How to Spread a Rumor: Call Your Neighbors or Take a Walk?

George Giakkoupis
Inria, Univ Rennes, CNRS, IRISA
Rennes, France

Frederik Mallmann-Trenn
Massachusetts Institute of Technology
Cambridge, MA, US

Hayk Saribekyan
University of Cambridge
Cambridge, UK

ABSTRACT

We study the problem of randomized information dissemination in networks. We compare the now standard PUSH-PULL protocol, with agent-based alternatives where information is disseminated by a collection of agents performing independent random walks. In the VISIT-EXCHANGE protocol, both nodes and agents store information, and each time an agent visits a node, the two exchange all the information they have. In the MEET-EXCHANGE protocol, only the agents store information, and exchange their information with each agent they meet.

We consider the broadcast time of a single piece of information in an n-node graph for the above three protocols, assuming a linear number of agents that start from the stationary distribution. We observe that there are graphs on which the agent-based protocols are significantly faster than PUSH-PULL, and graphs where the converse is true. We attribute the good performance of agent-based algorithms to their inherently fair bandwidth utilization, and conclude that, in certain settings, agent-based information dissemination, separately or in combination with PUSH-PULL, can significantly improve the broadcast time.

The graphs considered above are highly non-regular. Our main technical result is that on any regular graph of at least logarithmic degree, PUSH-PULL and VISIT-EXCHANGE have the same asymptotic broadcast time. The proof uses a novel coupling argument which relates the random choices of vertices in PUSH-PULL with the random walks in VISIT-EXCHANGE. Further, we show that the broadcast time of MEET-EXCHANGE is asymptotically at least as large as the other two's on all regular graphs, and strictly larger on some regular graphs.

As far as we know, this is the first systematic and thorough comparison of the running times of these very natural information dissemination protocols.

CCS CONCEPTS

• **Computing methodologies** → *Distributed algorithms*; • **Mathematics of computing** → *Stochastic processes.*

KEYWORDS

random walks, information dissemination, rumor spreading

PODC '19, July 29-August 2, 2019, Toronto, ON, Canada

ACM ISBN 978-1-4503-6217-7/19/07…$15.00
https://doi.org/10.1145/3293611.3331622

ACM Reference Format:
George Giakkoupis, Frederik Mallmann-Trenn, and Hayk Saribekyan. 2019. How to Spread a Rumor: Call Your Neighbors or Take a Walk?. In *2019 ACM Symposium on Principles of Distributed Computing (PODC '19), July 29-August 2, 2019, Toronto, ON, Canada.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3293611.3331622

1 INTRODUCTION

We investigate the problem of spreading information (or rumors) in a distributed network using randomized communication. The archetypal paradigm solution is the so-called, *randomized rumor spreading* protocol, where each informed node samples a random neighbor in each round, and sends the information to it. This is the PUSH version of rumor spreading, introduced by Demers *et al.* in the 80's [14], as a robust and lightweight protocol for distributed maintenance of replicated databases [14, 21].

The PUSH-PULL variant of rumor spreading, popularized by Karp *et al.* in 2000 [27], allows for bidirectional communication: In each round, every node calls a random neighbor and the two nodes exchange all information they have. PUSH-PULL was initially proposed as a way to reduce the message complexity of PUSH on the complete graph [27]. It was subsequently observed that it is significantly faster than PUSH in several families of graphs, including graph models of social networks [11, 16].

The above two protocols have been studied extensively over the past 15 years, and have also found several applications, including data aggregation [7, 28, 31], resource discovery [24], failure detection [35], and even efficient simulation of arbitrary distributed computations [9].

We compare the above well-established protocols for information spreading, with agent-based alternatives that have received almost no attention so far, even though they have very attractive properties, as we will see. These alternative protocols use a collection of agents performing independent random walks to disseminate information. In the VISIT-EXCHANGE protocol, both nodes and agents store information, and each time an agent visits a node, the two exchange all the information they have. In the MEET-EXCHANGE protocol, only the agents store information, and exchange their information with each agent they meet.

Independent parallel random walks have been studied since the late 70s [1], mainly as a way to speed-up cover and hitting times and related graph problems [2, 8, 19, 20]. As far as we know, VISIT-EXCHANGE has not been studied before. For MEET-EXCHANGE there is some limited previous work. It was studied for specific graph families, namely grids [29, 32] and random graphs [13]. Also, general bounds on the broadcast time of MEET-EXCHANGE with respect to the meeting time were shown [15].

In this paper, we restrict our attention to the case where the number of agents in the network is linear in the number of nodes n, and we assume that all agents start from the stationary distribution.

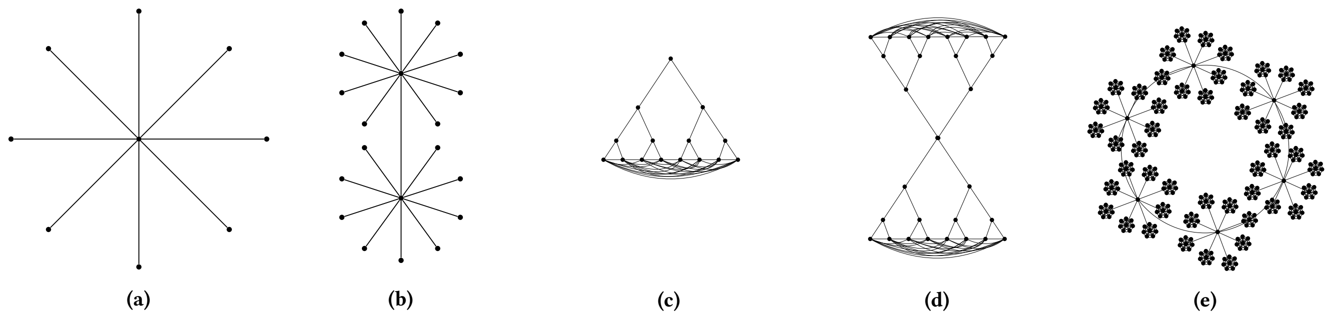

Figure 1 (a) Star S_n, on which $\mathbb{E}\left[T_{\text{push}}\right] = \Omega(n \log n)$ and all other processes take $O(\log n)$ time w.h.p. (b) Double-star S_n^2, on which $\mathbb{E}\left[T_{\text{ppull}}\right] = \Omega(n)$, and $T_{\text{visitx}}, T_{\text{meetx}} = O(\log n)$ w.h.p. (c) Heavy binary tree B_n (leaves are connected to a clique), on which $T_{\text{push}} = O(\log n)$ w.h.p., $\mathbb{E}[T_{\text{visitx}}] = \Omega(n)$, and, for a leaf source, $T_{\text{meetx}} = O(\log n)$ w.h.p. (d) Siamese heavy binary tree D_n, on which $T_{\text{push}} = O(\log n)$ w.h.p., and $\mathbb{E}[T_{\text{visitx}}], \mathbb{E}[T_{\text{meetx}}] = \Omega(n)$. (e) Cycle-of-stars-of-cliques ($n^{1/3}$ stars, $n^{1/3}$ nodes per clique), on which $\mathbb{E}[T_{\text{visitx}}] = O(n^{2/3})$ and $\mathbb{E}[T_{\text{meetx}}] = \Omega(n^{2/3} \log n)$.

Under the assumption that there is a linear number of agents, the agent-based protocols have similar amount of communication as the rumor spreading protocols, both in terms of the (maximum) total number of messages sent per round, which is linear, and the total number of bits. One can think of the agents simply as tokens passed between nodes, along with the actual information (if there is any). Agents need not be labeled, so each node only needs to send a counter of the number of agents in each message.

The assumption that agents start from the stationary distribution makes sense in a setting where several pieces of information (or rumors) are generated frequently and distributed in parallel over time by the same set of agents, which execute perpetual independent random walks. As discussed later, our results for regular graphs hold also in the case where there is exactly one agent starting from each node.

One distinct advantage of the agent-based protocols is their *locally fair* use of bandwidth, i.e., all edges are used with the same frequency, since the random walks are independent and start from stationarity. Interestingly, the superiority of PUSH-PULL over PUSH is commonly attributed to a similar fairness property: that nodes of larger degree contribute more to the dissemination — except that PUSH-PULL satisfies this property only for *some* graph topologies, and *approximately*, as we will see below. In the agent-based protocols, on the other hand, this property is satisfied in a very precise and exact way.

We will see that this fairness property results in a significant performance advantage of VISIT-EXCHANGE and MEET-EXCHANGE over PUSH and PUSH-PULL in certain families of graphs, on which the first two processes need only logarithmic time to spread an information, whereas the other two need polynomial time.

Contribution. We compare the broadcast times of a single piece of information, originated at an arbitrary node s of an n-node graph $G = (V, E)$, when PUSH (or PUSH-PULL), VISIT-EXCHANGE, and MEET-EXCHANGE are used. In the first three, the broadcast time is the time until all vertices are informed, while in MEET-EXCHANGE it is the time until all *agents* are informed. Also, for MEET-EXCHANGE, we assume that the first agent to visit the source s becomes informed, and from that point on, information is exchanged only between agents.[1] As mentioned before, we assume a linear number of agents, each starting from the stationary distribution.

We observe that in general graphs, the broadcast times of the above protocols are *incomparable*: For any pair of protocols, there are examples of graphs where the first protocol is significantly faster than the other, by a polynomial factor in most cases. The examples we use, depicted in Fig. 1, are fairly simple, mainly trees or superpositions of trees with cliques.

