24. *Suppose that a set of nodes $S \subseteq V$ in graph $G = (V, E)$ know a message at a given time. Then after applying r rounds of* PUSH–PULL, *where*

$$r := \left(\frac{24c \log(3n)}{\epsilon \phi^2}\right)^2 \cdot (4c + 8) \ln n, \tag{16}$$

$0 < \epsilon < 1/2$, and $c \geq 1$, the total volume of informed nodes has increased by at least

$$\frac{\phi_G^2}{4(1+\epsilon)^3} \cdot \min\{\mathrm{vol}(S), \mathrm{vol}(V-S)\},$$

with probability at least $1 - 2n^{-c}$.

PROOF. For each $u \in V$, let V_u denote the set of neighbors that u selects during the r rounds of PUSH–PULL. The edge sets $\{uv : v \in V_u\}$, $u \in V$, are distributed identically to the sets F_u selected by PP, when PP is applied to G with parameter $\Psi = r$. It follows that, if S' is the set of vertices $u \in V - S$ that either select or are selected by at east one neighbor from S during PUSH–PULL, then $\mathrm{vol}_G(S')$ has the same distribution as $\mathrm{vol}_G(\Gamma_{\mathrm{PP}(G)}(S))$. By the definition of volume expansion,

$$\mathrm{vol}_G(\Gamma_{\mathrm{PP}(G)}(S)) \geq \beta_{\mathrm{PP},G} \cdot \min\{\mathrm{vol}_G(S), \mathrm{vol}_G(V-S)\},$$

and applying Lemma 23, gives that

$$\mathrm{vol}_G(\Gamma_{\mathrm{PP}(G)}(S)) \geq \frac{\phi_G^2}{4(1+\epsilon)^3} \cdot \min\{\mathrm{vol}_G(S), \mathrm{vol}_G(V-S)\},$$

with probability at least $1 - n^{-c}$. The claim now follows, because all vertices $u \in S'$ get informed during PUSH–PULL, since they communicate with at least one informed neighbor from S. □

We are now ready to prove our main result.

PROOF OF THEOREM 14. In preparation to apply Lemma 24, we divide rumor spreading into phases of r rounds each, where r is as in (16), i.e.,

$$r := \left(\frac{24c\log(3n)}{\epsilon\phi^2}\right)^2 \cdot (4c+8)\ln n.$$

Recall that I_t denotes the set of informed nodes after the first t rounds and $U_t = V - I_t$. Let

$$\rho := \frac{\phi^2}{4(1+\epsilon)^3}.$$

For each $i \geq 1$, let X_i be a binary random variable that is 1 if

$$\mathrm{vol}\left(I_{(i-1)r} - I_{ir}\right) \geq \rho \cdot \min\left\{\mathrm{vol}\left(I_{(i-1)r}\right), \mathrm{vol}\left(U_{(i-1)r}\right)\right\};$$

otherwise, $X_i = 0$. Lemma 24 implies that for each $i \geq 1$, $\Pr(X_i = 1) \geq 1 - 2n^{-c}$. Moreover, it is easy to see that $I_{ir} \neq V$ (i.e., not all nodes are informed after ir rounds) implies

$$\sum_{1 \leq j \leq i} X_j < 2\ln_{1+\rho}(\mathrm{vol}(V)/2).$$

By a union bound, the probability that $X_j = 1$ for all $j \leq i$ is at least $1 - in^{-c}$. It follows that for $i = 2\ln_{1+\rho}(\mathrm{vol}(V)/2)$, we have $I_{ir} = V$ with probability at least

$$1 - \ln_{1+\rho}(\mathrm{vol}(V)/2) \cdot n^{-c} = 1 - \Omega(n^{-c}\log n/\rho) = 1 - \Omega(n^{-c+4}\log n),$$

where for the last equation we used that $\phi = \Omega(1/n^2)$ for any connected n-node graph. It follows that, with the above probability, the number of rounds before all nodes get informed is at most

$$r \cdot 2\ln_{1+\rho}(\mathrm{vol}(V)/2) = O(\log^4 n/\rho\phi^4) = O(\log^4 n/\phi^6). \qquad \square$$

REFERENCES

[1] Yossi Azar, Andrei Z. Broder, Anna R. Karlin, and Eli Upfal. 1999. Balanced allocations. *SIAM J. Comput.* 29, 1 (1999), 180–200.

