Rumour spreading and graph conductance*

Flavio Chierichetti, Silvio Lattanzi, Alessandro Panconesi
{chierichetti,lattanzi,ale}@di.uniroma1.it
Dipartimento di Informatica
Sapienza Università di Roma

October 12, 2009

Abstract

We show that if a connected graph with n nodes has conductance ϕ then rumour spreading, also known as randomized broadcast, successfully broadcasts a message within $O(\log^4 n/\phi^6)$ many steps, with high probability, using the PUSH-PULL strategy. An interesting feature of our approach is that it draws a connection between rumour spreading and the spectral sparsification procedure of Spielman and Teng [23].

1 Introduction

Rumour spreading, also known as randomized broadcast or randomized gossip (all terms that will be used as synonyms throughout the paper), refers to the following distributed algorithm. Starting with one source node with a message, the protocol proceeds in a sequence of synchronous rounds with the goal of *broadcasting* the message, i.e. to deliver it to every node in the network. At round $t \geq 0$, every node that knows the message selects a neighbour uniformly at random to which the message is forwarded. This is the so-called `PUSH` strategy. The `PULL` variant is specular. At round $t \geq 0$ every node that does not yet have the message selects a neighbour uniformly at random and asks for the information, which is transferred provided that the queried neighbour knows it. Finally, the `PUSH-PULL` strategy is a combination of both. In round $t \geq 0$, each node selects a random neighbour to perform a `PUSH` if it has the information or a `PULL` in the opposite case. These three strategies have been introduced by [7]. One of the most studied questions for rumour spreading concerns its completion time: how many rounds will it take for one of the above strategies to disseminate the information to all nodes in the graph, assuming a worst-case source? We will say that rumour spreading is *fast* if its completion time is poly-logarithmic in the size of the network regardless of the source, and that it is *slow* otherwise.

Randomized broadcast has been intensely investigated (see the related-work section). Our long term goal is to characterize a set of necessary and/or sufficient conditions for rumour spreading to be fast in a given network. In this work, we provide a very general sufficient condition– high conductance. Our main motivation comes from the study of social networks. Loosely stated, we are looking after a theorem of the form *"Rumour spreading is fast in social networks"*. Our result is a good step in this direction because there are reasons to believe that social networks have high conductance. This is certainly the case for preferential attachment models such as that of [18]. More importantly, there is some empirical evidence that this might be the case for real social networks; in particular the authors of [17] observe how in many different social networks there exist only cuts of small (logarithmic) size having small (inversely logarithmic) conductance – all other cuts appear to have larger conductance. That is, the conductance of the social networks they analyze is larger than a quantity seemingly proportional to an inverse logarithm. Knowing that rumour spreading is fast for social networks would have several implications. First, recently it has been realized that communication networks, especially ad-hoc and mobile, have a social structure. The advent of pervasive computing is likely to reinforce this trend. Rumour spreading is also a simplified form of viral mechanism. By understanding it in detail we might be able to say something about more complex and realistic epidemic processes, with implications that might go beyond the understanding of information dissemination in communication networks.

Another relevant context for or work is the relationship between rumour spreading and expansion properties that, intuitively, should ensure fast dissemination.

Perhaps surprisingly, in the case of edge expansion

*This research was partially supported by a grant of Yahoo! Research. For the role of the Italian Ministry of University and Research please see the Unacknowledgements.

there are classes of graphs for which the protocol is slow (see [5] for more details), while the problem remains open for vertex expansion. In this paper we show the following:

THEOREM 1.1. *Given any network G and any source node,* PUSH-PULL *broadcasts the message within $O(\log^4 n/\phi^6(G))$ many rounds, where n is the number of nodes of the input graph G and $\phi(G)$ is its conductance.*

Thus, if the conductance is high enough, say $\phi^{-1} = O(\log n)$ (as it has been observed to be in real social networks [17]), then, according to our terminology, rumour spreading is fast.

We notice that the use of PUSH-PULL is necessary, as there exist high conductance graphs for which neither the PUSH, nor the PULL, strategies are *fast* on their own. Examples can be found in [5] where it is shown that in the classical preferential attachment model PUSH and PULL by themselves are slow. Although it is not known if preferential attachment graphs have high conductance, the construction of [5] also applies to the "almost" preferential attachment model of [18], which is known to have high conductance.