The star graph in Fig. 1(a) is an example where PUSH is known to take $\Omega(n \log n)$ rounds, as the center must contact all leaves. VISIT-EXCHANGE and MEET-EXCHANGE, on the other hand, take only logarithmic time, as roughly half of the walks visit the center in each round, and a constant number visits each leaf on average.

In the star, PUSH-PULL is also (extremely) fast. The next example, the double-star in Fig. 1(b), is a graph where PUSH-PULL (and thus also PUSH) is slow, whereas VISIT-EXCHANGE and MEET-EXCHANGE are still fast. This demonstrates the advantages of the local fairness property we pointed out earlier, and the impact it can have on the broadcast time: Here PUSH-PULL selects the edge between the two stars only with probability $O(1/n)$, which results in an expected broadcast time of $\Omega(n)$. In VISIT-EXCHANGE and MEET-EXCHANGE, on the other hand, the probability that some agent crosses the edge in a round is constant, resulting in a logarithmic broadcast time.

Fig. 1(c) and Fig. 1(d) illustrate examples where rumor spreading protocols have an advantage over agent-based protocols. In both examples PUSH (and thus PUSH-PULL) has logarithmic broadcast time. For VISIT-EXCHANGE, at least linear time is needed: Since almost all the volume of the graph is concentrated on the leaves, it is likely that all agents are on the leaves at time zero, and then it takes linear time before the first walk reaches the root. For MEET-EXCHANGE, we have that it is fast in the first example, as all walks meet quickly in the clique induced by the leaves. However, in the second example, where agents are roughly split between the two induced cliques, the broadcast times of both MEET-EXCHANGE and VISIT-EXCHANGE is $\Omega(n)$.

[1] This is a technicality used to allow for direct comparison between the protocols, and has limited effect on our results.

The above results suggest that in certain settings, agent-based information dissemination, separately or in combination with PUSH-PULL, can significantly improve the broadcast time. We stress that, even though the examples presented may seem contrived, they are intentionally simple to demonstrate the principle reasons that make the protocols perform differently, and we expect that similar result can be observed in a wide range of networks. In particular, we believe that the observations for the double-star example of Fig. 1(b), extend to more general tree-like topologies with high-degree internal nodes.

All examples we have discussed so far, involve highly non-regular graphs. Our main technical result concerns regular graphs, and can be stated somewhat informally as follows. (For the formal, stronger statements see Sect. 4 and 5.)

THEOREM 1. *For any d-regular graph on n vertices, where $d = \Omega(\log n)$, and any source vertex, the broadcast times of PUSH and VISIT-EXCHANGE are asymptotically the same both in expectation and w.h.p.,[2] modulo constant multiplicative factors.*

Recall that PUSH and PUSH-PULL have asymptotically the same broadcast times on regular graphs [23]. Note also that the broadcast times of PUSH and PUSH-PULL on d-regular graphs can vary from logarithmic, e.g., in random d-regular graphs, to polynomial, e.g., in a path of d-cliques where the broadcast time is $\Omega(n)$.

The proof of Theorem 1 uses a novel coupling argument which relates the random choices of vertices in PUSH, with the random walks in VISIT-EXCHANGE. Roughly speaking, for each node u, we consider the list of neighbors that u samples in PUSH, and the list of neighbors to which informed agents move to in their next step after visiting u in VISIT-EXCHANGE. Our coupling just sets the two lists to be identical for each u. Even though the coupling is straightforward, its analysis is not. On the one direction of the proof, showing that the broadcast time of PUSH is dominated by the broadcast time of VISIT-EXCHANGE, the main step is to bound the congestion, i.e., the number of agents encountered along a path, for all possible paths through which information travels. On the reverse direction, we focus only on the fastest path through which information reaches each node in PUSH, and show that an equally fast path exists in VISIT-EXCHANGE. A useful trick we devise, to consider only every other round of VISIT-EXCHANGE in the coupling, simplifies the proof of this second direction. We expect that our proof ideas will be useful in other applications of multiple random walks as well.

In addition to Theorem 1, we observe that the broadcast time of MEET-EXCHANGE is asymptotically at least as large as VISIT-EXCHANGE's on any regular graph of at least logarithmic degree. The idea is that once all agents are informed it takes at most logarithmic time to cover the graph.

It is probably surprising that the converse direction is not true, i.e., there are regular graphs where MEET-EXCHANGE is strictly slower than VISIT-EXCHANGE. Fig. 1(e) presents one such example of a d-regular graph, where $d = n^{1/3}$, for which a logarithmic-factor gap exists between the broadcast times of the two protocols.

[2]By with high probability (w.h.p.) we mean with probability at least $1 - n^{-c}$, with some constant $c > 0$ that can be made arbitrary large, by adjusting the constants in the statement.

Road-map. In Sect. 2, we survey additional related work. In Sect. 3, we provide a formal description of the protocols we study. The first direction of Theorem 1, that PUSH is at least as fast as VISIT-EXCHANGE, is proved in Sect. 4; the other direction is proved in Sect. 5. The result that VISIT-EXCHANGE is at least as fast as MEET-EXCHANGE on regular graphs is provided in Sect. 6. Finally, some open problems are discussed in Sect. 7.

Due to space limitations, several proofs, including the analysis of the broadcast times of the example graphs in Fig. 1, are only available in the full version of the paper.

2 RELATED WORK

The PUSH variant of rumor spreading was first considered in [14]. It was subsequently analyzed on various graphs in [21], where also bounds with the degree and diameter were shown for general graphs. The PUSH-PULL variant was introduced in [27], and was studied initially on the complete graph. More recently, there has been a lot of work on showing that in several settings $O(\log n)$ rounds of rumor spreading suffice w.h.p. to broadcast information [4, 17, 18]. In addition, general bounds in terms of expansion parameters of the graph have been studied extensively, e.g., in [10, 22].

Another line of work compares synchronous and asynchronous versions of rumor spreading, where in the latter each node takes steps at the arrival times of an independent unit-rate Poisson process. In [34], it is shown that the asynchronous version of PUSH has the same broadcast time as standard PUSH on regular graphs. In [3, 23], tight bounds are given for the relation between the broadcast times of synchronous and asynchronous PUSH-PULL.

On the random walk literature, there has been some previous work on models related to MEET-EXCHANGE, motivated mainly by the study of the spread of infectious diseases. The earliest work considering a process equivalent to MEET-EXCHANGE is [15], which studies general graphs. It shows that the broadcast time of MEET-EXCHANGE is at most $O(\log n)$ times larger than the meeting time of two random walks in the graph, and that this upper bound is tight. Later, the authors of [13] studied MEET-EXCHANGE for the case of random regular graphs and $k \leqslant n^{\varepsilon}$ random walks. They showed that the expected broadcast time is $O(n \log k/k)$. In [32], the 2-dimensional grid was studied and a broadcast time of $\tilde{\Theta}(n/\sqrt{k})$ was shown for k random walks. This work was extended to d-dimensional grids in [29], where a tight lower bound up to a polylogarithmic factor was also shown.

A random process related to MEET-EXCHANGE is the frog model, where only one of the agents is active initially while the remaining agents are inactive. When an inactive agent is hit by an active one, it is activated and starts its own independent random walk. The model has been considered mostly on the infinite grid, where questions about the asymptotic shape of the set of active agents have been studied [25, 33].

Other superficially related processes include coalescing random walks [5, 26], and coalescing branching walks [6, 30]. See also [12] for a survey on multiple random walks.

3 PROTOCOL DESCRIPTIONS

We compare four information spreading protocols. The first two, PUSH and PUSH-PULL, are standard versions of randomized rumor spreading. The other two, VISIT-EXCHANGE and MEET-EXCHANGE, use a system of interacting agents performing independent random walks, and are less standardized. In PUSH and PUSH-PULL, information is communicated between adjacent vertices, whereas in VISIT-EXCHANGE and MEET-EXCHANGE information is passed between an agent and a vertex it visits, or between two agents when their random walks meet. All protocols proceed in a sequence of synchronous rounds. They are applied on a connected undirected graph $G = (V, E)$ with $|V| = n$ vertices, and the information originates from an arbitrary source vertex $s \in V$.

Push. In round zero, vertex s becomes informed. In each round $t \geqslant 1$, every vertex u that was informed in a previous round samples a random neighbor v to send the information to, and if v is not already informed, it becomes informed in this round. We denote by $T_{\text{push}}(G, s)$ the number of rounds before all vertices are informed.

Push-Pull. Similar to PUSH, vertex s is informed in round zero. In each round $t \geqslant 1$, every vertex $u \in V$ (informed or not) samples a random neighbor v to exchange information with, and if exactly one of u and v was informed before round t, then the other vertex becomes informed as well. The number of rounds before all vertices are informed is denoted $T_{\text{ppull}}(G, s)$.

Visit-Exchange. Let A be a set of *agents*. Every agent $g \in A$ performs an independent simple random walk on G, starting from a vertex sampled independently from the stationary distribution (i.e., each vertex v is sampled with probability $\deg(v)/(2|E|)$). In round zero, vertex s becomes informed, and every agent that is on vertex s becomes informed as well. In each subsequent round $t \geqslant 1$, all agents do a single step of their random walk in parallel. If an agent g, which was informed in an earlier round, visits vertex v that is not yet informed, then v becomes informed in this round. If, on the other hand, some not yet informed agent g visits vertex v which was informed either in an earlier round or the current one (by some other informed agent), then g becomes informed. We denote by $T_{\text{visitx}}(G, s)$ the number of rounds before all *vertices* (and thus all agents) are informed.