[2] Joshua Batson, Daniel A. Spielman, Nikhil Srivastava, and Shang-Hua Teng. 2013. Spectral sparsification of graphs: Theory and algorithms. *Commun. ACM* 56, 8 (2013), 87–94.

[3] P. Berenbrink, R. Elsässer, and T. Friedetzky. 2008. Efficient randomised broadcasting in random regular networks with applications in peer-to-peer systems. In *Proceedings of the 27th ACM Symposium on Principles of Distributed Computing (PODC'08)*. 155–164.

[4] S. P. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. 2006. Gossip algorithms: Design, analysis and applications. *IEEE Trans. Inf. Theory* 52, 6 (2006), 1653–1664.

[5] Keren Censor-Hillel, Bernhard Haeupler, Jonathan A. Kelner, and Petar Maymounkov. 2012. Global computation in a poorly connected world: Fast rumor spreading with no dependence on conductance. In *Proceedings of the 44th ACM Symposium on Theory of Computing (STOC'12)*. 961–970.

[6] Keren Censor-Hillel and Hadas Shachnai. 2011. Fast information spreading in graphs with large weak conductance. In *Proceedings of the 22nd ACM-SIAM Symposium on Discrete Algorithms (SODA'11)*. 440–448.

[7] J. Cheeger. 1970. A lower bound for the smallest eigenvalue of the laplacian. *Problems in Analysis*. 195–199.

[8] F. Chierichetti, S. Lattanzi, and A. Panconesi. 2009. Rumor spreading in social networks. In *Proceedings of the 36th International Colloquium on Automata, Languages and Programming (ICALP'09)*. 375–386.

[9] Flavio Chierichetti, Silvio Lattanzi, and Alessandro Panconesi. 2010. Almost tight bounds for rumour spreading with conductance. In *Proceedings of the 42nd ACM Symposium on Theory of Computing (STOC'10)*. 399–408.

[10] Flavio Chierichetti, Silvio Lattanzi, and Alessandro Panconesi. 2010. Rumour spreading and graph conductance. In *Proceedings of the 21st ACM-SIAM Symposium on Discrete Algorithms (SODA'10)*. 1657–1663.

[11] F. R. K. Chung. 1997. *Spectral Graph Theory*. American Mathematical Society.

[12] A. J. Demers, D. H. Greene, C. Hauser, W. Irish, J. Larson, S. Shenker, H. E. Sturgis, D. C. Swinehart, and D. B. Terry. 1987. Epidemic algorithms for replicated database maintenance. In *Proceedings of the 6th ACM Symposium on Principles of Distributed Computing (PODC'87)*.

[13] B. Doerr, M. Fouz, and T. Friedrich. 2011. Social networks spread rumors in sublogarithmic time. In *Proceedings of the 43rd ACM Symposium on Theory of Computing (STOC'11)*. 21–30.

[14] B. Doerr, T. Friedrich, and T. Sauerwald. 2008. Quasirandom rumor spreading. In *Proceedings of the 19th ACM-SIAM Symposium on Discrete Algorithms (SODA'08)*. 1964–1991.

[15] B. Doerr, T. Friedrich, and T. Sauerwald. 2009. Quasirandom rumor spreading: Expanders, push vs. pull, and robustness. In *Proceedings of the 36th International Colloquium on Automata, Languages and Programming (ICALP'09)*. 366–377.

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

[17] R. Elsässer. 2006. On the communication complexity of randomized broadcasting in random-like graphs. In *Proceedings of the 18th ACM Symposium on Parallel Algorithms and Architectures (SPAA'06)*. 148–157.

[18] Uriel Feige, David Peleg, Prabhakar Raghavan, and Eli Upfal. 1990. Randomized broadcast in networks. *Random Struct. Algori.* 1, 4 (1990), 447–460.