In terms of message complexity, we observe first that it has been determined precisely only for very special classes of graphs (cliques [15] and Erdös-Rényi random graphs [11]). Apart from this, given the generality of our class, it seems hard to improve the trivial upper bound on the number of messages– running time times number of nodes. For instance consider the "lollipop graph". Fix $\omega(n^{-1}) < \phi < o(\log^{-1} n)$, and suppose to have a path of length ϕ^{-1} connected to a clique of size $n - \phi^{-1} = \Theta(n)$. This graph has conductance $\approx \phi$. Let the source be any node in the clique. After $\Theta(\log n)$ rounds each node in the clique will have the information. Further it will take at least ϕ^{-1} steps for the information to be sent to the each node in the path. So, at least $n - \phi^{-1} = \Theta(n)$ messages are pushed (by the nodes in the clique) in each round, for at least $\phi^{-1} - \Theta(\log n) = \Theta(\phi^{-1})$ rounds. Thus, the total number of messages sent will be $\Omega(n \cdot \phi^{-1})$. Observing that the running time is $\Theta(\phi^{-1} + \log n) = \Theta(\phi^{-1})$, we have that the running time times n is (asymptotically) less than or equal to the number of transmitted messages.

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

As remarked, our result is based upon a connection with the spectral sparsification procedure of [23]. Roughly, the connection is as follows. The spectral sparsification procedure (henceforth ST) is a sampling procedure such that, given a graph G, it selects each edge uv independently with probability

$$p_{uv} := \min\left\{1, \frac{\delta}{\min\{\deg(u), \deg(v)\}}\right\} \tag{1.1}$$

where $\deg(u)$ denotes the degree of a node u and

$$\delta = \Theta\left(\frac{\log^2 n}{\phi^4}\right). \tag{1.2}$$

Spielman and Teng show that the eigenvalue spectrum of the sampled graph $ST(G)$ is, with high probability, a good approximation to that of G. In turn, this implies that $\phi(ST(G)) \geq \Omega(\phi^2(G))$ and that $ST(G)$ is connected (otherwise the conductance would be zero). The first thing we notice is that ST expands: after having applied ST, for each subset of vertices S of at most half the total volume of G, the total volume of the set of vertices reachable from S via edges sampled by ST is at least a constant fraction of the volume of S (the volume of a set of vertices is the sum of their degrees). Intuitively, if we were to use ST to send messages across the edges it samples, we would quickly flood the entire graph. Allowing for some lack of precision for sake of clarity, the second main component of our approach is that rumour spreading stochastically dominates ST, even if we run it for poly-logarithmically many rounds. That is to say, the probability that an edge is used by rumour spreading to pass the message is greater than that of being selected by ST.

In a broad sense our work draws a connection between the theory of spectral sparsification and the speed with which diffusion processes make progress in a network. This could potentially have deeper ramifications beyond the present work and seems to be worth exploring. For instance, recently in [1, 22] introduced a more efficient sparsification technique that is able to approximate the spectrum using only $O(n \log n)$, and $O(n)$, edges, respectively. Extending our approach to the new sampler appears challenging, but not without hope. The consequence would be a sharper bound on the diffusion speed. Of great interest would also be extending the approach to other diffusion processes, such as averaging. Finally, we remark that an outstanding open problem in this area is whether vertex expansion implies that rumour spreading is fast.

2 Related work

The literature on the gossip protocol and social networks is huge and we confine ourselves to what appears to be more relevant to the present work.

Clearly, at least diameter-many rounds are needed for the gossip protocol to reach all nodes. It has been shown that $O(n \log n)$ rounds are always sufficient for each connected graph of n nodes [12]. The problem has been studied on a number of graph classes, such as hypercubes, bounded-degree graphs, cliques and Erdös-Rényi random graphs (see [12, 14, 20]). Recently, there has been a lot of work on "quasi-regular" expanders (i.e., expander graphs for which the ratio between the maximum and minimum degree is constant) — it has been shown in different settings [2, 8, 9, 13, 21] that $O(\log n)$ rounds are sufficient for the rumour to be spread throughout the graph. See also [16, 19]. Our work can be seen as an extension of these studies to graphs of arbitrary degree distribution. Observe that many real world graphs (e.g., facebook, Internet, etc.) have a very skewed degree distribution — that is, the ratio between the maximum and the minimum degree is very high. In most social networks' graph models the ratio between the maximum and the minimum degree can be shown to be polynomial in the graph order.