Meet-Exchange. Similar to VISIT-EXCHANGE, we have a set A of agents that perform independent random walks starting from the stationary distribution. In round zero, all agents that are on vertex s become informed. If there is no agent on s in round zero, then the first agent to visit s after round zero becomes informed (if more than one agents visit s simultaneously, they all get informed). After this point, vertex s does not inform agents visiting it. In each subsequent round t, whenever two agent g, g' meet and exactly one of them was informed in a previous round, the other agent becomes informed as well. We denote by $T_{\text{meetx}}(G, s)$ the number of rounds before all agents are informed. Note that if G is a bipartite graph, then $T_{\text{meetx}}(G, s)$ can be infinite, and, in particular, we have $\mathbb{E}[T_{\text{meetx}}(G, s)] = \infty$. To avoid this complication we will sometimes assume that the random walks of the agents are lazy, i.e., a walk stays put in a round with probability $1/2$. This ensures that $\mathbb{E}[T_{\text{meetx}}(G, s)] < \infty$.

We will collectively refer to $T_{\text{push}}(G, s)$, $T_{\text{ppull}}(G, s)$, $T_{\text{visitx}}(G, s)$, and $T_{\text{meetx}}(G, s)$ as the *broadcast time* of the corresponding protocol. We will sometimes omit graph G and source vertex s in this notation, when they are clear from the context.

4 BOUNDING T_{push} BY T_{visitx} ON REGULAR GRAPHS

In this section, we prove the following theorem, which upper bounds the broadcast time of PUSH in a regular graph by the broadcast time of VISIT-EXCHANGE.

THEOREM 2. *For any constants $\varepsilon, \gamma, \lambda > 0$, there is a constant $c > 0$, such that for any d-regular graph $G = (V, E)$ with $|V| = n$ and $d \geqslant \varepsilon \log n$, and for any source vertex $s \in V$, the broadcast times of PUSH and VISIT-EXCHANGE, with $|A| \leqslant \gamma n$ agents, satisfy*

$$\mathbb{P}\left[T_{\text{push}} \leqslant ck\right] \geqslant \mathbb{P}\left[T_{\text{visitx}} \leqslant k\right] - n^{-\lambda},$$

for any $k \geqslant 0$.

From Theorem 2, it is immediate that if $T_{\text{visitx}} \leqslant T$ w.h.p., then $T_{\text{push}} = O(T)$ w.h.p. Moreover, using Theorem 2 and the known $O(n \log n)$ upper bound on T_{push} which holds w.h.p. [21], one can easily show that $\mathbb{E}\left[\,T_{\text{push}}\,\right] = O(\mathbb{E}\left[\,T_{\text{visitx}}\,\right])$.

Proof Overview of Theorem 2. The proof uses the following coupling of processes PUSH and VISIT-EXCHANGE: For each vertex u, let $\langle \pi_u(1), \pi_u(2), \ldots \rangle$ be the sequence of neighbors that u samples in PUSH. Similarly, for VISIT-EXCHANGE, consider all transitions of informed agents from u to neighbor vertices in chronological order (ordering transitions in the same round by, say, agent ID), and let $\langle p_u(1), p_u(2), \ldots \rangle$ be the destination vertices in those transitions. We couple the two processes by setting $\pi_u(i) = p_u(i)$, for all u, i.

The intuition behind this coupling is that in VISIT-EXCHANGE, at most a constant number of agents *in expectation* visits each vertex u in a round (since the graph is regular and $|A| = O(n)$), and thus the same number of agents leaves u per round in expectation. The coupling ensures that for each informed agent that moves from u to a neighbor v, vertex u samples the same neighbor v in PUSH. Thus, if we had a constant upper bound c on the *actual* number (rather than the expected number) of visits to each vertex on each round, then the coupling would immediately yield $T_{\text{push}} \leqslant c \cdot T_{\text{visitx}}$ for the coupled processes. In reality, however, a super-constant number of agents may visit a vertex in a round, and, moreover, the number of visits depends on the past history of the process.

An idea we use to tackle dependencies is to consider a tweaked version of VISIT-EXCHANGE, called T-VISIT-EXCHANGE. The only difference between this process and VISIT-EXCHANGE, is that it arbitrarily removes some agents after each round to ensure that the neighborhood of any vertex contains at most $O(d)$ agents. For $d = \Omega(\log n)$ and $|A| = O(n)$, we have that in the first poly(n) rounds the two processes are identical w.h.p. Therefore, we can consider T-VISIT-EXCHANGE in our proofs. The benefit of that is that since the neighborhood of any vertex u contains $O(d)$ agents in round t, at round $t + 1$ the number of agents that visit u will be bounded by binomial distribution $\text{Bin}(\Theta(d), 1/d)$, *independently of the past.*

To prove the theorem is suffices to show that under our coupling, we have w.h.p. (precisely, with probability at least $1 - n^{-\lambda}$) that if $T_{\text{visitx}} \leqslant k$ then $T_{\text{push}} \leqslant ck$. Further, it suffices to assume that k is at least $\Omega(\log n)$; for $k = o(\log n)$ the theorem follows by showing that $T_{\text{visitx}} = \Omega(\log n)$ w.h.p.

To show that w.h.p. $T_{\text{visitx}} \leqslant k$ implies $T_{\text{push}} \leqslant ck$, we consider all possible paths of length k through which information travels in VISIT-EXCHANGE, and for each path we count the total number of (non-distinct) agents encountered along this path, called the congestion of the path. Formally, we use the notion of a *canonical walk* θ, which is represented by a sequence of vertices $(\theta_0, \theta_1, \ldots, \theta_k)$ starting from $\theta_0 = s$: In each round $1 \leqslant t \leqslant k$, the walk either stays put and $\theta_t = \theta_{t-1}$, or it follows one of the agents g that leave θ_{t-1} in round t, and, in that case, θ_t is the new vertex that g moves to. For any round t, we count the agents that are in θ_t. The sum of these counts, for $0 \leqslant t < k$ is the *congestion* $Q(\theta)$ of the walk.

The congestion of a canonical walk is used to bound the time needed for information to travel along the *same* path in PUSH. Intuitively, larger congestion implies longer travel time for PUSH, for the following reason. Suppose there are m agents in u at some round after it is informed by VISIT-EXCHANGE. The coupled PUSH process, using the same random decisions for the choice of neighbors as VISIT-EXCHANGE, will take m rounds to "go through" them.

To relate the congestion of canonical walks with the time it takes for information to spread in PUSH, we introduce *C-counters*: For each vertex u, we maintain a counter C_u. The counter is initialized in the round t_u in which u becomes informed in VISIT-EXCHANGE. Its initial value is the value of the C-counter of the neighbor from which the first informed agent arrived to u. In each subsequent round $t > t_u$, C_u increases by the number of agents that visited u in round $t-1$. C-counters have the following two properties: If τ_u is the round when u gets informed in PUSH then $\tau_u \leqslant C_u(t_u)$; and for any $t \geqslant t_u$, there is a canonical walk θ of length t such that $C_u(t) = Q(\theta)$. Therefore, to show that w.h.p. $T_{\text{visitx}} \leqslant k$ implies $T_{\text{push}} \leqslant ck$, it suffices to show that the maximum congestion of *all* canonical walks of length k is at most ck w.h.p.

We can bound the congestion of a *single* canonical walk of length k using the property of T-VISIT-EXCHANGE, that the number of agents at a node is bounded by a binomial distribution with constant mean. This results in the desired bound of ck for a single walk with probability at least $1 - a^{-k}$, for some constant $a > 1$. We would like to take a union bound over all canonical walks, which would give the desired result. For this to work, however, we should also bound the total number of canonical walks of length k by at most a^k/n^λ.

We bound the number of canonical walks of length k by introducing a set of *descriptors* for these walks. A descriptor is represented by a matrix, which, together with a given execution of VISIT-EXCHANGE, uniquely defines a walk. Additionally, the set of descriptors suffices to encode all canonical walks, and therefore, it is at least as large as the set of all walks. Thus, we can use a bound on the number of descriptors that can be computed by a simple combinatorial argument involving the number of elements used in the matrix, and the values they can take. A naive construction of descriptors, however, is too wasteful giving us a much larger bound than the a^k/n^λ we need. A key idea here is that the majority of the descriptors represent walks only in executions that happen with low probability. So, we construct a set of *concise descriptors* that can describe all canonical walks in a random execution *w.h.p.* We show that the size of the set of concise descriptors can be bounded by a^k/n^λ, as desired. Next we give the detailed proof.