[19] Nikolaos Fountoulakis, Konstantinos Panagiotou, and Thomas Sauerwald. 2012. Ultra-fast rumor spreading in social networks. In *Proceedings of the 23rd ACM-SIAM Symposium on Discrete Algorithms (SODA'12)*. 1642–1660.

[20] A. Frieze and G. Grimmett. 1985. The shortest-path problem for graphs with random arc-lengths. *Discr. Appl. Math.* 1, 10 (1985), 57–77.

[21] George Giakkoupis. 2011. Tight bounds for rumor spreading in graphs of a given conductance. In *Proceedings of the 28th International Symposium on Theoretical Aspects of Computer Science (STACS'11)*. 57–68.

[22] George Giakkoupis. 2014. Tight bounds for rumor spreading with vertex expansion. In *Proceedings of the 25th ACM-SIAM Symposium on Discrete Algorithms (SODA'14)*. 801–815.

[23] George Giakkoupis and Thomas Sauerwald. 2012. Rumor spreading and vertex expansion. In *Proceedings of the 23rd ACM-SIAM Symposium on Discrete Algorithms (SODA'12)*. 1623–1641.

[24] George Giakkoupis, Thomas Sauerwald, and Alexandre Stauffer. 2014. Randomized rumor spreading in dynamic graphs. In *Proceedings of the 41st International Colloquium on Automata, Languages and Programming (ICALP'14)*. 495–507.

[25] George Giakkoupis, Thomas Sauerwald, He Sun, and Philipp Woelfel. 2012. Low randomness rumor spreading via hashing. In *Proceedings of the 29th International Symposium on Theoretical Aspects of Computer Science (STACS'12)*. 314–425.

[26] Bernhard Haeupler. 2013. Simple, fast and deterministic gossip and rumor spreading. In *Proceedings of the 24th ACM-SIAM Symposium on Discrete Algorithms (SODA'13)*. 705–716.

[27] Mark Jerrum and Alistair Sinclair. 1989. Approximating the permanent. *SIAM J. Comput.* 18, 6 (1989), 1149–1178.

[28] R. Karp, C. Schindelhauer, S. Shenker, and B. Vöcking. 2000. Randomized rumor spreading. In *Proceedings of the 41st IEEE Symposium on Foundations of Computer Science (FOCS'00)*. 565–574.

[29] J. Leskovec, K. J. Lang, A. Dasgupta, and M. W. Mahoney. 2008. Statistical properties of community structure in large social and information networks. In *Proceedings of the 17th International Conference on World Wide Web (WWW'08)*. 695–704.

[30] Milena Mihail, Christos H. Papadimitriou, and Amin Saberi. 2006. On certain connectivity properties of the Internet topology. *J. Comput. Syst. Sci.* 72, 2 (2006), 239–251.

[31] Michael Mitzenmacher and Eli Upfal. 2005. *Probability and Computing: Randomization and Probabilistic Techniques in Algorithms and Data Analysis*. Cambridge University Press.

[32] D. Mosk-Aoyama and D. Shah. 2008. Fast distributed algorithms for computing separable functions. *IEEE Trans. Inf. Theory* 54, 7 (2008), 2997–3007.

[33] B. Pittel. 1987. On spreading a rumor. *SIAM J. Appl. Math.* 47, 1 (1987), 213–223.

[34] Thomas Sauerwald. 2010. On mixing and edge expansion properties in randomized broadcasting. *Algorithmica* 56, 1 (2010), 51–88.

[35] Thomas Sauerwald and Alexandre Stauffer. 2011. Rumor spreading and vertex expansion on regular graphs. In *Proceedings of the 22nd ACM-SIAM Symposium on Discrete Algorithms (SODA'11)*. 462–475.

[36] D. A. Spielman and S. H. Teng. 2004. Nearly-linear time algorithms for graph partitioning, graph sparsification, and solving linear systems. In *Proceedings of the 36th ACM Symposium on Theory of Computing (STOC'04)*. 81–90.

[37] Daniel A. Spielman and Shang Hua Teng. 2011. Spectral sparsification of graphs. *SIAM J. Comput.* 40, 4 (2011), 981–1025.