How asynchrony affects rumor spreading time

George Giakkoupis, Yasamin Nazari, Philipp Woelfel

How Asynchrony Affects Rumor Spreading Time*

George Giakkoupis
INRIA Rennes
France
george.giakkoupis@inria.fr

Yasamin Nazari
University of Calgary
Canada
ynazari@ucalgary.ca

Philipp Woelfel
University of Calgary
Canada
woelfel@ucalgary.ca

ABSTRACT

In standard randomized (push-pull) rumor spreading, nodes communicate in synchronized rounds. In each round every node contacts a random neighbor in order to exchange the rumor (i.e., either push the rumor to its neighbor or pull it from the neighbor). A natural asynchronous variant of this algorithm is one where each node has an independent Poisson clock with rate 1, and every node contacts a random neighbor whenever its clock ticks. This asynchronous variant is arguably a more realistic model in various settings, including message broadcasting in communication networks, and information dissemination in social networks.

In this paper we study how asynchrony affects the rumor spreading time, that is, the time before a rumor originated at a single node spreads to all nodes in the graph. Our first result states that the asynchronous push-pull rumor spreading time is asymptotically bounded by the standard synchronous time. Precisely, we show that for any graph G on n-nodes, where the synchronous push-pull protocol informs all nodes within $T(G)$ rounds with high probability, the asynchronous protocol needs at most time $O(T(G) + \log n)$ to inform all nodes with high probability. On the other hand, we show that the expected synchronous push-pull rumor spreading time is bounded by $O(\sqrt{n})$ times the expected asynchronous time.

These results improve upon the bounds for both directions shown recently by Acan et al. (PODC 2015). An interesting implication of our first result is that in regular graphs, the weaker push-only variant of synchronous rumor spreading has the same asymptotic performance as the synchronous push-pull algorithm.

*This research was undertaken, in part, thanks to funding from the Canada Research Chairs program, the Discovery Grants program of NSERC, the INRIA Associate Team RADCON, Alberta Innovates Technology Futures, and MITACS.

PODC'16, July 25 - 28, 2016, Chicago, IL, USA

ISBN 978-1-4503-3964-3/16/07...$15.00

DOI: http://dx.doi.org/10.1145/2933057.2933117

Keywords

randomized rumor spreading; push/pull; asynchronous vs synchronous models

1. INTRODUCTION

Broadcasting information in large networks is a fundamental and well-studied problem. Desirable properties of broadcasting algorithms are efficiency, simplicity, decentralization, and tolerance to changes in the network topology. The classical abstraction is the randomized (synchronous) rumor spreading protocol [7, 22]: Initially, a piece of information, called rumor, is injected at a random or arbitrarily chosen node. After that, nodes communicate in synchronous rounds to inform each other of the rumor. In every round, each node calls a uniformly random neighbor and establishes a communication with its callee in order to possibly exchange the rumor: In the *push* protocol, an informed caller pushes the rumor to its callee, while in the *pull* protocol, a non-informed caller receives the rumor from its callee, if the callee is informed. The *push-pull* protocol combines the push and pull communication and allows a bi-directional rumor exchange in each round between each caller and its callee.

Besides being of fundamental interest, rumor spreading protocols have many direct applications, such as in the maintenance of distributed replicated database systems [7, 13], failure detection [26], resource discovery [20], and data aggregation [4]. As such, the rumor spreading time, i.e., the number of rounds until all nodes in a network have received the rumor (either in expectation or with high probability), has been studied intensively. A large body of research work deals with the question how the rumor spreading time is influenced by the network topology (e.g., [8, 12, 15, 16]), network parameters such as expansion [5, 6, 17–19, 25], or the communication modes push, pull, and push-pull (e.g., [24]).

The synchrony assumption, according to which all nodes establish connections simultaneously in a round-by-round fashion, has been criticized for not being plausible in many scenarios [9–11]. Real networks typically do not have a centralized clock, and individual links are affected by frequent changes in communication speed. Moreover, decentralization has been emphasized as one of the main advantages of the rumor spreading protocol, but this contradicts the model assumption of a centralized clock. More recently the performance of rumor spreading protocols in a natural asynchronous setting, initially proposed by Boyd, Ghosh, Prabhakar, and Shah [4], has been considered. Here, nodes establish communications with their neighbors at times determined by independent Poisson processes, rather than at fixed unit times. More precisely, each node is equipped with an independent Poisson clock with rate 1, and whenever a node's clock ticks, the node exchanges the rumor with a uniformly random neighbor (using push, pull, or push-pull communication).

On the hypercube, the asynchronous push-pull protocol corresponds to Richardson's model for the spread of a disease, and has been investigated in the study of first-passage percolation [3, 14]. But only recently has asynchronous rumor spreading found the interest of researchers in the area of networks, initially in order to understand information spread in social networks. It was observed that on common network topologies modelling social networks, such as Chung-Lu power law graphs [16] or preferential attachment graphs [9], the push-pull protocol spreads the rumor to a large fraction of the nodes significantly faster in the asynchronous than in the synchronous model. There are even graph topologies for which the asynchronous push-pull protocol has poly-logarithmic rumor spreading time, whereas the synchronous protocol requires a polynomial number of rounds [1]. On the other hand, there are simple networks, where synchrony allows for faster rumor spreading than asynchrony [1]: In an n-vertex star, it takes at most 2 rounds of the synchronous push-pull protocol to spread the rumor to all nodes (it takes at most one round for the centre node to get informed by push, and another round for all remaining nodes to pull the rumor from the center); whereas in the asynchronous model it takes with high probability $\Theta(\log n)$ time until sufficiently many different Poisson clocks have ticked for all nodes to get informed. Finally, for various classical graph topologies such as the hypercube, random graphs, and random regular graphs, both protocols have the same rumor spreading times within constant factors [2, 14, 21, 23].

These results raise the question how big the gap between the asynchronous and the synchronous rumor spreading times can be. In the following discussion we restrict ourselves to push-pull communication for graphs with n vertices, unless mentioned otherwise. Acan, Collevecchio, Mehrabian, and Wormald [1] showed that for any graph the high-probability rumor spreading time in the asynchronous model is at most a multiplicative $O(\log n)$ factor larger than that of the synchronous model. While this result is tight for the n-vertex star, it may not be tight for graphs that have super-constant synchronous rumor spreading time. In fact, Acan et al. conjectured that the high-probability asynchronous rumor spreading time can be at most by an additive $O(\log n)$ term larger than the synchronous one. Our first main result proves that this conjecture is true up to a constant factor.

Theorem 1. *Let G be a connected graph with n vertices, u be a vertex of G, and $T_{1/n}$ be the number of synchronous push-pull rounds before the rumor spreads from u to all other nodes with high probability (i.e., with probability $1 - 1/n$). Then the time in the asynchronous push-pull protocol before the rumor spreads from u to all other nodes is at most $O(T_{1/n} + \log n)$ with high probability.*

This bound is asymptotically tight, and in particular it shows that for most graphs, namely those that have at least logarithmic high-probability rumor spreading time, the asynchronous rumor spreading time is not asymptotically larger than the synchronous one. Further, it implies that several known upper bounds that have been shown to hold for the synchronous push-pull protocol carry over to the asynchronous, such as known bounds with graph expansion parameters [17, 18].

Acan et al. [1] showed also that the high-probability rumor spreading time in the synchronous model can be at most by a factor of $O(n^{2/3})$ larger than in the asynchronous model. They conjecture that this factor can be improved to $n^{1/2} \cdot (\log n)^{O(1)}$. Our second main result is a proof of this conjecture.

Theorem 2. *Let G be a connected graph with n vertices, u be a vertex of G, and T be the number of synchronous push-pull rounds before the rumor spreads from u to all other nodes. Then the time in the asynchronous push-pull protocol before the rumor spreads from u to all other nodes has expectation $\Omega(\mathrm{E}[T]/\sqrt{n})$.*

Using this theorem, the conjecture of Acan et al. follows from the fact that with high probability the synchronous rumor spreading time T is at most by a multiplicative $O(\log n)$ factor larger than its expectation; i.e., $T = O(\mathrm{E}[T] \cdot \log n)$ with high probability. We currently do not know if the above bound is tight. Acan et al. [1] described a graph in which the asynchronous push-pull protocol has logarithmic running time, whereas synchronous push-pull has a running time of $\Theta(n^{1/3})$. This implies that the lower bound of Theorem 2 may be smaller than the best possible bound by at most a factor of $\Theta(n^{1/6})$.

In the synchronous model, the push-pull protocol can be significantly faster than the push(-only) protocol (but clearly it cannot be slower). For example, for an n-vertex star it takes with high probability $\Theta(n \log n)$ synchronous rounds to inform all nodes using the push protocol, while it takes at most two rounds using the push-pull protocol, as discussed earlier. Theorem 1 has the interesting consequence that push-pull communication can only have performance benefits over push on *non-regular* graphs. This is immediate from the following observations: (1) For the push protocol, Sauerwald [24] showed that for any graph, the high-probability synchronous rumor spreading time is bounded by the asynchronous rumor spreading time within a constant multiplicative factor. (2) It is not hard to see that on regular graphs, the asynchronous rumor spreading time of the push protocol has the same distribution as twice the asynchronous rumor spreading time of the push-pull protocol (see the full version of the paper for details). Last, (3) Theorem 1 implies that on regular graphs, the high-probability asynchronous push-pull rumor spreading time is bounded by the synchronous push-pull rumor spreading time within a constant multiplicative factor.[1] To summarize, we obtain the following, informally stated relations for the asymptotic high-probability rumor spreading times on any regular graph:[2]

$$\text{synchronous push} \overset{(1)}{\leq} \text{asynchronous push} \overset{(2)}{\leq} \text{asynchronous push-pull} \overset{(3)}{\leq} \text{synchronous push-pull}.$$

Corollary 3. *Let G be a connected regular graph with n vertices, u be a vertex of G, and $T_{p,1/n}, T_{pp,1/n}$ be the*

[1] The logarithmic additive term in the bound of Theorem 1 can be omitted as the synchronous push-pull rumor spreading time is with high probability at least $\Omega(\log n)$ on any regular graph.