Mihail et al. [18] study the edge expansion and the conductance of graphs that are very similar to preferential attachment (PA) graphs. We shall refer to these as "almost" PA-graphs. They show that edge expansion and conductance are constant in these graphs.

Concerning PA graphs, the work of [5] shows that rumour spreading is fast in those networks. Although PA networks have high conductance, the present work does not supersede those results, for there it is shown a $O(\log^2 n)$ time bound.

In [3] it is shown that high conductance implies that *non-uniform* (over neighbours) rumour spreading succeeds. By non-uniform we mean that, for every ordered pair of neighbours i and j, node i will select j with probability p_{ij} for the rumour spreading step (in general, $p_{ij} \neq p_{ji}$). This results does not extend to the case of uniform probabilities studied in this paper. In our setting (but not in theirs), the existence of a non uniform distribution that makes rumour spreading fast is a rather trivial matter. A graph of conductance ϕ has diameter bounded by $O(\phi^{-1} \log n)$. Thus, in a synchronous network, it is possible to elect a leader in $O(\phi^{-1} \log n)$ many rounds and set up a BFS tree originating from it. By assigning probability 1 to the edge between a node and its parent one has the desired non uniform probability distribution. Thus, from the point of view of this paper the existence of non uniform problem is rather uninteresting.

3 Preliminaries

We introduce notation, definitions, and recall several facts for later use.

Given a graph $G = (V, E)$, we denote by $ST(G)$ the graph on the same vertex set of G whose edges have been selected by the ST-sparsification algorithm, i.e. with probability defined by Equation 1.1. We use $ST(E)$ to denote the edges of $ST(G)$.

In the spectral sparsification setting of [23] the weight of edge uv, surviving the sparsification procedure, is $w_{uv} := p_{uv}^{-1}$.

NOTATION 3.1. (WEIGHTS) *The* weight *of a set of edges $E' \subseteq E$ is defined as $w_G(E') := \sum_{e \in E'} w_e$. The* weight *of a vertex u in a graph G is defined as $w_G(v) := \sum_{e \ni v} w_e$. The weight of a set of vertices S is defined as $w_G(S) := \sum_{u \in S} w_G(u)$.*

Given a graph G, the degree of a node u is denoted as $\deg_G(u)$.

DEFINITION 3.1. (VOLUME) *The* volume *of a set of vertices S of a graph G is defined to be*

$$\mathrm{Vol}_G(S) = \sum_{v \in S} \deg_G(u).$$

DEFINITION 3.2. (VOLUME EXPANSION) *Let f be a randomized process selecting edges in a graph $G = (V, E)$. Given $S \subseteq V$, the set $f(S)$ is the union of S and the set of all vertices $u \in V - S$ such that there exists some $v \in S$ and $uv \in E$ was selected by f. We say that f α-*expands for *S if*

$$\mathrm{Vol}_G(f(S)) \geq (1 + \alpha) \cdot \mathrm{Vol}_G(S).$$

The set of edges across the cut $(S, V - S)$ will be denoted as $\partial_G(S)$

DEFINITION 3.3. (CONDUCTANCE) *A set of vertices S in a graph G has* conductance ϕ *if*

$$w_G(\partial_G(S)) \geq \phi \cdot w_G(S).$$

The conductance of G is the minimum conductance, taken over all sets S such that $w_G(S) \leq w_G(V)/2$.

We will make use of a deep result from [23]. Specifically, it implies that the spectrum of $ST(G)$ is approximately the same as the one of G. It follows from [4, 6] that:

THEOREM 3.1. (SPECTRAL SPARSIFICATION) *There exists a constant $c > 0$ such that, with probability at least $1 - O(n^{-6})$, for all $S \subseteq V$ such that $w_G(S) \leq w_G(V)/2$, we have*

$$w_{ST(G)}(\partial_{ST(G)}(S)) \geq c \cdot \phi^2(G) \cdot w_{ST(G)}(S).$$

We say that an event occurs *with high probability (whp)* if it happens with probability $1 - o(1)$, where the $o(1)$ term goes to zero as n, the number of vertices, goes to infinity.