4.1 Notation and Coupling Description

For each vertex $u \in V$, we denote by τ_u the round when u gets informed in PUSH. By $\pi_u(i)$, for $i \geqslant 1$, we denote the ith vertex that u samples, i.e., the vertex it samples in round $\tau_u + i$. Note that $\pi_u(i)$ gets informed in round $\tau_u + i$, if it is not already informed. In VISIT-EXCHANGE, we denote by t_u the round when vertex u gets informed. For any agent $g \in A$ and $t \geqslant 0$, we denote by $x_g(t)$, the vertex that g visits in round t. Thus, $\{x_g(t)\}_{t\geqslant 0}$ is a random walk on G. Let $Z_u(t)$ be the set of all agents that visit u in round t, i.e.,

$$Z_u(t) = \{g \in A \colon x_g(t) = u\}.$$

Equivalently, $Z_u(t)$ is the set of agents that depart from u in round $t+1$. Consider all visits to u in rounds $t \geqslant t_u$, in chronological order, ordering visits in the same round with respect to a predefined total order over all agents. For each $i \geqslant 1$, consider the agent g that does the ith one of those visits, and let $p_u(i)$ be the vertex that g visits next. Formally, let $X_u = \{(t, g) \colon t \geqslant t_u, x_g(t) = u\}$, and order its elements such that $(t, g) < (t', g')$ if $t < t'$, or $t = t'$ and $g < g'$. If (t, g) is the ith smallest element in X_u, then $p_u(i) = x_g(t+1)$.

Coupling. We couple processes PUSH and VISIT-EXCHANGE by setting $\pi_u(i) = p_u(i)$. Formally, let $\{w_u(i)\}_{u\in V, i\geqslant 1}$, be a collection of independent random variables, where $w_u(i)$ takes a uniformly random value from the set $\Gamma(u)$ of u's neighbors. Then, for every $u \in V$ and $i \geqslant 1$, we set $\pi_u(i) = p_u(i) = w_u(i)$.

4.2 Upper Bound on Agents and Tweaked Visit-Exchange

We will use the next simple bound on the number of agents that visit a given set S of vertices in some round t of VISIT-EXCHANGE. The proof is by a simple Chernoff bound, and relies on the assumption that agents execute independent walks starting from stationarity.

LEMMA 3. *For any $S \subseteq V$, $t \geqslant 0$, and $\beta \geqslant 2e \cdot |A|/n$,*

$$\mathbb{P}\left[\sum_{v\in S} |Z_v(t)| \leqslant \beta \cdot |S|\right] \geqslant 1 - 2^{-\beta\,|S|}.$$

We remark that the same bound holds also in the case where $|A| = n$, and exactly one walk starts from each vertex. This implies that Theorem 2 holds for that initial setting as well (the rest of the proof does not need any modifications).

In parts of the analysis, we will use a tweaked variant of VISIT-EXCHANGE, called T-VISIT-EXCHANGE, defined as follows. Let

$$\alpha \geqslant 2e \cdot |A|/n \tag{1}$$

be a (sufficiently large) constant to be specified later. If in some round $t \geqslant 0$, there is a vertex $u \in V$ for which the following condition is *not* true:

$$\sum_{v\in\Gamma(u)} |Z_v(t)| \leqslant \alpha \cdot d, \tag{2}$$

then before round $t+1$, we remove a minimal set of agents from the system in such a way that the above condition holds for all vertices u, when counting just the remaining agents.

It is an immediate corollary of Lemma 3 that the modified process is identical to the original one in the first polynomial number of rounds, if constant α is large enough, and $d = \Omega(\log n)$.

LEMMA 4. *The probability that no agent is removed in any of the first k rounds of T-VISIT-EXCHANGE is at least $1 - kn \cdot 2^{-\alpha d}$.*

PROOF. The claim follows by applying Lemma 3, for each $0 \leqslant t < k$ and each pair u, S, where $u \in V$ and $S = \Gamma(u)$, and then combining the results using a union bound. □

We will use the same definitions and notations for both VISIT-EXCHANGE and T-VISIT-EXCHANGE.

4.3 C-Counters

For each $u \in V$, let S_u be the set of neighbors v of u such that v was informed before u in VISIT-EXCHANGE, and some (informed) agent moved from v to u in round t_u, i.e.,

$$S_u = \{v \in \Gamma(u)\colon t_v < t_u,\ Z_v(t_u - 1) \cap Z_u(t_u) \neq \emptyset\}.$$

Thus, S_u contains those neighbors that "informed" u. For each $u \in V$ and $t \geqslant 0$, let

$$C_u(t) = \begin{cases} 0, & \text{if } t < t_u \text{ or } t = t_u = 0; \\ \min_{v \in S_u} C_v(t), & \text{if } t = t_u > 0; \\ C_u(t-1) + |Z_u(t-1)|, & \text{if } t > t_u. \end{cases} \quad (3)$$

That is, C_u is initialized in round t_u to the minimum counter value of the neighbors that informed u (or to zero if $u = s$), and $C_u(t) - C_u(t_u)$ is the number of visits of agents to u from round t_u until round $t-1$, or equivalently, the number of departures of agents from u in rounds $t_u + 1$ up to t.

LEMMA 5. *For any $u \in V$, $\tau_u \leqslant C_u(t_u)$.*

PROOF. Consider the following path through which information reaches u in VISIT-EXCHANGE. The path is $(v_0, v_1, \ldots, v_k)$, where $v_0 = s$, $v_k = u$, and for $0 < j \leqslant k$, $v_{j-1} \in S_{v_j}$ and $C_{v_j}(t_{v_j}) = C_{v_{j-1}}(t_{v_j})$. It is easy to verify that such a path exists. In the following, we prove by induction on $0 \leqslant j \leqslant k$ that

$$\tau_{v_j} \leqslant C_{v_j}(t_{v_j}). \quad (4)$$

This holds for $j = 0$, because $v_0 = s$, $t_s = 0$, and $\tau_s = 0 = C_s(0)$. Let $0 < j \leqslant k$, and suppose that $\tau_{v_{j-1}} \leqslant C_{v_{j-1}}(t_{v_{j-1}})$; we will show that $\tau_{v_j} \leqslant C_{v_j}(t_{v_j})$. We have that

$$\begin{aligned} C_{v_j}(t_{v_j}) &= C_{v_{j-1}}(t_{v_j}), \quad \text{by the path property} \\ &= C_{v_{j-1}}(t_{v_{j-1}}) + \sum_{t_{v_{j-1}} \leqslant t < t_{v_j}} |Z_{v_{j-1}}(t)|, \\ &\qquad \text{by recursive application of (3)} \\ &\geqslant \tau_{v_{j-1}} + \sum_{t_{v_{j-1}} \leqslant t < t_{v_j}} |Z_{v_{j-1}}(t)|, \quad \text{by induct. hypothesis.} \end{aligned}$$

Let $\ell = \min\{i\colon p_{v_{j-1}}(i) = v_j\}$, let g be the agent that does the ℓth visit to v_{j-1} since round $t_{v_{j-1}}$, and let r be the round when that visit takes place, thus $x_g(r) = v_{j-1}$ and $x_g(r+1) = v_j$. By the minimality of ℓ, $r+1$ is the first round when some informed agent moves to v_j from v_{j-1}. Since $v_{j-1} \in S_{v_j}$, it follows that $r + 1 = t_{v_j}$. Then

$$\ell \leqslant \sum_{t_{v_{j-1}} \leqslant t \leqslant r} |Z_{v_{j-1}}(t)| = \sum_{t_{v_{j-1}} \leqslant t < t_{v_j}} |Z_{v_{j-1}}(t)|.$$

Also, from the coupling, $\pi_{v_{j-1}}(\ell) = p_{v_{j-1}}(\ell) = v_j$, which implies

$$\tau_{v_j} \leqslant \tau_{v_{j-1}} + \ell.$$

Combining all the above we obtain $C_{v_j}(t_{v_j}) \geqslant \tau_{v_{j-1}} + \ell \geqslant \tau_{v_j}$, completing the inductive proof of (4). Applying (4) for $j = k$ we obtain $\tau_u \leqslant C_u(t_u)$. □

4.4 Canonical Walks and Congestion

Let $\theta = (\theta_0, \theta_1, \ldots, \theta_k)$ be a walk on G, where $\theta_0 = s$, and $\theta_i \in \Gamma(\theta_{i-1}) \cup \{\theta_{i-1}\}$, for $1 \leqslant i \leqslant k$. We construct θ from VISIT-EXCHANGE as follows. We start from vertex $\theta_0 = s$ in round zero, and in each round $1 \leqslant t \leqslant k$, we either stay put, in which case $\theta_t = \theta_{t-1}$, or we choose one of the agents $g \in Z_{\theta_{i-1}}(t-1)$, which visited θ_{i-1} in the previous round, and move to the same vertex as g in round t, i.e., $\theta_t = x_t(g)$. We call θ a *canonical walk* of length k. A *labeled canonical walk* is a canonical walk that specifies also the agent g_t that the walk follows in each step t, if $\theta_t \neq \theta_{t-1}$. Formally, a labeled canonical walk corresponding to θ is $\eta = (\theta_0, g_1, \theta_1, g_2, \ldots, g_k, \theta_k)$, where $g_t \in Z_{\theta_{t-1}}(t-1) \cap Z_{\theta_t}(t)$ if $\theta_t \neq \theta_{t-1}$, and $g_t = \perp$ if $\theta_t = \theta_{t-1}$. Note that different labeled canonical walks may correspond to the same (unlabeled) canonical walk. We define the *congestion* $Q(\theta)$ of a canonical walk θ as the total number of (non-distinct) agents encountered along the walk, not counting the last step, i.e.,

$$Q(\theta) = \sum_{0 \leqslant t < k} |Z_{\theta_t}(t)|.$$

The congestion of a *labeled* canonical walk is the same as the congestion of the corresponding unlabeled walk. We have the following simple connection between canonical walks and C-counters.