[2] Here, '$\leq$' reads as 'has with high probability smaller or equal rumor spreading time, modulo constant factors.'

number of synchronous push, respectively, push-pull rounds before the rumor spreads from u to all other nodes with high probability. Then $T_{p,1/n} = \Theta(T_{pp,1/n})$.

2. DEFINITIONS

Let $G = (V, E)$ be a connected undirected graph with $|V| = n$ nodes. For each node $u \in V$, $\deg_G(u)$ denotes the degree of u, and $\Gamma_G(u)$ is the set of neighbors of u in G; we will omit subscript G when there is no ambiguity.

Rumor Spreading Algorithms

We consider two randomized rumor spreading algorithms on graph G. The first is the standard *synchronous push-pull* algorithm, or simply *push-pull*, which proceeds in synchronized rounds; we will write *pp* to denote this algorithm. Initially, in round 0, a source node $u \in V$ generates a rumor. In each subsequent round $r = 1, 2, \ldots$, every node $v \in V$ initiates a communication channel with a uniformly random neighbor $w \in \Gamma(v)$ (we say v *contacts* w), and if before the round exactly one of v, w knows the rumor (is *informed*), then the other node gets informed in round r as well. In particular, if node v is informed before the round and w is not, we say that v *pushes* the rumor to w in round r, while if v is not informed and w is, we say that v *pulls* the rumor from w. Note that each node contacts exactly one other node, but may be contacted by several nodes in the same round. In this case, we assume that the communications take place in parallel and independently. In the analysis, we will represent the pairwise communications that take place in a round by a set $\{(v, w_v)\}_{v\in V}$ of n pairs, denoting that each node v contacts node w_v in the round.

The second algorithm we consider is the *asynchronous push-pull* algorithm, denoted *pp-a*. In this algorithm, each node $v \in V$ has its own independent Poisson clock with rate $\lambda = 1$, and each time v's clock ticks, v contacts a uniformly random neighbor $w \in \Gamma(v)$. As before, if only one of v, w knows the rumor before the communication takes place, then the other node gets informed as well. We refer to this communication as a *step* of the algorithm, and say that node v *takes* or *executes* this step. We will represent a step by a pair (v, w), denoting that node v contacts w in the step.

We will consider a couple of alternative, but equivalent views of asynchronous push-pull. Rather than assuming a Poisson clock with rate 1 on each node, we can assume that we have an independent Poisson clock for each (ordered) pair of adjacent nodes (v, w) with rate $1/\deg(v)$, and each time this clock ticks, v contacts w. A second alternative is to assume that we have a single Poisson clock with rate n, and each time this clock ticks, a uniformly random node is chosen to take a step, i.e., contact a random neighbor. The equivalence of these descriptions is immediate from the properties of the superposition of independent Poisson processes and the memoryless property of exponential random variables.

Rumor Spreading Time

Next we define the time complexity measures we will use. For a rumor spreading algorithm α, we define the *rumor spreading time of* α *on* $G = (V, E)$ *for source* $u \in V$, denoted $T(\alpha, G, u)$, to be the "time" before a rumor originated at node u spreads to all nodes in G using algorithm α. The notion of time is different for synchronous and asynchronous algorithms. For the former, time is measured in terms of *rounds*, while for asynchronous algorithms time is measured in terms of *time units*.[3] For $0 < q < 1$, we define

$$T_q(\alpha, G, u) = \min\{t\colon \Pr[T(\alpha, G, u) \le t] \ge 1 - q)\},$$

i.e., the time before all nodes are informed with probability $1 - q$. We will be particularly interested in $T_{1/n}(\alpha, G, u)$, the *high-probability rumor spreading time* of α.

Other Notation

We write $X \sim Y$ to denote that random variables X and Y have the same distribution, and $X \preccurlyeq Y$ to denote that Y stochastically dominates X.

$\text{Unif}(A)$ denotes the uniform distribution over the set A; $\text{Geom}(p)$ is the geometric distribution with success probability p; $\text{Exp}(\lambda)$ is the exponential distribution with rate λ; $\text{NegBin}(k, p)$ is the negative-binomial distribution, i.e., the distribution of the sum of k i.i.d. geometric random variables with probability parameter p; and $\text{Erl}(k, \lambda)$ is the Erlang distribution, i.e., the distribution of the sum of k i.i.d. exponential random variables with rate λ.

For any function $f(x)$, we define the set $\text{argmin}_x f(x) = \{x\colon \forall\, y, f(x) \le f(y)\}$. In case this set is a singleton set $\{a\}$, we will write $\text{argmin}_x f(x) = a$.

3. ANALYSIS OVERVIEW

Below we highlight the main ideas and techniques used in the analysis of our main results: the upper bound of Theorem 1, and the lower bound of Theorem 2.

Upper Bound

Our proof of Theorem 1 relies on a new coupling argument. This argument can be viewed as an extension of a basic coupling technique first used in [24] to relate the rumor spreading time of the asynchronous push algorithm (denoted *push-a*) with that of synchronous *push*. The following simple coupling was proposed there: Once it gets informed, each node v contacts its neighbors in the exact same order in both algorithms *push* and *push-a*. I.e., if r_v is the round when v gets informed in *push*, t_v is the time when v gets informed in *push-a*, and $t_{v,i}$ is the i-th time that v's clock ticks after time t_v, then v pushes the rumor to the same node in round $r_v + i$ of *push*, and at time $t_{v,i}$ in *push-a*. Consider a path $v_0 = u, v_1, \ldots, v_l = v$ through which v receives the rumor in *push*, where u is the initially informed node and node v_{i+1} learns the rumor from v_i, for $0 \le i < l$; let also $d_i = r_{v_{i+1}} - r_{v_i}$. The time before v gets informed is then $r_v = \sum_i d_i$. Consider now the same path in *push-a* and let $\tau_i = t_{v_{i+1}} - t_{v_i}$; then $t_v = \sum_i \tau_i$. The coupling implies that in *push-a*, v_{i+1} learns the rumor no later than in time t_{v_i,d_i}, when v_i pushes the rumor to it (but it may learn the rumor sooner, from another neighbor). And since for any j we have in expectation that $t_{v,j} - t_v = j$, it follows that $\text{E}[\tau_i \mid d_i] \le d_i$, and thus $\text{E}[t_v] \le \text{E}[r_v]$. This expectation bound can be turned into a high-probability bound, but we do not discuss the details here.

The above simple technique does not work for push-pull as there is no obvious way to couple pull operations. In fact, as

[3] An alterative measure for the asynchronous rumor spreading time would be the total number of *steps* before all nodes get informed. The ratio of this number of steps over n is equal in expectation to the number of time units before all nodes get informed.

far as we know, there is no coupling technique in the rumor spreading literature that achieves such a coupling between pull operations. Our analysis does exactly that: it provides a method to couple pull operations to achieve the same effect as the above natural coupling between push operations.

The coupling we propose and, especially, its analysis are somewhat involved, so we give the high-level ideas first. For each node v and each neighbor w of v we consider an independent exponential random variable $Y_{v,w}$ with rate $\lambda_v = 1/\deg(v)$. In *pp-a*, we set to be equal to $Y_{v,w}$ the time between the point t_w when w gets informed, and the point when v contacts w for the first time after t_w in order to pull the rumor, provided that v is still not informed by that time. For *pp*, we would like to set equal to $Y_{v,w}$ (precisely, to $\lceil Y_{v,w} \rceil$) the number of rounds after r_w, when v pulls the rumor from w, provided v is still not informed by that round. Doing so, however, results in a number of issues

(1) The probability for v to pull the rumor from any specific informed neighbor in a round of *pp* is $1 - e^{-1/\deg(v)}$ under this coupling, which is slightly smaller than $1/\deg(v)$. For this reason we set $\lambda_v = 2/\deg(v)$ (while we use $2Y_{v,w} \sim \mathrm{Exp}(1/\deg(v))$ rather than $Y_{v,w}$ in *pp-a*).

(2) Node v may have to contact more than one informed node in a round. However, this is not a real issue as it suffices to pull from just one of them to get informed.

(3) The overall probability that v succeeds in pulling the rumor in a given round is not exactly the right one: if fewer than some constant fraction of v's neighbors are informed then the probability is larger that it should be, otherwise it is smaller. Consider for example the two extreme cases: if v has only one informed neighbor the probability of pulling from it is $1 - e^{-2/\deg(v)} \approx 2/\deg(v)$ instead of $1/\deg(v)$; while if all of v's neighbors are informed, the probability of pulling from (at least) one of them is $1 - e^{-2}$, which is smaller than the correct value of 1. The former is not an actual problem, as it just speeds up *pp* which only makes our result stronger. The latter however is a problem. To solve it, we impose that as soon as v has at least $\deg(v)/2$ informed neighbors (and is still not informed), it pulls the rumor in the next round r^* with probability 1.