4 The proof

In this section we will prove Theorem 1.1. Before plunging into technical details, let us give an overview of the proof. The first thing we do is to show that ST-sparsification enjoys volume expansion. That is, there exists a constant $c > 0$ such that, for all sets S of volume at most $\mathrm{Vol}_G(V)/2$,

$$\mathrm{Vol}_G(ST(S)) > (1 + c \cdot \phi^2(G)) \mathrm{Vol}_G(S). \tag{4.3}$$

The second, more delicate, step in the proof is to show that rumour spreading (essentially) stochastically dominates ST-sparsification. Assume that S is the set of vertices having the message. If we run `PUSH-PULL` (henceforth PP, which plays the role of f in Definition 3.2) for $T = O(\log^3 n/\phi^4)$ rounds, then $\mathrm{Vol}_G(PP(S)) \succeq \mathrm{Vol}_G(ST(S))$, where $\succeq$ denotes stochastic domination. (Strictly speaking, this is not quite true, for there are certain events that happen with probability 1 in ST, and only with probability $1 - o(1)$ with PP.)

Consider then the sequence of sets $S_{i+1} := PP(S_i)$, and $S_0 := \{u\}$ where u is any vertex. These sets keep track of the diffusion via `PUSH-PULL` of the message originating from u (the process could actually be faster, in the sense that S_i is a subset of the informed nodes after $T \cdot i$ rounds). Then, for all i,

$$\begin{aligned}\mathrm{Vol}_G(S_{i+1}) = \mathrm{Vol}_G(PP(S)) \geq \\ \geq \mathrm{Vol}_G(ST(S)) > (1 + c\phi^2(G)) \mathrm{Vol}_G(S_i).\end{aligned}$$

The first inequality follows by stochastic domination, while the second follows from Equation 4.3. Since the maximum volume is $O(n^2)$, we have that $\mathrm{Vol}(S_t) > \mathrm{Vol}(G)/2$ for $t = O(\log n/\phi^2)$. This means that within $O(T \log n/\phi^2)$ many rounds we can deliver the message to a set of nodes having more than half of the network's volume. To conclude the argument we use the fact that PP is specular. If we interchange `PUSH` with `PULL` and viceversa, the same argument "backwards" shows that once we have S_t we can reach any other vertex within $O(T \log n/\phi^2)$ additional many rounds. After this informal overview, we now proceed to the formal argument. In what follows there is an underlying graph $G = (V, E)$. where $n := |V(G)|$, for which we run ST and PP.

4.1 Volume expansion of ST-sparsification

Our goal here is to show Equation 4.3. We begin by showing that the weight of every vertex in $ST(G)$ is concentrated around its expected value, namely its degree in G.

LEMMA 4.1. *With probability at least $1 - n^{-\omega(1)}$ over the space induced by the random ST-sparsification algorithm, for each node $v \in V(G)$ we have that*

$$w_{ST(G)}(v) = (1 \pm o(1)) \deg_G(v).$$

Proof. If $\deg_G(v) \leq \delta$, then $w_{ST(G)}(v)$ is a constant random variable with value $\deg_G(v)$. If we assume the opposite we have that $E[w_{ST(G)}(v)] = \deg_G(v)$, by definition of ST-sparsification. Recalling the definition of δ (Equation 1.2), let $X = w_{ST(G)}(v)\delta/\deg(v)$. Then, $E[X] = \delta$. By the Chernoff bound,

$$\begin{aligned}\Pr[|X - E[X]| \geq \epsilon E[X]] \leq \\ \leq 2 \exp\left(-\frac{\epsilon^2}{3} E[X]\right) = 2 \exp\left(-\frac{\epsilon^2}{3}\delta\right).\end{aligned}$$

Since $\delta = \Theta\left(\frac{\log^2 n}{\phi^4}\right)$, if we pick $\epsilon = \omega(\phi^2/\sqrt{\log n})$, the claim follows. □

COROLLARY 4.1. *Let $S \subseteq V$ be such that $\mathrm{Vol}_G(S) \leq \mathrm{Vol}_G(V)/2$. With probability at least $1 - n^{-\omega(1)}$ over the space induced by the random ST-sparsification algorithm, we have that*

$$w_{ST(G)}(S) = (1 \pm o(1)) \mathrm{Vol}_G(S).$$

Theorem 3.1 states that ST-sparsification enjoys weight expansion. By means of Lemma 4.1 and Corollary 4.1 we can translate this property into volume expansion. Recall that $ST(S)$ is S union the vertices reachable from S via edges sampled by ST.