LEMMA 6. *For any $u \in V$ and $t \geqslant t_u$, there is a canonical walk θ of length t with $Q(\theta) = C_u(t)$.*

PROOF. We consider the same path $(v_0, v_1, \ldots, v_k)$ as in the proof of Lemma 5, where $v_0 = s$, $v_k = u$, and for $0 < j \leqslant k$, $v_{j-1} \in S_{v_j}$ and $C_{v_j}(t_{v_j}) = C_{v_{j-1}}(t_{v_j})$. Consider the canonical walk θ obtained from this path by adding between each pair of consecutive vertices v_{j-1} and v_j, $t_{v_j} - t_{v_{j-1}} - 1$ copies of v_{j-1}, and also appending after v_k a number of $t - t_{v_k}$ copies of v_k. It is then easy to show by induction that $Q(\theta) = C_u(t)$. □

4.5 Concise Descriptors of Canonical Walks

In this section, we bound the number of distinct labeled canonical walks of a given length k. For that, we present a concise description for such walks, and bound the total number of the walks by the total number of different possible descriptions.

We start with a rather wasteful way to describe labeled canonical walks of length k, which we then refine in two steps. Let $\mathcal{A}_k$ denote the set of all $n \times k$ matrices $A_k = [a_{i,j}]$, where $a_{i,j} \in \{0, \ldots, i\}$. Let us fix the first k rounds of VISIT-EXCHANGE, and consider a labeled canonical walk $\eta = (\theta_0 = s, g_1, \theta_1, g_2, \ldots, g_k, \theta_k)$. For each $1 \leqslant t \leqslant k$, let

$$\delta_t = |Z_{\theta_{t-1}}(t-1)|$$

be the number of agents that visit θ_{t-1} in round $t-1$, and thus also the number of agents that depart from θ_{t-1} in round t. Let $\rho_t = 0$ if $g_t = \perp$, otherwise, ρ_t is equal to the *rank* of g_t in set $Z_{\theta_{t-1}}(t-1)$. We describe walk η by a matrix $A_k \in \mathcal{A}_k$ with the following entries: For each $1 \leqslant t \leqslant k$, if $\delta_t > 0$, then $a_{\delta_t, j} = \rho_t$, for

$j = |\{t' \leqslant t \colon \delta_{t'} = \delta_t\}|$; i.e., value ρ_t is stored in the first unused entry of row $A_k[\delta_t, \cdot]$. At most k of the entries of A_k are specified that way; the remaining entries can have arbitrary values. We call A_k a *non-concise descriptor* of η.

For any realization of VISIT-EXCHANGE, each $A_k \in \mathcal{A}_k$ describes exactly one labeled canonical walk of length k, and any labeled canonical walk of length k has at least one non-concise descriptor $A_k \in \mathcal{A}_k$ (in fact, several ones). The total number of different non-concise descriptors is $|\mathcal{A}_k| = \prod_{1 \leqslant i \leqslant n}(i+1)^k$, which is too large for our purposes.

A simple improvement is to use only entries in rows $A_k[i, \cdot]$ for which i is a power of 2 (we assume w.l.o.g. that n is also a power of 2). Roughly speaking, if δ_t is between $2^{\ell-1}$ and 2^ℓ then ρ_t is stored in raw $A_k[2^\ell, \cdot]$. Formally, let b be a large enough constant (to be specified later) that is a power of 2. The matrix $A_k \in \mathcal{A}_k$ we use to describes η has the following entries. For each $1 \leqslant t \leqslant k$:

(1) If $2^{\ell-1} < \delta_t \leqslant 2^\ell$, where $\ell \in \{1 + \log b, \ldots, \log n\}$, and $|\{t' \leqslant t \colon 2^{\ell-1} < \delta_{t'} \leqslant 2^\ell\}| = j$, then
   (a) if $\rho_t \neq 0$, $a_{2^\ell, j} = \rho_t$;
   (b) if $\rho_t = 0$, $a_{2^\ell, j}$ can take any value in $\{0\} \cup \{\delta_t + 1, \ldots, 2^\ell\}$;
(2) If $0 \leqslant \delta_t \leqslant b$ and $|\{t' \leqslant t \colon 0 < \delta_{t'} \leqslant b\}| = j$, then
   (a) if $\rho_t \neq 0$, $a_{b,j} = \rho_t$;
   (b) if $\rho_t = 0$, $a_{b,j}$ can take any value in $\{0\} \cup \{\delta_t + 1, \ldots, b\}$.

The purpose of subcases (b) is to maintain the convenient property that every A_k describes a labeled canonical walk, which would not be the case if we just set $a_{2^\ell, j} = 0$ or $a_{b,j} = 0$, since values greater than δ_t would not correspond to a walk. We call the above A_k a *semi-concise descriptor* of η.

A second modification we make is based on the observation that, in the $\log n$ rows of A_k used in the above scheme, most entries are very unlikely to be actually used. So, for each $i = 2^\ell$, we specify a threshold index $k_i \leqslant k$, such that the first k_i entries in each row $A_k[i, \cdot]$ suffice w.h.p. to describe all labeled canonical walks of length k. Let $\mathcal{B}_k$ be a subset of $\mathcal{A}_k$ defined as follows. Let

$$k_i = b \cdot k/i,$$

and recall that b is a constant power of 2. The set $\mathcal{B}_k$ consists then of all $A_k = [a_{i,j}] \in \mathcal{A}_k$ for which

$$\begin{aligned} &a_{i,j} \in \{0, \ldots, i\}, && \text{if } i \in \{2^\ell \colon \log b \leqslant \ell \leqslant \log n\}, \text{ and } j \leqslant k_i; \\ &a_{i,j} = 0, && \text{otherwise.} \end{aligned}$$

A *concise descriptor* of a labeled canonical walk η of length k is any semi-concise descriptor A_k of η that belongs to set $\mathcal{B}_k$.

Next we establish the following upper bound on the number of all possible concise descriptors of length k.

LEMMA 7. $|\mathcal{B}_k| \leqslant (4b)^{2k}$.

PROOF. From the definition of $\mathcal{B}_k$, we have

$$\begin{aligned} |\mathcal{B}_k| &\leqslant \prod_{\log b \leqslant \ell \leqslant \log n} (2^\ell + 1)^{bk/2^\ell} \\ &= \prod_{\log b \leqslant \ell \leqslant \log n} 2^{\ell bk/2^\ell} \cdot \prod_{\log b \leqslant \ell \leqslant \log n} (1 + 2^{-\ell})^{bk/2^\ell} \\ &\leqslant \frac{\prod_{\ell \geqslant 1} 2^{\ell bk/2^\ell}}{\prod_{\ell \leqslant \log b - 1} 2^{\ell bk/2^\ell}} \cdot \prod_{\ell \geqslant \log b} e^{bk/4^\ell} \\ &= \frac{2^{2bk}}{2^{(2(b - \log b - 1)k}} \cdot e^{(4/3)k/b} \\ &\leqslant 2^{2(\log b + 2)k}, \end{aligned}$$

where in the second-last line we used $\sum_{\ell \geqslant 1} \ell/2^\ell = 2$, $\sum_{\ell \leqslant y} \ell/2^\ell = 2^{-y}(2^{y+1} - y - 2)$, and $\sum_{\ell \geqslant 0} 1/4^\ell = 4/3$; and in the last line we used that $e^{(4/3)} < 4$. □

For any realization of VISIT-EXCHANGE, each $A_k \in \mathcal{B}_k$ is a concise descriptor of some labeled canonical walk of length k. However it is not always the case that a labeled canonical walk of length k has a concise descriptor. The next lemma shows that w.h.p. all labeled canonical walks of length k have concise descriptors for an appropriate choice of constant parameter b. Note that the lemma assumes the T-VISIT-EXCHANGE process.

LEMMA 8. *If* $b \geqslant \max\{2\alpha e^2, 64\}$ *then, with probability at least* $1 - 2^{-bk/4} \log n$*, all labeled canonical walks of length* k *in a random realization of* T-VISIT-EXCHANGE *have concise descriptors.*

4.6 Proof of Lemma 8

First, we bound the number of steps t in which more than i agents are encountered in a canonical walk of length k.

LEMMA 9. *Fix any* $A_k \in \mathcal{A}_k$*, and let* $\eta = (\theta_0, g_1, \theta_1, g_2, \ldots, g_k, \theta_k)$ *be the labeled canonical walk in* T-VISIT-EXCHANGE *that has non-concise (or semi-concise) descriptor* A_k*. For any* $i \geqslant e^2\alpha$ *and* $\beta \geqslant e^2\alpha$*,*

$$\mathbb{P}\left[|\{t \in \{1, \ldots, k\} \colon \delta_t > i\}| \geqslant \beta k/i\right] \leqslant 2^{-\beta k}.$$