This last modification requires some subtle handing in order to work: Among all its (at least $\deg(v)/2$) informed neighbors, we let v pull the rumor from the neighbor w^* that minimizes the quantity $t_w + Y_{v,w}$. We show that this implies that the value of $r_{w^*} + Y_{v,w^*}$ is not much larger than r^*, in particular, $r_{w^*} + Y_{v,w^*} = r^* + O(1)$ in expectation. Hence, this case is not very different from the setting without the last modification, i.e., the setting in which v pulls the rumor from w in a round r such that $r_w + \lceil Y_{v,w} \rceil = r$. The intuition why we have $r_{w^*} + Y_{v,w^*} - r^* = O(1)$ in expectation is simple: for each of the at least $\deg(v)/2$ informed neighbors w of v, the difference $r_w + Y_{v,w} - r^*$ is an independent exponential random variable with rate $2/\deg(v)$, and $r_{w^*} + Y_{v,w^*} - r^*$ is the minimum of them. This implies that $r_{w^*} + Y_{v,w^*} - r^*$ is exponentially distributed with a rate of at least 1, thus its expectation is at most 1.

The detailed analysis is given in Section 4

Lower Bound

Our proof of Theorem 2 can be viewed as a refinement of the analysis technique used in [1] to prove a lower bound of $\Omega(n^{-2/3})$ on the ratio of asynchronous over synchronous push-pull rumor spreading times. However, our analysis introduces several new ideas, in order to improve the lower bound to $\Omega(n^{-1/2})$.

We use a coupling argument which, roughly speaking, allows us to divide the sequence of steps in *pp-a* into *blocks*, and map each block to one or more consecutive rounds of *pp*, so that the set of informed nodes in *pp-a* after each block of steps is a subset of the set of informed nodes in *pp* after the last round corresponding to that block. The rounds mapped to a block contain all pairwise communications of the steps in the block. *On average* a block contains $\Theta(\sqrt{n})$ steps and is mapped to $O(1)$ rounds of *pp* (even though some blocks may be much smaller and mapped to a larger number of rounds as discussed below). Since the expected time between consecutive steps is $1/n$, the desired bound follows.

We have two types of blocks: *normal* and *special* blocks. A normal block consists of at most $\sqrt{n}$ steps and is mapped to a single round of *pp*, while a special block consists of a single step and may be mapped to more than one round. Let $S_1 = (x_1, y_1), S_2 = (x_2, y_2), \ldots$ be the sequence of steps in *pp-a*, where $S_i = (x_i, y_i)$ denotes that node x_i contacts node y_i in step i. The first of the blocks into which the sequence of S_i is partitioned is a normal block (starting from S_1). A normal block B starting from S_i consists of the steps $S_i, \ldots, S_{j-1}$, where $j > i$ is the smallest index for which some of the following three conditions hold:

(1) $j - i = \sqrt{n}$, i.e., B contains the maximum allowed number, $\sqrt{n}$, of steps.

(2) $x_j = x_k$ or $x_j = y_k$ for some $k \in \{i, \ldots, j-1\}$, i.e., in some of the steps i up to $j-1$, node x_j either contacted or was contacted by a node. We say that S_j is *left-incompatible* with $(S_i, \ldots, S_{j-1})$.

(3) Node y_j got informed in one of the steps i up to $j-1$. We say that S_j is *right-incompatible* with $(S_i, \ldots, S_{j-1})$.

Conditions (2) and (3) above ensure that all steps of a normal block can be mapped to a single round of *pp*. In particular, (2) implies that no node contacts more than one neighbor in a single block, while (2) and (3) together prevent the following undesirable scenario: a node v gets informed by some neighbor (via push or pull), and then during the same block a non-informed neighbor of v pulls the rumor from v. Clearly, this scenario cannot take place in just one round of *pp*. The single round of *pp* that corresponds to a normal block contains all the pairwise communications of the steps in that block.

If a normal block $B = (S_i, \ldots, S_{j-1})$ ends because condition (1) or (2) is met, then the next block is also normal. If this is not the case, and thus $S_j = (x_j, y_j)$ is right-incompatible with B, then the next block is a special block. The reason why we have to treat this case differently is the following. We want to ensure that in each round of *pp*, the neighbor that each node contacts is chosen independently of all other choices (in the same or previous rounds). If r is the round that corresponds to block B, treating the next block as normal despite knowing that S_j is right-incompatible with B would imply that (with probability 1) some of the nodes that got informed in round r will be contacted by at least one node in round $r+1$. And this could introduce dependencies between nodes' random choices in round $r+1$ and the past. Note that knowing instead that S_j is *left*-incompatible with B does not cause the same problems, because each round contains one communication pair (v, v') for *every* node v (and thus for node x_j).

We handle a special block as follows: We run a number of rounds in *pp* (independent of the sequence of steps so far) until we have a round that contains at least one communication pair (a, b) that is right-incompatible with the normal block B which is before the special block; these rounds will be the rounds of *pp* that correspond to the special block. Moreover we "discard" the right-incompatible pair S_j, and replace it by pair (a, b) in *pp-a*. If more than one such pair (a, b) exists in the round, then we choose one of them at random, from some appropriate distribution.

Establishing that the above coupling is a proper coupling is a bit technical. But once we have done so, proving the desired lower bound is not difficult. The proof goes roughly as follows. We argue that t steps of *pp-a* are mapped to at most $O(t/\sqrt{n} + \sqrt{n})$ rounds of *pp* in expectation: At most $t/\sqrt{n}$ rounds correspond to blocks of size exactly $\sqrt{n}$. Further, at most $t/\sqrt{n}$ rounds in expectation corresponding to blocks that end because a left-incompatible pair is encountered; the reason is that such a pair is encountered with probability $O(1/\sqrt{n})$ in a step, given that the block size is at most $\sqrt{n}$. It remains to bound the number of rounds corresponding to special blocks. By (3), a node v can result in a special block only if the following scenario occurs: v gets informed during a step in some block B, and after that, v is contacted by some neighbor before $\sqrt{n}$ additional steps are taken (otherwise B finishes due to (1)). Let $\pi(v)$ denote the probability that v is contacted in a random step (this probability depends on the degrees of v's neighbors). The probability that v causes a special block is then bounded by $\sqrt{n} \cdot \pi(v)$. Moreover, if that special block exists, the expected number of rounds mapped to the block is upper bounded by the expected number of rounds until v is contacted by some neighbor. The probability that v is contacted by at least one neighbor in a round can be easily shown to be $q_v \geq 1 - e^{-n\pi(v)}$. So, the expected number of rounds corresponding to the special block is at most $1/q_v$. It follows that the expected total number of special blocks is at most $\sum_v \frac{\sqrt{n}\cdot\pi(v)}{1-e^{-n\pi(v)}}$, which is easily shown to be $O(\sqrt{n})$.

Applying the above result, that t steps of *pp-a* are mapped to at most $O(t/\sqrt{n}+\sqrt{n})$ rounds of *pp* in expectation, to the total number of steps before all nodes get informed in *pp-a* (thus $t \geq n-1$), and using that the expected time between two consecutive steps is $1/n$, gives that the expected total number of rounds in *pp* is by at most a factor of $O(\sqrt{n})$ larger than the expected time in *pp-a*.

The detailed analysis can be found in Section 5.

4. UPPER BOUND ANALYSIS

We prove that on any graph, the high-probability rumor spreading time of asynchronous push-pull is bounded by the high-probability rumor spreading time of synchronous push-pull plus a logarithmic term.

THEOREM 4. *For any connected n-node graph $G = (V, E)$, and any vertex $u \in V$ of G, we have $T_{1/n}(pp\text{-}a, G, u) = O(T_{1/n}(pp, G, u) + \log n)$.*

The proof is based on a coupling argument. For the sake of comprehension, we define two auxiliary rumor spreading processes, *ppx* and *ppy*, and present the coupling in three steps: first we couple *pp* with *ppx*, then *ppx* with *ppy*, and finally *ppy* with *pp-a*.

Processes *ppx* and *ppy* are very similar to *pp*, except that they use different rules to decide from which neighbor a non-informed node tries to pulls the rumor in a round. We point out that *ppx* and *ppy* are not realistic rumor spreading algorithms, as they assume that at any time, a node knows the set of its informed neighbors; we introduce these processes just to facilitate our analysis. We describe process *ppx* first.

DEFINITION 5 (PROCESS *ppx*). *Process ppx is the following synchronous rumor spreading algorithm: For each round r and $v \in V$, (1) if v is informed before round r, then in round r, v pushes the rumor to a random neighbor; (2) if v is not informed before round r and has k informed neighbors at that time, then with probability*

$$p = \begin{cases} 1 - e^{-2k/\deg(v)}, & \text{if } k < \deg(v)/2; \\ 1, & \text{if } k \geq \deg(v)/2, \end{cases}$$

v pulls the rumor from a random informed neighbor in round r, while with the remaining probability, $1-p$, v does not pull the rumor in this round.

The rumor spreading time for *ppx* is dominated by that for *pp*. The proof (which can be found in the full version of the paper) relies on the observation that, for the same set of informed nodes, a non-informed node is more likely to pull the rumor in *ppx* than in *pp* in the next round.