LEMMA 4.2. (VOLUME EXPANSION OF ST) *There exists a constant c such that for each fixed $S \subseteq V$ having volume at most $\mathrm{Vol}_G(V)/2$, with high probability*

$$\mathrm{Vol}_G(\mathrm{ST}(S)) > (1 + c \cdot \phi^2(G)) \mathrm{Vol}_G(S).$$

Proof. By Theorem 3.1, $w_{ST(G)}(\partial_{ST(G)}(S)) \geq c \cdot \phi^2(G) \cdot w_{ST(G)}(S)$. Clearly,

$$w_{ST(G)}(ST(S)) \geq w_{ST(G)}(\partial_{ST(G)}(S)).$$

By Corollary 4.1 we have that $\mathrm{Vol}_G(ST(S)) = w_{ST(G)}(ST(S))(1 \pm o(1))$ and $\mathrm{Vol}_G(S) = w_{ST(G)}(S)(1 \pm o(1))$. The constant c in Theorem 3.1 and the error terms in Corollary 4.1 can be chosen in such a way that $\mathrm{Vol}_G(ST(S)) > (1 + c' \cdot \phi^2(G)) \mathrm{Vol}_G(S)$ for some $c' > 0$. The claim follows. □

We end this section by recording a simple monotonicity property stating that if a process enjoys volume expansion, then by adding edges expansion continues to hold.

LEMMA 4.3. *Let f and g be a randomized processes that select each edge e in G independently with probability p_e and p'_e, respectively, with $p'_e \geq p_e$. Then, for all $t > 0$ and S,*

$$\Pr(\mathrm{Vol}_G(g(S)) > t) \geq \Pr(\mathrm{Vol}_G(f(S)) > t).$$

Proof. The claim follows from a straightforward coupling, and by the fact that if $A \subseteq B$ then $\mathrm{Vol}(A) \leq \mathrm{Vol}(B)$. □

5 The road from ST-sparsification to Rumour Spreading

The goal of this section is to show that *PP*stochastically dominates *ST*. As stated the claim is not quite true and the kind of stochastic domination we will show is slightly different. Let us begin by mentioning what kind of complications can arise in proving a statement like this.

A serious issue is represented by the massive dependencies that are exhibited by *PP*. To tackle this we introduce a series of intermediate steps, by defining a series of processes that bring us from *ST* to *PP*. We will relax somewhat *PP* and *ST* by introducing two processes *PPW* and *DST* to be defined precisely later. In brief, *PPW* is the same as *PP* except that vertices select neighbours without replacement. *DST* differs from *ST* by the fact that edges are "activated" (we will come to this later) by both endpoints. Again, slightly simplifying a more complex picture for the sake of clarity, the main flow of the proof is to show that $ST \preceq DST \preceq PPW \preceq PP$, where $\preceq$ denotes stochastic domination. Let us now develop formally this line of reasoning.

The first intermediate process is called *double ST-sparsification* henceforth (DST) and it is defined as follows. *DST* is a process in which vertices select edges incident on them (similarly to what happens with *PP*). With DST each edge $e \ni u$ is *activated* independently by u with probability

$$p_e := \min\left\{1, \frac{\delta}{\deg_G(u)}\right\}. \tag{5.4}$$

An edge $e = uv$ is *selected* if it is activated by at least one of its endpoints. Clearly $DST \succeq ST$ and thus it follows immediately from Lemma 4.3 that DST expands. We record this fact for later use.

LEMMA 5.1. (VOLUME EXPANSION OF DST) *There exists a constant c such that for each fixed $S \subseteq V$ having volume at most $\mathrm{Vol}_G(V)/2$, with high probability*

$$\mathrm{Vol}_G(\mathrm{DST}(S)) > (1 + c \cdot \phi^2(G)) \,\mathrm{Vol}_G(S).$$

Therefore from now on we can forget about ST and work only with DST. The next lemma shows that with high probability after *DST*-sparsification the degree of all vertices is $O(\log^2 n/\phi^2)$.