PROOF. Recall that $\delta_t = |Z_{\theta_{t-1}}(t-1)|$ is the number of agents that visit vertex θ_{t-1} in round $t-1$, and thus also the number of agents that depart from θ_{t-1} in round t. We argue that for any $t \geqslant 1$, conditioned on $\delta_1, \ldots, \delta_t$, variable δ_{t+1} is stochastically dominated by the binomial variable $\mathrm{Bin}(\alpha d - 1, 1/d) + 1$: From (2), applied for vertex θ_t and round $t-1$, we get

$$\sum_{v \in \Gamma(\theta_t)} |Z_v(t-1)| \leqslant \alpha \cdot d,$$

thus, there are at most αd agents in the neighborhood of θ_t at the beginning of round t. If $g_t = \perp$ (thus $\theta_t = \theta_{t-1}$), then each one of those at most αd agents will visit θ_t in round t independently with probability $1/d$. If $g_t \neq \perp$ (thus $\theta_t \neq \theta_{t-1}$ and $g_t \in Z_{\theta_{t-1}}(t-1) \cap Z_{\theta_t}(t)$), then each of the at most αd agents will visit θ_t in round t independently with probability $1/d$, *except* for agent g_t who visits θ_t with probability 1. In both cases, the number δ_{t+1} of agents that visit θ_t is dominated by $\mathrm{Bin}(\alpha d - 1, 1/d) + 1$. It follows that, for $t > 1$ and $i \geqslant 1$,

$$\begin{aligned} \mathbb{P}[\delta_t > i \mid \delta_1, \ldots, \delta_{t-1}] &\leqslant \mathbb{P}[\mathrm{Bin}(\alpha d - 1, 1/d) > i - 1] \\ &\leqslant \mathbb{P}[\mathrm{Bin}(\alpha d, 1/d) \geqslant i] \\ &\leqslant \binom{\alpha d}{i} \cdot \frac{1}{d^i} \leqslant \left(\frac{e\alpha d}{i}\right)^i \cdot \frac{1}{d^i} = \left(\frac{e\alpha}{i}\right)^i. \end{aligned}$$

For the case of $t = 1$, since $|A| \leqslant \alpha n/(2e) < \alpha n$ by (1), we similarly have

$$\mathbb{P}[\delta_1 \geqslant i] \leqslant \mathbb{P}[\mathrm{Bin}(\alpha n, 1/n) \geqslant i] \leqslant \left(\frac{e\alpha}{i}\right)^i.$$

Let $p_i = \left(\frac{e\alpha}{i}\right)^i$. It follows from the above that, for any $\ell \geqslant 1$,

$$\mathbb{P}\left[|\{t \in \{1,\ldots,k\} : \delta_t > i\}| \geqslant \ell\right] \leqslant \mathbb{P}\left[\mathrm{Bin}(k, p_i) \geqslant \ell\right] \leqslant \binom{k}{\ell} \cdot p_i^\ell \leqslant \left(\frac{ekp_i}{\ell}\right)^\ell. \quad (5)$$

For $\ell \geqslant \beta k/i$ and $i \geqslant e^2\alpha$,

$$\begin{aligned}
\left(\frac{ekp_i}{\ell}\right)^\ell &\leqslant \left(\frac{ek(e\alpha/i)^i}{\beta k/i}\right)^\ell, && \text{by } p_i = \left(\frac{e\alpha}{i}\right)^i \text{ and } \ell \geqslant \beta k/i\\
&= \left(\frac{e^2\alpha}{\beta} \cdot \left(\frac{e\alpha}{i}\right)^{i-1}\right)^\ell \leqslant \left(\frac{e\alpha}{i}\right)^{(i-1)\ell}, && \text{by } \beta \geqslant e^2\alpha\\
&\leqslant \left(\frac{e\alpha}{i}\right)^{(1-1/i)\beta k}, && \text{by } \ell \geqslant \beta k/i\\
&\leqslant \left(\frac{1}{e}\right)^{(1-1/e^2)\beta k}, && \text{by } i \geqslant e^2\alpha \geqslant e^2\\
&\leqslant 2^{-\beta k}.
\end{aligned}$$

Substituting that to (5) completes the proof of Lemma 9. ◻

We proceed now to the proof of the main lemma. For any $A_k \in \mathcal{A}_k$, and for $\eta = (\theta_0, g_1, \theta_1, g_2, \ldots, \theta_k)$ the labeled canonical walk with semi-concise descriptor A_k, let $\mathcal{E}_{A_k}$ denote the event:

$$|\{t \in \{1,\ldots,k\} : 2^{\ell-1} < \delta_t \leqslant 2^\ell\}| \leqslant k_{2^\ell},\ \forall\, \ell \in \{\log b+1, \ldots, \log n\}.$$

Applying Lemma 9, for $i = 2^{\ell-1}$, $\beta = b/2$, and each $\ell \in \{\log b + 1, \ldots, \log n\}$, and then using a union bound, we obtain

$$\mathbb{P}\left[\mathcal{E}_{A_k}\right] \geqslant 1 - 2^{-bk/2}\log n.$$

By Lemma 7 and another union bound,

$$\begin{aligned}
\mathbb{P}\left[\bigcap_{A_k \in \mathcal{B}_k} \mathcal{E}_{A_k}\right] &\geqslant 1 - |\mathcal{B}_k| \cdot 2^{-bk/2}\log n\\
&\geqslant 1 - (4b)^{2k} \cdot 2^{-bk/2}\log n\\
&\geqslant 1 - 2^{-bk/4}\log n, \qquad (6)
\end{aligned}$$

where the last inequality holds when $b \geqslant 64$. Next we show that event $\bigcap_{A_k \in \mathcal{B}_k} \mathcal{E}_{A_k}$ implies that every labeled canonical walk η has a concise descriptor $A_k \in \mathcal{B}_k$. From this and (6), the lemma follows.

Fix a realization of T-VISIT-EXCHANGE conditioned on the event $\bigcap_{A_k \in \mathcal{B}_k} \mathcal{E}_{A_k}$. Suppose, for contradiction, that there is some labeled canonical walk $\eta' = (\theta'_0, g'_1, \theta'_1, \ldots, g'_k, \theta_k)$ that does *not* have a concise descriptor. Let $\eta = (\theta_0, g_1, \theta_1, \ldots, g_k, \theta_k)$ be a labeled canonical walk that *does* have a concise descriptor $A_k \in \mathcal{B}_k$, and shares a *maximal common prefix* with η'. Consider the first element where η' and η are different. We argue that this element is not a vertex: Suppose, for contradiction, that $(\theta'_0, \ldots, g'_i) = (\theta_0, \ldots, g_i)$ and $\theta'_i \neq \theta_i$, for some $0 \leqslant i \leqslant k$. Then $i \neq 0$, as $\theta'_0 = s = \theta_0$. Moreover, if $i > 0$, then by definition, $(\theta'_0, \ldots, g'_i) = (\theta_0, \ldots, g_i)$ implies $\theta'_i = \theta_i$, contradicting our assumption. Thus, the first element where η' and η are different must be an agent. Suppose $(\theta'_0, g'_1, \ldots, \theta'_{i-1}) = (\theta_0, g_1, \ldots, \theta_{i-1})$ and $g'_i \neq g_i$, for some $1 \leqslant i \leqslant k$. Then, by the maximal prefix assumption, the labeled canonical walk $(\theta_0, \ldots, \theta_{i-1}, g'_i, \theta'_i, \perp, \theta'_i, \perp, \ldots, \perp, \theta'_i)$, which stays at vertex θ'_i in rounds $i+1$ up to k, has no concise descriptor. This can only be true if $|\{t \in \{1, \ldots, i-1\} : 2^{\ell-1} < \delta_t \leqslant 2^\ell\}| > k_{2^\ell}$, for some $\ell \in \{\log b+1, \ldots, \log n\}$. But this contradicts event $\mathcal{E}_{A_k}$. Therefore, there exists no labeled canonical walk η' of length k such that η' has no concise descriptor.

4.7 Bound on Congestion

The next lemma gives un upper bound on the congestion of a single canonical walk of length k.

Lemma 10. *Fix any $A_k \in \mathcal{B}_k$, and let η be the labeled canonical walk in* T-VISIT-EXCHANGE *that has concise descriptor A_k. Then for any $\beta \geqslant 2e\alpha + 1$, $\mathbb{P}[Q(\eta) \leqslant \beta k] \geqslant 1 - 2^{-(\beta-1)k}$.*

Proof. Let $\eta = (\theta_0, g_1, \theta_1, \ldots, g_k, \theta_k)$. Then $Q(\eta) = \sum_{1 \leqslant t \leqslant k} \delta_t$, where $\delta_t = |Z_{\theta_{t-1}}(t-1)|$. From the same reasoning as in the proof of Lemma 9, we have that $Q(\eta)$ is stochastically dominated by $k + \sum_{1 \leqslant t \leqslant k} B_t$, where $B_1, \ldots, B_k$ are mutually independent binomial random variables, such that $B_1 \sim \mathrm{Bin}(\alpha n, 1/n)$ and, for $t > 1$, $B_t \sim \mathrm{Bin}(\alpha d, 1/d)$. Then $\mathbb{E}[\,Q(\eta) - k\,] \leqslant \sum_{1 \leqslant t \leqslant k} \mathbb{E}[\,B_t\,] = k\alpha$, and

$$\mathbb{P}[Q(\eta) \geqslant \beta k] = \mathbb{P}[Q(\eta) - k \geqslant (\beta - 1)k] \leqslant 2^{-(\beta-1)k},$$

where the last inequality follows from a standard Chernoff bound, using that $(\beta - 1)k \geqslant 2e \cdot \mathbb{E}[\,Q(\eta) - k\,]$. ◻

4.8 Putting the Pieces Together (Proof of Theorem 2)

We consider first the case where k is at most logarithmic. It is not hard to show that $T_{\text{visitx}} = \Omega(\log n)$ w.h.p., by arguing that some vertices are not visited by any agent (informed or not) during the first logarithmic number of rounds (see the full version of the paper). Thus, there is some constant $\epsilon > 0$ such that if $k \leqslant \epsilon \log n$, $\mathbb{P}[T_{\text{visitx}} \leqslant k] \leqslant n^{-\lambda}$. From this, the theorem's statement follows for $k \leqslant \epsilon \log n$. In the rest of the proof, we assume $k \geqslant \epsilon \log n$.