LEMMA 6. *$T(ppx, G, u) \preccurlyeq T(pp, G, u)$.*

The second auxiliary process we introduce, *ppy*, is identical to *ppx* except for the probability with which a node pulls the rumor in a round.

DEFINITION 7 (PROCESS *ppy*). *Process ppy is the following synchronous rumor spreading algorithm: For each round r and $v \in V$, (1) if v is informed before round r, then in round r, v pushes the rumor to a random neighbor; (2) if v is not informed before round r and has k informed neighbors at that time, then with probability*

$$p = 1 - e^{-2k/\deg(v)},$$

v pulls the rumor from a random informed neighbor in round r, while with the remaining probability, $1-p$, v does not pull the rumor in this round.

In Lemma 9 below, we bound the rumor spreading time for *ppy* in terms of the rumor spreading time for *ppx*. First we provide a technical lemma that we will need. This lemma computes the conditional distribution of the minimum of a collection of independent exponential random variables, given some limited information about them. (The proof can be found in the full version of the paper)

LEMMA 8. *Let $Z_1, \ldots, Z_k$ be i.i.d. random variables with $Z_i \sim Exp(\lambda)$, and let $J = \operatorname{argmin}_i Z_i$. For $\alpha_1, \ldots, \alpha_k$ arbitrary non-negative integers, let $Z = \min_i\{Z_i - \alpha_i\}$, and let $\mathcal{A}$ be the event: $\forall\, i, Z_i > \alpha_i$. Then $(Z \mid J = j, \mathcal{A}) \sim Exp(k\lambda)$, i.e., for any $t \geq 0$, $\Pr[Z \leq t \mid J = j, \mathcal{A}] = 1 - e^{-k\lambda t}$.*

We remark that from the memoryless property of exponential distribution it is immediate that $(Z \mid \mathcal{A}) \sim \text{Exp}(k\lambda)$. The lemma above says that conditioning also on J does not add any information. We now proceed to the main lemma.

LEMMA 9. *$T_\delta(ppy, G, u) = O\big(T_\delta(ppx, G, u) + \log(n/\delta)\big)$, for any $0 < \delta \leq 1/2$.*

PROOF. We define a coupling of the random choices in *ppx* and *ppy*. In this coupling, let r_v and r'_v denote the rounds in which node v gets informed in *ppx* and *ppy*, respectively (for the source u, $r_u = r'_u = 0$). We will show that for any $v \in V$, with probability at least $1 - \delta/2n$ we have $r'_v \leq 2r_v + O\big(\log(n/\delta)\big)$. A union bound then completes the proof.

To facilitate the coupling we define the following collection of random variables. For each $v \in V$, $i \geq 1$, and $w \in \Gamma(v)$, let $X_{v,i}$ and $Y_{v,w}$ be random variables with $X_{v,i} \sim \mathrm{Unif}(\Gamma(v))$ and $Y_{v,w} \sim \mathrm{Exp}(\lambda_v)$, with $\lambda_v = 2/\deg(v)$ (i.e., $X_{v,i}$ is a random neighbor of v, and $Y_{v,w}$ is an exponential random variable with rate λ_v). We assume that all these random variables are mutually independent.

For *push* operations, the coupling states that each node pushes the rumor to the same neighbor in both processes, in the i-th round after the node gets informed. Precisely, for each $v \in V$ and $i \geq 1$, v pushes the rumor to node $X_{v,i}$ in round $r_v + i$ of *ppx*, and in round $r'_v + i$ of *ppy*.

For *pull* operations, the coupling is more involved. For *ppy*, for each pair of adjacent nodes v, w, if w gets informed before v (i.e., $r'_w < r'_v$), and v is still not informed before round $r'_w + \lceil Y_{v,w} \rceil$, then we let v pull the rumor from w in round $r'_w + \lceil Y_{v,w} \rceil$. The formal definition takes also into account the possibility that $r'_w + \lceil Y_{v,w} \rceil = r'_x + \lceil Y_{v,x} \rceil$ for two distinct neighbors w, x of u, in which case the tie is broken based on the actual values of $Y_{v,w}$ and $Y_{v,x}$ (rather than their rounded up values). Precisely, for any $v \in V \setminus \{u\}$, if v does not get informed by a push operation before round $t = \min_{w\in\Gamma(v)}\{r'_w + \lceil Y_{v,w} \rceil\}$, then in round t, v pulls the rumor from node $\mathrm{argmin}_{w\in\Gamma(v)}\{r'_w + Y_{v,w}\}$, i.e., the neighbor w that minimizes $r'_w + Y_{v,w}$.[4] Clearly, for this neighbor w, $r'_w + \lceil Y_{v,w} \rceil = t$.

For *ppx*, we use a similar coupling rule except that we need to enforce that, as soon as half of v's neighbors get informed, v will pull the rumor in the next round with probability 1, if it is not already informed. The neighbor w from which v pulls the rumor in this case is the (currently informed) neighbor that minimizes $r_w + Y_{v,w}$. Precisely, for any $v \in V \setminus \{u\}$, if $t = \min_{w\in\Gamma(v)}\{r_w + \lceil Y_{v,w} \rceil\}$ and z is the first round by the end of which at least $\deg(v)/2$ of v's neighbors have been informed, we distinguish the following two cases:

(i) If $t \leq z$ and v does not get informed by a push operation before round t, then in round t, v pulls the rumor from node $\mathrm{argmin}_{w\in\Gamma(v)}\{r_w + Y_{v,w}\}$. So, in this case the rule is the same as for *ppy*.

(ii) If $t > z$ and v does not get informed by a push operation before round $z + 1$, then in round $z + 1$, v pulls the rumor from node $\mathrm{argmin}_{w\in\Gamma(v)\colon r_w \leq z}\{r_w + Y_{v,w}\}$.

It is not hard to verify that the above coupling is valid, in the sense that the marginal distributions of the two processes are the correct ones: For push operations there is nothing to argue about, so we focus on pull operations. In *ppy*, if before round r node v is still not informed, and its set of informed neighbors is S with $|S| = k$, then the probability that v pulls the rumor in round r is the same as the conditional probability that $\min_{w\in S}\{r'_w + Y_{v,w}\} \leq r$, given that $\min_{w\in S}\{r'_w + Y_{v,w}\} > r - 1$. Since $Y_{v,w} \sim \mathrm{Exp}(\lambda_v)$, from the memoryless property of the exponential distribution and the independence between random variables $Y_{v,w}$, it follows that the above conditional probability is

$$1 - \big(1 - \Pr[r'_w + Y_{v,w} > r \mid r'_w + Y_{v,w} > r - 1]\big)^k = 1 - e^{-k\lambda_v} = 1 - e^{-2k/\deg(v)},$$

which is the right probability according to Definition 7. Moreover, if v *does* pull the rumor in round r, then it is equally likely to pull it from any of its informed neighbors in S, as the conditional random variables $(r'_w + Y_{v,w} \mid r'_w + Y_{v,w} > r - 1)$, for $w \in S$, have the same distribution and are independent. A very similar argument shows that our coupling yields the correct distribution for the pull operations in *ppx*, as well.

Next we show that with probability at least $1 - \delta/2n$, $r'_v \leq 2r_v + O\big(\log(n/\delta)\big)$.

Let π_v be a path through which the rumor reaches node $v \in V \setminus \{u\}$ in *ppx*. Formally, $\pi_v = v_0 v_1 \ldots v_l$, where $v_0 = u$, $v_l = v$, and for each $0 \leq i < l$, v_i is a node from which v_{i+1} receives the rumor for the first time (i.e., in round $r_{v_{i+1}}$). We express r_v and r'_v as

$$r_v = \sum_{0\leq i<l}(r_{v_{i+1}} - r_{v_i}), \quad r'_v = \sum_{0\leq i<l}(r'_{v_{i+1}} - r'_{v_i}).$$

(Note that these equations hold for any collection of v_i, not just for the specific definition of π_v.) For $0 \leq i < l$, let $d_i = r_{v_{i+1}} - r_{v_i}$ and $d'_i = r'_{v_{i+1}} - r'_{v_i}$. We will show that the difference $d'_i - d_i$ is dominated by a geometric random variable with constant expectation. We distinguish three cases, depending on how v_{i+1} gets the rumor from v_i in *ppx*.

Case 1: v_i pushes the rumor to v_{i+1}, in round $r_{v_{i+1}}$ of *ppx*. In this case, we have $r_{v_{i+1}} = r_{v_i} + \min\{j\colon X_{v_i,j} = v_{i+1}\}$, and thus $d_i = \min\{j\colon X_{v_i,j} = v_{i+1}\}$. Similarly, in *ppy*, v_i pushes the rumor to v_{i+1} in round $r'_{v_i} + \min\{j\colon X_{v_i,j} = v_{i+1}\}$. Thus v_{i+1} gets informed in *ppy* no later than in this round, i.e., $r'_{v_{i+1}} \leq r'_{v_i} + \min\{j\colon X_{v_i,j} = v_{i+1}\}$, and so, $d'_i \leq \min\{j\colon X_{v_i,j} = v_{i+1}\} = d_i$.