LEMMA 5.2. *Let ξ be the event* "with *DST* no node will activate more than 2δ edges"*. Then,*

$$\Pr(\xi) = 1 - n^{-\omega(1)}.$$

Proof. The only case to consider is $\deg(v) > 2\delta$. Let X = (# of edges activated by v). Then $E[X] = \sum_{u \in N(v)} \frac{\delta}{\deg(v)} = \delta$. Invoking the Chernoff bound we get (see for instance [10]),

$$\Pr[X \geq 2E[X]] \leq 2\exp\left(-\Omega(\delta)\right) \leq n^{-\omega(1)}$$

for n large enough. □

Remark: For the remainder of the section, when dealing with *DST* we will work in the subspace defined by conditioning on ξ. We will do so without explicitly conditioning on ξ, for sake of notational simplicity.

The second step to bring *PP* "closer" to *ST* is to replace *PP* with a slightly different process. This process is dubbed PP *without replacement* (henceforth PPW) and it is defined as follows. If *PPW* runs for t rounds, then each vertex u will select $\min\{\deg_G(u), t\}$ edges incident on itself without replacement (while *PP* does it with replacement). The reason to introduce *PPW* is that it is much easier to handle than *PP*.

NOTATION 5.1. (TIME HORIZONS) *Given a vertex set $S \subseteq V$ we will use the notation $A := (A_u : u \in S)$ to denote a collection of vertex sets, where each A_u is a subset of the neighbours of u. A vector of integers $T = (t_u : u \in S)$ is called a* time horizon *for S. Furthermore we will use the notation $\|A\| := (|A_u| : u \in S)$, to denote the time horizon for S that corresponds to A.*

NOTATION 5.2. (BEHAVIOUR OF PPW) *Let S be a set of vertices in a graph G and let T be a time horizon for S.* $\mathrm{PPW}(T, S)$ *is the process where every vertex $u \in S$ activates a uniformly random subset of $\min\{\deg_G(u), t_u\}$ edges incident on itself, to perform a* `PUSH-PULL` *operation for each of them.*

Notice that *PPW* might specify different cardinalities for different vertices. This is important for the proofs to follow.

With this notation we can express the outcome of *DST* sampling. Focus on a specific set of vertices $S = (u_1, \ldots, u_k)$. We know that *DST*(S) expands with respect to S and we want to argue that so does *PPW*. The crux of the matter are the following two simple lemmas.

LEMMA 5.3. *Let u be a vertex in G and t a positive integer. And let* $\mathrm{DST}(u)$ *and* $\mathrm{PPW}(t,u)$ *denote, respectively, the subset of edges incident on u selected by the two processes. Then,*

$$\Pr(\mathrm{DST}(u) = A_u \cup \{u\} \mid |A_u| = t) = \\ = \Pr(\mathrm{PPW}(t, u) = A_u \cup \{u\}).$$

Proof. With DST each vertex activates (and therefore selects) edges incident on itself with the same probability. If we condition on the cardinality, all subsets are equally likely. Therefore, under this conditioning, $DST(u)$ simply selects a subset of t edges uniformly at random. But this is precisely what $PPW(t,u)$ does. $\square$

LEMMA 5.4. *Let $S = \{v_1, \ldots, v_{|S|}\}$ be a subset of vertices of G and $T = (t_1, \ldots, t_{|S|})$ a time horizon for S. Then,*

$$\Pr\left(\bigwedge_{i=1}^{|S|} \mathrm{DST}(v_i) = A_{v_i} \cup \{v_i\} \;\middle|\; \|A\| = T\right) = \\ = \Pr\left(\bigwedge_{i=1}^{|S|} \mathrm{PPW}(t_i, v_i) = A_{v_i} \cup \{v_i\}\right).$$

Proof. This follows from Lemma 5.3 and the fact that under both DST and PPW vertices activate edges independently. $\square$

In other words, for every realization A of DST there is a time horizon T_A such that the random choices of PPW are distributed exactly like those of DST. Said differently, if we condition on the cardinalities of the choices made by DST, then, for those same cardinalities, PPW is distributed exactly like DST. To interpret the next lemma refer to Definition 3.2.

LEMMA 5.5. *Let $T := (2\delta, \ldots, 2\delta)$. There exists $c > 0$ such that, for all sets $S \subseteq V$,*

$$\Pr(\mathrm{PPW}(S,T)\ (c\phi^2)\text{-}expands\ for\ S) \geq \\ \geq \Pr(\mathrm{DST}(S)\ (c\phi^2)\text{-}expands\ for\ S) = \\ = 1 - o(1).$$

Proof. For the first inequality, recall that we are assuming that DST operates under conditioning on ξ. Thus, by Lemma 5.2 we have that each $u \in V$ activates at most 2δ edges. Therefore T majorizes every time horizon T_S for which Lemma 5.4 holds. The last equality is derived from Lemma 5.2. $\square$

We conclude the series of steps by showing that, given any set S, PP also expands with high probability.