We have $T_{\text{push}} = \max_{u \in V} \tau_u$, and from Lemma 5,

$$T_{\text{push}} \leqslant \max_{u \in V} C_u(t_u).$$

Since for any fixed realization of VISIT-EXCHANGE and any $u \in V$, $C_u(t)$ is a non-decreasing function of t, and since $t_u \leqslant T_{\text{visitx}}$, it follows

$$T_{\text{push}} \leqslant \max_{u \in V} C_u(T_{\text{visitx}}).$$

By Lemma 6, for any $u \in V$, there is a canonical walk θ of length $t = T_{\text{visitx}}$ with congestion $Q(\theta) = C_u(T_{\text{visitx}})$. Thus, there is also a *labeled* canonical walk η of length T_{visitx} with $Q(\eta) = Q(\theta) = C_u(T_{\text{visitx}})$. It follows

$$T_{\text{push}} \leqslant \max_{\eta \in \mathcal{H}(T_{\text{visitx}})} Q(\eta), \qquad (7)$$

where $\mathcal{H}(t)$ denotes the set of all labeled canonical walks of length t in VISIT-EXCHANGE.

Next we bound $\max_{\eta \in \mathcal{H}(k)} Q(\eta)$. Consider T-VISIT-EXCHANGE, and for any $A_k \in \mathcal{B}_k$, let η_{A_k} be the labeled canonical walk in T-VISIT-EXCHANGE with concise descriptor A_k. From Lemma 10, for any $A_k \in \mathcal{B}_k$ and $\beta \geqslant 2e\alpha + 1$, $\mathbb{P}\left[Q(\eta_{A_k}) \leqslant \beta k\right] \geqslant 1 - 2^{-(\beta-1)k}$. Then

$$\begin{aligned}
\mathbb{P}\left[\max_{A_k \in \mathcal{B}_k} Q(\eta_{A_k}) \leqslant \beta k\right] &\geqslant 1 - 2^{-(\beta-1)k} \cdot |\mathcal{B}_k|\\
&\geqslant 1 - 2^{-(\beta-1)k} \cdot (4b)^{2k},
\end{aligned}$$

by Lemma 7. Choosing constant β large enough so that $(\beta-1)/2 \geqslant 2\log(4b)$, yields

$$\mathbb{P}\left[\max_{A_k \in \mathcal{B}_k} Q(\eta_{A_k}) \leqslant \beta k\right] \geqslant 1 - 2^{-(\beta-1)k/2}.$$

From Lemma 8, the probability that all labeled canonical walks of length k have concise descriptors is at least $1 - 2^{-bk/4}\log n$, if $b \geqslant \max\{2\alpha e^2, 64\}$. It follows

$$\mathbb{P}\left[\max_{A_k \in \mathcal{B}_k} Q(\eta_{A_k}) = \max_{\eta \in \mathcal{H}^*(k)} Q(\eta)\right] \geqslant 1 - 2^{-bk/4}\log n,$$

where $\mathcal{H}^*(t)$ is the set of all labeled canonical walks of length t in T-VISIT-EXCHANGE. By Lemma 4, however, VISIT-EXCHANGE and the coupled T-VISIT-EXCHANGE, which use the same random walks, are identical until round k, with probability at least $1 - kn \cdot 2^{-ad}$, thus

$$\mathbb{P}\left[\mathcal{H}(k) = \mathcal{H}^*(k)\right] \geqslant 1 - kn \cdot 2^{-\alpha d}.$$

Combining the last three inequalities above, we obtain

$$\mathbb{P}\left[\max_{\eta \in \mathcal{H}(k)} Q(\eta) \leqslant \beta k\right] \geqslant 1 - 2^{-(\beta-1)k/2} - 2^{-bk/4}\log n - kn \cdot e^{-\alpha d}.$$

Since $k \geqslant \epsilon \log n$ and $d \geqslant \varepsilon \log n$, for any given constant $\lambda > 0$ we can choose constants β, b, α large enough such that

$$\mathbb{P}\left[\max_{\eta \in \mathcal{H}(k)} Q(\eta) \leqslant \beta k\right] \geqslant 1 - n^{-\lambda}. \tag{8}$$

From (7) and (8), we obtain

$$\begin{aligned}
\mathbb{P}\left[T_{\text{push}} \leqslant \beta k\right] &\geqslant \mathbb{P}\left[\max_{\eta \in \mathcal{H}(T_{\text{visitx}})} Q(\eta) \leqslant \beta k\right], \qquad \text{by (7)} \\
&\geqslant \mathbb{P}\left[T_{\text{visitx}} \leqslant k \cap \max_{\eta \in \mathcal{H}(k)} Q(\eta) \leqslant \beta k\right] \\
&\geqslant \mathbb{P}\left[T_{\text{visitx}} \leqslant k\right] - \mathbb{P}\left[\max_{\eta \in \mathcal{H}(k)} Q(\eta) > \beta k\right] \\
&\geqslant \mathbb{P}\left[T_{\text{visitx}} \leqslant k\right] - n^{-\lambda}, \qquad \text{by (8)}.
\end{aligned}$$

This completes the proof of Theorem 2.

5 BOUNDING T_{visitx} BY T_{push} ON REGULAR GRAPHS

The following theorem upper bounds the broadcast time of VISIT-EXCHANGE in a regular graph by the broadcast time of PUSH.

THEOREM 11. *For any constants $\alpha, \beta, \lambda > 0$ with $\alpha \cdot \beta$ sufficiently large, there is a constant $c > 0$, such that for any d-regular graph $G = (V, E)$ with $|V| = n$ and $d \geqslant \beta \log n$, and for any source $s \in V$, the broadcast times of PUSH and VISIT-EXCHANGE, with $|A| \geqslant \alpha n$ agents, satisfy*

$$\mathbb{P}\left[T_{\text{visitx}} \leqslant ck\right] \geqslant \mathbb{P}\left[T_{\text{push}} \leqslant k\right] - n^{-\lambda},$$

for any $k \geqslant 0$.

From Theorem 11, it is immediate that if $T_{\text{push}} \leqslant T$ w.h.p., then $T_{\text{visitx}} = O(T)$ w.h.p., as well. Moreover, using Theorem 11 and the well-known $O(n^2 \log n)$ upper bound w.h.p. on the cover time for a single random walk on a regular graph, which also applies to T_{visitx}, it is easy to show that $\mathbb{E}[T_{\text{visitx}}] = O(\mathbb{E}[T_{\text{push}}])$.

Proof Overview Of Theorem 11. We use a coupling which is similar to that in the proof of the converse result, stated in Theorem 2, but with a twist (which we describe momentarily). Unlike in the proof of Theorem 2, were we essentially consider all possible paths through which information travels, here we focus on the first path by which information reaches each vertex. Let $P = (u_0 = s, u_1, \ldots, u_k = u)$ be such a path for vertex u in PUSH, where each vertex u_i in the path learns the information from u_{i-1}. Let δ_i be the number of rounds it takes for u_{i-1} to sample (and inform) u_i in PUSH. We consider the same path in VISIT-EXCHANGE, and compare δ_i with the number D_i of rounds until some informed agent moves from u_{i-1} to u_i, counting from the round when u_{i-1} becomes informed. Note that $\sum_i \delta_i$ is precisely the round when u is informed in PUSH, while $\sum_i D_i$ is an upper bound on the round when u is informed in VISIT-EXCHANGE.

The coupling we used in Sect. 4 seems suitable for this setup. Recall, in that coupling we let the list of neighbors that a vertex u samples in PUSH, be identical to the list of neighbors that informed agents visit in their next step after visiting u, in VISIT-EXCHANGE. The same intuition applies, namely, that on average each vertex is visited by $|A|/n = \Omega(1)$ agents per round, which suggests that D_i should be close to δ_i. We can even apply a similar trick as in Sect. 4 to avoid some dependencies: In each round, the number of agents in the neighborhood of a vertex is bounded below by $d \cdot |A|/n = \Omega(d)$, w.h.p. This should imply that the number of agents that visit a vertex in a round is bounded below by a geometric distribution with constant expectation. Let $\mathcal{E}$ denote the event that the above $\Omega(d)$ bound holds for all u, for polynomially many rounds.

There is, however, a problem with this proof plan. By fixing path P in advance, to be the first path to inform u in PUSH, we introduce *dependencies from the future*. So, when we analyse D_i and δ_i, we must condition on the event that the i-prefix of the path we have considered so far will indeed be a prefix of the first path to reach u. These kind of dependencies seem hard to deal with.

We use the following neat idea to overcome this problem. We only consider the *odd rounds* of VISIT-EXCHANGE in the coupling, i.e., we match the list of neighbors that a vertex v samples in PUSH (in all rounds), to the list of neighbors that informed agents visit in round $2k+1$ after visiting u in round $2k$, for all $k \geqslant 0$. In even rounds, agents take steps *independently* of the coupled PUSH process.