Case 2: v_{i+1} pulls the rumor from v_i, in round $r_{v_{i+1}}$ of *ppx*, and before that round fewer than half of v_{i+1}'s neighbors are informed. In this case, $r_{v_{i+1}} = r_{v_i} + \lceil Y_{v_{i+1},v_i} \rceil$, thus $d_i = \lceil Y_{v_{i+1},v_i} \rceil$. Similarly, in *ppy*, v_{i+1} gets informed no later than in round $r'_{v_i} + \lceil Y_{v_{i+1},v_i} \rceil$, because if v_{i+1} is still not informed before that round, it will pull the rumor in round $r'_{v_i} + \lceil Y_{v_{i+1},v_i} \rceil$ (from v_i or some other informed neighbor). It follows that $d'_i \leq \lceil Y_{v_{i+1},v_i} \rceil = d_i$.

Case 3: v_{i+1} pulls the rumor from v_i, in round $r_{v_{i+1}}$ of *ppx*, and before that round at least half of v_{i+1}'s neighbors are informed. This case is more involved. As in case 2, we have $d'_i \leq \lceil Y_{v_{i+1},v_i} \rceil$, but now it is possible that $d_i < \lceil Y_{v_{i+1},v_i} \rceil$. We will use Lemma 8 to bound $d'_i - d_i$.

Let z be the first round in *ppx* after which at least half of the neighbors of v_{i+1} are informed, and let S be the set of those informed neighbors (so $|S| \geq \deg(v_{i+1})/2$). Then $r_{v_{i+1}} = z + 1$. Also, from the coupling (case (ii)), $v_i = \mathrm{argmin}_{w\in S}\{r_w + Y_{v_{i+1},w}\}$, thus $r_{v_i} + Y_{v_{i+1},v_i} < r_w + Y_{v_{i+1},w}$ for all $w \in S \setminus \{v_i\}$. Notice that it is possible to have $r_{v_i} + Y_{v_{i+1},v_i} > z + 1$, which implies $d_i < Y_{v_{i+1},v_i}$.

Let us fix the pairwise communications that occur in *all* rounds of *ppx*, and for each $w \in S$, let $Z_w = r_w + Y_{v_{i+1},w} - z$. The set $\{Z_w\}_{w\in S}$ is then a collection of i.i.d. random variables with distribution $\mathrm{Exp}(\lambda_{v_{i+1}})$, and we know that $\mathrm{argmin}_{w\in S}\{Z_w\} = v_i$.

[4]Since $r'_w + Y_{v,w}$ is a continuous random variable, the probability this quantity is the same for two distinct w is 0.

Consider now process *ppy*. To simplify exposition we shift all round numbers in *ppy* by an appropriate offset, so that the round number in which v_i gets informed in *ppy* is the same as the one in *ppx*. E.g., if v gets informed after k rounds in *ppx* and after ℓ rounds in *ppy*, we add an offset of $(k-\ell)$ to all round numbers in *ppy*. So, the i-th round in *ppx* has number $i+(k-\ell)$, and the round when v gets informed is $r'_{v_i} = k = r_{v_i}$.

If $r'_{v_{i+1}} \le z$, then $d'_i = r'_{v_{i+1}} - r'_{v_i} = r'_{v_{i+1}} - r_{v_i} \le z - r_{v_i}$, and since $d_i = (z+1) - r_{v_i}$, we have $d'_i < d_i$. In the following we assume that $r'_{v_{i+1}} > z$, and bound the quantity $d'_i - d_i + 1 = r'_{v_{i+1}} - z$ using Lemma 8.

Let us fix all random communications in the first z rounds of *ppy*, in a way that respects our coupling (recall we have already fixed the communications in all rounds of *ppx*). Revealing these random choices in *ppy* discloses additional information about variables $Y_{v_{i+1},w}$, and thus also about the variables $Z_w = r_w + Y_{v_{i+1},w} - z$. Precisely, the additional information is that for each $w \in S$, $r'_w + Y_{v_{i+1},w} - z > 0$, which follows from the assumption $r'_{v_{i+1}} > z$. Therefore, $Z_w > r_w - r'_w$, and since it is also $Z_w > 0$ (from *ppx*), we have that for any $w \in S$, $Z_w > \alpha_w$, where $\alpha_w = \max\{0, r_w - r'_w\}$.

To summarize, we have that random variables Z_w, $w \in S$, are independent with distribution $\mathrm{Exp}(\lambda_{v_{i+1}})$, and we know that $\mathrm{argmin}_{w\in S}\{Z_w\} = v_i$ and $Z_w > \alpha_w = \max\{0, r_w - r'_w\}$. Letting $Z = \min_{w\in S}\{Z_w - \alpha_w\}$, we can apply Lemma 8 to obtain for $t \ge 0$,

$$\Pr[Z \le t] = 1 - e^{-|S|\lambda_{v_{i+1}} t} \ge 1 - e^{-t},$$

because $|S| \ge \deg(v_{i+1})/2$ and $\lambda_{v_{i+1}} = 2/\deg(v_{i+1})$.

We can now argue that $r'_{v_{i+1}} - z \le \lceil Z \rceil$, and use the result above to bound the distribution of $r'_{v_{i+1}} - z$.

$$\begin{aligned} Z &= \min_{w\in S}\{Z_w - \alpha_w\} \\ &= \min_{w\in S}\{(r_w + Y_{v_{i+1},w} - z) - \max\{0, r_w - r'_w\}\} \\ &= \min_{w\in S}\{\min\{r_w + Y_{v_{i+1},w},\, r'_w + Y_{v_{i+1},w}\}\} - z \\ &= \min\{r_{v_i} + Y_{v_{i+1},v_i},\, \min_{w\in S}\{r'_w + Y_{v_{i+1},w}\}\} - z \\ &= \min_{w\in S}\{r'_w + Y_{v_{i+1},w}\} - z, \end{aligned}$$

where the last equation holds because $r_{v_i} = r'_{v_i}$. Since $\min_{w\in S}\{r'_w + \lceil Y_{v_{i+1},w}\rceil\} \ge \min_{w\in\Gamma(v_{i+1})}\{r'_w + \lceil Y_{v_{i+1},w}\rceil\} = r'_{v_{i+1}}$, the equation above yields $\lceil Z \rceil \ge r'_{v_{i+1}} - z$. From this and the bound $\Pr[Z \le t] \ge 1 - e^{-t}$ shown earlier, it follows that for any *integer* t, $\Pr[r'_{v_{i+1}} - z \le t] \ge 1 - e^{-t}$.

Finally, substituting $d'_i = r'_{v_{i+1}} - r'_{v_i} = r'_{v_{i+1}} - r_{v_i}$ and $d_i = (z+1) - r_{v_i}$, yields that for any integer $t \ge 0$,

$$\Pr[d'_i - d_i + 1 \le t] \ge 1 - e^{-t}.$$

This completes the analysis of Case 3.

In each of the above cases, 1–3, we have showed that either (1) $d'_i \le d_i$, or (2) given all communications that take place in every round of *ppx*, and in the first $z = r'_{v_i} + d_i - 1$ rounds of *ppy*, we have $d'_i - d_i + 1 \le t$ with probability at least $1 - e^{-t}$, for any integer $t \ge 0$. Note that (1) implies (2).

Let us fix all random communications that take place in *ppx* (and thus all d_i). From (2) it follows that

$$\Pr[d'_i - d_i + 1 \le t \mid d'_1 \dots d'_{i-1}] \ge 1 - e^{-t}.$$

This says that random variable $Z_i = d'_i - d_i + 1$ is dominated by $\mathrm{Geom}(1/e)$, independently of $Z_1 \dots Z_{i-1}$. Using that $r_v = \sum_{0\le i<l} d_i$ and $r'_v = \sum_{0\le i<l} d'_i$, we obtain $r'_v - r_v + l = \sum_{0\le i<l}(d'_i - d_i + 1)$, and applying Lemma 15 (in the appendix) for the sum of the (dependent) random variables Z_i, $0 \le i < l$, we obtain that $r'_v - r_v + l \preccurlyeq \mathrm{NegBin}(l, 1-1/e)$. From this, it follows that with probability at least $1 - \delta/2n$, $r'_v - r_v + l \le 2l + O\big(\log(n/\delta)\big)$, and thus $r'_v \le 2r_v + O\big(\log(n/\delta)\big)$, as $l \le r_v$.

Taking the union bound over all v, gives that with probability at least $1-\delta/2$, we have for all v simultaneously that $r'_v \le 2r_v + O\big(\log(n/\delta)\big)$. And since by definition, with probability $1-\delta/2$, for all v we have $r_v \le T_{\delta/2}(ppx, G, u)$, another union bound gives that with probability at least $1-\delta$, we have for all v that $r'_v \le 2T_{\delta/2}(ppx, G, u) + O\big(\log(n/\delta)\big)$. This means $T_\delta(ppy, G, u) \le 2T_{\delta/2}(ppx, G, u) + O\big(\log(n/\delta)\big)$. Finally, observing that $T_{\delta/2}(ppx, G, u) \le 2T_\delta(ppx, G, u)$ for $\delta \le 1/2$, concludes the proof of Lemma 9. □

It remains to couple process *ppy* with the asynchronous push-pull rumor spreading algorithm, *pp-a*.

LEMMA 10. *$T_\delta(pp\text{-}a, G, u) = O\big(T_\delta(ppy, G, u) + \log(n/\delta)\big)$, for any $0 < \delta \le 1/2$.*

PROOF. The structure of the proof is similar to that for Lemma 9. The coupling is similar as well, even though the details and the justification are simpler in this case. Essentially, the coupling captures the intuition that *ppy* is just a discretized version of *pp-a*.