LEMMA 5.6. *Consider the* PP *process. Take any node v, and an arbitrary time t_0. Between time t_0 and $t_1 = t_0 + 9\delta \cdot \log n$, node v activates at least $\min(2\delta, \deg(v))$ different edges with high probability.*

Proof. We split the proof into two cases, $\deg(v) \leq 3\delta$ and $\deg(v) > 3\delta$. In the former case, a straightforward coupon collector argument applies.[1].

Otherwise $\deg(v) > 3\delta$, PP will either choose $> 2\delta$ different edges during the 9δ rounds, or it will choose at most $\leq 2\delta$ different edges. What is the probability of the latter event? In each round the probability of choosing a new edge will be $\geq \frac{\deg(v)-2\delta}{\deg(v)} \geq 1 - \frac{2\delta}{\deg(v)} \geq 1 - \frac{2}{3} = \frac{1}{3}$. But then by Chernoff bound, the probability of this event is at most $n^{-\omega(1)}$. $\square$

To summarize, if PP is run $t_1 - t_0 = O(\delta \log n)$ steps, with high probability every node in $ST(S)$ selects at least $\min\{\deg_G(u), 2\delta\}$ many edges, and therefore dominates PPW.

6 The speed of *PP*

In this section, we upper bound the number of steps required by PP to broadcast a message in the worst case. The basic idea is that, as we have seen in the previous section, a PP requires $(\log^3 n/\phi^4)$ rounds to expand out of a set. Suppose the information starts at vertex v. Since each expansion increases the total informed volume by a factor of $(1 + \Omega(\phi^2))$ we have that after $(\log^4 n/\phi^6)$ rounds, the information will have reached a set of nodes of volume greater than half the volume of the whole graph. Consider now another node w. By the symmetry of the `PUSH-PULL` process, w will be "told" the information by a set of nodes of volume bigger than half of the volume of the graph in $O(\log^4 n/\phi^6)$ many rounds. Thus the information will travel from v to w in $O(\log^4 n/\phi^6)$ many rounds, with high probability.

To develop the argument more formally, let us define a macro-step of PP as $2\delta \log n$ consecutive rounds. We start from a single node v having the information, $S_0 = \{v\}$. As we saw, in each macro-step, with probability $\geq 1 - O(n^{-6})$ the volume of the set of nodes that happen to have the information increases by a factor $1 + \Omega(\phi^2)$, as long as the volume of S_i is $\leq \frac{1}{2}\mathrm{Vol}_G(V)$.

Take any node $w \neq v$. If the information started at w, in $O(\log_{1+\Omega(\phi^2)} n) = O\left(\frac{1}{\phi^2}\log n\right)$ macro-steps the information will have reached a set of nodes $S = S_{O(\frac{1}{\phi^2}\log n)}$ of total degree strictly larger than $\frac{1}{2}\mathrm{Vol}_G(V)$ with probability $1 - O(n^{-6} \cdot \frac{\log n}{\phi^2}) \geq 1 - O(n^{-2}\log n)$. Note that the probability that the information, starting from some node in S, gets sent to w in $O(\frac{1}{\phi^2}\log n)$

[1]The probability of non-activation in $9\delta \log n$ rounds of some edge will be equal to $(1 - \frac{1}{\deg(v)})^{9\delta \log n} \leq (1 - \frac{1}{3\delta})^{9\delta \log n} \leq n^{-3}$. Thus, by union bounding over all its edges, the probability that event fails to happen is $O(n^{-2})$.

steps is greater than or equal to the probability that w spreads the information to the whole of S (we use PUSH-PULL, so each edge activation both sends and receive the information — thus by activating the edges that got the info from w to S in the reverse order, we could get the information from each node in S to w. Note that the probability of the two activation sequences are exactly the same).