Under this coupling, we proceed as follows. We condition on the high probability event $\mathcal{E}$ defined earlier (formally, we modify VISIT-EXCHANGE to ensure $\mathcal{E}$ holds). We then fix *all* random choices in PUSH, and thus the information path P to u. For each even round of VISIT-EXCHANGE, we have that vertex u_i in P is visited by at least one agent with constant probability, independently of the past and of the fixed choices in future odd rounds. If indeed some vertex visits u_i in an even round, then in the next round it will visit a vertex dictated by the coupling. This allows us to show that under this coupling, $\sum_i D_i \leqslant c(\sum_i \delta_i + \log n)$, w.h.p. We get rid of the $\log n$ term in the final bound, by using that $T_{\text{push}} = \Omega(\log n)$ w.h.p. The detailed proof can be found in the full version of the paper.

6 BOUNDING T_{visitx} BY T_{meetx} ON REG. GRAPHS

The next theorem bounds the broadcast time of VISIT-EXCHANGE in a regular graph by the broadcast time of MEET-EXCHANGE.

THEOREM 12. *For any constants $\alpha, \beta, \lambda > 0$ with $\alpha \cdot \beta$ sufficiently large, there is a constant $c > 0$, such that for any d-regular graph $G = (V, E)$ with $|V| = n$ and $d \geqslant \beta \ln n$, and for any source $s \in V$, the broadcast times of* VISIT-EXCHANGE *and* MEET-EXCHANGE*, both with $|A| \geqslant \alpha n$ agents, satisfy*

$$\mathbb{P}\left[T_{\text{visitx}} \leqslant k + c \ln n\right] \geqslant \mathbb{P}\left[T_{\text{meetx}} \leqslant k\right] - n^{-\lambda},$$

for any $k \geqslant 0$.

The proof of the theorem shows that once all agents are informed, it takes an additional $O(\log n)$ rounds w.h.p. before each vertex is visited at least once by some agent.

7 OPEN PROBLEMS

This work is the first systematic and thorough comparison of the running times of the standard PUSH and PUSH-PULL rumor spreading protocols with some very natural agent-based alternatives. Several open problems remain. The most obvious question to ask is whether our results for regular graphs hold also when the graph degree is sub-logarithmic. Another question is whether there are graphs where MEET-EXCHANGE is slower than VISIT-EXCHANGE by more than logarithmic factors. In this paper we assumed a linear number of agents. It would be interesting to study the performance of the protocols when a sub-linear number of agents is available.

The main attractive properties of standard rumor spreading protocols are simplicity, scalability, and robustness to failures [21]. Arguably, VISIT-EXCHANGE and MEET-EXCHANGE share the first two properties, but probably not the robustness property. In particular, it seems that faulty nodes or links can result in agents getting lost. It would be interesting to explore fault tolerant variants of these protocols. For example, it seems likely that the protocols could tolerate some number of lost agents, if a *dynamic* set of agents were used, where agents age with time and die, while new agents are born at a proportional rate.

ACKNOWLEDGMENTS

The authors would like to thank Thomas Sauerwald and Nicolás Rivera for helpful discussions. This research was undertaken, in part, thanks to funding from the ANR Project PAMELA (ANR-16-CE23-0016-01), the NSF Award Numbers CCF-1461559, CCF-0939370 and CCF-18107, the Gates Cambridge Scholarship programme, and the ERC grant DYNAMIC MARCH.

REFERENCES

[1] R. Aleliunas, R. M. Karp, R. J. Lipton, L. Lovász, and C. Rackoff. Random walks, universal traversal sequences, and the complexity of maze problems. In *Proc. 20th IEEE Symposium on Foundations of Computer Science, FOCS*, pages 218–223, 1979.

[2] N. Alon, C. Avin, M. Koucký, G. Kozma, Z. Lotker, and M. R. Tuttle. Many random walks are faster than one. *Combinatorics, Probability & Computing*, 20(4):481–502, 2011.

[3] O. Angel, A. Mehrabian, and Y. Peres. The string of diamonds is tight for rumor spreading. In *Proc. Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM*, pages 26:1–26:9, 2017.

[4] P. Berenbrink, R. Elsässer, and T. Friedetzky. Efficient randomised broadcasting in random regular networks with applications in peer-to-peer systems. *Distributed Computing*, 29(5):317–339, 2016.

[5] P. Berenbrink, G. Giakkoupis, A. Kermarrec, and F. Mallmann-Trenn. Bounds on the voter model in dynamic networks. In *Proc. 43rd International Colloquium on Automata, Languages, and Programming, ICALP*, pages 146:1–146:15, 2016.

[6] P. Berenbrink, G. Giakkoupis, and P. Kling. Tight bounds for coalescing-branching random walks on regular graphs. In *Proc. 29th ACM-SIAM Symposium on Discrete Algorithms, SODA*, pages 1715–1733, 2018.

[7] S. P. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Randomized gossip algorithms. *IEEE Trans. Information Theory*, 52(6):2508–2530, 2006.

[8] A. Z. Broder, A. R. Karlin, P. Raghavan, and E. Upfal. Trading space for time in undirected s-t connectivity. *SIAM J. Comput.*, 23(2):324–334, 1994.

[9] K. Censor-Hillel, B. Haeupler, J. A. Kelner, and P. Maymounkov. Rumor spreading with no dependence on conductance. *SIAM J. Comput.*, 46(1):58–79, 2017.

[10] F. Chierichetti, G. Giakkoupis, S. Lattanzi, and A. Panconesi. Rumor spreading and conductance. *J. ACM*, 65(4):17:1–17:21, 2018.

[11] F. Chierichetti, S. Lattanzi, and A. Panconesi. Rumor spreading in social networks. *Theor. Comput. Sci.*, 412(24):2602–2610, 2011.

[12] C. Cooper. Random walks, interacting particles, dynamic networks: Randomness can be helpful. In *Proc. Structural Information and Communication Complexity - 18th International Colloquium, SIROCCO*, pages 1–14, 2011.

[13] C. Cooper, A. M. Frieze, and T. Radzik. Multiple random walks in random regular graphs. *SIAM J. Discrete Math.*, 23(4):1738–1761, 2009.

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

[15] T. Dimitriou, S. E. Nikoletseas, and P. G. Spirakis. The infection time of graphs. *Discrete Applied Mathematics*, 154(18):2577–2589, 2006.

[16] B. Doerr, M. Fouz, and T. Friedrich. Social networks spread rumors in sublogarithmic time. In *Proc. 43rd ACM Symposium on Theory of Computing, STOC*, pages 21–30, 2011.

[17] B. Doerr, M. Fouz, and T. Friedrich. Social networks spread rumors in sublogarithmic time. In *Proc. 43rd ACM Symposium on Theory of Computing, STOC*, pages 21–30. ACM, 2011.

[18] B. Doerr, T. Friedrich, and T. Sauerwald. Quasirandom rumor spreading. *ACM Trans. Algorithms*, 11(2):9:1–9:35, 2014.

[19] K. Efremenko and O. Reingold. How well do random walks parallelize? In *Proc. Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, 12th International Workshop, APPROX, and 13th International Workshop, RANDOM*, pages 476–489, 2009.

[20] R. Elsässer and T. Sauerwald. Tight bounds for the cover time of multiple random walks. *Theor. Comput. Sci.*, 412(24):2623–2641, 2011.

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

[22] G. Giakkoupis. Tight bounds for rumor spreading with vertex expansion. In *Proc. 25th ACM-SIAM Symposium on Discrete Algorithms, SODA*, pages 801–815, 2014.

[23] G. Giakkoupis, Y. Nazari, and P. Woelfel. How asynchrony affects rumor spreading time. In *Proc. 35th ACM Symposium on Principles of Distributed Computing, PODC*, pages 185–194, 2016.

[24] M. Harchol-Balter, F. T. Leighton, and D. Lewin. Resource discovery in distributed networks. In *Proc. 18th ACM Symposium on Principles of Distributed Computing, PODC*, pages 229–237, 1999.

[25] J. Hermon. Frogs on trees? *Electron. J. Probab.*, 23, 2018.

[26] V. Kanade, F. Mallmann-Trenn, and T. Sauerwald. On coalescence time in graphs: When is coalescing as fast as meeting? In *Proc. 30th ACM-SIAM Symposium on Discrete Algorithms, SODA*, pages 956–965, 2019.

[27] R. M. Karp, C. Schindelhauer, S. Shenker, and B. Vöcking. Randomized rumor spreading. In *Proc. 41st IEEE Symposium on Foundations of Computer Science, FOCS*, pages 565–574, 2000.

[28] D. Kempe, A. Dobra, and J. Gehrke. Gossip-based computation of aggregate information. In *Proc. 44th IEEE Symposium on Foundations of Computer Science FOCS*, pages 482–491, 2003.

[29] H. Lam, Z. Liu, M. Mitzenmacher, X. Sun, and Y. Wang. Information dissemination via random walks in d-dimensional space. In *Proc. 23rd ACM-SIAM Symposium on Discrete Algorithms, SODA*, pages 1612–1622, 2012.

[30] M. Mitzenmacher, R. Rajaraman, and S. T. Roche. Better bounds for coalescing-branching random walks. *TOPC*, 5(1):2:1–2:23, 2018.

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

[32] A. Pettarin, A. Pietracaprina, G. Pucci, and E. Upfal. Infectious random walks. *CoRR*, abs/1007.1604, 2010.

[33] S. Y. Popov. Frogs and some other interacting random walks models. In *Proc. Discrete Random Walks, DRW*, pages 277–288, 2003.

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

[35] R. van Renesse, Y. Minsky, and M. Hayden. A gossip-style failure detection service. In *Proc. IFIP International Conference on Distributed Systems Platforms and Open Distributed Processing*, Middleware '98, pages 55–70, 1998.