In this coupling, for each $v \in V$, let t_v be the *time* in which v gets informed in *pp-a*, and let r_v be the *round* in which v gets informed in *ppy*. We will show that for any $v \in V$, with probability at least $1-\delta/2n$, $t_v \le 4r_v + O\big(log(n/\delta)\big)$. A union bound then completes the proof.

We define the same random variables as for the coupling in the proof of Lemma 9: For any $v \in V$, $i \ge 1$, and $w \in \Gamma(v)$, let $X_{v,i}$ and $Y_{v,w}$ be random variables with $X_{v,i} \sim \mathrm{Unif}(\Gamma(v))$ and $Y_{v,w} \sim \mathrm{Exp}(\lambda_v)$, where $\lambda_v = 2/\deg(v)$; all these random variables are mutually independent. For *ppy*, we use the same coupling rules as in the proof of Lemma 9, and for *pp-a* we use a continuous-time version of those rules.

Precisely, for *push* operations the coupling is as follows. For each $v \in V$ and $i \ge 1$, v pushes the rumor to node $X_{v,i}$ in round $r_v + i$ of *ppy*, and similarly v pushes the rumor to $X_{v,i}$ at time $t_{v,i}$ in *pp-a*, where $t_{v,i}$ is the time at which v takes its i-th step after time t_v.

For *pull* operations, the coupling is as follows. In *ppy*, for each $v \in V\setminus\{u\}$, if v does not get informed by a push operation before round $r = \min_{w\in\Gamma(v)}\{r_w + \lceil Y_{v,w}\rceil\}$, then in round r, v pulls the rumor from node $\mathrm{argmin}_{w\in\Gamma(v)}\{r_w + Y_{v,w}\}$. In *pp-a*, for each $v \in V\setminus\{u\}$, if v does not get informed by a push operation before time $t = \min_{w\in\Gamma(v)}\{t_w + 2Y_{v,w}\}$, then at time t, v pulls the rumor from node $\mathrm{argmin}_{w\in\Gamma(v)}\{t_w + 2Y_{v,w}\}$ (the factor 2 is justified below).

The above coupling yields the correct marginal distribution for the two processes: For *ppy* the same argument applies as in the proof of Lemma 9, so we just argue about *pp-a*. The distribution of the push operations is clearly the right one, so we just need to argue about pull operations. For that, it is convenient to consider the view of *pp-a* in which for each (ordered) pair v, w of adjacent nodes, there is an independent poisson clock $C_{v,w}$ with rate $1/\deg(v)$, and for each time t at which the clock ticks, if right before the tick

v is still not informed and w is informed, then v pulls the rumor from w at time t. Our coupling sets the length L_w of the interval between time t_w and the next tick of clock $C_{v,w}$ to be $L_w = 2Y_{v,w}$. Since $Y_{v,w} \sim \mathrm{Exp}(2/\deg(v))$ it follows that $2Y_{v,w} \sim \mathrm{Exp}(1/\deg(v))$, and this is the correct distribution for L_w.

Next we show that with probability at least $1 - \delta/2n$, $t_v \leq 4r_v + O\big(log(n/\delta)\big)$.

Let π_v be a path through which the rumor reaches node $v \in V \setminus \{u\}$ in *ppy*. Formally, $\pi_v = v_0 v_1 \dots v_l$, where $v_0 = u$, $v_l = v$, and for each $0 \leq i < l$, v_i is a node from which v_{i+1} receives the rumor for the first time (i.e., in round $r_{v_{i+1}}$). In line with the proof of Lemma 9, we compare the random variables $d_i = r_{v_{i+1}} - r_{v_i}$ and $\tau_i = t_{v_{i+1}} - t_{v_i}$. We have two cases:

Case 1: v_{i+1} pulls the rumor from v_i in round $r_{v_{i+1}}$ of *ppy*. In this case, $r_{v_{i+1}} = r_{v_i} + \lceil Y_{v_{i+1},v_i} \rceil$, thus $d_i = \lceil Y_{v_{i+1},v_i} \rceil$. Similarly, in *pp-a*, v_{i+1} gets informed no later than in time $t_{v_i} + 2Y_{v_{i+1},v_i}$, because if v_{i+1} is still not informed by that time, it will pull the rumor from v_i. It follows that $\tau_i = t_{v_{i+1}} - t_{v_i} \leq 2Y_{v_{i+1},v_i} \leq 2d_i$.

Case 2: v_i pushes the rumor to v_{i+1} in round $r_{v_{i+1}}$ of *ppy*. In this case, $r_{v_{i+1}} = r_{v_i} + x$, where $x = \min\{j\colon X_{v_i,j} = v_{i+1}\}$, thus $d_i = x$. Similarly in *pp-a*, v_i pushes the rumor to v_{i+1} at time $t_{v_i,x}$ (i.e., in its x-th step after it gets informed). Thus, $t_{v_{i+1}} \leq t_{v_i,x}$, and $\tau_i \leq t_{v_i,x} - t_{v_i}$. Given the random communications in every round of *ppy*, and the steps in *pp-a* up to time t_{v_i}, we have $(t_{v_i,x} - t_{v_i}) \sim \mathrm{Erl}(x, 1)$, [5] and thus $\tau_i \preccurlyeq \mathrm{Erl}(x, 1)$.

To summarise, in each case above, either (1) $\tau_i \leq 2d_i$, or (2) given the communications in every round of *ppy*, and the steps in *pp-a* up to time t_{v_i}, we have $\tau_i \preccurlyeq \mathrm{Erl}(d_i, 1)$. Hence in both cases it holds $\tau_i - 2d_i \preccurlyeq \mathrm{Erl}(d_i, 1)$, given all communications in *ppy*, and the steps in *pp-a* up to time t_{v_i}.

Let us fix all communications in *ppy* (and thus all d_i). Using the result above, and a similar reasoning as in the proof of Lemma 9 to tackle dependencies between variables $\tau_i - 2d_i$ for different i, we obtain that $\sum_i(\tau_i - 2d_i)$ is dominated by the sum $\sum_i Z_i$ of independent random variables $Z_i \sim \mathrm{Erl}(d_i, 1)$, for $0 \leq i < l$. It follows $\sum_i(\tau_i - 2d_i) \preccurlyeq \mathrm{Erl}\big(\sum_i d_i,\ 1\big)$. Substituting $\sum_i \tau_i = t_v$ and $\sum_i d_i = r_v$ we obtain that $t_v - 2r_v \preccurlyeq \mathrm{Erl}(r_v, 1)$. Using that $\mathrm{Erl}(k, \lambda) \preccurlyeq \mathrm{NegBin}(k, 1 - e^{-\lambda})$, yields $t_v - 2r_v \preccurlyeq \mathrm{NegBin}(r_v, 1 - 1/e)$. From this, it follows that $t_v \leq 4r_v + O\big(\log(n/\delta)\big)$, with probability at least $1 - \delta/2n$. Using now the union bound as in the end of the proof of Lemma 9, we finish the proof of Lemma 10. □

Proof of Theorem 4. Combining Lemma 10, Lemma 9, and Lemma 6, we obtain that for any $0 < \delta \leq 1/2$,

$$\begin{aligned} T_\delta(pp\text{-}a, G, u) &= O\big(T_\delta(ppy, G, u) + \log(n/\delta)\big) \\ &= O\big(T_\delta(ppx, G, u) + \log(n/\delta)\big) \\ &= O\big(T_\delta(pp, G, u) + \log(n/\delta)\big). \end{aligned}$$

Setting $\delta = 1/n$, yields the statement of the theorem. □

5. LOWER BOUND ANALYSIS

We prove that the expected rumor spreading time of asynchronous push-pull is at least $\Omega(1/\sqrt{n})$ times the expected rumor spreading time of synchronous push-pull.

[5]Recall, $\mathrm{Erl}(k, \lambda)$ is the distribution of the sum of k independent exponential random variables with rate λ.

Theorem 11. *For any connected n-node graph $G = (V, E)$, and any vertex $u \in V$ of G, we have* $\mathrm{E}[T(pp\text{-}a, G, u)] = \Omega\big((1/\sqrt{n}) \cdot \mathrm{E}[T(pp, G, u)]\big)$.

The proof is based on a coupling between *pp-a* and *pp*.

Let $H = ((u_i, v_i))_{1 \leq i \leq k}$ be a sequence of k pairs of adjacent nodes, and let $I \subseteq V$ be a set of nodes. We say that a pair (x, y) of adjacent nodes is *left-incompatible* with H, if $x \in \{u_i\}_{1 \leq i \leq k} \cup \{v_i\}_{1 \leq i \leq k}$. We say that (x, y) is *right-incompatible* with H and I, if the next two conditions hold:

(1) The pair (x, y) is *not* left-incompatible with H.

(2) If k steps of *pp-a* are executed such that in the i-th step node u_i contacts node v_i, and before the first of these steps the set of informed nodes is I, then node y gets informed in one of those steps (note, this implies $y \notin I$).

We say that H is *incompatible-free* with I if no pair (u_i, v_i) in H is left-incompatible with $H_{i-1} = ((u_j, v_j))_{1 \leq j < i}$, or right-incompatible with H_{i-1}, I.