Now take the originator node v, and let it send the information for $O\left(\frac{1}{\phi^2}\log n\right)$ macro-rounds (for a total of $O\left(\frac{\log^4 n}{\phi^6}\right)$ many rounds). With high probability, the information will reach a set of nodes S_v of volume strictly larger than $\frac{1}{2}\operatorname{Vol}_G(V)$. Take any other node w, and grow its S_w for $O(\frac{1}{\phi^2}\log n)$ rounds with the aim of letting it grab the information. Again, after those many rounds, w will have grabbed the information from a set of volume at least $\frac{1}{2}\operatorname{Vol}_G(V)$ with probability $1 - O(n^{-2}\log n)$. As the total volume is $\operatorname{Vol}_G(V)$ the two sets will intersect — so that w will obtain the information with probability $1 - O(n^{-2}\log n)$. Union bounding over the n nodes, gives us the main result: with probability $\geq 1 - O(n^{-1}\log n) = 1 - o(1)$, the information gets spread to the whole graph in $O\left(\frac{\log^4 n}{\phi^6}\right)$ rounds.

Acknowledgements

We thank the anonymous referees for their useful suggestions.

Unacknowledgements

This work is ostensibly supported by the the Italian Ministry of University and Research under the FIRB program, project RBIN047MH9-000. The Ministry however has not paid its dues and it is not known whether it will ever do.

References

[1] J. D. Batson, D. A. Spielman, and N. Srivastava. Twice-ramanujan sparsifiers. In *Proceedings of STOC*, 2009.

[2] P. Berenbrink, R. Elsässer, and T. Friedetzky. Efficient randomized broadcasting in random regular networks with appliations in peer-to-peer systems. In *Proceedings of PODC*, 2008.

[3] S. P. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Gossip algorithms: design, analysis and applications. *IEEE Transactions on Information Theory*, 52, 2006.

[4] J. Cheeger. A lower bound for smallest eigenvalue of laplacian. *Problems in Analysis*, 1970.

[5] F. Chierichetti, S. Lattanzi, and A. Panconesi. Rumor spreading in social networks. In *Proceedings of ICALP*, 2009.

[6] F. R. K. Chung. Spectral graph theory. In *CBMS Regional Conference Series in Mathematics*. American Mathematical Society, 1997.

[7] 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. In *Proceedings of PODC*, 1987.

[8] B. Doerr, T. Friedrich, and T. Sauerwald. Quasirandom broadcasting. In *Proceedings of SODA*, 2008.

[9] B. Doerr, T. Friedrich, and T. Sauerwald. Quasirandom rumor spreading: Expanders, push vs. pull, and robustness. In *Proceedings of ICALP*, 2009.

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

[11] R. Elsässer. On the communication complexity of randomized broadcasting in random-like graphs. In *Proceedings of SPAA*, 2006.

[12] U. Feige, D. Peleg, P. Raghavan, and E. Upfal. Randomized broadcast in networks. *Algorithms*, 1, 1990.

[13] T. Friedrich and T. Sauerwald. Near-perfect load balancing by randomized rounding. In *Proceedings of STOC*, 2009.

[14] A. Frieze and G. Grimmett. The shortest-path problem for graphs with random arc-lengths. *Algorithms*, 10, 1985.

[15] R. Karp, C. Schindelhauer, S. Shenker, and B. Vöcking. Randomized rumor spreading. In *Proceedings of FOCS*, 2000.

[16] D. Kempe, A. Dobra, and J. Gehrke. Gossip-based computation of aggregate information. In *Proceedings of FOCS*, 2003.

[17] J. Leskovec, K. J. Lang, A. Dasgupta, and M. W. Mahoney. Statistical properties of community structure in large social and information networks. In *Proceedings of WWW*, 2008.

[18] M. Mihail, C. H. Papadimitriou, and A. Saberi. On certain connectivity properties of the internet topology. *J. Comput. Syst. Sci.*, 72(2):239–251, 2006.

[19] D. Mosk-Aoyama and D. Shah. Fast distributed algorithms for computing separable functions. *Transactions on Information Theory*, 54, 2008.

[20] B. Pittel. On spreading a rumor. *SIAM Journal on Applied Mathematics*, 47, 1987.

[21] T. Sauerwald. On mixing and edge expansion properties in randomized broadcasting. In *Proceedings of ISAAC*, 2007.

[22] D. A. Spielman and N. Srivastava. Graph sparsification by effective resistances. In *Proceedings of STOC*, 2008.

[23] D. A. Spielman and S.-H. Teng. Nearly-linear time algorithms for graph partitioning, graph sparsification, and solving linear systems. In *Proceedings of STOC*, 2004.