Remark 12. *Suppose that H is incompatible-free with I, and we execute the pairwise push-pull communications described in H, assuming that at the beginning the set of informed nodes is I. Then the final set of informed nodes is the same whether these communications take place sequentially (as in pp-a), or in parallel (as in a round of pp).*

Let S denote a random variable that is a pair (x, y) with $x \sim \mathrm{Unif}(V)$ and $(y \mid x) \sim \mathrm{Unif}(\Gamma(x))$, i.e., x is a random node and y a random neighbor of x. Let R denote a random variable that is an n-set of pairs $\{(v, z_v)\}_{v \in V}$, with one pair per node, such that $z_v \sim \mathrm{Unif}(\Gamma(v))$ and the random variables z_v, $v \in V$, are independent.

For the coupling of *pp* and *pp-a* we will use the following notation. For $t \geq 1$, let a_t denote the node that takes the t-th step in *pp-a*, and let b_t be the neighbor that a_t contacts in that step. For $r \geq 1$ and $v \in V$, let $c_{r,v}$ denote the neighbor that node v contacts in round r of *pp*. The coupling must ensure that $(a_t, b_t) \sim S$ and $\{(v, c_{r,v})\}_{v \in V} \sim R$. We define also the following random variables, which will facilitate the coupling. For $i, j \geq 1$, let $S_i = (x_i, y_i) \sim S$ and $R_{i,j} = \{(v, z_{i,j,v})\}_{v \in V} \sim R$. We assume that all variables $S_i, R_{i,j}$ are mutually independent.

We partition the sequence $(S_i)_{i \geq 1}$ into *blocks* of consecutive elements. Each block corresponds to a number of steps of *pp-a* equal to the size of the block, and to one or more rounds on *pp*. We distinguish between *normal* and *special* blocks. The first block starts with element S_1 and is normal. A normal block B starting from S_i consists of the elements $S_i, \dots, S_{j-1}$, where $j > i$ is the smallest index for which at least one of the following conditions holds:

(1) $j - i = \sqrt{n}$;

(2) S_j is left-incompatible with $H = (S_i, \dots, S_{j-1})$;

(3) S_j is right-incompatible with H and the set I of nodes that are informed before the i-th step in *pp-a*.

Note that normal block B is incompatible-free with I, and contains at most $\sqrt{n}$ elements. If $|B| = \sqrt{n}$, or the element S_j after the last element of B is left-incompatible with B, then the next block is a normal block as well. Otherwise, S_j is *right*-incompatible with B, I, and the next block is a special block; this special block contains just the single pair S_j. The block right after a special block is always normal.

We describe now the steps in *pp-a* and the rounds in *pp* that correspond to each block. Let $(B_k)_{k \geq 1}$ be the sequence of blocks in which $(S_i)_{i \geq 1}$ is partitioned, as described above.

Suppose that $B_k = (S_i, \ldots, S_{j-1})$ is a normal block. Then we set $(a_t, b_t) = S_t$, for all $i \le t < j$, i.e., the steps in *pp-a* are described precisely by the sequence of pairs in the block. In *pp* we have a single round r corresponding to B_k, and for this round we set $c_{r,x_t} = y_t$ for all $i \le t < j$; for the remaining nodes $v \in V \setminus \{x_i, \ldots, x_{j-1}\}$, we can assume that they do not contact any node in the round, as this can only increase the rumor spreading time of *pp*.

Suppose now that $B_k = (S_i)$ is a special block, thus S_i is right-incompatible with B_{k-1} and the set I of informed nodes in *pp-a* before the first step in block B_{k-1}. In this case, the step (a_i, b_i) in *pp-a* that corresponds to B_k may be different than S_i, and we may have more than one rounds in *pp*; the communications in those rounds are described by the sets $R_{i,j}$, $j \ge 1$. Let j^* be the smallest j such that $R_{i,j}$ contains at least one element that is right-incompatible with B_{k-1}, I. Then for block B_k we will have j^* rounds in *pp*: if r is the index of the round corresponding to the previous block B_{k-1}, then for $1 \le j \le j^*$, we set $\{(v, c_{r+j,v})\}_{v\in V} = R_{i,j}$. For *pp-a* we set (a_i, b_i) to be a pair from R_{i,j^*} that is right-incompatible with B_{k-1}, I. If more than one such pair exists, we choose one as follows. Let A be the set containing all possible pairs of adjacent nodes that are right-incompatible with B_{k-1}, I. We set (a_i, b_i) to be an element of the set $D = A \cap R_{i,j^*}$ selected at random according to a distribution $\mu_{A|D}$ with the following property: Let

$$\mu_A(a,b) = \sum_{C\subseteq A} \Big(\mu_{A|C}(a,b) \cdot \Pr[A \cap R = C \mid A \cap R \ne \emptyset]\Big);$$

then

$$\forall\, (a,b) \in A, \quad \mu_A(a,b) = \Pr[S = (a,b) \mid S \in A]. \qquad (1)$$

I.e., $\mu_{A|D}$ is such that the pair (a,b) chosen has the same distribution as the pair S in a random step of *pp-a*, given that S is right-incompatible with B_{k-1}, I (μ_A averages over all possible rounds containing at least one right-incompatible pair.)

In the full version of the paper, we show that the above probability distributions $\mu_{A|D}$ exist, and establish that the coupling is valid, i.e., the marginal distribution of each process is the right one. In the full version, we also prove the next simple lemma. For $k \ge 0$, let $I_k(\textit{pp-a})$ be the set of informed nodes in *pp-a* after the steps corresponding to the first k blocks, and let $I_k(\textit{pp})$ be the set of informed nodes in *pp* after the rounds corresponding to the first k blocks.

LEMMA 13. *For any $k \ge 0$, $I_k(\textit{pp-a}) \subseteq I_k(\textit{pp})$.*

Next we bound the expected number of rounds in *pp* that correspond to the steps in *pp-a* before all nodes get informed. Let I_t be the set of informed nodes in *pp-a* after the first t steps, and let $\tau = \min\{t \colon I_t = V\}$ be the number of steps before all nodes get informed. Let ρ_t be the number of rounds in *pp* that correspond to the first k blocks, where k is the index of the block containing S_t.

LEMMA 14. $\mathrm{E}[\rho_\tau] = O\left(\mathrm{E}[\tau]/\sqrt{n} + \sqrt{n}\right)$.

PROOF. Let $B_1, \ldots, B_{k_t}$ denote the blocks into which the sequence $S_1, \ldots, S_t$ is partitioned. We decompose ρ_t into the following four terms:

The number $\rho_{t,full}$ of rounds corresponding to (normal) blocks B_k, $k \le k_t$, with $|B_k| = \sqrt{n}$.

The number $\rho_{t,left}$ of rounds corresponding to normal blocks $B_k = (S_i, \ldots, S_{j-1})$, with $k \le k_t$ and $|B_k| < \sqrt{n}$, such that S_j is left-incompatible with B_k.

The number $\rho_{t,right}$ of rounds corresponding to normal blocks $B_k = (S_i, \ldots, S_{j-1})$, with $k \le k_t$ and $|B_k| < \sqrt{n}$, such that S_j is right-incompatible with B_k and the set I of informed nodes in *pp-a* before step i.

The number $\rho_{t,special}$ of rounds corresponding to special blocks.

We have $\rho_t = \rho_{t,full} + \rho_{t,left} + \rho_{t,right} + \rho_{t,special}$. Since $\rho_{t,full} \le \frac{t}{\sqrt{n}}$, and $\rho_{t,right} \le \rho_{t,special} + 1$, we have that $\rho_t \le \frac{t}{\sqrt{n}} + \rho_{t,left} + 2\rho_{t,special} + 1$. Letting $t = \tau$ and taking the expectation of both sides yields

$$\mathrm{E}[\rho_\tau] = O\left(\mathrm{E}[\tau]/\sqrt{n} + \mathrm{E}[\rho_{\tau,left}] + \mathrm{E}[\rho_{\tau,special}] + 1\right).$$

We show $\mathrm{E}[\rho_{\tau,left}] \le 2\mathrm{E}[\tau]/\sqrt{n}$ and $\mathrm{E}[\rho_{\tau,special}] \le 2\sqrt{n}$. Substituting these above yields the claim.

The bound on $\rho_{\tau,left}$ is based on the following observation. For any $t \ge 1$, and any way of fixing the first $t-1$ steps of *pp-a*, the probability that $S_t = (x_t, y_t)$ is left-incompatible with $(S_i, \ldots, S_{t-1})$, where S_i is the first element in the block containing S_{t-1}, is at most $\frac{2(t-i)}{n} \le \frac{2}{\sqrt{n}}$. The reason is that x_t is chosen uniformly at random among all n nodes, and at most $2(t-i)$ distinct nodes appear in the pairs $S_i, \ldots, S_{t-1}$, while the number of those pairs is $t - i \le \sqrt{n}$.

For $t \ge 0$, define $Z_t = \rho_{t,left} - \frac{2t}{\sqrt{n}}$. The sequence $(Z_t)_{t\ge0}$ is a supermartingale with respect to $(X_t)_{t\ge1}$, where $X_t = (a_t, b_t)$, because: (1) $\rho_{t,left}$ (and thus Z_t) is a function of $X_1, \ldots, X_t$, as this sequence completely determines the collection of normal blocks into which $S_1, \ldots, S_t$ are divided; and (2) for $t \ge 1$, $\mathrm{E}[Z_t - Z_{t-1} \mid X_1 \ldots X_{t-1}] = \Pr[\rho_{t,left} - \rho_{t-1,left} = 1 \mid X_1 \ldots X_{t-1}] - \frac{2}{\sqrt{n}} \le 0$, as we argued above.

We use the optional stopping theorem for this supermartingale sequence and stopping time τ. Since for any $t \ge 1$, $|Z_t - Z_{t-1}| \le 1$, and $\mathrm{E}[\tau]$ is finite (bounded by $n^2 \log n$), the optional stopping theorem yields $\mathrm{E}[Z_\tau] \le \mathrm{E}[Z_0]$. Substituting $Z_\tau = \rho_{\tau,left} - \frac{2\tau}{\sqrt{n}}$ and $Z_0 = 0$, yields $\mathrm{E}[\rho_{\tau,left}] \le \frac{2\mathrm{E}[\tau]}{\sqrt{n}}$.

Next we bound $\rho_{\tau,special}$. The $O(\sqrt{n})$ bound we will show holds for any t, not just for $t = \tau$. For each node v, we bound the expected number of rounds that correspond to special blocks (S_t) with $y_t = v$, and then sum these expectations over all v to obtain a bound on $\mathrm{E}[\rho_{\tau,special}]$. For each v, there is at most one special block (S_t) with $y_t = v$, as this block must immediately follow the block in which v gets informed in *pp-a*. To prove the bound we just use a weaker fact: If t_v denotes the step when v gets informed in *pp-a*, and (S_t) is a special block with $y_t = v$, then it must be $t \le t_v + \sqrt{n}$, as the maximum block size is $\sqrt{n}$.

Let us fix $X_1, \ldots, X_{t-1}$, for some $t_v < t \le t_v + \sqrt{n}$, and let S_i be the first element in the block containing S_{t_v}. The probability that (S_t) is a special block with $y_t = v$, is the probability that v is contacted in the step by one of its neighbors in $\Gamma_t(v) = \Gamma(v) \setminus \{x_i, \ldots, x_{t-1}\} \cup \{y_i, \ldots, y_{t-1}\}$, and this probability is $\pi_t(v) = \frac{1}{n} \sum_{w\in\Gamma_t(v)} \frac{1}{\deg(w)}$ (we do not take into account v's neighbors that have already appeared in a pair in $(S_i, \ldots, S_{t-1})$, as S_t should not be left-incompatible). Given now that (S_t) is a special block with $y_t = v$, the expected number of rounds that correspond to this block is at most $1/q_t(v)$, where $q_t(v)$ is the probability that v is contacted by a neighbor from $\Gamma_t(v)$ in a given round. We have

$$\begin{aligned} q_t(v) &= 1 - \prod_{w\in\Gamma_t(v)} (1 - 1/\deg(w)) \\ &\ge 1 - e^{-\sum_{w\in\Gamma_t(v)} \frac{1}{\deg(w)}} = 1 - e^{-n\pi_t(v)}. \end{aligned}$$

Therefore, the expected number of rounds as a result of the possibility that (S_t) is a special block with $y_t = v$, is at most

$$\frac{\pi_t(v)}{q_t(v)} \le \frac{\pi_t(v)}{1-e^{-n\pi_t(v)}} = \frac{1}{n}\cdot\frac{n\pi_t(v)}{1-e^{-n\pi_t(v)}}$$
$$\le \frac{1}{n}\cdot(1+n\pi_t(v)) = \frac{1}{n}+\pi_t(v) \le \frac{1}{n}+\pi(v),$$

where $\pi(v) = \frac{1}{n}\sum_{w\in\Gamma(v)}\frac{1}{\deg(w)}$, and the first inequality in the last line is obtained using that $e^{-x} \le \frac{1}{1+x}$, for $x \ge 0$. Summing over all $t_v < t \le t_v + \sqrt{n}$, and over all v, we obtain

$$\mathrm{E}[\rho_{\tau,special}] \le \sqrt{n}\sum_{v\in V}\left(\tfrac{1}{n}+\pi(v)\right) = \sqrt{n}+\sqrt{n}\sum_{v\in V}\pi(v)$$
$$= 2\sqrt{n}.$$

This completes the proof of Lemma 14. □

From Lemmas 13 and 14, it follows that the expected number of rounds before all nodes get informed in *pp* is $\mathrm{E}[T(pp,G,u)] \le \mathrm{E}[\rho_\tau] = O\left(\mathrm{E}[\tau]/\sqrt{n}+\sqrt{n}\right)$. The expected time in *pp-a* until all nodes get informed is $\mathrm{E}[T(pp\text{-}a,G,u)] = \mathrm{E}[\tau]/n$, because the expected time between two consecutive steps is $1/n$, and the times between steps are independent and also independent of τ. It follows that $\mathrm{E}[T(pp,G,u)] = O\left(\sqrt{n}\cdot\mathrm{E}[T(pp\text{-}a,G,u)]+\sqrt{n}\right)$, which implies the bound of Theorem 11.

6. REFERENCES

[1] H. Acan, A. Collevecchio, A. Mehrabian, and N. Wormald. On the push&pull protocol for rumour spreading. In *Proc. 34th PODC*, pp. 405–412, 2015.

[2] H. Amini, M. Draief, and M. Lelarge. Flooding in weighted sparse random graphs. *SIAM J. Discrete Math.*, 27(1):1–26, 2013.

[3] B. Bollobás and Y. Kohayakawa. On Richardson's Model on the hypercube. In *Combinatorics, Geometry and Probability*, pp. 129–138. Cambridge University Press, 1997.

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

[5] K. Censor-Hillel and H. Shachnai. Fast information spreading in graphs with large weak conductance. In *Proc. 22nd SODA*, pp. 440–448, 2011.

[6] F. Chierichetti, S. Lattanzi, and A. Panconesi. Almost tight bounds for rumour spreading with conductance. In *Proc. 42nd STOC*, pp. 399–408, 2010.

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

[8] B. Doerr, M. Fouz, and T. Friedrich. Social networks spread rumors in sublogarithmic time. In *Proc. 43rd STOC*, pp. 21–30, 2011.

[9] B. Doerr, M. Fouz, and T. Friedrich. Asynchronous rumor spreading in preferential attachment graphs. In *Proc. 13th SWAT*, pp. 307–315, 2012.

[10] B. Doerr, M. Fouz, and T. Friedrich. Experimental analysis of rumor spreading in social networks. In *Proc. 1st MedAlg*, pp. 159–173, 2012.

[11] B. Doerr, T. Friedrich, M. Künnemann, and T. Sauerwald. Quasirandom rumor spreading: An experimental analysis. *ACM Journal of Experimental Algorithmics*, 16, 2011.

[12] R. Elsässer and T. Sauerwald. Broadcasting vs. mixing and information dissemination on Cayley graphs. In *Proc. 24th STACS*, pp. 163–174, 2007.

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

[14] J. Fill and R. Pemantle. Percolation, first-passage percolation and covering times for Richardson's Model on the *n*-cube. *The Annals of Applied Probability*, 3(2):593–629, 1993.

[15] N. Fountoulakis and K. Panagiotou. Rumor spreading on random regular graphs and expanders. In *Proc. 14th RANDOM*, pp. 560–573, 2010.

[16] N. Fountoulakis, K. Panagiotou, and T. Sauerwald. Ultra-fast rumor spreading in social networks. In *Proc. 23rd SODA*, pp. 1642–1660, 2012.

[17] G. Giakkoupis. Tight bounds for rumor spreading in graphs of a given conductance. In *Proc. 28th STACS*, pp. 57–68, 2011.

[18] G. Giakkoupis. Tight bounds for rumor spreading with vertex expansion. In *Proc. 25th SODA*, pp. 801–815, 2014.

[19] G. Giakkoupis and T. Sauerwald. Rumor spreading and vertex expansion. In *Proc. 23rd SODA*, pp. 1623–1641, 2012.

[20] M. Harchol-Balter, T. Leighton, and D. Lewin. Resource discovery in distributed networks. In *Proc. 18th PODC*, pp. 229–237, 1999.

[21] S. Janson. One, two and three times $\log(n)/n$ for paths in a complete graph with random weights. *Combinatorics, Probability & Computing*, 8(4):347–361, 1999.

[22] R. Karp, C. Schindelhauer, S. Shenker, and B. Vöcking. Randomized rumor spreading. In *Proc. 41st FOCS*, pp. 565–574, 2000.

[23] K. Panagiotou and L. Speidel. Asynchronous rumor spreading on random graphs. In *Proc. 24th ISAAC*, pp. 424–434, 2013.

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

[25] T. Sauerwald and A. Stauffer. Rumor spreading and vertex expansion on regular graphs. In *Proc. 22nd SODA*, pp. 462–475, 2011.

[26] R. van Renesse, Y. Minsky, and M. Hayden. A gossip-style failure detection service. In *Proc. 15th Middleware*, pp. 55–70, 1998.

APPENDIX

A. A DOMINATION LEMMA

LEMMA 15. *Let $Z_1,\ldots,Z_k$ be random variables such that for each $1 \le i \le k$ and $j \ge 0$, $\Pr[Z_i \le j \mid Z_1 \ldots Z_{i-1}] \ge 1-q^j$, for some $0 < q < 1$. Then $\sum_i Z_i \preccurlyeq$ NegBin$(k, 1-q)$.*

PROOF. A standard coupling argument shows that $\sum_i Z_i$ is dominated by the sum of k independent random variables $Z'_1,\ldots,Z'_k$ with $Z'_i \sim \mathrm{Geom}(1-q)$. The claim then follows because $\sum_i Z'_i \sim \mathrm{NegBin}(k, 1-q)$. □